I. COMBINATORIAL THEORY OF NONLINEAR GRAPHS, APPLIED TO THE VIRIAL EQUATION OF STATE

II. CHEMICAL THERMODYNAMICS OF OPEN SYSTEMS

Thesis by
Alvin Henry Larsen

In Partial Fulfillment of the Requirements
For the Degree of
Doctor of Philosophy

California Institute of Technology
Pasadena, California
1969

(Submitted October 24, 1968)

ACKNOWLEDGMENTS

First of all, I wish to express my sincere appreciation to Professor C. J. Pings, my research advisor, for his invaluable encouragement and assistance with this work. Thanks are also due to many other faculty members at the California Institute of Technology, especially those in the Department of Chemical Engineering, where the highly stimulating study and research environment motivated me, in large part, to accomplish this work. Helpful conversations with Professor G. R. Cokelet and with S. C. Smelser are acknowledged.

I am greatly indebted to the several sources of financial support which I received during the course of my graduate study: to the National Science Foundation, for a Graduate Traineeship for two years; to the California Institute of Technology, for teaching and research assistantships; and to the Air Force Office of Scientific Research, for a research assistantship and for funds to defray research expenses.

Finally, I am grateful to my wife, Sharon, for her patience and understanding during the time of research for and preparation of this thesis.

ABSTRACT

PART I

By extension of the concept of a linear graph, as a topological configuration of vertexes with lines (2-bonds) connecting certain pairs of them, a *nonlinear* graph is defined to include also associations of certain triplets of the vertexes, or 3-bonds, representable by topological "areas" of triangles; quadruplets, or 4-bonds, by "volumes" of tetrahedra; etc. The configurational part of the partition function, including nonadditivity effects or three-body and higher interactions in the potential energy of a configuration of molecules, may be expanded as a sum of integrals over products of cluster functions, corresponding to a sum of nonlinear graphs. Certain special types of graphs, called trees and stars, figure prominently in the combinatorial analysis. The nth virial coefficient corresponds to the sum of all stars on n vertexes. Since topologically equivalent graphs correspond to integrals which yield the same result upon integration, the nth virial coefficient corresponds also to the sum of all topologically distinct stars on n vertexes, each multiplied by the appropriate combinatorial coefficient.

Development of the combinatorial theory for nonlinear graphs, trees, and stars proceeds similarly to that for linear graphs. An explicit formula for counting nonlinear graphs on distinguishable vertexes is obtained, and generating functions relate the numbers of graphs to the numbers of corresponding nonlinear trees and stars. A novel term, the *cycle function*, is defined; the triplet cycle function, for 3-bonds, is derived; and cycle functions are used in generalizing Pólya's theorem to apply to nonlinear graphs with more than one type of bond present. A theorem is thus obtained which solves the problem of counting nonlinear graphs on indistinguishable vertexes, and a relationship between generating functions permits the number of corresponding trees to be calculated. All of these techniques are extended to apply to multicomponent systems and to rooted graphs and trees. Then the problem of counting stars on indistinguishable or multicomponent vertexes is solved by a systematic procedure. The numbers of such stars are also closely approximated by a simple formula. Including 2- and 3-bonds only, the number of distinct stars on $n = 3$ indistinguishable vertexes is 5; for $n = 4$, there are 72 topologically distinct stars; and for $n = 5$, 10,346.

Because of the rapidly increasing number and complexity of the calculations, a practical limit for actual calculations involving any nonadditivity effects for a pure substance is the fourth virial coefficient. For multicomponent systems, the numbers of topologically distinct stars are always greater than for systems of a single component; hence even more compelling reasons would then be necessary to justify calculation of the fourth virial coefficient including three-body nonadditivity effects.

PART II

General thermodynamic expressions for partial derivatives of extent of reaction with respect to external composition perturbations by a single species, for multicomponent, multiple-reaction systems constrained to paths of chemical equilibrium under various conditions, are obtained as the solution of a set of simultaneous, linear algebraic equations. The corresponding heat and temperature effects follow immediately. Derivatives of the extent of reaction are evaluated for ideal solutions. For multicomponent composition perturbations, the derivatives result from a linear combination of those for perturbations by a single species, each weighted by the net mole fraction of the given species in the streams crossing the boundary. Possible applications of the thermodynamic expressions include behavior of open systems under externally introduced composition perturbations, error analysis, and optimization of yield.

TABLE OF CONTENTS

TABLE OF CONTENTS—Continued

TABLE OF CONTENTS—Continued

LIST OF TABLES

LIST OF TABLES—Continued

LIST OF ILLUSTRATIONS

Figure Page

PART I

Proposition I

Proposition II

PART I

COMBINATORIAL THEORY OF NONLINEAR GRAPHS, APPLIED TO THE VIRIAL EQUATION OF STATE*

---

*Portions of Part I of this thesis have been published as a paper by A. H. Larsen and C. J. Pings, "Counting Graphs of Interest in Statistical Mechanics, Including Nonadditivity Effects," *J. Chem. Phys. 49*, 72 (1968).

A. INTRODUCTION

1. Intermolecular Potentials

The virial development of the equation of state for imperfect gases has historically involved many-body integrations over pairwise cluster functions. Extending the technique introduced by Ursell,(48) Mayer(30-32,34) expressed the coefficients in the density expansion of the compressibility factor in terms of certain cluster integrals, each of which may be represented by a sum of linear graphs,(40,47) in which the vertexes correspond to particles and the lines to interactions of pairs of the particles. Clarifications and generalizations of the technique have been given by Born,(8,9) Fuchs,(15) Husimi,(24) and Uhlenbeck and Ford.(47) The Mayer procedure, for simplicity, assumes pairwise additivity of intermolecular potentials; that is, that the potential energy of a configuration of molecules is equal to the sum of the potentials of pairs of molecules in the configuration. The cluster integrals are expressed successively in terms of certain "irreducible integrals," which are related very simply to the virial coefficients, and which correspond to the sum of all linear graphs of a certain type.

The Mayer procedure is not the only means developed for performing cluster expansions with the aid of linear graphs. For example, Stell(46) has developed cluster expansions by an alternative method of employing functional differentiation with linear graph theory. This work, however, involves the Mayer development.

The assumption of pairwise additivity of intermolecular potentials usually permits experimental data to be fitted fairly satisfactorily. Recently, however, evidence has accumulated to show that the properties of dense gases probably cannot be adequately expressed by pairwise additivity alone. Significant contributions to the intermolecular potential may occur from three-body interactions. Models to investigate the magnitude of a correction for nonadditivity have been given by a large number of investigators, including Axilrod and Teller,(2-4) Muto,(35) Bade and Kirkwood,(5) Kihara,(27) Kestner and Sinanoğlu,(26) Hutchinson,(25) Baer,(6) Williams *et al.*,(50) and Baxter.(7) Several investigators, including Kihara and coworkers,(33) Graben and Present,(16,17) Sherwood *et al.*,(43-45) and Chan and Dalgarno,(10) have calculated nonadditive corrections to the third virial coefficient for some simple potentials and compared the results with experimental data for a variety of substances.

2. Nonlinear Graph Theory

Friedman,(13,14) in application to ionic solution theory, has extended Mayer's graphical representation of two-body interactions to include nonadditivity effects as well, and Allnatt(1) has recently applied a similar representation to point-defect interactions in solids. Just as lines connecting pairs of vertexes of a graph correspond to two-body interactions, three-body interactions may be represented by topological "areas" of triangles formed by triplets of vertexes, four-body interactions by volumes of tetrahedra, etc. In contrast to the linear graphs representing cluster integrals involving only pairwise additive interactions, cluster integrals involving nonadditivity effects correspond to what may be termed *nonlinear* graphs.

Nonlinear as well as linear graphs may be classified according to topological type. Graphs of the same type, indistinguishable topologically, correspond to integrals which yield the same result upon integration. Hence the virial coefficients are obtained by evaluating the integrals corresponding to different types of graphs and multiplying each by the appropriate coefficient, which is the number of topologically indistinguishable graphs of the given type. The object of this work is to develop systematic methods for obtaining the numbers of nonlinear graphs corresponding to the virial coefficients.

In Sec. B, a number of graph theoretical terms are defined, including certain special types of graphs, called trees and stars. The relationship between nonlinear graphs and the integrals of cluster functions corresponding to the virial coefficients is indicated. Due to the rapidly increasing numbers of nonlinear graphs corresponding to the successive virial coefficients, as well as their increasing complexity, the associated combinatorial problems are formidable.

For the combinatorial analysis, techniques for counting graphs, trees, and stars on distinguishable vertexes are developed in Sec. C. Then in Sec. D the more difficult problem of counting graphs and trees on indistinguishable vertexes is considered. The general case, in which the vertexes are distinguishable by "species," is treated in Sec. E. This generalization includes the special cases for both distinguishable and indistinguishable vertexes, and applies to multicomponent systems. Sufficiently powerful methods are then available, in Sec. F, to count stars on indistinguishable vertexes, the numbers of which apply directly to the calculation of the virial coefficients.

B. NONLINEAR GRAPHS AND CLUSTER FUNCTIONS

1. Graph Theory—Definitions

For correspondence with particles and interactions between them, define a *nonlinear graph* as a set of *vertexes* and a set of associations of certain pairs, triplets, quadruplets, etc. of the vertexes. Denote the association of j vertexes as a *j-bond* and consider bonds of different j as bonds of different *types*, which occur independently of each other in a nonlinear graph. The vertexes may or may not be distinguishable (labeled). In general each j vertexes of a graph on n vertexes ($j = 2, 3, \ldots, n$) either may or may not be associated by a j-bond. Two vertexes which are associated by the same bond are said to be *adjacent*.(19) Two vertexes may therefore be adjacent by more than one type of bond, or by two or more 3-bonds, 4-bonds, etc., but by no more than one 2-bond. No bond connects a vertex to itself. A j-bond may be represented geometrically as a regular polyhedron of j vertexes in $j-1$ dimensions: a 2-bond by a line, a 3-bond by the area of a triangle, a 4-bond by the volume of a tetrahedron, etc. A graph on n vertexes is said to contain j-bonds (read j-set bonds) if j is a nonempty subset of the integers 2, 3, ..., n corresponding to the types of bonds in the graph. For a graph containing 2- and 3-bonds only, j = 2, 3.

A graph is disconnected (or disjoint) if it is possible to separate the vertexes of the graph into two or more subsets such that no two vertexes in different subsets are adjacent. Otherwise the graph is connected and is termed a *tree*. A vertex of a tree is termed a *cut vertex* (or articulation point(40,47)) if it is possible to separate the *remaining* vertexes into two or more subsets such that no two vertexes in different subsets are adjacent. A *star* denotes a tree with no cut vertexes. Considering two or more sets of vertexes nonadjacent if no two vertexes in different sets are adjacent, one may write the definitions of "tree" and "star" in this parallel form:

$$\left.\begin{array}{l}\text{Tree:}\\ \text{Star:}\end{array}\right\}\ \text{Can not separate the set of}$$

$$\left\{\begin{array}{l}\text{vertexes of a graph}\\ \text{all vertexes but one of a tree}\end{array}\right\}$$

into two or more nonadjacent subsets.

A graph consisting of a single vertex will be considered a tree but not a star. Any graph therefore consists of one or more disjoint trees, and any tree consists of one or more stars joined together at cut vertexes.

This generalization of linear trees and stars has been selected for compactness of expression and applicability to nonlinear graphs. The given definitions are simpler than the corresponding definitions of at-least-singly-connected (ALSC) and at-least-doubly-connected (ALDC) graphs by Friedman(14) and overcome the difficulty of applying the usual definition of "articulation point"(40) to nonlinear graphs.

Define the *symmetry number* of a graph as the order of the group of automorphisms (see Appendix B) of the graph.(47) If s is the symmetry number of a graph on n vertexes, then there are $n!/s$ permutations of the vertexes which yield topologically equivalent graphs. Define the *complement* of a graph containing j-bonds on n vertexes as the graph on the same vertexes, but containing all j-bonds *not* in the original graph. It follows that the complement of a graph has the same symmetry number as the given graph.

Define a *polygon* as a linear star of n 2-bonds (forming a single cycle) on n vertexes, for $n \geqq 3$, or a single 2-bond joining two vertexes, for $n = 2$. Polygons therefore include lines, triangles, quadrilaterals, pentagons, etc. Define a *cactus*(20) as a tree whose stars are polygons. If only lines are permitted, the result is a *Cayley tree*,(47) or a "tree" as the term has been ordinarily used.(10,28,41)

Define a *complete star* of j-bonds on n vertexes as a star containing all possible j-bonds. A star containing more than one type of bond is complete if and only if it is complete with respect to every type of bond present in the star. The vertexes of a complete star are therefore all symmetric with respect to each other, and all pairs of the vertexes are adjacent. A tree consisting of complete stars will be termed a *complete tree*. Different stars in a complete tree may themselves be complete with respect to different types of bonds.

As an example of nonlinear graphs, all different graphs on three indistinguishable vertexes are shown in Fig. 1. The coefficients given are the numbers of different graphs of each type which would be obtained if the vertexes were distinguishable. The first two graphs shown on the top line are disconnected, the third is a tree but not a star, and all the others are stars.

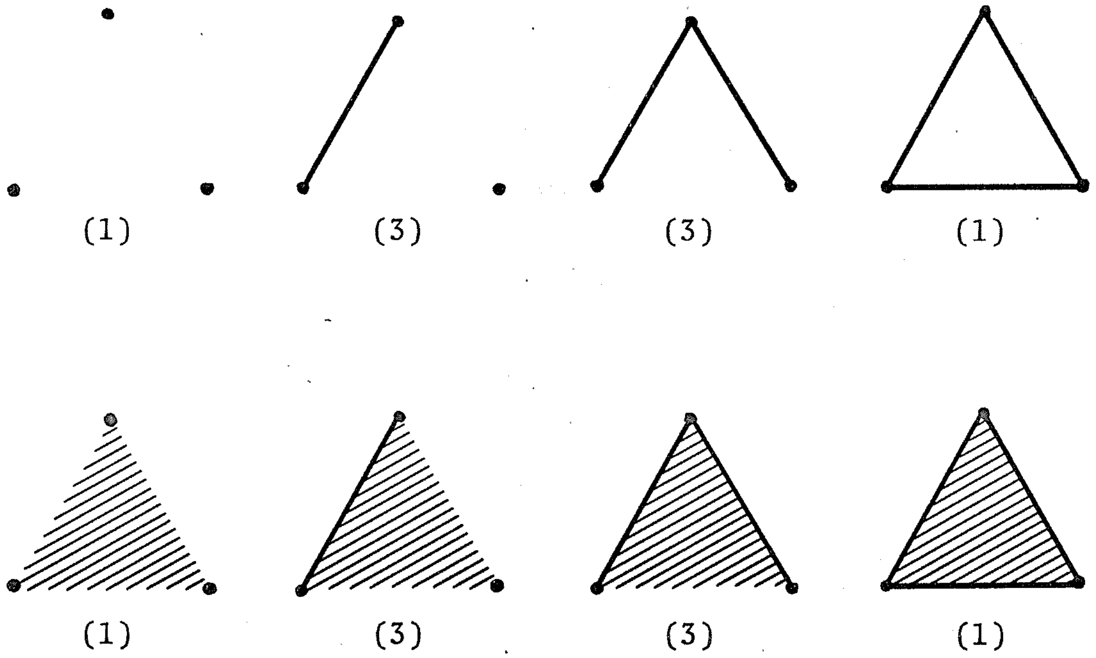
2. Cluster Functions

The potential energy of interaction of a configuration of three molecules is the sum of three pairwise interactions plus a small but nonnegligible correction term for the three-body interaction:(14)

$$U(\mathrm{r}_1,\mathrm{r}_2,\mathrm{r}_3) = u(\mathrm{r}_1,\mathrm{r}_2) + u(\mathrm{r}_1,\mathrm{r}_3) + u(\mathrm{r}_2,\mathrm{r}_3) + u(\mathrm{r}_1,\mathrm{r}_2,\mathrm{r}_3),$$

where r_i is the set of coordinates of molecule i. This defines the three-body interaction term $u(\mathrm{r}_1,\mathrm{r}_2,\mathrm{r}_3)$. Higher-order nonadditivity effects are defined similarly,

FIGURE 1.—Nonlinear graphs on three indistinguishable vertexes. The lines represent 2-bonds and the shaded areas, 3-bonds. The integers in parentheses give the number of different graphs which would be obtained from the given graph if the vertexes were distinguishable.



by considering them as successive corrections to the sum of all lower-order interaction terms. Denoting $U(\mathrm{r}_1,\mathrm{r}_2,\ldots,\mathrm{r}_N)$ by U_N and $u(\mathrm{r}_i,\mathrm{r}_j,\ldots)$ by $u_{ij\ldots}$, one has in general

$$U_N = \sum_{\text{pairs}} u_{ij} + \sum_{\text{triplets}} u_{ijk} + \sum_{\text{fours}} u_{ijk\ell} + \ldots + u_{12\ldots N}. \tag{1}$$

The correspondence of nonlinear graphs with cluster integrals is based on the definition of the cluster function(14)

$$f(\mathrm{r}_i,\mathrm{r}_j,\ldots) = f_{\mathrm{m}} = \exp\left(-\frac{u_{\mathrm{m}}}{kT}\right) - 1, \tag{2}$$

where $\mathrm{m} = i, j, \ldots$ is the *set* of interacting molecules, k is the Boltzmann constant, and T is the absolute temperature. If m is the number of molecules in the set $\mathbf{m}$, the cluster function f_{m} corresponds to an m-bond of a nonlinear graph. For pairwise interactions only, $m = 2$ and Eq. (2) defines the Mayer cluster function.(31)

3. Virial Coefficients

The configuration integral for N identical molecules,(22)

$$Z(N,V,T) = \int_V \cdots \int_V \exp\left(-\frac{U_N}{kT}\right) d\mathrm{r}_1 \ldots d\mathrm{r}_N, \tag{3}$$

may be expressed as the sum of integrals over products of cluster functions corresponding to all graphs on N distinguishable vertexes. The virial coefficients may then be obtained successively in terms of a sum of cluster integrals which may be represented by trees on distinguishable vertexes.(14) Furthermore, the virial equation in terms of the molar volume $\overline{V}$ may be expressed as

$$\frac{P\overline{V}}{RT} = 1 + \sum_{n=2}^{\infty} A_n(T)\overline{V}^{1-n}, \tag{4}$$

where the nth virial coefficient, $A_n(T)$ $(n \geqq 2)$, for systems of a single species, is explicitly(14)

$$A_n(T) = -(n-1)N_o^{n-1}\beta_n. \tag{5}$$

Here N_o is Avogadro's number, β_n is given by

$$\beta_n = \frac{1}{Vn!}\int_V \cdots \int_V S_n \mathrm{dr}_1 \ldots \mathrm{dr}_n, \tag{6}$$

and S_n is the sum of all cluster terms (products of cluster functions) which correspond to stars on n distinguishable vertexes. Neglecting surface effects, the integrations of the cluster terms are performed over all space, strictly in the limit as $V \to \infty$. The labeling of vertexes of the corresponding stars is arbitrary. For a pure substance, only topologically distinct stars on indistinguishable vertexes correspond to different values of the integrals. Hence the number of different graphs on distinguishable vertexes, and the number of these which could be formed from the same graph if the vertexes were indistinguishable, are desired. For multicomponent systems, the vertexes are distinguishable by species, and the numbers of topologically distinguishable graphs lie between those for indistinguishable vertexes and those for distinguishable vertexes, and depend on the distribution of the species present among the vertexes of the graphs under consideration. The coefficients as well as the numbers of inequivalent integrals increase rapidly with n; therefore the solution of the combinatorial problems associated with nonlinear graphs, trees, and stars is an important part of the calculation of virial coefficients including nonadditivity effects.

4. Multicomponent Systems

While the foregoing development has been given in a form to apply only to pure substances, it may be easily extended to the case of multicomponent systems. Instead of Eq. (5), the nth virial coefficient is then$^{(14)}$

$$A_n(T) = -(n-1)N_\circ^{n-1}\sum_{\mathrm{n}} \mathrm{x}^{\mathrm{n}}\beta_{\mathrm{n}}, \tag{7}$$

where

$$\mathrm{n} = n_1, n_2, \ldots, n_g, \tag{8}$$

$$\mathrm{x} = x_1,\ x_2,\ \ldots,\ x_s, \tag{9}$$

s is the number of species in the system, x_i is the mole fraction of species i, and

$$\mathrm{x}^{\mathrm{n}} = \prod_{i=1}^{s} x_i^{n_i}. \tag{10}$$

The sum in Eq. (7) is over all $\binom{n+s-1}{n}$ ordered partitions$^{(41)}$ of n into s elements, including zero elements, or equivalently, over all n, and

$$\beta_{\mathrm{n}} = \frac{1}{V\mathrm{n}!}\int_V\cdots\int_V S_{\mathrm{n}}\,\mathrm{d}\mathbf{r}_1\ldots\mathrm{d}\mathbf{r}_n, \tag{11}$$

$$\mathrm{n}! = \prod_{i=1}^{s} n_i!, \tag{12}$$

$$n = \sum_{i=1}^{s} n_i.$$

Then S_{n} corresponds to the sum of all stars on n vertexes distinguishable by species only; that is, all stars on n vertexes, of which n_i are of species i (i = 1, 2, ..., s).

C. COUNTING OF NONLINEAR GRAPHS ON DISTINGUISHABLE VERTEXES

Uhlenbeck and Ford(47) review systematic techniques for counting the number of linear graphs, trees, and stars on n distinguishable vertexes. These counting techniques, involving the use of generating functions, are extended here to nonlinear graphs.

1. Formulas for Counting Graphs

Different types of bonds of a nonlinear graph are independent of each other. The same four vertexes of a graph, for example, may involve any number from zero to six 2-bonds, from zero to four 3-bonds, and zero or one 4-bond simultaneously in any combination. For a graph on n vertexes, define a *bond set* k(n) by

$$\mathrm{k}(n) = 0,\ k_2,\ k_3,\ \ldots,\ k_n, \tag{13}$$

where k_j is the number of j-bonds in the graph, and $k_1 = 0$ by convention.

Vertexes of a linear graph are associated in pairs only, by 2-bonds, with no more than one 2-bond connecting any given pair of vertexes. The number $N(n,k_2)$ of such graphs of k_2 2-bonds on n distinguishable vertexes is given by(47)

$$N(n,k_2) = \binom{B_2(n)}{k_2}, \tag{14}$$

where

$$B_2(n) = \binom{n}{2}$$

is the number of pairs of vertexes (or the maximum number of 2-bonds) in a graph of n vertexes, and where

$$\binom{n}{k} = \begin{cases} \dfrac{n!}{(n-k)!k!} & n \geqq k, \\ 0 & n < k \end{cases} \tag{15}$$

is the binomial coefficient, the number of combinations of n objects taken k at a time, n and k being nonnegative integers. Equation (14) gives simply the number of ways of distributing k_2 2-bonds between $B_2(n)$ pairs of vertexes.

By generalization of Eq. (14), for graphs with j-bonds only ($j = 2, 3, \ldots, n$), one has

$$N(n,k_j) = \binom{B_j(n)}{k_j} \quad (j = 2, 3, \ldots, n), \tag{16}$$

where $B_j(n)$, the maximum number of j-bonds in a graph of n vertexes, is given by

$$B_j(n) = \binom{n}{j} \quad (j = 2, 3, \ldots, n), \tag{17}$$

since each j vertexes in the graph may have at most one j-bond associating them.

Since

$$(1 + t)^n = \sum_{m=0}^{n} \binom{n}{m} t^m, \tag{18}$$

the total number $N_j(n)$ of possible graphs with j-bonds on n distinguishable vertexes, upon setting $t = 1$, is

$$N_j(n) = \sum_{k_j=0}^{B_j(n)} N(n,k_j) = 2^{\binom{n}{j}} \quad (j = 2, 3, \ldots, n). \tag{19}$$

In general, several different types of bonds may be independently present on the same graph; hence the total number of graphs on n distinguishable vertexes, with k_2 2-bonds, k_3 3-bonds, etc., is expressed by a product of terms of the type given in Eq. (16), and may be written

$$N(n,\mathrm{k}) = \binom{\mathrm{B}}{\mathrm{k}}, \tag{20}$$

where the conventions

$$\mathrm{B} = B_1, B_2, \ldots, B_n, \tag{21}$$

$$\binom{\mathrm{B}}{\mathrm{k}} = \prod_{j=1}^{n} \binom{B_j}{k_j}, \tag{22}$$

have been adopted, and the B_j are understood to be functions of n. In this notation Eq. (20) for nonlinear graphs takes the same form as Eq. (14) for linear graphs. Linear graphs are therefore equivalent to nonlinear graphs for which the number of 3-bonds and higher-order bonds is zero.

The total possible number of graphs on n distinguishable vertexes, using all combinations of different types of bonds, follows from Eqs. (18) and (19), and is

$$N(n) = \prod_{j=2}^{n} N_j(n) = 2^{2^n - n - 1}. \qquad (23)$$

The total number of nonlinear graphs on n vertexes is readily seen to be 2^{n-1} times the *square* of the number of graphs on $n-1$ vertexes. It is apparent that the counting of nonlinear graphs by inspection rapidly becomes overwhelmingly cumbersome. The total numbers of graphs on $n \leqq 10$ distinguishable vertexes, from Eqs. (19) and (23), are given in Table 1. Values of $N(n,\mathrm{k})$ for 2- and 3-bonds only on $n = 3$, 4, and 5 vertexes, calculated from Eq. (20), are presented in Tables 2, 3, and 4, respectively, as the values of $\mu(\mathrm{n},\mathrm{k})$ for $\mathrm{n} = 1,1,\ldots,1$ (n elements of unity), since if all vertexes are of different species, they are all distinguishable.

2. Generating Functions for Counting Trees and Stars

Generating functions(42) are useful in the combinatorial theory of graphs, in simplifying operations involving the coefficients. To obtain the number of trees $C(n,\mathrm{k})$ and the number of stars $S(n,\mathrm{k})$ with given $\mathrm{k}(n)$ on n distinguishable vertexes, write the generating functions

TABLE 1.—Total numbers of graphs on *n* distinguishable vertexes.

n	2-bonds Only $N_2(n)$	3-bonds Only $N_3(n)$	All Bonds $N(n)$
1	1	1	1
2	2	1	2
3	8	2	16
4	64	16	2048
5	1024	1024	2^{26}
6	2^{15}	2^{20}	2^{57}
7	2^{21}	2^{35}	2^{120}
8	2^{28}	2^{56}	2^{247}
9	2^{36}	2^{84}	2^{502}
10	2^{45}	2^{120}	2^{1013}

TABLE 2.—Numbers of graphs, trees, and stars of k_2 2-bonds and k_3 3-bonds on n vertexes, distinguishable by species only, for n = 3.

k_2, k_3	n = 1,1,1			n = 1,2			n = 3		
	$\mu(n,k)$	$\gamma(n,k)$	$\sigma(n,k)$	$\mu(n,k)$	$\gamma(n,k)$	$\sigma(n,k)$	$\mu(n,k)$	$\gamma(n,k)$	$\sigma(n,k)$
0,0	1	0	0	1	0	0	1	0	0
1,0	3	0	0	2	0	0	1	0	0
2,0	3	3	0	2	2	0	1	1	0
3,0		1			1			1	
0,1		1			1			1	
1,1		3			2			1	
2,1		3			2			1	
3,1		1			1			1	

TABLE 3.—Numbers of graphs, trees, and stars of k_2 2-bonds and k_3 3-bonds on n vertexes, distinguishable by species only, for n = 4.

k_2,k_3	n=1,1,1,1			n=1,1,2			n = 2,2			n = 1,3			n = 4		
	μ	γ	σ	μ	γ	σ	μ	γ	σ	μ	γ	σ	μ	γ	σ
0,0	1	0	0	1	0	0	1	0	0	1	0	0	1	0	0
1,0	6	0	0	4	0	0	3	0	0	2	0	0	1	0	0
2,0	15	0	0	9	0	0	6	0	0	4	0	0	2	0	0
3,0	20	16	0	12	9	0	8	6	0	6	4	0	3	2	0
4,0	15	15	3	9	9	2	6	6	2	4	4	1	2	2	1
5,0	6	6	6	4	4	4	3	3	3	2	2	2	1	1	1
6,0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0,1	4	0	0	3	0	0	2	0	0	2	0	0	1	0	0
1,1	24	12	0	14	7	0	8	4	0	6	3	0	2	1	0
2,1	60	48	12	33	26	7	18	14	4	13	10	3	4	3	1
3,1	80	76	40	44	41	22	24	22	12	18	16	9	6	5	3
4,1	60	60	48	33	33	26	18	18	14	13	13	10	4	4	3
5,1	24			14			8			6			2		
6,1	4			3			2			2			1		
0,2	6			4			3			2			1		
1,2	36			20			12			8			3		
2,2	90			48			27			18			6		
3,2	120			64			36			24			8		
4,2	90			48			27			18			6		
5,2	36			20			12			8			3		
6,2	6			4			3			2			1		
0,3	4			3			2			2			1		
1,3	24			14			8			6			2		
2,3	60			33			18			13			4		
3,3	80			44			24			18			6		
4,3	60			33			18			13			4		
5,3	24			14			8			6			2		
6,3	4			3			2			2			1		
0,4	1			1			1			1			1		
1,4	6			4			3			2			1		
2,4	15			9			6			4			2		
3,4	20			12			8			6			3		
4,4	15			9			6			4			2		
5,4	6			4			3			2			1		
6,4	1			1			1			1			1		

TABLE 4.—Numbers of graphs, μ(n,k), of k_2 2-bonds and k_3 3-bonds on n vertexes, distinguishable by species only, for $n = 5$.

k_2,k_3	n						
	1,1,1,1,1	1,1,1,2	1,2,2	1,1,3	2,3	1,4	5
0,0	1	1	1	1	1	1	1
1,0	10	7	5	4	3	2	1
2,0	45	27	17	12	8	5	2
3,0	120	68	40	29	18	11	4
4,0	210	116	66	47	28	17	6
5,0	252	138	78	54	32	18	6
6,0	210	116	66	47	28	17	6
7,0	120	68	40	29	18	11	4
8,0	45	27	17	12	8	5	2
9,0	10	7	5	4	3	2	1
10,0	1	1	1	1	1	1	1
0,1	10	7	5	4	3	2	1
1,1	100	58	34	25	15	9	3
2,1	450	243	133	93	52	29	8
3,1	1200	632	336	233	126	69	18
4,1	2100	1094	574	395	210	113	28
5,1	2520	1308	684	468	248	132	32
6,1	2100	1094	574	395	210	113	28
7,1	1200	632	336	233	126	69	18
8,1	450	243	133	93	52	29	8
9,1	100	58	34	25	15	9	3
10,1	10	7	5	4	3	2	1
0,2	45	27	17	12	8	5	2

TABLE 4—Continued

k_2,k_3	n						
	1,1,1,1,1	1,1,1,2	1,2,2	1,1,3	2,3	1,4	5
1,2	450	243	133	93	52	29	8
2,2	2025	1053	553	378	202	108	27
3,2	5400	2772	1432	972	508	266	62
4,2	9450	4824	2474	1676	866	450	102
5,2	11340	5778	2958	1998	1032	534	120
6,2	9450	4824	2474	1676	866	450	102
7,2	5400	2772	1432	972	508	266	62
8,2	2025	1053	553	378	202	108	27
9,2	450	243	133	93	52	29	8
10,2	45	27	17	12	8	5	2
0,3	120	68	40	29	18	11	4
1,3	1200	632	336	233	126	69	18
2,3	5400	2772	1432	972	508	266	62
3,3	14400	7328	3744	2531	1303	675	151
4,3	25200	12776	6496	4379	2239	1151	251
5,3	30240	15312	7776	5232	2672	1368	296
6,3	25200	12776	6496	4379	2239	1151	251
7,3	14400	7328	3744	2531	1303	675	151
8,3	5400	2772	1432	972	508	266	62
9,3	1200	632	336	233	126	69	18
10,3	120	68	40	29	18	11	4
0,4	210	116	66	47	28	17	6
1,4	2100	1094	574	395	210	113	28
2,4	9450	4824	2474	1676	866	450	102
3,4	25200	12776	6496	4379	2239	1151	251
4,4	44100	22292	11292	7595	3863	1975	423

TABLE 4—Continued

k_2,k_3	n						
	1,1,1,1,1	1,1,1,2	1,2,2	1,1,3	2,3	1,4	5
5,4	52920	26724	13524	9084	4616	2352	500
6,4	44100	22292	11292	7595	3863	1975	423
7,4	25200	12776	6496	4379	2239	1151	251
8,4	9450	4824	2474	1676	866	450	102
9,4	2100	1094	574	395	210	113	28
10,4	210	116	66	47	28	17	6
0,5	252	138	78	54	32	18	6
1,5	2520	1308	684	468	248	132	32
2,5	11340	5778	2958	1998	1032	534	120
3,5	30240	15312	7776	5232	2672	1368	296
4,5	52920	26724	13524	9084	4616	2352	500
5,5	63504	32040	16200	10872	5520	2808	596
6,5	52920	26724	13524	9084	4616	2352	500
7,5	30240	15312	7776	5232	2672	1368	296
8,5	11340	5778	2958	1998	1032	534	120
9,5	2520	1308	684	468	248	132	32
10,5	252	138	78	54	32	18	6
0,6	210	116	66	47	28	17	6
1,6	2100	1094	574	395	210	113	28
2,6	9450	4824	2474	1676	866	450	102
3,6	25200	12776	6496	4379	2239	1151	251
4,6	44100	22292	11292	7595	3863	1975	423
5,6	52920	26724	13524	9084	4616	2352	500
6,6	44100	22292	11292	7595	3863	1975	423
7,6	25200	12776	6496	4379	2239	1151	251
8,6	9450	4824	2474	1676	866	450	102

TABLE 4—Continued

k_2,k_3	n						
	1,1,1,1,1	1,1,1,2	1,2,2	1,1,3	2,3	1,4	5
9,6	2100	1094	574	395	210	113	28
10,6	210	116	66	47	28	17	6
0,7	120	68	40	29	18	11	4
1,7	1200	632	336	233	126	69	18
2,7	5400	2772	1432	972	508	266	62
3,7	14400	7328	3744	2531	1303	675	151
4,7	25200	12776	6496	4379	2239	1151	251
5,7	30240	15312	7776	5232	2672	1368	296
6,7	25200	12776	6496	4379	2239	1151	251
7,7	14400	7328	3744	2531	1303	675	151
8,7	5400	2772	1432	972	508	266	62
9,7	1200	632	336	233	126	69	18
10,7	120	68	40	29	18	11	4
0,8	45	27	17	12	8	5	2
1,8	450	243	133	93	52	29	8
2,8	2025	1053	553	378	202	108	27
3,8	5400	2772	1432	972	508	266	62
4,8	9450	4824	2474	1676	866	450	102
5,8	11340	5778	2958	1998	1032	534	120
6,8	9450	4824	2474	1676	866	450	102
7,8	5400	2772	1432	972	508	266	62
8,8	2025	1053	553	378	202	108	27
9,8	450	243	133	93	52	29	8
10,8	45	27	17	12	8	5	2
0,9	10	7	5	4	3	2	1
1,9	100	58	34	25	15	9	3

TABLE 4—Continued

k_2, k_3	n						
	1,1,1,1,1	1,1,1,2	1,2,2	1,1,3	2,3	1,4	5
2,9	450	243	133	93	52	29	8
3,9	1200	632	336	233	126	69	18
4,9	2100	1094	574	395	210	113	28
5,9	2520	1308	684	468	248	132	32
6,9	2100	1094	574	395	210	113	28
7,9	1200	632	336	233	126	69	18
8,9	450	243	133	93	52	29	8
9,9	100	58	34	25	15	9	3
10,9	10	7	5	4	3	2	1
0,10	1	1	1	1	1	1	1
1,10	10	7	5	4	3	2	1
2,10	45	27	17	12	8	5	2
3,10	120	68	40	29	18	11	4
4,10	210	116	66	47	28	17	6
5,10	252	138	78	54	32	18	6
6,10	210	116	66	47	28	17	6
7,10	120	68	40	29	18	11	4
8,10	45	27	17	12	8	5	2
9,10	10	7	5	4	3	2	1
10,10	1	1	1	1	1	1	1

TABLE 5.—Numbers of trees, γ(n,k), of k_2 2-bonds and k_3 3-bonds on n vertexes, distinguishable by species only, for $n = 5$. For $k_3 > 4$, γ(n,k) = μ(n,k), as in Table 4.

k_2,k_3	n						
	1,1,1,1,1	1,1,1,2	1,2,2	1,1,3	2,3	1,4	5
0,0	0	0	0	0	0	0	0
1,0	0	0	0	0	0	0	0
2,0	0	0	0	0	0	0	0
3,0	0	0	0	0	0	0	0
4,0	125	67	37	26	15	9	3
5,0	222	120	67	46	27	15	5
6,0	205	112	63	44	26	15	5
7,0	120	68	40	29	18	11	4
8,0	45	27	17	12	8	5	2
9,0	10	7	5	4	3	2	1
10,0	1	1	1	1	1	1	1
0,1	0	0	0	0	0	0	0
1,1	0	0	0	0	0	0	0
2,1	150	81	45	31	18	10	3
3,1	780	408	216	148	80	43	11
4,1	1790	928	485	332	176	94	23
5,1	2400	1242	648	442	234	122	30
6,1	2080	1081	566	388	206	108	27
7,1	1200	632	336	233	126	69	18
8,1	450	243	133	93	52	29	8
9,1	100	58	34	25	15	9	3
10,1	10	7	5	4	3	2	1

TABLE 5—Continued

k_2, k_3	n						
	1,1,1,1,1	1,1,1,2	1,2,2	1,1,3	2,3	1,4	5
0,2	15	9	6	4	3	2	1
1,2	270	147	81	57	32	18	5
2,2	1575	819	430	294	157	84	21
3,2	4800	2460	1268	860	448	234	54
4,2	9000	4590	2351	1592	821	426	96
5,2	11160	5682	2906	1962	1012	523	117
6,2	9420	4806	2463	1668	861	447	101
7,2	5400	2772	1432	972	508	266	62
8,2	2025	1053	553	378	202	108	27
9,2	450	243	133	93	52	29	8
10,2	45	27	17	12	8	5	2
0,3	100	55	32	22	14	8	3
1,3	1080	566	300	207	112	61	16
2,3	5100	2613	1348	913	477	249	58
3,3	14000	7116	3632	2451	1261	651	145
4,3	24900	12617	6412	4320	2208	1134	247
5,3	30120	15246	7740	5206	2658	1360	294
6,3	25180	12763	6488	4372	2235	1148	250
7,3	14400	7328	3744	2531	1303	675	151
8,3	5400	2772	1432	972	508	266	62
9,3	1200	632	336	233	126	69	18
10,3	120	68	40	29	18	11	4
0,4	205	112	63	44	26	15	5
1,4	2070	1076	563	387	205	110	27
2,4	9375	4782	2450	1659	856	444	100
3,4	25100	12720	6464	4355	2225	1142	248

TABLE 5—Continued

k_2,k_3	n						
	1,1,1,1,1	1,1,1,2	1,2,2	1,1,3	2,3	1,4	5
4,4	44025	22250	11268	7578	3853	1969	421
5,4	52890	26706	13513	9076	4611	2349	499
6,4	44095	22288	11289	7592	3861	1973	422
7,4	25200	12776	6496	4379	2239	1151	251
8,4	9450	4824	2474	1676	866	450	102
9,4	2100	1094	574	395	210	113	28
10,4	210	116	66	47	28	17	6

TABLE 6.—Numbers of stars, σ(n,k), of k_2 2-bonds and k_3 3-bonds on n vertexes, distinguishable by species only, for $n = 5$. For $k_3 > 4$, σ(n,k) = μ(n,k), as in Table 4.

k_2,k_3	n						
	1,1,1,1,1	1,1,1,2	1,2,2	1,1,3	2,3	1,4	5
0,0	0	0	0	0	0	0	0
1,0	0	0	0	0	0	0	0
2,0	0	0	0	0	0	0	0
3,0	0	0	0	0	0	0	0
4,0	0	0	0	0	0	0	0
5,0	12	6	4	2	2	1	1
6,0	70	37	21	14	9	5	2
7,0	100	55	32	22	14	8	3
8,0	45	27	17	12	8	5	2
9,0	10	7	5	4	3	2	1
10,0	1	1	1	1	1	1	1
0,1	0	0	0	0	0	0	0
1,1	0	0	0	0	0	0	0
2,1	0	0	0	0	0	0	0
3,1	60	30	16	10	6	3	1
4,1	450	231	121	81	44	23	6
5,1	1200	618	322	218	116	59	15
6,1	1570	811	423	288	153	79	20
7,1	1120	586	310	213	115	62	16
8,1	450	243	133	93	52	29	8
9,1	100	58	34	25	15	9	3
10,1	10	7	5	4	3	2	1

TABLE 6—Continued

k_2,k_3	n						
	1,1,1,1,1	1,1,1,2	1,2,2	1,1,3	2,3	1,4	5
0,2	0	0	0	0	0	0	0
1,2	60	30	16	10	6	3	1
2,2	630	324	169	114	61	32	8
3,2	2700	1380	712	480	252	131	31
4,2	6375	3243	1658	1120	577	298	67
5,2	9270	4713	2409	1623	838	432	97
6,2	8685	4425	2265	1532	790	409	92
7,2	5280	2706	1396	946	494	258	60
8,2	2025	1053	553	378	202	108	27
9,2	450	243	133	93	52	29	8
10,2	45	27	17	12	8	5	2
0,3	100	55	32	22	14	8	3
1,3	1000	520	274	187	101	54	14
2,3	4620	2361	1216	821	429	223	52
3,3	12800	6498	3314	2233	1149	592	132
4,3	23300	11793	5988	4028	2058	1054	229
5,3	28920	14628	7422	4988	2546	1301	281
6,3	24700	12511	6356	4280	2187	1122	244
7,3	14320	7282	3718	2511	1292	668	149
8,3	5400	2772	1432	972	508	266	62
9,3	1200	632	336	233	126	69	18
10,3	120	68	40	29	18	11	4
0,4	205	112	63	44	26	15	5
1,4	2050	1063	555	380	201	107	26
2,4	9255	4716	2414	1633	842	436	98
3,4	24800	12561	6380	4296	2194	1125	244

TABLE 6—Continued

k_2, k_3	n						
	1,1,1,1,1	1,1,1,2	1,2,2	1,1,3	2,3	1,4	5
4,4	43625	22038	11156	7498	3811	1945	415
5,4	52590	26547	13429	9017	4580	2332	495
6,4	43975	22222	11253	7566	3847	1965	420
7,4	25180	12763	6488	4372	2235	1148	250
8,4	9450	4824	2474	1676	866	450	102
9,4	2100	1094	574	395	210	113	28
10,4	210	116	66	47	28	17	6

$$\begin{aligned} N'(x,\mathrm{y}) &= \sum_{n=1}^{\infty} N_n \frac{x^n}{n!} = \sum_{n=1}^{\infty} \sum_{\mathrm{k}(n)} \frac{1}{n!} N(n,\mathrm{k}) x^n \mathrm{y}^{\mathrm{k}(n)}, \\ C'(x,\mathrm{y}) &= \sum_{n=1}^{\infty} C_n \frac{x^n}{n!} = \sum_{n=1}^{\infty} \sum_{\mathrm{k}(n)} \frac{1}{n!} C(n,\mathrm{k}) x^n \mathrm{y}^{\mathrm{k}(n)}, \qquad (24) \\ S'(x,\mathrm{y}) &= \sum_{n=1}^{\infty} S_n \frac{x^n}{n!} = \sum_{n=1}^{\infty} \sum_{\mathrm{k}(n)} \frac{1}{n!} S(n,\mathrm{k}) x^n \mathrm{y}^{\mathrm{k}(n)}, \end{aligned}$$

where x and

$$\mathrm{y} = y_1, y_2, \ldots, y_n \qquad (25)$$

are arbitrary variables, $\mathrm{k}(n)$ takes on all possible values consistent with $k_1 = 0$, $0 \leqq k_j \leqq B_j(n)$ $(j = 2, 3, \ldots, n)$, and

$$\mathrm{y}^{\mathrm{k}(n)} = \prod_{j=1}^{n} y_j^{k_j} \qquad (26)$$

is a weight assigned to each graph, dependent on its bond set only. This weight is equal to the product of the weights assigned to each of the trees which make up a graph, or to each of the stars forming a tree. Since the coefficients $N(n,\mathrm{k})$ for graphs are known by an explicit formula, Eq. (20), relationships between the generating functions (24) will permit the corresponding coefficients for trees and stars to be obtained. The product theorem, as stated by Uhlenbeck and Ford,$^{(47)}$ will be used to derive the desired relationships between the generating functions.

3. The Product Theorem

Because of its importance in graph combinatorial theory, the product theorem will be stated and proved for the case of one variable. The statement of the theorem and a sketch of its proof are given by Uhlenbeck and Ford.(47)

Given two collections of graphs on distinguishable vertexes, defined by the generating functions

$$g(x) = \sum_{n=1}^{\infty} g_n \frac{x^n}{n!},$$

$$h(x) = \sum_{n=1}^{\infty} h_n \frac{x^n}{n!},$$

where g_n and h_n are total weights assigned to all graphs on n vertexes in the two collections, and where the weights are independent of any labeling of the vertexes, a product collection

$$H(x) = \sum_{n=2}^{\infty} H_n \frac{x^n}{n!}$$

can be constructed, which consists of all graphs which can be formed by selecting one graph from each collection. The weight of a graph in the product collection is taken to be the product of the weights of the two graphs forming it, and H_n is the total weight of all graphs on n vertexes in the product collection.

Theorem: $H(x) = g(x)h(x)$.

Proof: From the above definitions,

$$H_n = \sum_{n_1=1}^{n-1} \binom{n}{n_1} g_{n_1} h_{n-n_1},$$

where the factor $\binom{n}{n_1}$ accounts for the number of ways the n vertexes can be distributed between the two graphs of n and $n-n_1$ vertexes, respectively, from the two collections. Let $n_2 = n - n_1$. Then

$$\begin{aligned} g(x)h(x) &= \left(\sum_{n_1=1}^{\infty} g_{n_1} \frac{x^{n_1}}{n_1!}\right)\left(\sum_{n_2=1}^{\infty} h_{n_2} \frac{x^{n_2}}{n_2!}\right) \\ &= \sum_{n=2}^{\infty} \sum_{n_1+n_2=n} \frac{n!}{n_1!n_2!} g_{n_1} h_{n_2} \frac{x^{n_1+n_2}}{n!} \\ &= \sum_{n=2}^{\infty} \sum_{n_1=1}^{n-1} \binom{n}{n_1} g_{n_1} h_{n-n_1} \frac{x^n}{n!} \\ &= \sum_{n=2}^{\infty} H_n \frac{x^n}{n!} = H(x), \end{aligned}$$

which establishes the theorem.

Corollary 1: For two collections of graphs on indistinguishable vertexes, defined by the generating functions

$$g(x) = \sum_{n=1}^{\infty} g_n x^n,$$

$$h(x) = \sum_{n=1}^{\infty} h_n x^n,$$

the product collection $H(x)$ is given by

$$H(x) = \sum_{n=2}^{\infty} H_n x^n = g(x)h(x).$$

Proof: Since the vertexes are indistinguishable, the factorials in the proof of the product theorem are omitted, and the corollary follows directly.

Corollary 2 (Product Property):(47) The generating function $H(x)$ for the number of pairs of graphs from two *independent* collections, counted by $g(x)$ and $h(x)$, is given by $H(x) = g(x)h(x)$.

Proof: The weights g_n, h_n, and H_n in the generating functions of Corollary 1 are the total numbers of graphs on n vertexes in $g(x)$, $h(x)$, and $H(x)$, respectively. The vertexes may be either distinguishable or indistinguishable, but since the two collections of graphs are independent, the vertexes in the collection of pairs may be distributed between the two graphs of the pair in only the given way. Hence no factorials appear, and the corollary follows directly.

Corollary 1 is original with this work, and Corollary 2 is termed the Product Property, and is stated but not proved, by Uhlenbeck and Ford.(47)

4. Application to Counting Trees

Write the generating function for nonlinear graphs on distinguishable vertexes in the form

$$N'(x,y) = \sum_{t=1}^{\infty} N'_t(x,y),$$

where $N'_t(x,y)$ is the generating function for all graphs which consist of exactly t trees on distinguishable vertexes. Since any graph of t trees can be constructed by selecting a tree from each of t collections of all trees, it follows from the product theorem that

$$N'_t(x,y) = \frac{1}{t!}\{C'(x,y)\}^t,$$

where the $t!$ must be inserted since the t parts are really chosen from the same collection of trees, and any permutation of the t trees leads to the same graph. Then

$$\begin{aligned} N'(x,y) &= \sum_{t=1}^{\infty} \frac{1}{t!}\{C'(x,y)\}^t \\ &= \exp\{C'(x,y)\} - 1, \end{aligned}$$

or

$$C'(x,y) = \ln\{1+N'(x,y)\}, \qquad (27)$$

which relates the generating functions for trees and graphs. Equation (27) is a straightforward extension to nonlinear graphs of the corresponding theorem for linear graphs.(47)

By expanding the logarithm of Eq. (27), then substituting Eqs. (17) and (20) and equating coefficients of $x^n \mathrm{y}^{\mathrm{k}(n)}$, there results, in direct analogy to the result for linear graphs,$^{(47)}$

$$C(n,\mathrm{k}) = \sum_{m=1}^{n} \frac{(-1)^{m+1}}{m} \sum_{c(n)} \frac{n!}{\mathrm{n}!} \binom{\mathrm{K}}{\mathrm{k}}, \tag{28}$$

where the second summation is over all 2^{n-1} compositions (ordered partitions)$^{(29)}$ of n into n_1, n_2, ..., n_m such that $n_i = 1, 2, \ldots, n-m+1$ $(i = 1, 2, \ldots, m)$, and

$$\begin{aligned} \mathrm{n}! &= \prod_{i=1}^{m} n_i!, \\ \mathrm{K} &= K_1, K_2, \ldots, K_n, \\ K_j &= \sum_{i=1}^{m} \binom{n_i}{j} \qquad (j = 1, 2, \ldots, n). \end{aligned} \tag{29}$$

The combinatorial convention for $\binom{\mathrm{K}}{\mathrm{k}}$ is that of Eq. (18). For $j = 1$, the contribution to the product is unity, since $k_1 = 0$; hence the value of K_1 is immaterial. Equation (28) is derived in Appendix A, and numerical values of $C(n,\mathrm{k})$, calculated from Eq. (28) for 2- and 3-bonds only on $n = 3$, 4, and 5 vertexes, are given in Tables 2, 3, and 5, respectively, as $\gamma(\mathrm{n},\mathrm{k})$ for $\mathrm{n} = 1,1,\ldots,1$ (n elements of unity).

5. Application to Counting Stars

Consider topologically distinct stars on indistinguishable vertexes as stars of different *types*. Assign a weight y^{k_i} to a star of type i in a tree, where k_i is the bond set for a star of type i. The total weight of the tree depends only on the bond set of the tree, and is equal to the product of the weights assigned to the stars of which the tree is formed. To each tree on n distinguishable vertexes there correspond n *rooted trees*, in which one vertex, the *root*, is given a special designation. If the generating function for rooted trees on distinguishable vertexes is denoted by a superscript degree sign instead of a prime, it follows that

$$\begin{aligned} C^\circ(x,y) &= \sum_{n=1}^{\infty} \sum_{k(n)} \frac{nC(n,k)}{n!} x^n y^{k(n)} \\ &= x \frac{\partial C'(x,y)}{\partial x}. \qquad (30) \end{aligned}$$

In a rooted tree, call the stars which have the root vertex in common *trunks*, and let $C_m^\circ(x,y)$ be the generating function for rooted trees with exactly m trunks. Then

$$C^\circ(x,y) = \sum_{m=0}^{\infty} C_m^\circ(x,y),$$

with

$$C_0^\circ(x,y) = x.$$

Now a rooted tree with m trunks can be constructed by selecting m rooted trees with one trunk and joining them all at the root. From the product theorem it follows that

$$C_m^\circ(x,y) = \frac{x}{m!}\{C_1^\circ(x,y)/x\}^m,$$

since $C_1^\circ(x,y)/x$ is the generating function for rooted trees with one trunk, in which the root vertex is removed; the factor x accounts for the root, and the $m!$ is needed since the m root vertexes are identical. Hence

$$C^\circ(x,y) = x \exp\{C_1^\circ(x,y)/x\}. \tag{31}$$

To obtain $C_1^\circ(x,y)$, hang rooted trees on the vertexes of the trunk, other than the root. Then from the product theorem, there results

$$C_1^\circ(x,y) = x \sum_i \frac{n_i}{s_i}\, y^{k_i}\{C^\circ(x,y)\}^{n_i-1}, \tag{32}$$

where the trunk is a star of type i, having n_i vertexes. The factor x accounts for the root; the factor y^{k_i} accounts for the weight of the trunk; the factor n_i is the number of ways the root vertex can be selected; and the factor s_i, the symmetry number of a star of type i, is the number of equivalent arrangements of root and rooted trees on the trunk because of its symmetry. To distinguish between the different types of stars which can be the trunk of the rooted tree with one trunk, it is necessary to sum over all star types.

Since there are $n_i!/s_i$ stars of type i on n_i distinguishable vertexes, by summing over all the types of stars instead of over the number of vertexes of the stars, from Eq. (24) there results

$$S'(x,y) = \sum_i \frac{1}{s_i} x^{n_i} y^{k_i}.$$

Then

$$S'(C^\circ,y) = \sum_i \frac{1}{s_i} (C^\circ)^{n_i} y^{k_i},$$

and

$$\begin{aligned}\frac{\partial S'(C^\circ,y)}{\partial C^\circ} &= \sum_i \frac{n_i}{s_i}\, y^{k_i} (C^\circ)^{n_i-1} \\ &= C_1^\circ/x \\ &= \ell n\ (C^\circ/x), \qquad (33)\end{aligned}$$

from Eqs. (31) and (32). Equations (30) and (33), analogous to equations given by Uhlenbeck and Ford for linear graphs,$^{(47)}$ relate the generating functions $C'(x,y)$ and $S'(x,y)$. Upon substituting Eqs. (24) and (30) into Eq. (33), and expanding the logarithm in the same manner as for Eq. (28) (see Appendix A), there results

$$\sum_{n=1}^{\infty} \frac{1}{(n-1)!} \left\{ \sum_{m=1}^{\infty} \frac{x^m C_m}{(m-1)!} \right\}^{n-1} S_n = \sum_{m=1}^{\infty} \frac{(-1)^{m+1}}{m} \left\{ \sum_{n=1}^{\infty} \frac{x^n C_{n+1}}{n!} \right\}^m. \qquad (34)$$

By expanding and equating coefficients of like powers of x, the following relationships between the successive C_n and S_n are obtained:

$$\begin{aligned}
C_1 - S_1 &= 1 \quad (C_1 = 1),\\
C_2 - S_2 &= 0,\\
C_3 - S_3 &= 2C_2S_2 + C_2^2,\\
C_4 - S_4 &= 6C_2S_3 + 3C_3S_2 + 3C_2C_3 - 2C_2^3,\\
C_5 - S_5 &= 12C_2S_4 + 12(C_3+C_2^2)S_3 + 4C_4S_2 + 4C_2C_4\\
&\quad + 3C_3^2 - 12C_2^2C_3 + 6C_2^4,\\
C_6 - S_6 &= 20C_2S_5 + 30(C_3+2C_2^2)S_4 + 20(C_4+3C_2C_3)S_3\\
&\quad + 5C_5S_2 + 5C_2C_5 + 10C_3C_4 - 20C_2^2C_4\\
&\quad -30C_2C_3^2 + 60C_2^3C_3 - 24C_2^5,\\
C_7 - S_7 &= 30C_2S_6 + 60(C_3+3C_2^2)S_5 + 60(C_4+6C_2C_3+2C_2^3)S_4\\
&\quad + 30(C_5+4C_2C_4+3C_3^2)S_3 + 6C_6S_2 + 6C_2C_6\\
&\quad + 15C_3C_5 - 30C_2^2C_5 + 120C_2^3C_4 - 120C_2C_3C_4\\
&\quad + 20C_4^2 - 72C_2^4C_3 + 270C_2^2C_3^2 - 30C_3^3 + 120C_2^6.
\end{aligned}$$

From Eqs. (24), with $C(n,\mathrm{k})$ known from Eq. (28), one may obtain $S(n,\mathrm{k})$ successively. An explicit expression for $S(n,\mathrm{k})$, such as Eq. (28), is not possible because x and y are involved in power series inside the summations on both sides of Eq. (34). Numerical values of $S(n,\mathrm{k})$ are given in Tables 2, 3, and 6 as $\sigma(\mathrm{n,k})$ for n = 1,1,...,1 (n elements of unity).

D. COUNTING OF GRAPHS AND TREES ON INDISTINGUISHABLE VERTEXES

While the counting of nonlinear graphs, trees, and stars on distinguishable vertexes has been solved by explicit formulas or in terms of generating functions for the coefficients, the corresponding problem for graphs on indistinguishable vertexes is of much greater difficulty, and more powerful methods are necessary for the solution.

1. Generating Functions

To determine the numbers of topologically distinct nonlinear graphs, trees, and stars on indistinguishable vertexes, $\mu(n,k)$, $\gamma(n,k)$, and $\sigma(n,k)$, respectively, begin with the generating functions

$$\begin{aligned}
\mu'(x,y) &= \sum_{n=1}^{\infty} \mu_n x^n = \sum_{n=1}^{\infty} \sum_{k(n)} \mu(n,k) x^n y^{k(n)}, \\
\gamma'(x,y) &= \sum_{n=1}^{\infty} \gamma_n x^n = \sum_{n=1}^{\infty} \sum_{k(n)} \gamma(n,k) x^n y^{k(n)}, \qquad (35) \\
\sigma'(x,y) &= \sum_{n=1}^{\infty} \sigma_n x^n = \sum_{n=1}^{\infty} \sum_{k(n)} \sigma(n,k) x^n y^{k(n)}.
\end{aligned}$$

Use of the product theorem with the generating functions for graphs on distinguishable vertexes requires $1/n!$ in the definition of the generating functions (24). Since the vertexes of the graphs are now indistinguishable, there is no $1/n!$ in Eqs. (35).

2. Pólya's Theorem

Basic to combinatorial theory for graphs on indistinguishable vertexes is a powerful counting theorem developed by Pólya.(35) Various statements and proofs of the theorem are given by Harary,(19,21) Uhlenbeck and Ford,(47) Riordan,(41,42) and de Bruijn.(11) Since some concepts of group theory are required, Appendix B or Hall's book(18) may be helpful. Pólya's theorem may be stated briefly in the following way:

Let there be a collection of "figures," each with an assigned weight, or "content," defined by a figure generating function. Define a "configuration" by placing figures from the collection on a set S of n "points," one figure to a point, repetitions of the figures allowed. Define a *cycle index* $Z(G)$ of the group G of permutations of the set S by

$$Z(G) = \frac{1}{|G|} \sum_{g \in G} f^b, \tag{36}$$

where $|G|$ is the order of the group, the summation is over all elements of the group, f is a set of arbitrary functions f_i, f^b follows the convention of Eq. (39) below, and

b_i is the number of cycles of length i in the permutation of S corresponding to g.

Theorem: The configuration generating function is obtained by substituting the figure generating function, with its arbitrary variables each raised to the ith power, for f_i in $Z(G)$.

In this work the theorem is generalized to apply to nonlinear graphs with more than one type of bond present. A proof will be given for this special case, following the development of the necessary definitions and terminology.

3. Definitions of Terms for Pólya's Theorem

It is necessary to define the terms given above in quotation marks for Pólya's theorem as they apply specifically to counting nonlinear graphs. The "figures" for Pólya's theorem are j nonadjacent vertexes and j vertexes adjacent by a j-bond (j = 2, 3, ...). The figure collection for j-bonds then consists of two figures, with "contents" 0 and 1, respectively, and the figure collections for different types of bonds are independent of each other. The figure generating function $f(y_j)$ for j-bonds is then given by

$$f(y_j) = 1 + y_j \quad (j = 2, 3, \ldots), \tag{37}$$

where the coefficient of the k_jth power of y_j is the number of "figures" of k_j j-bonds in the collection ($k_j = 0,1$). The corresponding set of "points" for Pólya's theorem is the set of the $\binom{n}{j}$ j's of vertexes for graphs on n vertexes, for each $j = 2, 3, \ldots, n$. The term "point" is reserved exclusively for this usage with Pólya's theorem, so as not to be confused with a "vertex" of a graph, as may occur with Uhlenbeck and Ford's terminology.(47) The sets of points are independent for different j. A configuration is obtained by placing figures from the figure collection for j-bonds on the corresponding points, one figure to a point, for each j. The group G_j of Pólya's theorem is the group of permutations of the $\binom{n}{j}$ points induced by the symmetric group S_n of permutations of the vertexes. The group G_j is of degree $\binom{n}{j}$ and order $n!$, and is isomorphic to S_n, for all j.(18,47)

4. The Cycle Function

For convenience a number of graph-theoretical and combinatorial terms will be introduced.(41) The novel term "cycle function" is basic to the extension of Pólya's theorem to nonlinear graphs.

A *partition* of a positive integer n is a subdivision of the integer into a number k of positive integers ($k = 1, 2, \ldots, n$), the sum of which equals n. Define the *partition* *set* p(n) by

$$\mathrm{p}(n) = p_1, p_2, \ldots, p_n, \tag{38}$$

where p_i is the number of times the integer i occurs in the corresponding partition of n. It follows that

$$\sum_{i=1}^{n} i p_i = n.$$

A permutation of numbered elements may be described completely by a *cycle* notation,(41) in which each element in a cycle is replaced by the next element in the cycle, the last element in the cycle being replaced by the first. Thus the permutation taking the elements 12345 into 25431 may be represented uniquely by (125)(34). The *length* of a cycle is the number of elements in the cycle. In this example there is one cycle of length 3 and one cycle of length 2.

Define the *cycle function* $F_j(n,\mathrm{p})$ for j-bonds on n vertexes by

$$F_j(n,\mathrm{p}) = \mathrm{f}^{\mathrm{b}} = \prod_i f_i^{b_i}, \tag{39}$$

where f is a set of arbitrary functions f_i (i = 1, 2, ...) and b_i is the number of cycles of length i in the element of G_j corresponding to the permutation of S_n consisting of p_i cycles of length i = 1, 2, ..., n. Term G_2 the vertex pair group, $F_2(n,\mathrm{p})$ the pair cycle function; G_3 the vertex triplet group, $F_3(n,\mathrm{p})$ the triplet cycle function; etc.

The pair cycle function, for 2-bonds, has been obtained previously,(47) and may be expressed in the compact form

$$F_2(n,\mathrm{p}) = \prod_{i=1}^{n} f_i^{\,p_i[(i-1)/2]+\binom{p_i}{2}i+p_{2i}} \prod_{i<j} f_{m(i,j)}^{\,p_ip_jd(i,j)}, \tag{40}$$

where the square bracket indicates the integral part of the term contained within it, and where $m(i,j)$ and $d(i,j)$ are the least common multiple and greatest common divisor, respectively, of i and j. It follows that

$$m(i,j)d(i,j) = ij. \tag{41}$$

By convention, $p_k = 0$ if $k > n$.

To obtain the triplet cycle function, for 3-bonds, it is necessary to distinguish three cases, in which vertexes of the triplet occur in the same cycle, in two different cycles, or in three different cycles of the permutation in the vertex triplet group, G_3:

(1) If the vertexes occur in the same cycle, of length i, of the permutation in the triplet group, in each case the total number of triplets is $\binom{i}{3}$. After i cyclic permutations they will repeat themselves, and hence will induce cycles of length i in the triplet group. Therefore the triplets are divided into cycles of length i as far as

possible. Then for i not divisible by 3, there result $\frac{1}{3}\binom{i-1}{2}$ cycles of length i in the triplet group; for i divisible by 3, there are $\left[\frac{1}{3}\binom{i-1}{2}\right]$ cycles of length i plus one cycle of length $i/3$.

(2) If the three vertexes occur in two different cycles of equal length i, there are $\binom{i}{2}$ combinations of vertexes possible in the cycle containing two vertexes. After i cyclic permutations the cycles will repeat themselves, hence will induce $\binom{i}{2}$ cycles of length i in the triplet group.

If the two cycles are of unequal length, say i containing one vertex and j two vertexes ($i \neq j$), again calculate the total number of triplets in each case, the length of the cycles, and then the number of cycles. The number of pairs of vertexes in j is $\binom{j}{2}$. Each of these may be combined with each element of the cycle of length i, to yield $i\binom{j}{2}$ triplets in all. After at most $m(i,j)$ permutations, the cycles will repeat themselves. From Eq. (41), $j/d(i,j)$ is even if and only if i is a factor of $m(i,j)/2$. There are cycles of length $m(i,j)/2$ if both the cycles of i and j repeat themselves after $m(i,j)/2$ permutations. This can happen only if i is a factor of $m(i,j)/2$; that is, if $j/d(i,j)$ is even. Therefore if $j/d(i,j)$ is odd, there will be exactly $i\binom{j}{2}/m(i,j) = (j-1)d(i,j)/2$ cycles of length $m(i,j)$. If $j/d(i,j)$ is even, then j is even, and there will then be $d(i,j)$ cycles of length $m(i,j)/2$, corresponding to the different combinations of the added cycle in the pair group with the single other vertex to form the triplet group. The number of cycles of length $m(i,j)$ is then $(j/2 - 1)d(i,j)$.

(3) If the three vertexes occur in three different cycles, of lengths i, j, and k, the total number of different triplets is ijk. The length of the cycles is $m(i,j,k)$, since after $m(i,j,k)$ permutations, the cycles will repeat themselves. Because of their independence, no cycles will repeat themselves with fewer than $m(i,j,k)$ permutations. Therefore there result $ijk/m(i,j,k)$ cycles of length $m(i,j,k)$ in the triplet group.

Combination of these results for the triplet cycle function yields, from Eq. (39),

$$F_3(n,\mathrm{p}) = \prod_{i=1}^{n} f_i^{p_i[\frac{1}{3}\binom{i-1}{2}]} \prod_{i=3,6,9,\ldots} f_{i/3}^{p_i} \prod_{i=1}^{n} f_i^{2\binom{p_i}{2}\binom{i}{2}} \cdot$$

$$\cdot \prod_{\substack{i\neq j \\ j/d \text{ odd}}} f_{m(i,j)}^{p_i p_j (j-1)d(i,j)/2} \cdot$$

$$\cdot \prod_{\substack{i\neq j \\ j/d \text{ even}}} f_{m(i,j)}^{p_i p_j (j/2-1)d(i,j)} f_{m(i,j)/2}^{p_i p_j d(i,j)} \cdot$$

$$\cdot \prod_{i,j,k} f_{m(i,j,k)}^{p_i p_j p_k ijk/m(i,j,k)} \cdot \qquad (42)$$

In the first product, there are p_i cycles of length i in the symmetric group of the vertexes, and for each of these there result $\left[\frac{1}{3}\binom{i-1}{2}\right]$ cycles of length i in the triplet group. The second product represents the added cycle of length $i/3$ for i divisible by 3. In the third product, there are two cycles of equal length i out of p_i total cycles of length i, giving $\binom{p_i}{2}$ combinations. The factor 2 is needed since the cycle containing two vertexes and the cycle containing one vertex may be interchanged, giving twice as many combinations. For each of these, there are then induced $\binom{i}{2}$ cycles in the triplet group. In the fourth and fifth products, for two cycles of unequal lengths i and j, there are $p_i p_j$ different combinations. A factor of 2 does not occur here because the cycle of length j in each case is the one containing two vertexes. Each of these combinations gives rise to $(j-1)d(i,j)/2$ cycles of length $m(i,j)$ if $j/d(i,j)$ is odd, and $(j/2 - 1)d(i,j)$ cycles of length $m(i,j)$ plus $d(i,j)$ cycles of length $m(i,j)/2$ if $j/d(i,j)$ is even. In the sixth product, there is no restriction on equality or inequality of i, j, and k. Upon algebraic manipulation, Eq. (42) becomes

$$F_3(n,\mathrm{p}) = \prod_{i=1}^{n} f_i^{\,p_i[\frac{1}{3}\binom{i-1}{2}]+2\binom{p_i}{2}\binom{i}{2}+\binom{p_i}{3}i^3+p_{3i}} \cdot$$

$$\cdot \prod_{i<j} f_{m(i,j)}^{\,p_i p_j (p_i i + p_j j - 2) d(i,j)/2} \cdot$$

$$\cdot \prod_{i<j<k} f_{m(i,j,k)}^{\,p_i p_j p_k ijk/m(i,j,k)} \cdot$$

$$\cdot \prod_{\substack{i\neq j \\ j/d \text{ even}}} f_{m(i,j)}^{\,-p_i p_j d(i,j)/2}\; f_{m(i,j)/2}^{\,p_i p_j d(i,j)}, \qquad (43)$$

which is the complete expression for the triplet cycle function.

A number of relationships may be obtained which will simplify higher cycle functions. For n-bonds on n vertexes, the only possibilities are n unconnected vertexes or one n-bond. In each permutation of G_n there is only one cycle of length 1, hence from Eq. (39) there results

$$F_n(n,\mathrm{p}) = \mathrm{f}^{\mathrm{b}} = f_1^1 f_2^0 \ldots f_n^0 = f_1 \quad (n = 2,\ 3,\ \ldots), \qquad (44)$$

for all partition sets of n. If G_j is the symmetric group of degree n, then $\mathrm{b} = \mathrm{p}(n)$. This will be the case only if one has the symmetric group of the bonds as well as the vertexes, which occurs for $j = n-1$, since there is then one vertex not involved with each j-bond, and the numbers of j-bonds and vertexes correspond exactly. Hence

$$F_{n-1}(n,\mathrm{p}) = \mathrm{f}^{\mathrm{p}} = \prod_{i=1}^{n} f_i^{p_i} \quad (n = 3, 4, \ldots). \tag{45}$$

Similarly, corresponding to every j-bond in a graph on n vertexes, there are $n-j$ other vertexes in the graph, not connected by this j-bond. Considering this one-to-one correspondence, the number of graphs of $(n-j)$-bonds is equal to the number of graphs of j-bonds, both on n vertexes. Hence $\mathrm{G}_{n-j} = \mathrm{G}_j$, and

$$F_{n-j}(n,\mathrm{p}) = F_j(n,\mathrm{p}) \quad (j = 2, 3, \ldots, n-2). \tag{46}$$

All cycle functions for a single type of bond to and including $n = 5$ may therefore be obtained by using only the pair cycle function, Eq. (40), with Eqs. (44) through (46), and those for $n = 6$ and $n = 7$ using in addition only the triplet cycle function, Eq. (43). Since the nth virial coefficient depends only on graphs on n vertexes, the above cycle functions are sufficient to obtain in principle the combinatorial factors for all the virial coefficients to and including the seventh, including all nonadditivity effects. In Table 7 are given the cycle functions for 2- and 3-bonds on $n \leqq 7$ vertexes, for all partitions of n.

Since cycles of different lengths and different types of bonds are independent of each other, by the product property the cycle function for any combination of bonds on the same vertexes is the product of the cycle functions of the

TABLE 7.—Cycle functions for nonlinear graphs on *n* vertexes.

n	partition	p(n)	$F_2(n,\text{p})$	$F_3(n,\text{p})$
2	1,1	2,0	f_1	--
	2	0,1	f_1	--
3	1,1,1	3,0,0	f_1^3	f_1
	1,2	1,1,0	f_1f_2	f_1
	3	0,0,1	f_3	f_1
4	1,1,1,1	4,0,0,0	f_1^6	f_1^4
	1,1,2	2,1,0,0	$f_1^2f_2^2$	$f_1^2f_2$
	2,2	0,2,0,0	$f_1^2f_2^2$	f_2^2
	1,3	1,0,1,0	f_3^2	f_1f_3
	4	0,0,0,1	f_2f_4	f_4
5	1,1,1,1,1	5,0,0,0,0	f_1^{10}	f_1^{10}
	1,1,1,2	3,1,0,0,0	$f_1^4f_2^3$	$f_1^4f_2^3$
	1,2,2	1,2,0,0,0	$f_1^2f_2^4$	$f_1^2f_2^4$
	1,1,3	2,0,1,0,0	$f_1f_3^3$	$f_1f_3^3$
	2,3	0,1,1,0,0	$f_1f_3f_6$	$f_1f_3f_6$
	1,4	1,0,0,1,0	$f_2f_4^2$	$f_2f_4^2$
	5	0,0,0,0,1	f_5^2	f_5^2

TABLE 7—Continued

n	partition	p(n)	$F_2(n,\text{p})$	$F_3(n,\text{p})$
6	1,1,1,1,1,1	6,0,0,0,0,0	f_1^{15}	f_1^{20}
	1,1,1,1,2	4,1,0,0,0,0	$f_1^7 f_2^4$	$f_1^8 f_2^6$
	1,1,2,2	2,2,0,0,0,0	$f_1^3 f_2^6$	$f_1^4 f_2^8$
	1,1,1,3	3,0,1,0,0,0	$f_1^3 f_3^4$	$f_1^2 f_3^6$
	2,2,2	0,3,0,0,0,0	$f_1^3 f_2^6$	f_2^{10}
	1,2,3	1,1,1,0,0,0	$f_1 f_2 f_3^2 f_6$	$f_1^2 f_3^2 f_6^2$
	1,1,4	2,0,0,1,0,0	$f_1 f_2 f_4^3$	f_4^5
	3,3	0,0,2,0,0,0	f_3^5	$f_1^2 f_3^6$
	2,4	0,1,0,1,0,0	$f_1 f_2 f_4^3$	$f_2^2 f_4^4$
	1,5	1,0,0,0,1,0	f_5^3	f_5^4
	6	0,0,0,0,0,1	$f_3 f_6^2$	$f_2 f_6^3$

TABLE 7—Continued

n	partition	p(n)	$F_2(n,\mathrm{p})$	$F_3(n,\mathrm{p})$
7	1,1,1,1,1,1,1	7,0,0,0,0,0,0	f_1^{21}	f_1^{35}
	1,1,1,1,1,2	5,1,0,0,0,0,0	$f_1^{11}f_2^5$	$f_1^{15}f_2^{10}$
	1,1,1,2,2	3,2,0,0,0,0,0	$f_1^5f_2^8$	$f_1^7f_2^{14}$
	1,1,1,1,3	4,0,1,0,0,0,0	$f_1^6f_3^5$	$f_1^5f_3^{10}$
	1,2,2,2	1,3,0,0,0,0,0	$f_1^3f_2^9$	$f_1^3f_2^{16}$
	1,1,2,3	2,1,1,0,0,0,0	$f_1^2f_2^2f_3^3f_6$	$f_1^3f_2f_3^4f_6^3$
	1,1,1,4	3,0,0,1,0,0,0	$f_1^3f_2f_4^4$	$f_1f_2^3f_4^7$
	2,2,3	0,2,1,0,0,0,0	$f_1^2f_2^2f_3f_6^2$	$f_1f_2^2f_3^2f_6^4$
	1,3,3	1,0,2,0,0,0,0	f_3^7	$f_1^2f_3^{11}$
	1,2,4	1,1,0,1,0,0,0	$f_1f_2^2f_4^4$	$f_1f_2^3f_4^7$
	1,1,5	2,0,0,0,1,0,0	$f_1f_5^4$	f_5^7
	3,4	0,0,1,1,0,0,0	$f_2f_3f_4f_{12}$	$f_1f_4f_6f_{12}^2$
	2,5	0,1,0,0,1,0,0	$f_1f_5^2f_{10}$	$f_5^3f_{10}^2$
	1,6	1,0,0,0,0,1,0	$f_3f_6^3$	$f_2f_3f_6^5$
	7	0,0,0,0,0,0,1	f_7^3	f_7^5

different types of bonds in the combination:

$$F_{\mathrm{j}}(n,\mathrm{p}) = \prod_{j} F_{j}(n,\mathrm{p}). \qquad (47)$$

In performing calculations, the arbitrary functions f_i in cycle functions for different types of bonds must be kept distinct.

5. The Cycle Index

From Eq. (36) and the foregoing definitions, the cycle index of the group G_j is given by

$$Z(\mathrm{G}_j) = \frac{1}{n!} \sum_{\mathrm{p}(n)} g_{\mathrm{b}} F_{j}(n,\mathrm{p}), \qquad (48)$$

where the summation is over all partition sets of n, and g_{b} is the number of elements of $\mathbf{G}_j$ which consist of b_i cycles of length i $(i = 1, 2, \ldots)$. Since for nonlinear graphs on n vertexes, the group $\mathbf{G}_j$ is of order $n!$ and is isomorphic to the symmetric group S_n, for all j, the cycle index may be written

$$Z_{\mathrm{j}}(\mathrm{S}_n) = \frac{1}{n!} \sum_{\mathrm{p}(n)} \left\{ \frac{n!}{\mathrm{p}!} \prod_{i=1}^{n} i^{-p_i} \right\} F_{j}(n,\mathrm{p}), \qquad (49)$$

where

$$\mathrm{p}! = \prod_{i=1}^{n} p_i!.$$

The coefficient in braces in Eq. (49) is the number of elements of S_n consisting of p_i cycles of length $i = 1, 2, \ldots, n$, and is the same as the number of elements of $\mathbf{G}_j$ consisting of b_i cycles of length i.(41,47)

For graphs on $n \leqq 7$ vertexes, containing only a single type of bond, the cycle indexes can be obtained from Eq. (49) and Table 7, and are given in Table 8. For combinations of types of bonds on the same vertexes, the product of the respective cycle functions is taken in each term. For example,

$$Z_{2,3}(S_4) = \frac{1}{4!}(f_1^6 g_1^4 + 6f_1^2 f_2^2 g_1^2 g_2 + 3f_1^2 f_2^2 g_2^2 \\ + 8f_3^2 g_1 g_3 + 6f_2 f_4 g_4),$$

where the f_i corresponding to 3-bonds have been replaced by g_i, to keep the independent arbitrary functions distinct.

6. Theorem for Counting Nonlinear Graphs

With the above conventions, the generating function for nonlinear graphs on n indistinguishable vertexes is expressed as a generalization of Pólya's theorem. The proof presented below generally follows that given for Pólya's theorem by Uhlenbeck and Ford.(47)

TABLE 8.—Cycle indexes $Z_j(S_n)$ for nonlinear graphs containing only j-bonds on n vertexes.

n	j	$Z_j(S_n)$
2	2	f_1
3	2	$\frac{1}{3!}(f_1^3 + 3f_1f_2 + 2f_3)$
	3	f_1
4	2	$\frac{1}{4!}(f_1^6 + 9f_1^2f_2^2 + 8f_3^2 + 6f_2f_4)$
	3	$\frac{1}{4!}(f_1^4 + 6f_1^2f_2 + 3f_2^2 + 8f_1f_3 + 6f_4)$
	4	f_1
5	2	$\frac{1}{5!}(f_1^{10} + 10f_1^4f_2^3 + 15f_1^2f_2^4 + 20f_1f_3^3 + 20f_1f_3f_6$ $+ 30f_2f_4^2 + 24f_5^2)$
	3	Same as $Z_2(S_5)$
	4	$\frac{1}{5!}(f_1^5 + 10f_1^3f_2 + 15f_1f_2^2 + 20f_1^2f_3 + 20f_2f_3$ $+ 30f_1f_4 + 24f_5)$
	5	f_1
6	2	$\frac{1}{6!}(f_1^{15} + 15f_1^7f_2^4 + 60f_1^3f_2^6 + 40f_1^3f_3^4 + 120f_1f_2f_3^2f_6$ $+ 180f_1f_2f_4^3 + 40f_3^5 + 144f_5^3 + 120f_3f_6^2)$
	3	$\frac{1}{6!}(f_1^{20} + 15f_1^8f_2^6 + 45f_1^4f_2^8 + 80f_1^2f_3^6 + 15f_2^{10}$ $+ 120f_1^2f_3^2f_6^2 + 90f_4^5 + 90f_2^2f_4^4 + 144f_5^4 + 120f_2f_6^3)$
	4	Same as $Z_2(S_6)$

TABLE 8—Continued

n	j	$Z_j(S_n)$
6	5	$\frac{1}{6!}(f_1^6 + 15f_1^4f_2 + 45f_1^2f_2^2 + 40f_1^3f_3 + 15f_2^3$ $+ 120f_1f_2f_3 + 90f_1^2f_4 + 40f_3^2 + 90f_2f_4$ $+ 144f_1f_5 + 120f_6)$
	6	f_1
7	2	$\frac{1}{7!}(f_1^{21} + 21f_1^{11}f_2^5 + 105f_1^5f_2^8 + 70f_1^6f_3^5 + 105f_1^3f_2^9$ $+ 420f_1^2f_2^2f_3^3f_6 + 210f_1^3f_2f_4^4 + 210f_1^2f_2^2f_3f_6^2$ $+ 280f_3^7 + 630f_1f_2^2f_4^4 + 504f_1f_5^4$ $+ 420f_2f_3f_4f_{12} + 504f_1f_5^2f_{10} + 840f_3f_6^3 + 720f_7^3)$
	3	$\frac{1}{7!}(f_1^{35} + 21f_1^{15}f_2^{10} + 105f_1^7f_2^{14} + 70f_1^5f_3^{10}$ $+ 105f_1^3f_2^{16} + 420f_1^3f_2f_3^4f_6^3 + 840f_1f_2^3f_4^7$ $+ 210f_1f_2^2f_3^2f_6^4 + 280f_1^2f_3^{11} + 504f_5^7$ $+ 420f_1f_4f_6f_{12}^2 + 504f_5^3f_{10}^2 + 840f_2f_3f_6^5 + 720f_7^5)$
	4	Same as $Z_3(S_7)$
	5	Same as $Z_2(S_7)$
	6	$\frac{1}{7!}(f_1^7 + 21f_1^5f_2 + 105f_1^3f_2^2 + 70f_1^4f_3 + 105f_1f_2^3$ $+ 420f_1^2f_2f_3 + 210f_1^3f_4 + 210f_2^2f_3 + 280f_1f_3^2$ $+ 630f_1f_2f_4 + 504f_1^2f_5 + 420f_3f_4 + 504f_2f_5$ $+ 840f_1f_6 + 720f_7)$
	7	f_1

Theorem: $\mu_n = Z_{\mathrm{j}}\{S_n;\ f(y_j)\}$, (50)

where $Z_{\mathrm{j}}\{S_n;\ f(y_j)\}$ is obtained from $Z_{\mathrm{j}}(S_n)$ by replacing f_i for j-bonds by $1+y_j^i$ in $F_j(n,\mathrm{p})$, for all j in j.

Proof: The figure generating function for j-bonds is given by Eq. (37). There are B_j points on which j-bonds may be placed, and $\mathbf{G}_j$ is a (given) group of permutations of these points. Repetitions of the figures are allowed, but only one figure may be placed on each point. A configuration is a nonlinear graph on n vertexes. Two graphs are topologically equivalent if there is a permutation of the vertexes taking one into the other. The content of a configuration is the sum of the contents of its figures, but the contents for different types of bonds are not added together. Hence a "content set" which is equal to the bond set $\mathrm{k}(n)$, from Eq. (13), may be defined. Then $\mu(n,\mathrm{k})$ is the number of inequivalent configurations of content $\mathrm{k}(n)$ on n vertexes, and the configuration generating function is μ_n, defined by Eq. (35).

Consider a particular permutation P of the symmetric group S_n of n vertexes, and let $X_{\mathrm{k}}(P)$ be the number of configurations of content $\mathrm{k}(n)$ which are equivalent under P. Two equivalent configurations consist of the same set of figures. Furthermore, if in P there is a cycle of length i, then all the figures of each type of bond corresponding to the points in the cycle must be the same. Since the permutations are over the n vertexes of graphs on n vertexes, the group of permutations of the points in the cycle is isomorphic to S_n. The figures corresponding to different cycles may be different and can be chosen independently. When the content of a figure of j-bonds in a cycle of length i is k_j, then the content of all such figures together in the cycle is ik_j, since they are all the same. Their generating function will be therefore

$$f(y_j^i) = 1 + y_j^i.$$

Since the different types of bonds present are independent, the product property applies to the j present. Hence

$$\sum_{\mathrm{k}(n)} X_{\mathrm{k}}(P)\mathrm{y}^{\mathrm{k}(n)} = \prod_j \prod_{i=1}^{n} \{f(y_j^i)\}^{b_i}$$

$$= \prod_j \prod_{i=1}^{n} (1 + y_j^i)^{b_i}. \qquad (51)$$

Consider all configurations of content $\mathrm{k}(n)$ and let Q be one of these. Find all permutations of the vertexes which transform Q into configurations which are identical to Q. The identity is one such permutation, and if there are others, they form a subgroup H of S_n, say of order h. Divide S_n into the cosets(18) of the subgroup H:

$$S_n = H + P_1 H + P_2 H + \dots + P_r H,$$

with

$$r = n!/h - 1.$$

All permutations in the coset P_iH will change the configuration Q into an equivalent configuration Q_i, which is transformed into itself by a group of order h. If therefore all configurations of content k(n) are ordered into groups which are equivalent by the elements of S_n, that is, if the sum

$$\sum_{i=1}^{n!} X_{\mathrm{k}}(P_i)$$

is formed, then each configuration which is equivalent to Q will occur in h terms of this sum. Since each configuration Q gives rise to $n!/h$ equivalent configurations, including Q itself, each set of configurations which are equivalent to a particular one contributes $(n!/h)h = n!$ terms to the sum. Therefore

$$n!\mu(n,\mathrm{k}) = \sum_{P\in S_n} X_{\mathrm{k}}(P) y^{\mathrm{k}(n)},$$

and from Eqs. (35) and (51),

$$\begin{aligned}\mu_n &= \frac{1}{n!}\sum_{P}\sum_{\mathrm{k}(n)} \chi_{\mathrm{k}}(P)\mathrm{y}^{\mathrm{k}(n)} \\ &= \frac{1}{n!}\sum_{P}\prod_{j}\prod_{i=1}^{n}(1+y_j^i)^{b_i} \\ &= \frac{1}{n!}\sum_{\mathrm{p}(n)}\left\{\frac{n!}{\mathrm{p}!}\prod_{i=1}^{n} i^{-p_i}\right\}\prod_{j}\prod_{i=1}^{n}(1+y_j^i)^{b_i}, \quad (52)\end{aligned}$$

since the summation over P is the sum over all permutations of the n vertexes, and corresponds to the sum over all partition sets of n, with the appropriate coefficients, as in Eq. (49). Hence on substituting $1+y_j^i$ for f_i in the $F_j(n,\mathrm{p})$ in Eq. (49), Eq. (50) is obtained from Eq. (52), which establishes the theorem.

With μ_n obtained from Eq. (50), then $\mu(n,\mathrm{k})$ is the coefficient of $\mathrm{y}^{\mathrm{k}(n)}$ in μ_n, from Eq. (35). Values of $\mu(n,\mathrm{k})$ for graphs of 2- and 3-bonds on $n = 3$, 4, and 5 indistinguishable vertexes are given in Tables 2, 3, and 4, respectively, under $\mu(\mathrm{n},\mathrm{k})$ for $\mathrm{n} = n$.

7. Counting of Trees

A nonlinear graph consists of $t \geq 1$ trees. Any permutation of the trees does not change the graph. Since graphs with different numbers of trees are independent, write

$$\mu'(x,y) = \gamma'(x,y) + \sum_{t=2}^{\infty} \mu_t'(x,y), \tag{53}$$

where $\mu_t'(x,y)$ is the generating function for nonlinear graphs consisting of exactly t trees. For graphs consisting of only one tree, it follows that

$$\mu_1'(x,y) = \gamma'(x,y). \tag{54}$$

To obtain the number of trees on indistinguishable vertexes, $\gamma(n,\mathrm{k})$, Pólya's theorem may be used in a special form in which the "figures" are trees on indistinguishable vertexes, and the figure generating function is a *known function*, even though the coefficients $\gamma(n,\mathrm{k})$ are yet to be determined. The group of Pólya's theorem is the symmetric group of degree t, with cycle index

$$Z(S_t) = \frac{1}{t!} \sum_{\mathrm{p}(t)} \left\{ \frac{t!}{\mathrm{p}!} \prod_{i=1}^{t} i^{-p_i} \right\} \mathrm{f}^{\mathrm{p}}. \tag{55}$$

Then for nonlinear graphs consisting of t trees, Pólya's theorem yields

$$\mu_t'(x,y) = Z\{S_t;\ \gamma'(x,y)\} \quad (t = 1,\ 2,\ \ldots),$$

where $Z\{S_t;\ \gamma'(x,y)\}$ is the function obtained from $Z(S_t)$ by substituting $\gamma'(x^i,y^i)$ for f_i in Eq. (55). Hence from Eqs. (53) and (54) there results

$$\mu'(x,y) = \sum_{t=1}^{\infty} Z\{S_t;\ \gamma'(x,y)\}. \tag{56}$$

Substitution of Eq. (55) into Eq. (56) yields

$$\begin{aligned}\mu'(x,y) &= \sum_{t=1}^{\infty} \frac{1}{t!} \sum_{p(t)} \left\{\frac{t!}{p!} \prod_{i=1}^{t} i^{-p_i}\right\} \prod_{i=1}^{t} \left\{\gamma'(x^i,y^i)\right\}^{p_i} \\ &= \sum_{t=1}^{\infty} \sum_{p(t)} \prod_{i=1}^{t} \frac{1}{p_i!}\left\{\frac{\gamma'(x^i,y^i)}{i}\right\}^{p_i}. \end{aligned} \tag{57}$$

Since t ranges from 1 to ∞, so does i, and p_i ranges from 0 to ∞. Then every product of the form given in Eq. (57) will occur in

$$\prod_{i=1}^{\infty} \sum_{p_i=0}^{\infty} \frac{1}{p_i!}\left\{\frac{\gamma'(x^i,y^i)}{i}\right\}^{p_i}. \tag{58}$$

Conversely, this product can be written in the form

$$\prod_{i=1}^{\infty} (1 + T_i + T_i^2 + T_i^3 + \ldots),$$

and every term resulting from the expansion of the product is different and is of the form

$$\prod_{i=1}^{\infty} T_i^{p_i},$$

where only a finite number ($\leqq t$) of the p_i are not zero. Hence the summation in Eq. (57) is over all partition sets of all positive integers, and every term in (58) is represented by a term in Eq. (57), except for the unity obtained for all $p_i = 0$, since the summation over the partition sets of t does not allow all $p_i = 0$ simultaneously in any term. Therefore there must be a one-to-one correspondence between the right side of Eq. (57) and (58), except for the added unity. Hence Eq. (57) becomes

$$\mu'(x,y) = \prod_{i=1}^{\infty} \sum_{p_i=0}^{\infty} \frac{1}{p_i!}\left\{\frac{\gamma'(x^i,y^i)}{i}\right\}^{p_i} - 1$$

$$= \prod_{i=1}^{\infty} \exp\left\{\frac{\gamma'(x^i,y^i)}{i}\right\} - 1,$$

or

$$\sum_{i=1}^{\infty} \frac{\gamma'(x^i,y^i)}{i} = \ln\{1+\mu'(x,y)\}. \tag{59}$$

This equation is analogous to the corresponding result(19,47) for linear graphs. The logarithm may be expanded in the same manner as in Appendix A for Eq. (27). Then successive relationships between the generating functions for graphs and trees on indistinguishable vertexes may be obtained.

Write

$$\begin{aligned}\gamma'(x^i,y^i) &= \sum_{n=1}^{\infty} x^{in}\gamma_{n;i} \\ &= \sum_{n=1}^{\infty} \sum_{k(n)} \gamma(n,k)x^{in}y^{ik(n)}.\end{aligned} \tag{60}$$

Then, from Eqs. (35) and (59), it follows that

$$\sum_{i=1}^{\infty}\frac{1}{i}\sum_{n=1}^{\infty}x^{in}\gamma_{n;i} = \sum_{m=1}^{\infty}\frac{(-1)^{m+1}}{m}\left(\sum_{n=1}^{\infty}x^{n}\mu_{n}\right)^{m}. \qquad (61)$$

Now from Eqs. (35) and (60), $\gamma_{n;1} = \gamma_n$ for all n, and $\gamma_{1;i} = \gamma(1,k) = 1$ for all i. Hence Eq. (61) becomes

$$\sum_{n=1}^{\infty}x^{n}(\mu_n-\gamma_n) = \sum_{i=2}^{\infty}\frac{1}{i}\left\{\sum_{n=1}^{\infty}x^{in}\gamma_{n;i} + (-1)^{i}\left(\sum_{n=1}^{\infty}x^{n}\mu_n\right)^{i}\right\}.$$

Equating coefficients of like powers of x yields

$$\mu_n - \gamma_n = \sum_{i=2}^{n}\frac{1}{i}\left\{\gamma_{n/i;i} + (-1)^{i}\sum_{c(n)}\prod_{j=1}^{i}\mu_{n_j}\right\},$$

where the second summation is over all 2^{n-1} compositions of n into i positive integers n_j ($j = 1, 2, \ldots, i$). If i does not divide evenly into n, take $\gamma_{n/i;i} = 0$. Since $\mu_1 = \mu(1,k) = 1$, there results

$$\mu_1 - \gamma_1 = 0,$$

$$\mu_2 - \gamma_2 = 1,$$

$$\mu_3 - \gamma_3 = \mu_2,$$

$$\mu_4 - \gamma_4 = \mu_3 + \frac{1}{2}(\gamma_{2;2}+\mu_2^2+1) - \mu_2,$$

$$\mu_5 - \gamma_5 = \mu_4 + (\mu_2-1)(\mu_3-\mu_2),$$

$$\begin{aligned}\mu_6 - \gamma_6 = \mu_5 &+ (\mu_2+1)\mu_4 + \frac{1}{2}(\gamma_{3;2}+\mu_3^2+3\mu_2^2) + (2\mu_2+1)\mu_3 \\ &+ \frac{1}{3}(\gamma_{2;3}-\mu_2^3+1) - \mu_2,\end{aligned}$$

$$\mu_7 - \gamma_7 = \mu_6 + (\mu_2-1)\mu_5 + \mu_3\mu_4 - (2\mu_2-1)(\mu_2+\mu_4)$$
$$+ (3\mu_2-\mu_2^2-1)\mu_3 - \mu_3^2.$$

From these successive equations, with μ_n known, the γ_n, and hence the coefficients $\gamma(n,\mathrm{k})$, can be obtained. Values of $\gamma(n,\mathrm{k})$, for trees of 2- and 3-bonds on n = 3, 4, and 5 indistinguishable vertexes, are given in Tables 2, 3, and 5, respectively, under $\gamma(\mathrm{n},\mathrm{k})$ for $\mathrm{n} = n$.

E. GRAPHS AND TREES FOR MULTICOMPONENT SYSTEMS

In this section the techniques for counting graphs and trees are generalized for application to multicomponent systems, and in Sec. F these results will be applied to the counting of stars, which enters the combinatorial problem for the virial coefficients from Eqs. (5) and (7).

1. Generating Functions

The numbers $N(n,\mathrm{k})$, $C(n,\mathrm{k})$, and $S(n,\mathrm{k})$, for graphs, trees, and stars, respectively, apply to multicomponent systems just as to systems of a single component, since the vertexes are already distinguishable. If the vertexes are indistinguishable but restricted to n (n_i of species $i = 1, 2, \ldots, s$), the generating functions (35) may be generalized to take into account the different species involved:

$$\begin{aligned}
\mu'(\mathrm{x},\mathrm{y}) &= \sum_{\mathrm{n}} \mu_{\mathrm{n}}\mathrm{x}^{\mathrm{n}} = \sum_{\mathrm{n}} \sum_{\mathrm{k(n)}} \mu(\mathrm{n},\mathrm{k})\mathrm{x}^{\mathrm{n}}\mathrm{y}^{\mathrm{k(n)}}, \\
\gamma'(\mathrm{x},\mathrm{y}) &= \sum_{\mathrm{n}} \gamma_{\mathrm{n}}\mathrm{x}^{\mathrm{n}} = \sum_{\mathrm{n}} \sum_{\mathrm{k(n)}} \gamma(\mathrm{n},\mathrm{k})\mathrm{x}^{\mathrm{n}}\mathrm{y}^{\mathrm{k(n)}}, \qquad (62) \\
\sigma'(\mathrm{x},\mathrm{y}) &= \sum_{\mathrm{n}} \sigma_{\mathrm{n}}\mathrm{x}^{\mathrm{n}} = \sum_{\mathrm{n}} \sum_{\mathrm{k(n)}} \sigma(\mathrm{n},\mathrm{k})\mathrm{x}^{\mathrm{n}}\mathrm{y}^{\mathrm{k(n)}}.
\end{aligned}$$

The group of Pólya's theorem for j-bonds is now the group of permutations of the $\binom{n}{j}$ j's of vertexes induced by the direct product of the s symmetric groups(18,19) of n_i vertexes ($i = 1, 2, \ldots, s$). Let

$$S_n = S_{n_1} \times S_{n_2} \times \cdots \times S_{n_s} \tag{63}$$

denote this direct-product group.

2. Cycle Functions

First determine the cycle function for j-bonds in S_n, which may be indicated symbolically by $F_j(n)$. This cycle function is not equal to the product of the cycle functions for the separate n_i, since the j vertexes under consideration are not restricted to a single n_i, but may be distributed among the symmetric groups in the direct product. One must distinguish the cases in which the j vertexes all occur in the same n_i, and also the cases in which the j vertexes are distributed between 2, 3, ..., k different n_i, where k is the minimum of j and s.

Consider the pair cycle function $F_2(n)$ for the direct product of two symmetric groups, so that $n = n_1, n_2$. If the two vertexes of the pair both occur in S_{n_1}, either in the same cycle or in different cycles, the contribution to $F_2(n)$ will be $F_2(n_1, p)$, where $p = p(n_1)$. Similarly for S_{n_2}, the result is $F_2(n_2, p)$, where $p = p(n_2)$. There is no contribution for $n_1 < 2$ or $n_2 < 2$, and one may formally take $F_j(n, p) = 1$ for $n < j$, from Eq. (39), since then all the exponents $b_i = 0$. If one vertex occurs in S_{n_1} and the other in S_{n_2}, they are in different cycles. If the cycle in S_{n_1} is of length i and the cycle in S_{n_2} is of length j, their permutations will induce $d(i,j)$ cycles of length $m(i,j)$ in the pair group, whether or not $i = j$. Hence the contribution in this case will be

$$\prod_{i=1}^{n_1}\prod_{j=1}^{n_2} f_{m(i,j)}^{p_i(n_1)p_j(n_2)d(i,j)} .$$

The $p_i(n_1)p_j(n_2)$ occurs as an exponent since there are $p_i(n_1)$ different cycles of length i in S_{n_1} and $p_j(n_2)$ cycles of length j in S_{n_2}, independently, and hence $p_i(n_1)p_j(n_2)$ different combinations of the two.

Combining these results, one obtains

$$F_2(n_1,n_2) = F_2(n_1,\mathrm{p})F_2(n_2,\mathrm{p}) \prod_{i=1}^{n_1}\prod_{j=1}^{n_2} f_{m(i,j)}^{p_i(n_1)p_j(n_2)d(i,j)} .$$

By similar reasoning, consideration of all the possible cases results in

$$F_2(n_1,n_2,n_3) = F_2(n_1,\mathrm{p})F_2(n_2,\mathrm{p})F_2(n_3,\mathrm{p}) \cdot$$
$$\cdot \prod_{i=1}^{n_1}\prod_{j=1}^{n_2}\prod_{k=1}^{n_3} \left\{ f_{m(i,j)}^{p_i(n_1)p_j(n_2)d(i,j)} \, f_{m(i,k)}^{p_i(n_1)p_k(n_3)d(i,k)} \cdot \right.$$
$$\left. \cdot f_{m(j,k)}^{p_j(n_2)p_k(n_3)d(j,k)} \right\},$$

and, in general,

$$F_2(\mathrm{n}) = \left\{ \prod_{i=1}^{s} F_2(n_i,\mathrm{p}) \right\} \prod_{i<j} \prod_{r=1}^{n_i} \prod_{s=1}^{n_j} f_{m(r,s)}^{p_r(n_i)p_s(n_j)d(r,s)} . \qquad (64)$$

It is possible to obtain $F_2(1,1,\ldots,1,n)$ for $s = m + 1$ (m elements of unity) and $n \geqq 2$ either from Eq. (64) or by replacing p_1 by $p_1 + m$ in $F_2(n,\mathrm{p})$. In either case the result is

$$F_2(\underbrace{1,1,\ldots,1}_{m \text{ elements}},n) = f_1^{\binom{m}{2}} \mathrm{f}^{m\mathrm{p}(n)} F_2(n,\mathrm{p}) .$$

Correspondence of results by the two methods is a strong check on the calculations and on the validity of Eq. (64).

For the triplet cycle function $F_3(n_1,n_2)$, the following cases must be considered:

(1) If all three vertexes occur in S_{n_1} or in S_{n_2}, the result is $F_3(n_1,\mathrm{p})$ or $F_3(n_2,\mathrm{p})$, respectively.

(2) If two vertexes are in S_{n_1} and one in S_{n_2}, the two vertexes in S_{n_1} may be in either one cycle or in two different cycles. Hence this case will be equivalent to cases (2) and (3) of the derivation of Eq. (43). Upon algebraic manipulation, the result may therefore be expressed as

$$
\begin{aligned}
F_3(n_1,n_2;2,1) = \prod_{i=1}^{n_2} \Bigg\{ & \prod_{j=1}^{n_1} f_{m(i,j)}^{p_i(n_2)p_j(n_1)\{jp_j(n_1)-1\}d(i,j)/2} \cdot \\
& \cdot \prod_{j<k} f_{m(i,j,k)}^{p_i(n_2)p_j(n_1)p_k(n_1)ijk/m(i,j,k)} \cdot \\
& \cdot \prod_{j/d \text{ even}} \left(f_{m(i,j)}^{-p_j(n_1)p_i(n_2)d(i,j)} \; f_{m(i,j)/2}^{p_j(n_1)p_i(n_2)d(i,j)} \right) \Bigg\} .
\end{aligned}
\tag{65}
$$

(3) If two vertexes are in S_{n_2} and one in S_{n_1}, the result, $F_3(n_1,n_2;1,2)$, is the same as the right side of Eq. (65), with n_1 and n_2 interchanged wherever they appear. On combining these results, one obtains

$$
F_3(n_1,n_2) = F_3(n_1,\mathrm{p})F_3(n_2,\mathrm{p})F_3(n_1,n_2;2,1)F_3(n_1,n_2;1,2) .
$$

Similarly, it follows that

$$
\begin{aligned}
F_3(n_1,n_2,n_3) = {} & F_3(n_1,\mathrm{p})F_3(n_2,\mathrm{p})F_3(n_3,\mathrm{p})F_3(n_1,n_2;2,1) \cdot \\
& \cdot F_3(n_1,n_2;1,2)F_3(n_1,n_3;2,1)F_3(n_1,n_3;1,2) \cdot \\
& \cdot F_3(n_2,n_3;2,1)F_3(n_2,n_3;1,2) \cdot \\
& \cdot \prod_{i=1}^{n_1} \prod_{j=1}^{n_2} \prod_{k=1}^{n_3} f_{m(i,j,k)}^{p_i(n_1)p_j(n_2)p_k(n_3)ijk/m(i,j,k)} ,
\end{aligned}
$$

and, in general,

$$
\begin{aligned}
F_3(\mathrm{n}) = {} & \left\{ \prod_{i=1}^{s} F_3(n_i,\mathrm{p}) \right\} \left\{ \prod_{i<j} F_3(n_i,n_j;2,1)F_3(n_i,n_j;1,2) \right\} \cdot \\
& \cdot \prod_{i<j<k} \prod_{r=1}^{n_i} \prod_{s=1}^{n_j} \prod_{t=1}^{n_k} f_{m(r,s,t)}^{p_r(n_i)p_s(n_j)p_t(n_k)rst/m(r,s,t)} .
\end{aligned}
\tag{66}
$$

Just as for the pair cycle function, it is possible to obtain $F_3(1,1,\ldots,1,n)$ for m elements of unity and $n \geqq 3$ either from the general equation, Eq. (66), or by replacing p_1 by p_1+m in $F_3(n,\mathrm{p})$. The result, by either method, is

$$F_3(\underbrace{1,1,\ldots,1,n}_{m \text{ elements}}) = f_1^{\binom{m}{3}} \mathrm{f}^{\binom{m}{2}\mathrm{p}(n)} \{F_2(n,\mathrm{p})\}^m F_3(n,\mathrm{p}),$$

and provides a check on calculations and on the validity of Eqs. (43) and (66).

Since different types of bonds are independent of each other in graphs for multicomponent systems as well as for indistinguishable vertexes, Eq. (47) can be extended to multicomponent systems:

$$F_\mathrm{j}(\mathrm{n}) = \prod_j F_j(\mathrm{n}).$$

Values of pair and triplet cycle functions for all partitions of $n \leqq 5$ are given in Table 9.

3. Cycle Indexes

With the cycle functions for graphs corresponding to multicomponent systems obtained by Eqs. (64) and (66), the cycle index for j-bonds on n vertexes is given by

$$z_\mathrm{j}(\mathrm{S_n}) = \frac{1}{\mathrm{n}!}\sum_{\mathrm{p(n)}} \prod_{i=1}^{s} \left\{\frac{n_i!}{\mathrm{p}(n_i)!} \prod_{k=1}^{n_i} k^{-p_k(n_i)}\right\} F_\mathrm{j}(\mathrm{n}), \qquad (67)$$

TABLE 9.—Pair and triplet cycle functions for all partitions of $n \leqq 5$. $F_2(n,\mathrm{p})$ is given by Eq. (40). Asterisks indicate distinct partition sets, corresponding to distinct species.

n	$\mathbf{n}$	$F_2(\mathbf{n})$	$F_3(\mathbf{n})$
1	1	1	1
2	1,1	f_1	1
	2	f_1	1
3	1,1,1	f_1^3	f_1
	1,2	$f_1\mathbf{f}^{\mathbf{p}(2)}$	f_1
	3	$\mathbf{f}^{\mathbf{p}(3)}$	f_1
4	1,1,1,1	f_1^6	f_1^4
	1,1,2	$f_1^2\mathbf{f}^{2\mathbf{p}(2)}$	$f_1^2\mathbf{f}^{\mathbf{p}(2)}$
	2,2*	$f_1^2 \prod_{i=1}^{2}\prod_{j=1}^{2} f_{m(i,j)}^{p_i(2)p_j(2^*)d(i,j)}$	$\mathbf{f}^{\mathbf{p}(2)}\mathbf{f}^{\mathbf{p}(2^*)}$
	1,3	$\mathbf{f}^{2\mathbf{p}(3)}$	$f_1\mathbf{f}^{\mathbf{p}(3)}$
	4	$F_2(4,\mathrm{p})$	$\mathbf{f}^{\mathbf{p}(4)}$

TABLE 9—Continued

<table>
<tr><th>n</th><th>$\mathbf{n}$</th><th>$F_2(\mathbf{n})$</th><th>$F_3(\mathbf{n})$</th></tr>
<tr><td>5</td><td>1,1,1,1,1</td><td colspan="2">f_1^{10}</td></tr>
<tr><td></td><td>1,1,1,2</td><td colspan="2">$f_1^4\mathbf{f}^{3\mathbf{p}(2)}$</td></tr>
<tr><td></td><td>1,2,2*</td><td colspan="2">$f_1\mathbf{f}^{\mathbf{p}(2)}\mathbf{f}^{\mathbf{p}(2^*)}\prod_{i=1}^{2}\prod_{j=1}^{2}f_{m(i,j)}^{p_i(2)p_j(2^*)d(i,j)}$</td></tr>
<tr><td></td><td>1,1,3</td><td colspan="2">$f_1\mathbf{f}^{3\mathbf{p}(3)}$</td></tr>
<tr><td></td><td>2,3</td><td colspan="2">$f_1\mathbf{f}^{\mathbf{p}(3)}\prod_{i=1}^{2}\prod_{j=1}^{3}f_{m(i,j)}^{p_i(2)p_j(3)d(i,j)}$</td></tr>
<tr><td></td><td>1,4</td><td colspan="2">$F_2(4,\mathbf{p})\mathbf{f}^{\mathbf{p}(4)}$</td></tr>
<tr><td></td><td>5</td><td colspan="2">$F_2(5,\mathbf{p})$</td></tr>
</table>

where

$$p(n) = p(n_1),\ p(n_2),\ \ldots,\ p(n_s), \qquad (68)$$

and n! follows from Eq. (12), since the order of the direct product of a number of groups is the product of the orders of the separate groups. One must sum over all partition sets of the $n_i (i = 1, 2, \ldots, s)$ separately, since the permutation types of the direct product group S_n range over all the partition sets for the various n_i independently. Rather than perform the sum over each element of the product group, it is convenient to perform it over each permutation type. The number of permutations corresponding to the partition set $p(n_i)$ in S_{n_i} is given by the quantity in braces in Eq. (67). Hence the number of permutations of type $p(n_1)p(n_2)\cdots p(n_s)$ is the product of the quantity in braces over $i = 1, 2, \ldots, s$. Cycle indexes for 2- and 3-bonds separately on n vertexes for $n \leqq 5$ are presented in Table 10. Cycle indexes for graphs containing both 2- and 3-bonds are obtained by taking the product of the separate cycle functions, keeping the arbitrary functions for 2- and 3-bonds distinct.

4. Graphs and Trees

With these conventions, instead of Eq. (50), Pólya's theorem yields, symbolically,

TABLE 10.—Cycle indexes for 2- and 3-bonds on n vertexes, for $n \leqq 5$.

n	n	$Z_2(S_n)$	$Z_3(S_n)$
1	1	1	1
2	1,1	f_1	1
	2	f_1	1
3	1,1,1	f_1^3	f_1
	1,2	$\frac{1}{2}(f_1^3 + f_1 f_2)$	f_1
	3	$\frac{1}{6}(f_1^3 + 3f_1 f_2 + 2f_3)$	f_1
4	1,1,1,1	f_1^6	f_1^4
	1,1,2	$\frac{1}{2}(f_1^6 + f_1^2 f_2^2)$	$\frac{1}{2}(f_1^4 + f_1^2 f_2)$
	2,2	$\frac{1}{4}(f_1^6 + 3f_1^2 f_2^2)$	$\frac{1}{4}(f_1^4 + 2f_1^2 f_2 + f_2^2)$
	1,3	$\frac{1}{6}(f_1^6 + 3f_1^2 f_2^2 + 2f_3^2)$	$\frac{1}{6}(f_1^4 + 3f_1^2 f_2 + 2f_1 f_3)$
	4	$\frac{1}{24}(f_1^6 + 9f_1^2 f_2^2 + 8f_3^2 + 6f_2 f_4)$	$\frac{1}{24}(f_1^4 + 6f_1^2 f_2 + 3f_2^2 + 8f_1 f_3 + 6f_4)$

TABLE 10—Continued

<table>
<tr><th>n</th><th>n</th><th>$Z_2(S_n)$</th><th>$Z_3(S_n)$</th></tr>
<tr><td>5</td><td>1,1,1,1,1</td><td colspan="2">$\underbrace{\hphantom{Z_2(S_n)\qquad\qquad Z_3(S_n)}}$
f_1^{10}</td></tr>
<tr><td></td><td>1,1,1,2</td><td colspan="2">$\frac{1}{2}(f_1^{10} + f_1^4f_2^3)$</td></tr>
<tr><td></td><td>1,2,2</td><td colspan="2">$\frac{1}{4}(f_1^{10} + 2f_1^4f_2^3 + f_1^2f_2^4)$</td></tr>
<tr><td></td><td>1,1,3</td><td colspan="2">$\frac{1}{6}(f_1^{10} + 3f_1^4f_2^3 + 2f_1f_3^3)$</td></tr>
<tr><td></td><td>2,3</td><td colspan="2">$\frac{1}{12}(f_1^{10} + 4f_1^4f_2^3 + 3f_1^2f_2^4 + 2f_1f_3^3 + 2f_1f_3f_6)$</td></tr>
<tr><td></td><td>1,4</td><td colspan="2">$\frac{1}{24}(f_1^{10} + 6f_1^4f_2^3 + 3f_1^2f_2^4 + 8f_1f_3^3 + 6f_2f_4^2)$</td></tr>
<tr><td></td><td>5</td><td colspan="2">$\frac{1}{120}(f_1^{10} + 10f_1^4f_2^3 + 15f_1^2f_2^4 + 20f_1f_3^3 + 20f_1f_3f_6 + 30f_2f_4^2 + 24f_5^2)$</td></tr>
</table>

$$\mu_n = Z_{\mathrm{j}}\{S_n;\, f(y_j)\}, \tag{69}$$

where $Z_{\mathrm{j}}\{S_n;\, f(y_j)\}$ is obtained from $Z_{\mathrm{j}}(S_n)$ by replacing f_i for j-bonds by $1+y_j^i$ in $F_{\mathrm{j}}(n)$, for all j in j. Proof of Eq. (69) is analogous to the proof given for Eq. (50). Then $\mu(n,k)$ is the coefficient of $y^{k(n)}$ in μ_n.

Likewise, Eq. (59), relating the generating functions for graphs and trees, can also be generalized for multicomponent systems:

$$\sum_{i=1}^{\infty} \frac{\gamma'(x^i,y^i)}{i} = \ln\,\{1+\mu'(x,y)\}. \tag{70}$$

Derivation of Eq. (70) parallels that of Eq. (59). With μ_n known from Eq. (69), $\mu'(x,y)$ can be obtained from Eq. (62). Then Eq. (70) can be expanded and the coefficients of like powers of x equated to obtain the numbers of trees $\gamma(n,k)$ for multicomponent systems.

Write

$$\gamma'(x^i,y^i) = \sum_n x^{in}\gamma_{n;i} = \sum_n \sum_{k(n)} \gamma(n,k)x^{in}y^{ik(n)}. \tag{71}$$

Substitution of Eqs. (62) and (71) into Eq. (70) yields

$$\sum_{i=1}^{\infty} \frac{1}{i} \sum_n x^{in}\gamma_{n;i} = \sum_{m=1}^{\infty} \frac{(-1)^{m+1}}{m}\left(\sum_n x^n\mu_n\right)^m.$$

Since $\gamma_{n;1} = \gamma_n$ for all n, and $\gamma_{1;i} = \gamma_1 = 1$ for all i, one has

$$\sum_{\mathrm{n}} x^{\mathrm{n}}(\mu_{\mathrm{n}}-\gamma_{\mathrm{n}}) = \sum_{i=2}^{\infty}\frac{1}{i}\left\{\sum_{\mathrm{n}} x^{i\mathrm{n}}\gamma_{\mathrm{n};i} + (-1)^{i+1}\left[\sum_{\mathrm{n}} x^{\mathrm{n}}\mu_{\mathrm{n}}\right]^{i}\right\}.$$

Equating coefficients of x^{n} results in

$$\mu_{\mathrm{n}} - \gamma_{\mathrm{n}} = \sum_{i=2}\frac{1}{i}\left\{\gamma_{\mathrm{n}/i;i} + (-1)^{i}\sum_{c_i(\mathrm{n})}\prod_{j=1}^{i}\mu_{\mathrm{n}_j}\right\}, \qquad (72)$$

where the second summation is over all compositions (ordered partitions) of n into i sets $\mathrm{n}_1, \mathrm{n}_2, \ldots, \mathrm{n}_i$, such that

$$\sum_{j=1}^{i}\mathrm{n}_j = \mathrm{n}.$$

If all elements of n are not divisible by i, $\gamma_{\mathrm{n}/i;i} = 0$, and if $i > n$, at least one of the sets $\mathrm{n}_1, \mathrm{n}_2, \ldots, \mathrm{n}_i$ will be empty, for which the corresponding $\mu_{\mathrm{n}_j} = 0$. With μ_{n} known, γ_{n}, and hence $\gamma(\mathrm{n},k)$, can be obtained from Eqs. (62) and (72). Numerical values of $\mu(\mathrm{n},k)$ and $\gamma(\mathrm{n},k)$ are given in Tables 2 through 6 for $n = 3$, 4, and 5. If the vertexes are all distinguishable, $\mathrm{n} = 1,1,\ldots,1$, and $\mu(\mathrm{n},k) = N(n,k)$ and $\gamma(\mathrm{n},k) = C(n,k)$, as indicated previously. On the other hand, if the vertexes are all indistinguishable, $\mathrm{n} = n$, and the results are obtained by the techniques given in Sec. D.

5. Rooted Graphs and Trees

The coefficients for rooted graphs and trees are useful for obtaining the coefficients for graphs corresponding to the integrals of radial distribution and correlation functions.(14,47) Define a rooted or multiply rooted graph as one in which one or more vertexes are fixed and designated as roots; then in the corresponding integrals, integration over the coordinates of the root vertexes is not performed. The density expansion of the radial distribution function results in integrals which correspond to doubly rooted stars containing a 2-bond between the roots, from which this 2-bond is then removed. For a pure substance, coefficients for rooted and doubly rooted graphs and trees can be obtained from Eqs. (62), (69), and (70) by taking $\mathrm{n} = 1, n-1$ and $\mathrm{n} = 2, n-2$, respectively.

Corresponding results for rooted graphs for multicomponent systems are obtained similarly. The root vertexes are considered to be of different species than the remaining vertexes. The group of Pólya's theorem is then $S_{\mathrm{n}} = S_{n_1} \times S_{n_2} \times \cdots \times S_{\mathrm{n}_r}$, where n_i ($i = 1, 2, \ldots, s$) is the number of vertexes of species i, excluding the roots, and n_r is the corresponding composition set for the roots. For singly rooted graphs, $\mathrm{n}_r = 1$. For doubly rooted graphs, $\mathrm{n}_r = 2$ if the roots are of the same species, and $S_{\mathrm{n}_r} = S_1 \times S_1$ if the roots are of different species. For example, considering binary systems on four vertexes, two of each species, the group of Pólya's theorem is $S_{\mathrm{n}} = S_2 \times S_2$ if both roots are of the same species, and $S_{\mathrm{n}} = S_1 \times S_1 \times S_1 \times S_1$ if the roots are of different species. In this latter case, since the vertexes are all distinct, the desired coefficients are those obtained in Sec. C for graphs, trees, and stars on distinguishable vertexes.

Rooted graphs consist of either one rooted tree or one rooted tree and one or more nonrooted trees. Denoting the generating functions for rooted graphs and trees by a superscript degree sign instead of a prime, and making use of the product property, it follows that

$$\begin{aligned}\mu^\circ(x,y) &= \gamma^\circ(x,y) + \gamma^\circ(x,y)\mu_1'(x,y)\\ &\quad + \gamma^\circ(x,y)\mu_2'(x,y) + \ldots,\end{aligned}$$

where $\mu_t'(x,y)$ is the generating function for graphs which consist of exactly t trees, for multicomponent systems. Then considering each variable to be a function of the same (x,y), this equation, in shortened notation, becomes

$$\mu^\circ = \gamma^\circ(1 + \mu') \tag{73}$$

from which the coefficients in the generating function γ° for rooted trees may be obtained successively. Similarly, for doubly rooted trees,

$$\begin{aligned}\mu^{\circ\circ} &= \gamma^{\circ\circ} + \gamma^{\circ\circ}\mu_1' + \gamma^{\circ\circ}\mu_2' + \ldots\\ &\quad + (\gamma^\circ)^2 + (\gamma^\circ)^2\mu_1' + (\gamma^\circ)^2\mu_2' + \ldots\\ &= \{\gamma^{\circ\circ} + (\gamma^\circ)^2\}(1 + \mu').\end{aligned}$$

Substitution from Eq. (73) yields

$$\mu^{\circ\circ} = \gamma^{\circ\circ}(1 + \mu') + \mu^{\circ}\gamma^{\circ}, \qquad (74)$$

from which the coefficients in the generating function $\gamma^{\circ\circ}$ for doubly rooted trees for multicomponent systems may be obtained. Defining μ_n°, $\mu_n^{\circ\circ}$, γ_n°, and $\gamma_n^{\circ\circ}$ for singly and doubly rooted graphs and trees on n vertexes, respectively, in the same manner as μ_n and γ_n in Eqs. (62), there results from Eqs. (73) and (74)

$$\mu_n^{\circ} - \gamma_n^{\circ} = \sum_{i+j=n} \mu_i\gamma_j^{\circ},$$

$$\mu_n^{\circ\circ} - \gamma_n^{\circ\circ} = \sum_{i+j=n} (\mu_i\gamma_j^{\circ\circ} + \mu_i^{\circ}\gamma_j^{\circ}),$$

from which, with μ_n° and $\mu_n^{\circ\circ}$ known, γ_n° and $\gamma_n^{\circ\circ}$ may be obtained successively. The coefficients in μ°, $\mu^{\circ\circ}$, γ°, and $\gamma^{\circ\circ}$ may also be obtained by linear combinations of the results for μ_n and γ_n to take into account all possible ways of distributing the roots among the species involved.

F. COUNTING OF STARS ON INDISTINGUISHABLE VERTEXES

The development of systematic procedures for counting stars on indistinguishable vertexes has been deferred to this point, since the difficulty of this problem requires for its solution many of the techniques developed in preceding sections for graphs, trees, and multicomponent systems. A number of authors(17,20,40,47) have attempted solutions to the star-counting problem for linear graphs without success. Recently Harary(21) reported that a solution to the star-counting problem for linear graphs had been obtained by R. W. Robinson. Apparently this solution has not yet been published. Harary states:(21) "His method involves the computation of sums of cycle indexes, which leads to a recurrence relation giving the cycle index sum for the automorphism groups of the non-separable graphs [stars]." There is no indication that such a technique has been extended to nonlinear graphs. It is precisely the counting of stars on indistinguishable vertexes and, for multicomponent systems, on vertexes distinguishable by species only, that is essential to the theoretical calculation of the virial coefficients, from Eqs. (5) and (7).

The numbers $\sigma(n,k)$ of stars on indistinguishable vertexes have heretofore been obtained by inspection, for linear graphs.(47) For a given n, the correct values of $\sigma(n,k)$ can be obtained from the coefficients $\gamma(n,k)$ by subtracting the numbers of topologically distinct trees containing two or more stars each on n vertexes. A systematic procedure for doing this is developed, following which a simple formula to approximate very closely the numbers of nonlinear stars on indistinguishable vertexes is presented.

1. Systematic Procedure for Counting Stars Exactly

In the same manner as for Eqs. (60) and (71), define

$$\sigma_{n;i} = \sum_{k(n)} \sigma(n,k) y^{ik(n)}. \tag{75}$$

Then $\gamma_n - \sigma_n$ is the sum of all trees consisting of two or more stars each on n vertexes. With γ_n known from Eq. (72), by determining all trees of two or more stars each, one can then obtain σ_n, and hence $\sigma(n,k)$ from Eq. (62).

To obtain $\gamma_n - \sigma_n$, employ the following procedure:

1. Obtain all partitions of $(n+m-1)$ into m elements m_i $(m, m_i = 2, 3, \ldots, n-1)$. Since these partitions are equinumerous with those of $n-1$ into $m \geq 2$ elements, unity included, the number of these partitions$^{(41)}$ is the coefficient of t^{n+m-1} in the expansion of

$$\prod_{i=1}^{m} (1 - t^i)^{-1}.$$

For $n \leqq 7$, these partitions are given in Table 11.

TABLE 11.—Partitions of $(n+m-1)$ into $m \geqq 2$ elements $m_i \geqq 2$ $(m, m_i = 2, 3, \ldots, n-1)$, for $n \leqq 7$.

n	m	partitions
1	--	--
2	--	--
3	2	2,2
4	2	2,3
	3	2,2,2
5	2	2,4; 3,3
	3	2,2,3
	4	2,2,2,2
6	2	2,5; 3,4
	3	2,2,4; 2,3,3
	4	2,2,2,3
	5	2,2,2,2,2
7	2	2,6; 3,5; 4,4
	3	2,2,5; 2,3,4; 3,3,3
	4	2,2,2,4; 2,2,3,3
	5	2,2,2,2,3
	6	2,2,2,2,2,2

2. For each partition, form all possible distributions of these m elements into topologically distinct complete trees on n indistinguishable vertexes, for which m_i corresponds to a complete star on m_i vertexes. The number of these distributions can be obtained recursively by a generating-function technique described by Ford *et al.*(12) The trees may be represented by cacti (see Sec. B.1) consisting of lines and triangles if all $m_i \leqq 3$. If one or more $m_i \geqq 4$, these trees do not correspond to cacti, since all pairs of the vertexes of quadrilaterals, pentagons, etc. are not adjacent.

3. For each complete tree on n vertexes, consider the corresponding topologically distinct complete trees on n vertexes. The number of these is obtained systematically by the theory of distributions, by a generating-function procedure developed by MacMahon(29) and described in Appendix C.

4. Consider all possible distributions of bonds to form stars on m_i vertexes, for each element m_i in each of these trees. Denote the composition set of the m_i vertexes by $\mathbf{m}_i$, where cut vertexes are considered to be of different species than the remaining vertexes.

The generating function for a tree consisting of stars whose vertexes correspond to those of complete stars which are topologically distinct in the tree, considering the composition, is the product of the generating functions of the separate stars, since they are from independent sets. The generating function for the numbers of stars corresponding to complete star m_i is σ_{m_i}, hence the generating function for all trees corresponding to a complete tree consisting of m topologically distinct complete stars of composition m_i is simply

$$\prod_{i=1}^{m} \sigma_{\mathrm{m}_i}.$$

The product is more complicated if, in the complete tree, there are two or more complete stars which are symmetric with respect to each other, including the composition. In such cases an interchange of the complete stars would yield topologically equivalent trees, and this symmetry must be taken into account. This may be done by an application of Pólya's theorem. Consider σ_{m} as the figure generating function. For k symmetric complete stars in a tree, the group of Pólya's theorem is the symmetric group S_k. Then the configuration generating function for the symmetric stars is

$$Z(S_k;\sigma_{\mathrm{m}}) = \frac{1}{k!}\sum_{\mathrm{p}(k)} \frac{k!}{\mathrm{p}(k)!} \prod_{i=1}^{k} \left(\frac{\sigma_{\mathrm{m};i}}{i}\right)^{p_i}, \tag{76}$$

which reduces to the ordinary product σ_2^k for m = 1,1 or m = 2, since there is only one star on two vertexes, regardless of composition. The first few examples of Eq. (76) are the following:

$$\begin{aligned}
Z(S_1;\sigma_m) &= \sigma_m, \\
Z(S_2;\sigma_m) &= \tfrac{1}{2}(\sigma_m^2 + \sigma_{m;2}), \\
Z(S_3;\sigma_m) &= \tfrac{1}{6}(\sigma_m^3 + 3\sigma_m\sigma_{m;2} + 2\sigma_{m;3}), \\
Z(S_4;\sigma_m) &= \tfrac{1}{24}(\sigma_m^4 + 6\sigma_m^2\sigma_{m;2} + 3\sigma_{m;2}^2 \\
&\qquad + 8\sigma_m\sigma_{m;3} + 6\sigma_{m;4}).
\end{aligned}$$

Then, since $\gamma_n - \sigma_n$ corresponds to all trees of two or more stars each on n vertexes, and by the given procedure, these are represented by the distributions of complete stars to form trees consisting of two or more complete stars on n vertexes, with each complete star on m_i vertexes representing all stars on m_i vertexes, it follows that

$$\gamma_n - \sigma_n = \sum \prod_{i=1}^{m}{}^{*} \sigma_{m_i}, \tag{77}$$

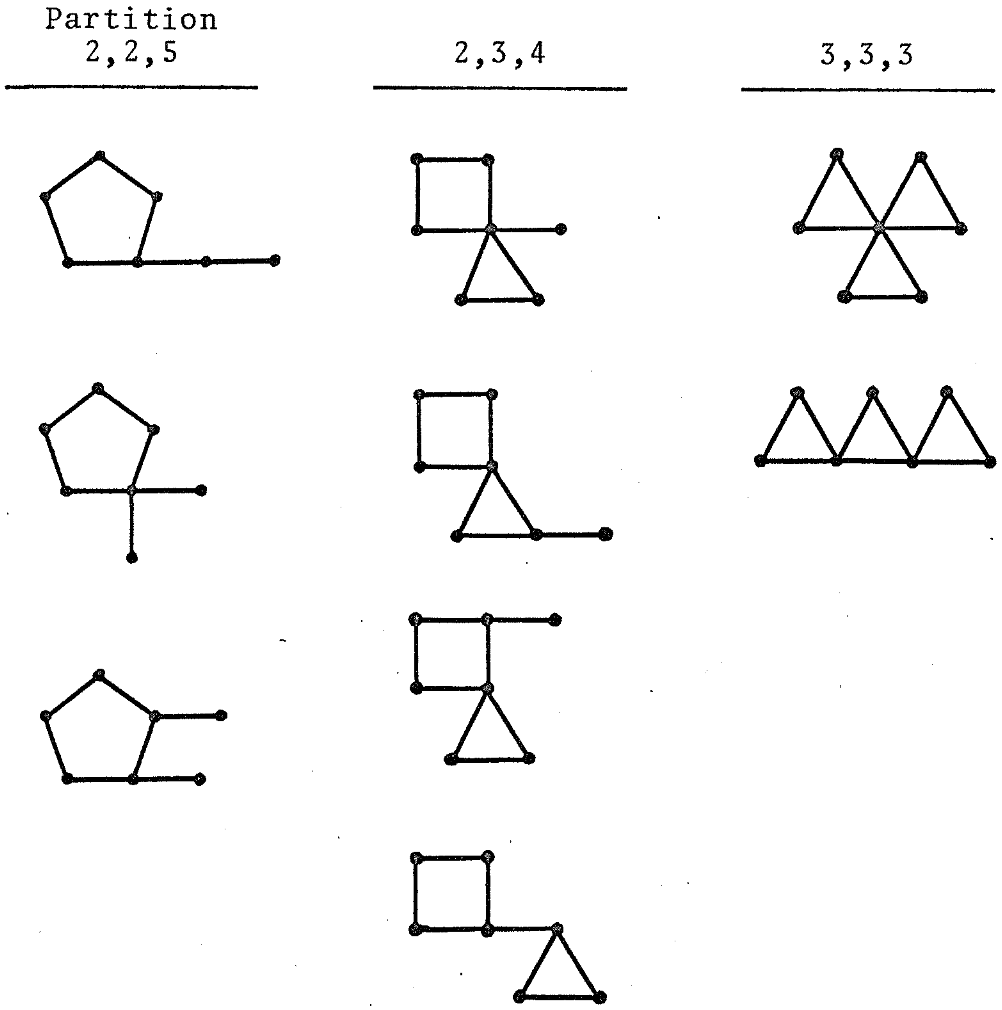
where the sum is taken over all complete trees on n vertexes, and the asterisk denotes that for $k \geqq 2$ symmetric complete stars of three or more vertexes each, including the composition, the product of the σ_{m_i} is replaced by $Z(S_k;\sigma_m)$ from Eq. (76), where m is the composition set of each of the symmetric complete stars. Since $m_i < n$, the numbers of stars on n vertexes may be obtained successively, from the numbers of stars on n—1 and fewer vertexes in all compositions.

A useful example is given in Fig. 2. For the first two trees shown under partition 2,2,5, the contribution to $\gamma_7 - \sigma_7$ is $\sigma_{11}^2\sigma_{14}$, where σ_{11} is the generating function for stars on two distinguishable vertexes, and σ_{14} is that for stars on five vertexes, one distinguishable from the other four. For the third tree, the contribution is $\sigma_{11}^2\sigma_{23}$. Similarly, for the four trees under partition 2,3,4, the contributions to $\gamma_7 - \sigma_7$ are $\sigma_{11}\sigma_{12}\sigma_{13}$, $\sigma_{11}\sigma_{111}\sigma_{13}$, $\sigma_{11}\sigma_{12}\sigma_{112}$, and $\sigma_{11}\sigma_{12}\sigma_{13}$, respectively. The three triangles in the first tree under partition 3,3,3 are symmetric, hence the generating function for the numbers of trees corresponding to this complete tree is given by

$$Z(S_3;\sigma_{12}) = \frac{1}{6}(\sigma_{12}^3 + 3\sigma_{12}\sigma_{12;2} + 2\sigma_{12;3}).$$

For the remaining tree in Fig. 2, however, the three triangles are not symmetric. Consider trees corresponding to this complete tree. If the two end stars are *identical*, the generating function for the center star is σ_{12}, since the two cut vertexes are then symmetric. If the end stars are not identical, the two cut vertexes are not symmetric, and the generating function for the center star is σ_{111} instead. Now the generating function for the combinations of the two identical end stars is $\sigma_{12;2}$, as can be seen from Eq. (75). Hence the generating function for the required numbers of trees is

FIGURE 2.—Complete trees corresponding to the three partitions of 9 into three elements, unity excluded, as given in Table 11 for $n = 7$. The polygons represent complete stars.



$$\sigma_{12}\sigma_{12;2} + \{Z(S_2;\sigma_{12}) - \sigma_{12;2}\}\sigma_{111}, \qquad (78)$$

where the quantity in braces is the generating function for the combinations of two nonidentical end stars. Now (78) reduces to

$$\frac{1}{2}\{\sigma_{111}\sigma_{12}^2 + (2\sigma_{12}-\sigma_{111})\sigma_{12;2}\},$$

which bears a resemblance to $Z(S_2;\sigma_{12})$, but in which the terms σ_{12}^2 and $\sigma_{12;2}$ are "weighted" by the factors σ_{111} and $(2\sigma_{12}-\sigma_{111})$, respectively.

The generating functions which are obtained by Eq. (77) for $\gamma_n - \sigma_n$, for $n \leqq 5$, are presented in Table 12, and numerical results for σ_n are given in Tables 2, 3, and 6.

2. Approximation Formula

It is possible to closely approximate the numbers of stars $\sigma(n,k)$ in a simple manner. Since $S(n,k)$ is the number of stars with bond set $k(n)$ on n distinguishable vertexes, one has[47]

$$S(n,k) = \sum_{i=1}^{\sigma(n,k)} \frac{n!}{s_i},$$

where s_i is the symmetry number of the star of type i ($i = 1, 2, \ldots, \sigma(n,k)$) with bond set $k(n)$ on n vertexes. The ratio $S(n,k)/\sigma(n,k)$ is then equal to $n!$ times the average of the reciprocals of the symmetry numbers of the $\sigma(n,k)$

TABLE 12.—Generating functions for $\gamma_n - \sigma_n$, the number of trees consisting of two or more stars each, on $n \leqq 5$ vertexes.

n	n	$\gamma_n - \sigma_n$
3	1,1,1	$3\sigma^2_{11}$
	1,2	$2\sigma^2_{11}$
	3	$\frac{1}{2}(\sigma^2_{11} + \sigma_{11;2}) = \sigma^2_{11}$
4	1,1,1,1	$12\sigma_{111}\sigma_{11} + 16\sigma^3_{11}$
	1,1,2	$(5\sigma_{111}+2\sigma_{12})\sigma_{11} + 9\sigma^3_{11}$
	2,2	$2(\sigma_{111}+\sigma_{12})\sigma_{11} + 4\sigma^3_{11} + 2\sigma^2_{11}\sigma_2$
	1,3	$(\sigma_{111}+2\sigma_{12})\sigma_{11} + 4\sigma^3_{11}$
	4	$\sigma_{12}\sigma_{11} + \sigma^3_{11} + \sigma^2_{11}\sigma_2$
5	1,1,1,1,1	$20\sigma_{1111}\sigma_{11} + 15\sigma^2_{111} + 150\sigma_{111}\sigma^2_{11}$ $+ 125\sigma^4_{11}$
	1,1,1,2	$(7\sigma_{1111}+6\sigma_{112})\sigma_{11} + 6\sigma^2_{111} + 3\sigma_{111}\sigma_{12}$ $+ 9(8\sigma_{111}+\sigma_{12})\sigma^2_{11} + 67\sigma^4_{11}$
	1,2,2	$2(\sigma_{1111}+3\sigma_{112})\sigma_{11} + \sigma^2_{12}$ $+ 2(\sigma_{111}+\sigma_{12})\sigma_{111} + \frac{1}{2}(\sigma^2_{111}+\sigma_{111;2})$ $+ 3(11\sigma_{111}+4\sigma_{12})\sigma^2_{11} + 37\sigma^4_{11}$

TABLE 12—Continued

n	n	$\gamma_{n} - \sigma_{n}$
5	1,1,3	$(\sigma_{1111}+4\sigma_{112}+2\sigma_{13})\sigma_{11} + \sigma_{111}^2 + 3\sigma_{111}\sigma_{12} + (22\sigma_{111}+9\sigma_{12})\sigma_{11}^2 + 26\sigma_{11}^4$
	2,3	$(3\sigma_{112}+\sigma_{13})\sigma_{11} + \sigma_{111}\sigma_{12} + \sigma_{12}^2 + \frac{1}{2}(\sigma_{111}^2+\sigma_{111;2}) + 9(\sigma_{111}+\sigma_{12})\sigma_{11}^2 + 15\sigma_{11}^4$
	1,4	$(\sigma_{112}+2\sigma_{13})\sigma_{11} + \sigma_{111}\sigma_{12} + \frac{1}{2}(\sigma_{12}^2+\sigma_{12;2}) + 2(2\sigma_{111}+3\sigma_{12})\sigma_{11}^2 + 9\sigma_{11}^4$
	5	$\sigma_{13}\sigma_{11} + \frac{1}{2}(\sigma_{12}^2+\sigma_{12;2}) + 3\sigma_{12}\sigma_{11}^2 + 3\sigma_{11}^4$

stars, and similarly for $C(n,\mathrm{k})/\gamma(\mathrm{n,k})$ for trees and $N(n,\mathrm{k})/\mu(\mathrm{n,k})$ for graphs. If the proportion of stars is small, the complements of most or all stars will be disconnected graphs, and, for a given (n,k), the ratio S/σ will be nearly equal to the ratio N/μ. If the proportion of stars is large, however, the complements of many (or most) stars will be trees or stars themselves, and S/σ will generally be closer to C/γ than to N/μ. If $S = 0$, then $\sigma = 0$, and if $S = N$, $\sigma = \mu$. Arbitrarily selecting $S = N/2$ as the dividing line between these two cases, one may write

$$\sigma \cong \begin{cases} \mu S/N & S \leqq N/2, \\ \gamma S/C & S > N/2, \end{cases} \tag{79}$$

where each variable is a function of the same (n,k). This approximation will usually yield either the correct value or a result very close to the correct value for $\sigma(\mathrm{n,k})$. For example, all nonzero deviations from the correct values of $\sigma(\mathrm{n,k})$ for $n = 5$, as calculated from Eq. (79), are given in Table 13. Of 121 different bond sets possible, the exact values are calculated by Eq. (79) in 92 cases (95 cases, for n = 5). For the nonzero deviations, generally, with increasing $\sigma(\mathrm{n,k})$, the absolute error from Eq. (79) increases slowly, but the percentage error decreases markedly.

TABLE 13.—Comparison of approximate values of σ(n,k) from Eq. (79) with exact values, for n = 5, for cases in which a nonzero difference occurs.

n = 1,1,1,2				
k_2,k_3	σ(n,k)		Difference	Percentage Difference
	Exact	From Eq. (79)		
5,0	6	6.57	0.57	9.52
6,0	37	38.67	1.67	4.50
7,0	55	56.67	1.67	3.03
3,1	30	31.60	1.60	5.33
4,1	231	234.43	3.43	1.48
5,1	618	622.86	4.86	0.79
6,1	811	815.95	4.95	0.61
7,1	586	589.87	3.87	0.66
1,2	30	32.40	2.40	8.00
2,2	324	327.60	3.60	1.11
3,2	1380	1386.00	6.00	0.43
4,2	3243	3251.25	8.25	0.25
5,2	4713	4719.73	6.73	0.14
6,2	4425	4431.01	6.01	0.14
7,2	2706	2710.40	4.40	0.16
1,3	520	524.07	4.07	0.78
2,3	2361	2367.07	6.07	0.26
3,3	6498	6506.06	8.06	0.12
4,3	11793	11806.27	13.27	0.11
5,3	14628	14638.59	10.59	0.07
6,3	12511	12519.70	8.70	0.07
7,3	7282	7287.29	5.29	0.07
1,4	1063	1065.60	2.60	0.24
2,4	4716	4720.79	4.79	0.10
3,4	12561	12567.97	6.97	0.06
4,4	22038	22047.84	9.84	0.04
5,4	26547	26554.52	7.52	0.03
6,4	22222	22227.34	5.34	0.02
7,4	12763	12765.86	2.86	0.02

TABLE 13—Continued

n = 1,2,2				
k_2,k_3	σ(n,k) Exact	σ(n,k) From Eq. (79)	Difference	Percentage Difference
5,0	4	3.71	-0.29	-7.14
6,0	21	22.00	1.00	4.76
7,0	32	33.33	1.33	4.17
3,1	16	16.80	0.80	5.00
4,1	121	123.00	2.00	1.65
5,1	322	325.71	3.71	1.15
6,1	423	427.22	4.22	1.00
7,1	310	313.60	3.60	1.16
1,2	16	17.73	1.73	10.83
2,2	169	172.04	3.04	1.80
3,2	712	716.00	4.00	0.56
4,2	1658	1665.29	7.29	0.44
5,2	2409	2413.86	4.86	0.20
6,2	2265	2270.82	5.82	0.26
7,2	1396	1400.18	4.18	0.30
1,3	274	277.78	3.78	1.38
2,3	1216	1221.13	5.13	0.42
3,3	3314	3320.69	6.69	0.20
4,3	5988	5999.98	11.98	0.20
5,3	7422	7431.63	9.63	0.13
6,3	6356	6364.32	8.32	0.13
7,3	3718	3723.20	5.20	0.14
1,4	555	557.56	2.56	0.46
2,4	2414	2418.64	4.64	0.19
3,4	6380	6386.74	6.74	0.11
4,4	11156	11165.62	9.62	0.09
5,4	13429	13436.35	7.35	0.05
6,4	11253	11258.28	5.28	0.05
7,4	6488	6490.84	2.84	0.04

TABLE 13—Continued

n = 1,1,3				
k_2,k_3	$\sigma(n,k)$		Difference	Percentage Difference
	Exact	From Eq. (79)		
5,0	2	2.57	0.57	28.57
6,0	14	15.67	1.67	11.90
7,0	22	24.17	2.17	9.85
3,1	10	11.65	1.65	16.50
4,1	81	84.64	3.64	4.50
5,1	218	222.86	4.86	2.23
6,1	288	292.86	4.86	1.69
7,1	213	217.47	4.47	2.10
1,2	10	12.40	2.40	24.00
2,2	114	117.60	3.60	3.16
3,2	480	486.00	6.00	1.25
4,2	1120	1127.67	7.67	0.68
5,2	1623	1629.73	6.73	0.41
6,2	1532	1537.85	5.85	0.38
7,2	946	950.40	4.40	0.47
1,3	187	191.67	4.67	2.50
2,3	821	827.07	6.07	0.74
3,3	2233	2240.91	7.91	0.35
4,3	4028	4042.41	14.41	0.36
5,3	4988	4998.59	10.59	0.21
6,3	4280	4288.66	8.66	0.20
7,3	2511	2516.94	5.94	0.24
1,4	380	383.26	3.26	0.86
2,4	1633	1637.76	4.76	0.29
3,4	4296	4302.95	6.95	0.16
4,4	7498	7509.15	11.15	0.15
5,4	9017	9024.52	7.52	0.08
6,4	7566	7571.34	5.34	0.07
7,4	4372	4375.52	3.52	0.08

TABLE 13—Continued

n = 2,3				
k_2,k_3	σ(n,k)		Difference	Percentage Difference
	Exact	From Eq. (79)		
5,0	2	1.52	-0.48	-23.81
6,0	9	9.33	0.33	3.70
7,0	14	15.00	1.00	7.14
3,1	6	6.30	0.30	5.00
4,1	44	45.00	1.00	2.27
5,1	116	118.10	2.10	1.81
6,1	153	155.49	2.49	1.63
7,1	115	117.60	2.60	2.26
1,2	6	6.93	0.93	15.56
2,2	61	62.84	1.84	3.02
3,2	252	254.00	2.00	0.79
4,2	577	581.54	4.54	0.79
5,2	838	840.61	2.61	0.31
6,2	790	793.82	3.82	0.48
7,2	494	496.71	2.71	0.55
1,3	101	103.70	2.70	2.68
2,3	429	432.11	3.11	0.72
3,3	1149	1152.91	3.91	0.34
4,3	2058	2066.12	8.12	0.39
5,3	2546	2552.10	6.10	0.24
6,3	2187	2192.40	5.40	0.25
7,3	1292	1295.76	3.76	0.29
1,4	201	203.02	2.02	1.00
2,4	842	845.04	3.04	0.36
3,4	2194	2198.41	4.41	0.20
4,4	3811	3817.99	6.99	0.18
5,4	4580	4584.85	4.85	0.11
6,4	3847	3850.49	3.49	0.09
7,4	2235	2237.22	2.22	0.10

TABLE 13—Continued

n = 1,4				
k_2,k_3	$\sigma(n,k)$		Difference	Percentage Difference
	Exact	From Eq. (79)		
5,0	1	0.86	-0.14	-14.29
6,0	5	5.67	0.67	13.33
7,0	8	9.17	1.17	14.58
3,1	3	3.45	0.45	15.00
4,1	23	24.21	1.21	5.28
5,1	59	62.86	3.86	6.54
6,1	79	81.52	2.52	3.19
7,1	62	64.40	2.40	3.87
1,2	3	3.87	0.87	28.89
2,2	32	33.60	1.60	5.00
3,2	131	133.00	2.00	1.53
4,2	298	301.75	3.75	1.26
5,2	432	434.43	2.43	0.56
6,2	409	412.12	3.12	0.76
7,2	258	260.09	2.09	0.81
1,3	54	56.48	2.48	4.60
2,3	223	225.56	2.56	1.15
3,3	592	595.20	3.20	0.54
4,3	1054	1061.13	7.13	0.68
5,3	1301	1305.82	4.82	0.37
6,3	1122	1126.12	4.12	0.37
7,3	668	671.25	3.25	0.49
1,4	107	108.94	1.94	1.81
2,4	436	438.32	2.32	0.53
3,4	1125	1128.35	3.35	0.30
4,4	1945	1951.11	6.11	0.31
5,4	2332	2335.68	3.68	0.16
6,4	1965	1967.63	2.63	0.13
7,4	1148	1150.09	2.09	0.18

TABLE 13—Continued

n = 5				
k_2,k_3	$\sigma(n,k)$		Difference	Percentage Difference
	Exact	From Eq. (79)		
5,0	1	0.29	-0.71	-71.43
7,0	3	3.33	0.33	11.11
3,1	1	0.90	-0.10	-10.00
5,1	15	15.24	0.24	1.59
6,1	20	20.38	0.38	1.90
7,1	16	16.80	0.80	5.00
1,2	1	1.07	0.07	6.67
2,2	8	8.40	0.40	5.00
4,2	67	68.00	1.00	1.49
5,2	97	97.18	0.18	0.19
6,2	92	93.12	1.12	1.22
7,2	60	60.62	0.62	1.04
1,3	14	14.82	0.82	5.82
2,3	52	52.54	0.54	1.04
3,3	132	132.57	0.57	0.43
4,3	229	231.13	2.13	0.93
5,3	281	282.29	1.29	0.46
6,3	244	245.23	1.23	0.51
7,3	149	150.16	1.16	0.78
1,4	26	26.74	0.74	2.84
2,4	98	98.72	0.72	0.73
3,4	244	245.04	1.04	0.42
4,4	415	417.18	2.18	0.52
5,4	495	496.17	1.17	0.24
6,4	420	420.85	0.85	0.20
7,4	250	250.80	0.80	0.32

3. Application to Virial Coefficients

The desired combinatorial factor for the integrals corresponding to the *n*th virial coefficient for a pure substance is then, for each of the $\sigma(n,k)$ stars involved, the number of stars on distinguishable vertexes which correspond to the given star on indistinguishable vertexes, and is equal to $n!$ divided by the symmetry number for the star. The sum of all these combinatorial factors for stars of bond set $k(n)$ on n vertexes is $S(n,k)$. Hence by obtaining the symmetry numbers, the combinatorial factors can be calculated.

The numbers of topologically different stars for multicomponent systems are always greater than for systems of a single component, as shown in Table 14, and are obtained by Eq. (77) or may be closely approximated by Eq. (79). The desired combinatorial factor for each of the $\sigma(n,k)$ stars is then the number of stars on distinguishable vertexes which corresponds to the given star on n vertexes distinguishable by species only. This number is $n!$ divided by the symmetry number for the star.

TABLE 14.—Total numbers of different graphs, trees, and stars on **n** vertexes, using 2- and 3-bonds only.

n	n	Graphs	Trees	Stars
1	1	1	1	1
2	1,1	2	1	1
	2	2	1	1
3	1,1,1	16	12	9
	1,2	12	9	7
	3	8	6	5
4	1,1,1,1	1024	966	842
	1,1,2	576	539	471
	2,2	336	312	274
	1,3	240	221	194
	4	90	81	72
5	1,1,1,1,1	1048576	1043480	1023950
	1,1,1,2	532480	529730	519557
	1,2,2	271360	269862	264552
	1,1,3	183056	181944	178323
	2,3	93920	93332	91409
	1,4	48400	48059	47022
	5	10688	10592	10346
6	1,1,1,1,1,1	34359738638		
	6	48565952		

G. SUMMARY AND DISCUSSION

1. Summary and Implications of This Work

This thesis develops the combinatorial theory of nonlinear graphs for specific application to calculating the virial coefficients. Definitions and terminology for nonlinear graphs and the relationship of the graphs to cluster functions and to the virial coefficients are given. An explicit formula for counting nonlinear graphs on distinguishable vertexes, Eq. (20), is obtained, and generating functions relate the numbers of graphs to the numbers of nonlinear trees and stars, by Eqs. (27), (30), and (33). The *cycle function* is defined, and the triplet cycle function, for 3-bonds on n vertexes, is derived in general, and is given by Eq. (43). The problem of counting nonlinear graphs on indistinguishable vertexes is solved by generalizing Pólya's theorem to apply to nonlinear graphs with more than one type of bond present, in Eq. (50). Generating functions for graphs and trees on indistinguishable vertexes are used to obtain Eq. (59), permitting the numbers of trees to be calculated. All of these techniques are extended to apply to multicomponent systems and to rooted graphs and trees. Then with this entire combinatorial development at hand, the problem of counting stars on indistinguishable or multicomponent vertexes is solved by a systematic procedure, which is symbolically represented by Eq. (77). A fortuitously simple but surprisingly accurate approximation formula for stars, Eq. (79), permits close approximation to be easily made before or instead of embarking on the detailed procedure. Tables 2 through 6 give the numbers of all graphs, trees, and stars for 2- and 3-bonds on n = 3, 4, and 5 vertexes.

The utility of the results obtained in this work is primarily in the extension of systematic methods for counting linear graphs to nonlinear graphs as well, and in a solution of the formidable problem of counting stars, which has immediate application to the theoretical development of the virial coefficients including nonadditivity effects. By considering only topologically distinct stars instead of all stars, and multiplying each by the appropriate combinatorial coefficient, the calculation of the nth virial coefficient is simplified by a reduction in the number of distinct integrations, from the number of stars on n distinguishable vertexes to the number of stars on n indistinguishable vertexes, for a system of a single component. For example, considering 2- and 3-bonds only, for n = 4 the number of integrals is reduced from 842 to 72; and for n = 5, from 1,023,950 to 10,346, as shown in Table 14.

With the development of high-speed computers, the tables of calculated numbers of graphs, trees, and stars will be useful in determining the feasibility of making particular calculations. At the present time, the practical limit for calculations involving even the simplest nonadditivity effects is apparently at $n = 4$, for the fourth virial coefficient. Compelling reasons, depending on the knowledge of accurate two-body potentials, and to a lesser degree, accurate three-body potentials, would be necessary to justify even the calculation of the 72 different integrals required for $n = 4$, for a pure substance. For multicomponent systems, the number of distinct integrations is greater, so that the practical limit for calculation of the integrals is reduced from that for pure substances, and even more compelling reasons would be necessary to justify the calculation of the fourth virial coefficient including three-body nonadditivity effects.

2. Suggestions for Future Research

This research uncovers a number of important problems and topics for future investigation:

1. By suitable transformation of Mayer cluster functions, Ree and Hoover(23,37-39) have been able to substantially reduce the number of integrations required to calculate virial coefficients assuming pairwise additivity. It should be possible to perform corresponding transformations when nonadditivity effects are included, but the transformations themselves may be considerably more complicated.

2. As indicated in Sec. E.5, the radial distribution function corresponds to all doubly rooted stars containing a 2-bond between the roots, from which this 2-bond is then removed. Combinatorial theory to obtain the numbers of such graphs which are topologically distinct would be useful, especially for cases in which nonadditivity effects are included, which become too cumbersome to count by inspection for stars on $n \geqq 4$ vertexes. Systematic procedures for counting related types of graphs, corresponding to correlation functions,$^{(7,14,49)}$ would likewise be of value.

3. With the use of suitably accurate potential functions, the third and fourth virial coefficients, including three-body nonadditivity effects, could be calculated and compared to results from experimental data. A better appraisal than is presently possible of the effect of various approximations in potential functions and of the contribution of nonadditivity effects could then be made.

4. Since the combinatorial results of this work have been calculated manually, but the systematic procedures are easily amenable to computer calculation, these results may be extended to $n \geqq 6$, or special cases calculated as desired. Particularly convenient estimates of the numbers of stars may be made by Eq. (79), before obtaining the exact numbers by the detailed procedure, since questions of feasibility may be important, due to the large numbers of stars involved.

During the past two decades, applications of graph theory have increased rapidly in number and may well continue to do so. Wherever interrelationships or interactions occur, whether of molecules, men, or machines, graph theory may prove to be a useful analytical tool in studying the interrelationships themselves. It is therefore by no means unlikely that the combinatorial theory of nonlinear graphs, developed in this work, can be applied in yet unforeseen ways to help in gaining scientific understanding and in expanding an increasingly sophisticated technology.

NOMENCLATURE

$A_n(T)$	nth virial coefficient in Eq. (4), defined by Eq. (5) or Eq. (7)
B	Variable defined by Eq. (21)
$B_j(n)$	Maximum number of j-bonds in a graph on n vertexes, given by Eq. (17)
b_i	Number of cycles of length i in an element of G in the cycle index $Z(\mathrm{G})$
C_n	Function defined by Eq. (24)
$C(n,\mathrm{k})$	Number of trees of bond set $\mathrm{k}(n)$ on n distinguishable vertexes, defined by Eq. (24)
$C'(x,\mathrm{y})$	Generating function for trees on distinguishable vertexes, defined by Eq. (24)
$d(i,j)$	Greatest common divisor of i and j
$F_j(n,\mathrm{p})$	Cycle function for j-bonds on n vertexes, defined by Eq. (39)
$F_j(\mathrm{n})$	Cycle function for j-bonds on S_n
f	Set of arbitrary functions, in Eq. (36)
f_m	Cluster function, defined by Eq. (2)
G	Group of permutations of elements of the set S
G_j	Permutation group for j-bonds on graphs of vertexes
g_b	Number of elements of G_j which consist of b_i cycles of length i (i = 1, 2, ...)
j	Type of bond in a graph, associating j vertexes
K	Variable defined by Eqs. (29)
k	Boltzmann constant, in Eq. (2)
$\mathrm{k}(n)$	Bond set for a graph on n vertexes, defined by Eq. (13)

$\mathrm{k}(\mathrm{n})$	Bond set for a graph on n vertexes
k_j	Number of j-bonds in a graph
$m(i,j)$	Least common multiple of i and j
N	Number of particles in a configuration
$N_\circ$	Avogadro's number
$N(n,\mathrm{k})$	Number of graphs of bond set $\mathrm{k}(n)$ on n distinguishable vertexes, given by Eq. (20)
$N_j(n)$	Total number of possible graphs with j-bonds on n distinguishable vertexes, given by Eq. (19)
$N(n)$	Total number of graphs on n distinguishable vertexes, using all combinations of different types of bonds
$N'(x,\mathrm{y})$	Generating function for graphs on distinguishable vertexes, defined by Eq. (24)
n	Number of vertexes in a graph
n	Variable defined by Eq. (8)
n_i	Number of vertexes of species i in a graph
n_r	Composition set for root vertexes in a rooted graph
P	Pressure
$\mathrm{p}(n)$	Partition set of the integer n, defined by Eq. (38)
$\mathrm{p}(\mathrm{n})$	Variable defined by Eq. (68)
R	Molar gas constant
r_i	Set of coordinates of particle i
S_n	Sum of all cluster terms which correspond to stars on n distinguishable vertexes, in Eq. (6)
S_n	Function defined by Eq. (24)

S_{n}	Sum of all cluster terms which correspond to stars on n vertexes distinguishable by species only
S_{n}	Symmetric group of permutations of n elements
$\mathrm{S}_{\mathbf{n}}$	Direct product of s symmetric groups corresponding to n, defined by Eq. (63)
$S(n,\mathbf{k})$	Number of stars of bond set $\mathbf{k}(n)$ on n distinguishable vertexes
$S'(x,\mathbf{y})$	Generating function for stars on distinguishable vertexes, defined by Eq. (24)
s	Number of species in a configuration
s_i	Symmetry number of a star of type i, in Eq. (32)
T	Absolute temperature
U_N	Potential energy of interaction of a configuration of N particles, given by Eq. (1)
u_{m}	Potential energy of interaction of the set m of particles, in Eq. (2)
V	Volume
$\overline{V}$	Molar volume
x	Arbitrary variable, in Eqs. (24)
x	Variable defined by Eq. (9)
$\mathbf{x}$	Set of arbitrary variables, in Eqs. (62)
x_i	Mole fraction of species i, in Eq. (9)
$\mathbf{y}$	Set of arbitrary variables, defined by Eq. (25)
$\mathbf{y}^{\mathbf{k}(n)}$	Weight assigned to a graph with bond set $\mathbf{k}(n)$, defined by Eq. (26)
$Z(N,V,T)$	Configuration integral, defined by Eq. (3)
$Z(\mathrm{G})$	Cycle index of the permutation group G, defined by Eq. (36)
$Z(\mathrm{G}_j)$	Cycle index of G_j, given by Eq. (48)

$Z_j(S_n)$	Cycle index for nonlinear graphs on *n* vertexes, given by Eq. (49)
$Z_j(S_{\mathrm{n}})$	Cycle index for j-bonds on n vertexes, given by Eq. (67)
β_n	Variable defined by Eq. (6)
β_{n}	Variable defined by Eq. (11)
$\gamma(\mathrm{n},\mathrm{k})$	Number of trees of bond set k(n) on n vertexes
$\gamma'(x,y)$	Generating function for trees, defined by Eq. (62)
γ_{n}	Function defined by Eq. (62)
$\gamma_{\mathrm{n};i}$	Function defined by Eq. (71)
$\mu(\mathrm{n},\mathrm{k})$	Number of graphs of bond set k(n) on n vertexes
$\mu'(x,y)$	Generating function for graphs, defined by Eq. (62)
$\mu_t'(x,y)$	Generating function for graphs which consist of exactly *t* trees
$\prod$	Product operator
$\sum$	Summation operator
$\sigma(\mathrm{n},\mathrm{k})$	Number of stars of bond set k(n) on n vertexes
$\sigma'(x,y)$	Generating function for stars, defined by Eq. (62)
σ_{n}	Function defined by Eq. (62)
$\sigma_{\mathrm{n};i}$	Function defined by Eq. (75)

Superscripts

$'$	Denotes generating function for graphs, trees, or stars
$^{\circ}$	Denotes generating function for rooted graphs, trees, or stars
$^{\circ\circ}$	Denotes generating function for doubly rooted graphs, trees, or stars

Conventions

$\mathbf{n}!$	Product of factorials of the elements n_i of $\mathbf{n}$, defined by Eq. (12)
$\mathbf{x}^{\mathbf{n}}$	Product of elements x_i of $\mathbf{x}$, raised to the powers n_i of $\mathbf{n}$, respectively, defined by Eq. (10)
$[a]$	Denotes maximum integer not greater than a
$\binom{n}{k}$	Binomial coefficient, defined by Eq. (15)
$\binom{\mathbf{B}}{\mathbf{k}}$	Product of binomial coefficients of respective elements of $\mathbf{B}$ and $\mathbf{k}$, defined by Eq. (22)

REFERENCES

1. Allnatt, A. R., "Statistical Mechanics of Point-Defect Interactions in Solids," *Advan. Chem. Phys. 11*, 1 (1967).

2. Axilrod, B. M., and E. Teller, "Interaction of the van der Waals Type between Three Atoms," *J. Chem. Phys. 11*, 299 (1943).

3. ————, "The Triple-Dipole Interaction between Atoms and Cohesion in Crystals of the Rare Gases," *J. Chem. Phys. 17*, 1349 (1949).

4. ————, "Triple-Dipole Interaction. I. Theory," *J. Chem. Phys. 19*, 719 (1951); "II. Cohesion in Crystals of the Rare Gases," *ibid.*, p. 724.

5. Bade, W. L., "Drude-Model Calculation of Dispersion Forces. I. General Theory," *J. Chem. Phys. 27*, 1280 (1957); W. L. Bade and J. G. Kirkwood, "II. The Linear Lattice," *ibid.*, p. 1284; W. L. Bade, "III. The Fourth-Order Contribution," *ibid. 28*, 282 (1958).

6. Baer, S., "Nonadditive Interactions between Molecules in a Statistical Treatment of Dilute Gases," *J. Chem. Phys. 47*, 2889 (1967).

7. Baxter, R. J., "Three-Particle Hyper-Netted Chain Approximation," *Ann. Phys. 46*, 509 (1968).

8. Born, M., "The Statistical Mechanics of Condensing Systems," *Physica 4*, 1034 (1937).

9. ————, and K. Fuchs, "The Statistical Mechanics of Condensing Systems," *Proc. Roy. Soc.* (London) *166*, 391 (1938).

10. Chan, Y. M., and A. Dalgarno, "Long-Range Interactions between Three Hydrogen Atoms," *Mol. Phys. 9*, 525 (1965).

11. De Bruijn, N. G., "Pólya's Theory of Counting," in *Applied Combinatorial Mathematics*, ed. by E. F. Beckenbach, Wiley, New York (1964), Chap. 5.

12. Ford, G. W., and G. E. Uhlenbeck, "Combinatorial Problems in the Theory of Graphs. I," *Proc. Nat. Acad. Sci. U.S.A. 42*, 122 (1956); G. W. Ford, R. Z. Norman, and G. E. Uhlenbeck, "II," *ibid.*, p. 203; G. W. Ford and G. E. Uhlenbeck, "III," *ibid.*, p. 529; "IV," *ibid. 43*, 163 (1957).

13. Friedman, H. L., "On Mayer's Ionic Solution Theory," *Mol. Phys. 2*, 23 (1959).

14. ————, *Ionic Solution Theory, Based on Cluster Expansion Methods*, Interscience, New York (1962).

15. Fuchs, K., "The Statistical Mechanics of Many-Component Systems," *Proc. Roy. Soc.* (London) *179A*, 408 (1941).

16. Graben, H. W., and R. D. Present, "Evidence for Three-Body Forces from Third Virial Coefficients," *Phys. Rev. Letters 9*, 247 (1962).

17. ————, R. D. Present, and R. D. McCulloch, "Intermolecular Three-Body Forces and Third Virial Coefficients," *Phys. Rev. 144*, 140 (1966).

18. Hall, M. Jr., *The Theory of Groups*, Macmillan, New York (1959), Chaps. 2, 3, and 5.

19. Harary, F., "The Number of Linear, Directed, Rooted, and Connected Graphs," *Trans. Am. Math. Soc. 78*, 445 (1955).

20. ————, "Combinatorial Problems in Graphical Enumeration," in *Applied Combinatorial Mathematics*, ed. by E. F. Beckenbach, Wiley, New York (1964), p. 207.

21. ————, "Graphical Enumeration Problems," in *Graph Theory and Theoretical Physics*, ed. by F. Harary, Academic Press, London (1967), Chap. 1.

22. Hill, T. L., *Statistical Mechanics*, McGraw-Hill, New York (1956), p. 123.

23. Hoover, W. G., and A. G. De Rocco, "Sixth and Seventh Virial Coefficients for the Parallel Hard-Cube Model," *J. Chem. Phys. 36*, 3141 (1962).

24. Husimi, K., "Note on Mayer's Theory of Cluster Integrals," *J. Chem. Phys. 18*, 682 (1951).

25. Hutchinson, P., "The Superposition Approximation for the Triplet Distribution Function in a Classical Fluid," *Proc. Phys. Soc. 91*, 506 (1967).

26. Kestner, N. R., and O. Sinanoğlu, "Effective Intermolecular Pair Potentials in Nonpolar Media," *J. Chem. Phys. 38*, 1730 (1963).

27. Kihara, T., "Intermolecular Forces and Equation of State of Gases," *Advan. Chem. Phys. 1*, 267 (1958).

28. Knuth, D. E., *Fundamental Algorithms*, Vol. 1 of *The Art of Computer Programming*, Addison-Wesley, Reading, Mass. (1968), Chap. 2.

29. MacMahon, P. A., *Combinatory Analysis*, Cambridge University Press (1915), Vol. 1.

30. Mayer, J. E., "The Statistical Mechanics of Condensing Systems. I," *J. Chem. Phys. 5*, 67 (1937); J. E. Mayer and P. G. Ackermann, "II," *ibid.*, p. 74; J. E. Mayer and S. F. Harrison, "III," *ibid. 6*, 87 (1938); S. F. Harrison and J. E. Mayer, "IV," *ibid.*, p. 101.

31. ————, and M. G. Mayer, *Statistical Mechanics*, Wiley, New York (1940), Chap. 13.

32. ————, and E. Montroll, "Molecular Distribution," *J. Chem. Phys. 9*, 2 (1941).

33. Midzuno, Y., and T. Kihara, "Nonadditive Intermolecular Potential in Gases. I. van der Waals Interactions," *J. Phys. Soc. Japan 11*, 1045 (1956); S. Koba, S. Kaneko, and T. Kihara, "II. Cluster Integrals," *ibid.*, p. 1050.

34. Montroll, E. W., and J. E. Mayer, "Statistical Mechanics of Imperfect Gases," *J. Chem. Phys. 9*, 626 (1941).

35. Muto, Y., *Nihon Sugaku Butsuri Gakkaishi* (*J. Phys.-Math. Soc. Japan*) *17*, 629 (1943). This reference has been erroneously reported by numerous authors(6,11,17,26,27,45) as *Proc. Phys.-Math. Soc. Japan* (*Nippon Suugaku-Buturigakkwai Kizi*). See *Chem. Abstr. 42*:27b (1948).

36. Pólya, G., "Kombinatorische Anzahlbestimmungen für Gruppen, Graphen und chemische Verbindungen," *Acta. Math. 68*, 145 (1937).

37. Ree, F. H., and W. G. Hoover, "Fifth and Sixth Virial Coefficients for Hard Spheres and Hard Disks," *J. Chem. Phys. 40*, 939 (1964).

38. ————, and W. G. Hoover, "Reformulation of the Virial Series for Classical Fluids," *J. Chem. Phys. 41*, 1635 (1964).

39. ————, and W. G. Hoover, "Seventh Virial Coefficients for Hard Spheres and Hard Disks," *J. Chem. Phys. 46*, 4181 (1967).

40. Riddell, R. J. Jr., and G. E. Uhlenbeck, "On the Theory of the Virial Development of the Equation of State of Monatomic Gases," *J. Chem. Phys. 21*, 2056 (1953).

41. Riordan, J., *An Introduction to Combinatorial Analysis*, Wiley, New York (1958).

42. ————, "Generating Functions," in *Applied Combinatorial Mathematics*, ed. by E. F. Beckenbach, Wiley, New York (1964), Chap. 3.

43. Sherwood, A. E., and J. M. Prausnitz, "Third Virial Coefficient for the Kihara, Exp-6, and Square-Well Potentials," *J. Chem. Phys. 41*, 413 (1964).

44. ————, and J. M. Prausnitz, "Intermolecular Potential Functions and the Second and Third Virial Coefficients," *J. Chem. Phys. 41*, 429 (1964).

45. ————, A. G. De Rocco, and E. A. Mason, "Nonadditivity of Intermolecular Forces: Effects on the Third Virial Coefficient," *J. Chem. Phys. 44*, 2984 (1966).

46. Stell, G., "Cluster Expansions for Classical Systems in Equilibrium," in *The Equilibrium Theory of Classical Fluids*, ed. by H. L. Frisch and J. L. Lebowitz, Benjamin, New York (1964), p. II-171.

47. Uhlenbeck, G. E., and G. W. Ford, "The Theory of Linear Graphs with Applications to the Theory of the Virial Development of the Properties of Gases," in *Studies in Statistical Mechanics*, ed. by J. de Boer and G. E. Uhlenbeck, North-Holland, Amsterdam (1962), Vol. I, Part B.

48. Ursell, H. D., "The Evaluation of Gibbs' Phase Integral for Imperfect Gases," *Proc. Cambridge Phil. Soc.* *23*, 685 (1927).

49. Verlet, L., "On the Theory of Classical Fluids," *Nuovo Cimento 43*, 77 (1960); L. Verlet and D. Levesque, "II," *Physica 28*, 1124 (1962); L. Verlet, "III," *ibid. 30*, 95 (1964); "IV," *ibid. 31*, 959 (1965); "VI," *ibid. 36*, 254 (1967).

50. Williams, D. R., L. J. Schaad, and J. N. Murrell, "Deviations from Pairwise Additivity in Intermolecular Potentials," *J. Chem. Phys. 47*, 4916 (1967).

APPENDIX A

DERIVATION OF EQUATION (28)

Equation (28) is an explicit formula for counting nonlinear trees on distinguishable vertexes. The general procedure for its derivation is to expand the logarithm of Eq. (27), perform a number of algebraic manipulations, and equate coefficients of $x^n \mathrm{y}^{\mathbf{k}(n)}$ on both sides of the resulting equation.

Formally one may write

$$\ell\mathrm{n}\ (1 + x) = \sum_{m=1}^{\infty} \frac{(-1)^{m+1}}{m} x^m .$$

Hence Eq. (27) yields

$$C'(x,\mathrm{y}) = \sum_{m=1}^{\infty} \frac{(-1)^{m+1}}{m} \{N'(x,\mathrm{y})\}^m .$$

Substitution from Eqs. (24) results in

$$\sum_{n=1}^{\infty} C_n \frac{x^n}{n!} = \sum_{m=1}^{\infty} \frac{(-1)^{m+1}}{m} \left\{ \sum_{n=1}^{\infty} N_n \frac{x^n}{n!} \right\}^m . \qquad (\mathrm{A.1})$$

On the left side of Eq. (A.1), let $n = n^*$. Then for $m > n^*$, all powers of x on the right will be greater than n^*, and none of these will correspond with x^{n^*} on the left. Hence $m \leqq n^*$, and one may write

$$\sum_{n^*=1}^{\infty} C_{n^*} \frac{x^{n^*}}{n^*!} = \sum_{m=1}^{n^*} \frac{(-1)^{m+1}}{m} \left\{ \sum_{n=1}^{n^*-m+1} N_n \frac{x^n}{n!} \right\}^m . \quad (A.2)$$

The upper limit on the summation over n in Eq. (A.2) is determined by the largest value of n for which there will be a term in x^{n^*} to correspond with the one on the left side of the equation. The coefficient of x^{n^*} in the quantity in braces is

$$\sum_{c(n^*)} \prod_{i=1}^{m} \frac{1}{n_i^*!} N_{n_i^*} ,$$

where the summation is over all compositions (ordered partitions)(29) of n^* into m nonzero elements n_i^* ($i = 1, 2, \ldots, m$) such that $n_i^* = 1, 2, \ldots, n^*-m+1$. Permutations of the partitions (of the subscripts i in the n_i^*) will yield identical contributions but must be considered separately in the summation process. From Eqs. (20), (22), and (24), one obtains

$$N_{n_i^*} = \sum_{\mathrm{k}(n_i^*)} \prod_{j=1}^{n_i^*} \binom{B_j(n_i^*)}{k_j(n_i^*)} \mathrm{y}^{\mathrm{k}(n_i^*)} .$$

The coefficient of x^{n^*} on the right side of Eq. (A.2) then becomes

$$\sum_{m=1}^{n} \frac{(-1)^{m+1}}{m} \sum_{c(n^*)} \frac{1}{\mathrm{n}^*!} \prod_{i=1}^{m} \prod_{j=1}^{n_i^*} \sum_{\mathrm{k}(n_i^*)} \binom{B_j(n_i^*)}{k_j(n_i^*)} \mathrm{y}^{\mathrm{k}(n_i^*)} , \quad (A.3)$$

where the product over j and the summation over $\mathrm{k}(n_i^*)$ have been interchanged since their variables are independent.

Now the coefficient of $x^{n*}\mathrm{y}^{\mathrm{k}(n*)}$ on the right side of Eq. (A.2) is obtained, for each m and for each ordered partition of $n*$, from those terms in (A.3) for which

$$\mathrm{y}^{\mathrm{k}(n*)} = \prod_{i=1}^{m} \mathrm{y}^{\mathrm{k}(n_i^*)},$$

or

$$k_j(n*) = \sum_{i=1}^{m} k_j(n_i^*) \qquad (j = 2,\ 3,\ \ldots,\ n*). \qquad \text{(A.4)}$$

The maximum value of n_i^* is $n*$. For $j > n_i^*$, $k_j(n_i^*) = B_j(n_i^*) = 0$ and only a factor of unity is contributed from the products in (A.3). Therefore the upper limit in the product over j may be extended from n_i^* to $n*$, and the products over i and j may be interchanged, since their ranges are then independent. For a particular j, the summation over $\mathrm{k}(n_i^*)$ involves only that j. Hence one may write

$$\prod_{i=1}^{m} \prod_{j=1}^{n_i^*} \sum_{\mathrm{k}(n_i^*)}{}' \begin{pmatrix} B_j(n_i^*) \\ k_j(n_i^*) \end{pmatrix} = \prod_{j=1}^{n*} \prod_{i=1}^{m} \sum_{k_j(n_i^*)}{}' \begin{pmatrix} B_j(n_i^*) \\ k_j(n_i^*) \end{pmatrix}, \qquad \text{(A.5)}$$

where the primes indicate the condition (A.4). Writing

$$(1 + t)^{K_j} = \prod_{i=1}^{m} (1 + t)^{B_j(n_i)},$$

where K_j is defined by Eq. (29), from Eq. (18) one has

$$\sum_{k_j(n^*)=0}^{K_j} \begin{pmatrix} K_j \\ k_j(n^*) \end{pmatrix} t^{k_j(n^*)} = \prod_{i=1}^{m} \sum_{k_j(n_i^*)}^{B_j(n_i^*)} \begin{pmatrix} B_j(n_i^*) \\ k_j(n_i^*) \end{pmatrix} t^{k_j(n_i^*)} .$$

Equating coefficients of $t^{k_j(n^*)}$, one obtains

$$\begin{pmatrix} K_j \\ k_j(n^*) \end{pmatrix} = \prod_{i=1}^{m} \sum_{k_j(n_i^*)}{}' \begin{pmatrix} B_j(n_i^*) \\ k_j(n_i^*) \end{pmatrix} . \tag{A.6}$$

Substituting Eqs. (A.5) and (A.6) into (A.3), and equating coefficients of $x^{n^*}\mathbf{y}^{\mathbf{k}(n^*)}$ in Eq. (A.2), results in

$$\frac{C\{n^*,\mathbf{k}(n^*)\}}{n^*!} = \sum_{m=1}^{n^*} \frac{(-1)^{m+1}}{m} \sum_{c(n^*)} \frac{1}{\mathbf{n}^*!} \prod_{j=1}^{n^*} \begin{pmatrix} K_j \\ k_j(n^*) \end{pmatrix} ,$$

or, upon dropping the asterisks and making use of Eqs. (22) and (29),

$$C(n,\mathbf{k}) = \sum_{m=1}^{n} \frac{(-1)^{m+1}}{m} \sum_{c(n)} \frac{n!}{\mathbf{n}!} \begin{pmatrix} \mathbf{K} \\ \mathbf{k} \end{pmatrix} ,$$

which is Eq. (28).

APPENDIX B

CONCEPTS AND DEFINITIONS IN GROUP THEORY

A *group* may be defined as a set of elements and a binary operation (denoted hereafter by †) on pairs of the elements such that:

1. The operation performed on every ordered pair of elements of the group yields a unique element of the group.

2. The binary operation is associative (but not necessarily commutative).

3. An identity element of the group exists such that the operation performed on any given element of the group and the identity element, in either order, yields the given element.

4. For every element of the group, there exists an inverse element such that the operation performed on the given element and its inverse, in either order, yields the identity element.

More briefly, a group is a set of elements, including identity and all inverse elements, and a well defined, associative binary operation on these elements.

A *permutation* is a one-to-one mapping of a set onto itself. A set of permutations themselves may form a group, termed a *permutation group*. The number of elements being permuted (in each permutation) is the *degree* of the permutation group, while the *order* of a permutation group is the

APPENDIX B

CONCEPTS AND DEFINITIONS IN GROUP THEORY

A *group* may be defined as a set of elements and a binary operation (denoted hereafter by †) on pairs of the elements such that:

1. The operation performed on every ordered pair of elements of the group yields a unique element of the group.

2. The binary operation is associative (but not necessarily commutative).

3. An identity element of the group exists such that the operation performed on any given element of the group and the identity element, in either order, yields the given element.

4. For every element of the group, there exists an inverse element such that the operation performed on the given element and its inverse, in either order, yields the identity element.

More briefly, a group is a set of elements, including identity and all inverse elements, and a well defined, associative binary operation on these elements.

A *permutation* is a one-to-one mapping of a set onto itself. A set of permutations themselves may form a group, termed a *permutation group*. The number of elements being permuted (in each permutation) is the *degree* of the permutation group, while the *order* of a permutation group is the number of permutations in the group. The order of any group is the number of elements in the group. The group of all permutations of a set of elements is called the *symmetric group*. If there are n elements in the set, the symmetric group is designated by S_n.

If g is an element of the group G, write $g \varepsilon G$. A one-to-one mapping of the elements of two groups onto each other is termed an *isomorphism* if, for $g_1 \varepsilon G$, $g_2 \varepsilon G$, and $h_1 \varepsilon H$, $h_2 \varepsilon H$, whenever g_1 is mapped to h_1 and g_2 to h_2, then $g_1 \dagger g_2$ is mapped to $h_1 \dagger h_2$. Every group is isomorphic to a permutation group of its own elements.

This work is concerned only with finite groups, that is, groups of finite order. For a finite group, an *automorphism* is an isomorphism mapping the group onto itself. Two graphs are said to be automorphic, or topologically equivalent, if there is a one-to-one mapping of their vertexes which preserves all adjacencies by bonds. The group of a graph is the group of automorphisms of the graph.

The set of ordered pairs of elements of two groups, G and H, form a new group, the *direct product* G × H, with a group operation defined by

$$(g_1,h_1) \dagger (g_2,h_2) = (g_1 \dagger g_2, h_1 \dagger h_2),$$

where g_1 and g_2 are elements of G, and h_1 and h_2 are elements of H.

A *subgroup* H of the group G is a subset of the elements of G, and itself forms a group with respect to the binary operation defined for G. The identity element of a subgroup must be the same as that of the whole group. A nonempty subset of a group is a subgroup if the group operation performed on any two elements of the subset yields an element of the subset, and the inverse of each element of the subset is also in the subset.

The set of elements formed from the group operation of all elements of a subgroup H of the group G, in turn with a fixed element of G, is called a *coset* of H. A coset of H contains the same number of elements as H, and any two cosets of H are either disjoint or identical subsets of the elements of G.

The above terms utilized in group theory are employed in this work. Only a brief description of them has been given here. For a more complete development, refer to the book by M. Hall.(18)

APPENDIX C

CALCULATION OF THE NUMBER OF TOPOLOGICALLY DISTINCT GRAPHS ON n VERTEXES CORRESPONDING TO A GIVEN GRAPH ON *n* INDISTINGUISHABLE VERTEXES

Define a similarity class(47) of vertexes as a subset of the vertexes of a graph such that, if the vertexes were labeled, interchange of any vertexes within the subset would result in a topologically equivalent graph. In general the n vertexes of a graph are partitioned into a number of similarity classes, which may be denoted by s. That is, there are s_1 vertexes in class 1, s_2 in class 2, etc.

The number of topologically distinct graphs on n vertexes with s similarity classes of the vertexes, corresponding to a given graph containing r similarity classes of n vertexes, is obtained directly from the theory of distributions developed by MacMahon.(29) This number is the coefficient of $b^{p(n,\mathrm{s})}$ in $D^{p(n,\mathrm{n})}c^{p(n,\mathrm{r})}$, where

$$c_k = \sum_{(\mathrm{k})} (\mathrm{k}) b^{p(k,\mathrm{k})};$$

$p(k,\mathrm{k})$ is the partition set corresponding to the partition k of k; (k) is the partition representation of a symmetric function of the elements of x, in general of the form

$$\sum x_i^{k_1} x_j^{k_2} x_k^{k_3} \ldots,$$

the summation being taken over the indexes $i \neq j \neq k \ldots$; and D_k is a Hammond obliteration operator,(29) defined by

$$D_k = \frac{1}{k!}\left(\sum_{i=1}^{\infty} a_{i-1} \frac{d}{da_i}\right)^k ,$$

where a_i is the coefficient of t^i in the expansion of

$$\prod_{k=1}^{s} (1 + x_k t),$$

and, from Pólya's theorem, is given by

$$a_i = Z\left\{S_i ; -\left(-\sum_{m=1}^{s} x_m\right)\right\},$$

which is obtained from $Z(S_i)$ by replacing $f_i^{p_i}$ in f^p by

$$-\left(-\sum_{m=1}^{s} x_m^i\right)^{p_i} .$$

It follows that

$$D_{n_i}(n) = (n/n_i),$$

$$D_m(n) = 0 \quad \text{if } m \text{ is not in n},$$

$$D_m(m) = 1,$$

where (n/n_i) represents (n) with the element n_i deleted (obliterated). For convenience of calculation, write

$$D_k = \sum_{(k)} D(k),$$

where the sum is taken over all partitions (k) of k, and $D_{(\mathrm{k})}$ is a partition-obliterating operator.

For example, suppose

$$n = 5,$$
$$\mathrm{n} = 1,2,2,$$
$$\mathrm{r} = 1,1,3,$$
$$\mathrm{s} = 1,1,1,2.$$

Then

$$\mathrm{D}^{\mathrm{p}(5,\mathrm{n})}\mathrm{C}^{\mathrm{p}(5,\mathrm{r})} = D_1 D_2^2 c_1^2 c_3 .$$

But

$$\begin{aligned} c_1 &= (1)b_1, \\ c_3 &= (3)b_3 + (1,2)b_1 b_2 + (1,1,1)b_1^3, \\ D_1 &= D_{(1)} \\ D_2 &= D_{(2)} + D_{(1,1)}. \end{aligned}$$

Hence

$$\begin{aligned} D_1 D_2^2 c_1^2 c_3 &= D_1 D_2^2\{(1)^2(3)b_1^2 b_3 + (1)^2(1,2)b_1^3 b_2 \\ &\qquad + (1)^2(1,1,1)b_1^5\} \\ &= D_2^2\{2(1)(3)b_1^2 b_3\} \\ &\qquad + D_2^2\{2(1)(1,2) + (1)^2(2)\}b_1^3 b_2 \\ &\qquad + D_2^2\{2(1)(1,1,1) + (1)^2(1,1)\}b_1^5 \\ &= D_2\{2(1)^2 + (1)^2 + 2(2) + (2)\}b_1^3 b_2 \\ &\qquad + D_2\{2(1,1) + (1,1) + 2(1)^2\}b_1^5 \\ &= D_2\{3(1)^2 + 3(2)\}b_1^3 b_2 \\ &\qquad + D_2\{3(1,1) + 2(1)^2\}b_1^5 \\ &= 6b_1^3 b_2 + 2b_1^5 . \end{aligned}$$

The coefficient of $b^{p(5,s)} = b_1^3 b_2$ is 6, which is the coefficient of $\sigma_{112}\sigma_{11}$ in Table 12 for n = 1,2,2.

In general n, r, and s can be any three partitions of n, and n and r may be interchanged if desired for convenience. To calculate the total number of graphs without regard to s, take all $b_i = 1$ and calculate $D^{p(n,n)}C^{p(n,r)}$.

PART II

CHEMICAL THERMODYNAMICS OF OPEN SYSTEMS*

---

*Some results of Part II of this thesis have been revised from part of a paper by A. H. Larsen, C. S. Lu, and C. J. Pings, "Thermodynamics of Chemical Equilibrium—V. Multiple-Reaction and Open Systems," *Chem. Eng. Sci.* *23*, 289 (1968). Previous papers in the series(9) considered closed systems only.

A. INTRODUCTION AND BACKGROUND

Classical thermodynamics is a phenomenological theory relating certain macroscopic properties of matter under various conditions. It is usually limited to describing the behavior of closed systems. Furthermore, it is most often restricted to equilibrium or near-equilibrium conditions, in that any two states of a system are assumed to be associated by some reversible path, and the thermodynamic properties of these states are rigorously defined only under equilibrium conditions.(3)

1. Closed and Open Systems

While a closed system permits only an exchange of energy with the surroundings, an open system permits a transfer of both matter and energy across the boundary of the system. In an open system, the thermodynamic properties and the masses of the various constituents may be varied in an arbitrary manner.

For closed systems, the first law of thermodynamics establishes the existence of a state function E, called the internal energy. Assume that this function exists also when the composition varies in an arbitrary manner.(10) Similarly, the second law of thermodynamics establishes the existence of a state function S, the entropy, for a closed system. Assume that this function exists also for an open system.

2. The First Law of Thermodynamics

For a closed system, the first law of thermodynamics(7) may be written

$$\mathrm{d}E = \delta q - \delta w,$$

where δ is a variation operator used to indicate that q and w are not state variables, but are path-dependent functions instead. Here q represents heat transferred to the system, and w work done by the system.

Assume that E exists and is a state function for open systems also. Define q to be the heat transferred to an open system due to a temperature gradient, and w to be the work done by the system in addition to that required to introduce or remove material at the boundary.* Then material crossing the boundary, either entering or leaving the system, has an additional contribution to the differential change in internal energy:

$$\mathrm{d}E = \delta q - \delta w + \sum_{k=1}^{c} \overline{H}_k^{(e)} \mathrm{d}n_k^{(e)}, \tag{1}$$

---

*This definition of w for an open system differs from the definition of work used by Sage,(12) who includes the work involved in transferring material across the boundary.

where $\overline{H}_k^{(e)}$ is the partial molal enthalpy of species k entering or leaving the system, and includes the work required to transfer material across the boundary. Kinetic and potential energy changes are neglected. If the only work done by the system is expansion work, one has

$$\delta w + \delta D_v = P\mathrm{d}V, \tag{2}$$

where δD_v is the differential of viscous dissipation of energy. Here the time derivative $\mathrm{d}D_v/\mathrm{d}t$ is the same quantity as denoted by E_v by Bird *et al.*(2)

3. Chemical Potential

From the definition of Gibbs (free) energy G and the first and second laws of thermodynamics for closed systems, one has for open systems in the absence of external force fields(10)

$$\mathrm{d}G = -S\mathrm{d}T + V\mathrm{d}P + \sum_{k=1}^{c} \mu_k \mathrm{d}n_k, \tag{3}$$

which is termed Gibbs' equation. The chemical potential μ_k is a state variable, and is defined by

$$\mu_k = \left(\frac{\partial G}{\partial n_k}\right)_{T,P,n_\ell} \tag{4}$$

where the subscript n_ℓ means that the number of moles of each species except species k is held constant. Hence μ_k may be considered as either a function of temperature, pressure, and composition, or a function of temperature, volume, and composition, so that

$$d\mu_k = -\overline{S}_k dT + \overline{V}_k dP + \sum_{\ell=1}^{c} \mu_{k\ell} dn_\ell \quad (k = 1, 2, \ldots, c), \tag{5}$$

or

$$d\mu_k = -\tilde{S}_k dT - \tilde{P}_k dV + \sum_{\ell=1}^{c} \mu'_{k\ell} dn_\ell \quad (k = 1, 2, \ldots, c), \tag{6}$$

since from Eqs. (3) and (4) and the definition of Helmholtz energy A, one obtains$^{(10)}$

$$\left(\frac{\partial \mu_k}{\partial T}\right)_{P,n} = -\overline{S}_k, \tag{7}$$

$$\left(\frac{\partial \mu_k}{\partial P}\right)_{T,n} = \overline{V}_k, \tag{8}$$

$$\left(\frac{\partial \mu_k}{\partial T}\right)_{V,n} = -\tilde{S}_k, \tag{9}$$

$$\left(\frac{\partial \mu_k}{\partial V}\right)_{T,n} = -\tilde{P}_k. \tag{10}$$

Also $\mu_{k\ell}$ and $\mu'_{k\ell}$ are defined by

$$\mu_{k\ell} = \left(\frac{\partial \mu_k}{\partial n_\ell}\right)_{T,P,n_m} = \left(\frac{\partial^2 G}{\partial n_k \partial n_\ell}\right)_{T,P,n_m}, \tag{11}$$

$$\mu'_{k\ell} = \left(\frac{\partial \mu_k}{\partial n_\ell}\right)_{T,V,n_m} = \left(\frac{\partial^2 A}{\partial n_k \partial n_\ell}\right)_{T,V,n_m} , \qquad (12)$$

in which the subscript n_m means that the number of moles of all species except k and ℓ is to be held constant, and $\tilde{S}_k$ and $\tilde{P}_k$, defined by$^{(9)}$

$$\tilde{S}_k = \left(\frac{\partial S}{\partial n_k}\right)_{T,V,n_\ell} , \qquad (13)$$

$$\tilde{P}_k = \left(\frac{\partial P}{\partial n_k}\right)_{T,V,n_\ell} , \qquad (14)$$

possess a formal similarity to partial molal quantities, but at constant volume instead of constant pressure.

4. Material Balance

Consider differential changes occurring in a homogeneous multicomponent open system. The differential change dn_k in the number of moles of species k consists of two parts: an internal part, $dn_k^{(i)}$, due to chemical reaction, and an external part, $dn_k^{(e)}$, due to matter crossing the boundary of the system. Suppose the system contains c species involved in r independent chemical reactions, which may be expressed in the form

$$\sum_{k=1}^{c} \nu_{ik} A_k = 0 \quad (i = 1, 2, \ldots, r), \qquad (15)$$

where ν_{ik} is the stoichiometric coefficient for species k in reaction i, and A_k represents the chemical symbol for species k. By convention ν_{ik} is taken to be positive for products, negative for reactants, and zero for species not contributing to reaction i. The internal contribution to the differential change in the number of moles of species k may then be written$^{(10)}$

$$\mathrm{d}n_k^{(i)} = \sum_{i=1}^{r} \nu_{ik}\mathrm{d}\xi_i,$$

where ξ_i is termed the molar extent of reaction i. Hence

$$\mathrm{d}n_k = \sum_{i=1}^{r} \nu_{ik}\mathrm{d}\xi_i + \mathrm{d}n_k^{(e)} \quad (k = 1, 2, \ldots, c). \qquad (16)$$

This equation is the basic differential material balance (conservation of mass) for an open system in which chemical reactions occur.

5. Chemical Equilibrium in Open Systems

The behavior of the system will be constrained to paths of chemical equilibrium, for which the necessary and sufficient condition is, for each reaction,$^{(10)}$

$$\sum_{k=1}^{c} \nu_{ik}\mu_k = 0 \quad (i = 1, 2, \ldots, r) \qquad [\mathrm{eq}]. \; (17)$$

Details of the derivation for open systems are given in the Appendix. Differentiating in general, Eq. (17) becomes

$$\sum_{k=1}^{c} \nu_{ik} d\mu_k = 0 \qquad (i = 1, 2, \ldots, r) \qquad [\mathrm{eq}]. \ (18)$$

This differential form of the condition for chemical equilibrium is convenient since it is linear in the chemical potentials, while an equivalent criterion, in terms of an equilibrium constant, is nonlinear in the activities of the species in the system.

6. Assumptions and Definitions

Certain assumptions and limitations are imposed when extending classical thermodynamics to open systems. The state functions E and S, established for closed systems, are assumed to apply to open systems as well. This assumption is strictly applicable only when describing equilibrium states. An equilibrium state can be rigorously maintained in an open system during the transfer of material across the boundary only if the material transferred is at the same state as the system, or if the rate of transfer is infinitesimal. Assume therefore that a locus of equilibrium states can be approached as closely as desired in an open system.

It is difficult to make an unambiguous distinction between heat crossing the boundary due to a temperature gradient and that flowing with material, and also between work done by the system to transfer material across the boundary and other work done by the system.(4) In this thesis, by definition, the heat *q* is restricted to that energy crossing the boundary of the system due to a temperature gradient. The work *w* is considered to be limited to work done by the system in addition to that required to transfer material across the boundary under the same conditions as in the system. Only "simple systems"(5) will be considered, for which the effect of external force fields, capillary action, distortion of solid phases, and internal adiabatic walls is neglected. If kinetic energy changes are also neglected, the work done by the system is expansion work only.

7. Scope of This Work

The independent relationships describing the behavior of open systems constrained to paths of chemical equilibrium, expressed by Eqs. (1), (5) or (6), (16), and (18), constitute the framework from which the effect of externally introduced composition perturbations may be calculated. General thermodynamic expressions will be obtained for the partial derivatives of the extent of chemical reaction with respect to external composition perturbations by a single species for multiple-reaction systems constrained to paths of chemical equilibrium under various conditions. Consideration will be given to heat and temperature effects, evaluation of derivatives of the extent of reaction for ideal solutions, and extension of the results to the general case of multicomponent perturbations.

B. ISOBARIC CONDITIONS

It is convenient to develop the composition derivatives first for isobaric conditions, by use of Eq. (5), and then similarly for isochoric conditions, using Eq. (6) instead. In this section the desired derivatives for isobaric conditions are derived, and the results are presented in Tables 1 and 2.

1. Derivation of Equations

From the definitions of Gibbs energy and partial molal quantities, one has

$$-\bar{S}_k = \frac{\mu_k - \bar{H}_k}{T}.$$

Substituting the resulting Eq. (5) into Eq. (18), the condition of chemical equilibrium, one obtains at constant pressure

$$\sum_{k=1}^{c} \nu_{ik}\left\{\frac{\mu_k - \bar{H}_k}{T}\mathrm{d}T + \sum_{\ell=1}^{c} \mu_{k\ell}\mathrm{d}n_\ell\right\} = 0 \quad (i = 1, 2, \ldots, r) \quad [P,\mathrm{eq}]. \quad (19)$$

Then, using Eqs. (16) and (17), Eq. (19) results in

$$\sum_{i=1}^{r} \sum_{k,\ell} \nu_{hk}\nu_{i\ell}\mu_{k\ell}\mathrm{d}\xi_i + \sum_{k,\ell} \nu_{hk}\mu_{k\ell}\mathrm{d}n_\ell^{(e)} = \frac{(\Delta H)_h}{T}\,\mathrm{d}T \quad (h = 1, 2, \ldots, r) \quad [P,\mathrm{eq}], \quad (20)$$

where $(\Delta H)_i$ is the molal "heat of reaction *i*," defined by$^{(9)}$

$$(\Delta H)_i = \sum_{k=1}^{c} \nu_{ik}\overline{H}_k. \tag{21}$$

One may similarly define

$$(\Delta V)_i = \sum_{k=1}^{c} \nu_{ik}\overline{V}_k. \tag{22}$$

The first law of thermodynamics, Eq. (1), in the absence of viscous dissipation, from Eq. (2), becomes

$$dE = \delta q - P dV + \sum_{k=1}^{c} \overline{H}_k^{(e)} dn_k^{(e)} \qquad [E_v]. \tag{23}$$

Now from the definition of enthalpy, at constant pressure, Eq. (23) results in

$$dH = \delta q + \sum_{k=1}^{c} \overline{H}_k^{(e)} dn_k^{(e)} \qquad [P,E_v]. \tag{24}$$

Considering the enthalpy as a function of temperature, pressure, and composition, one has at constant pressure

$$dH = nC_{P,n}dT + \sum_{k=1}^{c} \overline{H}_k dn_k \qquad [P], \tag{25}$$

where $C_{P,n}$ is the isobaric molal heat capacity of the system at constant composition, defined by$^{(7)}$

$$nC_{P,n} = \left(\frac{\partial H}{\partial T}\right)_{P,n} , \qquad (26)$$

and n is the total number of moles present. Equating the expressions for dH in Eqs. (24) and (25), and substituting from Eqs. (16) and (21), one obtains

$$\delta q = nC_{P,n}\mathrm{d}T + \sum_{i=1}^{r} (\Delta H)_i \mathrm{d}\xi_i + \sum_{k=1}^{c} \{\overline{H}_k - \overline{H}_k^{(e)}\}\mathrm{d}n_k^{(e)} \qquad [P,E_v]. \quad (27)$$

The independent equations, Eqs. (20) and (27), obtained from the condition of chemical equilibrium and Gibbs' equation, and from the first law of thermodynamics for isobaric conditions, respectively, may be divided by d$n_j^{(e)}$, with only species j crossing the boundary of the system, to yield

$$\sum_{i,k,\ell} \nu_{hk}\nu_{i\ell}\mu_{k\ell}\left(\frac{\partial \xi_i}{\partial n_j^{(e)}}\right)_{n_k^{(e)}} + \sum_k \nu_{hk}\mu_{jk} = \frac{(\Delta H)_h}{T}\left(\frac{\partial T}{\partial n_j^{(e)}}\right)_{n_k^{(e)}} \quad (h = 1, 2, \ldots, r) \; [P,\mathrm{eq}], \quad (28)$$

$$\left.\frac{\delta q}{\mathrm{d}n_j^{(e)}}\right|_{n_k^{(e)}} = nC_{P,n}\left(\frac{\partial T}{\partial n_j^{(e)}}\right)_{n_k^{(e)}} + \sum_i (\Delta H)_i \left(\frac{\partial \xi_i}{\partial n_j^{(e)}}\right)_{n_k^{(e)}} + \overline{H}_j - \overline{H}_j^{(e)} \qquad [P,E_v]. \quad (29)$$

The summations over i are from 1 to r, and those over k and ℓ are from 1 to c. The symbol δq has been retained to emphasize that it is not an exact differential.

2. Results

The partial derivatives of ξ_i, q, and T with respect to $n_j^{(e)}$ under isothermal-isobaric or adiabatic-isobaric conditions are obtained by solving Eqs. (28) and (29) simultaneously. The resulting equations for the derivatives of the chemical reaction parameter ξ_i may be expressed in the matrix form

$$\mathrm{Bz} = \mathrm{y}, \tag{30}$$

where z is a column matrix, the elements of which are the partial derivatives of ξ_i ($i = 1, 2, \ldots, r$). The set of desired derivatives may be obtained by inversion:

$$\mathrm{z} = \mathrm{B}^{-1}\mathrm{y}. \tag{31}$$

The elements of B and y for various cases are given in Table 1. Corresponding heat and temperature derivatives, from Eq. (29), are presented in Table 2.

For a single reaction, at constant temperature and pressure, Eq. (31) becomes

$$\left(\frac{\partial \xi}{\partial n_j^{(e)}}\right)_{T,P,n_k^{(e)}} = \frac{-\sum\limits_{k} \nu_k \mu_{jk}}{\sum\limits_{k,\ell} \nu_k \nu_\ell \mu_{k\ell}}\,, \tag{32}$$

where the subscripts h and i pertaining to the reaction have been dropped, since both these subscripts are equal to unity. For two chemical reactions, one obtains

TABLE 1.—Partial derivatives of the chemical reaction parameter ξ_i with respect to $n_j^{(e)}$ under various constraints.

Constraints	B_{hi}	y_{hj}
$[T,P]$	I	J
$[T,P,\mathrm{is}]$	I^*	J^*
$[q,P,E_v]$	$I + K(\Delta H)_i$	$J - K\{\overline{H}_j - \overline{H}_j^{(e)}\}$
$[q,P,E_v,\mathrm{is}]$	$I^* + K(\Delta H)_i$	$J^* - K\{\overline{H}_j - \overline{H}_j^{(e)}\}$
$[T,V]$	I'	J'
$[T,V,\mathrm{is}]$	$I^* + K_1(\Delta V)_i$	$J^* - K_1\overline{V}_j$
$[T,V,\mathrm{pg}]$	$I^* + K_2\nu_i$	$J^* - K_2$
$[q,V,E_v]$	$I' + K'(\Delta E)_i$	$J' - K'\{\tilde{E}_j - \overline{H}_j^{(e)}\}$
$[q,V,E_v,\mathrm{is}]$	$I^* + K'(\Delta E)_i$ $+ K_1(\Delta V)_i$	$J^* - K'\{\tilde{E}_j - \overline{H}_j^{(e)}\}$ $- K_1\overline{V}_j$
$[q,V,E_v,\mathrm{pg}]$	$I^* + K'(\Delta E)_i$ $+ K_2\nu_i$	$J^* - K'\{\tilde{E}_j - \overline{H}_j^{(e)}\}$ $- K_2$

Table 1—Continued

In each case $z = B^{-1}y$, where z is the column matrix of the set of derivatives of ξ_i $(i = 1, 2, \ldots, r)$ with respect to $n_j^{(e)}$. The elements of B and y are shown.

$$I = \sum_{k,\ell} \nu_{hk}\nu_{i\ell}\mu_{k\ell}$$

$$J = -\sum_{k} \nu_{hk}\mu_{jk}$$

$$I' = \sum_{k,\ell} \nu_{hk}\nu_{i\ell}\mu'_{k\ell}$$

$$J' = -\sum_{k} \nu_{hk}\mu'_{jk}$$

$$I^* = \frac{RT}{n}\left(\sum_{k} \frac{\nu_{hk}\nu_{ik}}{x_k} - \underset{\sim}{\nu}_h\underset{\sim}{\nu}_i\right)$$

$$J^* = \frac{RT}{n}\left(\underset{\sim}{\nu}_h - \frac{\nu_{hj}}{x_j}\right)$$

$$K = \frac{(\Delta H)_h}{nC_{P,n}T}$$

$$K_1 = (\Delta V)_h\left(\frac{\partial P}{\partial V}\right)_{T,n}$$

$$K' = \frac{(\Delta E)_h}{nC_{V,n}T}$$

$$K_2 = \frac{RT}{n}\,\underset{\sim}{\nu}_h$$

TABLE 2.—Heat and temperature effects from composition perturbations.*

$$\left.\frac{\delta q}{\mathrm{d}n_j^{(e)}}\right|_{T,P,n_k^{(e)}} = \sum_{i=1}^{r} (\Delta H)_i \left(\frac{\partial \xi_i}{\partial n_j^{(e)}}\right)_{T,P,n_k^{(e)}} + \overline{H}_j - \overline{H}_j^{(e)}$$

$$\left(\frac{\partial T}{\partial n_j^{(e)}}\right)_{q,P,n_k^{(e)}} = -\frac{1}{nC_{P,n}}\left\{\sum_{i=1}^{r} (\Delta H)_i \left(\frac{\partial \xi_i}{\partial n_j^{e}}\right)_{q,P,n_k^{(e)}} + \overline{H}_j - \overline{H}_j^{(e)}\right\}$$

$$\left.\frac{\delta q}{\mathrm{d}n_j^{(e)}}\right|_{T,V,n_k^{(e)}} = \sum_{i=1}^{r} (\Delta E)_i \left(\frac{\partial \xi_i}{\partial n_j^{(e)}}\right)_{T,V,n_k^{(e)}} + \tilde{E}_j - \overline{H}_j^{(e)}$$

$$\left(\frac{\partial T}{\partial n_j^{(e)}}\right)_{q,V,n_k^{(e)}} = -\frac{1}{nC_{V,n}}\left\{\sum_{i=1}^{r} (\Delta E)_i \left(\frac{\partial \xi_i}{\partial n_j^{(e)}}\right)_{q,V,n_k^{(e)}} + \tilde{E}_j - \overline{H}_j^{(e)}\right\}$$

---

*Viscous dissipation is neglected.

$$\left(\frac{\partial \xi_1}{\partial n_j^{(e)}}\right)_{T,P,n_k^{(e)}} = \frac{B_{22}y_{1j} - B_{12}y_{2j}}{B_{11}B_{22} - B_{12}B_{21}} \quad [T,P] \qquad (33)$$

$$= \frac{\sum\limits_{k,\ell,m} \nu_{2\ell}D_{km}\mu_{k\ell}\mu_{jm}}{\sum\limits_{k,\ell,m,n} \nu_{1k}\nu_{2m}D_{\ell n}\mu_{k\ell}\mu_{mn}} \; [T,P], \qquad (34)$$

where $D_{k\ell}$ may be conveniently defined in determinant form by

$$D_{k\ell} = \begin{vmatrix} \nu_{1k} & \nu_{1\ell} \\ \nu_{2k} & \nu_{2\ell} \end{vmatrix}. \qquad (35)$$

The corresponding derivative of ξ_2 is given by a similar expression, by interchanging the subscripts 1 and 2 wherever they appear in Eqs. (33), (34), and (35).

C. ISOCHORIC CONDITIONS

The development of the desired derivatives for isochoric conditions proceeds in a manner similar to that for isobaric conditions, but by using Eq. (6) instead of Eq. (5) for the differential of the chemical potentials.

From the definition of Helmholtz energy and Eq. (13), one may write

$$-\tilde{S}_k = \frac{\mu_k - \tilde{E}_k}{T}, \tag{36}$$

where

$$\tilde{E}_k = \left(\frac{\partial E}{\partial n_k}\right)_{T,V,n_\ell} . \tag{37}$$

Substitution of Eqs. (6) and (36) into Eq. (18), at constant volume, results in

$$\sum_{k=1}^{c} \nu_{ik}\left\{\frac{\mu_k - \tilde{E}_k}{T}\,\mathrm{d}T + \sum_{\ell=1}^{c} \mu'_{k\ell}\mathrm{d}n_\ell\right\} = 0$$

$$(i = 1,\ 2,\ \ldots,\ r) \qquad [V]. \tag{38}$$

Then from Eqs. (16) and (38), one obtains

$$\sum_{i,k,\ell} \nu_{hk}\nu_{i\ell}\mu'_{k\ell}\mathrm{d}\xi_i + \sum_{k,\ell} \nu_{hk}\mu'_{k\ell}\mathrm{d}n_\ell^{(e)} = \frac{(\Delta E)_h}{T}\,\mathrm{d}T$$

$$(h = 1,\ 2,\ \ldots,\ r) \qquad [V,\mathrm{eq}], \tag{39}$$

by the use of Eq. (17), where$^{(9)}$

$$(\Delta E)_i = \sum_{k=1}^{c} \nu_{ik}\tilde{E}_k . \tag{40}$$

Similarly, one may define(9)

$$(\Delta P)_i = \sum_{k=1}^{c} \nu_{ik}\tilde{P}_k. \tag{41}$$

Now considering the internal energy E as a function of temperature, volume, and composition, and using Eq. (37), one has at constant volume

$$dE = nC_{V,n}dT + \sum_{k=1}^{c} \tilde{E}_k dn_k \qquad [V], \tag{42}$$

where $C_{V,n}$ is the isochoric molal heat capacity of the system at constant composition, defined by(7)

$$nC_{V,n} = \left(\frac{\partial E}{\partial T}\right)_{V,n} . \tag{43}$$

From Eqs. (23) and (42), upon substitution of Eq. (16), one obtains, at constant volume,

$$\delta q = nC_{V,n}dT + \sum_{i=1}^{r} (\Delta E)_i d\xi_i + \sum_{k=1}^{c} \{\tilde{E}_k - \overline{H}_k^{(e)}\} dn_k^{(e)} \qquad [V,E_v]. \tag{44}$$

Now Eqs. (39) and (44) may be divided by $dn_j^{(e)}$, with only species j crossing the boundary of the system, to obtain

$$\sum_{i,k,\ell} \nu_{hk}\nu_{i\ell}\mu'_{k\ell}\left(\frac{\partial \xi_i}{\partial n_j^{(e)}}\right)_{n_k^{(e)}} + \sum_{k} \nu_{hk}\mu'_{jk}$$

$$= \frac{(\Delta E)_h}{T}\left(\frac{\partial T}{\partial n_j^{(e)}}\right)_{n_k^{(e)}} \qquad (h = 1, 2, \ldots, r) \qquad [V,\text{eq}], \tag{45}$$

$$\left.\frac{\delta q}{\mathrm{d}n_j^{(e)}}\right|_{n_k^{(e)}} = nC_{V,n}\left(\frac{\partial T}{\partial n_j^{(e)}}\right)_{n_k^{(e)}} + \sum_{i=1}^{r}(\Delta E)_i\left(\frac{\partial \xi_i}{\partial n_j^{(e)}}\right)_{n_k^{(e)}} + \tilde{E}_j - \overline{H}_j^{(e)} \qquad [V,E_v]. \quad (46)$$

The simultaneous solution of Eqs. (45) and (46), under either isothermal-isochoric or adiabatic-isochoric conditions, in much the same manner as for isobaric conditions, yields solutions of the form of Eq. (31) for the derivatives of ξ_i with respect to $n_j^{(e)}$. The solutions are given in Table 1, with the corresponding heat and temperature effects in Table 2.

D. IDEAL SOLUTIONS

Explicit evaluation of several of these formulas for a simple equation of state is informative. Define an ideal solution$^{(7)}$ by

$$\mu_k = \mu_k^\circ(T,P) + RT\,\ell\mathrm{n}\,x_k \qquad [\mathrm{is}], \quad (47)$$

for species k, where $\mu_k^\circ(T,P)$ is the chemical potential of pure species k and is a function of temperature and pressure alone, and x_k is the mole fraction of species k. Then

$$d\mu_k = \left(\frac{\partial\mu_k^\circ}{\partial T}\right)_P dT + \left(\frac{\partial\mu_k^\circ}{\partial P}\right)_T dP + d(RT\,\ell\mathrm{n}\,x_k) \qquad [\mathrm{is}], \quad (48)$$

and

$$\begin{aligned}\mu_{k\ell} &= RT\left(\frac{\partial\,\ell\mathrm{n}\,x_k}{\partial n_\ell}\right)_{T,P,n_m} \\ &= RT\left(\frac{\delta_{k\ell}}{n_k} - \frac{1}{n}\right) \qquad [\mathrm{is}], \quad (49)\end{aligned}$$

where $\delta_{k\ell}$ is the Kronecker delta function, defined by

$$\delta_{k\ell} = \begin{cases} 1 & k = \ell, \\ 0 & k \neq \ell. \end{cases} \qquad (50)$$

Also, from Eq. (47), one obtains

$$\left(\frac{\partial\mu_k}{\partial P}\right)_{T,n} = \left(\frac{\partial\mu_k^\circ}{\partial P}\right)_T \qquad [\mathrm{is}].$$

Then Eq. (48) becomes, at constant temperature,

$$d\mu_k = \overline{V}_k dP + RTd\ \ln x_k \qquad [T,\text{is}],$$

from Eq. (8). Hence

$$\mu'_{k\ell} = RT\left(\frac{\partial\ \ln x_k}{\partial n_\ell}\right)_{T,V,n_m} + \overline{V}_k\tilde{P}_\ell \qquad [\text{is}]$$

$$= RT\left(\frac{\delta_{k\ell}}{n_k} - \frac{1}{n}\right) + \overline{V}_k\tilde{P}_\ell \qquad [\text{is}],\ (51)$$

with the first term on the right side following in the same manner as Eq. (49). Then by substituting Eqs. (49) and (51) into the general terms B_{hi} and y_{hj}, as given in Table 1, one obtains the corresponding terms for ideal solutions, which terms are also shown in the table. In the case of a single reaction at constant temperature and pressure, Eq. (32) becomes

$$\left(\frac{\partial\xi}{\partial n_j^{(e)}}\right) = \frac{\underset{\sim}{\nu} - \frac{\nu_j}{x_j}}{\sum_{k=1}^{c}\frac{\nu_k^2}{x_k} - \underset{\sim}{\nu}^2} \qquad [\text{is}],$$

where $\underset{\sim}{\nu}$ is given by

$$\underset{\sim}{\nu} = \sum_{k=1}^{c}\nu_k,$$

and for two reactions, one has from Eq. (33)

$$\left(\frac{\partial \xi_1}{\partial n_j^{(e)}}\right)_{T,P,n_k^{(e)}} =$$

$$\frac{\left(\underset{\sim}{\nu}_1 - \frac{\nu_{1j}}{x_j}\right)\left(\sum_k \frac{\nu_{2k}^2}{x_k} - \underset{\sim}{\nu}_2^2\right) - \left(\underset{\sim}{\nu}_2 - \frac{\nu_{2j}}{x_j}\right)\left(\sum_k \frac{\nu_{1k}\nu_{2k}}{x_k} - \underset{\sim}{\nu}_1\underset{\sim}{\nu}_2\right)}{\left(\sum_k \frac{\nu_{1k}^2}{x_k} - \underset{\sim}{\nu}_1^2\right)\left(\sum_k \frac{\nu_{2k}^2}{x_k} - \underset{\sim}{\nu}_2^2\right) - \left(\sum_k \frac{\nu_{1k}\nu_{2k}}{x_k} - \underset{\sim}{\nu}_1\underset{\sim}{\nu}_2\right)^2}$$

[is], (52)

where $\underset{\sim}{\nu}_i$ is defined by

$$\underset{\sim}{\nu}_i = \sum_{k=1}^{c} \nu_{ik}. \qquad (53)$$

The result for the derivatives of ξ_2 is obtained similarly, with the subscripts 1 and 2 interchanged wherever they appear in Eq. (52). If an isochoric system also obeys the perfect gas law, the solutions are simplified further, and the results are given in Table 1.

E. MULTICOMPONENT COMPOSITION PERTURBATIONS

1. Extent of Reaction

These results may easily be extended to the case of any number of species entering or leaving the system simultaneously. Let

$$\dot{n}^{(e)} = \frac{dn^{(e)}}{dt} = \sum_{j=1}^{c} \dot{n}_j^{(e)},$$

and

$$x_j^{(e)} = \frac{\dot{n}_j^{(e)}}{\dot{n}^{(e)}},$$

where $x_j^{(e)}$ is defined as the net mole fraction of species j in the mass crossing the boundary of the system, and $\dot{n}_j^{(e)}$ is the molar rate of species j entering the system. Then

$$\sum_{j=1}^{c} x_j^{(e)} = 1.$$

Under the constraint to paths of chemical equilibrium, for example at constant temperature and pressure,

$$\dot{\xi}_i = \frac{d\xi_i}{dt} = \dot{\xi}_i\{n_1^{(e)}, n_2^{(e)}, \ldots, n_c^{(e)}\} \qquad [T,P].$$

Hence

$$\frac{d\xi_i}{dt} = \sum_{j=1}^{c} \left(\frac{\partial \xi_i}{\partial n_j^{(e)}}\right)_{n_k^{(e)}} \dot{n}_j^{(e)} \qquad [T,P],$$

and

$$\left(\frac{\partial \xi_i}{\partial n^{(e)}}\right)_{T,P} = \sum_{j=1}^{c} x_j^{(e)} \left(\frac{\partial \xi_i}{\partial n_j^{(e)}}\right)_{T,P,n_k^{(e)}} \quad . \qquad (54)$$

Similar equations may be written for adiabatic-isobaric, isothermal-isochoric, and adiabatic-isochoric conditions, by simply replacing the subscripts T and P in Eq. (54) by q and P, T and V, or q and V, respectively.

2. Heat and Temperature Effects

Heat and temperature effects from multicomponent composition perturbations may also be expressed as the sum of the effects from individual species, as in Table 2, each weighted by the net mole fraction $x_j^{(e)}$ of that species in the material crossing the boundary.

As an example, for isobaric conditions, dividing Eq. (27) by dt, one has

$$\dot{q} = nC_{P,n}\dot{T} + \sum_{i=1}^{r} (\Delta H)_i \dot{\xi}_i + \sum_{k=1}^{c} \{\overline{H}_k - \overline{H}_k^{(e)}\} \dot{n}_k^{(e)} \quad [P].$$

Then dividing by $\dot{n}^{(e)}$ results in

$$\frac{\delta q}{\mathrm{d}n^{(e)}} = nC_{P,n}\frac{\mathrm{d}T}{\mathrm{d}n^{(e)}} + \sum_{i=1}^{r} (\Delta H)_i \frac{\mathrm{d}\xi_i}{\mathrm{d}n^{(e)}} + \sum_{k=1}^{c} x_k^{(e)} \{\overline{H}_k - \overline{H}_k^{(e)}\} \quad [P]. \qquad (55)$$

At constant temperature and pressure, one has

$$\left.\frac{\delta q}{dn^{(e)}}\right|_{T,P} = \sum_{i=1}^{r} (\Delta H)_i \left(\frac{\partial \xi_i}{\partial n^{(e)}}\right)_{T,P} + \sum_{k=1}^{c} x_k^{(e)} \{\overline{H}_k - \overline{H}_k^{(e)}\} .$$

Then upon substituting from Eq. (54) and the first line of Table 2, one obtains

$$\left.\frac{\delta q}{dn^{(e)}}\right|_{T,P} = \sum_{j=1}^{c} x_j^{(e)} \left.\frac{\delta q}{n_j^{(e)}}\right|_{T,P,n_k^{(e)}} . \qquad (56)$$

Now from Eq. (55), one has also

$$\left(\frac{\partial T}{\partial n^{(e)}}\right)_{q,P} = -\frac{1}{nC_{P,n}}\left\{\sum_{i=1}^{r} (\Delta H)_i \left(\frac{\partial \xi_i}{\partial n^{(e)}}\right)_{q,P} + \sum_{k=1}^{c} x_k^{(e)} \{\overline{H}_k - \overline{H}_k^{(e)}\}\right\}.$$

Substitution from Eq. (54) for adiabatic-isobaric conditions and from the second line of Table 2 results in

$$\left(\frac{\partial T}{\partial n^{(e)}}\right)_{q,P} = \sum_{j=1}^{c} x_j^{(e)} \left(\frac{\partial T}{\partial n_j^{(e)}}\right)_{q,P,n_k^{(e)}} . \qquad (57)$$

Equations analogous to Eqs. (56) and (57), for isochoric conditions, may be derived in a similar manner, by starting with Eq. (44) instead of Eq. (27).

3. Vanishing of Reaction Derivatives

Now if

$$x_j^{(e)} = x_j \quad (j = 1, 2, \ldots, c), \tag{58}$$

Eq. (54) becomes

$$\left(\frac{\partial \xi_i}{\partial n^{(e)}}\right)_{T,P} = \sum_{j=1}^{c} x_j \left(\frac{\partial \xi_i}{\partial n_j^{(e)}}\right)_{T,P,n_k^{(e)}} . \tag{59}$$

The solutions of Eq. (31) for the derivatives in the summation of Eq. (59), at constant temperature and pressure, may be written in the form

$$\left(\frac{\partial \xi_i}{\partial n_j^{(e)}}\right)_{T,P,n_k^{(e)}} = \frac{|\mathrm{B}_i'|}{|\mathrm{B}|}, \tag{60}$$

where $|\mathrm{B}|$ is the determinant having elements B_{hi}, which do not depend on species j, and $|\mathrm{B}_i'|$ is the substituted determinant with the ith column replaced by the elements y_{hj}. Each term in the numerator of Eq. (60) contains a single factor of the y_{hj} type, which depends on species j. This factor must be taken into account in performing the summation in Eq. (59), which then becomes

$$\begin{aligned}\left(\frac{\partial \xi_i}{\partial n^{(e)}}\right)_{T,P} &= \sum_{j=1}^{c} \frac{n_j}{n|\mathrm{B}|} \sum_{h=1}^{r} y_{hj} |\mathrm{B}_{hi}''| \\ &= \frac{1}{n|\mathrm{B}|} \sum_{h=1}^{r} \left(\sum_{j=1}^{c} n_j y_{hj}\right) |\mathrm{B}_{hi}''|, \end{aligned} \tag{61}$$

where $|\mathrm{B}_{hi}''|$ is the cofactor of y_{hj} in $|\mathrm{B}_i'|$, and is the same as the cofactor of B_{hi} in $|\mathrm{B}|$.

Now considering the Gibbs energy G as a function of temperature, pressure, and composition, one may write$^{(10)}$

$$G = \sum_{k=1}^{c} n_k \mu_k .$$

Differentiating with respect to n_ℓ, holding temperature, pressure, and the number of moles of all other species constant, one obtains

$$\sum_{k=1}^{c} n_k \mu_{k\ell} = 0 \quad (\ell = 1, 2, \ldots, c), \tag{62}$$

which is a fundamental relationship between the derivatives of the chemical potential for any thermodynamic system, not dependent on the laws of thermodynamics or on the condition of chemical equilibrium.

Multiplying y_{hj} for isothermal-isobaric conditions by n_j and summing over all the species j, one obtains

$$\begin{aligned}\sum_{j=1}^{c} n_j y_{hj} &= - \sum_{j,k} n_j \nu_{hk} \mu_{jk} \quad (h = 1, 2, \ldots, r) \\ &= - \sum_{k=1}^{c} \nu_{hk} \sum_{j=1}^{c} n_j \mu_{jk} \\ &= 0 \quad (h = 1, 2, \ldots, r),\end{aligned} \tag{63}$$

from Eq. (62), by setting $\ell = j$, and since $\mu_{k\ell} = \mu_{\ell k}$. Now substitution of Eq. (63) into Eq. (61) yields

$$\left(\frac{\partial \xi_i}{\partial n^{(e)}}\right)_{T,P} = 0,$$

under the condition of Eq. (58). This condition holds in general for a stream entering or leaving the system when the stream has the same composition as the system. Under conditions other than constant temperature and pressure, stricter constraints are necessary in order that derivatives of the form given on the left side of Eq. (54) vanish. Constraints which are sufficient for this purpose may be obtained in a manner similar to that given above for isothermal-isobaric conditions and are given in Table 3.

TABLE 3.—Constraints to assure the vanishing of the derivatives of ξ_i with respect to $n^{(e)}$.

Equation	Sufficient Constraints
$\left(\frac{\partial \xi_i}{\partial n^{(e)}}\right)_{T,P} = 0$	$x_j^{(e)} = x_j \quad (j = 1, 2, \ldots, c)$
$\left(\frac{\partial \xi_i}{\partial n^{(e)}}\right)_{q,P} = 0$	$x_j^{(e)} = x_j \quad (j = 1, 2, \ldots, c)$, $[E_v]$, $\sum_{j=1}^{c} n_j \overline{H}_j^{(e)} = H$
$\left(\frac{\partial \xi_i}{\partial n^{(e)}}\right)_{T,V} = 0$	$x_j^{(e)} = x_j \quad (j = 1, 2, \ldots, c)$, $(\Delta P)_h = 0 \quad (h = 1, 2, \ldots, r)$
$\left(\frac{\partial \xi_i}{\partial n^{(e)}}\right)_{q,V} = 0$	$x_j^{(e)} = x_j \quad (j = 1, 2, \ldots, c)$, $[E_v]$, $\frac{(\Delta E)_h}{C_{V,n}T} \sum_{j=1}^{c} x_j \{\tilde{E}_j - \overline{H}_j^{(e)}\}$ $+ V(\Delta P)_h = 0 \quad (h = 1, 2, \ldots, r)$

F. DISCUSSION

Possible applications of the derivatives of extent of reaction with respect to external composition perturbations and the corresponding heat and temperature derivatives include (1) behavior of multicomponent systems in open tanks, reservoirs, and reaction vessels; (2) error analysis for composition of material entering or leaving the system; and (3) correlation of yield data and optimization of yield.

The condition of chemical equilibrium in terms of an equilibrium constant, while equivalent to Eq. (17), is nonlinear in the activities of the species in the system. Equilibrium conditions appear as separate nonlinear constraints on the variables in the system. Implementation of Lagrange multipliers or the method of gradients(6) to obtain a numerical result becomes tedious for complex reaction systems.

On the other hand, the condition of chemical equilibrium in the form of Eq. (17) is linear in the chemical potentials, hence the set of derivatives of extent of reaction is obtained by simple matrix inversion, as Eq. (31). Since the condition of chemical equilibrium is included in Eq. (31), no separate constraints on the variables in the system appear, and yield as a function of input composition may be easily optimized numerically by the method of gradients.

It may be more convenient to make use of the fugacities f_k than the derivatives $\mu_{k\ell}$ and $\mu'_{k\ell}$. For the molal chemical potential, μ_k, one has

$$\mu_k = \mu_k^*(T) + RT\ \ell n\ f_k, \tag{64}$$

where $\mu_k^*(T)$ is the chemical potential of species k at unit fugacity and is a function of temperature alone.(10) Hence Eqs. (11) and (12) become

$$\mu_{k\ell} = RT\left(\frac{\partial\ \ell n\ f_k}{\partial n_\ell}\right)_{T,P,n_m},$$

$$\mu'_{k\ell} = RT\left(\frac{\partial\ \ell n\ f_k}{\partial n_\ell}\right)_{T,V,n_m},$$

which may then be obtained from PVT data.

Examination of Table 1 illustrates some features of the various derivatives of ξ_i. The elements B_{hi} and y_{hj} of the determinants form the solutions for the derivatives. The first terms of these elements, designated by I and J, respectively, are similar for all four types of constraints. In the case of ideal solutions, the dependence upon the derivatives $\mu_{k\ell}$ of the chemical potential vanish, and the first terms of the elements, I^* and J^*, are identical in all four cases. The added terms, designated by K as factors, depend upon the constraints on the system. In practice, a system at constant pressure, for example, may be neither isothermal nor adiabatic. By considering the K terms as perturbations upon the I and J terms, when this assumption is justifiable, approximations to the derivatives of ξ_i may be obtained for a range of conditions at chemical equilibrium, but not corresponding exactly to any of the four types of constraints for which the derivatives are solved explicitly.

A logical extension of this analysis would be to the steady state. The general equilibrium criterion of maximum entropy, with zero entropy production, may be replaced in the steady state near equilibrium by the criterion of minimum entropy production.(11) Topics for future investigation might include the corresponding analysis of steady-state chemical systems under near-equilibrium conditions, the effect of external composition perturbations in steady-state chemical systems, the range of validity of linear phenomenological relations in the steady state, and the combination of linearized kinetic expressions(9) with thermodynamic relationships to describe perturbations of the steady state.

NOMENCLATURE

A	Total Helmholtz energy, $E - TS$
A_k	Chemical symbol for species k, in Eq. (15)
A_i	Affinity of reaction i, given by Eq. (A.12)
B	Matrix in Eq. (30), having elements given in Table 1
$C_{P,n}$	Isobaric molal heat capacity of the system at constant composition, defined by Eq. (26)
$C_{V,n}$	Isochoric molal heat capacity of the system at constant composition, defined by Eq. (43)
c	Number of chemical species in the system
$D_{k\ell}$	Variable defined by Eq. (35)
D_v	Viscous dissipation of energy
d	Differential operator
E	Total internal energy
$\tilde{E}_k$	Variable defined by Eq. (37)
$(\Delta E)_i$	Variable defined by Eq. (40)
f_k	Fugacity of species k, defined by Eq. (64)
G	Total Gibbs energy, $H - TS$
H	Total enthalpy
$\overline{H}_k$	Partial molal enthalpy of species k
$(\Delta H)_i$	Molal "heat of reaction i," defined by Eq. (21)
n	Total number of moles in the system
n_k	Number of moles of species k
P	Pressure
$\tilde{P}_k$	Variable defined by Eq. (14)
$(\Delta P)_i$	Variable defined by Eq. (41)

q	Energy crossing the boundary of the system due to a temperature gradient
R	Molar gas constant
r	Number of independent chemical reactions occurring in the system
S	Total entropy
$\overline{S}_k$	Partial molal entropy of species k
$\tilde{S}_k$	Variable defined by Eq. (13)
T	Absolute temperature
t	Time
V	Total volume
$\overline{V}_k$	Partial molal volume of species k
$(\Delta V)_i$	Variable defined by Eq. (22)
w	Net work done by the system
x_k	Mole fraction of species k
y	Column matrix in Eq. (30), having elements given in Table 1
z	Column matrix of partial derivatives of ξ_i with respect to $n_j^{(e)}$ $(i = 1, 2, \ldots, r)$
δ	Variation operator
$\delta_{k\ell}$	Kronecker delta function, defined by Eq. (50)
μ_k	Chemical potential of species k, defined by Eq. (4)
$\mu_{k\ell}, \mu'_{k\ell}$	Variables defined by Eqs. (11) and (12)
$\underset{\sim}{\nu}_i$	Variable defined by Eq. (53)
ν_{ik}	Stoichiometric coefficient for species k in reaction i
ξ_i	Molar extent of reaction i
$\sum$	Summation operator

Subscripts

h, i	Denote chemical reactions
irr	Irreversible process
j, k, ℓ, m, n	Denote chemical species
n	Constant number of moles of all species
$n_k^{(e)}$	Restriction to species j only, crossing the boundary
P	Constant pressure
q	Adiabatic conditions
T	Constant temperature
V	Constant volume

Superscripts

(e)	Denotes material crossing the boundary of the system
(i)	Denotes occurrence due to chemical species in the system
$^{\circ}$	Denotes pure species
—	Denotes partial molal quantities
~	Denotes partial differentiation with respect to n_k, at constant temperature and volume
•	Denotes differentiation with respect to time

Brackets

$[E_v]$	No viscous dissipation
[eq]	Constraint to paths of chemical equilibrium
[is]	Ideal solution behavior
$[P]$	Constant pressure process

[pg]	Perfect gas behavior
[q]	Adiabatic process
[T]	Constant temperature process
[V]	Constant volume process

REFERENCES

1. Aris, R., *Introduction to the Analysis of Chemical Reactors*, Prentice-Hall, Englewood Cliffs, N. J. (1965), p. 41.

2. Bird, R. B., W. E. Stewart, and E. N. Lightfoot, *Transport Phenomena*, Wiley, New York (1960), p. 214.

3. Denbigh, K. G., *The Thermodynamics of the Steady State*, Wiley, New York (1951), p. 42.

4. ————, *The Principles of Chemical Equilibrium*, 2d ed., University Press, Cambridge (1966), p. 81.

5. Hatsopoulos, G. N., and J. H. Keenan, *Principles of General Thermodynamics*, Wiley, New York (1965).

6. Kelley, H. J., "Method of Gradients," in *Optimization Techniques: with Applications to Aerospace Systems*, ed. by G. Leitman, in *Mathematics in Science and Engineering*, ed. by R. Bellman, Vol. 5, Academic Press, New York (1962), Chap. 6.

7. Lewis, G. N., and M. Randall, *Thermodynamics*, 2d ed., rev. by K. S. Pitzer and L. Brewer, McGraw-Hill, New York (1961).

8. Moore, W. J., *Physical Chemistry*, 3d ed., Prentice-Hall, Englewood Cliffs, N. J. (1962), p. 171.

9. Pings, C. J., "Thermodynamics of Chemical Equilibrium—I. Effect of Temperature and Pressure," *Chem. Eng. Sci. 16*, 181 (1961); "II. Effect of Volume, Entropy, and Enthalpy," *ibid. 18*, 671 (1963); "III. Effect of Heat, Work, and Viscous Dissipation," *ibid. 21*, 693 (1966); C. S. Lu and C. J. Pings, "IV. Dynamic Response of Near-Equilibrium Systems," *ibid.*, p. 697; A. H. Larsen, C. S. Lu, and C. J. Pings, "V. Multiple-Reaction and Open Systems," *ibid. 23*, 289 (1968).

10. Prigogine, I., and R. Defay, *Chemical Thermodynamics*, trans. by D. H. Everett, Wiley, New York (1954).

11. ————, *Introduction to Thermodynamics of Irreversible Processes*, 2d ed., Interscience, New York (1961), p. 75.

12. Sage, B. H., *Thermodynamics of Multicomponent Systems*, Reinhold, New York (1965), p. 47.

13. Tolman, R. C., *The Principles of Statistical Mechanics*, Oxford University Press, London (1938), p. 163.

14. Van Ness, H. C., *Classical Thermodynamics of Non-Electrolyte Solutions*, Macmillan, New York (1964), p. 117.

15. Waser, J., *Basic Chemical Thermodynamics*, Benjamin, New York (1966), p. 160.

APPENDIX

THE CONDITION FOR CHEMICAL EQUILIBRIUM

For an open system, the second law of thermodynamics may be written

$$T\mathrm{d}S = \delta q + T\mathrm{d}S_{\mathrm{irr}}^{(i)} + T\mathrm{d}S^{(e)}, \tag{A.1}$$

where $\mathrm{d}S_{\mathrm{irr}}^{(i)}$ and $\mathrm{d}S^{(e)}$ are internal and external contributions to the differential change in the entropy of the system, due to irreversible chemical reaction and viscous dissipation, and to material entering or leaving under different conditions than in the system. Assuming that a state function S, established for closed systems, also exists for open systems, $\mathrm{d}S$ is a perfect differential. The internal entropy generation term $\mathrm{d}S_{\mathrm{irr}}^{(i)}$ must be positive. For the case of a single chemical reaction, write

$$T\mathrm{d}S_{\mathrm{irr}}^{(i)} = \mathrm{A}\mathrm{d}\xi + \delta D_v \geqq 0, \tag{A.2}$$

where A is called the affinity and $\mathrm{d}\xi$ is the differential extent of the reaction. In the absence of viscous dissipation, Eq. (A.2) is essentially a defining relationship for the affinity A.

Define *chemical equilibrium* by the vanishing of the chemical reaction part of $\mathrm{d}S_{\mathrm{irr}}^{(i)}$, independent of the existence of irreversibilities introduced by viscous dissipation or by material entering or leaving the system. Then

$$\mathrm{A}d\xi = 0 \qquad [\mathrm{eq}], \quad (\mathrm{A}.3)$$

where [eq] indicates constraint to paths of chemical equilibrium. Since the variation dξ is arbitrary, the necessary and sufficient condition for chemical equilibrium becomes

$$\mathrm{A} = 0 \qquad [\mathrm{eq}], \quad (\mathrm{A}.4)$$

which means that the driving force for the reaction is zero. If A $\neq$ 0 even though no reaction occurs (dξ = 0), the system is not at stable equilibrium. For example, a mixture of hydrogen and oxygen at room temperature does not react at any measurable rate, but may be made to do so by introducing a spark or a suitable catalyst, such as spongy platinum. The spontaneous reaction is highly irreversible, showing that $dS^{(i)}_{\mathrm{irr}}$ is positive or, from Eq. (A.2), that A $\neq$ 0. The unreactive mixture is not at true chemical equilibrium, but its state is termed "metastable equilibrium"(5) or "false equilibrium."(10)

Now Eqs. (A.1) and (A.2) result in

$$TdS = \delta q + \mathrm{A}d\xi + \delta D_v + TdS^{(e)}. \qquad (\mathrm{A}.5)$$

Substituting Eq. (1), the first law of thermodynamics for an open system, Eq. (2), and Eq. (A.5) into the differential of Gibbs energy, one obtains

$$dG = -SdT + VdP - \mathrm{A}d\xi + \sum_{k=1}^{c} \overline{H}_k^{(e)} dn_k^{(e)} - TdS^{(e)}. \quad (\mathrm{A}.6)$$

Since $dS^{(e)}$ depends on material crossing the boundary, and vanishes for $dn_k^{(e)} = 0$ $(k = 1, 2, \ldots, c)$, one has

$$A = -\left(\frac{\partial G}{\partial \xi}\right)_{T,P,n^{(e)}}, \tag{A.7}$$

where the subscript $n^{(e)}$ indicates that the differentiation is performed under the constraint of no material crossing the boundary. Hence the affinity of a chemical reaction in an open system is the same as that in a closed system. If the Gibbs energy G is considered to be a function of temperature, pressure, and the numbers of moles of all species present, Eq. (A.7) becomes

$$A = -\sum_{k=1}^{c} \left(\frac{\partial G}{\partial n_k}\right)_{T,P,n_\ell} \left(\frac{\partial n_k}{\partial \xi}\right)_{T,P,n^{(e)}} . \tag{A.8}$$

Now Eq. (16) for a single reaction is

$$dn_k = \nu_k d\xi + dn_k^{(e)}, \tag{A.9}$$

so that

$$\left(\frac{\partial n_k}{\partial \xi}\right)_{n^{(e)}} = \nu_k, \tag{A.10}$$

and substitution of Eqs. (4) and (A.10) into Eq. (A.8) results in

$$A = -\sum_{k=1}^{c} \nu_k \mu_k . \tag{A.11}$$

Invoking the principle of microscopic reversibility,(13) each reaction in a multiple-reaction system at chemical equilibrium must be independently in equilibrium. Therefore from Eqs. (A.4) and (A.11), it follows that

$$A_i = -\sum_{k=1}^{c} \nu_{ik}\mu_k = 0 \quad (i = 1, 2, \ldots, r) \quad [\mathrm{eq}], \tag{A.12}$$

from which one has Eq. (17), the condition of chemical equilibrium, valid for open systems.

Using an alternate derivation, some authors(1,8,14,15) imply that Eq. (17) is restricted to conditions of constant temperature and pressure, and fail to indicate its generality. From the above development, with the condition of chemical equilibrium defined by Eq. (A.3), such a restriction is unnecessary. In fact Eq. (17) is valid for *arbitrary* thermal and mechanical variations in an open system.

Further information can now be obtained regarding the nature of $\mathrm{d}S^{(e)}$, the contribution to $\mathrm{d}S$ due to material crossing the boundary. Equating two independent expressions for $\mathrm{d}G$, from Eqs. (3) and (A.6), one has

$$\sum_{k=1}^{c} \mu_k \mathrm{d}n_k = -\mathrm{A}\mathrm{d}\xi + \sum_{k=1}^{c} \overline{H}_k^{(e)} \mathrm{d}n_k^{(e)} - T\mathrm{d}S^{(e)}.$$

Substituting Eqs. (A.9) and (A.11) results in

$$Td S^{(e)} = \sum_{k=1}^{c} \{\overline{H}_k^{(e)} - \mu_k\} dn_k^{(e)} . \tag{A.13}$$

Now

$$\mu_k = \overline{H}_k - T\overline{S}_k ,$$

and

$$\mu_k^{(e)} = \overline{H}_k^{(e)} - T^{(e)}\overline{S}_k^{(e)} ,$$

where $\mu_k^{(e)}$ and $\overline{S}_k^{(e)}$ are the chemical potential and partial molal entropy, respectively, of species k crossing the boundary, and $T^{(e)}$ is the temperature of the material crossing the boundary. Hence Eq. (A.13) becomes

$$\begin{aligned} Td S^{(e)} &= T \sum_{k=1}^{c} \overline{S}_k dn_k^{(e)} + \sum_{k=1}^{c} \{\overline{H}_k^{(e)} - \overline{H}_k\} dn_k^{(e)} \\ &= T^{(e)} \sum_{k=1}^{c} \overline{S}_k^{(e)} dn_k^{(e)} + \sum_{k=1}^{c} \{\mu_k^{(e)} - \mu_k\} dn_k^{(e)} . \end{aligned}$$

Therefore $dS^{(e)}$ includes a term to account for the differences between either partial molal enthalpies or chemical potentials of the species crossing the boundary and those in the system, depending on whether the partial molal entropy terms are described in terms of properties of the system or of properties of the material crossing the boundary.

PROPOSITIONS

Proposition I

Equations involving arbitrary variable-volume and flow-rate policies in isothermal stirred-tank chemical reactors may be solved explicitly for the component concentrations for any system of homogeneous first-order reactions, however complex.

Elimination of the restriction to constant volume for stirred-tank chemical reactors provides additional flexibility of approach for economic optimization. Constant mean residence times, as in steady-state systems, are replaced by variable pseudo residence times. Both batch and steady-state reactors are special cases of the variable-volume problem. Consider homogeneous reaction systems with negligible volume change due to reaction. For first-order "non-coupled" reactions, the concentrations of reactive species, as a function of time, may be explicitly obtained, either analytically in closed form or in terms of numerically evaluated integrals. For isothermal variable-volume reactors, by the technique of matrix diagonalization, any system of first-order reactions of arbitrary complexity may be transformed to one of simultaneous, independent, irreversible first-order reactions, for which the solution is immediate.

Consider a general stirred-tank chemical reactor, as shown in Fig. 1, with input and output volumetric flow rates $Q^\circ(t)$ and $Q(t)$ and input and output molar concentrations $c_j^\circ(t)$ and $c_j(t)$, respectively, of species j. These variables and the volume $V(t)$ of the reacting mixture are in general functions of time. The volume and flow rates are related by

$$\frac{dV}{dt} = Q^\circ - Q + \left(\frac{\partial V}{\partial t}\right)_{\text{reactions}} .$$

A set of R simultaneous chemical reactions involving S species may be expressed by

$$\sum_{j=1}^{S} \nu_{ij} A_j = 0 \quad (i = 1, 2, \ldots, R),$$

where A_j represents the chemical symbol for species j, and the stoichiometric coefficients ν_{ij} are positive for products, negative for reactants, and zero for species not participating in reaction i. A common form of rate dependence for the forward and backward reactions is

$$r_i = k_i \prod_{j=1}^{S} c_j^{\alpha_{ij}} \quad (i = 1, 2, \ldots, R),$$

and

$$r_i' = k_i' \prod_{j=1}^{S} c_j^{\beta_{ij}} \quad (i = 1, 2, \ldots, R),$$

respectively, where k_i and k_i' are the rate constants, in

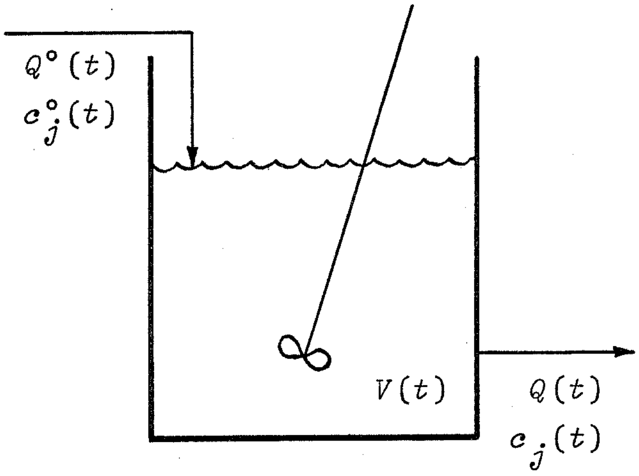


FIGURE 1.—Schematic diagram of a general stirred-tank reactor, to illustrate the nomenclature involved.

general functions of temperature and hence of time, but independent of c_j; and α_{ij} and β_{ij} are experimentally determined constants, defined$^{(1)}$ as the orders of the forward and backward i reactions, respectively, with respect to species j. The equilibrium constant in terms of concentrations is defined by$^{(2)}$

$$K_{c_i} = \prod_{j=1}^{S} c_j^{\nu_{ij}} \quad (i = 1,\ 2,\ \ldots,\ R);$$

and from the principle of microscopic reversibility,

$$r_i = r_i' \quad (i = 1,\ 2,\ \ldots,\ R),$$

and

$$K_{c_i} = \frac{k_i}{k_i'} \quad (i = 1,\ 2,\ \ldots,\ R).$$

If the order and the molecularity of the reactions coincide, it follows that one may write without loss of generality

$$\nu_{ij} = \alpha_{ij} - \beta_{ij}, \tag{1}$$

where $\alpha_{ij} = |\nu_{ij}|$ for reactants, $\beta_{ij} = |\nu_{ij}|$ for products, and both are zero otherwise, since the stoichiometric coefficients are arbitrary to within a multiplicative constant. Then the "overall reaction rate" R_j may be expressed as

$$R_j = \sum_{i=1}^{R} \nu_{ij} \left(k_i \prod_{k=1}^{S} c_k^{\alpha_{ik}} - k_i' \prod_{k=1}^{S} c_k^{\beta_{ik}} \right) \quad (j = 1,\ 2,\ \ldots,\ S), \tag{2}$$

or

$$R_j = \sum_{i=1}^{R} \nu_{ij} \left(k_i \prod_{\substack{k=1 \\ \nu_{ik}<0}}^{S} c_k^{|\nu_{ik}|} - k_i' \prod_{\substack{k=1 \\ \nu_{ik}>0}}^{S} c_k^{\nu_{ik}} \right)$$

$$(j = 1, 2, \ldots, S).$$

Now a material balance on the reactor yields

$$\frac{\mathrm{d}}{\mathrm{d}t}(c_j V) = Q^\circ c_j^\circ - Qc_j - VR_j,$$

and substitution from Eq. (1) results in

$$\frac{\mathrm{d}c_j}{\mathrm{d}t} = \frac{Q^\circ}{V}(c_j^\circ - c_j) - R_j - \frac{c_j}{V}\left(\frac{\partial V}{\partial t}\right)_{\text{reactions}}, \qquad (3)$$

where $Q^\circ(c_j^\circ - c_j)/V$ is the time rate of change of concentration due to flow, and the remaining terms on the right side of Eq. (3) represent the change of concentration due to reaction. Define a *pseudo residence time* $\theta(t)$ by

$$\theta(t) = \frac{V(t)}{Q^\circ(t)}. \qquad (4)$$

It follows that $\theta(t)$ is equal to the mean residence time for a steady-state system. Make the following assumptions: complete mixing (uniform c_j, T); mixing time negligible in comparison with $\theta(t)$; negligible volume change due to reaction; first-order "non-coupled" reactions, so that the rate equation for species j does not involve unknown concentrations of other species, and the equations can be solved independently, or at least successively. Then Eq. (3) becomes

$$\frac{dc_j}{dt} = \frac{c_j^\circ - c_j}{\theta} - R_j, \tag{5}$$

where R_j can be expressed in the form

$$R_j = \phi_j c_j + \psi_j. \tag{6}$$

Here ϕ_j and ψ_j are, in general, functions of time but not of c_j. Upon multiplying Eq. (5) by the integrating factor

$$F_j = \exp\left[\int_0^t \left(\frac{1}{\theta} + \phi_j\right) dt\right], \tag{7}$$

its general solution is

$$c_j(t) = F_j^{-1}\left[c_{j0} + \int_0^t \left(\frac{c_j}{\theta} - \psi_j\right) F_j dt\right], \tag{8}$$

including the initial condition $c_j(0) = c_{j0}$. Examples of the explicit solution, Eq. (8), are given in Table 1.

For a single irreversible reaction, Eqs. (7) and (8) are simplified to the usual results(2,3) for $c_1(t)$ under special conditions:

1. Isothermal batch reaction:

$$F_1 = e^{k_1 t},$$

$$c_1(t) = c_{10} e^{-k_1 t}.$$

TABLE 1.—Solutions for non-coupled first-order reaction systems in a general stirred-tank reactor.

Reaction	Equations	j	ϕ_j	ψ_j
Irreversible	$A_1 \xrightarrow{k_1}$ prod.	1	k_1	0
Pseudo first-order	$A_1+A_2 \xrightarrow{k_1}$ prod. ($c_2(t)$ known)	1	k_1c_2	0
Reversible, constant total conc.	$A_1 \overset{k_1}{\rightleftarrows} A_2$	1	k_1+k_1'	$-k_1'(c_{10}+c_{20})$
		2	k_1+k_1'	$-k_1(c_{10}+c_{20})$
Consecutive	$A_1 \xrightarrow{k_1} A_2 \xrightarrow{k_2} A_3$	1	k_1	0
		2	k_2	$-k_1c_1$
		3	0	$-k_2c_2$
Competing	$\left.\begin{array}{l} A_1+A_2 \xrightarrow{k_1} A_3 \\ A_1+A_3 \xrightarrow{k_2} A_4 \\ (c_1(t)\text{ known}) \end{array}\right\}$	2	k_1c_1	0
		3	k_2c_1	$-k_1c_1c_2$
		4	0	$-k_2c_1c_3$

$$\frac{dc_j}{dt} = \frac{c_j^\circ - c_j}{\theta} - R_j \qquad R_j = \phi_j c_j + \psi_j$$

$$c_j(t) = F_j^{-1}(t)\left[c_{j0} + \int_0^t \left(\frac{c_j^\circ}{\theta} - \psi_j\right) F_j(t)\,dt \right]$$

$$F_j(t) = \exp\left[\int_0^t \left(\frac{1}{\theta} + \phi_j\right) dt \right]$$

2. Steady-state reaction:

$$F_1 = \exp\left[\left(\frac{1}{\theta} + k_1\right)t\right],$$

$$c_1 = \frac{c_1^\circ}{1 + k_1\theta}.$$

The above analysis will yield solutions for any stirred-tank reactor in which the given assumptions hold. No restriction to constant input concentrations, constant volume, or isothermal conditions is required. If the reactions are coupled, however, the concentrations of all species may still be obtained as a function of time, by using matrix techniques. But the reaction must then be restricted to isothermal conditions, or cases where temperature variations are insignificant.

For a system of first-order reactions of arbitrary complexity, write the input and output concentrations in vector form as

$$\mathrm{c}^\circ(t) = \begin{bmatrix} c_1^\circ(t) \\ \vdots \\ c_S^\circ(t) \end{bmatrix}, \tag{9}$$

and

$$\mathrm{c}(t) = \begin{bmatrix} c_1(t) \\ \vdots \\ c_S(t) \end{bmatrix}. \tag{10}$$

The overall reaction rate may be expressed as Kc(t), where K is the S by S matrix of the rate constants, having elements given by

$$K_{ij} = -k_{ij} \quad (i \neq j),$$
$$K_{jj} = \sum_{\substack{i=1 \\ i \neq j}}^{S} k_{ij}, \qquad (11)$$

where k_{ij} is the rate constant for the reaction

$$A_j \xrightarrow{k_{ij}} A_i \quad (i,j = 1, 2, \ldots, S;\ i \neq j).$$

Then Eq. (5) becomes

$$\frac{\mathrm{dc}}{\mathrm{d}t} = \frac{\mathrm{c}^\circ - \mathrm{c}}{\theta} - \mathrm{Kc}, \qquad (12)$$

which may be transformed to a characteristic system by the similarity transformation(5)

$$\Lambda = \mathbf{X}^{-1}\mathbf{KX},$$

where Λ is a diagonal matrix composed of the eigenvalues of K. Wei and Prater(4) have shown that the eigenvalues of K are real and nonnegative, and that the corresponding eigenfunctions x_i form a complete set. The matrix X is formed from the eigenfunctions x_i written as column vectors.

If b(t) and b°(t) are defined by

$$\mathrm{b}(t) = \mathbf{X}^{-1}\mathrm{c}(t), \qquad (13)$$

and

$$b^\circ(t) = X^{-1}c^\circ(t), \tag{14}$$

Eq. (12) becomes

$$X^{-1}\frac{dc}{dt} = \frac{b^\circ - b}{\theta} - \Lambda b, \tag{15}$$

since

$$\Lambda b = X^{-1}KXb = X^{-1}Kc.$$

In general the rate constants, and hence K and X, are functions of temperature and therefore of time in nonisothermal systems. If temperature variations are small enough that they do not significantly affect K, X^{-1} may be considered constant. Then Eq. (15) becomes

$$\frac{db}{dt} = \frac{b^\circ - b}{\theta} - \Lambda b, \tag{16}$$

which is completely analogous to the case of a single, irreversible first-order reaction. The solution of Eq. (16), including the initial conditions, is then

$$b = F^{-1}\left[\, b(0) + \int_0^t \left(\frac{1}{\theta}Fb^\circ\right)dt \,\right],$$

where

$$F = \exp\left[\, \int_0^t \left(\frac{1}{\theta}I + \Lambda\right)dt \,\right] \tag{17}$$

is the integrating factor matrix and I is the unit matrix.

Then, from Eq. (13), the desired concentration vector $c(t)$ is

$$c(t) = \mathbf{X}b(t).$$

Matrix techniques for the analysis of complex systems of first-order chemical reactions are described by Wei.(5) His analysis of non-steady-state systems is restricted to those having fixed values of θ. The technique is extended here to arbitrary variable-volume and flow-rate policies for isothermal systems. The concentrations of all species are obtained as explicit functions of time. For non-coupled reaction systems, the restriction to isothermal conditions may be removed, and explicit solutions for the concentrations may be obtained by integrating factors. In certain simple cases, the solutions are analytic; in even the most complex cases, explicit solutions are expressed in terms of numerically evaluated integrals.

Nomenclature

A_j	Chemical symbol for species j
$b(t)$	Vector defined by Eq. (13)
$b^\circ(t)$	Vector defined by Eq. (14)
$c_j(t)$	Molar concentration of species j in reaction vessel
$c_j^\circ(t)$	Input molar concentration of species j
$c(t)$	Concentration vector, defined by Eq. (10)
$c^\circ(t)$	Input concentration vector, defined by Eq. (9)
F_j	Integrating factor, defined by Eq. (1)
F	Integrating factor matrix, defined by Eq. (17)
I	Unit matrix
K_{c_i}	Equilibrium constant based on concentrations, for reaction i
K	Rate constant matrix, defined by Eq. (11)
k_i	Rate constant for forward reaction i
k_i'	Rate constant for backward reaction i
$Q(t)$	Output volumetric flow rate
$Q^\circ(t)$	Input volumetric flow rate
R	Number of simultaneous chemical reactions taking place in reaction vessel
R_j	Overall reaction rate, defined by Eq. (2)
r_i	Forward rate of reaction i
r_i'	Backward rate of reaction i
S	Number of chemical species present
t	Time
$V(t)$	Volume of reacting mixture

$\mathbf{X}$	Modal matrix formed of x_i as column vectors
x_i	Eigenfunctions corresponding to the eigenvalues of K
α_{ij}	Order of forward reaction i with respect to species j
β_{ij}	Order of backward reaction i with respect to species j
$\theta(t)$	Pseudo residence time, defined by Eq. (4)
Λ	Diagonal matrix composed of λ_i
λ_i	Eigenvalue of K
ν_{ij}	Stoichiometric coefficient of species j in reaction i
$\phi_j(t)$	Coefficient of c_j in R_j, from Eq. (6)
$\psi_j(t)$	Term independent of c_j in R_j, from Eq. (6)

References

1. Aris, R., *Introduction to the Analysis of Chemical Reactors*, Prentice-Hall, Englewood Cliffs, N. J. (1965), p. 77.

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

3. Walas, S. M., *Reaction Kinetics for Chemical Engineers*, McGraw-Hill, New York (1959), p. 88.

4. Wei, J., and C. D. Prater, "The Structure and Analysis of Complex Reaction Systems," *Advances in Catalysis 13*, Academic Press, New York (1962), p. 364.

5. ————, "On Reactor Design for Complex Systems of First-Order Chemical Reactions," *Can. J. Chem. Eng. 44*, 31 (1966).

Proposition II

The yield obtained with arbitrary periodic alternating input and output, from a single, irreversible, first-order reaction in an isothermal stirred-tank chemical reactor, is calculated in terms of the corresponding steady-state residence time and the parameters defining the period. A special case is considered in detail.

For a given stirred-tank reactor, the equation to calculate the ratio of the yield under periodic alternating input and output conditions to that for steady-state operation is derived, and the results are shown graphically for the special case of zero initial volume, equal input and output times, and zero down time. In this special case, the ratio of the yield with periodic operation to the steady-state yield passes through a maximum with respect to the dimensionless steady-state residence time, τ_{ss}. The value of τ_{ss} at which the maximum occurs depends on the filling-time fraction of each period. In the limiting case, with negligible filling, emptying, and down time, the maximum yield is 29.8% greater than the steady-state yield, at $\tau_{ss} = 1.793$.

Consider an isothermal stirred-tank reactor of fixed maximum volume V_m, as shown in Fig. 1, in which the

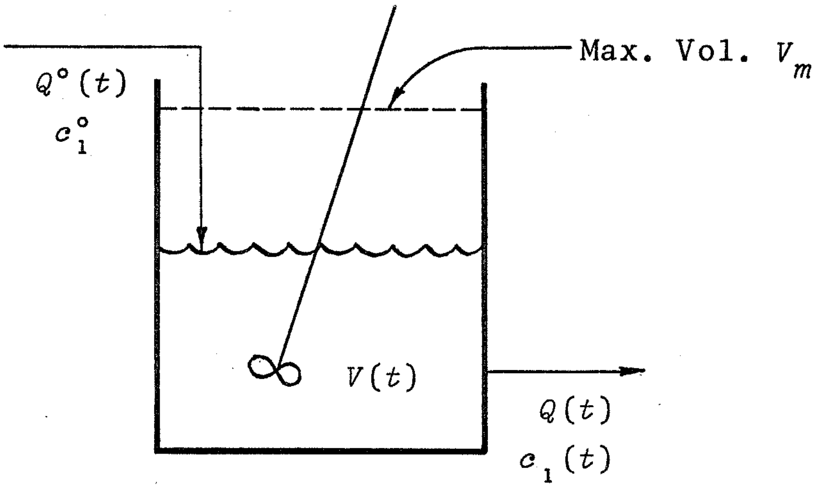


FIGURE 1.—Schematic diagram of an isothermal stirred-tank chemical reactor, with fixed maximum volume V_m and constant input concentration of species 1.

first-order reaction

$$A_1 \xrightarrow{k_1} \text{products}$$

occurs. Assume that the time-average volumetric flow rate Q_{ss} and the input molar concentration c_1° of species 1 are constant, and that the volume change due to reaction is negligible. Consider the case of periodic alternating input and output, depicted in Fig. 2. The conditions during each time interval within one period are given in Table 1. Define a pseudo residence time $\theta(t)$ by

$$\theta(t) = \frac{V(t)}{Q^\circ(t)}. \tag{1}$$

For a single irreversible first-order reaction in which volume changes due to the reaction are negligible, a material balance on the reactor gives

$$\frac{d}{dt}(c_1 V) = Q^\circ c_1^\circ - Qc_1 - Vk_1 c_1,$$

or

$$\frac{dc_1}{dt} + \left(\frac{1}{\theta} + k_1\right)c_1 = \frac{c_1^\circ}{\theta}. \tag{2}$$

Consider isothermal conditions. Then for the input time interval, from Table 1, the integrating factor technique applied to Eq. (2) yields

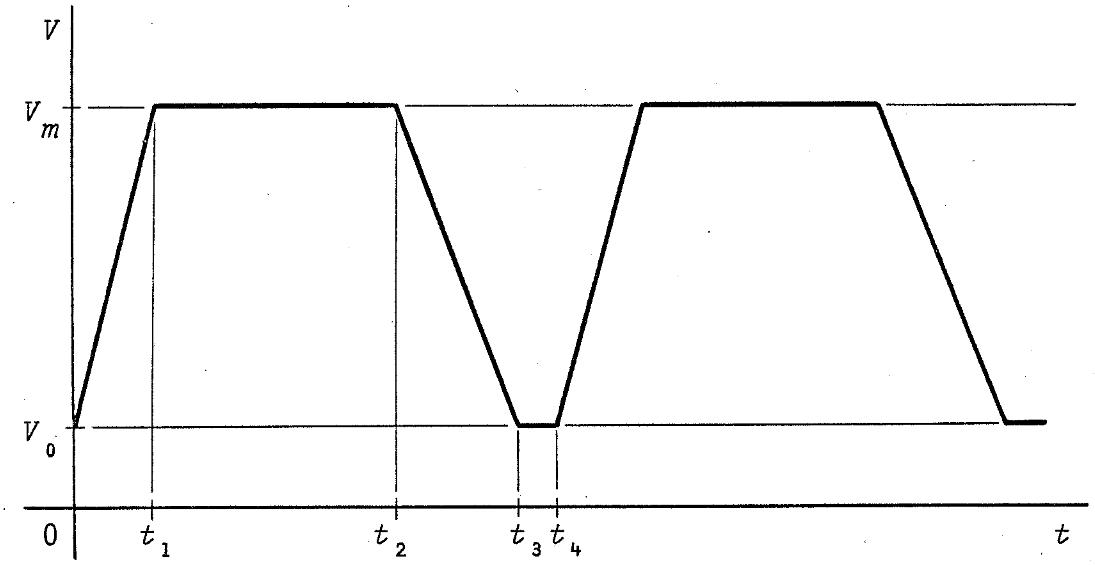


FIGURE 2.—Graph of periodic alternating input and output volume policy for a stirred-tank chemical reactor.

TABLE 1.—Conditions of alternating input and output for a stirred-tank reactor, as shown in Fig. 2.

Time Interval	$0 \leqq t \leqq t_1$	$t_1 \leqq t \leqq t_2$	$t_2 \leqq t \leqq t_3$	$t_3 \leqq t \leqq t_4$
Process	filling	reacting	emptying	down
Input flow rate	$Q°$	0	0	0
Output flow rate	0	0	Q	0
Volume	$V_0+tQ°$	V_m	$V_m-(t-t_2)Q$	V_0
Pseudo residence time	θ_0+t	∞	∞	∞

$$c_1(t) = \frac{\theta_0}{\theta_0 + t}\left[c_{10}e^{-k_1 t} + \frac{c_1}{k_1\theta_0}\left(1 - e^{-k_1 t}\right)\right] \qquad (0 \leqq t \leqq t_1),$$

where the subscripts 0 indicate initial conditions. For the remainder of each period, the reciprocal of θ is zero, so that integration of Eq. (4) results in

$$c_1(t) = c_1(t_1)e^{-k_1(t-t_1)} \qquad (t_1 \leqq t \leqq t_4). \tag{3}$$

For periodic flow, where the volume and concentration profiles repeat themselves in each cycle, it follows that

$$V(t_3) = V(t_4) = V_0,$$

and

$$c_1(t_4) = c_{10} \qquad (V_0 \neq 0). \tag{4}$$

Then the volume conditions in Table 1 and Eq. (4) give

$$\frac{Q^\circ}{Q} = \frac{t_3 - t_2}{t_1}. \tag{5}$$

Furthermore, Eqs. (6) and (8) yield

$$\frac{c_{10}}{c_1^\circ} = \frac{e^{-k_1 t_4}(e^{k_1 t_1} - 1)}{k_1 t_1 + k_1\theta_0(1 - e^{-k_1 t_4})} \qquad (V_0 \neq 0).$$

For $V_0 = 0$, the initial concentration c_{10} is taken equal to the input concentration c_1°:

$$c_{10} = \lim_{t \to 0} c_1(t) = c_1^\circ.$$

Define the overall yield as the fraction of the feed converted to the desired product.(1) For alternating input and output, the yield is given by

$$\eta_p = \frac{\int_0^{t_1} Q^\circ c_1^\circ dt - \int_{t_2}^{t_3} Q c_1 dt}{\int_0^{t_1} Q^\circ c_1^\circ dt},$$

or

$$\eta_p = 1 - \frac{Q c_1(t_1) e^{k_1 t_1} (e^{-k_1 t_2} - e^{-k_1 t_3})}{Q^\circ c_1^\circ k_1 t_1},$$

where the subscript p indicates periodic alternating input and output conditions. For steady-state operation, as in a continuous stirred-tank reactor,(2) it is well known that

$$c_1 = \frac{c_1^\circ}{1 + k_1 \theta_{ss}}.$$

Then

$$\eta_{ss} = \frac{c_1^\circ - c_1}{c_1^\circ}$$

$$= \frac{k_1 \theta_{ss}}{1 + k_1 \theta_{ss}}.$$

From the conditions on volume in Table 1, it follows that

$$V_m = V_0 + Q^\circ t_1.$$

Equating the average flow rates in the two cases, one obtains

$$Q_{ss} t_4 = Q^\circ t_1.$$

Define the dimensionless yield, η, as the ratio of the periodic variable-volume yield to that obtained under steady-state conditions:

$$\eta = \frac{\eta_p}{\eta_{ss}}. \tag{6}$$

Define also the dimensionless variables

$$\begin{aligned} \tau_i &= k_1 t_i \quad (i = 1, 2, 3, 4), \\ \tau_{ss} &= k_1 \theta_{ss}. \end{aligned} \tag{7}$$

Making the necessary substitutions, straightforward manipulations result in

$$\eta = \left(1 + \frac{1}{\tau_{ss}}\right)\left[1 - \left(\frac{e^{\tau_1}-1}{\tau_1}\right)\left(\frac{e^{\tau_2}-e^{\tau_3}}{\tau_3-\tau_2}\right)\left(\frac{\tau_4}{\tau_{ss}-(\tau_{ss}-\tau_4)e^{-\tau_4}}\right)\right], \tag{8}$$

which gives the dimensionless yield for an arbitrary periodic alternating input and output policy under the assumptions given. If the initial volume is zero, it follows that

$$\tau_4 = \tau_{ss} \quad (V_0 = 0).$$

Consider the special case where the initial volume is zero, the input and output flow rates are equal, and the down time is zero, as shown in Fig. 3. Let σ be the dimensionless filling time (the filling-time fraction of each period):

$$\sigma = \frac{\tau_1}{\tau_{ss}} = \frac{t_1}{\theta_{ss}} \quad (0 \leqq \sigma \leqq \tfrac{1}{2}). \tag{9}$$

Then in this special case, Eq. (8) becomes

$$\eta = \left(1 + \frac{1}{\tau_{ss}}\right)\left[1 - \left(\frac{e^{\sigma\tau_{ss}}-1}{\sigma\tau_{ss}}\right)^2 e^{-\tau_{ss}}\right], \tag{10}$$

with

$$\eta_p = 1 - \left(\frac{e^{\sigma\tau_{ss}}-1}{\sigma\tau_{ss}}\right)^2 e^{-\tau_{ss}}, \tag{11}$$

and

$$\eta_{ss} = \frac{\eta_{ss}}{1 + \tau_{ss}}.$$

In Fig. 4, the yield in the periodic case, η_p, as calculated from Eq. (11), is plotted with respect to τ_{ss} at values of σ in intervals of 0.1 throughout its range. For comparison, the corresponding steady-state yield, from Eq. (7), is also shown. Yields significantly greater than that obtained in the steady-state case are possible for appropriate values of σ and τ_{ss}. The dimensionless yield, η, from Eq. (10), is

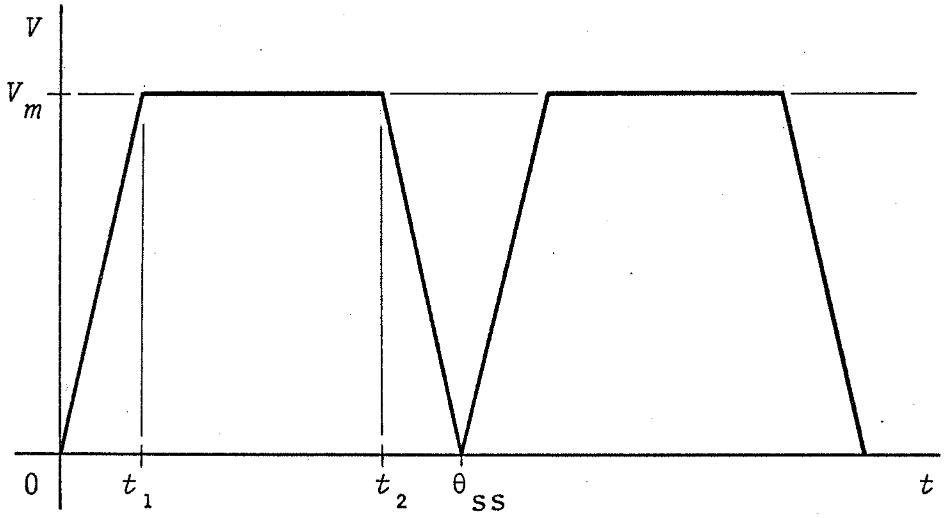


FIGURE 3.—Special case of periodic alternating input and output policy for a stirred-tank chemical reactor. Here $\theta_{ss} - t_2 = t_1$.

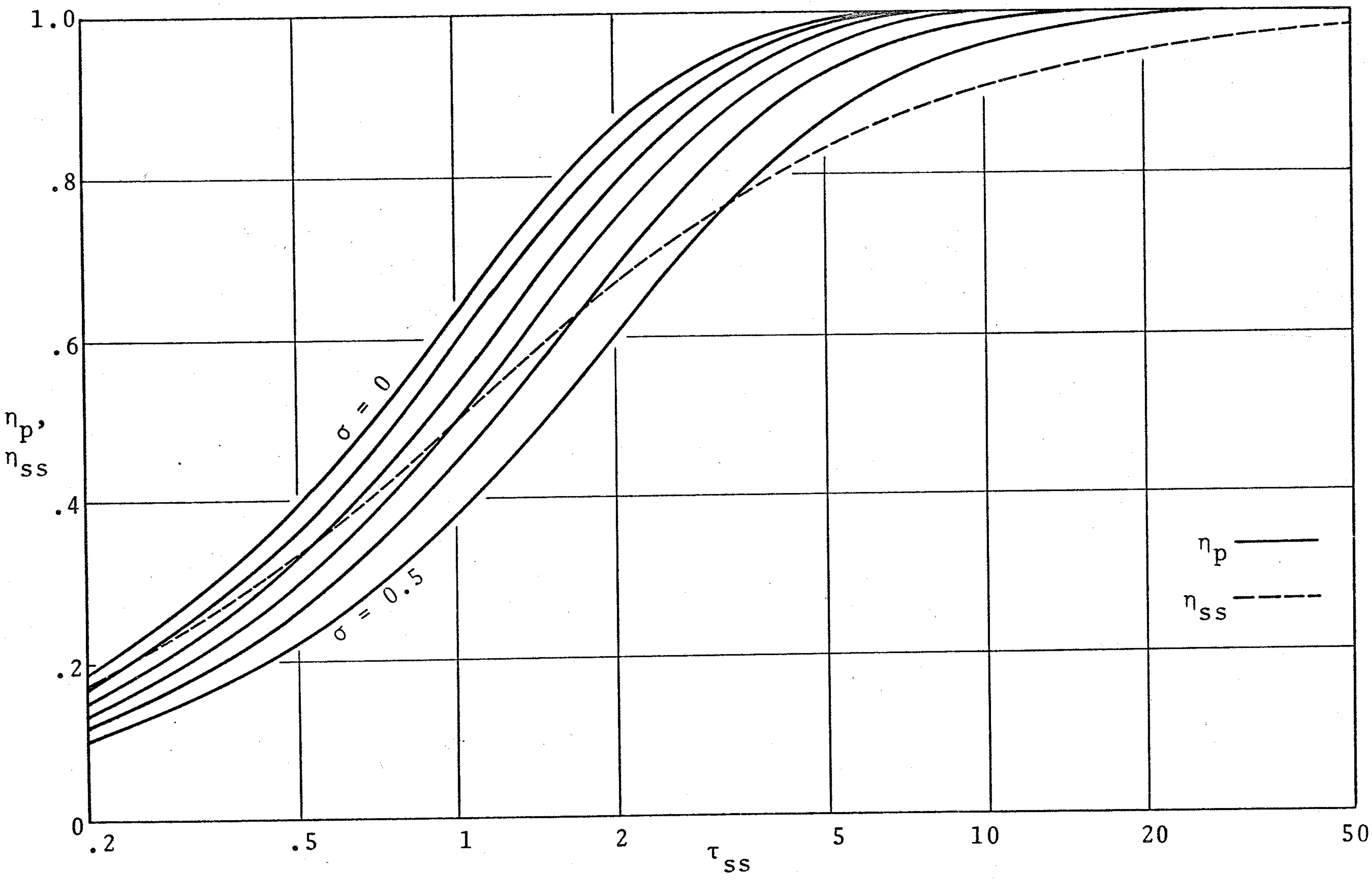


FIGURE 4.—Yield in the periodic and steady-state cases vs. the dimensionless steady-state residence time τ_{ss}. Sigma, plotted in intervals of 0.1, is the filling-time fraction of each period in the periodic cases.

plotted with respect to τ_{ss} at various values of σ in Fig. 5. In each case there is a maximum yield efficiency at some particular value of τ_{ss}. These maxima occur within the range of τ_{ss} from 1 to 10, for which the steady-state yield is from 0.5 to 0.9. For negligible filling and emptying time ($\sigma = 0$), the maximum increase in yield occurs for $\tau_{ss} = 1.793$. Then $\eta_p = 0.8336$, which is 29.8% greater than the corresponding steady-state value $\eta_{ss} = 0.6420$.

The yield from a continuous, isothermal stirred-tank reactor in which a single, irreversible, first-order reaction occurs may be substantially improved by changing the process to the periodic alternating input and output policy of this special case, provided the following conditions are met: (1) complete mixing may be assumed; (2) the volume of the reactor, the time-average flow rate, and the rate constant are such that τ_{ss} is in the desired range; (3) the filling (or emptying) time is a sufficiently small part of each cycle (say $\sigma < 0.2$); and (4) the effect of temperature changes on the reaction rate in the periodic case is negligible. Without regard to condition (2), it is known that, for $\sigma = 0$, some improvement in yield occurs in the periodic case, which is then simply an idealized batch reaction scheme. Conversely, condition (2) insures that the improvement will be substantial under the appropriate conditions, even while allowing for small filling and emptying times. Figure 5 illustrates this feature vividly.

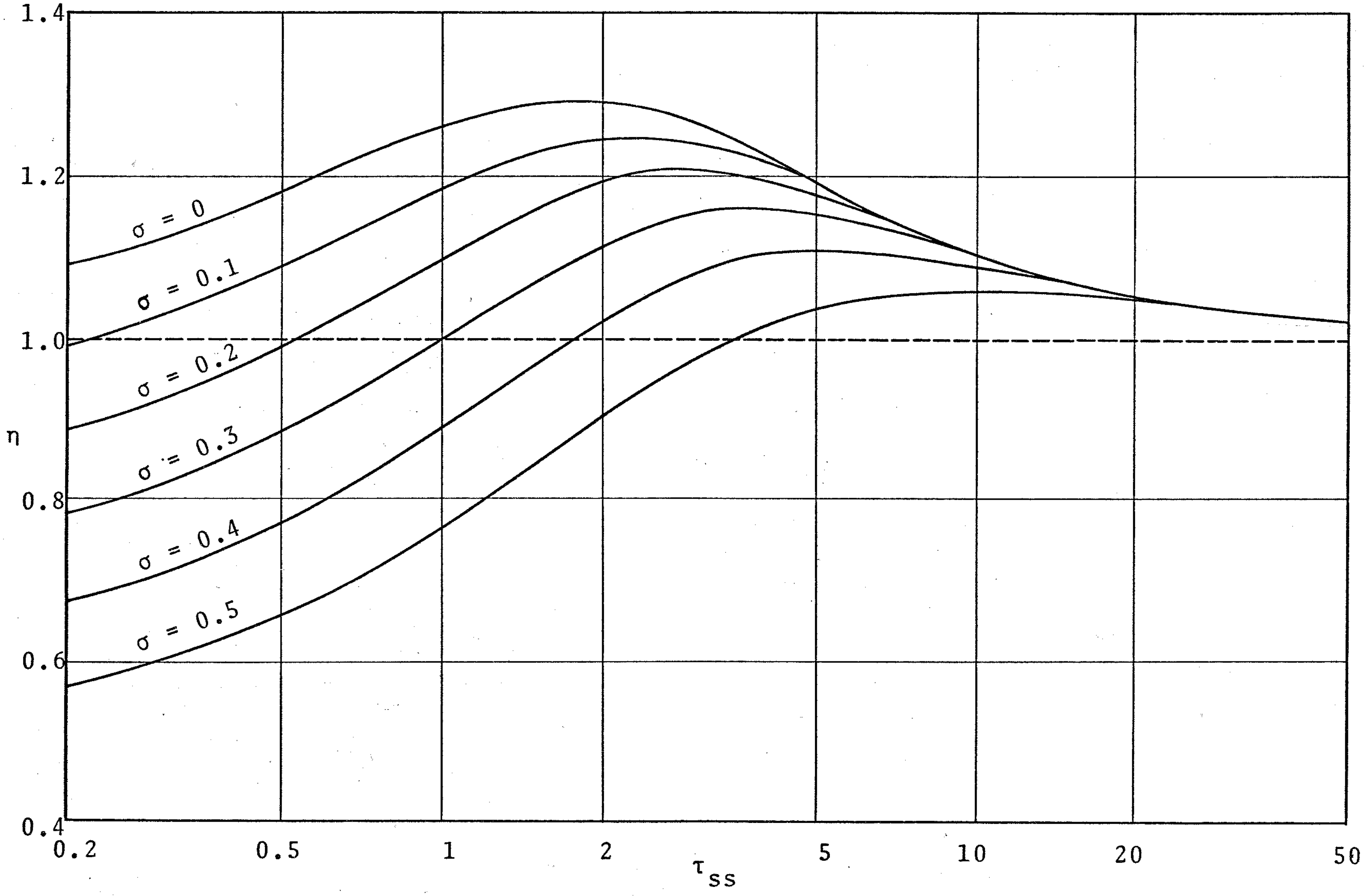


FIGURE 5.—Dimensionless yield, η, vs. the dimensionless steady-state residence time, τ_{ss}, at various values of σ, the filling-time fraction of each period. The broken line represents the steady-state yield.

Changing a steady-state process to this special case of periodic flow conditions would prove economically feasible provided the value of the increased yield would more than offset additional handling and storage costs in the periodic case.

For an arbitrary alternating input and output policy for an isothermal stirred-tank reactor, Eq. (8) gives the dimensionless yield in terms of τ_{ss} and the dimensionless variables defining the period. If input and output occur simultaneously during any part of the period, the equations become more complicated, but the yield is not further improved.

This analysis may be extended to an isothermal system of first-order reactions of arbitrary complexity by matrix diagonalization, as shown in Proposition I.

Nomenclature

$c_1(t)$	Molar concentration of reactive species in reactor
$c_1^{\circ}(t)$	Input molar concentration of reactive species
k_1	First-order reaction rate constant
$Q(t)$	Output volumetric flow rate
$Q^{\circ}(t)$	Input volumetric flow rate
t	Time
t_i	Times defining periodic conditions, given in Fig. 2 (i = 1, 2, 3, 4)
$V(t)$	Volume of reacting mixture
V_m	Maximum volume of reacting mixture
η	Dimensionless yield, defined by Eq. (5)
η_{p}	Yield obtained under periodic alternating input and output conditions, given by Eq. (4)
η_{ss}	Yield obtained under steady-state conditions, given by Eq. (5)
$\theta(t)$	Pseudo residence time, defined by Eq. (1)
σ	Dimensionless filling time, defined by Eq. (9)
$\tau_i, \tau_{\mathrm{ss}}$	Dimensionless variables defined by Eqs. (7) (i = 1, 2, 3, 4)

Subscripts

0	Initial conditions
ss	Steady-state conditions

References

1. Kramers, H., and K. R. Westerterp, *Elements of Chemical Reactor Design and Operation*, Academic Press, New York (1963), p. 22.

2. Walas, S. M., *Reaction Kinetics for Chemical Engineers*, McGraw-Hill, New York (1959), p. 88.

Proposition III

An algebra of polynomial arrays reduces the problem of algebraic operations on multivariable polynomials to one of operations on arrays of the coefficients. In particular, addition and multiplication of polynomials are thereby made easily amenable to computer calculation.

Write the polynomial $P(\mathbf{x})$ in n variables in the form

$$P(\mathbf{x}) = \sum_{\mathbf{m}} P_{\mathbf{m}} \mathbf{x}^{\mathbf{m}},$$

where

$$\mathbf{x} = x\ ,\ x\ ,\ \ldots,\ x_n,$$

$$\mathbf{m} = m\ ,\ m\ ,\ \ldots,\ m_n,$$

$$m_i = 0,\ 1,\ 2,\ \ldots,\ p_i,$$

$$\mathbf{x}^{\mathbf{m}} = \prod_{i=1}^{n} x_i^{m_i},$$

and $P_{\mathbf{m}}$ is a numerical coefficient. Any polynomial in n variables $x_1, x_2, \ldots, x_n$ can be expressed uniquely as a finite n-dimensional array of its coefficients, such that the coefficient of $\mathbf{x}^{\mathbf{m}}$ is in position $\mathbf{m}$ in the array; that is, in position m_i in dimension i ($i = 1, 2, \ldots, n$). Then $P_{\mathbf{m}}$ in the array P for $P(\mathbf{x})$ is the element in position $\mathbf{m}$, and the array extends in dimension i from position 0 to position p_i.

Addition of two polynomials $P(x)$ and $Q(x)$ may be written in the form

$$P(x) + Q(x) = \sum_{m} (P_m + Q_m)x^m,$$

from which the corresponding operation on the coefficient arrays is defined by

$$(P + Q)_m = P_m + Q_m. \tag{1}$$

The addition holds even if the arrays are of different size or dimensionality, since elements for positions corresponding to powers of the variables which do not occur in a particular polynomial may be taken equal to zero.

Similarly, multiplication of the two polynomials is given by

$$\begin{aligned} P(x)Q(x) &= \sum_{k} \sum_{\ell} P_k Q_\ell x^{k+\ell} \\ &= \sum_{k} \sum_{m-k} P_k Q_{m-k} x^m. \end{aligned}$$

Therefore multiplication of the arrays can be defined by

$$\begin{aligned} (PQ)_m &= \sum_{k+\ell=m} P_k Q_\ell \\ &= \sum_{k \leqq m} P_k Q_{m-k}, \end{aligned} \tag{2}$$

where the sum over $k \leqq m$ is taken over all *values* of the elements of k up to and including the values of the elements of m; that is,

$$k_i = 0,\ 1,\ 2,\ \ldots,\ m_i \qquad (i = 1,\ 2,\ \ldots,\ n).$$

As an example, for two variables only, one has

$$
\begin{aligned}
(PQ)_{ij} &= \sum_{k=0}^{i} \sum_{\ell=0}^{j} P_{k\ell} Q_{i-k,j-\ell} \\
&= P_{00}Q_{ij} + P_{01}Q_{i,j-1} + \ldots + P_{0j}Q_{i0} \\
&+ P_{10}Q_{i-1,j} + P_{11}Q_{i-1,j-1} + \ldots + P_{1j}Q_{i-1,0} \\
&+ \ldots \\
&+ P_{i0}Q_{0j} + P_{i1}Q_{0,j-1} + \ldots + P_{ij}Q_{00}.
\end{aligned}
$$

Since the product array PQ will in general be larger than either P or Q, some of the terms in this expansion will be identically zero. Let p_i and q_i be the highest values of m_i in P and Q, respectively. Then the limits on the summations can be changed, so that

$$(PQ)_{ij} = \sum_{k=r_1}^{s_1} \sum_{\ell=r_2}^{s_2} P_{k\ell} Q_{i-k,j-\ell},$$

where

$$r_n = \max\ (0,\ i-q_n),$$

$$s_n = \min\ (i,\ p_n).$$

Let P, Q, and R be any three members of the set of all polynomial arrays. This set contains a zero member, an array in which all elements are zero; and a unity member, an array in which the element $P_{\mathrm{m}} = 1$ for $\mathrm{m} = 0,0,\ldots,0$ and all other elements are zero. All zero members, of whatever size or dimensionality, are equivalent, and similarly for all unity members. For every member the set also contains a negative member, an array every element of which is the negative of that in the given array. But since the values of m_i $(i = 1, 2, \ldots, n)$ are taken to be nonnegative integers, no inverse members are defined.

Algebraic properties of the arrays are the following:

(1) Addition is obviously associative and commutative, since from Eq. (1), it follows that

$$\begin{aligned}[(P+Q) + R]_{\mathrm{m}} &= (P+Q)_{\mathrm{m}} + R_{\mathrm{m}} \\ &= P_{\mathrm{m}} + Q_{\mathrm{m}} + R_{\mathrm{m}} \\ &= P_{\mathrm{m}} + (Q+R)_{\mathrm{m}} \\ &= [P + (Q+R)]_{\mathrm{m}},\end{aligned}$$

and

$$\begin{aligned}(P + Q)_{\mathrm{m}} &= P_{\mathrm{m}} + Q_{\mathrm{m}} \\ &= Q_{\mathrm{m}} + P_{\mathrm{m}} \\ &= (Q + P)_{\mathrm{m}}.\end{aligned}$$

(2) Multiplication is associative and commutative:

$$\begin{aligned}
[P(QR)]_m &= \sum_{i+j=m} P_i (QR)_j \\
&= \sum_{i+j=m} P_i \sum_{k+\ell=j} Q_k R_\ell \\
&= \sum_{i+k+\ell=m} P_i Q_k R_\ell \\
&= \sum_{j+\ell=m} \sum_{i+k=j} P_i Q_k R_\ell \\
&= \sum_{j+\ell=m} (PQ)_j R_\ell \\
&= [(PQ)R]_m,
\end{aligned}$$

and

$$\begin{aligned}
(PQ)_m &= \sum_{k+\ell=m} P_k Q_\ell \\
&= \sum_{k+\ell=m} P_\ell Q_k \\
&= \sum_{k+\ell=m} Q_k P_\ell \\
&= (QP)_m.
\end{aligned}$$

(3) Multiplication is distributive with respect to addition:

$$\begin{aligned}
[P(Q+R)]_m &= \sum_{k\leq m} P_k (Q+R)_{m-k} \\
&= \sum_{k\leq m} P_k (Q_{m-k} + R_{m-k}) \\
&= \sum_{k\leq m} P_k Q_{m-k} + \sum_{k\leq m} P_k R_{m-k} \\
&= (PQ)_m + (PR)_m.
\end{aligned}$$

The set of all polynomial arrays, the given definitions of addition and multiplication, zero, unity and negative members of the set, and the above properties, together, are sufficient to define an algebra.(1,2)

In the case of polynomials of one or two variables only, small arrays may be manipulated by hand, and more easily than the polynomials themselves. Even for polynomials of several variables, the operations can be easily performed by computer, by the use of Eqs. (1) and (2). As an example, consider the following polynomials:

$$P(x,y) = (x+1)^2(x+y) = x^3 + x^2y + 2x^2 + 2xy + x + y,$$

$$Q(x,y) = (x^2+1)(y+2)^2 = x^2y^2 + 4x^2y + 4x^2 + y^2 + 4x + 4.$$

To obtain the coefficient of x^3y in $P(x,y)[P(x,y) + Q(x,y)]$, rather tedious but direct manipulation of the polynomials gives

$$\begin{aligned} &P(x,y)[P(x,y) + Q(x,y)] \\ &\quad = (x+1)^2(x+y)[(x+1)^2(x+y) + (x^2+1)(y+2)^2] \\ &\quad = (x^3+x^2y+2x^2+2xy+x+y)(x^3+5x^2y+6x^2+2xy+x+5y+x^2y^2 \\ &\qquad\qquad +y^2+4) \end{aligned}$$

$$\begin{aligned}
&= x^6 + 5x^5y + 6x^5 + 2x^4y + x^4 + 5x^3y + x^5y^2 + x^3y^2 \\
&\quad + 4x^3 + x^5y + 5x^4y^2 + 6x^4y + 2x^3y^2 + x^3y + 5x^2y^2 \\
&\quad + x^4y^3 + x^2y^3 + 4x^2y + 2x^5 + 10x^4y + 12x^4 + 4x^3y \\
&\quad + 2x^3 + 10x^2y + 2x^4y^2 + 2x^2y^2 + 8x^2 + 2x^4y + 10x^3y^2 \\
&\quad + 12x^3y + 4x^2y^2 + 2x^2y + 10xy^2 + 2x^3y^3 + 2xy^3 \\
&\quad + 8xy + x^4 + 5x^3y + 6x^3 + 2x^2y + x^2 + 5xy + x^3y^2 \\
&\quad + xy^2 + 4x + x^3y + 5x^2y^2 + 6x^2y + 2xy^2 + xy + 5y^2 \\
&\quad + x^2y^3 + y^3 + 4y
\end{aligned}$$

$$\begin{aligned}
&= x^6 + 8x^5 + 6x^5y + x^5y^2 + 14x^4 + 20x^4y + 7x^4y^2 \\
&\quad + x^4y^3 + 12x^3 + 28x^3y + 14x^3y^2 + 2x^3y^3 + 9x^2 \\
&\quad + 24x^2y + 16x^2y^2 + 2x^2y^3 + 4x + 14xy + 13xy^2 \\
&\quad + 2xy^3 + 4y + 5y^2 + y^3,
\end{aligned}$$

from which the coefficient of x^3y is 28. Array algebra alone requires only the following:

$$P(P+Q) = \begin{pmatrix} 0 & 1 & 2 & 1 \\ 1 & 2 & 1 & 0 \end{pmatrix} \left[\begin{pmatrix} 0 & 1 & 2 & 1 \\ 1 & 2 & 1 & 0 \end{pmatrix} + \begin{pmatrix} 4 & 0 & 4 \\ 4 & 0 & 4 \\ 1 & 0 & 1 \end{pmatrix} \right]$$

$$= \begin{pmatrix} 0 & 1 & 2 & 1 \\ 1 & 2 & 1 & 0 \end{pmatrix} \begin{pmatrix} 4 & 1 & 6 & 1 \\ 5 & 2 & 5 & 0 \\ 1 & 0 & 1 & 0 \end{pmatrix} \tag{3}$$

$$= \begin{pmatrix} 0 & 4 & 9 & 12 & 14 & 8 & 1 \\ 4 & 14 & 24 & 28 & 20 & 6 & 0 \\ 5 & 13 & 16 & 14 & 7 & 1 & 0 \\ 1 & 2 & 2 & 2 & 1 & 0 & 0 \end{pmatrix},$$

in which, as can be readily seen, the element in row $i+1$ and column $j+1$ is the coefficient of $x^i y^j$ in the expansion of $P(x,y)[P(x,y) + Q(x,y)]$. If only the single coefficient of x^3y is desired, it is not necessary to perform the final multiplication, since from Eq. (3), Eq. (2) gives immediately

$$\begin{aligned}[P(P+Q)]_{3,1} &= 0\cdot 0 + 1\cdot 5 + 2\cdot 2 + 1\cdot 5 + 1\cdot 1 + 2\cdot 6 \\ &\quad + 1\cdot 1 + 0\cdot 4 \\ &= 28.\end{aligned}$$

The usefulness of this polynomial algebra appears in two major ways:

(1) For manual calculations, by keeping track of the variables by their position in the arrays, instead of by a collection of products of variables raised to various powers, the mind is free to concentrate on the actual computations to be performed, rather than cluttered with a preponderance of the variables themselves. For polynomials of one variable, this feature may be compared to synthetic division in place of long division.

(2) A computer can be efficiently programmed to add or multiply polynomials by manipulating the coefficient arrays. Present methods for adding or multiplying polynomials by computer, for which algorithms exist,(3) use a searching technique with circularly linked lists to perform the computations, and can not be easily adapted to obtaining the coefficient of a single term without performing the

entire set of calculations for a given product or sum of polynomials. On the other hand, Eqs. (1) and (2) are expressed in terms of a single element of the sum or product array, which element can then be obtained individually.

References

1. Albert, A. A., *Modern Higher Algebra*, University of Chicago Press, Chicago (1937), p. 17.
2. Aleksandrov, A. D., A. N. Kolmogorov, and M. A. Lavient'ev, editors, *Mathematics, Its Content, Methods and Meaning*, American Mathematical Society, Providence, R. I. (1963), Part 6, p. 141.
3. Knuth, D. E., *Fundamental Algorithms*, Vol. 1 of *The Art of Computer Programming*, Addison-Wesley, Reading, Mass. (1968), p. 272.

Proposition IV

Explicit expressions for the effect of volume on the chemical reaction parameter, for multiple-reaction closed systems at chemical equilibrium, are obtained as solutions of a set of linear algebraic equations. Adiabatic as well as isothermal conditions are considered, and simplifications for the special case of perfect gases are indicated.

In his Ph.D. thesis, C. S. Lu(2) obtained expressions for the effect of pressure on the equilibrium displacement of multiple-reaction closed systems. Corresponding expressions for the effect of volume, including adiabatic conditions, are derived here. All these results have recently been published jointly.(1) For single-reaction systems, expressions for the effect of pressure and volume have been previously given by Pings.(3)

The notation here is simpler than that employed by Lu(3) since this notation is given entirely in molar instead of in mass form, does not require dividing the chemical potentials by temperature before performing the derivations, and does not involve differentiating with respect to an arbitrary variable.

Consider c chemical species involved in r independent reactions in a closed system, where ν_{ik} is the stoichiometric coefficient for species k in reaction i. The following basic definitions and thermodynamic equations for closed systems are employed. Constraints on equations are indicated in brackets.

Helmholtz energy: $A = E - TS$. (1)

Chemical potential: $\mu_k = \left(\frac{\partial A}{\partial n_k}\right)_{T,V,n_\ell}$. (2)

Material balance: $\mathrm{d}n_k = \sum_{i=1}^{r} \nu_{ik}\mathrm{d}\xi_i$. (3)

First law of thermodynamics: $\mathrm{d}E = \delta q - P\mathrm{d}V$ $[E_v]$. (4)

Gibbs' equation: $\mathrm{d}A = -P\mathrm{d}V - S\mathrm{d}T + \sum_{k=1}^{c} \mu_k \mathrm{d}n_k$. (5)

Chemical equilibrium: $\sum_{k=1}^{c} \nu_{ik}\mu_k = 0 \quad (i = 1, 2, \ldots, r)$ [eq]. (6)

The following quantities are introduced for convenience:(1,3)

$$\tilde{E}_k = \left(\frac{\partial E}{\partial n_k}\right)_{T,V,n_\ell} , \tag{7}$$

$$(\Delta E)_i = \sum_{k=1}^{c} \nu_{ik}E_k , \tag{8}$$

$$\tilde{P}_k = \left(\frac{\partial P}{\partial n_k}\right)_{T,V,n_\ell} , \tag{9}$$

$$(\Delta P)_i = \sum_{k=1}^{c} \nu_{ik}\tilde{P}_k, \tag{10}$$

$$\tilde{S}_k = \left(\frac{\partial S}{\partial n_k}\right)_{T,V,n_\ell}, \tag{11}$$

$$\beta_{hi} = \frac{1}{RT}\sum_{k,\ell} \nu_{hk}\nu_{i\ell}\mu'_{k\ell}, \tag{12}$$

$$\mu'_{k\ell} = \left(\frac{\partial \mu_k}{\partial n_\ell}\right)_{T,V,n_m} = \left(\frac{\partial^2 A}{\partial n_k \partial n_\ell}\right)_{T,V,n_m} = \left(\frac{\partial \ \ell n \ f_k}{\partial n_\ell}\right)_{T,V,n_m}, \tag{13}$$

$$\underset{\sim}{\nu}_i = \sum_{k=1}^{c} \nu_{ik}. \tag{14}$$

Substitution of partial derivatives of μ_k into $d\mu_k$, where $\mu_k = \mu_k(T,V,\xi_1,\xi_2,\ldots,\xi_r)$ and the derivatives are obtained by the use of Eqs. (3), (5), (9), (11), and (13), with an interchange in the order of differentiation, results in

$$d\mu_k = -\tilde{S}_k dT - \tilde{P}_k dV + \sum_{i=1}^{r}\sum_{\ell=1}^{c} \nu_{i\ell}\mu'_{k\ell} d\xi_i. \tag{15}$$

Combination of Eqs. (1), (2), (7), and (11) gives

$$\mu_k = \tilde{E}_k - T\tilde{S}_k. \tag{16}$$

Then substitution of Eqs. (15) and (16) into the differential form of Eq. (6) and simplification by the use of Eqs. (8), (10), and (12) yields

$$-\frac{(\Delta E)_h}{T}\,\mathrm{d}T - (\Delta P)_h \mathrm{d}V + RT\sum_{i=1}^{r} \beta_{hi}\mathrm{d}\xi_i = 0$$

$$(h = 1,\ 2,\ \ldots,\ r)\ [\mathrm{eq}].\quad (17)$$

Hence

$$RT\sum_{i=1}^{r}\beta_{hi}\left(\frac{\partial \xi_i}{\partial T}\right)_V = \frac{(\Delta E)_h}{T} \qquad (h = 1,\ 2,\ \ldots,\ r)$$

$$[\mathrm{eq}],\quad (18)$$

$$RT\sum_{i=1}^{r}\beta_{hi}\left(\frac{\partial \xi_i}{\partial V}\right)_T = (\Delta P)_h \qquad (h = 1,\ 2,\ \ldots,\ r)$$

$$[\mathrm{eq}].\quad (19)$$

Since Eqs. (18) and (19) are systems of linear equations in the unknown derivatives of ξ_i, they can each be solved simultaneously by Cramer's rule or by matrix inversion, giving the derivatives of ξ_i with respect to temperature at constant volume, and with respect to volume at constant temperature, explicitly for $i = 1, 2, \ldots, r$.

Considering $E = E(T, V, n_1, n_2, \ldots, n_c)$, making use of Eqs. (1), (3), (4), (5), and (8), dividing by dV, and constraining the derivatives to adiabatic conditions, one obtains

$$nC_{V,\xi}\left(\frac{\partial T}{\partial V}\right)_q + \left(\frac{\partial P}{\partial V}\right)_{T,\xi} + \sum_{i=1}^{r}(\Delta E)_i\left(\frac{\partial \xi_i}{\partial V}\right)_q = 0$$

$$[E_v].\quad (20)$$

Now Eq. (17) yields

$$\left(\frac{\partial T}{\partial V}\right)_q = \frac{T}{(\Delta E)_h}\left[RT\sum_{i=1}^{r}\beta_{hi}\left(\frac{\partial \xi_i}{\partial V}\right)_q - (\Delta P)_h\right]$$

$$(h = 1,\ 2,\ \ldots,\ r)\ [\text{eq}].\quad (21)$$

Substitution of Eq. (21) into Eq. (20) therefore gives

$$\sum_{i=1}^{r}\left[\beta_{hi}RT + \frac{(\Delta E)_h(\Delta E)_i}{nC_{V,\xi}T}\right]\left(\frac{\partial \xi_i}{\partial V}\right)_q = (\Delta P)_h - \frac{(\Delta E)_h}{nC_{V,\xi}}\left(\frac{\partial P}{\partial T}\right)_{V,\xi}$$

$$(h = 1,\ 2,\ \ldots,\ r)\qquad [E_v,\text{eq}],$$

from which the derivatives of ξ_i with respect to volume under adiabatic conditions can be obtained explicitly. The corresponding derivatives with respect to pressure can be obtained in an analogous manner, and are presented in the published paper$^{(1)}$ which is based in part on these results.

For a perfect gas, one has

$$f_k = \frac{n_k RT}{V}\qquad [\text{pg}],$$

which, when substituted into Eq. (13), reduces Eq. (12) to

$$\beta_{hi} = \sum_{k=1}^{c}\frac{\nu_{hk}\nu_{ik}}{n_k}\qquad [\text{pg}].\quad (22)$$

Also for a perfect gas, from Eq. (10), it follows that

$$(\Delta P)_h - \frac{(\Delta E)_h}{nC_{V,\xi}}\left(\frac{\partial P}{\partial T}\right)_{V,\xi} = \frac{P}{n}\left[\underset{\sim}{\nu}_h - \frac{(\Delta E)_h}{C_{V,\xi}T}\right]\qquad [\text{pg}].\quad (23)$$

With the use of Eqs. (22) and (23), some simplifications can be made in the terms of Eqs. (18), (19), and (21) for systems which can be considered as perfect gases.

Nomenclature

A	Total Helmholtz energy, defined by Eq. (1)
$C_{V,\xi}$	Molar heat capacity at constant volume and composition
c	Number of species in the system
d	Differential operator
E	Total internal energy
$\tilde{E}_k$	Variable defined by Eq. (7)
$(\Delta E)_i$	Variable defined by Eq. (8)
f_k	Fugacity of species k
n	Total number of moles in the system
n_k	Number of moles of species k in the system
P	Pressure
$\tilde{P}_k$	Variable defined by Eq. (9)
$(\Delta P)_i$	Variable defined by Eq. (10)
q	Heat transferred to the system
R	Molar gas constant
r	Number of independent chemical reactions occurring in the system
S	Total entropy
$\tilde{S}_k$	Variable defined by Eq. (11)
T	Absolute temperature
V	Total volume
β_{hi}	Variable defined by Eq. (12)
δ	Variation operator
μ_k	Chemical potential of species k, defined by Eq. (2)

$\mu'_{k\ell}$	Variable defined by Eq. (13)
ν_{ik}	Stoichiometric coefficient of species k in reaction i
$\underset{\sim}{\nu}_i$	Variable defined by Eq. (14)
$\sum$	Summation operator
ξ_i	Extent of reaction i, defined by Eq. (3)

Subscripts

h,i	Denote chemical reactions
k,ℓ,m	Denote chemical species in the system
n_ℓ	Denotes restriction to constant numbers of moles of all species except species k
q	Denotes adiabatic conditions
T	Denotes constant temperature
V	Denotes constant volume
ξ	Denotes constant composition

Brackets

[eq]	Chemical equilibrium
$[E_v]$	No viscous dissipation
[pg]	Perfect gas behavior

References

1. Larsen, A. H., C. S. Lu, and C. J. Pings, "Thermodynamics of Chemical Equilibrium—V. Multiple-Reaction and Open Systems," *Chem. Eng. Sci. 23*, 289 (1968).

2. Lu, C. S., "A Thermodynamic Study of Multiple-Reaction Systems at and near Equilibrium," Ph.D. thesis submitted to California Institute of Technology, Pasadena, Calif. (1967), Part II.

3. Pings, C. J., "Thermodynamics of Chemical Equilibrium—I. Effect of Temperature and Pressure," *Chem. Eng. Sci. 16*, 181 (1961); "II. Effect of Volume, Entropy, and Enthalpy," *ibid. 18*, 671 (1963).