Ind. Eng. Chem. Fundam. 1985, 24, 153–170

Structure of Complex Catalytic Reactions: Thermodynamic Constraints in Kinetic Modeling and Catalyst Evaluation

Reuel Shinnar* and Cheng A. Feng

Department of Chemical Engineering, The City College of New York, The City University of New York, New York, New York 10031

The paper examines the thermodynamic constraints on chemical reaction trajectories. It shows how thermodynamic concepts can be used to organize and analyze the results of kinetic studies in complex reaction systems where several reactions can occur simultaneously. It defines rigorous criteria for setting up a set of stoichiometric relations to obtain an empirical kinetic model for the system. This can be done without any kinetic calculations by inspecting the measured trajectories in composition space. The paper also defines the concept of coupling between overall ostensible reactions and explains the thermodynamic advantages by using shape-selective catalysts. The concept of thermodynamic constraint is defined, and it is shown that the mechanism of the kinetic reactions has thermodynamic consequences far more restrictive than the Second Law itself. The results should be useful in the modeling of complex reactions systems as well as in the testing and evaluation of new catalysts.

I. Introduction

The Second Law of thermodynamics imposes many limits on the performance of chemical reactors. Strictly speaking, the only thermodynamic constraint on any reactor that is a direct result of the law is that the overall feee energy change is negative definite. However, available catalysts impose much stronger constraints on the reachable set of compositions or composition spaces. The free energy must not only decrease for the total system, but it must also decrease separately for each kinetically independent ostensible reaction. Despite great interest in understanding the constraints that these considerations place on reactor performance and design, the subject remains only partly solved (Krambeck, 1970; Wei and Prater, 1962; Aris, 1969; Sellers, 1967; Bowen, 1968). In this paper a systematic approach is presented which enables one to apply thermodynamic constraints to reactor design and performance.

The thermodynamic constraints can be classed into three types: (1) thermodynamic constraints which are the results of stoichiometric relations including enthalpy and free energy relations; these are defined as "hard" constraints; (2) thermodynamic consequences of the kinetic properties of available catalysts; (3) thermodynamic consequences of design decisions.

An example of a thermodynamic constraint which is a consequence of a design decision is the limited thermal efficiency of a fuel-burning, steam-fired power plant. The limited efficiency is not a direct result of the second law but a result of the design decision to limit the maximum temperature of the superheated steam. An example of a thermodynamic constraint which is a consequence of available technology is the production of methanol by the reaction

$$CO + 2H_2 \rightleftharpoons CH_3OH \qquad (1)$$

Since there is no available catalyst that promotes the reaction at room temperature, the reactor is forced to operate at a higher temperature, commonly 520 °F. The constraint is dictated by the properties of available catalysts and by other design considerations.

In each of the examples, the constraints are not a direct result of the second law. However, we can use the second law to analyze the consequences of these imposed constraints. The advantage of using thermodynamic relations instead of rate equations to analyze chemical reactions is that we do not need exact kinetic data. It is sufficient to know or be able to estimate the relative magnitudes of the reaction rates and the range of temperatures over which these relations apply.

The purpose of this article is to define such a thermodynamic approach and to demonstrate its applicability to a wide set of problems. This approach can also be used to obtain performance bounds for complex multiple reactions in catalytic reactor design. As a result it can provide guidelines for rational process development.

The paper is divided into several parts. Section II discusses principles of modeling chemical reactors and defines some of the terminology which will be used later. In section III this approach will be outlined by discussion of first-order systems similar to the ones considered by Wei and Prater (1962). Sections IV and V extend the analysis to nonlinear systems and show how the set of reachable composition depends on the mechanism and upon the intermediates of the individual chemical reactions. Section VI shows how this can be used to define rigorously a set of overall chemical reactions, each of which is both thermodynamically and kinetically independent of the other reactions, and to test their consistency with experimental data. In section VII we discuss the simple nonlinear example of steam gasification of carbon. Finally, the last part of this paper deals with the effect of transport properties on thermodynamic constraints. We will discuss shape selective catalysis and evaluate the impact of shape selectivity. It will be shown that introducing shape selectivity is equivalent to introducing additional kinetic pathways and can achieve results that are similar to those obtained by introducing new reactions (or in other words, new catalytic functions to a catalyst surface).

II. Definition of the Problem and Overview of the Method

In this section we address some of the problems encountered in modeling a complex reaction. In most cases, one does not model molecular events. Rather, one tries to describe complex events involving intermediates by a simplified overall model. To model a chemical reactor, we do not need to know all the individual steps of each reaction. A phenomenological approach using overall reactions is often adequate unless one is dealing with dynamic processes which are outside the scope of this paper. It is fortunate that one can lump intermediate steps together

0196-4313/85/1024-0153$01.50/0

because often these steps are unknown and may be hard to elucidate. The question, however, remains: how does one set up a set of overall phenomenological reactions if several reactions occur simultaneously which cannot be studied independently? If one can identify all the compounds involved (neglecting free radicals and all intermediates which are not present in measurable concentrations in the bulk phase), one can always set up a set of independent chemical reactions describing all the changes (Aris, 1969). These independent reactions correspond to a set of vectors in species space which are linearly independent and span the vector space of all the observed chemical changes.

Aris (1969) has shown that if there are N species and E independent element balances, the number of stoichiometrically independent reactions S is

$$S = N - E \tag{2}$$

Each of these sets of stoichiometrically independent reactions can describe all stoichiometric changes and is useful in computing heat and free energy balances and allows us to compute global equilibrium. However, these sets of linear independent stoichiometric relations are not unique, nor are they necessarily suitable for deriving kinetic rate equations. They are also not equivalent when one proposes phenomenological models to describe this system. One can realize this by just looking at a very simple system, such as the isomerization of three xylene components. Stoichiometrically, there are only two linearly independent reactions. However, intuitively we realize that at least three overall reaction expressions are needed in order to model the system. This result can be obtained without using any time-dependent dynamic relations if one analyzes the changes in composition space or trajectories. We define a *trajectory* as a set of compositions that occur, as a result of reactions, as a function of time. Such a set of compositions or trajectory has properties independent of time and it is these properties on which this paper will focus. The fact that some of the properties of the measured trajectory are time independent is very crucial to the approach, as it allows us to use thermodynamic relations. For example, such a trajectory must have the property that free energy decreases along it. We will postulate that if such a trajectory is the result of several reactions occurring simultaneously *each* reaction contributing to the change in composition must have a decreasing free energy. It is this condition which prevents one from describing the changes in the xylene isomerization system by two reactions, even if one is not concerned about their dynamic form.

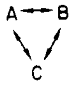
The paper outlines a method for obtaining suitable sets of chemical reactions for the study of rate equations and a simple way of verifying if these sets are in agreement with all known trajectories. All trajectories should be realizable from this set of reactions without violating the free energy constraint. It will be shown that the minimum number of reactions required to describe a complex reaction system can be larger than the $N - E$, which is the number of stoichiometrically independent reactions. A set of reactions that fulfill these requirements is not necessarily unique, and all one can conclude is whether or not a certain postulated set is consistent with the data or not. We will also show that often the smallest set of reactions which is consistent with the data requires more reactions than stoichiometric independence allows. Stoichiometric independence (Aris, 1969) and kinetic independence are two separate concepts, as will be explained in the paper. The term "kinetically independent reaction" is used here for an observable overall reaction that is made up of individual elementary steps and occurs independently of all other observable reactions that occur simultaneously. A simple method for efficiently choosing a set of overall kinetic equations suitable for reactor modeling will be presented later.

Define a "composition space" as the set of all possible compositions. This space does not have all the properties of a formal vector space but is a useful construct. The previous results have two important implications for the reaction engineer and the catalyst chemist. The first is that if one knows the set of reactions that a catalyst promotes, one can predict the entire set of compositions reachable from a given starting position, independent of the magnitudes of the reaction rates. This bounds the selectivity and conversion achievable by improvements of the catalyst, if these improvements only change the relative magnitudes of reaction rates without a change in the mechanism that the catalyst promotes. If an improved catalyst leads to compositions which are still in the space accessible by the reactions promoted by existing catalysts, then the improved catalyst does not promote new independent reactions. Conversely, if a new catalyst shows that it can reach a composition forbidden by the known reaction mechanism of an available catalyst, it merits further investigation even if the preliminary results are not economically promising. Thus, such a procedure provides a screening tool for evaluating new catalysts. This analysis can be done using only thermodynamic data, by reasonable computations. A large part of the paper will therefore discuss how to compute the set of compositions reachable from a specified initial condition given a set of overall reactions.

III. Bounds on the Reachable Composition Space for Monomolecular Reaction

We begin by analyzing a monomolecular reaction system involving three species. Such a simple system allows us to present the basic concepts of our method in a simple graphical way. Later we will extend it to more general cases.

For linear systems the properties of the system have been studied in detail by Wei and Prater (1962). The fact that a reaction system is monomolecular in the sense that all reactions can be expressed in the form $A_i \rightleftharpoons A_j$ does not imply that the reaction rates are a linear function of concentration. Our approach will not involve any assumptions of linearity. For simplicity, we start with the linear approach of Wei and Prater. Consider the monomolecular system.



(3)

As the system has three components and the reactions are all monomolecular, there is only one independent element balance. Therefore, it follows from eq 2 that there are only two stoichiometrically independent reactions. However, it is intuitively clear that, at times, three reactions are needed to describe the kinetic behavior of the system (see Wei and Prater, 1962).

In our first example, we assume that the equilibrium constants of all three reactions are equal (see Figure 1a). One can then ask what constraints are imposed by the second law on the reachable composition space. In this system the total number of moles remains constant. Therefore, the Helmholz and Gibbs free energies are equal to each other. The example was intentionally chosen such that pure A, B, and C have the same molar free energy and all points inside the triangle ABC have a free energy less

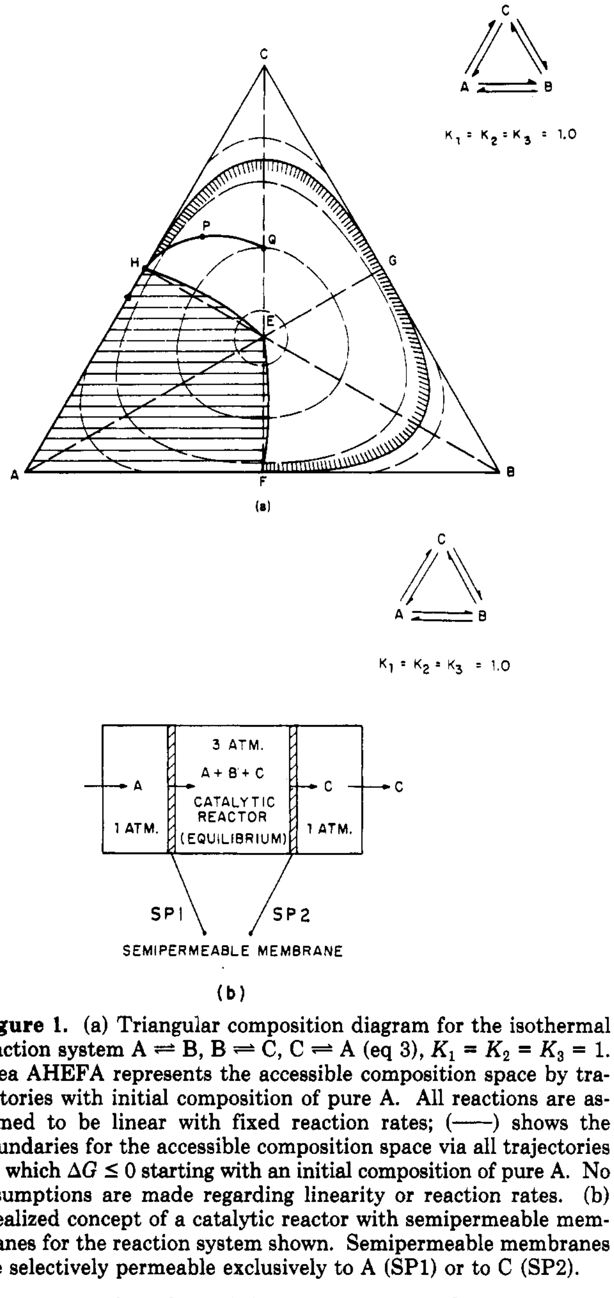


Figure 1. (a) Triangular composition diagram for the isothermal reaction system $A \rightleftharpoons B$, $B \rightleftharpoons C$, $C \rightleftharpoons A$ (eq 3), $K_1 = K_2 = K_3 = 1$. Area AHEFA represents the accessible composition space by trajectories with initial composition of pure A. All reactions are assumed to be linear with fixed reaction rates; (——) shows the boundaries for the accessible composition space via all trajectories for which $\Delta G \leq 0$ starting with an initial composition of pure A. No assumptions are made regarding linearity or reaction rates. (b) Idealized concept of a catalytic reactor with semipermeable membranes for the reaction system shown. Semipermeable membranes are selectively permeable exclusively to A (SP1) or to C (SP2).

than or equal to that of the pure compound. Thus, if the initial composition is pure A, then there is no thermodynamic constraint to go to pure B or C at constant temperature and pressure.

Every composition point in Figure 1a is in principle reachable starting from pure A. However, if we limit ourselves to an isothermal chemical reactor, we impose an additional constraint. All reaction trajectories must have a decreasing free energy along the trajectory. In this example, there is no trajectory leading from A to B in which the free energy stays constant or monotonically decreases. This is shown in Figure 1a where the lines of constant free energy are plotted. All possible trajectories from A to B have a definite minimum. Therefore B is outside the accessible composition space from A. Using the iso-free energy lines, one can easily check which compositions can be reached from A, via a trajectory which satisfies the requirement that along the trajectory the free energy change is always negative definite. All reachable compositions are inside the area bounded by the curve AHGFA in Figure 1a. (In computing this bound it was assumed that the change in free energy, while negative, could be infinitesimally small.)

The limitations on the reachable composition space are not direct consequences of the second law. They are thermodynamic consequences of the design decision to use an isothermal reactor without any selective membranes. In theory, one can devise a hypothetical reactor which will force pure A to go to pure C without violating thermodynamic constraints. Consider the hypothetical reactor illustrated in Figure 1b, in which the feed enters the reactor through a selective membrane permeable only to A and the product is removed through another selective membrane permeable only to C. The reactor itself is never more than infinitesimally displaced from its global equilibrium. Inside the reactor the total pressure (or molar concentration) is larger than that of the feed or product, but, since the fugacity (or partial pressure) of A is differentially smaller inside the reactor than in the feed, this does not cause any problem for this hypothetical reversible reactor. If the partial pressure of the product inside the reactor is slightly smaller than the pressure of pure A in the feed, one can provide a small driving force to fulfill the demand that the free energy change should be negative definite. For real life such selective membranes are not usually available. In the absence of the membrane, constraints are imposed by the stoichiometric relations in the system which are independent of catalyst properties. The most important of these is the requirement that the free energy along any trajectory must be negative definite. If the reactions are really linear kinetic reactions, the reachable composition space is much smaller than AHGFA in Figure 1a.

The requirement for detailed balancing at equilibrium imposes an additional and more restrictive constraint. In the case of a monomolecular reaction which implies the linearity of the reaction rate expressions, detailed balancing also holds far from equilibrium. Therefore, for the linear case all the reaction rates r_i can be written in the form

$$r_{A\rightleftharpoons B} = k([A] - [B]/K_{AB}) \qquad (3a)$$

where K_{AB} is the equilibrium constant of the reaction $A \rightleftharpoons B$ (see Wei and Prater, 1962). If one now imposes on the reactor system the further constraint of a catalyst which has uniform properties and therefore the same rate constants for all compositions, there are even more severe restrictions on the accessible composition space (see Faith and Vermeulen, 1967).

Consider the monomolecular system in eq 3 with the above constraints. If one starts with pure compound A, then the accessible composition space for any reaction rate combination is the shaded area AHEFA shown in Figure 1a. The bounds AHEFA are derived as follows. In this system there are assumed to be three independent first-order reactions. As the backward reactions are related to the forward reactions by eq 3a, there are only three independent rate constants. We are interested only in the trajectories (or set of compositions) that result from these reactions. However, the trajectory in the composition space remains constant if all reaction rates are multiplied by an arbitrary constant. We can therefore designate one of the reaction rates as unity and change the two others over all possible values. Each set of three rate constants yields a specific trajectory. When this is done, all possible values of the rate constants and all compositions that can be reached by any trajectory from A lie in the region AHEFA. The detailed computations for this system are given in Faith and Vermeulen (1967). Krambeck (1982) has developed a simple method to obtain such limits for linear reactions of monomolecular systems of arbitrary species.

The boundary AHGFA is not a direct result of the

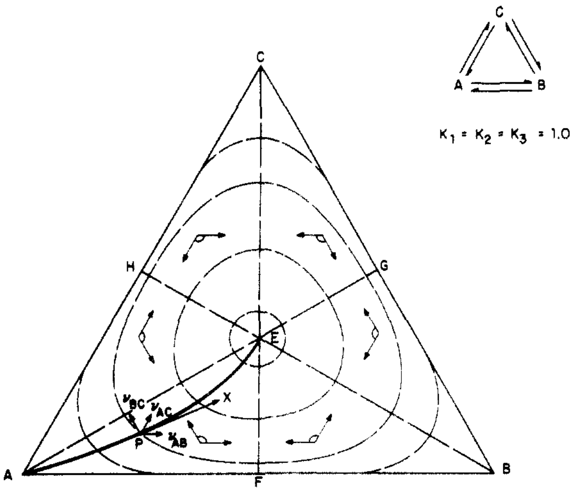


Figure 2. Thermodynamic limitations on kinetic trajectories for the isothermal monomolecular reaction system shown. Each individual reaction separately fulfills the conditions of decreasing free energy: (E) global equilibrium of system; (APE) hypothetical trajectory from pure A to equilibrium.

second law. It is derived from assuming that, in this system, there are no selective semipermeable membranes and the catalyst is uniformly accessible by reactants and products. This will be called a "stoichiometric constraint". Note that the constraint is more severe than merely requiring $\Delta G \leq 0$, as the isocline of free energy passing through A extends beyong the triangle area. However, the reachable composition space is much larger than the shaded area AHEFA which was derived assuming fixed reaction rates. Let us now examine the reasons for the differences in the accessible composition space and consider how the larger permitted space can be reached.

We begin by removing the requirement that the ratio between all reaction rates must remain constant. Assume that one could find different catalysts that allow one to vary the relative magnitudes of the reaction rates at will. This catalyst is then placed in a prescribed order along the length of an isothermal plug flow reactor. For this case a rigorous limit on permissible reaction trajectories can be derived without invoking the concept of detailed balancing by demanding that the criterion that free energy changes must be negative will be fulfilled for each individual reaction. Thus, the reaction vector

$$A_i \rightleftharpoons A_j \tag{4}$$

must satisfy these criteria at each point in the trajectory. Moreover, no trajectory following along this vector can pass any point at which this reaction is in equilibrium. The vector A → B cannot pass the line CF which forms an attractor for the reaction vector A → B, regardless of starting composition (see Figure 2). For a first-order reaction system in which eq 3 holds, this simply means that all rate coefficients are positive. Along CF the reaction $A \rightleftharpoons B$ is in equilibrium. If only the reaction $A \rightleftharpoons B$ occurs and one starts at an arbitrary composition, either to the left or right of this line, one will always reach a composition very close to the line CF.

Consider now the case of an experimentally measured trajectory of composition changes as given by the line APE in Figure 2 and one looks at the composition P on this trajectory. The tangent to the trajectory is a vector that gives the local change of composition. The direction can be expressed independently of the rate at which the change occurs (Aris and Mah, 1963; Aris, 1969). Assume that there is a set of chemical reactions involving N compounds. If one examines any composition based on these N compounds, then the number of moles for each compound A_j form a vector $\mathbf{N}_j$ (j = 1, 2, ..., N). Any stoichiometric relation can then be expressed as

$$\sum_{j=1}^{N} \nu_j A_j = 0 \tag{5a}$$

where ν_j is the coefficient of species A_j in the stoichiometric relation. The sign of ν_j is used to designate the coefficient of products as positive and reactants as negative. If a single reaction occurs an initial composition defined by a vector $\mathbf{N}_{j0}$ changes to another composition $\mathbf{N}_j$. The reaction expression 5a defines the direction of the vector that connects the two composition vectors. If there are R reactions, then (5a) can be rewritten in the form

$$\sum_{j=1}^{N} \nu_{ij} A_j = 0; \quad (i = 1, 2, ..., R) \tag{5b}$$

where the free index i indicates the number of the reaction chosen. If the set of reactions is chosen such that they are linearly independent, then R is equivalent to S in eq 2. However, eq 5b is valid even if the equations are not all linearly independent. If one starts with a given initial composition which contains N_{j0} moles of compound A_j, we can then express any composition resulting from this initial composition

$$\mathbf{N}_j = \mathbf{N}_{j0} + \sum_{i=1}^{R} \nu_{ij} \xi_i \tag{6}$$

where ξ_i is the relative magnitude of the conversion for each of the reactions. (ξ_i could be either positive or negative depending on how one writes the stoichiometric reaction.) The set of $\nu_{ij}\xi_i$ again forms a vector that connects the vector $\mathbf{N}_{j0}$ with the composition vector $\mathbf{N}_j$. For purely stoichiometric descriptions one chooses a set of reactions for which $S = N - E$, but for kinetic purposes we can choose a larger R and eq 6 will still be valid. The only difference is that if the dimension of ξ_i is equal the number of reactions, equal to $N - E$ then the vectors $\nu_{ij}\xi_i$ are linearly independent of each other. In case the number of reactions R is greater than S, this is not true. Each $\nu_{ij}\xi_i$ will be a vector that represents the composition change due to the reaction i. The number of actual reactions that can occur is not limited by any considerations of linear independence. Thus there is no limit to the number of reactions that enter into the vector $\nu_{ij}\xi_i$. Linear independence is only a convenient tool to describe stoichiometric changes and compute global equilibrium. It does not enter the choice of reactions if one wants to describe physical changes by observable overall reactions.

Most chemical reactions consist of a number of intermediate steps involving free radicals or, in case of catalytic reactions, adsorbed species. We often do not observe them and are content to describe the system by overall kinetic relations, the stoichiometry of which can be described by an expression such as eq 5a. We call a reaction kinetically independent of it fulfills two criteria: (a) It occurs independent of the other reactions in the sense that the only way the other reactions affect it is by removing or adding to the total system species involved in the reaction. It does not matter that other species not appearing in (5a) catalyze it or inhibit it. We require only that no other reaction has to occur simultaneously to promote it. (b) It has to fulfill the criterion that $\Delta G \leq 0$ separately for the vector $\nu_{ij}\xi_i$. This is a direct result of assumption (a). A more rigorous definition of a kinetic independent set of reactions will be given later.

For our model system, $S = 2$ as $N = 3$ and $E = 1$. For the linear case it is obvious that kinetically we can take three independent kinetic reactions ($R = 3$). The fact that R is larger than S does not change the way in which we use the system of such equations to describe a change in composition. If R is larger than $(N - E)$, the decomposition of the vector into the individual reaction vectors $\nu_{ij}\xi_i$ is indeterminate, because the reactions are no longer stoichiometrically independent. However, given $\mathbf{N}_{j0}$ and $\mathbf{N}_j$, $\sum^R_{i=1}\nu_{ij}\xi_i$ is completely determined; only its components are not.

The tangent to a trajectory is defined as the vector $\sum^R_{i=1}\nu_{ij}\Delta\xi_i$ at a point of the trajectory. For our set of three reactions, the term ν_{ij} represents three unit vectors, each with a direction parallel to one side of the triangle. From a thermodynamic point of view, there are no limitations on the relative magnitudes of the $\Delta\xi_i$, as long as they are small. The second law of thermodynamics only requires that the free energy along the trajectory is negative definite, or

$$\sum_{i=1}^{R}\Delta\xi_i\sum_{j=1}^{N}\nu_{ij}\mu_j \leq 0 \tag{7a}$$

for each point, where μ_j is the molar chemical potential of A_j at that point. However, if the reactions are kinetically independent in the sense defined above, then each individual reaction must fulfill this condition separately, or

$$\Delta\xi_i\sum_{j=1}^{N}\nu_{ij}\mu_j \leq 0 \tag{7b}$$

for each i. Thus, thermodynamic constraints determine only the signs of the $\Delta\xi_i$; they do not determine their relative magnitudes.

Equation 7b is simply a statement that each reaction occurs separately in a direction of decreasing free energy along its path. That is, the monomolecular reaction A $\rightleftharpoons$ B cannot transfer its free energy decrease to the separate reaction B $\rightleftharpoons$ C, or A $\rightleftharpoons$ C. (See Boudart, 1983; Gorban et al., 1982.) In a linear system, eq 7b is equivalent to the statement that the rate coefficient in eq 3 must be positive. If one accepts eq 7b, then one can define bounds on possible trajectories. At each point any permissible tangent, made up from the three real reactions, must be composed of three vectors, $\nu_{ij}\Delta\xi$, in a way consistent with eq 7b.

In our example we give the directions as $\nu_{(A\rightarrow B)}$, $\nu_{(B\rightarrow C)}$, and $\nu_{(A\rightarrow C)}$ at point P. These are the directions of the three individual reactions in composition space, and are parallel to the sides of the triangle. The unit vector corresponding to the reaction A $\rightleftharpoons$ C is between those corresponding to the two other reaction vectors. Thus the vectors corresponding to the reaction A $\rightleftharpoons$ B and B $\rightleftharpoons$ C constrain the possible direction of the overall reaction vector at point P. As shown in Figure 2, only kinetic trajectories passing P between the reaction vectors ν_{AB} and ν_{BC} can be constructed from these three reactions such that they fulfill the requirements that free energy decreases for each individual reaction separately. Therefore the vectors corresponding to the reactions A $\rightarrow$ B and B $\rightarrow$ C are called limiting vectors at point P as all permissible trajectories must have a tangent vector that points in a direction between them. If one considers all points inside the triangle AFE, they all share the same limiting reaction vectors. The total composition space is divided into six subsections, each of which has different sets of limiting reaction vectors. The dividing lines are the equilibrium lines for the individual reactions, which are called "attractors". A trajectory that goes along the attractor in the direction toward global equilibrium is always permissible.

Consider all possible trajectories starting at the point P (see Figure 3). In the subsection AFE the line PL forms one bound for the possible trajectories. For any point on the line PL, the trajectories passing through it can either go along the line PL or turn to the inside of the space delineated by the lines PL, LE, ET, TP. Consider a point N arbitrarily close to PL on the outside of the shaded region. A trajectory through this point can either be parallel to PL or can enter the shaded space. Any trajectory that has this property is a limiting trajectory in the sense that all trajectories starting from inside the limited space cannot cross this trajectory. The line PL therefore defines a limiting trajectory, and in our terms, it is a minimal bound for the space of permissible trajectories, as every point in the PLET trapezoid is reachable by a straight trajectory from P. Such a trajectory is permissible under the constraint of decreasing free energy for each of the three reactions chosen as the basis of our set (A $\rightleftharpoons$ B, B $\rightleftharpoons$ C, A $\rightleftharpoons$ C).

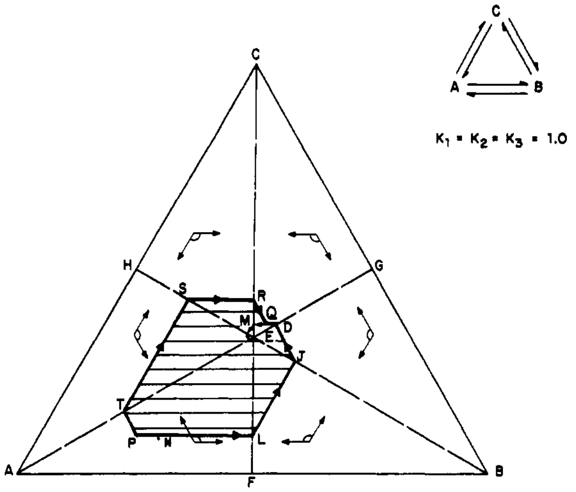


Figure 3. Accessible composition space (shaded area) from initial composition P. Ratio of rate constants k_1 to k_2 to k_3 can be changed at will at each point of the trajectory (catalyst properties change); system assumed to be isothermal: (E) global equilibrium.

Can one obtain a minimum bound for all trajectories starting at P? (A minimum bound is here defined such that no composition outside the bounded set of compositions is reachable from P and all compositions inside it are reachable.) This can be done as follows: L is the intersection of the limiting line PL with the attractor CF. From L the line LJ gives a bound for all trajectories in the triangle LEJ. The line LJ is a true minimum bound. At any point close to LJ the permissible trajectory must either enter the triangle or proceed parallel to it because the tangent to any permissible trajectory at a point in the triangle FEB must either be parallel to LJ or point to the inside of the triangle LEJ. At J one reaches another attractor, and by the same argument one can reach D. From D the limiting trajectory counterclockwise goes to M and from this point it will continue to encircle the equilibrium point E. Thus, if one continues from point P to the right, one forms the limiting trajectory given in Figure 3. This limiting trajectory circles around the equilibrium point in a way that is not permitted for a closed mixed linear system near equilibrium (Onsager, 1931).

One can also construct a limiting trajectory traveling around E from P going clockwise to the left. It will intersect the trajectory from the right at Q. The shaded area will then give the total bound on all compositions reachable by trajectories starting from P. Using our criteria, all points inside the shaded area are reachable by a permitted trajectory but no points outside this area are accessible. The limiting trajectory is in theory realizable if one has a set of three catalysts, each promoting a single reaction

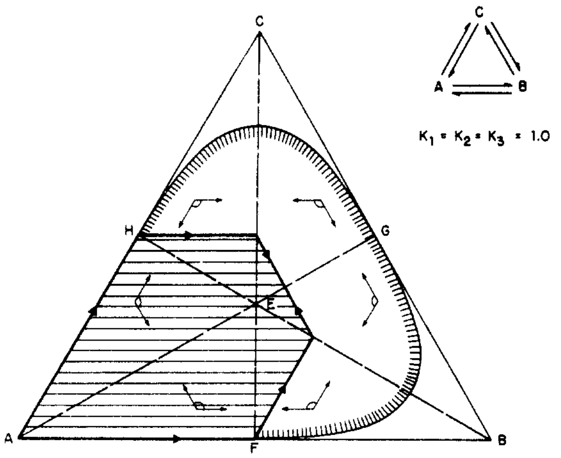


Figure 4. Shaded area represents accessible composition space from initial composition A for the same system shown in Figure 3: (——) shows boundaries of accessible composition space via all kinetic trajectories for which $\Delta G \leq 0$ starting from initial composition A. No assumptions are made regarding linearity of reaction rates.

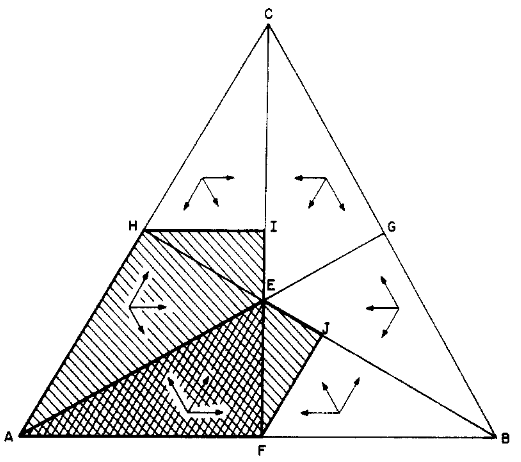


Figure 5. Area AEFA—accessible composition space realizable by the reaction set $A \rightleftharpoons B$, $B \rightleftharpoons C$ with initial composition A; area AHIEJF—accessible composition space realizable by the reaction set $A \rightleftharpoons B$, $A \rightleftharpoons C$ with initial composition A. System is isothermal but reaction rates may be varied arbitrarily.

of the three permitted reactions. A plug-flow reactor could be built with different sections, each containing a different catalyst permitting the single reaction which gives the limiting trajectory. With a single catalyst, kinetic constants would not change abruptly in the way our limiting trajectory does.

The limiting trajectory in Figure 3 also bounds all kinetic trajectories that can occur using a single catalyst promoting these three reactions. However, for this case it is not a minimal bound in the sense that there will be compositions inside the shaded area that cannot be reached from P by using realizable trajectories regardless of what the properties of this single catalyst use. It is only a minimum bound for the hypothetical reactor in which we were free to change catalysts at each position. The advantage of this bound is that it can be computed without any knowledge of the reaction rates using thermodynamic constraints. All we needed was the knowledge of the direction of the reactions (which are obtained here from the fact that the reactions are monomolecular). Figure 4 shows the limiting trajectories when the initial composition is pure A.

Until now we have assumed that all three reactions occur. In some cases some reactions are much less likely to occur than others or are very slow compared to the rates of the other reactions. This knowledge can be included in the analysis. For example, if the reaction $A \rightleftharpoons C$ is slow compared to the other two reactions, the accessible composition space is reduced (see Figure 5). The reason for this is that at any point the trajectory must be between the two vectors defined by the two first reactions $A \rightleftharpoons B$ and $B \rightleftharpoons C$. This follows from the assumption that $A \rightleftharpoons C$ is always small compared to $A \rightleftharpoons B$ and $B \rightleftharpoons C$. Figure 5 also shows the composition space accessible from pure A for the reaction set $A \rightleftharpoons B$, $A \rightleftharpoons C$. We note that the space is much larger than for the set $A \rightleftharpoons C$, $B \rightleftharpoons C$ and is only slightly smaller than the space reachable by all three reactions. Computations such as given in Figure 5 allow one to estimate limits on the potential value of new catalysts that promote an additional more desirable reaction path in a complex system.

IV. Set of Reachable Compositions for Nonlinear Reactions

In deriving the limits on the accessible composition space using reaction trajectories, it was assumed that the reactions are both monomolecular and linear. However, the results can be extended to the more general case where these conditions may not apply. The derivation outlined in Figure 3 assumed only two properties of linear systems: (1) The only reactions that occur are the three reactions $A \rightleftharpoons B$, $A \rightleftharpoons C$, and $B \rightleftharpoons C$. Therefore, the overall reaction rate vector is a sum of the vectors representing these three reactions. (2) Each of the individual reactions is separately governed by the demand that free energy decreases in the direction of this individual reaction.

These conditions are automatically fulfilled (see Wei and Prater, 1962) if the reactions are truly linear. In this case the three reactions are also the only possible reactions. However, linearity is not a necessary condition. The same three reactions with strongly nonlinear reaction rates will have exactly the same constraints on their trajectories.

If the reactions were all second order, nonlinear events and if the overall reaction were $2A \rightleftharpoons 2B$, the direction of the reaction vector would be identical with the reaction $A \rightleftharpoons B$ and so would be the line at which these reactions are in equilibrium. All the arguments in the previous section would apply, if any one of the three reactions, or all of them, should be second order. If the reaction rate expression is more complex, such as a Langmuir expression, then the reaction expressed by the stoichiometric expression $A \rightleftharpoons B$ is not a nonlinear event but the sum of a series of complex events involving adsorbed series or free radicals. But the overall reaction would still go in the same direction toward equilibrium and the equilibrium itself would be the same, provided the concentrations of the intermediates are such that they do not change the equilibrium.

That the nonlinear system still satisfies the second requirement can be deduced from the following argument. Assume the reaction $A \rightleftharpoons B$ which occurs in the presence of several other compounds, all of which are nonreactive, i.e., their individual concentrations do not change. The reaction will always go toward the equilibrium composition. This is true even if the observed reaction $A \rightleftharpoons B$ does not describe a real molecular event but rather the sum of a large number of intermediate steps involving free radicals or special states at a catalytic surface. The exact mechanism is unimportant as long as the concentration of intermediates is very small. For example, one could assume a mechanism, such as

$$\begin{gathered} A + A \rightleftharpoons A + A^* \rightleftharpoons A + B \\ A \rightleftharpoons M \rightleftharpoons B \\ 2A \rightleftharpoons P \rightleftharpoons 2B \end{gathered} \qquad (8)$$

If the concentrations of M and P remain small compared to those of A and B, then the equilibrium

$$[B]/[A] = K_{AB} \qquad (9)$$

is unaffected by the presence of both P and M and is independent of the mechanism or sequence of reactions. The reactions described by eq 8, occurring by any similar mechanism with an arbitrarily large number of steps between A and B, have one important property. Any trajectory based solely on these reactions cannot cross the line or surface in composition space for which the reaction A $\rightleftharpoons$ B is in equilibrium. This follows directly from the previous statement that the observed reaction A $\rightleftharpoons$ B will always proceed toward equilibrium. It is not dependent on the system being linear. In the case of non-first-order systems, there may be more than three rate-determined reactions. It can be shown, however, that the requirement of decreasing free energy in the direction of each individual reaction must still apply.

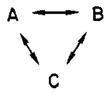
In a homogeneous isothermal system the free energy is a continuously differentiable function of composition. If a very small amount of P is formed in the system, the change in free energy due to its formation will also be very small as there are no known intermediates for which the molar free energy is very large compared to the other molar free energies. Therefore, a very small amount of P cannot modify the free energy of system along the path A $\rightleftharpoons$ B, and the free energy constraints of the single reaction A $\rightleftharpoons$ B will be independent of the mechanism of the reaction. This is not true, however, if the intermediate P appears in large concentration. For example, one can regard C in Figure 1a as an intermediate of the reaction A $\rightleftharpoons$ B. As long as [C] is small the reaction trajectory is very close to the line AB, limited by the free energy constraint of the reaction A $\rightleftharpoons$ B and cannot pass the equilibrium line FC. On the other hand, if the net amount of C is large, one can find a permissible trajectory that passes this line. However, if the concentration of any intermediate is large, it should be measurable and should be included in the coordinates of the composition space.

One should note that, near the equilibrium of the reaction A $\rightleftharpoons$ B, the detailed reaction mechanism must also fulfill the constraint of detailed balancing for the forward and backward reactions. This has led some authors (Aris, 1969) to suggest that complex reaction rate expressions should always be written such that detailed balancing near equilibrium is automatically satisfied. This is an advantageous procedure for obtaining empirical fitting functions, since otherwise they would only be valid far from equilibrium. However, for our purposes, the only requirement for the single reaction is that the free energy must decrease along the reaction path. For the bounds derived in Figure 3, it is therefore not important if the reaction rates are linear or are more complex functions of concentration. What is important is that the overall reaction is of the form given.

There is, however, one essential difference between the linear system with three compounds and the case in which the same compounds undergo nonlinear reactions. For a linear system the three reactions in eq 3 are the only possible reactions. In a nonlinear system one can write many other overall reactions, such as

$$2A \rightleftharpoons B + C$$

While stoichiometrically one can decompose such a reaction into the three reactions given, these are kinetically not equivalent. This will be discussed in the next section. If the reactions are nonlinear, then for the constraints in Figure 3 to apply to the system of reactions one further assumption is needed, namely, that the only way that reaction B $\rightleftharpoons$ C affects reaction A $\rightleftharpoons$ B is via the stoichiometric changes in the system. A semipermeable membrane, removing or adding A or B, would have exactly the same effect on the reaction A $\rightleftharpoons$ B. If this holds we say that the reaction A $\rightleftharpoons$ B is not coupled to the reaction B $\rightleftharpoons$ C. As the observed reactions are not necessarily monomolecular but may involve a series of complex reactions, the concepts of coupling and independence require a precise definition. This is presented in the next section.



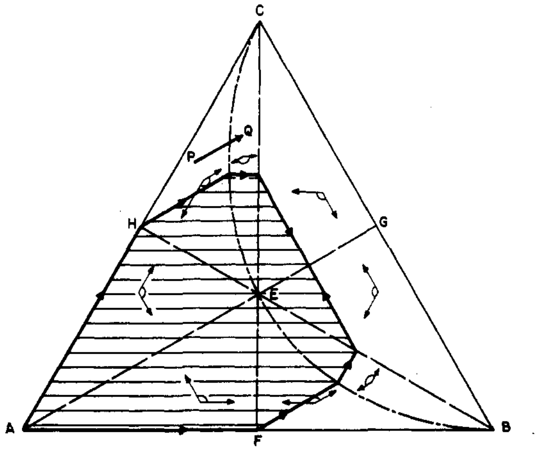


Figure 6. Accessible composition space (shaded area) from an initial composition A realizable by the reaction set A $\rightleftharpoons$ B, B $\rightleftharpoons$ C, A $\rightleftharpoons$ C, 2A $\rightleftharpoons$ B + C. System is isothermal but reaction rates may be varied arbitrarily. The line (---) shows the locus of compositions at chemical equilibrium according to eq 11.

V. Coupling of Catalytic Reactions

Consider the reaction mechanism

$$2A \rightleftharpoons M \rightleftharpoons B + C \tag{10}$$

M is an intermediate that is present only in very small amounts, such that its influence on the free energy of the system is negligible and the topography of the free energy isoclines in the volume ABCM are practically identical with that in the plane ABC. The direction of reaction 10 is given by the vector PQ in Figure 6. This reaction can proceed as long as the free energy declines along this vector. The equilibrium line for which

$$\frac{[B][C]}{[A][A]} = K \tag{11}$$

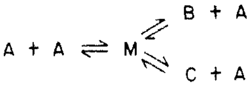
is given also in Figure 6. Reaction 10 is stoichiometrically the sum of the reactions A $\rightleftharpoons$ B, and A $\rightleftharpoons$ C. As it is not linearly independent in a stoichiometric sense, there is no reason to add it to the system if one merely wishes to compute equilibria or heat and mass balances.

However, kinetically, reaction 10 is completely different from the pair of separate reactions A $\rightleftharpoons$ B and A $\rightleftharpoons$ C. Consider, for example, the composition at point H. Here, the reaction A $\rightleftharpoons$ C is at equilibrium. A reaction trajectory starting from composition H in the direction P → Q has a decreasing free energy. As this trajectory corresponds to the overall reaction 2A $\rightleftharpoons$ B + C, there is no thermodynamic reason why such an overall reaction could not occur. However, the two individual reactions A $\rightleftharpoons$ B and A $\rightleftharpoons$ C could not result in such a trajectory. The reaction A $\rightleftharpoons$ C has an increase in free energy along this trajectory, whereas the reaction A $\rightleftharpoons$ B has a decrease in free energy. In an isothermal system there is no way the independent reaction A $\rightleftharpoons$ B can transfer its release of free energy to the independent reaction A $\rightleftharpoons$ C (Bouduart, 1983). The intermediate M permits one to achieve this goal. *It couples* *the two reactions.* If M is measurable and occurs in significant concentration, then the reaction $2A \rightleftharpoons M$ would have to be included as a separate reaction. Often in catalysis, however, the intermediate is unknown or occurs in small quantities. Earlier we specified that we deal with phenomenologically observable reactions and not with elementary steps. Therefore, there is nothing wrong in writing reaction 10 as $2A \rightleftharpoons B + C$, since this represents the real, phenomenologically observable reaction better than the pair $A \rightleftharpoons B$, $B \rightleftharpoons C$. Reaction 10 therefore describes a permissible form of coupling, as all real steps have a decreasing free energy. This example illustrates the difference between linear independence in stoichiometric relations and kinetic independence of phenomenological reactions defined in this paper.

Let us look more carefully at the effect of including the added reaction $2A \rightleftharpoons B + C$. The intermediate M opens a trajectory that is not allowed by the two separate reactions or by the three reactions used previously. As the concentration of M is small, it does not change the topography of the free energy relations. However, it couples the two reactions, $A \rightleftharpoons B$ and $A \rightleftharpoons C$, and creates a new permissible trajectory. For a kinetic description of a system having such a trajectory, one must add reaction 10 (or another reaction with similar properties) to the set of the first three reactions.

Figure 6 shows the new reaction space boundary, including reaction 10. In comparison with the boundaries for the first three reactions only, there are now eight regions in which the directions of the limiting vectors can change. Note that the accessible composition space is larger than the one permitted using only the three first-order noncoupled reactions (Figure 4).

The coupling of the two reactions in eq 10 is due to a joint intermediate. The terms coupling and independence were previously described in terms of semipermeable membranes. Two reactions are independent if the only effects on each other are equivalent to the effect of a selective semipermeable membrane that can change the concentration of the reactants by adding or removing a reactant. One can now define coupling in another way. Consider an overall reaction that can be broken down into two simpler reactions. The two reactions are called coupled if the overall reaction permits trajectories which are thermodynamically not permitted by the two separate reactions. Elementary reactions cannot be coupled. Coupling requires a more complex mechanism involving joint intermediates. However, joint intermediates do not always lead to coupling. For example, the reactions



$$A + A \rightleftharpoons M \begin{array}{l} \nearrow\!\!\swarrow\ B + A \\ \searrow\!\!\nwarrow\ C + A \end{array} \qquad (12)$$

have a joint intermediate M, but are not coupled in the thermodynamic sense discussed here. The restrictions on the direction of reaction trajectories for reactions 12 are identical with those of the two individual reactions $A \rightleftharpoons B$, $A \rightleftharpoons C$.

In the above definition no constraints are imposed on the nature of the actual reaction rate expression. It may be quite complex as it represents the sum of a number of molecular events. In addition, the individual steps of the phenomenological reaction are not required to be in steady state or even pseudo steady state. The intermediate compounds may well change with time, which will change the overall reaction rate. It also makes no difference if the reaction rate expression contains concentrations of compounds that do not appear in the overall reaction (or for which ν_{ij} is zero in reaction i). Such rate expressions are common in catalytic reactions, in which some catalyst sites become unavailable due to preferential adsorption. This will slow reactions in which the adsorbed compound is not involved but does not necessarily lead to coupling in our sense. The only requirement is that the concentrations of these intermediates are sufficiently small that they do not affect the free energy isoclines.

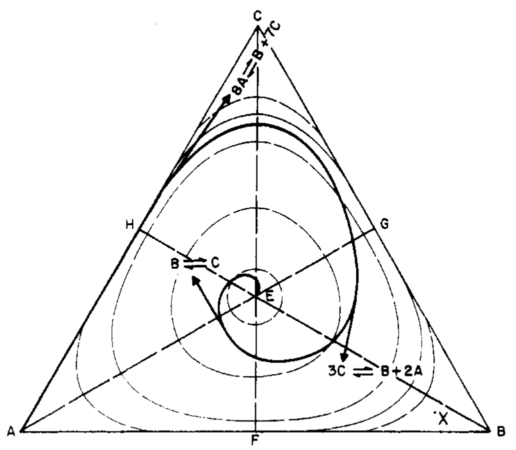


Figure 7. A hypothetical trajectory in composition space with an initial composition pure A. The only requirement is that free energy decrease along the trajectory. The different tangents show the directions of the overall reaction at each point, which can be realized if the ostensible reaction is equal to the overall reaction.

The model system used here has only two linearly independent stoichiometric relations. However, in Figures 3–5, three kinetically (or thermodynamically) independent reaction vectors are assumed. In Figure 6 a fourth reaction vector is added. This system could have a large number of such thermodynamically independent overall "ostensible" reactions. For each such reaction (as for example the one shown in eq 10), there could be a large number of underlying mechanisms. These may even lead to ostensible reactions with different stoichiometric coefficients (Happel and Sellers, 1982). As long as the direction of the overall reaction vector is the same, these mechanisms result in the same thermodynamic constraints and are therefore equivalent for our purposes. This allows one to illustrate the concept of thermodynamic independence of kinetic reactions defined below. A reaction $2A \rightleftharpoons B + C$ is thermodynamically independent of the three reactions $A \rightleftharpoons B$, $B \rightleftharpoons C$, and $A \rightleftharpoons C$ if it allows trajectories that cannot be realized with the three single reactions. The assumption that such an additional independent reaction exists or is required depends on experimental evidence that a trajectory that is not permissible by the simpler set exists. Each new reaction would add another possible vector. If enough reactions are added, one can, in principle, realize any trajectory for which the overall free energy is decreasing.

In Figure 7, a hypothetical trajectory is shown. Overall reactions which have a direction tangential to the trajectory vector are shown for several points. For the isothermal system ABC, the only constraint imposed by the second law is that the free energy should decrease. Owing to stoichiometric constraints, however, not all points in the composition space with a lower free energy than that of the starting point are accessible. For example, point X in Figure 7 has a lower free energy than the initial composition of pure A but cannot be reached by any trajectory with continuously decreasing free energy because the free energy isocline passing X crosses the stoichiometric boundary. However, any stoichiometrically permissible kinetic trajectory along which the free energy decreases can, in theory, be realized by a series of plug-flow reactors with different catalysts provided one can find catalysts that promote a set of suitable coupled reactions.

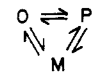
In an isothermal reactor system with a single catalyst there are additional constraints. Reaction rates of different reactions cannot be altered at will. For a given catalyst they have a constant relationship which introduces kinetic constraints much narrower than free energy limits.

For these reasons one often uses more than one catalyst to achieve a specific overall reaction. However, even with multiple catalysts the number of possible reactions with available catalysts is finite and much smaller than is required to achieve trajectories close to the free energy constraint. Therefore, one is usually dealing with thermodynamic constraints which are the results of kinetic properties of available catalysts. Our method allows one to learn something about these thermodynamic constraints by looking at the thermodynamic implications of catalyst properties. If one knows which reactions are occurring at reasonable rates (rates fast enough to have an impact on the concentrations in the time-span of observation), one can derive bounds for the set of accessible compositions, using only information from thermodynamic tables. In some sense, this is similar to computing a global equilibrium in complex reactions. Then one only includes compounds which are formed at observable rates and which are stable enough to exist in measurable concentrations. Knowledge of kinetics enters through the choice of the compounds used to compute the global equilibrium composition. One does need more information to compute bounds for the accessible space by our methods; one needs to know which ostensible reactions actually occur.

The method described here can be extended to an arbitrary number of compounds. If the number of stoichiometrically independent reactions is larger than 2, the graphical description becomes more complex and requires a machinery that is outside the scope of the paper. Attractors for a single reaction become surfaces in multidimensional space and divide the space into regions in which the limiting vectors are constant. However, the principles remain exactly the same and the proofs hold for the multidimensional case as well. No single reaction can pass through the attractor surface defined by its equilibrium relation. If the number of stoichiometrically independent reactions is larger than 3, the graphical representation is too complex. But one can always use eq 7b to compute bounds on permissible trajectories and to check whether an experimental trajectory is allowable by an assumed set of reactions.

VI. Defining a Sufficient Set of Chemical Reactions and a Sufficient and Kinetically Independent Set

We can now deal with a problem raised in the Introduction, namely: how can one go about setting up a set of overall phenomenological reactions that can serve as a base for obtaining rate equations in a complex multicomponent system? Consider the simple case of the isomerization of three xylenes which has been dealt with extensively by Wei and Prater (1962).



(13)

There are two linearly independent stoichiometric reactions. However, a set of reaction rate expressions is needed to model the system adequately. The set given by these three reaction rates is sufficient in the sense that all available data can be fitted using three linear reaction rate expressions. By using the methods outlined, one can obtain the necessary results in two ways without kinetic fitting. To do this one needs a set of observed trajectories in composition space. These trajectories should have different initial compositions and ideally will cover the composition space of interest over the actual range of temperatures and pressures. One can then verify if the following two conditions are met. (1) Each experimentally accessible composition should be reachable from the initial composition by using these three reactions and the criteria outlined in section V. (2) Equation 7b must be fulfilled at each point of each trajectory. In other words, the tangent of the trajectory at each point must be decomposable into the assumed reactions in such a way that each of the individual reactions goes in the direction of decreasing free energy.

At each point one can construct a tangent which can be expressed as a vector connecting two adjacent compositions. The magnitude of this vector can be chosen arbitrarily. If the number of reactions R chosen is equal to the number of linear stoichiometric reactions S, one can decompose the tangents uniquely into the direction of the reactions as the difference between the compositions is given by the vector

$$\sum_{i=1}^{R} \nu_{ij} \Delta\xi_i$$

Here, the $\Delta\xi_i$ are uniquely given and one can check if

$$\Delta\xi_i \sum_{j=1}^{N} \nu_{ij}\mu_j \leq 0 \tag{7b}$$

for all i. If the number of reactions, R, is larger than S, then the $\Delta\xi_i$ are not unique. One only requires that there is one set of $\Delta\xi_i$ which fulfills these conditions. If $S = 2$ it can be done by inspection. If S is larger, one can take subsets of R with the same dimension as S and check at each point for all subsets. If R is larger than 4, this becomes tedious. There is a straightforward method based on linear programming called the Simplex Method which allows one to solve this for artibrary dimension of R and S. It is well described in the literature (Bazaraa and Jarvis, 1977). A detailed explanation of the method is, however, outside the scope of this paper.

The second condition is more stringent than the first. In fact, it implies the first. If eq 7b is satisfied at each point of the trajectory than all the compositions in the trajectory are accessible. Therefore, the first condition is merely a convenient and valuable fast screening technique. Note that the order of the reactions was not important. If one cannot find for each point a subset for which eq 7b holds, then the total set of reactions is insufficient to describe the system. Conversely, if both conditions are met for a given set of reactions for all trajectories, then the set is called sufficient. If a set is sufficient and no subset is sufficient, then the set is a *minimum sufficient set* of kinetically independent reactions. Note that these definitions apply only to a given set of experimental data which define the accessible composition space.

In contrast, *linear independence* of a set of reactions in the stoichiometric sense can be defined without any experimental data, if one knows only the number of compounds involved. A minimum sufficient set of kinetically independent reactions is not necessarily unique, and different minimum sets may have a different number of reactions. However, the concept still remains useful as it allows one to start with a proposed set of reactions, test for sufficiency, and, if desired, reduce that set to a minimum.

To describe reactions in multicomponent systems, one needs to define a sufficient set such that

$$\sum_{j=1}^{N} \nu_{ij} A_j = 0; \quad (i = 1, ..., R) \tag{5b}$$

and fit reaction rate expressions to the experimentally observed trajectories. Remember that one is not dealing with molecular events but with overall ostensible reaction rates. The methods outlined in this paper do not offer an easy way to set up a set of reactions to fit the kinetic data. In general, one does this studying the reactions in isolation and by a combination of experience, intuition, and knowledge of possible mechanisms. Once one sets up a set of reactions, the techniques allow one to check whether the set is sufficient to explain the set of experimentally observed trajectories. For any given set of trajectories one verifies condition 2 by checking each trajectory separately. For several points along the trajectory one computes the tangent vector and checks if it is decomposable into the unit vectors of the individual reactions of the set assumed in a way that eq 7b is fulfilled.

The set is sufficient if this can be done at all points of all trajectories. If the number of reactions in the set chosen is equal to the number of linearly independent reactions, S, then a sufficient set will also be a minimum set as the minimum number of reactions required to describe the reactions underlying the trajectories is equal to the number of linearly independent reactions. If the number of reactions in the set chosen is larger than S, then one can check if there are any subsets which are sufficient. If not, the set is a minimum set. Sufficient minimum sets of chemical reactions are not unique, as several sets may be consistent with a given set of data. A good kineticist may have clues for the choice of a preferable set, but one cannot make such choices based on the trajectories. For purposes of modeling these sets are equivalent. If the test shows that a set of proposed reactions is not sufficient, then one must either enlarge it by adding reactions or modify it by substituting other reactions. Thermodynamic analysis can provide guidelines and clues. A set of reactions used for setting up rate equations by fitting experimental trajectories must be sufficient. There is no requirement for it to be minimum, though this may be advisable if there is no independent evidence that larger sets are justified.

In section VII detailed examples of applications will be presented. However, to illustrate the concepts outlined above, we will elaborate a little further on the simple hypothetical model shown in Figures 3–7. Consider the trajectory APE shown in Figure 8. One first tries to use the set of the three simple overall reactions ($A \rightleftharpoons B$, $A \rightleftharpoons C$, $B \rightleftharpoons C$). This set is sufficient to explain the trajectory. As there are only two independent stoichiometric reactions, one wants to check if any two-member subsets are sufficient. The only two-member subset which is sufficient is $A \rightleftharpoons B$, $B \rightleftharpoons C$. While the trajectory could be decomposed at each point into two other subsets ($A \rightleftharpoons C$, $B \rightleftharpoons C$) and ($A \rightleftharpoons C$, $A \rightleftharpoons B$), these decompositions would violate eq 7b at some point along the trajectory. This is clear for the subset $A \rightleftharpoons C$, $B \rightleftharpoons C$ as the corresponding trajectory is outside the accessible space for this subject, which is the triangle AHE. Even the less stringent first condition is not met. However, the trajectory is in the space accessible by the set $A \rightleftharpoons B$, $A \rightleftharpoons C$. Close inspection of the trajectory will confirm that eq 7b is also violated at some points of the trajectory for this subset.

Now consider the trajectory AQE. It is not contained in the reachable composition space for the subsets ($A \rightleftharpoons B$, $B \rightleftharpoons C$) and ($A \rightleftharpoons C$, $B \rightleftharpoons C$). It is, however, in the reachable space of the subset ($A \rightleftharpoons B$, $A \rightleftharpoons C$). It also fulfills eq 7b for this subset at each point along the trajectory. Thus, whether or not a two-reaction subset of these three reactions can be sufficient depends on the experimental data. If the only observed trajectory is APE (or AQE) then the subset ($A \rightleftharpoons B$, $B \rightleftharpoons C$) (or $A \rightleftharpoons B$, $A \rightleftharpoons C$) constitutes a minimum sufficient set. If, however, both trajectories APE and AQE have been experimentally measured, even if at different pressures or different catalysts states, then the three-reaction set ($A \rightleftharpoons B$, $A \rightleftharpoons C$, $B \rightleftharpoons C$) is the minimum sufficient set as no two reaction subsets will fulfill eq 7b at all points for both trajectories.

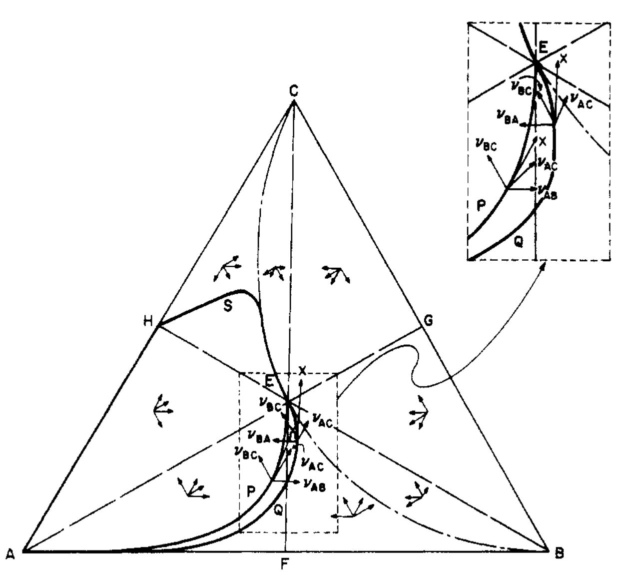


Figure 8. HSE, APE, AQE—hypothetical trajectories with different initial compositions for the reaction system $A \rightleftharpoons B$, $B \rightleftharpoons C$, $A \rightleftharpoons C$ shown in Figure 2. The tangents to the trajectory at each point are the reaction vectors. HSE cannot be realized without a fourth permissible reaction $2A \rightleftharpoons B + C$. For further discussion see text.

Consider now the case in which the trajectory HSE is also observed. This trajectory is outside the composition space reachable by the basic three reactions. Therefore, the basic set is no longer sufficient and additions or modifications are needed. For example, if one adds the reaction $2A \rightleftharpoons B + C$ to the basic set, then a sufficient set is again obtained. The choice of the additional reaction is usually based on some knowledge of the observed overall reactions.

If one wishes to determine reliable kinetics for a given reaction system, it is preferable to have trajectories with different initial compositions over a range of temperatures and pressures. For the three-compound system used in this paper, the local equilibrium lines and the sets of limiting vectors are independent of pressure but change with temperatures. Plotting figures such as Figures 2–8 for different temperatures and pressures and different sets of reactions is useful in the study of complex reactions as it provides the researcher with insight as to the constraints. One can then test the sufficiency of a set of proposed reactions for different pressures and temperatures directly by using eq 7b at different points of the observed trajectories. This technique is especially useful in screening new catalysts. Assume, for example, that with a given class of catalysts in our model system, the three-member set of reactions ($A \rightleftharpoons B$, $B \rightleftharpoons C$, $A \rightleftharpoons C$) is always sufficient. If one now finds a catalyst that promotes the trajectory HSE, a trajectory that is outside the composition space accessible by these reactions, then this provides strong evidence that this new catalyst operates by a different mechanism than the previous catalyst. Alternatively, it may accelerate reactions that, while they may have occurred using the previous catalysts, were too slow to have an impact on the observed trajectories. Such a catalyst would therefore have the potential of achieving selectivities and compositions

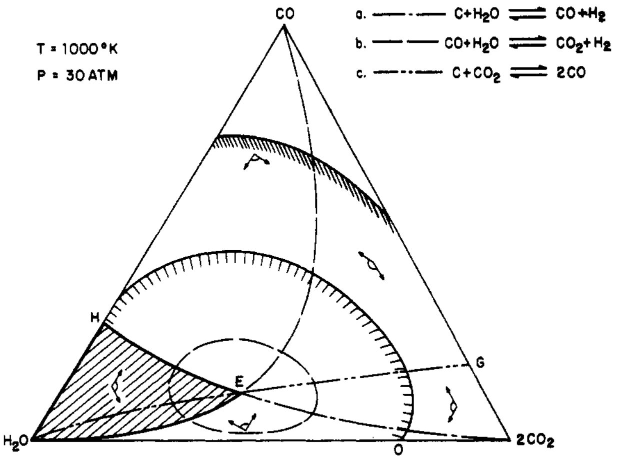


Figure 9. Accessible composition space for steam gasification of carbon starting from pure steam at 1000 K and 30 atm. The only compounds formed are CO, CO_2, H_2: (——) composition space for which the total free energy change is zero or negative; (——) composition space reachable by a kinetic trajectory along which the free energy change is negative definite. (Arbitrary kinetics as in Figure 7.) Area H_2O–H–E–H_2O: composition space reachable by the set of reactions 14a, 14b, and 14c.

unobtainable by the previous class of catalysts.

VII. Steam Gasification of Carbon

In this section we attempt to further illustrate the concepts of a sufficient set of kinetically independent reactions by examining in detail the gasification of carbon or coal chars. As a specific example, we consider the steam-gasification of carbon under conditions where the only products are H_2, CO, and CO_2 (no methane is formed). Since there are five compounds and three elements, two independent stoichiometric reactions are sufficient to describe the system in a two-dimensional space. Starting with pure steam, an element balance of oxygen can describe the system completely in a triangular composition diagram (Figure 9). The H_2 yield is simply equal to the steam conversion and can be obtained from the plot. It is normally assumed that the overall reactions most likely to take place are

$$C + H_2O \rightleftharpoons CO + H_2 \qquad (14a)$$

$$CO + H_2O \rightleftharpoons CO_2 + H_2 \qquad (14b)$$

$$C + CO_2 \rightleftharpoons 2CO \qquad (14c)$$

However, our analysis will include another kinetically independent ostensible reaction

$$C + 2H_2O \rightleftharpoons CO_2 + 2H_2 \qquad (14d)$$

We further assume that the desired product is pure hydrogen. Therefore, the desired overall reaction is (14d), which occurs along the axis H_2O–CO_2 in Figure 9.

Using the methods previously outlined, one can bound the accessible composition space at a given temperature. We choose as our first example a temperature of 1000 K and 30 atm. If one considers only the constraint that the final composition has to have a lower free energy than the initial composition, the accessible space covers most of the triangle in Figure 9. It is shown by the isocline $\Delta G = 0$.

A much smaller permitted region results if one now includes the additional stoichiometric constraint that, in order to reach a given composition, one must find a trajectory inside the accessible triangle along which the free energy G decreases at each point. The reachable composition space is then bounded by the free energy isocline passing through point H (which is the equilibrium composition of the reaction $C + H_2O \rightleftharpoons CO + H_2$.

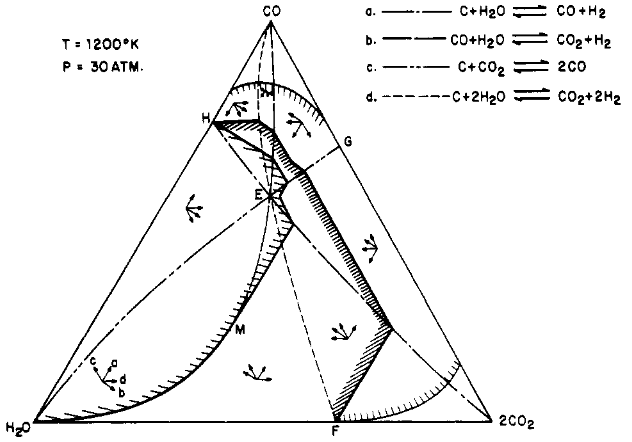


Figure 10. Accessible composition space for the steam gasification of carbon reaction system of eq 14 at 1200 K and 30 atm: (——) accessible composition space realizable by the subset of 3 reactions a, b, and c; (——) accessible composition space realizable by the subset of 4 reactions a, b, c, and d.

Figure 9 also shows the accessible composition space for the reaction set 14a, 14b, 14c at 1000 K and 30 atm. Note that the limitation of only three reactions further limits the accessible composition space. A larger accessible composition space is obtained either by changing the temperature of the reaction or by adding a fourth permissible reaction as shown in Figure 10.

Since the desired product is pure H_2, the desired overall reaction is (14d). Reaction 14d does not occur in practice, at least not by itself. However, one can achieve the desired end result (almost pure H_2) by first gasifying at high temperature and getting close to global equilibrium and subsequently cooling the product gas and shifting it with steam over a catalyst (reaction 14b). In a nonisothermal reactor one can then get practically complete conversion of steam to $H_2 + CO_2$ (reaction 14d) despite the fact that reaction 14d by itself may be negligibly slow. In a single-stage isothermal reactor one cannot do this unless reaction 14d dominates. Such a two-stage operation is a standard way in which the reaction engineer overcomes constraints imposed by the kinetic nature of the reaction system. The final composition reached by such a two-stage nonisothermal scheme is not reachable in any isothermal reactor, which promotes only the reactions 14a, 14b, and 14c. (A similarly simplified case was discussed by Shinnar et al., 1982.) The advantage of the two-temperature schemes is based on the fact that reaction 14a is endothermic and 14b is exothermic. At high temperatures the conversion of steam by (14a) is high but the global equilibrium composition has a high CO content, as the equilibrium of reaction 14b is not favorable. If one now reduces the temperature, this results in a starting composition which is not reachable at all at a low temperature. From this new starting composition one can obtain a high conversion to H_2. This trajectory is not permissible at any one temperature in an isothermal reactor via reaction paths based on (14a) and (14b). This principle is widely used to overcome thermodynamic constraints such as in the production of methanol (Shinnar, 1983), SNG (Shinnar et al., 1982), and other processes.

The reaction set 14a, 14b, 14c has been successfully used to fit experimental data for such gasifiers (Johnson, 1974). There is independent evidence that each of these reactions is a real ostensible reaction. Reactions 14b and 14c can be realized separately. Reaction 14a can never occur separately as (14b) always also occurs. We will now try and use this set of reactions to illustrate the method of

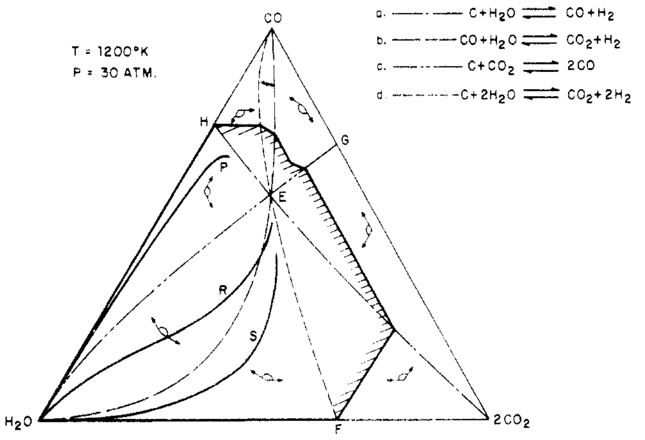


Figure 11. Three hypothetical experimental trajectories, H_2O–P, H_2O–R, and H_2O–S, realizable by the system of reactions a, b, c, and d.

choosing a set of reactions to fit experimental trajectories. We will initially limit ourselves to isothermal reactors.

In Figure 11, three alternate hypothetical experimental trajectories are shown. In the presence of different catalysts, all three trajectories are theoretically possible even with fixed temperature and pressure. Assume that the reaction set 14a, 14b, 14c, 14d is a reasonable set with which to start the fitting process. First one checks whether any of the reactions is much faster than the others. If this were the case, then any trajectory would proceed close to the equilibrium line of this reaction. If one could assume pseudo-equilibrium for this reaction, then after a short induction period, one would only have one degree of freedom left and, therefore, only one possible trajectory. In this case one could not deduce any information about the nature of the other reactions. None of the three hypothetical trajectories fulfills these conditions, though trajectory P is close to the equilibrium line of reaction 14a.

As no one reaction dominates by virtue of its rate, one proceeds to the next step. Inspection shows that all three trajectories are consistent with the total set of four reactions. The question is how many reactions are needed to be consistent with each specific trajectory.

First consider three reaction subsets. Since only two of these reactions, 14a and 14d, are primary gasification reactions, one initially examines the two reaction subsets which contain only one primary reaction. This is illustrated in Figure 12a,b. Note that the accessible space in either figure does not cover all three trajectories. (Trajectory P is outside the space in Figure 12a, and trajectory S is outside the space in Figure 12b.) In both cases eq 7b is fulfilled for all trajectories which are inside the reachable space. These two three-reaction sets cannot fit all trajectories. Therefore, if the observed trajectory is P, the primary gasification reaction must be (14a). On the other hand, if the measured trajectory is S, then the primary gasification reaction would have to be (14d). Trajectory R is consistent with both sets. In reality, observed trajectories are closer to P than to S, indicating that (14a) is the primary reaction.

One can also look at smaller subsets containing both primary gasification reactions. The simplest subset is (14a, 14d) which is shown in Figure 13. The permissible space covers all trajectories, but trajectory S violates eq 7b. For trajectory S to be realizable one would therefore need at least one additional reaction. If one adds the shift reaction 14b, then the set becomes sufficient. However, no trajectory similar to S has been reported. If one could find a catalyst that promoted a trajectory similar to S, it would indicate that the reaction set assumed by Johnson (see Figure 12b) is insufficient and that the catalyst promotes reaction 14d faster than reaction 14a or promotes another more complex reaction as the primary gasification reaction.

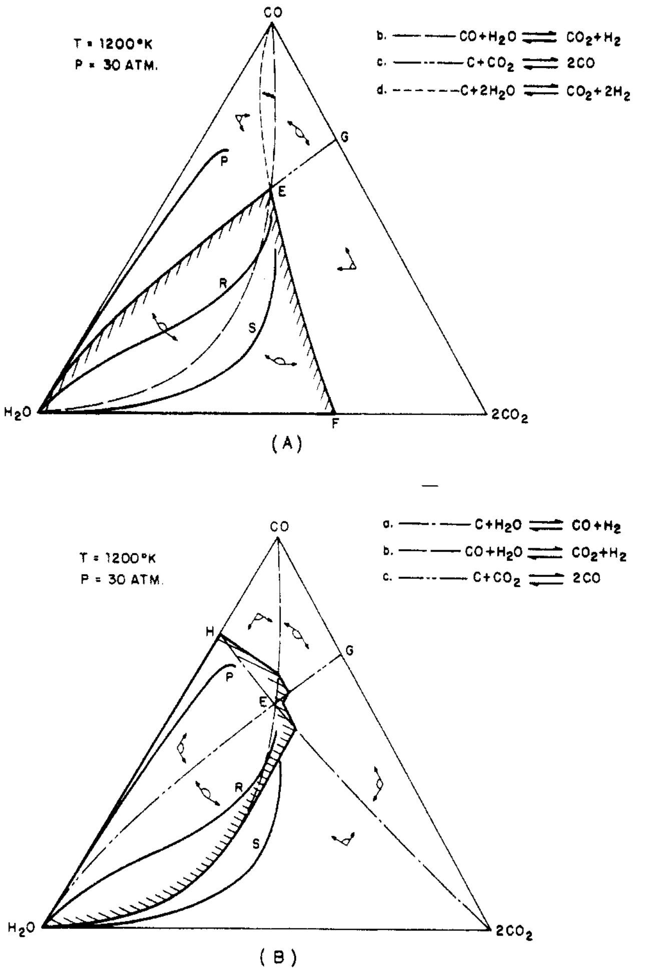


Figure 12. Evaluating minimum sufficient reaction subsets for the steam gasification of carbon system, trajectories shown in Figure 11: (——) accessible composition space realizable by the reaction subset b, c, d (12a) and a, b, c (12b) for the system with initial composition H_2O. The trajectory H_2O–P is not realizable by the reaction subset b, c, d, whereas the trajectory H_2O–S is not realizable by the reaction subset a, b, c.

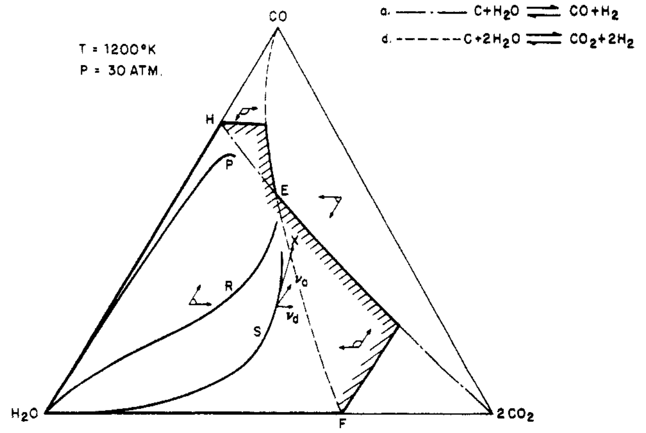


Figure 13. Evaluating reaction subsets for steam gasification of carbon system shown in Figure 11: (——) accessible composition space realizable by the reaction subset a, d, although all three trajectories H_2O–P, H_2O–R, H_2O–S are inside the accessible composition space. Trajectory H_2O–S violates eq 7b and is therefore not permitted in this system.

Let us now look at a case where trajectories P and R have been measured (either for different chars or in the presence of different catalysts), and one is looking for a

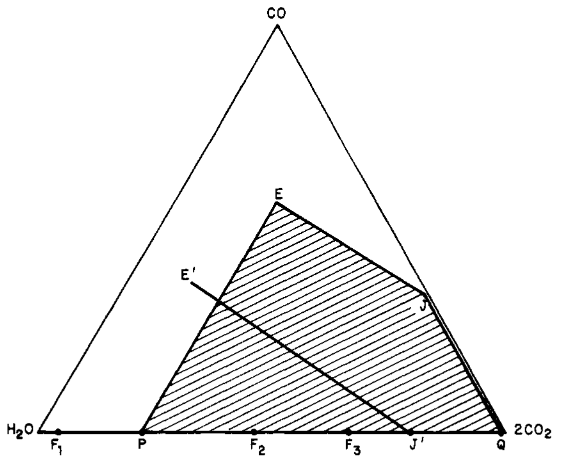


Figure 14. Modification of accessible space in char gasification by nonisothermal trajectories (reaction system 14a, 14b, 14c). Shaded area—composition space accessible by the two-step process: (a) gasification at 1200 K reaching equilibrium point E; (b) cooling the products to 500 K and continuing all reactions at 500 K; F_1, F_2, F_3 are the equilibrium points for the reaction $C + 2H_2O \rightleftharpoons CO_2 + 2H_2$ at 500, 1000, and 1200 K, respectively; trajectory EJ, shift reaction only, at 500 K; trajectory E′J′, equilibrium composition at 1200 K (E), diluted with pure steam (E′), and shifted at 500 K.

sufficient reaction set which will correspond to both trajectories. The three-reaction set shown in Figure 12b is sufficient. One can now look at two-member reaction subsets. Of the three possible subsets the only one meeting the necessary criteria is the set containing reaction 14a (water gas reaction) and reaction 14b (the shift reaction). Reaction 14c is not required for this minimum or kinetically independent set. In practice it is included because there is independent evidence for its occurrence which can be obtained by gasifying carbon with CO_2, where it is the only possible overall reaction. However, any of the published data can always be fitted by the first two reactions (14a and 14b). In the same way one could also include reaction 14d in a set used for fitting these trajectories. Trajectories P and R provide evidence only that reaction 14d is not needed and, therefore, must be slow compared to reaction 14a. Real trajectories cannot provide conclusive evidence whether the reaction occurs or not. The example illustrates all the points made in the previous section.

One can use the same example to illustrate another important point, namely, the way in which thermodynamic constraints imposed by catalysts can be modified by using a nonisothermal reactor. In the above example, in an isothermal reactor (14d) is not the dominant reaction, and therefore large amounts of CO will be formed. As the desired product is hydrogen, we are really interested in exclusively promoting reaction 14d if that is possible. This cannot be done in a single reactor but can be achieved in two steps. In practice what is done in hydrogen production is to try to reach global equilibrium at high temperature and then to cool the mixture and to shift it at lower temperature in a separate shift reactor using a separate shift catalyst. Let us look at this procedure in terms of our thermodynamic constraints.

Figure 14 shows the accessible space that can be reached from point E in Figure 13 at 500 K, the lowest temperature used for shift, using only reactions 14a, 14b, and 14c. One limiting trajectory is the shift reaction 14b, which occurs along the line EJ. This reaction is achievable and reaches the greatest H_2 yield of any point reachable from point E. The other limiting trajectory is determined by reaction 14a, $CO + H_2 \rightleftharpoons C + H_2O$ and is given by the line EP. Point P is actually slightly removed from the line H_2O–CO_2, but is so close that the difference cannot be plotted on an observable scale. The global equilibrium point at 500 K is so close to the equilibrium point F_1 of reaction 14d, $C + 2H_2O \rightleftharpoons CO_2 + 2H_2$, that one cannot distinguish it from F_1 on the scale of Figure 14. This global equilibrium point is reachable from point P by a combination of reactions 14a and 14b or 14b and 14c. Here, the thermodynamic constraints are such that the sum of reactions 14a and 14b (or 14b and 14c) is practically equivalent to reaction 14d. The third limiting trajectory, JQ, is given by reaction 14c, $2CO \rightleftharpoons CO_2 + C$. It is very slow at 500 K and in practive one tries to avoid it because it deactivates the shift catalyst.

Also shown in the same plot are the sets of accessible compositions that can be reached starting from only steam and carbon at 500 K. This set is so small and so close to the line H_2O–F_1 that it is impossible to draw it accurately on the scale of the figure. For practical purposes, it is represented by the line H_2O–F_1. Point J has a higher overall steam conversion than any equilibrium point based on reaction 14d alone in the temperature range 500–1200 K (these are shown as points F_1, F_2, and F_3). However, the mixture at point J still contains significant amounts of CO. In practice, most of the CO is shifted to CO_2 by adding steam to the mixture at point E and moving the composition to point E′, which now can be shifted to point J′. Note that, at the point J′, the CO yield is very small and CO_2 removal will lead to almost pure H_2. This allows one to achieve a high conversion of the total steam into H_2 and CO_2. Kinetically, this step is also possible because at 500 K reaction 14b is very fast compared to reactions 14a and 14c.

The above example illustrates that, in a multicomponent system, the ability to vary the temperature increases the accessible composition space and the permissible trajectories over those that exist in an isothermal system of the same dimension in compound space. These additional permitted trajectories may allow one to improve the yield of the desired product. This principle is widely used in chemical reactor technology.

The gasification of carbon also provides a convenient example of a three-dimensional system to further illustrate the methodology outlined in this paper. A detailed discussion for the technical implications of this case can be found in Shinnar et al., (1982). If carbon is gasified with steam at a temperature below 1300 K, methane is also formed. Methane can also be formed from CO and H_2 by the methanation reaction

$$CO + 3H_2 \rightleftharpoons CH_4 + H_2O \qquad (14e)$$

Steam gasification and subsequent methane formation is the basis of most SNG processes (Shinnar et al., 1982). This leads to the overall reaction

$$2C + 2H_2O \rightleftharpoons CH_4 + CO_2 \qquad (14f)$$

In SNG production, there is a great advantage in maximizing methane formation inside the gasification zone. Reaction 14f has a small heat of reaction, whereas reaction 14a is strongly endothermic and reaction 14e is highly exothermic. Heat is expensive to transfer at high temperature. The heat requirements of the gasification reactor are reduced if both reactions occur simultaneously. Therefore, it is preferable to have a catalyst that directly promotes reaction 14f. At low temperatures the global equilibrium yield is mostly CH_4 and CO_2 (see Figure 15). If a catalyst can be found that promotes any set of reactions that leads to formation of CH_4 and CO_2 at low temperatures, the mechanism of the catalytic reaction will have very little impact on the reachable composition space. At low temperatures the global equilibrium composition contains only small amounts of CO and H_2 and therefore

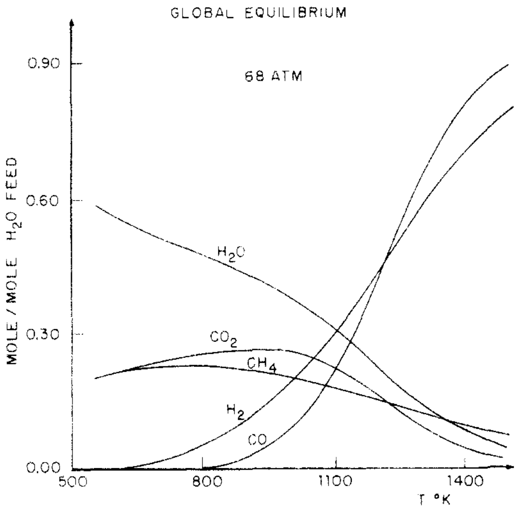


Figure 15. Product yields at equilibrium for the steam–carbon gasification system; products formed are CH_4, CO, CO_2, H_2, and H_2O.

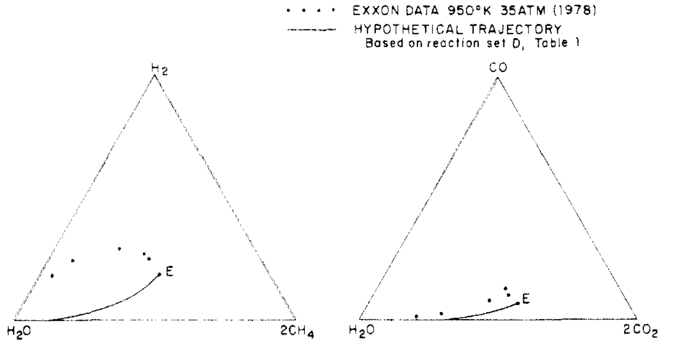


Figure 16. Kinetic trajectories for catalytic (K_2CO_3) steam gasification of char (Shinnar et al., 1982).

all trajectories will be close to that promoted by (14f). Unfortunately, no such catalyst is available. Figure 16 shows an experimental trajectory for the best available catalyst (K_2CO_3 from Exxon, 1978).

In this example one is dealing with six compounds, three element balances, and three independent stoichiometric reactions. However, as it was assumed that carbon conversion in the reactor is incomplete, two element balances (or two triangular composition diagrams) are sufficient to describe the system completely. While such a presentation allows easy plotting of any trajectory, it is harder to present limiting surfaces as these are two-dimensional. We will limit ourselves to the following simple problem. One can find minimum sets of reactions consistent with the experimental trajectory given in Figure 16 by writing down different sets and verifying their consistency with the trajectory using eq 7b. Several such sets are given in Table I. The experimental trajectory is consistent with either set A or B in both of which the primary reaction gasifying carbon is reaction 14a. The data are not consistent with either set D or E, which omit reaction 14a.

The data are also consistent with set F, which contains both primary gasification reactions. The fact that one can form a sufficient set without using any of the primary gasification reactions in sets D and E is of interest to the kineticist and sheds some light on the possible mechanism.

For purposes of kinetic fitting, the most likely set is C, which includes all reactions included in A and B. If one knows which reactions occur and has independent estimates for reaction rates, there is no need to limit oneself to a minimum set of reactions. Sets A, B, and C have a common property. The maximum yield of CH_4 per mole of steam is at global equilibrium, which for these sets gives optimum gasifier conditions.

Table I. Alternative Sets of Reactions for Steam Gasification of Carbon

$C + H_2O \rightleftharpoons CO + H_2$	$C + H_2O \rightleftharpoons CO + H_2$
$CO + H_2O \rightleftharpoons CO_2 + H_2$	$CO + H_2O \rightleftharpoons CO_2 + H_2$
$CO + 3H_2 \rightleftharpoons CH_4 + H_2O$	$C + 2H_2 \rightleftharpoons CH_4$
Set A	Set B
$CO + H_2O \rightleftharpoons CO_2 + H_2$	$2C + 2H_2O \rightleftharpoons CH_4 + CO_2$
$CO + 3H_2 \rightleftharpoons CH_4 + H_2O$	$CO + H_2O \rightleftharpoons CO_2 + H_2$
$C + 2H_2 \rightleftharpoons CH_4$	$CO + 3H_2 \rightleftharpoons CH_4 + H_2O$
$C + CO_2 \rightleftharpoons 2CO$	
Set C	Set D
$3C + 2H_2O \rightleftharpoons CH_4 + 2CO$	$2C + 2H_2O \rightleftharpoons CH_4 + CO_2$
$CH_4 + H_2O \rightleftharpoons CO + 3H_2$	$CO + H_2O \rightleftharpoons CO_2 + H_2$
$CO + H_2O \rightleftharpoons CO_2 + H_2$	$CO + 3H_2 \rightleftharpoons CH_4 + H_2O$
Set E	Set F

Sets D and E are different because both permit methane yields that exceed that of global equilibrium and also allow high methane yield at conditions where the overall heat of reaction is low. It is possible that the basic gasification reactions in set D and E occur as independent kinetic reactions. The data in Figure 16 cannot be used to disprove this. However, it is not necessary to postulate these reactions in order to explain the data. The fact that the data are inconsistent with set D does imply that reaction 14f is not the dominant gasification reaction. If it occurs it is slow compared to (14a).

In the above example and the previous one, the minimum set of independent reactions used has the same dimension as the minimum set of independent stoichiometric reactions. This is not always true as was shown in Figure 2. However, while sets A, B, D, and E are all linearly independent, A and B are the only linearly independent sets which are also kinetically sufficient to describe the reported data. A and B contribute minimum sufficient sets which are also linearly independent.

For efficient modeling of rate equations, it is important to have sets of reactions which are kinetically sufficient. Our method permits a fast check without any calculation of actual reaction rates. The method also provides for efficient screening of new catalysts. One is interested in a catalyst that promotes reaction 14f or 14g (set E, Table I). Initially, such a catalyst may not give a large methane yield, but if an experimental trajectory is consistent with sets D or E, this would indicate that the catalyst operates with a different mechanism than that of K_2CO_3. A hypothetical trajectory is shown in Figure 16 that is consistent with set D. The most interesting catalysts in multiple reaction systems are always those that operate with different mechanisms because they provide a hope that with further development, they may give significantly improved results. The method presented here not only allows efficient and simple screening, but it also allows one to bound the likely improvements. The example also illustrates how to set up a set of reactions that are useful for reactor modeling.

VIII. Effect of Mass Transfer on Thermodynamic Constraints

In the previous sections the reacting phase was assumed to be locally uniform with no restrictions due to mass transfer. However, the results apply to heterogeneous

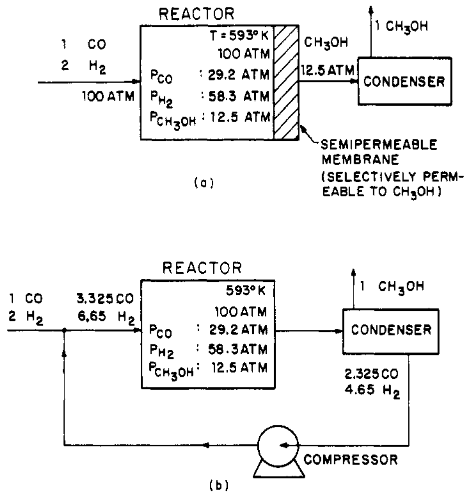


Figure 17. Hypothetical reaction system for converting syngas to methanol: (a) using a semipermeable membrane, selectively permeable to methanol; (b) using separation and recycling.

catalysts as well, so long as the active sites are uniformly accessible to all species. In this section we will examine the effect of transport processes.

As a simple example of a system in which transport processes play a major role, consider the hypothetical process shown in Figure 17a. In general, conversion of syngas in a methanol reactor is strongly limited by global equilibrium. However, if one had a semipermeable membrane that removes methanol only, one could get complete conversion of the syngas to methanol, provided the pressure of the syngas fed to the reactor is high enough. The partial pressure of the methanol inside the reactor must be equal to the partial vapor pressure of the methanol removed through the membrane. The overall process has a decreasing free energy and overcomes the limitations of equilibrium inside the reactor. A similar example was given in Figure 1a, in which we showed how semipermeable membranes can completely overcome thermodynamic constraints, which are the results of kinetic properties of catalysts. Matson (1981) has incorporated semiselective membranes into biochemical reactors and has shown that these reactors are able to reach a product composition not obtainable otherwise.

We use a similar concept, removing syngas from the reactor, condensing the methanol by cooling it and recycling the syngas (Figure 17b). Unfortunately, these selective removal processes are either slow or energy inefficient. The scheme shown in Figure 17b is a highly inefficient, energy intensive "membrane". A more efficient way to accomplish the same task is to use shape-selective zeolites (Weisz and Frillette, 1960). In shape-selective catalysts, pore size is very uniformly controlled and active sites are inside the pores. Diffusion into and out of the pores is, therefore, a strong function of the size and shape of the molecules, and the equivalent diffusion coefficients can vary by a factor of 1000 or more for molecules of different sizes. If one considers each active site as a microreactor, then one has a reactor with selective feed and removal of reactants and products. Again, this has a strong effect on the constraints of reachable composition space discussed earlier as shown in Figure 18.

Assume that for purposes of modeling we can represent the shape-selective catalyst as a reaction which is sheilded from the outside by a semipermeable membrane. This model is illustrated in Figure 18b. Only a limited number of reactants can enter, and only specific products can leave. If one is dealing with a single reaction as in Figure 17, this arrangement has no effect on conversion. In that sense, it is inferior to a true semipermeable membrane. However, in more complex reaction systems, there are strong potential advantages for such a catalyst. Consider the case described in eq 3, and assume that the only reactions occurring with a significant rate are $A \rightleftharpoons B$ and $B \rightleftharpoons C$. In a uniformly accessible reactor, the set of compositions reachable from pure A is contained in the triangle AEF

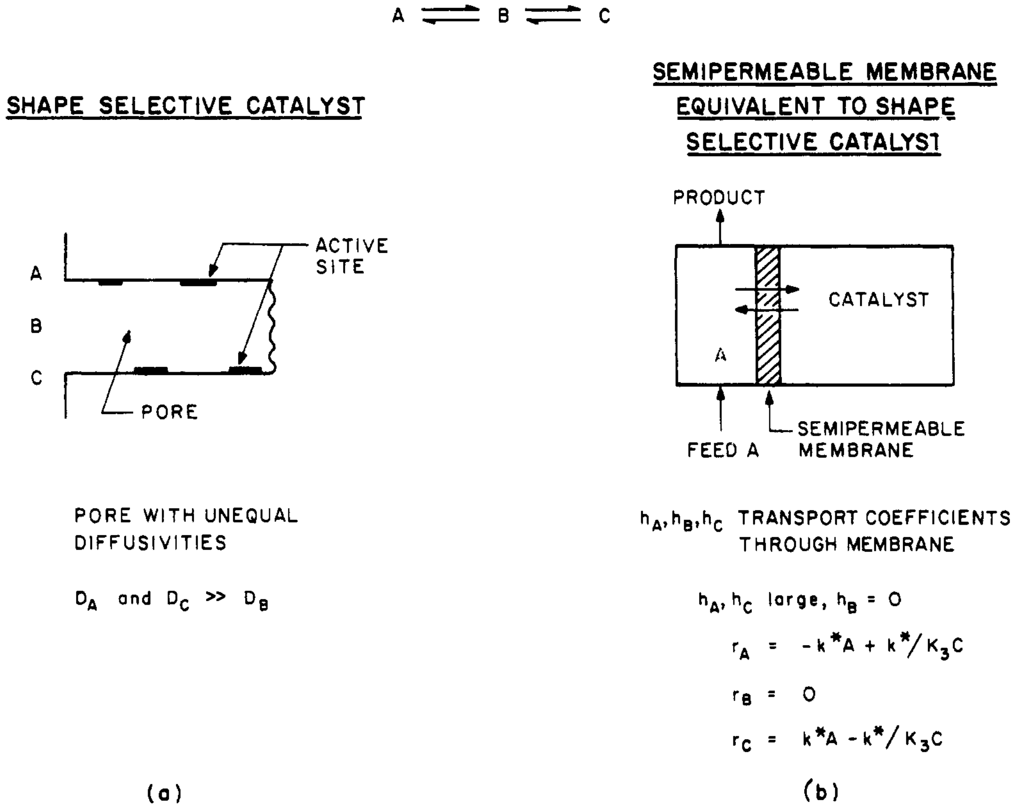


Figure 18. (a) Simplified model for shape selective catalyst. (b) A semipermeable membrane model equivalent to a shape selective catalyst.

in Figure 5. What happens inside the catalyst pore of Figure 18a? If all transport processes are equal, the diffusional resistance of the catalyst pore will not change the reachable space. If the transport processes are very slow, all trajectories will converge to the line AE, regardless of kinetics.

Consider the case where A and C can diffuse rapidly but B has a zero diffusion rate inside the pore and therefore cannot escape from the pore. One can analyze this effect by looking at the equivalent case described by Figure 18b. The diffusion rates inside the pore can be replaced by transport coefficients through the membrane (h_A, h_B, h_C). The difference in diffusion rates is equivalent to saying that the membrane is selective such that h_A and h_C are much larger than h_B. As long as the system is not in equilibrium, A will diffuse into the pore and C will diffuse out of it but B will stay inside the pore. The overall reaction observed outside the pore is A $\rightleftharpoons$ C, which is on the line AH in Figure 5. This trajectory, while permissible by free energy considerations, is totally inaccessible with the same catalyst if the catalytic sites are uniformly accessible by all reactants and products.

In practice, the diffusivity of B will not be zero, but it could be several magnitudes smaller than that of A or of C. Any strong difference will modify the permissible trajectories and the accessible composition space. The transport processes inside a shape-selective catalyst are too complex to be described by simple overall diffusion processes, but a simple overall diffusion model illustrates the way in which they modify the thermodynamic constraints of regular catalysts. Such selective transport processes may exist in other catalysts too, but shape-selective catalysts utilize them in a more efficient manner.

An example of the way such a catalyst modifies the set of accessible compositions is described by Chen et al. (1979) and Wei (1982). Toluene is alkylated over a shape-selective zeolite with methanol to give xylene. In a regular zeolite catalyst the alkylation reaction results directly in an equilibrium mixture of all three isomers of xylene. If the para isomer is preferentially desired, it has to be separated from the mixture, which will be isomerized again. Zeolite catalysts used for alkylation reactions are also good isomerization catalysts. If this reaction is carried out over a shape-selective zeolite, the reaction goes selectively to pure *p*-xylene. An equilibrium mixture is still formed at the active site inside the pore, but *p*-xylene can diffuse out of the pore much faster than the other two xylene isomers. This upsets the equilibrium inside the pore. If transport processes are slow compared to the isomerization reaction itself, the removed para isomer is continuously replaced by further reaction of the two other isomers. With a truly semipermeable membrane that removes only *p*-xylene and water, the reaction could be driven to completion. Use of the shape-selective catalyst allows one, at least in theory, to reach the equilibrium composition of the reaction

$$C_7H_8 + CH_3OH \rightleftharpoons p\text{-}C_8H_{10} + H_2O$$

In practice one would have to be content with somewhat lower selectivity as the membrane is imperfect and the diffusion rate of the undesired product is small but not zero. The importance of shape-selective zeolites is that they allow better selectivity and admit effective reaction pathways that would not be permitted if the same catalysts were uniformly accessible.

Consider, for example, the case shown in Figure 5. The constants are modified in such a way that the global equilibrium contains only traces of C, despite the fact that K_{AC} is large. Again one assumes that the only reactions are A $\rightleftharpoons$ B and B $\rightleftharpoons$ C. In a uniformly accessible catalyst the set of reachable concentrations of species C is very small (see Figure 19, point E). To obtain large concentrations of C one must find a catalyst that directly promotes A $\rightleftharpoons$ C. Shape-selective catalysts provide an alternative way to achieve exactly the same goal without changing the nature of the catalytic reaction at the active site. If one places the catalytic sites inside a shape-selective pore from which C escapes preferentially compared to B, the reaction will be driven toward more C. This is equivalent to finding a catalyst that promotes the reaction A $\rightleftharpoons$ C directly. The changes in the accessible composition space are shown in Figure 19. Clearly, one cannot drive reactions into a direction forbidden by overall free energy considerations. However, selective transport processes remove thermodynamic restrictions that are due to the nature of the catalyst.

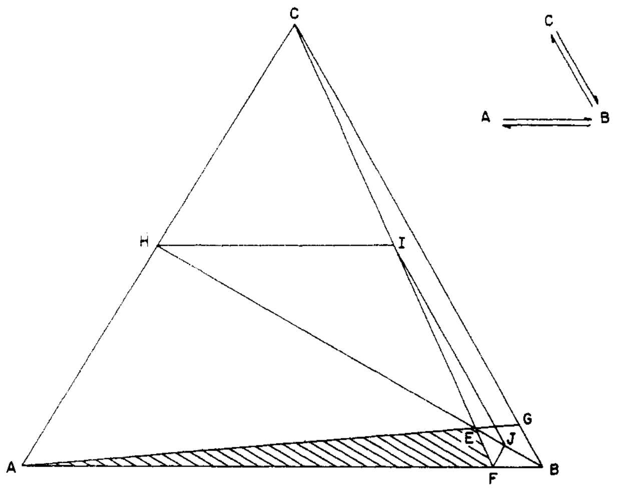


Figure 19. Monomolecular reaction system (A, B, C) with K_{AB} = 10, K_{BC} = 0.1, and initial composition pure A: area AFE, composition space accessible by regular catalyst; AHEF, composition space accessible by shape selective catalyst; AHIJF, composition space theoretically reachable by combination of shape selective and regular catalysts where reaction rates K_{AB} and K_{BC} may be arbitrarily vaired.

In a single reaction the equilibrium constraint is not due to the second law. In Figure 1a pure B has the same free energy as pure A. However, the reaction A $\rightleftharpoons$ B cannot reach pure B from pure A as the trajectory has a minimum at point F. This constraint is a result of free energy topography of the kinetically and stoichiometrically accessible composition space. A semipermeable membrane removes this constraint and changes it to the hard thermodynamic constraint that free energy must decrease. Shape selective catalysts also modify thermodynamic constraints but in a different way. In Figure 5 the restriction that the accessible composition space was the triangle AEF was not a hard constraint, in the sense of an unavoidable consequence of the second law, but a thermodynamic consequence of the kinetic properties of the catalyst. A shape-selective catalyst can remove constraints due to kinetic pathways and is completely equivalent to finding a catalyst that promotes a different set of reactions. It therefore opens completely new avenues for developing new catalysts.

Shape-selective catalysts cannot remove topographical constraints in the composition space in the same ways that truly semipermeable membranes do. All permissible overall reaction trajectories must fulfill the criterion that the free energy decreases along the trajectory. However, they have the advantage of being cheaper and allowing fast overall reactions. In that sense, shape selective catalysts immitate nature. A large fraction of the catalysts utilized by nature are enzymes embedded in semipermeable membranes. A large number of papers have been written showing how diffusion and mixing modify kinetic trajectories, but very little has been written as to how diffusion and, especially, selective diffusion processes modify thermodynamic constraints or kinetic trajectories. This should be an important area for further research, especially since it offers an understandable approach to the modification of catalysts to improve selectivity and yields.

IX. Discussion and Summary

The main goal of this paper was to investigate and demonstrate the usefulness of thermodynamic analysis in chemical reaction engineering, especially in the study of catalytic reactions. Some aspects of thermodynamic analysis, such as computation of global equilibrium, are well known. Thermodynamics is a powerful tool for investigating the impact of design decisions (or design alternatives) in power plants, turbines, and engines. Properly used, it can be just as powerful a tool in the study of chemical reactions and chemical reactors. One must always remember that, in practice, one is dealing with thermodynamics in the context of other outside constraints, whereas energy analysis is thermodynamics without any outside constraints. The properties of a catalyst limit the accessible composition space at different temperatures and pressures far more than the demand of decreasing free energy. While these constraints are kinetic in nature, thermodynamics allows one to obtain reliable bounds in a simple way. The bounds can be computed from a minimum amount of kinetic information. All that one needs to know is the overall stoichiometry of the actually occurring ostensible or overall chemical reactions.

The main concept that allows us to derive such bounds is the demand that every independent chemical reaction must proceed in the direction of decreasing free energy. The only way that one reaction can "transfer its free energy" to another reaction is via heat transfer. This introduces severe constraints on the accessible composition space. A simple method was presented to evaluate the consequences of this constraint. This method can be used in two ways. If one knows what reactions actually occur, one can derive bounds on catalyst performance. On the other hand, if one has the experimental trajectories, one can use thermodynamic data to derive minimum sets of kinetic reactions that are consistent with the experimental results.

This is an important problem in studying and modeling complex chemical reactions. One knows, a priori, that usually one is not describing molecular events, but rather overall complex events, each probably involving many intermediates. There are methods for choosing a set of independent stoichiometric relations to describe all changes. However, such a set is not necessarily suitable for deriving kinetic rate expressions. In many cases, real trajectories cannot be fitted this way. The rate expressions obtained would result in some individual reactions going in the directions of increasing free energy. This paper outlines a method for obtaining sets of chemical reactions for the study of rate equations and for testing in a simple way whether these sets are in agreement with all known trajectories. All trajectories should be realizable from this set of reactions without violating the free energy constraint. This method should provide an efficient tool in setting up empirical rate equations in multicompound systems.

This approach should also have important implications for screening of new catalysts. For a given catalyst, an appropriate set of ostensible reactions describing its kinetic properties can be established from experimental reaction trajectories. From this one can derive bounds on the kinetically accessible composition space. If a single composition obtained via a new modified catalyst falls outside this space, the catalyst has novel properties that should be further investigated. The beauty of this method is that it avoids complex trajectory computations and combines existing kinetic information with available thermodynamic data. This can be a powerful tool for analyzing complex reactions and will be discussed in more detail in a subsequent paper.

Such analyses provide a structured engineering approach to development problems. The chemical engineer relies heavily on multistage processes, each at a different temperature. Temperature and pressure are used to overcome thermodynamic constraints. One can look at temperature as an additional dimension in compound space which provides new permissible pathways. Nature operates isothermally and achieves its high efficiency and selectivity in several ways: (a) coupling of reactions via common intermediates, (b) enlargement of the compound space, and (c) embedding catalysts (enzymes) inside semipermeable selective membranes. Coupling of reactions is a very efficient tool for transfer of free energy between two reactions. It occurs when the detailed mechanism or reaction path involves a common intermediate, thus creating a permissible trajectory along which the free energy decreases. The concentration of the common intermediate is not important. The nature of the reaction mechanism strongly determines the practical thermodynamic constraints of a system. In that sense catalysis affects thermodynamics. While chemical engineers make use of coupling in some reactions ("oxo" intermediates are an example), its use is still limited and may justify greater attention.

Use of selective membranes is now finding increased attention in catalysis. The most interesting and probably most important recent innovation in catalysis, shape-selective catalysts, is based on this principle. The catalytic site is embedded in a narrow uniform pore, in which mass transfer processes depend on molecular size and shape. Such selective mass transfer processes strongly modify the thermodynamic constraints imposed by the nature of the catalytic sites. Understanding these constraints can provide a conceptual framework and an efficient tool for further advances in the development of new catalysts.

Acknowledgment

The work was supported partially by DOE Grant No. DEFG-22-80-PC-30276 and partially by DOE Contract No. DE AC-01-79ET-14811, and the authors want to express their gratitude for this support.

Literature Cited

Aris, R. "Elementary Chemical Reactor Analysis"; Prentice Hall: Englewood Cliffs, NJ, 1969.

Aris, R.; Meh, R. H. S. *Ind. Eng. Chem. Fundam.* **1963**, *2*, 90.

Bazaraa, M. S.; Jarvis, J. J. "Linear Programming and Network Flows"; Wiley: New York, 1977; p 108.

Boudart, M. *J. Phys. Chem.* **1983**, *87*, 2786.

Bowen, R. M. *J. Chem. Phys.* **1968**, *49*, 1625.

Chen, N. Y.; Koed, W.; Dwyer, R. G. *J. Am. Chem. Soc.* **1979**, *101*, 6783.

Exxon Research & Engineering Co., Exxon Catalytic Coal Gasification Process, FE-2369-24, 1978.

Faith, E.; Vermeulen, T. *AIChE J.* **1967**, *13*, 936.

Gorban, A. N.; Yabonskii, G. S.; Bykov, V. I. *Int. Chem. Eng.* **1982**, *22*, 368.

Happel, J.; Sellers, P. H. *Ind. Eng. Chem. Fundam.* **1982**, *21*, 67.

Johnson, J. L. *Adv. Chem. Ser.* **1974**, *131*, 145.

Krambeck, F. J. *Arch. R. Mech. Anal.* **1970**, *38*, 317.

Krambeck, F. J. "Accessible Composition Domain for Chemically Reacting Systems", 4th International Conference on Physicochemical Hydrodynamics; New York, June 1982.

Matson, S. L.; Quinn, J. A. "Coupling Separation and Enrichment to Chemical Conversion in a Membrane Reactor"; AIChE Annual Meeting, New Orleans, Nov 1981.

Onsager, L. *Phys. Rev. Lett.* **1931**, *37*, 405.

Sellers, D. *SIAM J. Appl. Math.* **1967**, *15*, 13.
Shinnar, R.; Fortuna, G.; Shapira, D. *Ind. Eng. Chem. Process Des. Dev.* **1982**, *21*, 728.
Shinnar, R. *Ann. N.Y. Acad. Sci.* **1983**, *404*, 432.
Wei, J.; Prater, C. D. *Adv. Catal.* **1962**, *13*, 204.
Wei, J. *J. Catal.* **1982**, *76*, 433.
Weize, P. W.; Frillette, V. J. *J. Phys. Chem.* **1960**, *69*, 382.

Received for review October 25, 1982
Revised manuscript received May 29, 1984
Accepted November 19, 1984

Fluid Mechanics of Mixing in a Single-Screw Extruder

Ravindran Chella and Julio M. Ottino*

Department of Chemical Engineering, University of Massachusetts, Amherst, Massachusetts 01003

An analysis is presented of mixing in cavity flows and in the single-screw extruder, valid for realistic degrees of mixing (stretching). The limitations of conventional numerical tracking schemes and the general characteristics of flows with weak reorientation are discussed. The analysis gives insight into the physics of the process and suggests ways for improved operation by highlighting the role of operating conditions, the mode of feed introduction, and the role of mixing sections.

Introduction

The objective of this work is to present and analyze a detailed physical model of mixing in a single-screw extruder. For simplicity our results will be presented in terms of a Newtonian fluid, but the extension to other constitutive models should be clear. When possible, our results will be compared with experimental observations and other theoretical analyses. However, one of the main problems of mixing research is the almost total absence of systematic experimental work. (For a review of laminar mixing of polymeric fluids see Ottino and Chella, 1983; the existing experimental work relevant to extruders is reviewed by Arimond, 1984.) From a general viewpoint, there are two principal conceptual problems: (i) how to measure mixing, and (ii) how to implement the measure, i.e., to relate the mixing measure to the velocity field. It is desirable that the analysis be valid, without breakdown or unreasonable computational requirements, up to realistic degrees of mixing.

Usually the starting point for the mixing of immiscible fluids is a two-phase system consisting of thick striations or large blobs. Mechanical mixing causes the striation thicknesses to decrease by several orders of magnitude (10^4). A decrease in length scales of segregation from centimeters to microns is reasonable, and eventually breakup into droplets or ribbons will occur (Figure 1). The stretching and breakup are related in a complicated way to the velocity field.

According to Aref and Tryggvason (1984), the mixing of immiscible fluids may be conveniently divided into two classes: mixing with *passive* and *active* interfaces. For passive interfaces, the motion is *topological*; i.e. there is no breakup, and the interfaces act as markers of the flow. In particular, this is the case of mixing of two immiscible fluids having similar properties and negligible interfacial tension. For active interfaces, generally occurring at small scales where interfacial forces become important, the interfaces interact with the flow and modify it. A conceptual representation of this process is shown in Figure 1. It is convenient to describe mixing in terms of deformation and stretching of passive interfaces and then to add at a small scale the effect of active interfaces (Khakhar et al., 1984).

Given these definitions, the scope of this work (a) and its possible generalizations (b) are given respectively as follows: (1) (a) similar Newtonian fluids; (b) non-Newtonian fluids: trivial provided that the velocity field can be obtained; (2) (a) passive interfaces; (b) active interfaces: deformation and breakup can be added at a small scale (see Khakhar et al., 1984); (3) (a) immiscible fluids; (b) diffusing and reacting fluids: can be treated in terms of the lamellar model (see Chella and Ottino, 1982, for reactions in an extruder; different reaction schemes are treated by Chella and Ottino, 1984).

There are several goals to the present work. In a narrow sense the objective is to use the single-screw extruder, a prototype case of mixing in a reasonably complicated flow field, as a test ground for the application of the lamellar mixing model (Ottino et al., 1981). However, the byproducts of the application are equally important. We provide new insights into the physics and guidance for improved operation. We will be able to compare our approach with other methods: the use of *Weighted Average Total Strain* (WATS) to model mixing in extruders, the *Marker-and-Cell technique* (MAC) to describe mixing in cavity flows, and general numerical methods based on tracking interfaces by means of small displacements.

The Marker-and-Cell technique provides a good means of dispelling a commonly held belief *that once the velocity field is obtained, the mixing problem is essentially solved.* This is far from being true, and actually fairly simple, deterministic, flows can lead to extremely intricate, *chaotic*, mixing patterns (Aref, 1984; Khakhar et al., 1984). Although we do not consider chaotic mixing here, calculating large degrees of mixing in the extruder flow can easily exceed computational limits.

Laminar Mixing

One of the earliest theoretical treatments of laminar mixing was by Spencer and Wiley (1951). They found that the deformation of an interface subject to large unidirectional shear is proportional to the imposed strain, the proportionality constant being a function of the initial orientation. This result has been empirically extrapolated to the analyses of mixing in the extruder (Mohr et al., 1957; McKelvey, 1962; Pinto and Tadmor, 1970; Erwin, 1978).

0196-4313/85/1024-0170$01.50/0