

## COMBINATORIAL GROUP THEORY FOR CHEMICAL REACTION SYSTEMS

Tetsuo Morikawa

School of Pharmaceutical Science, Toho University,  
Miyama 2 - 2, Funabashi 274, Japan

(Summary) A combinatorial method is developed for construction of a group, called "a reaction group" hereafter, from a given set of molecules which represents a chemical reaction system. The composition matrix of the molecules with respect to chemical elements determines a group-theoretical structure that is written as a direct sum of cyclic subgroups. All reactions in a reaction group, and consequently all reaction groups are classified in types by use of an equivalence relation having some connection with order of reaction. It is also shown that chemical equilibrium constants for coupled reactions have a cyclic group structure.

Introduction. A chemical process is usually expressed as an ordered pair of molecules (reactants and products) or as a reaction equation, and the inverse process as the inversely written pair. A complete list of such reaction equations for a given set of molecules can be made by use of

combinatorial methods, and a real reaction system can obviously be written as a subset of the list.

The word "equilibrium" has been used by chemists to represent a state of chemical process when the rate of a forward reaction is equal to the rate of the backward reaction or when the forward and the backward reactions are balancing. A chemically reacting system is said to be in a state of equilibrium when equilibria are reached in all of the reactions in the system. The combinatorial set thus not only corresponds to a chemical reaction system in a state of dynamic equilibrium, but also gives foundations of non-equilibrium reaction theory. It is of great importance to develop combinatorial/stoichiometric reaction theory for a given set of molecules.

A previous note<sup>1</sup> showed that chemically reacting systems have a group-theoretical structure as a special case. The purpose of the present note is to actually construct such a group from a given set of molecules, to classify such reaction groups in types from the group-theoretical view, and to shed light on the algebraic structure of thermodynamic variables that can be added together in the same manner as reaction equations.

Construction of Reaction Groups. Let  $\underline{M} = (M(1) M(2) \dots M(n_M))$  be the collection of  $n_M$  distinct molecular species (molecules, free radicals, ions, etc.), and let  $\underline{E} = (E(1) E(2) \dots E(n_E))$  be the collection of  $n_E$  distinct elemental species (chemical elements, atoms, chemical functional groups,

electrons, etc.). Assume that the k-th molecular species  $M(k)$  consists of  $E(j)$ 's, of which the number of each is  $c_{jk}$  (a positive integer or zero). Each row/column in the composition matrix  $\underline{c} = (c_{jk})$  contains at least one non-zero entry. We define a set of reaction equations in terms of ordered pairs of molecules as

$$R_E = \{ q \mid q = "M\underline{x} \rightarrow M\underline{y}", \underline{c}\underline{x} = \underline{c}\underline{y} \}$$

which represents a chemically reacting system.  $M\underline{x} = x_1 M(1) + \dots + x_{n_M} M(n_M)$  indicates the chemical reactants, and  $M\underline{y}$  the chemical products.  $\underline{x}$  is a column matrix of which the entries are called stoichiometric coefficients (positive integers or zero). The condition  $\underline{c}\underline{x} = \underline{c}\underline{y}$  on  $q$  is the  $n_E$  simultaneous Diophantine equations with unknowns  $\underline{x}$  and  $\underline{y}$ , and states the law of conservation of matter in chemical reaction equations. It follows from this condition that if " $M\underline{x} \rightarrow M\underline{y}$ " =  $q \in R_E$ , then " $M\underline{y} \rightarrow M\underline{x}$ " =  $-q \in R_E$  (the inverse reaction of  $q$ ). The set  $R_E$  is an abelian semigroup with respect to the additive operation because  $q + q' = "M\underline{x} \rightarrow M\underline{y}" + "M\underline{x}' \rightarrow M\underline{y}'" = "M(\underline{x} + \underline{x}') \rightarrow M(\underline{y} + \underline{y}')"$  =  $q' + q \in R_E$ , and  $q + (q' + q'') = (q + q') + q''$  for  $q, q', q'' \in R_E$ .

By a manner similar to that described in Reference 1 all  $q$  in  $R_E$  are classified; then the quotient set of  $R_E$  becomes a group, and further it is decomposed as a direct sum:

$$R_E/\cong = \{ r, r', \dots \} = [r_1] \oplus \dots \oplus [r_{n_G}] = R_G$$

The sign  $\equiv$  denotes an equivalence relation which means that  $q \equiv q'$  when  $q = \underline{Mx} \rightarrow \underline{My}$  and  $q' = \underline{M(x+z)} \rightarrow \underline{M(y+z)}$ . The notation  $[r_j] = \{\dots, -2r_j, -r_j, r_0, r_j, 2r_j, \dots\}$  indicates a cyclic subgroup with the generating reaction  $r_j$ .  $r_0$  is the unit reaction:  $r_0 = \{q \mid q = \underline{Mx} \rightarrow \underline{Mx}, q \in R_E\}$ . The inverse reaction of  $r_j$  is expressed as  $-r_j$ :  $r_j + (-r_j) = r_0$ . The base transformation of  $R_G$  is carried out by use of the relations  $[r_j] = [-r_j]$  and  $[r_j] \oplus [r_k] = [r_j] \oplus [r_j + r_k]$  ( $j \neq k$ ). Note: The reaction group of a real system is a subgroup in the corresponding (combinatorial/stoichiometric)  $R_G$ .

Example 1.  $\underline{M} = (H \ H_2 \ H_3)$  and  $\underline{E} = (H)$ ; then

$$\underline{c} = \begin{pmatrix} H & H_2 & H_3 \\ 1 & 2 & 3 \end{pmatrix} H.$$

$$R_E = \{H \rightarrow H, 2H \rightarrow H_2, H + H_2 \rightarrow H_3, 3H \rightarrow H_3, \dots \\ \dots(\text{and their inverse equations})\}.$$

$$R_G = R_E / \equiv = [r_1] \oplus [r_2] = [r_1'] \oplus [r_2] = [r_1] \oplus [r_2'] = \dots$$

$$\begin{aligned} r_0 &= \{H \rightarrow H, H_2 \rightarrow H_2, H + H_2 \rightarrow H + H_2, \dots\}, \\ r_1 &= \{2H \rightarrow H_2, 3H \rightarrow H + H_2, 2H + H_3 \rightarrow H_2 + H_3, \dots\}, \\ -r_1 &= \{H_2 \rightarrow 2H, H + H_2 \rightarrow 3H, H_2 + H_3 \rightarrow 2H + H_3, \dots\}, \\ r_2 &= \{H + H_2 \rightarrow H_3, 2H + H_2 \rightarrow H + H_3, \dots\}, \\ -r_2 &= \{H_3 \rightarrow H + H_2, H + H_3 \rightarrow 2H + H_2, \dots\}, \\ r_1' &= -r_1 + r_2 = \{2H_2 \rightarrow H + H_3, \dots\}, \\ r_2' &= r_1 + r_2 = \{3H \rightarrow H_3, \dots\}. \end{aligned}$$

Example 2.  $\underline{M} = (\text{H}^+ \text{ e}^- \text{ H}_2 \text{ H}_2^+)$  and  $\underline{E} = (\text{H}^+ \text{ e}^-)$

(the hydrogen-electrode reaction); then

$$\underline{c} = \begin{matrix} & \text{H}^+ & \text{e}^- & \text{H}_2 & \text{H}_2^+ & \\ \begin{matrix} \text{C} \\ \text{H} \end{matrix} & \begin{pmatrix} 1 & 0 & 2 & 2 \\ 0 & 1 & 2 & 1 \end{pmatrix} & \begin{matrix} \text{H}^+ \\ \text{e}^- \end{matrix} \end{matrix}$$

$$R_G = [2\text{H}^+ + \text{e}^- \rightarrow \text{H}_2^+] \otimes [\text{H}_2^+ + \text{e}^- \rightarrow \text{H}_2].$$

(The representatives were used.)

(Contraction of composition matrices) The rank of  $\underline{c}$  is different from the rank of  $R_G$ . When one of the columns of  $\underline{c}$  contains common divisors, and/or when all of the rows of  $\underline{c}$  have the same form, the  $\underline{c}$  is contractible. The following  $\underline{c}$ , for example, gives the same reaction group as for the contracted matrices  $\underline{c}'$ ,  $\underline{c}''$ , and  $\underline{c}'''$ :

$$\begin{matrix} & \text{C}_2\text{H}_2 & \text{C}_4\text{H}_4 & \text{C}_6\text{H}_6 & & \text{C}_2\text{H}_2 & \text{C}_4\text{H}_4 & \text{C}_6\text{H}_6 \\ \begin{matrix} \text{C} \\ \text{H} \end{matrix} & \begin{pmatrix} 2 & 4 & 6 \\ 2 & 4 & 6 \end{pmatrix} & = \underline{c}, & \text{C}_2 \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} & = \underline{c}', \end{matrix}$$

$$\begin{matrix} & \text{C}_2\text{H}_2 & \text{C}_4\text{H}_4 & \text{C}_6\text{H}_6 & & \text{C}_2\text{H}_2 & \text{C}_4\text{H}_4 & \text{C}_6\text{H}_6 \\ \text{CH} & (2 & 4 & 6) & = \underline{c}'', & \text{C}_2\text{H}_2 (1 & 2 & 3) & = \underline{c}'''. \end{matrix}$$

If a given  $\underline{c}$  can be transformed into a block-diagonal form after the exchange of columns/rows, it is enough to construct a group for each block because there is no reaction between such blocks except  $r_0$ . We hereafter consider  $\underline{c}$  non-block-diagonal. Even if  $\underline{c}$  is non-block-diagonal, there may be molecules that cannot react with others in the block. Such "lonely" molecules are excluded from  $\underline{c}$  hereafter.

The Rank of Reaction Groups. The rank of  $R_G$  ( $= n_G =$  the cardinal number of a base) is constant, although many bases can be set for  $R_G$  in general. Assuming that  $R_G$  is written in terms of two bases as  $R_G = [r_1] \oplus \dots \oplus [r_n] = [r'_1] \oplus \dots \oplus [r'_{n'}]$ , we prove that  $n = n'$ . Define the subgroup  $R_G^{(m)} = \{ m r \mid r \in R_G \}$ , ( $m$ , a natural number). We have

$$R_G^{(m)} = [mr_1] \oplus \dots \oplus [mr_n] = [mr'_1] \oplus \dots \oplus [mr'_{n'}]$$

The factor groups are then isomorphic:

$$\begin{aligned} R_G/R_G^{(m)} &\cong [r_1]/[mr_1] \oplus \dots \oplus [r_n]/[mr_n] \\ &\cong [r'_1]/[mr'_1] \oplus \dots \oplus [r'_{n'}]/[mr'_{n'}] \end{aligned}$$

Here  $[r_j]/[mr_j]$  and  $[r'_k]/[mr'_k]$  are cyclic groups of order  $m$ , and so  $m^n = m^{n'}$ . Therefore,  $n = n'$ .

Setting the rank of a reaction group for a subset of  $\underline{M}$  with  $k$  distinct molecules as  $n(k)$ , and adding a new kind of molecule to the composition matrix of the group, we obtain  $n(k+1) \leq n(k) + 1$ . This inequality is valid for the reason that the new kind of molecule allows the rank either to increase if a reaction including it is independent, or to hold if dependent. The summation from  $k=1$  to  $k=n_M - 1$  yields  $n_G \leq n_M - 1$ . (Note:  $n(1) = 0$ .)

Let  $k_0$  be the maximum number of  $k$ 's for  $n(k) = 0$ . If a new kind of molecule is added to the group of  $k_0$ , then there is an increase in the rank. In a similar manner the addition of  $n_M - k_0$  kinds of molecule increases the rank; therefore,  $n_G = n_M - k_0$  ( $k_0 \geq 1$ ). If  $k_1$  denotes the maximum number

of  $k$ 's for  $n(k) = 1$ , then  $k_1 = k_0 + 1$  because, if a kind of molecule is removed from one of the groups of  $k_1$ , the resulting group is that with  $k_0$ . For Example 1,  $k_0 = 1$ ,  $k_1 = 2$ , and  $n_G = 3 - 1 = 2$ ; for Example 2,  $k_0 = 2$ ,  $k_1 = 3$ , and  $n_G = 4 - 2 = 2$ .

Classification of Reactions and Reaction Groups. Let  $d(r)$  be the difference between the number of reactants and that of products for  $r$ ; namely,  $d(r) = \sum y_k - \sum x_k$ , " $\sum x_k M(k) \rightarrow \sum y_k M(k)$ " =  $q \in r$ . The integer  $\sum x_k$  for reactants is often termed by chemists stoichiometric/kinetic order of reaction.

(Elementary reaction) An elementary reaction equation in chemical theory is not only a statement in symbols of a stoichiometric chemical change, but also a reaction that takes place as a quantum-chemical molecular/atomic process. It is known that in an elementary reaction the number of reactants ( $= \sum x_k$ ) is at most 3, and the minimum number of products ( $= \sum y_k$ ) is obviously 1; therefore,  $d(r) \geq -2$  for elementary reactions. By use of this inequality elementary reactions including chemically unusual ones can be classified in types as the following. (i)  $d(r) = -2 = 1 - 3$ . The reaction of third order yields one molecule. (ii)  $d(r) = -1 = 2 - 3 = 1 - 2$ . The reaction of third order gives two molecules, and the reaction of second order one molecule. The latter is an association reaction (e.g.,  $A + B \rightarrow AB$ ). (iii)  $d(r) = 0 = 1 - 1 = 2 - 2 = 3 - 3$ . Only three kinds of reaction are possible. The first is an intramolecular rearrangement reaction (e.g.,  $A \rightarrow B$ ), and the second is a

substitution reaction (e.g.,  $A + BC \rightarrow AB + C$ ). (iv)  $d(r) = 1 = 2 - 1 = 3 - 2 = 4 - 3$ . The first is a decomposition reaction (e.g.,  $AB \rightarrow A + B$ ). (v)  $d(r) = 2 = 3 - 1$ , and so on.

The following remarks are obvious consequences of the definition of  $d(r)$ :  $d(-r_j) = -d(r_j)$ ,  $d(mr_j) = m d(r_j)$ ,  $d(r_0) = 0$ ,  $d(r_j + r_k) = d(r_j) + d(r_k)$ , and a result,  $d(r_{OV}) = \sum m_j d(r_j)$  if  $r_{OV} = \sum m_j r_j$ . Two reactions  $r$  and  $r'$  are regarded as equivalent when  $d(r) = d(r')$ ; then we write  $r \stackrel{d}{\equiv} r'$ . Using this equivalence relation we can classify all  $r$  in  $R_G$  as equivalence classes:

$$R_G / \stackrel{d}{\equiv} = \{s, s', \dots\} = R_D$$

This set becomes a group with respect to addition because the equivalence relation is compatible with the additive operation on  $r$ 's; we write this homomorphism as  $R_G \sim R_D$ .

Two values  $d(r_j)$  and  $d(r_k)$  clearly become 0 and  $m$  ( $m$ , integer), after repetitions of the appropriate base-transformation of  $r_j$  and  $r_k$ , and the use of the above remarks. Applying this procedure to the sequence of  $d(r_j)$ 's for a base  $r_1, r_2, \dots, r_{n_G}$  of  $R_G$ , we obtain the form  $0, 0, \dots, 0, m$ . Hence  $R_D = [s_m] \cong [m]$ , where  $s_m = \{r \mid d(r) = m, r \in R_G\}$  is the generating element for  $R_D$ . (Note:  $[m]$  is a cyclic subgroup in  $[1]$ .) Therefore,  $R_G \sim [m]$ ; this homomorphism shows a method for classification of reaction groups of rank  $n_G$  according to the generating element  $m$  of  $d(r)$ 's. For Example 1 ( $n_G = 2$ ),  $R_G \sim [1]$ ; for  $\underline{M} = \begin{pmatrix} A & -B \\ \swarrow & \searrow \\ C & \end{pmatrix} \quad \begin{pmatrix} A & -B \\ \swarrow & \searrow \\ C & \end{pmatrix} \quad \begin{pmatrix} A & -B \\ \swarrow & \searrow \\ C & \end{pmatrix}$



(intramolecular rearrangement,  $n_G = 2$ ),  $R_G \sim [0] = \{0\}$ .

Note that  $s_0$  is a normal subgroup in  $R_G$  because  $d(r + r') = d(r) + d(r') = 0$ , and  $d(-r) = -d(r) = 0$  for  $r, r' \in s_0$ .

The classification by means of  $\equiv$  is finer/stronger than that by  $\equiv_d$  because if  $q \equiv_d q'$ , then necessarily  $q \equiv q'$ .

Relations similar to the homomorphism between  $R_G$  and  $R_D$  can also be obtained for  $R_G$  and one of the thermodynamic variables that can be added together in a similar manner as reaction equations. As was discussed in Reference 1, for example,  $R_G \sim \{\ln K, \ln K', \dots\} = R_K$ , where  $K$ 's are chemical equilibrium constants. This homomorphism suggests that  $R_K$  equals the cyclic group with the generating element  $\ln K_g$ ;  $R_K = [\ln K_g]$ ; in other words,  $(\ln K)/m = (\ln K')/m' = \dots = \ln K_g$ .

Note. There are many equivalence relations other than  $\equiv_d$ . For example, the value  $v(r) =$  (the number of kinds of molecule in  $\underline{M}_y$ ) - (the number of kinds of molecule in  $\underline{M}_x$ ) for  $q$  defines the third equivalence relation by  $r \equiv_v r'$  when  $v(r) = v(r')$ . This relation is not compatible with the addition of  $r$ 's because  $v(r + r') \neq v(r) + v(r')$ , so that  $R_G/\equiv_v$  no longer becomes a group with respect to addition.

### References

1. T. Morikawa, Semigroup-Theoretical Structure of Chemical Reaction Systems, submitted for publication in MATCH.
2. On recent theory of groups, refer to D. J. S. Robinson, "A Course in the Theory of Groups", Springer-Verlag, New York(1982).