Effective Kinetic Modeling of Multistep Homogeneous Reactions

The Bodenstein approximation for trace-level intermediates is used to reduce multistep reaction-kinetic networks of arbitrary topology and size to simpler forms with only pseudosingle steps between nodes and to derive explicit equations for formation or consumption rates of end members. Networks with nontrace intermediates or steps of higher orders in intermediates must be broken at the respective steps into portions to which the procedure can then be applied separately. A recipe is provided with which rate and yield-ratio equations are readily compiled for networks of arbitrary configuration and complexity. The method is illustrated with an example of a reaction of industrial interest.

Jia-Ming Chern
Friedrich G. Helfferich
Department of Chemical Engineering
Pennsylvania State University
University Park, PA 16802

Introduction

Competitive pressures in chemical industry have made it highly desirable to scale up directly from the laboratory bench, in order to shorten the time between inception of a new process and start-up of the full-scale plant. Safe scale-up of reactors by factors so large calls for reaction-kinetic modeling that correctly reflects the events at the molecular level, that is, mathematics based on reaction networks or pathways. In industrial practice, many reactions of interest have complex multistep networks. Today's computers are sufficiently large, fast, and cheap in calculation time that such reactions can easily be modeled in terms of the differential equations of all molecular steps. However, an *inordinate amount of time and experimental effort* would then be required to verify the postulated network or pathway and to obtain the values of all coefficients and their activation energies. Moreover, design, scale-up, optimization, and on-line control of reactors have to contend with many additional facets, so that streamlined, fast, yet reasonably accurate mathematical models of reaction kinetics are at a premium.

The majority of recent work on kinetics of multistep homogeneous reactions has been concerned with:

1. Identifying what independent reactions between given participants are mathematically possible
2. Classifying such reactions by types
3. Establishing whether and how a possible network or pathway can be distinguished from others

(Sinanoglu, 1978, 1981; Lee and Sinanoglu, 1981; Vajda and Rabitz, 1988; Happel and Sellers, 1989; Vajda et al., 1989). The present approach differs in that it addresses the problem of providing a convenient tool for the practical reaction engineer, facilitating his or her job of arriving at reliable and manageable mathematics for design, scale-up, optimization, and control. The basis is the same as for the classical treatment of single-cycle catalytic and enzymatic reactions (Christiansen, 1953; King and Altman, 1956) and its more recent extensions (Cha, 1968; Segel, 1975). However, the method presented here leads to simpler mathematics (for systems of comparable complexity) and is more general in that it includes a recipe with which explicit rate and yield-ratio equations for networks of arbitrary configuration and complexity can be written without need for derivation.

As an example to illustrate its power, the method will be applied to a moderately complex reaction of industrial interest: hydroformylation of olefin to aldehyde with subsequent hydrogenation to alcohol and side reactions producing paraffin and an alcohol of higher carbon number, Figure 1. Brute-force modeling would require solving 17 simultaneous rate equations (for the 18 participants) with 25 rate coefficients whose values and activation energies would have to be determined. With the method described here, the problem can be reduced to three simultaneous rate equations and three additional algebraic equations with ten more easily determined coefficients (only six coefficients if the gas-cap composition remains constant).

The basis of the method has been described in a previous publication (Helfferich, 1989). The essential elements are:

- Use of the Bodenstein (1913) approximation of quasistationary states for all trace-level intermediates

---

Correspondence concerning this paper should be addressed to F. G. Helfferich.

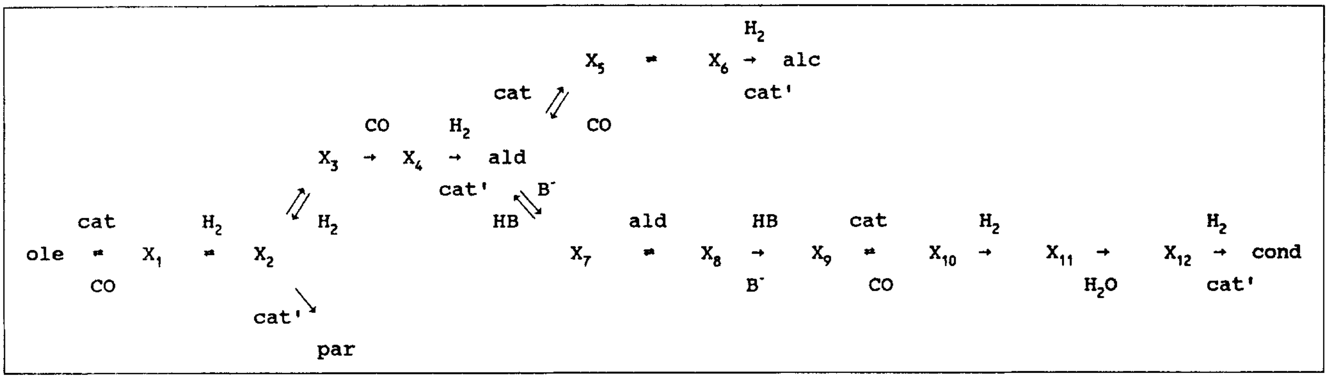


Figure 1. Network for olefin hydroformylation including reactions to paraffin, alcohol, and condensation product.

The dehydration and hydrogenation of aldol ($X_8 \rightarrow$ cond) is probably oversimplified, but without effect of mathematical model

alc = alcohol
ald = aldehyde
cond = condensation product (alcohol)
cat = $HCo(CO)_3PR_3$ hydroformylation catalyst
cat′ = $HCo(CO)_2PR_3$ CO-deficient hydroformylation catalyst
ole = olefin
par = paraffin
B^- = base catalyst
HB = conjugate acid

- Reduction of the actual network to one with only single steps between nodes, end points, and, if any, non-trace intermediates or steps of higher order in trace intermediates
- Establishment of general, explicit rate and yield-ratio equations applicable regardless of the number of steps, locations of nodes, and points of coreactant entries and coproduct exits

This procedure serves two purposes: to reduce to a minimum the number of coefficients and of simultaneous rate equations to be solved, and to facilitate network verification. The latter application has been discussed previously (Helfferich, 1989). The present contribution addresses the reduction of mathematical complexity for effective modeling.

The presentation here is restricted to ordinary homogeneous reactions. The extension to reactions with trace or heterogeneous catalysts is possible in principle at some loss of simplicity, as will be outlined briefly in closing. The presentation is also restricted to the reduction of mathematical complexity and does not include discussions of uniqueness, techniques for parameter fitting, or constraints such as thermodynamic consistency and microscopic reversibility (Sorensen and Stewart, 1980), to which the method contributes nothing new.

Kinetic Basis

For any molecular step, reaction orders are assumed to correspond strictly to reactant molecularities. Thus, for a step $X_i + A \rightarrow X_j$, the rate contribution is taken to be

$$r_{i \rightarrow j} = k_{ij}[X_i][A] \tag{1}$$

where square brackets denote concentrations.

Extensive use is made of the Bodenstein approximation of quasistationary behavior, which postulates that the net rate r_j of concentration variation of an intermediate X_j is small in absolute magnitude compared with the formation and decay rates of that intermediate (Bodenstein, 1913); provided the intermediate remains at trace level, the approximation is valid except during a short initial time period (Aris, 1970).

Simple and Nonsimple Networks

For convenience, we shall call a network or network portion "simple" if

1. All intermediates are and remain at trace level, and
2. No step is of higher than first order in intermediates

If both conditions are met, explicit equations for rates and yield ratios of all end members can be given, regardless of the actual complexity of the network.

To be sure, a simple network or network portion may be of arbitrary size and topology, and may contain steps of higher order as long as the coreactants are not themselves intermediates:

Steps $X_j + A \rightarrow X_k$, $X_j + 2A \rightarrow X_k$, or $X_j + A + B \rightarrow X_k$ are allowed

Steps $2X_j \rightarrow X_k$ or $X_j + X_k \rightarrow X_l$, are not allowed ($X_i$ being trace-level intermediates)

A nonsimple network can be broken at the offending point or points into simple portions, whose equations then have to be solved simultaneously. This will be taken up after the discussion of simple networks and network portions.

Network Reduction

Network reduction replaces the actual network by a much simpler one, with only single, pseudofirst-order steps between adjacent nodes (or nodes and adjacent end points).

Linear simple segments

As shown previously (Helfferich, 1989), any simple network segment between nodes, or between a node and an end member, can be reduced to a pseudosingle reversible or irreversible step with forward and reverse coefficients that are functions of the actual rate coefficients and the concentrations of any coreactants and coproducts, but are independent of the concentrations of the intermediates. Thus, the linear, reversible, simple segment

$$X_j \leftrightarrows X_{j+1} \rightleftarrows \cdots \rightleftarrows X_k \tag{2}$$

with arbitrary number of steps is reduced to a pseudosingle step

$$X_j \rightleftarrows X_k$$

with rate contribution

$$r_{j \rightleftarrows k} = \Lambda_{jk}[X_j] - \Lambda_{kj}[X_k] \tag{3}$$

and with pseudofirst-order forward and reverse segment coefficients Λ_{jk} and Λ_{kj}, which are functions only of the rate coefficients, not of the concentrations of the intermediates; see Eq. 4.

Usually, at least some steps involve coreactants or coproducts. For any such step, the product of the actual, higher order rate coefficient and the concentration(s) of the coreactant(s) is replaced by a pseudofirst-order rate coefficient. For example, for a step

$$X_j + B \rightarrow X_k$$

the rate contribution is

$$r_{j \rightarrow k} = \lambda_{jk}[X_j] \quad (\lambda_{jk} = k_{jk}[B])$$

where k_{jk} is the actual, second-order rate coefficient and λ_{jk} is the pseudofirst-order coefficient. For steps of first order, $k_{jk} = \lambda_{jk}$.

In terms of the λ coefficients, the segment coefficients of the arbitrary linear network segment, Eq. 2, are given by

$$\begin{aligned} \Lambda_{jk} &= \prod_{i=j}^{k-1} \lambda_{i,i+1}/D_{jk} \\ \Lambda_{kj} &= \prod_{i=j}^{k-1} \lambda_{i+1,i}/D_{jk} \end{aligned} \tag{4}$$

where

$$D_{jk} = \sum_{i=j+1}^{k} \left(\prod_{m=j+1}^{i-1} \lambda_{m,m-1} \prod_{m=i}^{k-1} \lambda_{m,m+1} \right) \quad (k > j) \tag{5}$$

(set product $\Pi = 1$ if lower limit exceeds upper) (Helfferich, 1979; Wilkinson, 1980). A derivation is given in the Appendix.

In Eqs. 4, the numerators are the products of all forward or all reverse coefficients, respectively. The denominator D_{jk}, common to both segment coefficients, is found in a way most easily remembered as the sum of the products of all rows of the square matrix of order $k - j$ with elements 1 along the diagonal, with λ coefficients of the mth step in the mth column, and with forward coefficients above and reverse coefficients below the diagonal:

$$\begin{bmatrix} 1 & \lambda_{j+1,j+2} & \lambda_{j+2,j+3} & \cdots & \lambda_{k-2,k-1} & \lambda_{k-1,k} \\ \lambda_{j+1,j} & 1 & \lambda_{j+2,j+3} & \cdots & \lambda_{k-2,k-1} & \lambda_{k-1,k} \\ \lambda_{j+1,j} & \lambda_{j+2,j+1} & 1 & \cdots & \lambda_{k-2,k-1} & \lambda_{k-1,k} \\ & & \cdots & & & \\ \lambda_{j+1,j} & \lambda_{j+2,j+1} & \lambda_{j+3,j+2} & \cdots & 1 & \lambda_{k-1,k} \\ \lambda_{j+1,j} & \lambda_{j+2,j+1} & \lambda_{j+3,j+2} & \cdots & \lambda_{k-1,k-2} & 1 \end{bmatrix} \tag{6}$$

For example, for $j = 0$, $k = 4$ (linear pathway with four reversible steps),

$$D_{jk} = \lambda_{12}\lambda_{23}\lambda_{34} + \lambda_{10}\lambda_{23}\lambda_{34} + \lambda_{10}\lambda_{21}\lambda_{34} + \lambda_{10}\lambda_{21}\lambda_{32}$$

If the segment contains an irreversible step, the respective reverse coefficient is zero. In consequence, the reverse segment coefficient is also zero. Moreover, the forward coefficients of all steps following the irreversible step cancel:

Steps following the first irreversible step in a simple, linear segment do not contribute to the rate behavior.

Equations 3 to 5 for the step sequence, Eq. 2, remain valid if either or both segment end members X_j and X_k are also network end members (that is, reactants or products) or nontrace intermediates (Helfferich, 1989). For any linear, simple pathway with arbitrary number of steps and arbitrary coreactants and coproducts, the equations immediately give the rate of reactant consumption and product formation.

Example. For illustration, consider homogeneous liquid-phase hydroformylation olefin + H_2 + CO → aldehyde with the network shown in Figure 1, disregarding for the time being the side reaction to paraffin and the subsequent reactions of aldehyde. With Eqs. 3 to 5 the rate law of aldehyde formation is:

$$r_{\text{ald}} = \frac{\lambda_{01}\lambda_{12}\lambda_{23}\lambda_{34}[\text{ole}]}{\lambda_{12}\lambda_{23}\lambda_{34} + \lambda_{10}\lambda_{23}\lambda_{34} + \lambda_{10}\lambda_{21}\lambda_{34} + \lambda_{10}\lambda_{21}\lambda_{32}}$$

(steps beyond X_4 do not contribute and the reverse segment coefficient is zero). After replacement of the λ coefficients, collection of terms, and cancellation of [CO]:

$$r_{\text{ald}} = \frac{k_{01}k_{12}k_{23}k_{34}[\text{ole}][\text{cat}][\text{H}_2]}{(k_{12}k_{23}k_{34} + k_{01}k_{21}k_{32})[\text{H}_2] + k_{10}(k_{23} + k_{21})k_{34}[\text{CO}]}$$

which is of the algebraic form of the well-known two-constant Heck-Breslow rate law (Heck and Breslow, 1961):

$$r_{\text{ald}} = \frac{k[\text{ole}][\text{cat}]}{1 + k'[\text{CO}]/[\text{H}_2]}$$

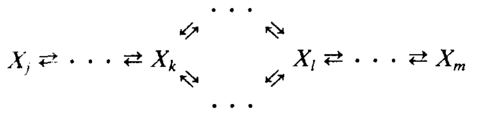
with lumped (constant) coefficients

$$\begin{aligned} k &= k_{01}k_{12}k_{23}k_{34}/(k_{12}k_{23}k_{34} + k_{10}k_{21}k_{32}) \\ k' &= [k_{10}(k_{23} + k_{21})k_{34}]/(k_{12}k_{23}k_{34} + k_{10}k_{21}k_{32}) \end{aligned}$$

Brute-force modeling of the original network would require solving six simultaneous rate equations with nine coefficients (for aldehyde or olefin, the four intermediates, and cat'; the other rates follow from stoichiometry). This has been reduced to one rate equation with two coefficients.

Looped simple segments

As was shown previously (Helfferich, 1989), a looped segment between nodes X_j and X_k, such as



(7)

can be reduced to a pseudosingle step $X_j = X_m$ with collective coefficients Γ_{jm} and Γ_{mj} and rate contribution

$$r_{j\rightleftarrows m} = \Gamma_{jm}[X_j] - \Gamma_{mj}[X_m] \tag{8}$$

where

$$\Gamma_{jm} = \Lambda_{jk}L_{kl}\Lambda_{lm}/(L_{kl}\Lambda_{lm} + \Lambda_{kj}\Lambda_{lm} + \Lambda_{kj}L_{lk})$$
$$\Gamma_{mj} = \Lambda_{kj}L_{lk}\Lambda_{ml}/(L_{kl}\Lambda_{lm} + \Lambda_{kj}\Lambda_{lm} + \Lambda_{kj}L_{lk})$$

and where the loop coefficients L_{kl} and L_{lk} are the sums of the Λ coefficients of the respective parallel pathways connecting X_k and X_l.

Loops are relatively rare. Therefore, in the following, segment coefficients Λ will be used, with the understanding that they are to be replaced by collective coefficients Γ where called for.

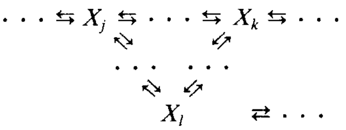
Rate and Yield-Ratio Equations for Branched Simple Networks

Explicit equations for rates of product formation and ratios of such rates (yield ratios) can be obtained for simple networks or network portions of any size and topology. While generally applicable equations can be constructed, they would require a highly complex notation. It is easier to compile the equations for a given network for a set of rules, stated below and derived in the Appendix.

Rate equations

The rate law of formation of any product P_j (or of consumption of any reactant P_j) in a single- or multinode simple network of arbitrary configuration can be compiled with the following rules:

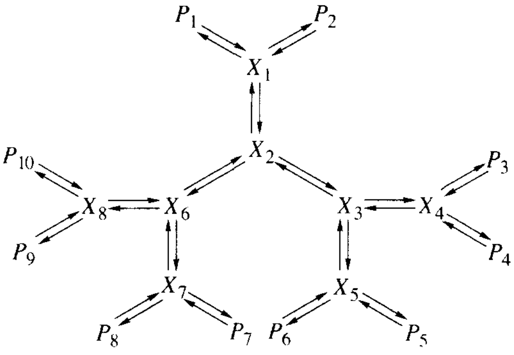
1. The rate law r_{P_j} is composed of additive contributions from all $P_i \rightleftarrows P_j$ ($i \neq j$), each expressing the net rate at which P_j is formed from (or, if negative, reverts to) the respective P_i.
2. Each contribution $P_i \rightleftarrows P_j$ consists of a factor $(\Pi_{ij}[P_i] - \Pi_{ji}[P_j])$, a factor det M_{ij}, and a factor $1/\det M$. Here, Π_{ij} is the product of all forward Λ coefficients (or Γ coefficients, for segments containing loops) on the path from P_i to P_j; Π_{ji} is the product of all reverse Λ coefficients on that path; M is a matrix of order n, with elements S_i ($i = 1, 2, \ldots, n$) along its diagonal, elements Λ_{km} in row k and column m for all X_k and X_m directly connected with one another, and elements zero in all other positions; $M_{ij} = M_{ji}$ is a matrix similarly compiled, but with mth rows and columns omitted, where the m are the indices of all intermediates on the path from P_i to P_j; and S_i is the sum of the Λ coefficients of the reduced steps $X_i \longrightarrow J$, the J being the immediate neighbors of X_i in the network (intermediates or end members).
3. In networks containing loops with branches, such as



the contributions of both parallel pathways to an affected rates $P_i \rightleftarrows P_j$ are additive.

The rules are applicable regardless of whether the end members are reactants and products or are themselves intermediates in a still larger network.

Example. For a fully reversible, simple, ten-ended, eight-node, snowflake-type network with seventeen reduced reversible steps.



(9)

the rate law r_{P_1} is composed of nine contributions of net reaction rates between P_1 and the other nine end members $P_2, P_3, \ldots, P_{10}$. As an example of compilation of these contributions, that of $P_5 \rightleftarrows P_1$ has the form

$$r_{P5\rightleftarrows P1} = (\Pi_{51}[P_5] - \Pi_{15}[P_1])\boldsymbol{detM}_{51}/\boldsymbol{detM}$$

with

$$\Pi_{51} = \Lambda_{P5,5}\Lambda_{53}\Lambda_{32}\Lambda_{21}\Lambda_{1,P1} \quad \Pi_{15} = \Lambda_{P1,1}\Lambda_{12}\Lambda_{23}\Lambda_{35}\Lambda_{5,P5}$$
$$S_1 = \Lambda_{1,P1} + \Lambda_{1,P2} + \Lambda_{12} \quad S_2 = \Lambda_{21} + \Lambda_{23} + \Lambda_{26}$$
$$S_3 = \Lambda_{32} + \Lambda_{34} + \Lambda_{35} \quad S_4 = \Lambda_{4,P3} + \Lambda_{4,P4} + \Lambda_{43}$$
$$S_5 = \Lambda_{5,P5} + \Lambda_{5,P6} + \Lambda_{53}, \quad S_6 = \Lambda_{62}+\Lambda_{67}+\Lambda_{68},$$
$$S_7 = \Lambda_{7,P7} + \Lambda_{7,P8} + \Lambda_{76}, \quad S_8 = \Lambda_{8,P9} + \Lambda_{8,P10} + \Lambda_{86},$$

and determinants

$$\det M = \begin{vmatrix} S_1 & \Lambda_{12} & 0 & 0 & 0 & 0 & 0 & 0 \\ \Lambda_{21} & S_2 & \Lambda_{23} & 0 & 0 & \Lambda_{26} & 0 & 0 \\ 0 & \lambda_{32} & S_3 & \Lambda_{34} & \Lambda_{35} & 0 & 0 & 0 \\ 0 & 0 & \Lambda_{43} & S_4 & 0 & 0 & 0 & 0 \\ 0 & 0 & \Lambda_{53} & 0 & S_5 & 0 & 0 & 0 \\ 0 & \Lambda_{62} & 0 & 0 & 0 & S_6 & \Lambda_{67} & \Lambda_{68} \\ 0 & 0 & 0 & 0 & 0 & \Lambda_{76} & S_7 & 0 \\ 0 & 0 & 0 & 0 & 0 & \Lambda_{86} & 0 & S_8 \end{vmatrix}$$

(Λ coefficients for all existing connections: $X_1 \rightleftarrows X_2$, $X_2 \rightleftarrows X_3$, $X_3 \rightleftarrows X_4$, $X_3 \rightleftarrows X_5$, $X_2 \rightleftarrows X_6$, $X_6 \rightleftarrows X_7$, and $X_6 \rightleftarrows X_8$; all other nondiagonal elements zero) and

$$\det M_{51} = \begin{vmatrix} S_4 & 0 & 0 & 0 \\ 0 & S_6 & \Lambda_{67} & \Lambda_{68} \\ 0 & \Lambda_{76} & S_7 & 0 \\ 0 & \Lambda_{86} & 0 & S_8 \end{vmatrix}$$

(rows and columns $m = 1, 2, 3$, and 5 omitted because these X_m are on the path connecting P_1 and P_5).

In the particular case of the network in 9, the equations can be further condensed by pairing the contributions of end members with common first intermediates, such as P_5 and P_6, which differ only in one $[P_i]$ and two Λ factors. Additional pairing at interior nodes such as X_3 is possible but may be beyond the point of diminishing returns. Lastly, one end-member rate can always be simply calculated from the material balance requirement $\Sigma(r_{Pi}) = 0$.

Yield-ratio equations

In mathematical modeling of branched networks, some rate equations can be replaced by simpler, algebraic equations for ratios of product formation rates (yield ratios). Such equations were previously derived for various configurations of reduced networks (Helfferich, 1989). They assume their simplest and easiest to use forms if the pathways to the respective products are irreversible. If so, only the pathways to the respective two products from the last common intermediate contribute, the rest of the network having no effect. For example, if in the snowflake-type network, 9, the steps $X_5 \rightarrow P_5$ and $X_5 \rightarrow P_6$ were irreversible, the P_6:P_5 yield ratio would simply be

$$R_{65} \equiv r_{P6}/r_{P5} = \Lambda_{5,P6}/\Lambda_{5,P5} \qquad (10)$$

Example. As an example, reconsider the hydroformylation reaction of olefin to aldehyde, this time with side reaction to paraffin, but still without subsequent conversion of aldehyde to end products, Figure 1. The paraffin: aldehyde yield ratio is

$$R_{\text{par,ald}} = \Lambda_{2,\text{par}}/\Lambda_{2,\text{ald}} \qquad (11)$$

and can replace the more complex rate equation of aldehyde formation. The rate of paraffin formation, compiled by the general rules for arbitrary networks, is

$$r_{\text{par}} = \frac{\Lambda_{\text{ole},2}\Lambda_{2,\text{par}}[\text{ole}]}{\Lambda_{2,\text{ole}} + \Lambda_{2,\text{par}} + \Lambda_{2,\text{ald}}} \qquad (12)$$

and reaction stoichiometry provides

$$-r_{\text{ole}} = r_{\text{ald}} + r_{\text{par}} \qquad (13)$$

The four Λ coefficients, obtained with Eqs. 3 to 5, are

$$\Lambda_{\text{ole},2} = \frac{k_{\text{ole},1}k_{12}[\text{cat}][\text{H}_2]}{k_{12}[\text{H}_2] + k_{1,\text{ole}}[\text{CO}]} = \frac{k_{\text{ole},1}[\text{cat}]}{1 + k_a[\text{CO}]/[\text{H}_2]} \qquad (14)$$

$$\Lambda_{2,\text{ole}} = \frac{k_{1,\text{ole}}k_{21}[\text{CO}]}{k_{12}[\text{H}_2] + k_{1,\text{ole}}[\text{CO}]} = \frac{k_{21}}{1 + [\text{H}_2]/k_a[\text{CO}]} \qquad (15)$$

$$\Lambda_{2,\text{ald}} = \frac{k_{23}k_{34}[\text{CO}]}{k_{34}[\text{CO}] + k_{32}[\text{H}_2]} = \frac{k_{23}}{1 + k_b[\text{H}_2]/[\text{CO}]} \qquad (16)$$

$$\Lambda_{2,\text{par}} = k_{2,\text{par}} \qquad (17)$$

where

$$\begin{aligned} k_a &\equiv k_{1,\text{ole}}/k_{12} \\ k_b &\equiv k_{32}/k_{34} \end{aligned} \qquad (18)$$

The Λ coefficients are functions only of the rate coefficients, the (constant) catalyst concentration, and the H_2:CO ratio. At given temperature and constant concentrations of H_2 and CO (constant gas pressure and composition, no mass-transfer limitation) the system is thus reduced to one rate equation, Eq. 12, and two algebraic equations, Eqs. 11 and 13, with four constant Λ coefficients. The dependence of these coefficients on the concentrations of catalyst, H_2, and CO is given by six constant coefficients: $k_{\text{ole},1}$, k_{21}, k_{23}, $k_{2,\text{par}}$, k_a, and k_b. This compares with seven simultaneous rate equations and ten coefficients in the original network.

Nonsimple Networks

The establishment of explicit rate and yield-ratio equations for end members is possible for simple networks or network segments because the restriction to trace-level intermediates and absence of steps of higher order in intermediates results in linear algebra. If either premise is violated, the linearity is destroyed and the formulas for simple networks become inapplicable.

A nonsimple network can be broken into simple portions. The division or divisions are made at any nontrace intermediate and at the reactant of any step of higher order in intermediates. For example, the nonsimple network

$$\begin{tikzcd}[row sep=small] & P_2 \\ P_1 \rightleftarrows X_1 \rightleftarrows K \rightleftarrows X_2 \arrow[ur, shift left] \arrow[ur, leftarrow, shift right] \arrow[dr, shift left] \arrow[dr, leftarrow, shift right] & \\ & P_3 \end{tikzcd}$$

is broken into the simple portions

$$\begin{tikzcd}[row sep=small] & & P_2 \\ P_1 \rightleftarrows X_1 \rightleftarrows K & \text{and} \quad K \rightleftarrows X_2 \arrow[ur, shift left] \arrow[ur, leftarrow, shift right] \arrow[dr, shift left] \arrow[dr, leftarrow, shift right] & \\ & & P_3 \end{tikzcd}$$

or the nonsimple network

$$P_1 \rightleftarrows X_1 \rightleftarrows X_2 \overset{X_1}{\rightleftarrows} P_2$$

into

$$P_1 \rightleftarrows X_1 \rightleftarrows X_2 \quad \text{and} \quad X_2 \overset{X_1}{\rightleftarrows} P_2$$

A reduction to single, explicit rate and yield-ratio equations for the end members of the complete network is no longer achieved.

Rather, the rate equations for the end members of the "piecewise simple" portions must be solved simultaneously. Nevertheless, in many cases of practical interest the reduction in number of simultaneous equations (which often are differential equations) and coefficients is considerable, with corresponding advantages for numerical modeling.

In some cases it is possible to justify an additional assumption that permits explicit rate and yield-ratio equations to be obtained even if the network is nonsimple. A typical example is the classical treatment of chain reactions with second-order termination (Frost and Pearson, 1961; Hill, 1977; Fogler, 1986). Here, the assumption of dissociation equilibrium in the initiation/termination step (e.g., $Br_2 = 2Br\bullet$) permits the rate equation of the nonsimple termination step to be replaced by an algebraic equation, thereby removing its interference with linear network algebra.

Example. As an example of a moderately complex reaction, consider the entire hydroformylation network in Figure 1, now including hydrogenation of aldehyde to alcohol and base-catalyzed aldehyde condensation eventually leading to an alcohol with twice the number of carbon atoms. The figure shows the most probable pathways of the contributing reactions (Heck and Breslow, 1961; Falbe, 1970; Helfferich, 1979; Ingold, 1969; Van Winkle, unpublished results, 1977). However, so as not to obscure the principle of the method with too much detail, double-bond isomerization of the olefin under reaction conditions and the resulting formation of different aldehyde and alcohol isomers as well as the effect of equilibria between different cobalt complexes have not been included.

The network is nonsimple for two reasons: Aldehyde is a nontrace intermediate, and also is a coreactant in another step, $X_7 + \text{ald} = X_8$. Breaking the network at the aldehyde and reducing the fragments one obtains two simple, reduced portions

$$\begin{tikzcd} \text{ole} \arrow[r, shift left] & X_2 \arrow[l, shift left] \arrow[r] \arrow[d] & \text{ald} & \text{and} & \text{ald} \arrow[r] \arrow[d] & \text{alc} \\ & \text{par} & & & \text{cond} & \end{tikzcd} \qquad (19)$$

with five end members and six Λ coefficients.

The rate equations for the intermediate and the three products are

$$r_{\text{ald}} = \frac{\Lambda_{\text{ole},2}\Lambda_{2,\text{ald}}[\text{ole}]}{\Lambda_{2,\text{ole}} + \Lambda_{2,\text{ald}} + \Lambda_{2,\text{par}}} - \Lambda_{\text{ald,alc}}[\text{ald}] - 2\Lambda'_{\text{ald,cond}}[\text{ald}]^2 \qquad (20)$$

$$r_{\text{alc}} = \Lambda_{\text{ald,alc}}[\text{ald}] \qquad (21)$$

$$r_{\text{cond}} = \Lambda_{\text{ald,cond}}[\text{ald}] = \Lambda'_{\text{ald,cond}}[\text{ald}]^2 \qquad (22)$$

and Eq. 12 for r_{par}. (since $\Lambda_{\text{ald,cond}}$ contains [ald] as a factor, a new coefficient $\Lambda'_{\text{ald,cond}}$ without this factor proves more convenient.) The two new Λ coefficients in addition to those in Eqs. 14 to 17 are given by

$$\Lambda_{\text{ald,alc}} = \frac{k_{\text{ald},5}k_{56}k_{6,\text{alc}}[\text{cat}][H_2]}{k_{56}k_{6,\text{alc}}[H_2] + k_{5,\text{ald}}k_{6,\text{alc}}[H_2][CO] + k_{5,\text{ald}}k_{65}[CO]}$$

$$= \frac{k_c[\text{cat}][H_2]/[CO]}{1 + k_d[H_2] + k_e[H_2]/[CO]} \qquad (23)$$

and

$$\Lambda'_{\text{ald,cond}} = \frac{k_{\text{ald},7}k_{78}k_{89}[B^-][HB]}{k_{78}k_{89}[\text{ald}][HB] + k_{7,\text{ald}}k_{89}[HB]^2 + k_{7,\text{ald}}k_{87}[HB]}$$

$$= \frac{k_f[B^-]}{1 + k_g[\text{ald}] + k_h[HB]} \qquad (24)$$

where

$$k_c \equiv k_{\text{ald},5}k_{56}k_{6,\text{alc}}/k_{5,\text{ald}}k_{65} \qquad k_d \equiv k_{6,\text{alc}}/k_{65}$$
$$k_e \equiv k_{56}k_{6,\text{alc}}/k_{5,\text{ald}}k_{65} \qquad k_f \equiv k_{\text{ald},7}k_{78}k_{89}/k_{7,\text{ald}}k_{87}$$
$$k_g \equiv k_{78}k_{89}/k_{7,\text{ald}}k_{87} \qquad k_h \equiv k_{89}/k_{87}$$

The rates of consumption or formation of olefin, H_2, CO, and water can be calculated from the reaction stoichiometries. Of the four simultaneous rate equations, Eqs. 12, 20, 21, 22, that for paraffin can be replaced by a yield-ratio equation

$$\frac{r_{\text{par}}}{r_{\text{ald}} + r_{\text{alc}} + 2r_{\text{cond}}} = \frac{\Lambda_{2,\text{par}}}{\Lambda_{2,\text{ald}}} = \frac{k_{2,\text{par}}}{k_{23}}\left(1 + k_b\frac{[H_2]}{[CO]}\right) \qquad (25)$$

which involves no rates and no concentrations other than of H_2 and CO. With known k coefficients and at constant concentrations of H_2 and CO, modeling will thus require solving three simultaneous rate equations, Eqs. 20 to 22, with four stoichiometric equations and yield-ratio Eq. 25 to obtain the net rates of formation or consumption of all participants.

The equations above are in terms of only the six Λ coefficients of the reduced network, 19. These are functions of the concentrations of the coreactants H_2 and CO and the catalyst species *cat*, B^-, and HB; and $\Lambda'_{\text{ald,cond}}$ in addition, of the aldehyde concentration. However, since catalyst is not consumed, the three catalyst-species concentrations remain constant (granted the premise that *cat'* is a trace intermediate while *cat* is at much higher concentration). The same is true for the concentrations of H_2 and CO if the reaction is conducted under conditions of constant gas-cap composition and no gas–liquid mass-transfer limitation, as is often the case. $\Lambda'_{\text{ald,cond}}$ as a function of the aldehyde concentration could be expressed by two constants, but is found in practice to be constant itself (see below). Thus, the reaction under the stated conditions can be described by only six constant coefficients.

The dependence of the six Λ coefficients on the concentrations of H_2, CO, and the catalyst species is described by twelve coefficients in Eqs. 14 to 17, 23, and 24: the four rate coefficients k_{01}, k_{21}, k_{23}, and $k_{2,\text{par}}$, and the eight lumped coefficients k_a to f_h. Of these twelve, two turn out not to be needed: As is found by experiment, formation of alcohol from aldehyde is of order -1 in CO, and formation of condensation product is second order in aldehyde, so that k_e and k_g must be negligible; see Eqs. 23 and 24, respectively. This ensures the constancy of $\Lambda'_{\text{ald,cond}}$ anticipated above.

Experimentally determined yield ratios of paraffin to the other products, given by Eq. 29, can be used as an easy way to determine k_b and the ratio $k_{2,\text{par}}/k_{23}$ (plot yield ratio vs. $[H_2]/[CO]$, evaluate slope and intercept). Similarly, k_c and k_d as well as k_f and k_h are readily determined from plots of $[\text{ald}][\text{cat}][H_2]/[CO]r_{\text{alc}}$ vs. $[H_2]$ and of $[\text{ald}]^2[B^-]/r_{\text{cond}}$ vs. [HB] with Eqs. 23 and 24. This leaves only the five coefficients

k_{01}, k_{21}, k_a, k_{23}, and $k_{2,\mathrm{par}}$ still to be fitted, with the ratio of the last two already established.

Brute-force modeling of the original network would require solving 17 simultaneous rate and three stoichiometric equations with 25 coefficients, for which values and activation energies would have to be determined. This has been reduced to three rate equations and five stoichiometric and yield-ratio equations with ten coefficients (six if gas-cap is constant).

Application to Trace-level and Heterogenous Catalysis

The method described is applicable in principle to catalytic reactions. In fact, the practical example chosen was one of homogeneous catalysis. As long as the fractional amount of catalyst bound in the form of reaction intermediates is negligible (as in many reactions of homogeneous catalysis including the hydroformylation example), the free catalyst can be treated in the same way as the reactants, and the equations apply without change. However, if this is not true (as is the rule in enzyme catalysis), the rate equations must be expressed in terms of total concentration of catalyst (free plus bound). The resulting equations are of the same form, except that the denominator D_{jk} in Eq. 4 then contains a much larger number of terms (Christiansen, 1953; King and Altman, 1956). Simpler mathematics is obtained if this approach is combined with the assumption that most steps are at equilibrium (Cha, 1968; Segel, 1975).

In heterogeneous catalysis the situation is similar: The mathematics of the surface reaction is analogous, free and bound catalyst here corresponding to the unoccupied and occupied portions of the catalyst surface, respectively, and in most cases of practical interest the surface is largely covered. Again, the introduction of additional approximations of rate-controlling steps, equilibrium steps, and most abundant surface intermediates can reduce the complexity (Temkin, 1965; Boudart, 1972). An added complication is that equilibria and rates of adsorption and desorption and possible surface heterogeneity must be considered.

In both instances, the general formalism for compilation of rate and yield-ratio equations in the present approach offers the opportunity to extend current work systematically to networks of arbitrary configurations and complexity.

Conclusions

The method presented here is seen to reduce the set of rate equations for all participants to just those for the end members, nontrace intermediates, and reactants of steps (if any) of higher order in intermediates. For irreversible reactions, it also allows some of the remaining rate equations to be replaced by simpler yield-ratio equations. Moreover, for simple networks or for the simple portions into which any nonsimple network can be divided, the method is seen to provide general equations from which the rate and yield-ratio equations of any reaction with network of arbitrary configuration and complexity can be directly recovered by appropriate setting of indices, collection of terms, and cancellations.

The method is at its best when applied to reactions with long strings of trace intermediates. Its most serious limitation is its lesser effectiveness for reactions in which a majority of steps involve nontrace intermediates or are of higher order in intermediates.

Acknowledgment

This work was conducted under the auspices of the Pennsylvania State University, whose support we gratefully acknowledge.

Notation

A, B = coreactants and coproducts
$[i]$ = concentration of i, mol V^{-1}
k_{ij} = rate coefficient of step $i \rightarrow j$, dimension depends on molecularity
k, k', k_i = lumped coefficients, dimension depends on network
K, L = nontrace intermediates
$\boldsymbol{M}, \boldsymbol{M}^{\circ}, \boldsymbol{M}_{ij}$ = matrices for arbitrary branched networks
P_i = network end members (reactants and products)
r_i = net rate of formation of i, mol $\cdot V^{-1} \cdot t^{-1}$
$r_{i \rightarrow j}$ = rate contribution of step or segment $i \rightarrow j$, t^{-1}
$r_{i \rightleftarrows j}$ = net rate contribution of step or segment $i \rightleftarrows j$, t^{-1}
$R_{ij} \equiv r_i / r_j$ = yield ratio of products i and j
S_i = sum of Λ coefficient of all segments originating from node i in branched networks, t^{-1}
X_i = trace-level intermediates (X suppressed in subscripts)

Greek letters

$\Gamma_{i \rightarrow j}$ = collective coefficient of looped segment $i \rightarrow j$, Eq. 8, t^{-1}
λ_{ij} = pseudo-first order coefficient of step $i \rightarrow j$, t^{-1}
Λ_{ij} = segment coefficient of segment $i \rightarrow j$, Eqs. 4, t^{-1}
Π_{ij} = product of Λ coefficients along path $i \rightarrow j$, t^{j-i}

Literature Cited

Aris, R., "Chemical Kinetics and the Ecology of Mathematics," *Am. Sci.*, **58**, 419 (1970).

Bodenstein, M., "Eine Theorie der photochemischen Reaktionsgeschwindigkeiten," *Z. Phys. Chem.*, **85**, 329 (1913).

Boudart, M., "Two-Step Catalytic Reactions," *AIChE J.*, **18**, 465 (1972).

Cha, S., "A Simple Method for Derivation of Rate Equations for Enzyme-Catalyzed Reactions Under the Rapid Equilibrium Assumption or Combined Assumptions of Equilibrium and Steady State," *J. Biol. Chem.*, **243**, 820 (1968).

Christiansen, J. A., "The Elucidation of Reaction Mechanisms by the Method of Intermediates in Quasistationary Concentrations," *Adv. Catal.*, **6**, 311 (1953).

Falbe, J., *Carbon Monoxide in Organic Synthesis*, Springer, Berlin (1967).

Fogler, H. S., *Elements of Chemical Reaction Engineering*, Prentice-Hall, Englewood Cliffs, NJ, sec. 7.1.2 (1986).

Frost, A. A., and R. G. Pearson, *Kinetics and Mechanism*, 2d ed., Wiley, New York, ch. 10 (1961).

Happel, J., and P. H. Sellers, "The Characterization of Complex Systems of Chemical Reactions," *Chem. Eng. Commun.*, **83**, 221 (1989).

Heck, R. F., and D. S. Breslow, "The Reaction of Cobalttetracarbonyl with Olefins," *J. Am. Chem. Soc.*, **83**, 4023 (1961).

Helfferich, F. G., "Hydroformylation Model Development: Reaction Network and Kinetics," Rept. WCR 68–79, Shell Development Co., Houston, TX (1979).

———, "Systematic Approach to Elucidation of Multistep Reaction Networks," *J. Phys. Chem.*, **93**, 6676 (1989).

Hill, C. G., Jr., *An Introduction to Chemical Engineering Kinetics and Reactor Design*, Wiley, New York, sec. 4.1.2, 4.2 (1977).

Ingold, C. K., *Structure and Mechanism in Organic Chemistry*, 2d ed., Cornell Univ. Press, Ithaca (1969).

King, E. L., and C. A. Altman, "A Schematic Method of Deriving the Rate Laws of Enzyme-Catalyzed Reactions," *J. Phys. Chem.*, **60**, 1375 (1956).

Lee, L.-S., and O. Sinanoglu, "Reaction Mechanisms and Chemical Networks—Types of Elementary Steps and Generation of Laminar Mechanisms," *Z. Phys. Chem.* (Munich), **125**, 129 (1981).

Segel, I. H., *Enzyme Kinetics: Behavior and Analysis of Rapid Equilibrium and Steady State Enzyme Systems*, Wiley, New York, 1975.

Sinanoglu, O., "Finding All Possible a priori Mechanisms for a Given Type of Overall Reaction," *Theor. Chim. Acta*, **48**, 287 (1978).

———, "1- and 2-Topology of Reaction Networks," *J. Math. Phys.* (NY), **22**, 1504 (1981).

Sorensen, J. P., and W. E. Stewart, "Structural Analysis of Multicomponent Reaction Models," *AIChE J.*, **26**, 98 (1980).

Temkin, M. I., "Two-Step Kinetics on Inhomogeneous Surfaces," *Dokl. Akad. Nauk SSSR*, **161**, 166 (1965).

Vajda, S., and H. Rabitz, "Identifiability and Distinguishability of First-Order Reaction Systems," *J. Phys. Chem.*, **92**, 701 (1988).

Vajda, S., H. Rabitz, E. Walter, and Y. Lecourtier, "Qualitative and Quantitative Identifiability Analysis of Nonlinear Chemical Kinetic Models," *Chem. Eng. Commun.*, **83**, 191 (1989).

Wilkinson, F., *Chemical Kinetics and Reaction Mechanisms*, Van Nostrand Reinhold, New York, sec. 3.6, 3.10 (1980).

Appendix

Derivation of formula for linear simple segments

The Bodenstein approximation of quasistationary behavior of an intermediate X_i in a step sequence $X_j \rightleftarrows \ldots \rightleftarrows X_k$ assumes that the net rate of formation of the intermediate is negligible compared with the net rates of its formation from the previous intermediate and of decay to the next intermediate:

$$r_i = r_{i-1\rightleftarrows i} - r_{i\rightleftarrows i+1} \simeq 0$$

so that

$$r_{i-1\rightleftarrows i} = r_{i\rightleftarrows i+1} \simeq 0$$

that is, the net rate is the same for every successive step. In matrix form:

$$M^* \begin{bmatrix} 1/r \\ [X_{j+1}]/r \\ [X_{j+2}]/r \\ \cdots \\ [X_{k-1}]/r \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ \cdots \\ 1 \end{bmatrix}$$

where M *is a square matrix of order* $k = j\text{-}l$ *given by*

$$M = \begin{vmatrix} \lambda_{j,j+1}[X_j] & -\lambda_{j+1,j} & 0 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_{j+1,j+2} & -\lambda_{j+2,j+1} & 0 & \cdots & 0 & 0 \\ 0 & 0 & \lambda_{j+2,j+3} & -\lambda_{j+3,j+2} & \cdots & 0 & 0 \\ & & & \cdots & & & \\ 0 & 0 & 0 & 0 & \cdots & \lambda_{k-2,k-1} & -\lambda_{k-1,k-2} \\ -\lambda_{k,k-1}[X_k] & 0 & 0 & 0 & \cdots & 0 & \lambda_{k-1,k} \end{vmatrix}$$

Using Cramer's rule to solve for $1/r$ one finds:

$$1/r = \det M^\circ/\det\ M \qquad \text{(A1)}$$

where

$$M^\circ = \begin{vmatrix} 1 & -\lambda_{j+1,j} & 0 & \cdots & 0 & 0 \\ 1 & \lambda_{j+1,j+2} & -\lambda_{j+2,j+1} & \cdots & 0 & 0 \\ & & & \cdots & & \\ 1 & 0 & 0 & \cdots & \lambda_{k-2,k-1} & -\lambda_{k-1,k-2} \\ 1 & 0 & 0 & \cdots & 0 & \lambda_{k-1,k} \end{vmatrix}$$

The two determinants are easily evaluated by expansion into their minors along their first columns:

$$\det M = \prod_{i=j}^{k-1} \lambda_{i,i+1}[X_j] - \prod_{i=j}^{k-1} \lambda_{i+1,i}[X_k]$$

$$\det M^\circ = \sum_{i=j+1}^{k} \left(\prod_{m=j+1}^{i-1} \lambda_{m,m-1} \prod_{m=j}^{k-1} \lambda_{m,m+1} \right)$$

The use of these expressions in Eq. A1 gives Eq. 3.

The derivation follows the lines of that by Christiansen (1953) for catalytic "closed" step sequences, that is, for cases in which X_j and X_k are identical and are the free catalyst. However, while Christiansen based his derivation on steady-state behavior of the catalytic closed sequence, he did not invoke the restriction to trace-level intermediates. His final result implicitly gives the rate as a function of the rate coefficients and the total amount of catalyst (free and bound in form of the intermediates). This formula is more complex, containing $(k\text{-}j)^2$ terms in the denominator rather than only the $k\text{-}j$ terms of Eq. 3 (see also Appendix of Helfferich, 1989).

Derivation of equations and rules for arbitrary networks

The equations and rules for branched simple networks can be derived by lengthy matrix algebra. Here, instead, a much simpler reduction technique is shown, one that can be applied to networks of any kind of topology.

As derived previously (Helfferich, 1989), the rate of formation of a product P_j in a single-node, reduced, simple network

$$\begin{array}{ccccc} & & P_{j+1} & & \\ & & \upharpoonleft\downharpoonright & & \\ P_{j+2} & \rightleftarrows & X_m & \rightleftarrows & P_j \end{array} \qquad \text{(A2)}$$

is

$$r_{P_j} = \sum_{i=j+1}^{j+2} \left(\Pi_{\mathrm{Pi},Pj}\,[P_i] - \Pi_{\mathrm{Pj},Pi}\,[P_j] \right) / S_m$$

(or with summation to higher value of i if X_m is connected to more end members). This equation can instead be written

$$r_{Pj} = \sum_{i=j+1}^{j+2} (G_{ij}[P_i] - G_{ji}[P_j]) \tag{A3}$$

with

$$\begin{aligned} G_{ij} &\equiv \Pi_{Pi,Pj}/S_m, \\ G_{ji} &\equiv \Pi_{Pj,Pi}/S_m \end{aligned} \tag{A4}$$

corresponding to a still farther reduced network with G coefficients for its steps:

$$\begin{tikzcd} & P_{j+1} \arrow[d, shift left] \\ P_{j+2} \arrow[r, shift left] & P_j \arrow[l, shift left] \arrow[u, shift left] \end{tikzcd} \tag{A5}$$

Equation A3 also holds if A2 is a network portion with an intermediate X_j as connecting node instead of a product P_j. For two simple network portions linked by a common intermediate, X_j, the actual reduced network with Λ coefficients, say,

$$\begin{tikzcd} & P_{j+1} \arrow[d, shift left] & & P_{j+3} \arrow[d, shift left] & \\ P_{j+2} \arrow[r, shift left] & X_m \arrow[l, shift left] \arrow[u, shift left] \arrow[r, shift left] & X_j \arrow[l, shift left] \arrow[r, shift left] & X_n \arrow[l, shift left] \arrow[u, shift left] \arrow[r, shift left] & P_{j+4} \arrow[l, shift left] \end{tikzcd} \tag{A6}$$

is further reduced to

$$\begin{tikzcd} P_{j+1} \arrow[dr, shift left] & & P_{j+3} \arrow[dl, shift left] \\ & X_j \arrow[ul, shift left] \arrow[ur, shift left] \arrow[dl, shift left] \arrow[dr, shift left] & \\ P_{j+2} \arrow[ur, shift left] & & P_{j+4} \arrow[ul, shift left] \end{tikzcd} \tag{A7}$$

with G coefficients and rates

$$r_{P_{j+1}} = \sum_{i \neq j+1} (\Pi^*_{ij+1}[P_i] - \Pi^*_{j+1,i}[P_{j+1}])/S^*_j \tag{A8}$$

(summation over all other end members P_i to which the common intermediate is connected). Here, the $\Pi^*_{i,k}$ coefficients are defined as the products of the G coefficients along the path from P_i to P_k; for a reduced network such as A7 with one common intermediate X_j:

$$\Pi^*_{ik} \equiv G_{ij}G_{jk} \tag{A9}$$

and similarly, any S^*_j is the sum of the G coefficients of all pseudosingle steps connecting the intermediate X_j with its immediate neighbors in the further reduced network. For instance, for the reduced network A6, further reduced to A7 and its rates A8:

$$S^*_j = \sum_{i=j+1}^{j+4} G_{ji} \tag{A10}$$

In this fashion, any single- or multiple-node simple network portion can be linked with others and rate equations in terms of Λ coefficients can be obtained by appropriate substitutions according to Eqs. A9, A10, and A4. This further reduction of networks eliminates all node intermediates and leads to explicit equations for the product formation and reactant consumption rates in terms of the concentrations of network end members and coreactants only. For example, in the snowflake-type reduced network, 9, two pairs of single-node portions

$$\begin{tikzcd} P_i \arrow[dr, shift left] & & \\ & X_m \arrow[ul, shift left] \arrow[dl, shift left] \arrow[r, shift left] & X_n \arrow[l, shift left] \\ P_j \arrow[ur, shift left] & & \end{tikzcd}$$

are linked at their respective common intermediate X_n, and the two resulting larger portions are then linked with a fifth single-node portion, to give rate equations and coefficients as shown in the main text.

Manuscript received Mar. 2, 1990, and revision received June 4, 1990.