Theoret. Chim. Acta (Berl.) 45, 215–223 (1977)

THEORETICA CHIMICA ACTA

A Possible Construction of a Complex Chemical Reaction Network

I. Definitions and Procedure for Construction

István Nemes, Tamás Vidóczy, László Botár and Dezsö Gál

Central Research Institute for Chemistry of the Hungarian Academy of Sciences, Pusztazeri út 59–67, H-1025 Budapest, Hungary

A procedure is suggested for the construction of chemical reaction networks. We define the kinetic communication as a transfer of atoms or atomic groups between two species and determine all the kinetic communications occurring in the possible mechanism of a complex chemical process. The set of kinetic communications is the basis of the communication matrices resulting in the complete network of the overall reaction.

Limiting the consideration for certain types of kinetic communications we obtain the reaction subnetworks and selecting arbitrarily species among those participating in the possible mechanism we introduced the concept of the partial subnetworks which correspond to subsets of the complete network.

By the simple analysis of the subnetworks it is easy to obtain the sequence network indicating the pathways *via* which the selected species are formed in the course of the overall process, by the transfer of chosen atoms or atomic groups.

Key words: Reaction network, complex chemical ~ – Kinetic communication

1. Introduction

New work has been published recently on the problem of rationalizing reaction mechanisms in a formal way. Sellers worked out a mathematical formalism for mechanisms using finitely generated Abelian groups [1], which – though of a general character – if applied to complex chemical processes requires a rather wearisome mathematical treatment. Sinanoğlu introduced a theory allowing all possible mechanisms to be figured out *a priori* and to be drawn as simple networks [2]. Thus the Sinanoğlu theory which includes all the chemically important elementary reaction types is of much practical value especially when used for mechanisms up to 5–8 elementary processes.

In the present work we develop the methodology for chemical reactions with some possible mechanisms consisting of any number of elementary processes. This formalization would be of importance in 1) planning the experimental study of the reaction, 2) excluding certain network elements during experimental verification of the network resulting in reducing (or eventually extending) the possible mechanism, and 3) proper selection of the elementary processes included in the possible mechanism, the detailed study of which is essential concerning the overall reaction.

With respect to degenerate branching chain reactions (such as hydrocarbon oxidation processes) reaction networks have been used earlier in the form of reaction schemes comprising sequence relations of the formation of the selected products. Their "construction", however, was based mainly on the intuition of the authors and not on the mathematical remodeling of the possible mechanism [3–5].

Our first paper presents a method which allows drawing the different chemical reaction networks derived from the possible complex mechanism of the overall process in a quite simple way while the second paper shows the application of the procedure to two complex systems.

2. Definitions

Study of the complex process starts with compilation of the possible mechanism which, by definition, includes all possible elementary processes not excluded experimentally and collected from literature data as well as by using chemical evidence [6, 7].

A single elementary event is referred to as an elementary step and the corresponding macroscopic transformation an elementary process governed by a cumulative statistical rate law. Two reversibly coupled elementary processes are called an elementary reaction. Consequently the possible mechanism is the set of elementary processes. The species participating in the possible mechanism are denoted by $A_1 \ldots A_j \ldots A_p \ldots A_N$.

Definition 1. There is direct *kinetic communication* between two species A_j and A_p if, among the elementary processes of the possible mechanism there is at least one, the initial component and the product of which are A_j and A_p, respectively, and A_p contains *at least* one atom of A_j. (Thus according to Definition 1, we do not consider an electron transfer between charged particles or energy transfer as a kinetic communication, although the definition can be easily extended to such phenomena.)

Let a simple arrow denote the kinetic communication $A_j \rightarrow A_p$ and, according to Definition 1 direct kinetic communication is a *one-step communication.*

If the possible mechanism contains reversible elementary processes kinetic communication exists in both directions (e.g. elementary reaction $A_j \rightleftharpoons A_p$ determines the following communications: $A_j \rightarrow A_p$ *and* $A_p \rightarrow A_j$).

Definition 2. If we can choose m reactions from the possible mechanism in such a manner that they represent the following one-step communications: $A_j \rightarrow A_{j+1}$

$\rightarrow \cdots \rightarrow A_{j+m-1} \rightarrow A_{j+m}$ then between A_j and A_{j+m} we have an *m-step kinetic communication.*

The following remarks refer to the properties of the many-step kinetic communications:

Remark 1. The many-step kinetic communications reflect the sequence properties of the elementary processes included in the possible mechanism.

Remark 2. In contrast to mathematical communications where an entity has no communication with itself, an $A_p \rightarrow A_p$ kinetic communication might have a definite meaning. A good example is the following elementary process occurring in hydrocarbon oxidations:

$$HROO^{\cdot} + HROOH \rightarrow HROOH + {}^{\cdot}ROOH$$

where $HROO^{\cdot}$ and HROOH mean a peroxy radical and a hydroperoxide molecule, respectively.

In this case we have a

$$HROOH \rightarrow HROOH$$

kinetic communication.

Remark 3. A definite minimal step-number L can be attached to the elementary processes of any possible mechanism sufficient to reveal kinetic communications between any two species as an *at most* L-step communication. The number L denotes the characteristic step-number. Since the number of species is finite, it is obvious that the characteristic step-number cannot exceed the number of species, that is $L \leqslant N$.

Based on the above concepts the complete reaction network can be defined.

Definition 3. The entirety of the at most L-step communications determined by the possible mechanism of a complex process is the complete reaction network.

The one-step communications can be represented by a matrix called the kinetic communication matrix, denoted by K.

Definition 4. The kinetic communication matrix is a matrix where the p'th element of the j'th row, $k_{jp}=1$ if species A_j and A_p are in direct communication, otherwise $k_{jp}=0$ ($j=1, 2, \ldots N$; $p=1, 2, \ldots N$; independently).

Communication matrices used in pure mathematics have similar characteristics [8] with the sole difference that their main diagonal contains only 0, while – due to Remark 2 – the present matrix may also contain 1 in its main diagonal. Therefore, theorems proved for the former cannot be applied without further considerations.

The kinetic communication matrices defined above have important characteristics as to the network construction given by the following theorems.

Theorem 1. If K is a one-step communication matrix, then sgn K^2 represents two-step communications between the species (that is, if $k_{jp}=1$, there is and if $k_{jp}=0$, there is no two-step communication between species A_j and A_p).

The proof of Theorem 1 will be given in the appendix. Based on this evidence it is obvious that Theorem 1 is valid also for the S'th power of the one-step communication matrix:

Theorem 2. If K is a one-step communication matrix, then sign K^s ($s = 1, 2, 3 \ldots L$) represents s-step communications.

The Appendix verifies the following theorem:

Theorem 3. The matrix-sum $K' = \text{sgn} \sum_{s=1}^{L} K^s$ represents the entirety of the at most L-step kinetic communications.

Thus it can be seen that while K represents one-step communications, the matrix sum K' corresponds to the network defined in Definition 3. The latter, naturally, does not contain more information than K but has certain practical advantages in constructing the specific networks discussed later.

The communication matrices can be used in a more descriptive way, represented graphically where the numbers "1" in the matrix correspond to directed arrows. Unfortunately, in the case of "big" possible mechanisms the graphic interpretation is pointless, since it becomes too immense.

In order to illustrate the construction of the different matrices we take a very simplified "mechanism structure" of three elementary processes where the corresponding A, B and C letters can be easily replaced by others, more realistic chemically (in fact, this we intend to do in Part II), without an undue change in the mechanism structure.

$$AB_2 + C_2 \rightarrow AB + BC_2$$

$$AB_2 + BC_2 \rightarrow AB + B_2C_2$$

$$AB + C_2 \rightarrow ABC_2$$

Here A, B, and C stand for atoms or atomic groups. If these processes are considered as a possible mechanism of a certain overall reaction, the corresponding one-step kinetic communications are:

$$AB_2 \rightarrow AB \qquad C_2 \rightarrow BC_2$$

$$AB_2 \rightarrow BC_2 \qquad AB \rightarrow ABC_2$$

$$AB_2 \rightarrow B_2C_2 \qquad C_2 \rightarrow ABC_2$$

$$BC_2 \rightarrow B_2C_2$$

The respective two-step communications are:

$$AB_2 \rightarrow B_2C_2$$

$$C_2 \rightarrow B_2C_2$$

$$AB_2 \rightarrow ABC_2$$

while the one-step communication matrix is:

	AB_2	AB	C_2	BC_2	B_2C_2	ABC_2
AB_2	0	1	0	1	1	0
AB	0	0	0	0	0	1
C_2	0	0	0	1	0	1
BC_2	0	0	0	0	1	0
B_2C_2	0	0	0	0	0	0
ABC_2	0	0	0	0	0	0

$= K$

According to Definition 3 the communication matrix defines unambiguously the graph and thus further on the "graph" and "network" can be used as synonymic expressions. Consequently, numbers 1 in the matrix can be represented by directed arrows as shown in Fig. 1.

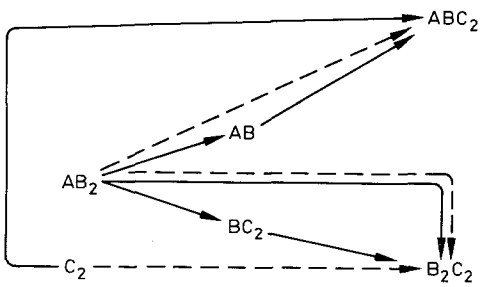

Fig. 1.

The dotted lines in Fig. 1 represent two-step communications according to the sgn K^2 matrix, though – as can be seen from Fig. 1 – these do not increase the amount of information since the precursor-product relations are already given by the one-step communications.

Practically, for real complex systems, the following further aspects should be taken into account:

a) Certain elements of the species-space of the possible mechanism are only hypothetic, their analytical determination is not always possible.
b) The experimental procedures allow observation of pathways of only certain atoms or atomic groups.

With respect to a) and b), Definition 3 is of excessively general character and thus the complete reaction network lacks significance. Its application to real systems requires more specific networks. These can be achieved by taking into account that

a) the number of the different types of atoms participating in a complex process is finite
b) the communication of two species differs essentially depending on the atoms or atomic groups realizing it (e.g. in the elementary process $RH_2 + O_2 \rightarrow HR^{\cdot}$

$+HO_2^{\cdot}$ – where RH_2 means a hydrocarbon molecule – a kinetic communication $RH_2 \rightarrow HO_2^{\cdot}$ exists with respect to the hydrogen atom, but there is no communication with respect to the carbon skeleton).

In order to distinguish the type of kinetic communication we introduce the "table of atomic indices";

1. Based on the possible mechanism we determine the number of atoms or atomic groups n participating in the kinetic communications and choose an arbitrary succession of them: $T_1, T_2 \ldots T_n$ where the T's stand for the chemical symbols of the atoms or atomic groups.

2. We construct a table the rows and columns of which contain the species in identical succession.

3. In each square belonging to the j'th row and p'th column we write n indices in the form: $t_1^{jp}, t_2^{jp}, \ldots t_n^{jp}$ as follows: $t_i^{jp} = 1 (i = 1, 2, 3, \ldots n)$, if the kinetic communication between A_j and A_p is realized by the atom or atomic group T_i, otherwise $t_i^{jp} = 0$ ($j, p = 1, 2, \ldots N$, independently). For a given square i goes from 1 to n.

It is obvious that if in matrix K, k_{jp} is 1, then at least one element of the indices t_1^{jp}, $t_2^{jp}, \ldots t_n^{jp}$ is 1 and if $k_{jp} = 0$, each element of the indices must be zero. The table of the atomic indices allows construction of the *reaction subnetworks.*

Definition 5. The entirety of one, two-, ... many-step kinetic communications realized by atoms or atomic groups T_i is the reaction subnetwork generated by the atom (or atomic group) T_i.

The number of subnetworks is obviously n. The subnetwork belonging to T_i represents the *pathway* of the atom (or atomic group) T_i *via* the different species specified by the possible mechanism.

The one-step subnetwork generated by T_i can be represented by a communication matrix $K(T_i)$ constructed in the following manner:

a) Using a succession of species chosen arbitrarily we produce a matrix the elements of which are the numbers t_i^{jp}.

b) We obtain $K(T_i)$ from this matrix abandoning those rows and columns which correspond to species not containing atoms (or atomic groups) T_i.

The reaction subnetwork according to Definition 5 is represented by the matrix sum $K'(T_i) = \operatorname{sgn} \sum_{s=1}^{L_i} K^s(T_i)$, where L_i means the characteristic step-number of the subnetwork generated by the atom (or atomic group) T_i, that is, Theorems 1, 2 and 3 are valid also for $K(T_i)$.

Some of the general properties of the kinetic communication matrices will be described in the Appendix.

Even more importance can be attributed to the specific subnetworks derived from the reaction subnetworks especially concerning the planning of the experimental work.

Such specific subnetwork could comprise the pathways of a given atom (or atomic

group) *via* the selected species. Selection of the species varies depending on the purpose of the investigations (e.g. stable components; identifiable compounds, etc....).

The specific subnetwork produced for the selected species is the *partial network*. Its communication matrix is the minor-matrix of $K(T_i)$ preserving the rows and columns which correspond to the selected species. A graphical representation of a simple partial network is given in Fig. 2.

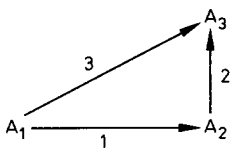

Fig. 2.

This means that a given atom is transferred from species A_1 to A_3 and to A_2 *and* from A_2 to A_3. (It should be noted that in the course of constructing the partial network from its subnetwork analysis of the one-, two-,... L-step communications yields information on the species that take part in the transfer, but have not been selected to follow.)

A basic insufficiency of the above formalization lies in the fact e.g. that it does not enable us to decide whether transfer of the atom T_i *via* route 3 can be realized without participation of the species A_2 or not? In order to answer this problem further analysis is required:

a) We analyse the one-step subnetwork generated by T_i (that is the $K(T_i)$ communication matrix which contains all the species), or
b) We omit the row and the column of species A_2 in the one-step $K(T_i)$ matrix and construct a matrix sum $K'(T_i)$ preserving the rows and columns corresponding to the species selected. If this "defective" partial subnetwork preserves arrow 3 (see Fig. 2), then T_i is transferred from A_1 to A_3 without the intermediate participation of A_2.

Production of the partial network for the species selected and its specification according to a) or b) leads to the *sequence network*.

1. The sequence network describes the possible pathways of the atom or atomic group T_i *via* selected species.
2. The sequence network (or the directed arrows in the graphic interpretation) represents definite sets of kinetic communications generated by the possible mechanism.
3. The sequence network informs us about those selected species which participate necessarily in the transfer of the atom or atomic group T_i between the species selected along the given pathway.

The sequence networks facilitate planning the experimental work and its evaluation in order to reduce the possible mechanism towards the probable mechanism.

Appendix

1. Evidence of Theorem 1

Evidence of Theorem 1, according to which if K is a one-step communication matrix, then sgn K^2 is the corresponding two-step communication matrix.

If $E=\operatorname{sgn} K^2$ and $K=(k_{ij})_{N,N}$ then the element of matrix E in the i'th row of the j'th column will be:

$$e_{ij}=\operatorname{sgn}(k_{i1}k_{1j}+k_{i2}k_{2j}+\cdots+k_{iN}k_{Nj}) \tag{1}$$

The members in the brackets are not zero unless any of the multiplicators is zero. If, however $k_{i1}=1$, species A_i has a one-step kinetic communication with A_1 and similarly if $k_{1j}=1$ species A_1 has a one-step kinetic communication with A_j. This means that $A_i \rightarrow A_1$ *and* $A_1 \rightarrow A_j$. Consequently if $k_{i1} \cdot k_{1j} \neq 0$, there exists an $A_i \rightarrow A_j$ two-step communication. Since the brackets of the expression (1) cannot contain a negative member, the signum of the sum is 1 only if at least one member differs from zero and it equals 0 only if each member is zero, that is, $e_{ij}=1$, if communication $A_i \rightarrow A_j$ is a two-step communication and $e_{ij}=0$ if there is no two-step communication between A_i and A_j at all.

2. Evidence of Theorem 3

The sum $K^1=\operatorname{sgn} \sum K^s$ gives the entirety of L-step kinetic communications. Let us discuss the j'th element of the i'th row in the $K=(k_{ij})_{N,N}$ matrix. It is evident that

$$k_{ij}=\operatorname{sgn}(k_{ij,1}+k_{ij,2}+\cdots+k_{ij,s}+\cdots+k_{ij,L})$$

where $k_{ij,1}$ is the j'th element of the i'th row in the one-step communication matrix raised to the power s (in other words it is the corresponding element of the s-step matrix). Since we have only positive members in the brackets $k_{ij}=0$ only in case all members are zero, that is, if among species A_i and A_j there is no communication whatsoever, neither one-, nor two-, or L-step communication. Furthermore $k_{ij}=1$ if at least one member differs from zero and this means that between A_i and A_j there is a certain-step communication. Thus the matrix-sum reflects, indeed, the entirety of the at most L-step communications.

3. Some Properties of the Kinetic Communication matrices

3.1.

The characteristic step-number described in Remark 3 can easily be determined for any kinetic communication matrix. Raised to higher powers and summed the one-step kinetic communication matrices may offer two possibilities:

a) Either the $(L+1)$'th power of the one-step communication matrix will be a O matrix;

b) or the sum of the matrices raised to higher powers remains invariant after the L'th power, that is:

$$\operatorname{sgn}\left[\sum_{s=1}^{L} K^s\right] = \operatorname{sgn}\left[\sum_{s=1}^{L+1} K^s\right]$$

3.2.

If the communication matrix-sum representing the complete reaction network has a row with exclusively zero-elements, this shows that the species corresponding to this row does not participate as a reactant in any of the elementary processes included in the possible mechanism.

The same for a column indicates that the corresponding species is not formed in any of the elementary processes included in the possible mechanism (this might be true for an initial substrate which is not reforming in any elementary process).

References

1. Sellers, P. H.: Arch. Ratl. Mech. Anal. **44**, 23 (1971)
2. Sinanoğlu, O.: J. Am. Chem. Soc. **97**, 2309 (1975); Acta Chim. Turcica **3**, 155 (1976); Sinanoğlu, O., Lee, L. S.: In preparation
3. Norrish, R. G. W.: Disc. Faraday Soc. **10**, 269 (1951)
4. Fish, A.: Proc. Roy. Soc. **A298**, 204 (1967)
5. Szabó, Z. G., Galiba, I., Gál, D.: Acta Chim. Hung. **74**, 239 (1972)
6. Nemes, I., Botár, L., Gál, D.: Z. Phys. Chem. **255**, 371 (1974)
7. Nemes, I., Botár, L., Gál, D.: Z. Phys. Chem. **255**, 378 (1974)
8. Kemény, J. G., Snell, J. L., Thompson, G. L.: Introduction to finite mathematics. New York: Prentice-Hall, Englewood Cliffs 1957

Received January 28, 1977/March 10, 1977