Optimization of Complex Chemical Plants by a Gradient Technique

E. STANLEY LEE

Kansas State University, Manhattan, Kansas

The gradient technique and Lagrange multiplier are used to obtain the optimum of complex chemical plants. The advantage of this approach is its ability to handle nearly all types of complex stages in a natural way. This approach is used to solve a heterogeneous complex chemical process with recycle. It is shown that although there are many different iteration loops, the convergence rate is fast even with rough starting values and only 0.3 minute is needed to obtain the optimum operating conditions.

This paper describes the combined use of the gradient technique and the Lagrange multiplier to optimize complex chemical plants. In a previous paper (*1*), the functional gradient technique has been used to solve continuous and simple stagewise processes. It has been shown that the Lagrange multiplier can be used to extend the gradient technique to the stagewise processes with implicit equations. This paper shows that this technique is a useful tool in obtaining optimum conditions of a complex plant.

There are two classes of optimization techniques. The mathematical programming techniques such as linear and nonlinear programming ignore the structure of the process completely and optimize the various stages simultaneously. Since these techniques treat an N stage process as if the process were one stage, we shall refer to this class of techniques as the *single-stage optimization techniques.* The other class of techniques is known as the *multistage optimization techniques.* The dynamic programming (*2*) and the discrete maximum principle (*3*) are multistage techniques in which a certain relationship is used to isolate the interconnections between the various stages and thus permit us to search one stage at time. This relationship is the principle of optimality in dynamic programming and the adjoint equations in the maximum principle.

Both of the above two classes of techniques are not suited for optimizing nonlinear complex processes in which the stages are interconnected in a complex manner. Although the single-stage optimization techniques are useful for solving linear complex processes by linear programming, no efficient single-stage technique exists for solving nonlinear problems. Since the computational effort depends exponentially on the number of variables, treating an N stage process as a one stage process is certainly not desirable. The multistage techniques avoid this difficulty. However, other difficulties arise.

The multistage techniques are developed for simple serial processes. However, the stagewise serial structure is seldom seen in a complex chemical plant. For example, in a modern refinery, each operating unit has its unique purpose. Not only are the equations that govern the operating conditions usually different for different operating units, but the control and state variables may also be different. If one considers an operating unit as a stage, then even a simple homogeneous serial arrangement of only three stages is seldom seen. Thus the concept of cut state based on decomposition (*16, 17*), which has been recently developed for optimizing complex processes by the multistage techniques, is also not suited for this type of complex chemical plants.

A second difficulty in applying the multistage techniques to complex chemical processes arises from the fact that a chemical plant is not only nonlinear and has a very complex structure, but also has a large number of variables. For example, because of the dimensionality difficulty, dynamic programming cannot conveniently handle optimization problems with more than three state variables with the currently available computers.

The gradient technique is an elementary concept for the solution of optimization problems. A detailed discussion of this technique can be found in the literature (*4, 5*). Because of the computational appeal of the method, it has been adopted to the optimization of various problems. It is a valuable tool for calculating maximum or minimum in multivariable ordinary calculus problems (*5* to *7*). It has also been adopted to the solution of nonlinear programming problems, where the treatment of constraints is emphasized (*8* to *10*). Even in the numerical solution of systems of algebraic equations, the gradient technique has been found to be useful (*11*).

A version of the gradient technique, which is known as the functional gradient technique, has been applied extensively to aerospace systems (*12* to *14*). The functional gradient technique was developed for variational problems. This paper generalizes and applies the functional gradient technique to complex nonlinear systems. As has been discussed in a previous paper (*1*), this generalization cannot be obtained in a straightforward manner from the functional gradient technique.

The gradient technique is especially suited for optimizing the nonlinear complex chemical process with many variables. The advantage of this approach is its ease in handling heterogeneous complex processes. There is no additional effort involved in handling branching or interconnected streams. Note that this is not the case when decomposition is used (*16*). The convergence rate in handling recycle by iteration has been quite fast for the few problems solved.

The computational advantages of this technique are that it constitutes an approximation in policy space (*2*), and, in general, the convergence property is monotone.

The gradient technique has several disadvantages. The most serious one is that convergence to an extreme which is not the optimum may occur. However, for practical problems, this is not a serious disadvantage. The results can be judged by the physical situation of the problem. Another disadvantage is that it cannot conveniently handle problems with inequality constraints on the state variable. Generally, iteration must be used to handle this kind of constraint.

After a short introduction concerning the terminology of a complex chemical process, the approach is introduced by an actual example. It is shown that although there are many different iteration loops, the convergence rate is fast even with rough starting values and only 0.3 min. is needed to obtain the optimum operating conditions for the problem solved in this work.

COMPLEX STAGEWISE PROCESSES

Generally, a stage represents a unit in which a certain transformation takes place. A stage may be a complete unit such as a distillation column, or it may be part of a unit such as a plate in a distillation column. The equations governing these transformations will be called process equations. The state variables of a stage are those quantities that must be known in order to render an optimal decision. The control variables are those quantities that are considered to be under control and which are to be chosen optimally. In any stage, the number of state variables must be equal to the number of process equations. If the process equations and the state and control variables are the same for all the stages involved, this stagewise process is called a *homogeneous stagewise process.* Otherwise, it is called a *heterogeneous stagewise process.* In this paper, all the processes considered will be heterogeneous processes. There are basically three kinds of heterogeneous processes depending on whether this heterogeneity is due to differences in process equations, control variables, or state variables. For the present approach, only the heterogeneity due to differences in state variables needs special attention. All the other heterogeneities can be treated the same way as a homogeneous process.

There are three basic types of stages. A simple stage is a stage with one entering stream and one leaving stream; a branching stage is a stage with one entering and several leaving streams; and a combining stage is one with several entering streams and only one leaving stream. More complex stages can be formed by combining these three basic stages.

A simple stagewise process consists of simple stages only. A complex process may consist of simple stages and various complex stages which form various complex topological arrangements. The most frequently encountered complex processes are processes with branching streams, combining streams, interconnected streams, and recycle or feedback streams. For the present technique, only feedback streams need to be treated separately.

For the ease of description, a complex process will be divided into sections. Each section is a part of a complex process which, by itself, forms a simple homogeneous serial arrangement of stages. A section may consist of one or any number of these stages. A new section may be created either by a topologically complex stage, such as a branching stage or a combining stage, or by a change of the control variables, state variables, or process equations.

A fictitious stage will be added whenever a new state variable is created or whenever the feed values of the state variables to any stage are different from the effluent values of the same state variables of the previous stage. For example, the latter situation may occur when the process has combining stages. The purpose of the fictitious stages is to adjust the state variables.

REACTORS FOLLOWED BY EXTRACTORS WITH RECYCLE

Since we are dealing with a variety of complex topological situations, it would be impossible to derive an algorithm covering all situations. Furthermore, most complex situations can be treated by the use of the same basic principle. For the purpose of concrete illustration, a simple but certainly not trivial example will be considered in detail in this section.

The process shown in Figure 1 is a heterogeneous stagewise process with recycle. Stages 1 to 5 and 7 to 9 form two sections. The stages represented by dotted lines are either fictitious stages or stages which do not need to be optimized. Figure 1*a* shows the actual material flow and Figure 1*b* shows the quantities needed for optimization calculations.

Stages 1 to 5 are five continuous flow stirred tank reactors in which the following reactions take place:

$$A \overset{k_a}{\rightarrow} B \overset{k_b}{\rightarrow} C \tag{1}$$

where B is the more valuable product and C is a waste product. Both reactions are first order and are irreversible. If one assumes that the volume of the reaction mixture remains constant, the reaction rate equations can be obtained by material balances around the reactors

$$x_i - x_{i-1} + t_i k_{ai} x_i = 0, \qquad i = 1, \ldots, 5 \tag{2}$$

$$y_i - y_{i-1} + t_i k_{bi} y_i - t_i k_{ai} x_i = 0, \quad i = 1, \ldots, 5 \tag{3}$$

where x and y represent the concentrations of A and B, respectively; t represents the holding time or residence time of the reaction mixture in the reactors; and k_a and k_b are the rate constants for reactions 1 and 2, respectively:

$$k_{ai} = G_a \exp\left(-\frac{E_a}{RT_i}\right), \quad k_{bi} = G_b \exp\left(-\frac{E_b}{RT_i}\right), \quad i = 1, \ldots, 5 \tag{4}$$

The variable T_i is the temperature of the reaction mixture in reactor i. The other symbols are defined in the table of notations. The state variables are x and y. The variable T is the control variable.

Stages 7 to 9 are three crosscurrent extractors. Assuming the waste product C can be separated very easily from the reaction mixture, no consideration for optimization is necessary for this separation. One possible situation would be that the waste product C has a much lower boiling point than the other components. The component C is separated in stage 6. Stage 6 also serves as a fictitious stage. There are changes in the state variables and in the units used for the state variables between stages 5 and 7.

The mixture containing A and B is being extracted by a miscible solvent, W. The concentrations of A and B in the extract streams are represented by x_E and y_E, respectively. The flow rates of the original mixture and the raffinate streams are represented by q. The extract stream has an inlet flow rate w of pure solvent W and an outlet rate u. The streams q and w are expressed in total weight and all concentrations are in weight fractions.

To obtain the process equations for stage 6, note that there is a change in the units used between the reactors and the extractors. The concentrations x and y in the reactors are in moles/liter while the concentrations x and y in the extractors are in weight fractions. If one assumes the volumetric feed rate to stage 0 is one liter of x_f and y_f/hr., then the volumetric feed rate to stage 1 is

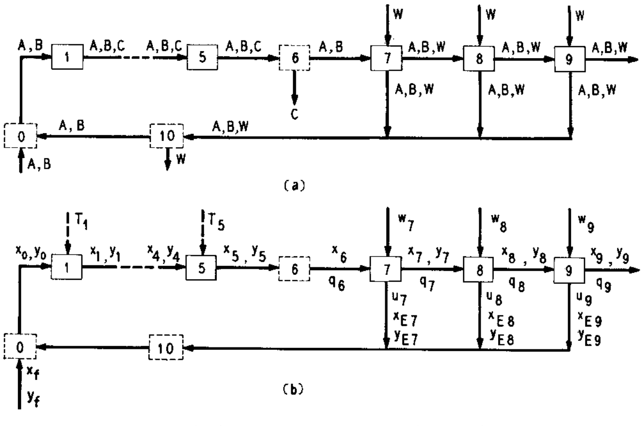

Fig. 1. A complex process with recycle.

$$V(u_i, x_i) = 1 + \frac{\sum_{7}^{9} u_n x_{En}}{M_x V_x} + \frac{\sum_{7}^{9} u_n y_{En}}{M_y V_y}, \quad i = 7, 8, 9 \tag{5}$$

where M_x and M_y are the molecular weights of A and B respectively, and V_x and V_y are moles per unit volume of A and B. The second and third terms in Equation (5) represent the volumetric recycle rates of A and B.

Since there is no change in volume in reactors 1 to 5, the volumetric outlet rate from stage 5, based on 1 liter of original feed, is also represented by Equation (5). Thus, the flow rate q_6 in weight unit is

$$q_6 - (x_5 M_x + y_5 M_y) \left[1 + \frac{\sum_{7}^{9} u_n x_{En}}{M_x V_x} + \frac{\sum_{7}^{9} u_n y_{En}}{M_y V_y} \right] = 0 \tag{6}$$

and the variables x_6 and y_6 are

$$x_6 - \frac{x_5 M_x}{x_5 M_x + y_5 M_y} = 0 \tag{7}$$

$$y_6 - \frac{y_5 M_y}{x_5 M_x + y_5 M_y} = 0 \tag{8}$$

The equations for the extractors can be obtained by mass balances on the total flow rates and the components x and y:

$$q_n + u_n - q_{n-1} - w_n = 0, \qquad n = 7, 8, 9 \tag{9}$$

$$q_n x_n - q_{n-1} x_{n-1} + u_n x_{En} = 0, \qquad n = 7, 8, 9 \tag{10}$$

$$q_n y_n - q_{n-1} y_{n-1} + u_n y_{En} = 0, \qquad n = 7, 8, 9 \tag{11}$$

The purpose of stages 7 to 9 is to extract as much of the component A as possible and then recycle it back to the original feed mixture. To simplify the recycle problem, we have assumed that the solvent W can be separated very easily from the extract streams and this separation is performed in stage 10. Because this separation is so easy, stage 10 is not considered in the optimizing calculations. Even if this separation is not easy and must be considered for optimizing calculations, the following approach can still be used.

With given initial feed concentrations, x_f and y_f, the feed concentrations to stage 1 in moles per liter are

$$x_o - \frac{x_f + \sum_{7}^{9} u_n x_{En}/M_x}{1 + \sum_{7}^{9} u_n x_{En}/(M_x V_x) + \sum_{7}^{9} u_n y_{En}/(M_y V_y)} = 0 \tag{12}$$

$$y_o - \frac{y_f + \sum_{7}^{9} u_n y_{En}/M_y}{1 + \sum_{7}^{9} u_n x_{En}/(M_x V_x) + \sum_{7}^{9} u_n y_{En}/(M_y V_y)} = 0 \tag{13}$$

All the above derivations are based upon 1 liter of the original feed mixture.

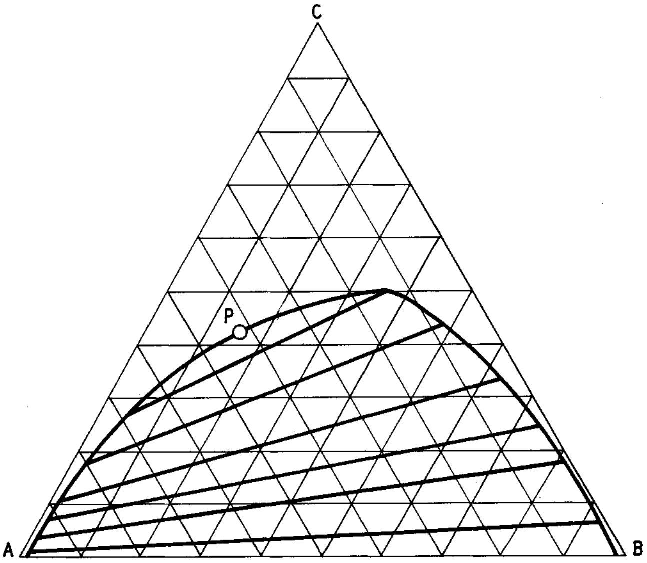

Fig. 2. Equilibrium diagram.

There are six unknowns, q, u, x, y, x_E, and y_E in Equations (9) to (11). In addition to these three equations, three more equations can be obtained from the equilibrium relationship. Typical equilibrium data are used for the extraction and are shown in Figure 2 by the use of triangular coordinates. From the saturation curve in the raffinate layer, the following relationship is obtained:

$$y = f_1(x) \tag{14}$$

The tie line relationship gives

$$x_E = f_2(x) \tag{15}$$

By the combined use of both the tie line and the saturation curve relationships the third equation is obtained:

$$y_E = f_3(x) \tag{16}$$

The control variable for the extractors is the amount of solvent used, w_i, $i = 7, 8, 9$. The variables q_i, u_i and x_i, $i = 7, 8, 9$ are state variables. Equations (9) to (11) are the process equations for these three state variables. Once x is known, the variables y, x_E, and y_E can be obtained by the use of Equations (14) to (16). Thus, y, x_E, and y_E are not state variables but are functions of the state variable x.

The two state variables for stage 6 are q_6 and x_6. Once q_6 and x_6 are known, y_6 is fixed. Thus, y_6 is not a state variable and Equation (8) is not needed for the optimization calculations.

No consideration is necessary for stage 10. Equations (12) and (13) are the two process equations for stage 0 with x_o and y_o as the state variables.

It is interesting to compare Figures 1*a* and 1*b* and to note what kind of variable is considered as state variable. The state variables are essentially the unknown variables which are controlled by the process equations. In certain cases, the choice of a state variable is very arbitrary. Thus, y_6 could have been considered as the state variable instead of x_6. According to this definition, Equations (4), (5), and (14) to (16) are not process equations.

The quantity to be maximized is

$$\phi = q_9 y_9 - \lambda_1 \sum_{i=7}^{9} w_i - \lambda_2 \sum_{i=7}^{9} u_i x_{Ei} \tag{17}$$

where λ_1 is the cost of the extracting solvent per unit value of the desired product B and λ_2 is the recycle cost.

Both λ_1 and λ_2 have the meaning of Lagrange multipliers and have been used extensively in this manner in the literature (*1*). The recycle cost is assumed proportional to the amount of x recycled.

In establishing Equation (17), all the operating costs have been neglected. These costs can be added to Equation (17) with only minor modifications in the following derivations.

The problem is to find the control variables, T_i and w_j, where $i = 1, \ldots, 5$, and $j = 7, 8, 9$ so that the quantity ϕ in Equation (17) is maximized. The process equations [Equations (2), (3), (6), (7), (9) to (13)] are the constraints for this optimization. The symbols t_i, $i = 1, \ldots, 5$, G_a, G_b, E_a, E_b, R, M_x, M_y, V_x, and V_y in the process equations are all constants.

Following the standard approach in the use of Lagrange multipliers in the classical methods of optimization, we shall introduce the Lagrange multipliers μ_i, θ_i, and η_j, where $i = 0, 1, \ldots, 9$ and $j = 7, 8, 9$ and consider the equation:

$$\begin{aligned}\phi = {} & q_9y_9 - \lambda_1\sum_7^9 w_i - \lambda_2\sum_7^9 u_ix_{Ei} \\ & - \sum_1^5[\mu_i(x_i - x_{i-1} + t_ik_{ai}x_i)] - \sum_1^5[\theta_i(y_i - y_{i-1} \\ & + t_ik_{bi}y_i - t_ik_{ai}x_i)] - \mu_6\left(x_6 - \frac{x_5M_x}{x_5M_x + y_5M_y}\right) \\ & - \theta_6[q_6 - (x_5M_x + y_5M_y)V] - \sum_7^9[\mu_i(q_ix_i \\ & - q_{i-1}x_{i-1} + u_ix_{Ei})] \\ & - \theta_7\left[q_7y_7 - q_6\frac{y_5M_y}{x_5M_x + y_5M_y} + u_7y_{E7}\right] \\ & - \sum_8^9[\theta_i(q_iy_i - q_{i-1}y_{i-1} + u_iy_{Ei})] \\ & - \sum_7^9[\eta_i(q_i + u_i - q_{i-1}w_i)] \\ & - \mu_o[x_o - (x_f + \sum_7^9 u_ix_{Ei}/M_x)/V] \\ & - \theta_o[y_o - (y_f + \sum_7^9 u_iy_{Ei}/M_y)/V] \qquad (18)\end{aligned}$$

where V represents the volumetric feed rate to stage 1 and is defined by Equation (5). Note that V is not a state variable but is a function of u_i and x_i, $i = 7, 8, 9$. Note also that k_a and k_b are functions of T through Equation (4).

Next consider the differential changes of ϕ due to differential changes in the control variables

$$\begin{aligned}d\phi = {} & \frac{\partial\phi}{\partial x_f}dx_f + \frac{\partial\phi}{\partial y_f}dy_f + \sum_0^9\frac{\partial\phi}{\partial x_i}dx_i + \sum_0^5\frac{\partial\phi}{\partial y_i}dy_i \\ & + \sum_6^9\frac{\partial\phi}{\partial q_i}dq_i + \sum_7^9\frac{\partial\phi}{\partial u_i}du_i + \sum_1^5\frac{\partial\phi}{\partial T_i}dT_i + \sum_7^9\frac{\partial\phi}{\partial w_i}dw_i \qquad (19)\end{aligned}$$

After the differentiations are carried out, Equation (19) becomes a very complex expression. To save space, Equation (19) has been written in an abstract form. Choose the Lagrange multipliers in Equation (19) so that the coefficients of dx_i, dy_j, dq_k, and du_n; where $i = 0, 1, \ldots, 9$; $j = 0, 1, \ldots, 5$; $k = 6, \ldots, 9$; and $n = 7, 8, 9$; vanish:

$$\mu_1 - \mu_o = 0 \qquad (20)$$

$$-\mu_i + \mu_{i+1} - \mu_it_ik_{ai} + \theta_it_ik_{ai} = 0, \qquad i = 1, \ldots, 4 \qquad (21)$$

$$\begin{aligned}& -\mu_5 - \mu_5t_5k_{a5} + \theta_5t_5k_{a5} + \mu_6\frac{y_5M_xM_y}{(x_5M_x + y_5M_y)^2} \\ & + \theta_6M_xV - \theta_7q_6\frac{y_5M_yM_x}{(x_5M_x + y_5M_y)^2} = 0 \qquad (22)\end{aligned}$$

$$-\mu_6 + \mu_7q_6 = 0 \qquad (23)$$

$$\begin{aligned}& -\lambda_2u_i\frac{\partial x_{Ei}}{\partial x_i} - \mu_iq_i + \mu_{i+1}q_i - \mu_iu_i\frac{\partial x_{Ei}}{\partial x_i} \\ & + \theta_{i+1}q_i\frac{\partial y_i}{\partial x_i} - \theta_iq_i\frac{\partial y_i}{\partial x_i} - \theta_iu_i\frac{\partial y_{Ei}}{\partial x_i} \\ & + \theta_6(x_5M_x + y_5M_y)\frac{\partial V}{\partial x_i} + \left\{\mu_o\left[Vu_i/M_x\frac{\partial x_{Ei}}{\partial x_i}\right.\right. \\ & \left. - \left(x_f + \sum_7^9 u_nx_{En}/M_x\right)\frac{\partial V}{\partial x_i}\right] + \theta_o\left[Vu_i/M_y\frac{\partial y_{Ei}}{\partial x_i}\right. \\ & \left.\left. - \left(y_f + \sum_7^9 u_ny_{En}/M_y\right)\frac{\partial V}{\partial x_i}\right]\right\}\Big/ V^2 = 0, \quad i = 7, 8 \qquad (24)\end{aligned}$$

$$\theta_1 - \theta_o = 0 \qquad (25)$$

$$-\theta_i + \theta_{i+1} - \theta_it_ik_{bi} = 0, \qquad i = 1, \ldots, 4 \qquad (26)$$

$$\begin{aligned}& -\theta_5 - \theta_5t_5k_{b5} \\ & + \theta_6VM_y + (\theta_7q_6 - \mu_6)\frac{x_5M_xM_y}{(x_5M_x + y_5M_y)^2} = 0 \qquad (27)\end{aligned}$$

$$-\theta_6 + \mu_7x_6 + \theta_7\frac{y_5M_y}{x_5M_x + y_5M_y} + \eta_7 = 0 \qquad (28)$$

$$-\mu_ix_i + \mu_{i+1}x_i + \theta_{i+1}y_i - \theta_iy_i - \eta_i + \eta_{i+1} = 0, \quad i = 7, 8 \qquad (29)$$

$$\begin{aligned}& -\lambda_2x_{Ei} - \mu_ix_{Ei} - \theta_iy_{Ei} - \eta_i + \theta_6(x_5M_x + y_5M_y)\frac{\partial V}{\partial u_i} + \left\{\mu_o\left[Vx_{Ei}/M_x - \left(x_f + \sum_7^9 u_nx_{En}/M_x\right)\frac{\partial V}{\partial u_i}\right]\right. \\ & \left. + \theta_o\left[Vy_{Ei}/M_y - \left(y_f + \sum_7^9 u_ny_{En}/M_y\right)\frac{\partial V}{\partial u_i}\right]\right\}\Big/ V^2 = 0 \quad i = 7, 8 \qquad (30)\end{aligned}$$

and at the last stage

$$q_9\frac{\partial y_9}{\partial x_9} - \lambda_2u_9\frac{\partial x_{E9}}{\partial x_9} - \mu_9q_9 - \mu_9u_9\frac{\partial x_{E9}}{\partial x_9}$$

$$- \theta_9 q_9 \frac{\partial y_9}{\partial x_9} - \theta_9 u_9 \frac{\partial y_{E9}}{\partial x_9} + \theta_6 (x_5 M_x + y_5 M_y) \frac{\partial V}{\partial x_9} + \left\{ \mu_o \left[V u_9 / M_x \frac{\partial x_{E9}}{\partial x_9} - \left(x_f + \sum_7^9 u_n x_{En} / M_x \right) \frac{\partial V}{\partial x_9} \right] + \theta_o \left[V u_9 / M_y \frac{\partial y_{E9}}{\partial x_9} - \left(y_f + \sum_7^9 u_n y_{En} / M_y \right) \frac{\partial V}{\partial x_9} \right] \right\} \Big/ V^2 = 0 \quad (31)$$

$$y_9 - \mu_9 x_9 - \theta_9 y_9 - \eta_9 = 0 \quad (32)$$

$$- \lambda_2 x_{E9} - \mu_9 x_{E9} - \theta_9 y_{E9} - \eta_9 + \theta_6 (x_5 M_x + y_5 M_y) \frac{\partial V}{\partial u_9} + \left\{ \mu_o \left[V x_{E9} / M_x - \left(x_f + \sum_7^9 u_n x_{En} / M_x \right) \frac{\partial V}{\partial u_9} \right] + \theta_o \left[V y_{E9} / M_y - \left(y_f + \sum_7^9 u_n y_{En} / M_y \right) \frac{\partial V}{\partial u_9} \right] \right\} \Big/ V^2 = 0 \quad (33)$$

and Equation (19) is reduced to:

$$d\phi = \frac{\mu_o}{V} dx_f + \frac{\theta_o}{V} dy_f + \sum_1^5 \frac{\partial \phi}{\partial T_i} dT_i + \sum_7^9 \frac{\partial \phi}{\partial w_i} dw_i \quad (34)$$

where

$$\frac{\partial \phi}{\partial T_i} = [(\theta_i - \mu_i) k_{ai} x_i E_a - \theta_i k_{bi} y_i E_b] t_i / (R T_i^2), \qquad i = 1, \ldots, 5 \quad (35)$$

$$\frac{\partial \phi}{\partial w_i} = - \lambda_1 + \eta_i, \qquad i = 7, 8, 9 \quad (36)$$

The partial expressions $\partial V/\partial x$ and $\partial V/\partial u$ can be obtained by partial differentiation of Equation (5). The other partial expressions, $\partial y/\partial x$, $\partial y_E/\partial x$, and $\partial x_E/\partial x$ for stages 7 to 9 can be obtained by the use of the equilibrium relations, Equations (14) to (16). Note that y is a state variable in stages 0 to 5 but it is not a state variable in stages 7 to 9.

As shown in an earlier paper (*1*), the Lagrange multiplier Equations (20) to (33) serve the same role here as that of the multiplier rule in the calculus of variations (*15*). The optimality condition is obtained when $\partial\phi/\partial T_i = 0$ and $\partial\phi/\partial w_j = 0$, where $i = 1, \ldots, 5$, and $j = 7, 8, 9$, as no more improvement would be possible. Note that the number of Lagrange multipliers is equal to the number of Lagrange multiplier equations for each stage. Thus, the values of the unknown Lagrange multipliers can be obtained from Equations (20) to (33).

There are various ways for obtaining the optimum condition for this problem from the above set of equations. The most obvious way would be to set Equations (35) and (36) equal to zero and then eliminate the control variables from the process and Lagrange multiplier equations by the use of these two equations. Finally the values for the state variables and the Lagrange multipliers can be obtained from the process and Lagrange multiplier equations. However, there are several difficulties connected with this procedure. Since the control variable T is present implicitly in Equation (35), this variable can not be eliminated easily from the process and Lagrange multiplier equations. Even if T could be conveniently eliminated, there is still the problem of obtaining the solutions for such a complex set of equations. Owing to the presence of recycle, there is no way to solve these equations except by iteration. Another difficulty associated with this approach is that the solutions obtained may not be the true optimum solution.

A more effective way for obtaining the optimum condition is by assuming a set of reasonable values for the control variables and then improving these values by the use of the gradient direction. The sequence $\partial\phi/\partial T_i$ and $\partial\phi/\partial w_j$, where $i = 1, \ldots, 5$ and $j = 7, 8, 9$ is the gradient of ϕ with respect to the control variables, T and w. Hence, for a given change in the control vector $\partial\phi/\partial T_1, \ldots, \partial\phi/\partial T_5, \partial\phi/\partial w_7, \partial\phi/\partial w_8, \partial\phi/\partial w_9$, the greatest change of ϕ will occur if the individual change of each element of this control vector is obtained by

$$\delta T_i = \frac{\dfrac{\partial \phi}{\partial T_i}}{\sum_{n=1}^{5} \left(\dfrac{\partial \phi}{\partial T_n} \right)^2 + \sum_{n=7}^{9} \left(\dfrac{\partial \phi}{\partial w_n} \right)^2} \Delta\phi, \qquad i = 1, \ldots, 5 \quad (37)$$

$$\delta w_i = \frac{\dfrac{\partial \phi}{\partial w_i}}{\sum_{n=1}^{5} \left(\dfrac{\partial \phi}{\partial T_n} \right)^2 + \sum_{n=7}^{9} \left(\dfrac{\partial \phi}{\partial w_n} \right)^2} \Delta\phi, \qquad i = 7, 8, 9 \quad (38)$$

where $\Delta\phi$ is the desired improvement in ϕ. Thus, the improved control variable sequence is

$$T_{i,\,\text{new}} = T_{i,\,\text{old}} + \delta T_i, \quad i = 1, \ldots, 5 \quad (39)$$

$$w_{i,\,\text{new}} = w_{i,\,\text{old}} + \delta w_i, \quad i = 7, 8, 9 \quad (40)$$

If there were no recycle, the state variables can be obtained from the process equations in a forward recursive fashion with assumed values of the control variables. The Lagrange multipliers can then be obtained by solving the Lagrange multiplier equations in a backward recursive fashion. However, due to the presence of recycle, the terms $\Sigma u_n x_{En}$ and $\Sigma u_n y_{En}$ are unknowns in the process Equations (6), (12), and (13) during the forward recursive calculations. In order to solve the process equations the values of these two unknowns must be assumed. Examination of Equations (20) to (33) for the Lagrange multipliers will reveal that the multipliers μ_o, θ_o, and θ_6 are unknowns in Equations (24), (30), (31), and (33) during the backward recursive calculations. Thus the values for these unknowns also must be assumed. To summarize, this optimization problem can be solved by the following procedure:

1. Assume a reasonable set of values for the control variables T and w.
2. Assume reasonable values for $\Sigma u_n x_{En}$ and $\Sigma u_n y_{En}$.
3. Obtain the state variables starting from stage 0 in a forward recursive fashion until stage 9.
4. Obtain improved values for $\Sigma u_n x_{En}$ and $\Sigma u_n y_{En}$.
5. Repeat steps 3 and 4 until the desired accuracy on $\Sigma u_n x_{En}$ and $\Sigma u_n y_{En}$ is obtained.
6. Assume reasonable values for μ_o and θ_o.
7. Assume reasonable value for θ_6.
8. Obtain the Lagrange multipliers starting from stage 9 in a backward recursive fashion until stage 7.
9. Obtain θ_6 and μ_6 by solving Equations (23) and (28).
10. Repeat steps 8 and 9 until the desired accuracy for θ_6 is obtained.
11. Obtain the Lagrange multipliers starting from stage 5 in a backward recursive fashion until stage 1.

12. Obtain μ_o and θ_o by solving Equations (20) and (25).

13. Repeat steps 8 to 12 until the desired accuracy for μ_o and θ_o is obtained.

14. Obtain improved values for the control variables T and w by the use of Equations (35) to (40).

15. Repeat steps 3 to 5 and 8 to 14 until $\sum_{1}^{5} (\partial\phi/\partial T_n)^2 + \sum_{7}^{9} (\partial\phi/\partial w_n)^2$ becomes so small that further improvement is not significant.

The above procedure is essentially a modification of the procedure listed in an earlier paper (*1*). This modification is made to take care of the five unknowns in the above equations. Although there are many different iteration loops, we shall see that the numerical solution of this problem is not very difficult. All the iterations involved converge fairly fast even if the initially assumed values are quite different from the correct ones. Note that steps 1, 2, 6, and 7 are needed only for the first iteration. Instead of assuming these values, the values from the previous iteration are used for second and latter iterations.

In solving Equations (9) to (11) and (14) to (16) in step 3, some kind of iteration has to be used. This is because the following equation is in implicit form

$$x = f(w, x) \tag{41}$$

This equation is obtained by eliminating q, u, y, x_E, and y_E from Equations (9) to (11) and (14) to (16). With known values for w, iteration must be used to obtain the value of x. The Newton-Raphson method is used alternately with a simple iteration scheme. First, a few iterations are calculated by the Newton-Raphson method. If no convergence is obtained, another number of iterations is then calculated by the simple iteration scheme. This alternation is continued until the desired accuracy of x is obtained. With this alternative scheme, no convergence problem has been encountered in solving Equation (41).

Detailed discussion for the simple iteration scheme can be found in any numerical analysis book. With an assumed value of x, the improved x can be obtained from Equation (41). It has been found that a faster convergence rate can be obtained if the value of x, used for the next iteration, is the average value of the assumed and the improved values of x.

The desired accuracy, as has been used in the above computation procedure, can be defined as follows:

$$|z_{i+1} - z_i| \leq \epsilon_z \tag{42}$$

where ϵ_z is the maximum error allowed for the variable z and the subscript i denotes the number of iteration. The variable z represents any of the unknowns such as x in Equation (41), $\Sigma u_n x_{En}$ or $\Sigma u_n y_{En}$ in step 4, θ_6 in step 10, or μ_o or θ_o in step 13.

The equilibrium relationships which are usually given in tabular form, can be correlated into the following equations by linear regression:

$$y = a + bx + cx^2 \tag{43}$$

$$y_E = d + ex + fx^2 \tag{44}$$

$$x_E = g + hx + ix^2 + jx^3 \tag{45}$$

which correspond to Equations (14) to (16). These equations can be used only within the range for which they are correlated. They can not be used in the completely miscible region of the three components, as shown in Figure 2. The reason for correlating the equilibrium data into the above equations is to avoid numerical differentiation which is known to be inaccurate.

This problem has been solved numerically. The equilibrium diagram shown in Figure 2 has been correlated into Equations (43) to (45). The coefficients obtained for these equations are:

$a = 0.9865$
$b = -0.8295$
$c = -1.2203$
$d = 0.009$
$e = -0.1156$
$f = 2.6043$
$g = 0.0150$
$h = 4.1882$
$i = -11.23$
$j = 8.481$

The equilibrium data used are completely the same as that used in an earlier paper (*1*).

The other numerical values used are:

$\lambda_1 = 0.1$
$\lambda_2 = 0.1$
$M_x = 33.5$
$M_y = 31.5$
$x_f = 0.95$
$y_f = 0.05$
$V_x = 1.0$
$V_y = 1.0$
$G_a = 0.535 \times 10^{11}$
$G_b = 0.461 \times 10^{18}$
$E_a = 18{,}000$
$E_b = 30{,}000$
$R = 2$
$t_i = 6, \quad i = 1, \ldots, 5$

This problem has been solved on an IBM 7094 computer. The results are shown in Table 1. The maximum value of ϕ, $\phi_{\max}$, obtained is 24.767. For the first iteration, the following values have been assumed for steps 2, 6, and 7:

$\Sigma u_n x_{En} = 3$
$\Sigma u_n y_{En} = 3$
$\mu_o = 25$
$\theta_o = 25$
$\theta_6 = 0.008$

Note that these values are quite different from the final results. However, even with these rough starting values, steps 3 and 4 need to be repeated but nine times, steps 8 and 9 nine times, and steps 8 to 12 five times for the first iteration to obtain the following accuracy:

$\epsilon_{\mu_o} = 0.1 \times 10^{-4}$
$\epsilon_{\theta_o} = 0.1 \times 10^{-4}$
$\epsilon_{\theta_6} = 0.1 \times 10^{-5}$
$\epsilon_{\Sigma u_n x_{En}} = 0.1 \times 10^{-4}$
$\epsilon_{\Sigma u_n y_{En}} = 0.1 \times 10^{-4}$

TABLE 1. OPTIMUM CONDITIONS WITH RECYCLE

Stage	Control variable T or w	State Variables x	y	q	u	Lagrange Multipliers μ	θ	η
0		0.937	0.0629			32.5	34.9	
1	327.12	0.688	0.301			32.5	34.9	
2	325.45	0.524	0.454			31.7	36.1	
3	324.41	0.407	0.558			30.3	37.1	
4	323.70	0.321	0.631			28.4	37.9	
5	323.19	0.255	0.684			25.8	38.7	
6	—	0.284	0.716	39.7		16.2	0.824	
7	6.998	0.166	0.815	33.9	12.7	0.408	0.849	0.1
8	4.72	0.099	0.892	30.8	7.85	0.330	0.865	0.1
9	4.82	0.057	0.935	29.3	6.31	0.179	0.882	0.1

The number of iterations for steps 3 and 4, 8 and 9, and 8 to 12 are reduced tremendously for the latter iterations. The initially assumed values for the control variables for step 1 are

$$T_i = 325, \qquad i = 1, \ldots, 5$$
$$w_i = 10, \qquad i = 7, 8, 9$$

With these initially assumed values and with $\Delta\phi = 0.005$, the results listed in Table 1 have been obtained in approximately 0.3 min. This computation time can be greatly reduced if a program with less print out were to be used.

In carrying out step 3, Equation (41) is solved by using the following initially assumed value for x:

$$x = 0.1$$

The accuracy required for x is $\epsilon_x = 0.1 \times 10^{-5}$.

To show that the maximum point obtained is global, various different starting values for the control variables have been assumed. The above calculations are repeated with these different sets of starting values. It has been found that the same maximum point is obtained in all these calculations.

Without Penalty for Recycle

It is interesting to examine the influence of the penalty function, λ_2, on the optimum results. The same problem discussed above is solved without any penalty for recycle. This is done by setting $\lambda_2 = 0$. All other numerical values remain the same. The optimum results are shown in Table 2. The maximum profit obtained is $\phi_{\max} = 25.959$. The starting values used for this calculation are

$$T_i = 329, \qquad i = 1, \ldots, 5$$
$$w_i = 4, \qquad i = 7, 8, 9$$

The value of ϕ at these starting values is $\phi = 24$
In less than 0.2 min., the optimum results listed in Table 2 are obtained from these starting values.

TABLE 2. OPTIMUM CONDITIONS WITHOUT PEALTY FOR RECYCLE

Stage	Control Variable T or w	State Variables x	y	q	u	Lagrange Multipliers μ	θ	η
0		0.922	0.077			39.6	41.6	
1	323.9	0.723	0.271			39.6	41.6	
2	322.3	0.584	0.404			39.0	42.5	
3	321.2	0.479	0.501			38.2	43.2	
4	320.5	0.398	0.575			37.1	43.9	
5	319.9	0.333	0.632			35.7	44.4	
6	—	0.359	0.641	47.5		26.8	0.837	
7	10.48	0.194	0.779	36.8	21.2	0.564	0.833	0.1
8	6.29	0.105	0.886	32.0	11.0	0.481	0.854	0.1
9	6.55	0.0536	0.939	30.1	8.45	0.293	0.877	0.1

DISCUSSION

A fairly effective technique for obtaining numerical solutions of complex optimization problems has been obtained by the combined use of the gradient technique and Lagrange multiplier. The advantage of this approach is its ability to handle nearly all types of complex topological situations. The interconnections between sections are handled in a very natural way.

Although only one control variable is assumed for each stage, processes with more than one control variable in each stage can also be treated in essentially the same way. However, it should be emphasized that the computational difficulties increase rather rapidly with the increase of the number of control variables.

For the purpose of simplicity, only crosscurrent extraction is used for the separation stages. Obviously, other separation equipment, such as distillation columns, can be treated essentially in the same way. Countercurrent processes can also be handled. Processes with both continuous and stagewise operating units can be treated by the combined use of the approaches outlined in this paper and an earlier paper (*1*).

The approach can handle almost any optimization criterion. Frequently, this criterion includes the terminal values of certain state variables, the operating cost of each stage, and the cost of recycle. The terminal value of a state variable is the value of the state variable which does not need to be processed further. It does not necessarily mean the value of the state variable in the last stage. The very general criterion in which the function, ϕ, is a function of all the state variables and all the control variables in all the stages can also be handled.

ACKNOWLEDGMENT

This work was partly supported by the Office of Water Resources Research, U.S. Dept. of the Interior Grant No. 14-01-0001-1962. Part of the work was done while the author was with Phillips Petroleum Company.

NOTATION

E_a, E_b = activation energies of reactions
G_a, G_b = frequency factor constants in Arrhenius equations
k_a, k_b = reaction rate constants
M_x, M_y = molecular weights
q = flow rate of raffinate stream
R = gas constant
T = temperature
t = residence time
u = flow rate of extract stream
V = a variable defined by Equation (5)
V_x, V_y = moles per unit volume of components x and y
W = extracting solvent
w = feed rate of extracting solvent
x, y, z = concentrations in raffinate stream
x_E, y_E, z_E = concentrations in extract stream
x_f, y_f = feed concentrations
$a, b, \ldots, j$ = constants
$\Delta\phi$ = value controlling step size in the gradient direction iteration
ϕ = profit function
ϵ_z = maximum error allowed for the variable z
λ_1 = relative cost of extracting solvent
λ_2 = relative recycle cost
μ, θ, η = Lagrange multipliers

LITERATURE CITED

1. Lee, E. S., *Ind. Eng. Chem. Fundamentals*, **3**, 373 (1964).
2. Bellman, R., and S. E. Dreyfus, "Applied Dynamic Programming," Princeton Univ. Press, New Jersey (1962).
3. Fan, L. T., "The Continuous Maximum Principle," John Wiley, New York (1966).
4. Courant, R., *Bull. Am. Math. Soc.*, **49**, 1 (1943).
5. Tompkins, C. B., "Methods of Steep descent," *Modern Mathematics for the Engineer*, E. F. Beckenbach, ed., Chap. 18, McGraw-Hill, New York (1956).
6. Box, G. E. P., and K. B. Wilson, *J. Roy. Statist. Soc. Ser.*, B13, 1 (1951).
7. Fletcher, R., and C. M. Reeves, *Computer J.*, **7**, 149 (1964).
8. Arrow, K. J., L. Hurwicz, and H. Uzawa, "Studies in Linear and Nonlinear Programming," Stanford Univ. Press, Stanford, Calif. (1958).
9. Wolfe, P., "Methods of Nonlinear Programming," *Recent Advances in Mathematical Programming*, R. L. Graves and P. Wolfe, ed., McGraw-Hill, New York (1963).
10. Rosen, J. B., *J. Soc. Ind. Appl. Math.*, **8**, 181 (1960), **9**, 514 (1961).
11. Hestenes, M. R. and E. Stiefel, *J. Res. Natl. Bur. Std.*, **49**, 409 (1952).
12. Kelley, H. J., *Am. Rocket Soc. J.*, **30**, 947 (1960).
13. ———, "Methods of Gradients," *Optimization Techniques*, G. Leitman, ed., Academic Press, New York (1962).
14. Bryson, A. E. and W. F. Denham, *J. Appl. Mech.*, **29**, 247 (1962).
15. Bliss, G. A., "Lectures on the Calculus of Variations," Univ. of Chicago Press, Illinois (1946).
16. Aris, R., A. L. Nemhauser, and D. J. Wilde, *AIChE J.*, **10**, 913 (1964).
17. Niederlinski, A., *ibid.*, **12**, 805 (1966).

Manuscript received November 29, 1967; revision received March 25, 1968; paper accepted March 27, 1968. Paper presented at AIChE St. Louis meeting.

The Effects of Internal Mass Transfer on the Hydrogenation of Benzene Over Nickel-Alumina Catalyst

FRANTISEK JIRACEK, JOSEF HORAK, and JOSEF PASEK

Institute of Chemical Technology, Prague, Czechoslovakia

The influence of mass transfer on reaction kinetics, within the porous structure of a catalyst, was investigated during the hydrogenation of benzene. The temperature range for a nickel catalyst at atmospheric pressure was 99 to 160°C. The rate equation and the apparent activation energy on a cylindrical catalyst pellet (4.5 mm. in diameter and 5.5 mm. in height) and on the catalyst particles (0.5 to 0.63 mm. in diameter) were evaluated.

The effect of internal diffusion on the dependence of the reaction rate, on the temperature, and on the partial pressures of benzene and hydrogen was satisfactorily explained by means of theoretical relations. The values of the effective diffusion coefficient of benzene were compared by calculating the kinetic data and temperature difference in the cylindrical pellet.

Many papers have been published relating to the theoretical effect of internal diffusion on the rate of a catalytic reaction. Few papers, however, aim at verifying such theories. The object of this paper (*4*) is to verify the theory of mass transfer effect on the reaction order within the porous structure of a catalyst, that is, on the rate equation and on the temperature dependence of the catalytic reaction rate. Furthermore, a kinetic measurement was made of the effective diffusion coefficient of a reaction component in the porous structure and its value was compared with that of the temperature difference in a catalyst pellet.

EXPERIMENTAL PROCEDURE

Reaction Used

The reaction chosen for this study was the hydrogenation of benzene on a nickel catalyst in the benzene-hydrogen-nitrogen mixture at atmospheric pressure and 99 to 160°C. The partial pressure in the inlet stream varied from 22 to 93 Torr for benzene and 100 to 700 Torr for hydrogen. The conversion of benzene to cyclohexane in the outlet stream varied from 5 to 95%.

Catalyst

An industrial nickel-alumina catalyst was used containing 49% nickel and 4.3% graphite by weight (*1*). The specific area per unit mass was 256 sq.m./g., the apparent pellet density was 1.17 g./cc., the mean solid density was 3.84 g./cc., the micropore volume (pore size below 150 Å.) was 0.308 cc./g., the macropore volume (pore size above 150 Å.) was 0.288 cc./g., and the thermal conductivity was 3.9×10^{-4} cal./(cm.)(sec.)(°C.). A mean pore radius of 50Å. was determined from the specific surface, the apparent pellet density, and the mean density of a solid phase catalyst.

Before carrying out the measurements, the catalyst was activated by heating it in a hydrogen stream (20 liters/hr.) to 180°C. for 3 hr. The effective thermal conductivity of the catalyst pellet was determined experimentally as described by Masamune and Smith (*6*).

Apparatus

A schematic diagram of the experimental apparatus is shown in Figure 1. Kinetic measurements were carried out in a recycle reactor. The hydrogen and nitrogen were metered by flow meters (*2*), the mixed and combined feed streams were purified in a series of columns. The first column was purified by phosphoric acid on a kieselguhr to absorb alkaline impurities, in the second column a copper catalyst was used