Generalized Cumulant Expansion Method*

Ryogo KUBO

Department of Physics, University of Tokyo

(Received April 11, 1962)

The moment generating function of a set of stochastic variables defines the cumulants or the semi-invariants and the cumulant function. It is possible, simply by formal properties of exponential functions, to generalize to a great extent the concepts of cumulants and cumulant function. The stochastic variables to be considered need not be ordinary c-numbers but they may be q-numbers such as used in quantum mechanics. The exponential function which defines a moment generating function may be any kind of generalized exponential, for example an ordered exponential with a certain prescription for ordering q-number variables. The definition of average may be greatly generalized as far as the condition is fulfilled that the average of unity is unity. After statements of a few basic theorems these generalizations are discussed here with certain examples of application. This generalized cumulant expansion provides us with a point of view from which many existent methods in quantum mechanics and statistical mechanics can be unified.

§ 1. Introduction

It is well known that the cumulants or the semi-invariants of a distribution play very important roles in probability theory.[1] If they exist, they determine the nature of the random variable which we are concerned with. They are particularly useful for discussion of such asymptotic properties of distribution functions as the well-known example of central limit theorem. In quantum-mechanics and in statistical mechanics, similar concepts seem to be very essential. This fact has been noticed at least partially in many of the existent work in the above mentioned fields,

but it seems that its significance has not been fully recognized. The purpose of the present paper is to show how the concept of cumulants can be widely applied to various problem of quantum-mechanics and statistical physics, and that many known formal properties of expansion series can be most clearly understood from such a general point of view which is developed in the following. The author hopes in particular that this point of view will be of some advantage in order to clarify the stochastic nature of many-body systems. He believes that any useful theory of many-body systems has necessarily to involve certain stochastic assumptions, which are not always made clear. This will not be discussed, however, in any detail in the present paper, but a lecture of the author[2] is quoted with this respect. More detailed treatment will be given elsewhere.

§ 2. Moments and Cumulants

The moments and cumulants (semi-invariants) for a random variable X are defined by the moment generating function* $M(\xi)$ and the cumulant function $K(\xi)$:

$$\langle e^{\xi X}\rangle = \sum_{n=0}^{\infty} \frac{\xi^n}{n!}\mu_n$$

$$= \exp\left\{\sum_{n=1}^{\infty}\frac{\xi^n}{n!}\kappa_n\right\} \equiv \exp K(\xi) \tag{2.1}$$

where the bracket $\langle A\rangle$ means the expectation of a random variable A. μ_n is the n-th moment and κ_n is the n-th cumulant. These definitions may easily be generalized to multi-variant distributions. Let us suppose that we have N random variables $X_1, \cdots X_N$. For these, the moments and cumulants are generally defined by the moment generating function

$$M(\xi_1, \cdots, \xi_N) \equiv \left\langle \exp \sum_{j=1}^{N}\xi_j X_j\right\rangle = \sum_{\nu_1,\cdots,\nu_N=0}\cdots\sum\left(\prod_j \frac{\xi_j{}^{\nu_j}}{\nu_j!}\right)\mu(\nu_1,\cdots\nu_N)$$

$$= \exp\left[\sum_{\nu_1,\cdots\nu_N}'\left(\prod_j \frac{\xi_j{}^{\nu_j}}{\nu_j!}\right)\kappa(\nu_1,\cdots\nu_N)\right]$$

$$= \exp\left[K(\xi_1, \cdots, \xi_N)\right] \tag{2.2}$$

where $\displaystyle\sum_{\nu_1,\cdots\nu_N}'$ means summation over $\nu_1,\cdots$ and ν_N excluding $\nu_1=\cdots=\nu_N=0$.

The following notations are found to be useful:

$$\mu(\boldsymbol{\nu}) \equiv \mu(\nu_1, \cdots, \nu_N)$$
$$\equiv \mu(X_1{}^{\nu_1}\cdots X_N{}^{\nu_N}) \equiv \langle X_1{}^{\nu_1}\cdots X_N{}^{\nu_N}\rangle \tag{2.3}$$

$$\kappa(\boldsymbol{\nu}) \equiv \kappa(\nu_1, \cdots, \nu_N)$$
$$\equiv \kappa(X_1{}^{\nu_1}\cdots X_N{}^{\nu_N}) \equiv \langle X_1{}^{\nu_1}\cdots X_N{}^{\nu_N}\rangle_c \tag{2.4}$$

The suffix c in (2.4) indicates that the cumulant $\kappa(\boldsymbol{\nu})$ is a certain average of $X_1,\cdots X_N$, which we call the *cumulant average*. (c might as well be interpreted as *connected* in the sense of the word used in well-known techniques of graphical representation. This will become apparent later.) A cumulant average is not a simple average, but is defined by (2.2). For instance, we have

$$\langle X_1 X_2\rangle_c \equiv \langle X_1 X_2\rangle - \langle X_1\rangle\langle X_2\rangle .$$

Therefore (2.2) may be written as

$$\left\langle\exp\sum_{j=1}^{N}\xi_j X_j\right\rangle = \sum_{\nu_1,}\cdots\sum_{\nu_N=0}^{\infty}\prod_j\frac{\xi_j{}^{\nu_j}}{\nu_j!}\langle X_1{}^{\nu_1}\cdots X_N{}^{\nu_N}\rangle$$

$$= \exp\left\{\sum_{\nu_1}\cdots\sum_{\nu_N}'\prod_j\frac{\xi_j{}^{\nu_j}}{\nu_j!}\langle X_1{}^{\nu_1}\cdots X_N{}^{\nu_N}\rangle_c\right\} . \tag{2.5}$$

In the last expression the term with $\nu_1=\cdots\nu_j=\cdots=\nu_N=0$ is omitted from the sum. It is sometimes very convenient to write

$$\left\langle\exp\sum_{j=1}^{N}\xi_j X_j\right\rangle = \exp\left\langle\exp\left(\sum\xi_j X_j\right)-1\right\rangle_c \tag{2.6}$$

where $\langle\quad\rangle_c$ means that the exponential function in it is expanded and each product is averaged by the *cumulant average*.

Now an important property of cumulants is that a cumulant can be explicitly represented only by the lower (not higher) moments and vice versa. General formulae for this relation have been given in literatures. In particular, Meeron[3] gave the relationship,

$$\kappa(\boldsymbol{\nu})=-\prod_j \nu_j! \sum_{l=1}^n \sum_{\substack{l \\ \sum_{i=1} k_i m_{ij}=\nu_j}} (\sum k_i-1)!(-)^{\Sigma k_i} \prod_{i=1}^l \frac{1}{k_i!} \left\{ \frac{\mu(\boldsymbol{m}_i)}{\prod_j m_{ij}!} \right\}^{k_i}. \tag{2.7}$$

In Eq. (2.7), each term corresponds to a decomposition of $n=\sum_{j=1}^N \nu_j$ objects into $\sum_{i=1}^l k_i$ subsets which may be expressed schematically by

$$(X_1{}^{\nu_1}\cdots X_j{}^{\nu_j}\cdots X_N{}^{\nu_N}) \rightarrow \prod_{i=1}^l (X_1{}^{m_{i1}}\cdots X_j{}^{m_{ij}}\cdots X_N{}^{m_{iN}})^{k_i}$$

$$(\sum_{i=1}^l k_i m_{ij}=\nu_j)$$

and

$$\mu(\boldsymbol{m}_i)=\langle \prod_{j=1}^N X_j{}^{m_{ij}} \rangle, \qquad (\boldsymbol{m}_i=(m_{i1}, m_{i2}\cdots m_{iN}))$$

is the moment for a subset of the i-th kind. The cumulant $\kappa(\boldsymbol{\nu})$ is a linear combination of the products of these subset moments, so that it is represented by lower order moments, $\mu(\boldsymbol{m}_i)$, with $m_{ij}\leq \nu_j$ $(j=1,\cdots N)$.

We shall not make much use of this explicit form of cumulants, but for convenience of later reference let us note a few examples:

$$\langle X_j \rangle_c = \langle X_j \rangle$$

$$\langle X_j{}^2 \rangle_c = \langle X_j{}^2 \rangle - \langle X_j \rangle^2$$

$$\langle X_j X_l \rangle_c = \langle X_j X_l \rangle - \langle X_j \rangle\langle X_l \rangle$$

$$\langle X_j X_k X_l \rangle_c = \langle X_j X_k X_l \rangle$$
$$-\{\langle X_j \rangle\langle X_k X_l \rangle+\langle X_k \rangle\langle X_l X_j \rangle+\langle X_l \rangle\langle X_j X_k \rangle\}$$
$$+2\langle X_j \rangle\langle X_k \rangle\langle X_l \rangle$$

$$\langle X_j X_k X_l X_m \rangle_c = \langle X_j X_k X_l X_m \rangle$$
$$-\{\langle X_j \rangle\langle X_k X_l X_m \rangle+\langle X_k \rangle\langle X_j X_l X_m \rangle+\langle X_l \rangle\langle X_j X_k X_m \rangle+\langle X_m \rangle\langle X_j X_k X_l \rangle\}$$
$$-\{\langle X_j X_k \rangle\langle X_l X_m \rangle+\langle X_j X_l \rangle\langle X_k X_m \rangle+\langle X_j X_m \rangle\langle X_k X_l \rangle\}$$
$$+2\{\langle X_j \rangle\langle X_k \rangle\langle X_l X_m \rangle+\langle X_j \rangle\langle X_l \rangle\langle X_k X_m \rangle+\langle X_j \rangle\langle X_m \rangle\langle X_k X_l \rangle$$
$$+\langle X_j X_k \rangle\langle X_l \rangle\langle X_m \rangle+\langle X_j X_l \rangle\langle X_k \rangle\langle X_m \rangle+\langle X_j X_m \rangle\langle X_k \rangle\langle X_l \rangle\}$$
$$-6\langle X_j \rangle\langle X_k \rangle\langle X_l \rangle\langle X_m \rangle \tag{2.8}$$

The expression (2.7) becomes simpler if none of ν_j's is greater than one. For example, for

$$\nu_1=\cdots=\nu_n=1$$

the cumulant $\kappa\{n\}$ is defined for the set of the variables $(X_1, X_2, \cdots X_n)\equiv\{n\}$, which is represented in terms of lower order moments as

$$\kappa\{n\}=\sum (l-1)! (-)^{l-1} \sum_{\substack{l \\ \sum_{i=1}\{m_i\}=\{n\}}} \prod_{i=1}^l \mu\{m_i\} \tag{2.9}$$

where $\mu\{m_i\}$ is the moment for a subset $\{m_i\}$ of the set $\{n\}$ which consists of m_i variables and contains any X_j at most only once.

Generalizations of the concepts of moments and cumulants are discussed in the following, but before such generalizations a few basic theorems will be remarked.

§3. Some Basic Theorems

Theorem I: A cumulant,

$$\kappa(X_i X_j \cdots) \equiv \langle X_i X_j \cdots \rangle_c$$

is zero if the elements X_i, X_j, $\cdots$ are divided into two or more groups which are statistically independent.

Collorary: A cumulant is zero if one of the variables in it is independent of the others. Conversely, a cumulant is not zero if and only if the variables in it are statistically *connected*.

The *proof* is very simple. If the variables $(X_i, X_j \cdots)$ are divided into two groups, i.e.

$$\{X\} = \{X'\} + \{X''\}$$

which are statistically independent, the moment generating function is factorized as

$$\langle \exp \sum \xi X \rangle = \langle \exp \sum \xi' X' \rangle \langle \exp \sum \xi'' X'' \rangle \tag{3.1a}$$

because of the assumed independence. Thus the *cumulant function* takes the form

$$K\{\xi\} = K_1\{\xi'\} + K_2\{\xi''\} \tag{3.1b}$$

so that the powers of ξ' and those of ξ'' will never mix. This implies that any cumulant, in which the variables from the two groups appear, does vanish identically. The collorary immediately follows from the theorem.

Theorem II: For a stochastic variable $X(t)$, which is a function of a continuous parameter t, we have

$$\left\langle \exp \int_a^b X(t)\xi(t)dt \right\rangle$$
$$= \exp \left\{ \sum_{n=1}^{\infty} \frac{1}{n!} \int_a^b dt_1 \cdots \int_a^b dt_n \langle X(t_1) \cdots X(t_n) \rangle_c \xi(t_1) \cdots \xi(t_n) \right\} \tag{3.2a}$$
$$= \exp \left\{ \sum_{n=1}^{\infty} \int_a^b dt_1 \int_a^{t_1} dt_2 \cdots \int_a^{t_{n-1}} dt_n \langle X(t_1) \cdots X(t_n) \rangle_c \xi(t_1) \cdots \xi(t_n) \right\} \tag{3.2b}$$

where the cumulant averages are those defined by (2.4).

Collorary: Without losing generality, Eqs. (3.2a,b) may be written as

$$\left\langle \exp \int_a^b X(t)dt \right\rangle$$
$$= \exp \left\{ \sum_{n=1}^{\infty} \frac{1}{n!} \int_a^b dt_1 \cdots \int_a^b dt_n \langle X(t_1) \cdots X(t_n) \rangle_c \right\} \tag{3.3a}$$
$$= \exp \left\{ \sum_{n=1}^{\infty} \int_a^b dt_1 \int_a^{t_1} dt_2 \cdots \int_a^{t_{n-1}} dt_n \langle X(t_1) \cdots X(t_n) \rangle_c \right\} \tag{3.3b}$$

by putting $\xi(t) \equiv 1$. Eq. (3.3a) may also be written as

$$\left\langle \exp \int_a^b X(t)dt \right\rangle = \exp \left\{ \left\langle \exp \int_a^b X(t)dt - 1 \right\rangle_c \right\} \tag{3.4}$$

symbolically with the convention introduced by (2.6).

For the *proof* of (3.3a), let us write

$$\left\langle \exp \int_a^b X(t)dt \right\rangle = \lim_{\text{Max } \delta t \to 0} \left\langle \exp \sum_{j=1}^{N} X(t_j)\delta t_j \right\rangle$$

where δt_j's are small intervals which cover the whole interval of the integration, and t_j are chosen in each interval δt_j. Now the set of variables $X(t_j)\delta t_j (j=1, \cdots N)$ replaces the set of variables $X_j \xi_j$'s in Eq. (2.5). But the cumulants with any of ν_j larger than 1 vanish in the limit of Max $\delta t \to 0$. For instance

$$\lim \sum_j \langle X(t_j)^2 \rangle_c \delta t_j^2 = \lim 0(\delta t) \times \int_a^b \langle X(t)^2 \rangle_c dt = 0 \ .$$

Therefore Eq. (2.5) reduces to (3.3a) in this particular case. Eq. (3.2a) is equivalent to (3.3a) since $X(t)$ in the latter may be replaced by $X(t)\xi(t)$. Eqs. (3.2b) and (3.3b) follow from Eqs. (3.2a) and (3.3a) by restricting the integration domain to $1/n!$ of the domain $(a, b)^n$.

Eq. (3.3) may be generalized to

$$\left\langle \exp \int_a^b \sum_j X_j(t) \, dt \right\rangle$$

$$= \exp \left\{ \sum_{n=1}^\infty \frac{1}{n!} \int_a^b dt_1 \cdots \int_a^b dt_n \sum_{j_1} \cdots \sum_{j_n} \langle X_{j_1}(t_1) \cdots X_{j_n}(t_n) \rangle_c \right\} \tag{3.5}$$

and

$$\left\langle \exp \iint X(s, t) \, ds \, dt \right\rangle$$

$$= \exp \left\{ \sum_{n=1}^\infty \frac{1}{n!} \int ds_1 \cdots \int ds_n \int dt_1 \cdots \int dt_n \langle X(s_1 t_1) \cdots X(s_n t_n) \rangle_c \right\} \tag{3.6}$$

and to variables depending on more parameters.

In the above expressions the averages denoted by $\langle \quad \rangle_c$ are the cumulant averages and so may be expressed in terms of lower order moments as given by Eq. (2.9). More important is to notice that the variables in each cumulant average must be statistically *connected* or *linked*. Otherwise it vanishes by THEOREM I. This is the most general logical ground to introduce the concept of *linked* or *connected* clusters or diagrams as used in various methods in quantum-mechanical or statistical-mechanical theories.

§4. Rearrangement of Cumulant Expansion. Cluster Expansion

It is sometimes necessary to rearrange the cumulant series $K(\xi_1, \xi_2, \cdots)$ in another form,

$$K(\xi_1, \xi_2, \cdots) = \sum{}' \prod_i \frac{\xi_i^{\nu_i}}{\nu_i!} \kappa(\nu) = \sum_l K_l \ . \tag{4.1}$$

There are a great variety of the principles for such rearrangement. The general purpose is to improve the convergence of the series in order to secure a better approximation by retaining only the lower order terms in the new series, $\sum_l K_l$. The choice of the rearrangement principle depends on the nature of the particular problem. The improvement of convergence must be, in principle, proved mathematically, but such a mathematical proof is rather seldom to be given. Mostly it is just anticipated by a physical intuition.

Here we discuss only an example of rearrangement procedure of the cumulant series in a formal way without examining convergence. This is a kind of " cluster " expansion as we shall see immediately. Namely, we first collect all the terms in the cumulant series which contain a particular variable X_i. We call this $K_1(X_i)$. This is given by

$$M_1(X_i) \equiv \langle \exp \xi_i X_i \rangle = \exp K_1(X_i) \ . \tag{4.2}$$

Therefore the first term in $K = \sum_l K_l$ may be chosen as

$$K_1 = \sum_i K_1(X_i) \ . \tag{4.3}$$

In Eq. (4.2) the notation $M_1(X_i)$ is introduced for the moment generating function of the single variable X_i.

The next term K_2 will be of the form

$$K_2 = \sum_{(i,j)} K_2(X_i, X_j) \tag{4.4}$$

where $K_2(X_i, X_j)$ is the collection of all the terms in the cumulant series that contain two variables X_i and X_j. It is easy to see that

$$\exp K_2(X_i, X_j) = \frac{M_2(X_i, X_j)}{M_1(X_i) M_1(X_j)} \tag{4.5}$$

because

$$M_2(X_i, X_j) = \langle \exp(\xi_i X_i + \xi_j X_j) \rangle$$
$$= \exp\{K_1(X_i) + K_1(X_j) + K_2(X_i, X_j)\} .$$

We may proceed in this way to obtain an expansion of the form

$$K = \sum_{n=1}^{N} \sum_{\{n\}_N} K_n(\{n\}_N) \tag{4.6}$$

by introducing the functions $K_n(\{n\}_N)$ for the collection of all the terms involving the cumulants which contain any of the variables

$$\{n\}_N \equiv (X_{i_1}, X_{i_2}, \cdots, X_{i_n}) \tag{4.7}$$

at least once. $\{n\}_N$ denotes a set of n variables selected from N variables $X_1, \cdots, X_N$. The expression (4.6) is simply a rearrangement, so that there is no danger in the convergence problem as far as N is finite and the convergence of the moments is assured. There occurs the question of convergence when we go to the limit of $N \to \infty$.

The same principle of rearrangement applies to more general cases where we have a hierarchy of functions $U_n(\{n\}_N)$ which are functions of n variables,

$$\{n\}_N = (X_{i_1}, X_{i_2}, \cdots, X_{i_n}) \tag{4.8}$$

selected from a given set of N variables

$$\{N\} = (X_1, X_2, \cdots X_N) . \tag{4.9}$$

Since $U_n(\{n\}_N)$ is defined for any set of variables $\{n\}$ in $\{N\}$, we may define its moment generating function $M_n(\{n\}_N)$ by

$$M_n(\{n\}_N) = \langle \exp U_n(\{n\}_N) \rangle . \tag{4.10}$$

Then the functions $K_n(\{n\}_N)$ may be introduced exactly in the same way as before and the cumulant function $K(\{N\})$ defined by

$$M(\{N\}) \equiv \langle \exp U(N) \rangle = \exp K(\{N\}) \tag{4.11}$$

is expanded in the form

$$K(\{N\}) = \sum_{n=1}^{N} \sum_{\{n\}_N} K_n(\{n\}_N) . \tag{4.12}$$

For the explicit from of $K_n(\{n\}_N)$ we have the following theorem:

Theorem III: In the cluster expansion (4.12) of a cumulant function $K(\{n\})$, the cluster cumulant function $K_n(\{n\}_N)$ for a set of n variables, (4.8), is given explicitly by

$$K_n(\{n\}_N) = \sum_{l=1}^{n} (-)^{n-l} \sum_{\{l\}_n} \log M_l(\{l\}_N) \tag{4.13}$$

or

$$\exp K_n(\{n\}_N) = \frac{M_n(\{n\}_N) \, \Pi \, M_{n-2}(\{n-2\}_N) \cdots \cdots \Pi \, M_2(i, j)}{\Pi \, M_{n-1}(\{n-1\}_N) \, \Pi \, M_{n-3}(\{n-3\}_N) \cdots \Pi \, M_1(i)} \tag{4.14a}$$

if n is even and

$$\exp K_n(\{n\}_N) = M_n(\{n\}_N) \frac{\Pi M_{n-2}(\{n-2\}_N) \cdots \Pi M_1(i)}{\Pi M_{n-1}(\{n-1\}_N) \cdots \Pi M_2(i, j)} \tag{4.14b}$$

if n is odd. In the above expression, (4.13), $\{l\}_n$ is a set of l variables selected from the set $\{n\}_N$.

Eq. (4.13) is proved by mathematical induction. Let us consider a set of variables, $\{n+1\}_N$. Our definition of K_n's is such that

$$\log M_{n+1}(\{n+1\}_N) = \sum_{m=1}^{n+1} \sum_{\{m\}_{n+1}} K_m(\{m\}_N)$$

$$= \sum_{m=1}^{n} \sum_{\{m\}_{n+1}} K_m(\{m\}_N) + K_{n+1}(\{n+1\}_N). \tag{4.15}$$

Eq. (4.13) is now inserted into (4.15) to give

$$K_{n+1}(\{n+1\}_N) = - \sum_{m=1}^{n} \sum_{\{m\}_{n+1}} \sum_{l=1}^{m} (-)^{m-l} \sum_{\{l\}_m} \log M_l(\{l\}_N)$$

$$+ \log M_{n+1}(\{n+1\}_N) . \tag{4.16}$$

On the right hand side, each term, $\log M_l(\{l\}_N)$, appears repeatedly in the sum. Its coefficient is found to be

$$1 - \binom{n-l+1}{1} + \binom{n-l+1}{2} + \cdots \quad + (-)^{n-l}\binom{n-l+1}{n-l}$$

$$= (1-1)^{n-l+1} - (-)^{n-l+1} = (-)^{n-l}$$

which verifies Eq. (4.13).

Eqs. (4.3), (4.5) and (4.6) correspond to the simplest example for which the functions $U(\{n\}_N)$ are

$$U(\{n\}_N) = \sum_{i \in \{n\}} \xi_i X_i .$$

Corresponding to Theorem I, we have the following theorem:

Theorem IV: If the set $\{n\}_N$ is divided into independent sets $\{n'\}_N$ and $\{n''\}_N$, namely if

$$\{n\}_N = \{n'\}_N + \{n''\}_N$$

and

$$M_n(\{n\}_N) = M_{n'}(\{n'\}_N) \times M_{n''}(\{n''\}_N) \tag{4.17}$$

then

$$K_n(\{n\}_N) \equiv K_{n'+n''}(\{n'+n''\}) = 0 , \tag{4.18}$$

and more generally

$$K_{m'+m''}(\{m'\}_{n'} + \{m''\}_{n''}) = 0 \tag{4.19}$$

if neither of $\{m'\}$ and $\{m''\}$ is empty.

The proof of this theorem is almost evident, because we have

$$M_n(\{n\}_N) = M_{n'}(\{n'\}_N) \times M_{n''}(\{n''\}_N)$$

$$= \exp\left[\sum_{m=1}^{n} \sum_{\{m'\}_{n'}} \sum_{\{m''\}_{n''}} K_m(\{m'\}_{n'} + \{m''\}_{n''})\right]$$

$$= \exp\left[\sum_{m'=1}^{n'} \sum_{\{m'\}_{n'}} K_{m'}(\{m'\}_{n'})\right] \times \exp\left[\sum_{m''=1}^{n''} \sum_{\{m''\}_{n''}} K_{m''}(\{m''\}_{n''})\right] . \tag{4.20}$$

We may call a set of variables $\{n\}$, (4.8), a *connected set* if (4.17) does not hold for any separation of it. If a certain choice of $\{n'\}$ and $\{n''\}$ makes (4.17) valid, then it is called an

unconnected set. A *cluster cumulant function* $K_n(\{n\}_N)$ defined by Eq. (4.13) for a connected set or cluster $\{n\}_N$ may be called a (connected) cluster cumulant function. Theorem IV shows that the cluster expansion (4.12) consists only of connected cluster cumulant functions, the unconnected ones being automatically dropped out.

§ 5. Generalized Exponential Functions

The concepts of moment generating functions and cumulant functions as introduced in § 2 will be generalized in the following. The generalization may be made in two directions. The first is generalization of the definition of exponential functions and the second is generalization of the definition of average.

Let us begin with the first problem. An exponential function exp x, is defined by a power series,

$$e^x = \sum_{n=0}^{\infty} \frac{1}{n!} x^n . \tag{5.1}$$

This definition gives the relation,

$$e^{x+y} = \sum_{n=0}^{\infty} \frac{1}{n!} (x+y)^n \tag{5.2a}$$

$$= \sum \frac{1}{p!} x^p \sum \frac{1}{q!} y^q \tag{5.2b}$$

$$= e^x \cdot e^y . \tag{5.2c}$$

Now let x and y be q-number variables rather than c-numbers. Then x and y are not in general commutable, so that some prescription is needed in defining a mixed product of x and y. For example, a totally symmetrized product of x and y may be defined by

$$\{x^p y^q\}_s = \frac{p! \, q!}{(p+q)!} \sum_P P(x^p y^q) \tag{5.3}$$

where the operation P means permutation of the order of x and y in all possible ways. The total number of such permutations is $(p+q)!/p!q!$. Ordinarily an exponential function exp $(x+y)$ for non-commuting x and y is defined by (5.2a), which means that any product involving x and y is totally symmetrized, so that the equality (5.2c) does not hold. It is, however, important to notice here that the equalities (5.2) could be recovered if we introduce the symmetrization operation S by the definition

$$S(x^p y^q) \equiv \{x^p y^q\}_s \tag{5.4}$$

and interpret the equation

$$e^{x+y} = e^x e^y$$

as

$$e^{x+y} \equiv e_S^{x+y} \equiv S(e^{x+y}) = S(e^x \cdot e^y) . \tag{5.5}$$

Thus a symmetrized exponential may be explicitly denoted by the subscript S to e or exp. Eq. (5.2) is interpreted by this prescription as

$$e_S^{x+y} = \sum \frac{1}{n!} S(x+y)^n$$

$$= \sum_p \sum_q \frac{1}{p!q!} S(x^p y^q) = S(e^x \cdot e^y) . \tag{5.6}$$

There can be many ways to generalize exponential functions by introducing a certain prescription of ordering. For instance, instead of symmetrization, a total separation of x

and y may be prescribed. That is, in a product involving x and y, x comes always to the left of y. If this ordering is denoted by O, such an exponential is defined by

$$e_O^{x+y} \equiv e^x e^y .$$ (5.7)

Eq. (5.2) then means

$$e_O^{x+y} = \sum \frac{1}{n!} O(x+y)^n$$

$$= \sum \frac{1}{p!} x^p \sum \frac{1}{q!} y^q = e^x \cdot e^y .$$

A well-known example of ordered exponential is that ordered in chronological order. If the variable $X(t)$ is a function of the real parameter t, the chronological order gives the prescription that $X(t)$'s should be ordered in increasing t from the left or from the right. Thus, for instance, we have

$$\exp_O \int_0^t X(t')dt' = \sum_{n=0}^\infty \frac{1}{n!} \int_0^t dt_1 \cdots \int_0^t dt_n O(X(t_1) \cdots X(t_n))$$

$$= \sum_{n=0}^\infty \int_0^t dt_1 \int_0^{t_1} dt_2 \cdots \int_0^{t_{n-1}} dt_n X(t_1)X(t_2) \cdots X(t_n)$$ (5.8)

if the prescription O means the ordering of $X(t)$'s from the right in increasing t.

Another example of ordering is given by

$$\exp_O \int \int \phi^+(x)\phi(y)A(x, y)dxdy$$

$$\equiv \sum_{n=0}^\infty \frac{1}{n!} \int dx_1 \cdots \int dx_n \int dy_1 \cdots \int dy_n O(\phi^+(x_1)\phi(y_1)\phi^+(x_2)\phi(y_2) \cdots \phi^+(x_n)\phi(y_n)A(x_1 y_1) \cdots A(x_n y_n))$$

$$\equiv \sum_{n=0}^\infty \frac{1}{n!} \int dx_1 \cdots \int dx_n \int dy_1 \cdots \int dy_n \phi^+(x_1) \cdots \phi^+(x_n)\phi(y_n) \cdots \phi(y_1)A(x_1 y_1) \cdots A(x_n y_n)$$ (5.9)

where ϕ^+ and ϕ are quantized wave functions and $A(x, y)$ is simply a c-number function of x and y. This is useful in statistical quantum-statistical mechanics of many-body systems.

More generally an ordering prescription may be defined with respect to a discrete parameter, a continuous parameter and the different classes of variables. For such a generalized ordered exponential function, Eq. (5.2) may be regarded as valid by a suitable interpretation.

We may also introduce different kinds of generalized exponential functions. An example is provided by

$$\exp_L (\sum_{i=1}^N X_i) = \sum_{n=0}^\infty \frac{1}{n!} L(\sum X_i)^n$$

$$= 1 + \sum_i X_i + \sum_{(ij)} X_i X_j + \cdots + \sum_{\{n\}_N} X_{i_1} \cdots X_{i_n} + \cdots$$

$$= \prod_{i=1}^N (1+X_i)$$ (5.10)

where the operation L may be called the *leveling* because it levels off a product of X_i's by *erasing* out those terms in which any of X's appears with a power higher than one. For this *leveled* exponential function we have

$$\exp_L (\sum X_i + \sum Y_j) = \exp_L (\sum X_i) \exp_L (\sum Y_j)$$

$$= \prod (1+X_i) \prod (1+Y_j)$$ (5.11)

Thus Eq. (5.2) again is applicable by interpreting the products always to follow the leveling operation.

In Eq. (5.10) we did not give a prescription for the ordering of X's. But in some cases

it becomes necessary to define an ordering prescription, so that it may be a leveled and ordered exponential.

§ 6. Generalized Moments, Cumulants and Cumulant Functions

If a certain *average* operation A is defined for the variables $(X_1, X_2, \cdots, X_N)$ in such a way to allow convergent moments and a moment generating function, we may write

$$A\left(\exp \sum_{i=1}^{N} \xi_i X_i\right) \equiv \left\langle \exp \sum_{i=1}^{N} \xi_i X_i \right\rangle$$

$$= \sum_{\nu_1=0}^{\infty} \cdots \sum_{\nu_N=0}^{\infty} \Pi \frac{\xi_j{}^{\nu_j}}{\nu_j!} A(X_1{}^{\nu_1} \cdots X_N{}^{\nu_N}). \tag{6.1}$$

This average is assumed to satisfy the *normalization* condition

$$A(1)=1. \tag{6.2}$$

When X_j's are q-numbers, it is sometimes necessary to introduce an average operation which leaves the averages of these variables still q-numbers. In such cases the moment generating function, (6.1), is still a q-number and the unity on the right hand side of Eq. (6.2) must be the unity in the field of such q-numbers. Thus it is important to note that the concept of *average* may be quite freely extended as far as a certain condition of convergence and the normalization condition (6.2) are fulfilled.

With this possibility in mind we now remark that the moments and cumulants may be widely generalized. Let X_j's be c-number or q-numbers and let the exponential function

$$\exp_Q \left(\sum \xi_i X_i\right) \equiv Q \exp \left(\sum \xi_i X_i\right) \tag{6.3}$$

be a generalized exponential function which is defined by a prescription Q, which operates on any product of X's to give, for instance, certain ordering or leveling. The parameters $\xi_1, \cdots, \xi_N$ are regarded simply as ordinary c-numbers. Then Eq. (2.2) may now be interpreted as

$$\langle \exp_Q \sum \xi_i X_i \rangle = \sum_{\nu_1} \cdots \sum_{\nu_N} \Pi_i \frac{\xi_i{}^{\nu_i}}{\nu_i!} A(Q \Pi X_i{}^{\nu_i}) \tag{6.4a}$$

$$= \exp_Q \left\{ \sum_{\nu_1 \cdots \nu_N}' \Pi \frac{\xi_i{}^{\nu_i}}{\nu_i!} \kappa(\nu_1, \cdots, \nu_N) \right\} \tag{6.4b}$$

$$= \exp_Q \{K(\xi_1, \cdots, \xi_N)\}. \tag{6.4c}$$

Eq. (6.4a) defines the generalized moments

$$\mu(\nu_1, \cdots, \nu_N) = A(Q \Pi X_i{}^{\nu_i}). \tag{6.5}$$

The second equality, (6.4a, b) defines generalized cumulants. If the average operation A is such that gives simply c-number moments, there is no need to retain the subscript Q for the exponential in (6.4b) since it is an ordinary exponential, and so Eq. (2.7) as it is holds for the cumulants $\kappa(\nu)$. If, on the other hand, the average operation A leaves the moments still q-numbers, then it becomes necessary to specify the exponential in (6.4b) by the subscript Q, for it actually means

$$\exp_Q \sum_{\nu_1 \cdots \nu_N}' \Pi \frac{\xi_i{}^{\nu_i}}{\nu_i!} \kappa(\nu) \equiv \sum_{n=0}^{\infty} \frac{1}{n!} Q \left\{ \sum \Pi \frac{\xi_i{}^{\nu_i}}{\nu_i!} \kappa(\nu) \right\}^n$$

$$= \sum_{\nu_1 \cdots \nu_N} \Pi \frac{\xi_i{}^{\nu_i}}{\nu_i!} \mu(\nu). \tag{6.6}$$

The operation Q here takes care of necessary arrangements of q-number $\kappa(\nu)$'s. By solving Eq. (6.6) we have again Eq. (2.7) but with the complication that all products of X's in the

moments and products of q-number moments have to follow the prescription Q.

Therefore Eqs. (6.4a, b, c) can be used even with these generalizations of exponential functions and the definition of average. It now remains to see if the Theorems I–IV also allow such generalizations, but it is actually almost evident.

The concept of *connected* or *unconnected* is based upon the relation (3.1) for unconnected sets $\{X'\}$ and $\{X''\}$. But we have seen already in § 5 that Eq. (5.2) may be regarded as true if the prescription is kept in mind in order to define the exponential functions properly. Thus Eqs. (3.1a, b) are interpreted as

$$\langle \exp_Q \sum \xi X \rangle = Q \langle \exp_Q \sum \xi' X' \rangle \langle \exp_Q \sum \xi'' X'' \rangle$$
$$= \exp_Q \{ K_1(\xi') + K_2(\xi'') \} \tag{6.7}$$

which shows that the powers of ξ' and ξ'' will never mix in the cumulant function if the sets $\{X'\}$ and $\{X''\}$ are unconnected. Thus Theorem I holds quite generally.

Theorem II holds evidently irrespective of the nature of the variable $X(t)$ or the exponential function or the definition of average, since it only depends on the continuous nature of the parameter t.

Theorem III states a particular kind of rearrangement of the cumulant series. Any kind of rearrangement can be made for generalized cumulant series if the proper prescription is always followed strictly. Thus (4.13) may be applied if $\log M_l(\{l\}_N)$ in it are properly constructed. This can be done by expanding the logarithmic functions in powers of the moments and all kinds of products in the expressions of the moments and the products of moments are dealt with the given prescription.

The above arguments are summarized by

Theorem V: The concepts of moments, cumulants and cumulant functions can be generalized to c-number or q-number variables for which an average operation is defined and a certain prescription for defining their products is given. Theorems I–IV allow such generalization.

Eq. (6.4) may be written as follows:

$$\left\langle \exp_Q \sum_{i=1}^{N} \xi_i X_i \right\rangle = \exp_Q \left\langle \exp_Q \left(\sum \xi_i X_i \right) - 1 \right\rangle_c \tag{6.8a}$$

$$= \exp_Q \left\{ \sideset{}{'}\sum_{\nu_1 \cdots \nu_N} \Pi \frac{\xi_i^{\nu_i}}{\nu_i !} \langle Q \, \Pi \, X_i^{\nu_i} \rangle_c \right\} \tag{6.8b}$$

where $\langle Q \, \Pi \, X_i^{\nu_i} \rangle_c$ is a cumulant which may be explicitly expressed in terms of the moments as

$$\langle Q \, \Pi_j \, X_j^{\nu_j} \rangle_c = - \Pi \, \nu_j ! \sum_{l=1}^{n} \sum_{\substack{l \\ \sum\limits_{i=1} k_i m_{ij} = \nu_j}} \left(\sum k_i - 1 \right)! (-)^{\Sigma k_i}$$

$$\times Q \prod_{i=1}^{l} \frac{1}{k_i !} \left\{ \frac{\langle Q \, \Pi \, X_j^{m_{ij}} \rangle}{\Pi \, m_{ij}!} \right\}^{k_i} \tag{6.9}$$

by generalizing Eq. (2.6). Thus interpreting the symbolical expression (6.8a), the operation " c ", which constructs cumulants out of the variables inside, must be done first, then the operation Q for the products of the variables, next the average operations, and finally the operation Q is again considered, if necessary. In this way one can construct formal expressions of the cumulants and cumulant functions.

A few example may serve for illustration. Let the variable $X(t)$ be a q-number for which the ordering (5.8) is prescribed. If an average operation A is defined which gives c-numbers after average, we may write

$$\left\langle \exp_O \int_0^t X(t')dt' \right\rangle = \exp \left\langle \exp_O \int_0^t X(t')dt' - 1 \right\rangle_c = \exp K(t). \tag{6.10}$$

Therefore the cumulant function is given by

$$K(t)=\sum_{n=1}^{\infty}\frac{1}{n!}\int_0^t dt_1\cdots\int_0^t dt_n\langle OX(t_1)\cdots X(t_n)\rangle_c$$

$$=\sum_{n=1}^{\infty}\int_0^t dt_1\int_0^{t_1}dt_2\cdots\int_0^{t_{n-1}}dt_n\langle X(t_1)\cdots X(t_n)\rangle_c . \qquad (6.11)$$

The cumulant averages may be represented by moments as shown by (2.9). For example,

$$\langle OX(t_1)X(t_2)\rangle_c=\langle X(t_1)X(t_2)\rangle-\langle X(t_1)\rangle\langle X(t_2)\rangle \qquad \text{if} \quad t_1>t_2$$

$$=\langle X(t_2)X(t_1)\rangle-\langle X(t_1)\rangle\langle X(t_2)\rangle \qquad \text{if} \quad t_1<t_2 ,$$

$$\langle OX(t_1)X(t_2)X(t_3)\rangle_c=\langle X(t_1)X(t_2)X(t_3)\rangle-\langle X(t_1)\rangle\langle X(t_2)X(t_3)\rangle$$

$$-\langle X(t_2)\rangle\langle X(t_1)X(t_3)\rangle-\langle X(t_3)\rangle\langle X(t_1)X(t_2)\rangle$$

$$+2\langle X(t_1)\rangle\langle X(t_2)\rangle\langle X(t_3)\rangle , \qquad \text{if} \quad t_1>t_2>t_3 . \qquad (6.12)$$

If we have

$$X(t)=Y(t)Z(t) \qquad (6.13)$$

where $Y(t)$ and $Z(t)$ are both q-numbers and the average operation A is defined only for $Y(t)$'s, then the expressions (6.10)-(6.12) become q-numbers so that one has to keep a certain prescription for ordering $Z(t)$'s. Thus, if

$$O=O_Y O_Z \qquad (6.14)$$

is an ordering prescription, O_Y and O_Z being those for $Y(t)$'s and $Z(t)$'s respectively, we have

$$\left\langle \exp_O\left\{\int_0^t Y(t')Z(t')dt'\right\}\right\rangle=\exp_{O_Z}\left\langle \exp_O\left(\int_0^t Y(t')Z(t')dt'\right)-1\right\rangle_c$$

$$=\exp_{O_Z}K(t) . \qquad (6.15)$$

The cumulant function $K(t)$ is defined here for the exponential function ordered with respect to $Z(t)$'s as

$$K(t)=\sum_{n=1}^{\infty}\frac{1}{n!}\int_0^t dt_1\cdots\int_0^t dt_n\langle OY(t_1)Z(t_1)\cdots Y(t_n)Z(t_n)\rangle_c$$

$$=\sum_{n=1}^{\infty}\int_0^t dt_1\int_0^{t_1}dt_2\cdots\int_0^{t_{n-1}}dt_n\left\langle Y(t_1)Z(t_1)\cdots Y(t_n)Z(t_n)\right\rangle_c . \qquad (6.16)$$

For a leveled exponential (5.10) with a c-number average, we have the following example:

$$\langle \exp_L \sum \xi_i X_i\rangle=\exp\langle \exp_L \sum \xi_i X_i-1\rangle_c$$

$$=\exp K(\xi) .$$

$$K(\xi)=\sum_{\nu_1}\cdots\sum_{\nu_N}{}'\prod\frac{\xi_i^{\nu_i}}{\nu_i!}\langle L \prod X_i^{\nu_i}\rangle_c. \qquad (6.17)$$

Following the instruction given by Eq. (6.9) the cumulants are explicitly calculated. For example,

$$\langle LX_j\rangle_c=\langle X_j\rangle$$

$$\langle LX_j^2\rangle_c=\langle LX_j^2\rangle-\langle LX_j\rangle^2=-\langle X_j\rangle^2$$

$$\langle LX_j^3\rangle_c=\langle LX_j^3\rangle-3\langle LX_j\rangle\langle LX_j^2\rangle+2\langle LX_j\rangle^3=2\langle X_j\rangle^3$$

$$\langle LX_iX_j\rangle_c=\langle X_iX_j\rangle-\langle X_i\rangle\langle X_j\rangle \qquad \text{if} \quad i\neq j$$

$$\langle LX_i^2X_j\rangle_c=\langle LX_i^2X_j\rangle-2\langle LX_i\rangle\langle LX_iX_j\rangle-\langle LX_i^2\rangle\langle LX_j\rangle+2\langle LX_i\rangle^2\langle LX_j\rangle$$

$$=-2\langle X_i\rangle\langle X_iX_j\rangle+2\langle X_i\rangle^2\langle X_j\rangle \qquad (6.18)$$

and so on. But this example may be calculated more conveniently using the cluster cumulants by THEOREM III. Note that

$$\exp_{\mathrm{L}} \left(\sum_{i=1}^{N} \xi_i X_i \right) = \Pi \, \exp_{\mathrm{L}} \xi_i X_i = \Pi \, (1 + \xi_i X_i) \tag{6.19}$$

and so

$$M_n(\{n\}_N) = \langle \Pi_{\{n\}}(1 + \xi_i X_i) \rangle \tag{6.20}$$

By THEOREM III, we have

$$\left\langle \exp_{\mathrm{L}} \sum_{i}^{N} \xi_i X_i \right\rangle = \exp K_N = \exp \sum K_n \tag{6.21}$$

where

$$K_1 = \sum_j \log M_1(X_j) = \sum_j \log (1 + \xi_j \langle X_j \rangle)$$

$$= \sum_{r=1}^{\infty} (-)^{r-1} \xi_j{}^r \langle X_j \rangle^r / r \,,$$

$$K_2 = \sum_{(ij)} \log \{ M_2(X_i X_j) / M_1(X_i) M_1(X_j) \}$$

$$= \sum_{(ij)} \log \left\{ 1 + \xi_i \xi_j \frac{\langle X_i X_j \rangle_{\mathrm{c}}}{(1 + \xi_i \langle X_i \rangle)(1 + \xi_j \langle X_j \rangle)} \right\} . \tag{6.22}$$

§7. Applications to Physical Problems

In order to illustrate the usefulness of the present method, we shall briefly sketch a few examples of physical application. More detailed treatments will be published elsewhere.

(1) *Ursell-Mayer Expansion of Classical Gases*[4] :   For a classical gas, which is assumed for simplicity to consist of structureless mass points, the partition function Z_N is written as

$$Z_N = Z_N{}^0 \langle \exp (-\beta U\{N\}) \rangle \tag{7.1}$$

where

$$Z_N{}^0 = \left(\frac{2\pi m k T}{h^2} \right)^{(3/2)N} V^N / N! $$

is the kinetic part of partition function and $U\{N\}$ represents the interaction potential of N molecules. The average in (7.1) is defined by

$$\langle \exp (-\beta U\{N\}) \rangle = V^{-N} \int \cdots \int \exp (-\beta U\{N\}) d\{N\} \tag{7.2}$$

where $d\{N\}$ means a volume element in the configuration space of N molecules. The cumulant function for this is nothing but the interaction free energy F_N', so that the free energy may be expanded in cluster cumulant functions by Eq. (4.12). The result must be identical with the known Ursell-Mayer virial expansion.

If the molecular interaction consists of two particle interactions, it is very convenient to write

$$\exp (-\beta U\{N\}) = \exp (-\beta \sum_{(ij)} u_{ij})$$

$$= \Pi_{(ij)} (1 + f_{ij})$$

$$= \exp_{\mathrm{L}} (\sum_{(ij)} f_{ij}) \tag{7.3}$$

where u_{ij} is the pair interaction potential and f_{ij} is defined by

$$f_{ij} = \exp (-\beta u_{ij}) - 1$$

as usual. In the last expression of (7.3) we have introduced a leveled exponential. Therefore one may write

$$
\begin{aligned}
-\beta F_{N}' &= \log \langle \exp(-\beta U\{N\}) \rangle \\
&= \log \exp \langle \exp_{\mathrm{L}} (\sum_{(ij)} f_{ij}) - 1 \rangle_{\mathrm{c}} \\
&= \langle \exp_{\mathrm{L}} (\sum_{(ij)} f_{ij}) - 1 \rangle_{\mathrm{c}} .
\end{aligned}
\tag{7.4}
$$

This is similar to the example (6.17) discussed previously. The general term of expansion of (7.4) has the form

$$
\langle \Pi \, f_{ij} \rangle_{\mathrm{c}}
\tag{7.5}
$$

where the product Π is taken over a specified set of pair bonds. By Theorem IV, the set of these pair bonds must be *connected*, for otherwise (7.5) will vanish. It is obvious that the set of bonds is unconnected if the bonds do not connect all of the particles involved, but a *connected set* of bonds means more in this case. Since the pair interaction u_{ij} is assumed to depend only on the relative position of two particles, the condition (4.17) for an unconnected set allows us to identify the concept of a *connected set* with the *irreducibility* of a cluster integral familiar in Mayer's theory. Furthermore one sees easily that

$$
\langle \Pi \, f_{ij} \rangle_{\mathrm{c}} = \langle \Pi \, f_{ij} \rangle \{ 1 + O(V^{-1}) \}
\tag{7.6}
$$

in the limit of $V \to \infty$, so that the subscript c on (7.5) may be dropped by remembering, however, that the bonds are connected. If n molecules are involved in the set of bonds in question, the expression (7.5) is in fact of the order of V^{-n+1} as is seen from the definition of average: i. e.

$$
\begin{aligned}
\langle \Pi \, f_{ij} \rangle &= V^{-N} \int \cdots \int \Pi \, f_{ij} d\{N\} \\
&= V^{-n} \int \cdots \int \Pi \, f_{ij} d\{n\} \\
&= V^{-(n-1)} \int \cdots \int \Pi \, f_{ij} d\{n-1\} .
\end{aligned}
\tag{7.7}
$$

In the expansion of the cumulant (7.5) in lower moments by Eq. (2.9) those lower order moments give only the lower order terms in $1/V$, because some molecules must be counted at least twice or more in different sets of bonds corresponding to a decomposition to a number of lower moments. Therefore Eq. (7.4) gives

$$
-\beta F_{N}' = \sum_{n=1}^{N-1} \binom{N}{n+1} \sum \langle \Pi \, f_{ij} \rangle_{\text{connected}}^{(n+1)}
\tag{7.8}
$$

where the superscript $(n+1)$ indicates the number of molecules in the connected set of bonds defining $\Pi \, f_{ij}$. Taking the limit of $N \to \infty$ and $V \to \infty$, Eq. (7.8) will give

$$
-\beta F_{N}' = N \sum_{n=1} \frac{\beta_n}{n+1} v^{-n} , \qquad (v = V/N)
\tag{7.9}
$$

where the irreducible cluster integral β_n is defined by

$$
\sum \langle \Pi \, f_{ij} \rangle_{\text{connected}}^{(n+1)} = \frac{n!}{V^n} \beta_n .
\tag{7.10}
$$

This in fact coincides with the usual definition of β_n.

   (2) *Ursell-Mayer Expansion for Quantum Gases:* The quantum-mechanical partition function may be written as[5]

$$Z_N = Z_N{}^0 \left\langle \exp_+ \left(-\int_0^\beta U_N(\lambda) d\lambda \right) \right\rangle .$$ (7.11)

The kinetic part of the partition function is defined by

$$Z_N{}^0 = \text{Trace} \, (\exp -\beta K_N)$$

where K_N is the pure kinetic energy. $U_N(\lambda)$ in (7.11) is defined by

$$U_N(\lambda) = e^{\lambda K_N} U_N \, e^{-\lambda K_N}$$ (7.12)

and there the exponential function is ordered *chronologically* in λ. The average in (7.11) is defined as

$$\left\langle \exp_+ \left(-\int_0^\beta U_N(\lambda) d\lambda \right) \right\rangle = (Z_N{}^0)^{-1} \, \text{Trace} \, e^{-\beta K_N} \exp_+ \left(-\int_0^\beta U_N(\lambda) d\lambda \right)$$ (7.13)

so we have

$$-\beta F_N' = \log Z_N / Z_N{}^0$$

$$= \left\langle \exp_+ \left(-\int_0^\beta U_N(\lambda) d\lambda \right) - 1 \right\rangle_c$$ (7.14)

for the interaction free energy. One may introduce here a leveled exponential if U_N consists of pair interactions. Namely Eq. (7.14) can be written as

$$-\beta F_N' = \langle \Pi_+ (1 + f_{ij}) - 1 \rangle_c$$
$$= \langle \exp_{+\text{L}} (\sum_{(ij)} f_{ij}) - 1 \rangle_c$$ (7.15)

with

$$f_{ij} = \exp \left(-\int_0^\beta u_{ij}(\lambda) d\lambda \right) - 1.$$ (7.16)

Here the subscript $+$ takes care of ordering in λ and L means the leveling. In Eq. (7.15) f_{ij}'s actually involve the integration parameters λ, and so they ought to be ordered. The general principle of constructing the cluster cumulant functions is similar to the classical case, but the concept of irreducibility has to be modified, because the interference effects are present due to the wave nature of the particles. Thus, for instance, $f_{12} f_{23}$ is not reducible any more. We shall not go into further details here, but it is rather easy to write down some formal expressions, if necessary, for the first few terms of the virial expansion from (7.15).

(3) *Perturbation Series in Quantum Mechanics*: If the Hamiltonian of a system is given by

$$\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_1$$ (7.17)

where $\mathcal{H}_0$ is the unperturbed Hamiltonian and $\mathcal{H}_1$ is the the perturbation, the quantum-mechanical equation of the state vector takes the form

$$i \frac{\partial \phi}{\partial t} = \mathcal{H}_1(t) \phi \,, \qquad \mathcal{H}_1(t) = \exp (it \mathcal{H}_0) \mathcal{H}_1 \exp (-it \mathcal{H}_0)$$ (7.18)

in the interaction representation. This is integrated to

$$\phi(t) = U(t) \phi(0)$$ (7.19)

with the use of the transformation $U(t)$ as given by, with the use of an ordered exponential,

$$U(t) = \exp (it \mathcal{H}_0) \exp (-it \mathcal{H}) = \exp_+ \left(i \int_0^t \mathcal{H}_1(t') dt' \right).$$ (7.20)

The diagonal matrix element of $U(t)$ is, in the representation diagonalizing $\mathscr{H}_0$,

$$(\alpha|U(t)|\alpha) = \left(\alpha \left| \exp_+\left(i\int_0^t \mathscr{H}_1(t')dt' \right) \right| \alpha \right)$$

$$= \left\langle \exp i\int_0^t \mathscr{H}_1(t')dt' \right\rangle \qquad (7.21)$$

where the average $\langle \ \rangle$ is defined as the diagonal element to the unperturbed state α. Thus we may write

$$(\alpha|U(t)|\alpha) = \exp\left\langle \exp_+\left(i\int_0^t \mathscr{H}_1(t')dt' \right) - 1 \right\rangle_c = e^{\gamma(t)}. \qquad (7.22)$$

The cumulant function $\gamma(t)$ is now written as

$$\gamma(t) = \sum_{n=1}^{\infty} i^n \int_0^t dt_1 \int_0^{t_1} dt_2 \cdots \int_0^{t_{n-1}} dt_n \langle \mathscr{H}_1(t_1) \cdots \mathscr{H}_1(t_n) \rangle_c. \qquad (7.23)$$

By Laplace transformation, it becomes

$$\int_0^\infty \gamma(t)\, e^{-(s-iE_\alpha)t} dt = \int_0^\infty e^{-(s-iE_\alpha)t} dt \sum_{n=1}^{\infty} i^n \int_0^t dt_1 \cdots \int_0^{t_{n-1}} dt_n\, e^{-it_1 E_\alpha}$$

$$\times \langle (\alpha|\mathscr{H}_1|\beta)\, e^{it_1 E_\beta - it_2 E_\beta} \beta(\beta|\mathscr{H}_1|\gamma) \cdots e^{-it_n E_\mu} \mu(\mu|\mathscr{H}_1|\alpha) \rangle_c\, e^{it_n E_\alpha}$$

$$= -\frac{1}{(s-iE_\alpha)^2}(\alpha|\Gamma(s)|\alpha) \qquad (7.24)$$

where $\Gamma(s)$ is defined by

$$\Gamma(s) = \frac{1}{i} \sum_{n=0}^{\infty} \left\langle \mathscr{H}_1 \left(\frac{i}{s-i\mathscr{H}_0} \mathscr{H}_1 \right)^n \right\rangle_c \qquad (7.25)$$

The subscript c means that the operators inside are connected. By Theorem I one sees easily that α never appears in the intermediate states. But this concept of connected cumulants can be greatly generalized if $\mathscr{H}_1$ consists of many terms as one often meets in treating many-body systems. In such cases, some graphs are seen to be unconnected even when the initial or the final state never appears at the intermediate stages[6]. Generalizations of the concept of *connected* and *unconnected* originate from Theorem I but depend greatly on the particular type of the system in question. Eq. (7.25) is a well-known formula in perturbation theories[7].

  4) *Random Frequency Modulation*: A simple oscillator with modulated frequency may be represented by the equation of motion,

$$\dot{x} = i\omega(t)x \qquad (7.26)$$

where

$$\omega(t) = \omega_0 + \omega_1(t)$$

is the frequency, ω_0 being the average frequency and $\omega_1(t)$ the frequency modulation. Therefore $\omega_1(t)$ may be assumed to be averaged to zero, i.e.,

$$\langle \omega_1(t) \rangle = 0 . \qquad (7.27)$$

Since $\omega(t)$ is now regarded as a stochastic process, the solution $x(t)$ of Eq. (7.26) as given by

$$x(t) = x(0) \exp\left(i \int_0^t \omega(t')dt' \right) \qquad (7.28)$$

defines a stochastic process derived from the basic process $\omega(t)$.

The correlation function of $x(t)$ is defined by

$$\langle x(t)x^*(0)\rangle = \langle |x(0)|^2\rangle \left\langle \exp i \int_0^t \omega(t')dt' \right\rangle \tag{7.29}$$

where we assumed that the stochastic process $\omega(t)$ is independent of the distribution of the initial value $x(0)$. We are now interested in the normalized correlation function $x(t)$, namely,

$$\varphi(t) = \langle x(t)x^*(0)\rangle / \langle |x(0)|^2\rangle$$

$$= \left\langle \exp i \int_0^t \omega(t')dt' \right\rangle \tag{7.30}$$

$$= \exp\{i\omega_0 t + \phi(t)\} . \tag{7.31}$$

Here $\phi(t)$ is the cumulant function for the stochastic process $\omega_1(t)$ as defined by

$$\phi(t) = \left\langle \exp\left(i \int_0^t \omega_1(t')dt'\right) - 1 \right\rangle_{\mathrm{c}}$$

$$= \sum_{n=1}^\infty i_n \int_0^t dt_1 \int_0^{t_1} dt_2 \cdots \int_0^{t_{n-1}} dt_n \langle \omega_1(t_1)\cdots\omega_1(t_n)\rangle_{\mathrm{c}}. \tag{7.32}$$

The effect of frequency modulation is determined by the nature of this function $\phi(t)$, which depends essentially on the relative magnitudes of the modulation amplitude of $\omega_1(t)$ and the speed of its temporal change. A well known physical example is provided by the narrowing phenomena in magnetic resonance spectrum. The modulation amplitude of $\omega_1(t)$ may be measured roughly by

$$\Delta = \langle \omega_1^2(t)\rangle^{1/2} \tag{7.33}$$

and the speed of the modulation by

$$\tau_c = \int_0^\infty \langle \omega_1(t)\omega_1(0)\rangle dt / \Delta^2 \tag{7.34}$$

which may be called the correlation time of the modulation. If the modulation is *slow* as represented by the condition

$$\Delta \cdot \tau_c \gg 1 , \tag{7.35}$$

it may be called quasi-static, and the spectrum of $x(t)$ reflects directly the distribution of the modulated frequency $\omega(t)$. On the other hand, if the modulation is *fast* so that the condition

$$\Delta \cdot \tau_c \ll 1 \tag{7.36}$$

is fulfilled, the spectrum of $x(t)$ is *narrowed* and it approaches to a Lorentzian form with the half-width,

$$1/\gamma = \Delta^2 \cdot \tau_c . \tag{7.37}$$

More detailed treatments of this problem are referred to preceding articles of the present author.[2,8]

5) *Random Perturbations in Dynamical Systems*: Let us briefly remark an important resemblance between the foregoing example of random frequency modulation and the dynamical behaviour of a system which suffers a perturbation of some stochastic nature. The dynamical motion of a system may be described in terms of the Hamiltonian equation of motion,

$$\frac{dp_j}{dt} = -\frac{\partial \mathscr{H}}{\partial q_j} , \qquad \frac{dq_j}{dt} = \frac{\partial \mathscr{H}}{\partial p_j} . \tag{7.38}$$

One may, however, equally use the Liouville equation in the phase space,

$$\frac{\partial f(p, q; t)}{\partial t} = iLf \equiv \sum_j \left(\frac{\partial \mathscr{H}}{\partial q_j} \frac{\partial}{\partial p_j} - \frac{\partial \mathscr{H}}{\partial p_j} \frac{\partial}{\partial q_j} \right) f \tag{7.39}$$

for the distribution function $f(p, q; t)$. L in (7.39) is a linear differential operator which constructs the Poisson bracket for $\mathscr{H}$ and f. If the unperturbed motion of the system is determined by the Hamiltonian $\mathscr{H}_0$, the interaction representation introduced by the transformation

$$f(p, q; t) = e^{iL_0 l} \hat{f}(p, q; t) \tag{7.40}$$

transforms Eq. (7.39) into

$$\frac{\partial \hat{f}}{\partial t} = i \hat{L}_1(t) \hat{f}. \tag{7.41}$$

L_0 in Eq. (7.40) is the unperturbed Liouville operator, and $\hat{L}_1(t)$ in Eq. (7.41) is a modulation,

$$\hat{L}_1(t) = e^{-iL_0 t} L_1 e^{iL_0 t} \equiv e^{-L_0 t}(L - L_0) e^{iL_0 t}. \tag{7.42}$$

Similarly, the equation of motion of the density matrix of a quantum-mechanical system,*

$$i \frac{\partial \rho}{\partial t} = [\mathscr{H}, \rho] \equiv \mathscr{H}^\times \rho \tag{7.43}$$

may be transformed into the interaction representation

$$\frac{\partial \hat{\rho}}{\partial t} = \frac{1}{i} [\hat{\mathscr{H}}_1(t), \hat{\rho}] \equiv \frac{1}{i} \hat{\mathscr{H}}_1(t)^\times \hat{\rho} \tag{7.44}$$

by the transformation

$$\rho(t) = \exp(-i \mathscr{H}_0^\times t) \hat{\rho}(t) \equiv \exp(-i \mathscr{H}_0 t) \hat{\rho} \exp(i \mathscr{H}_0 t). \tag{7.45}$$

Eqs. (7.41) and (7.44) are quite similar to Eq. (7.26). The only difference is that $\omega_1(t)$ in Eq. (7.26) is now replaced by operators operating on functions in the phase space or the space of density matrices. The formal integral of (7.41), for instance, is written as

$$\hat{f}(t) = \exp_0 \left(i \int_0^t \hat{L}_1(t') dt' \right) \cdot f(0). \tag{7.46}$$

Similarly for Eq. (7.44) we have

$$\hat{\rho}(t) = \exp_0 \left(i^{-1} \int_0^t \hat{\mathscr{H}}_1(t')^\times dt' \right) \cdot \rho(0). \tag{7.47}$$

In the above expressions the exponential functions have to be ordered chronologically. If the perturbation $\hat{L}_1(t)$ or $\hat{\mathscr{H}}_1(t)$ can be regarded as stochastic in some sense and if the average can be defined for the ensembles of such stochastic processes, $\hat{f}(t)$ or $\hat{\rho}(t)$ as given by (7.46) or (7.47) may be averaged. The results may be expressed in the form

$$\langle \hat{f}(t) \rangle = (\exp_0 \mathfrak{L}(t)) \cdot f(0) \tag{7.48}$$

or

$$\langle \hat{\rho}(t) \rangle = (\exp_0 \mathfrak{R}(t)) \cdot \hat{\rho}(0) \tag{7.49}$$

where $\mathfrak{L}(t)$ or $\mathfrak{R}(t)$ is a generalized cumulant function which is generally a q-number and the exponential in each case has to be ordered by a certain prescription. The temporal

---

* The operator $a^\times$ is defined by $a^\times b \equiv [a, b]$. This is conveniently used in dealing with quantum-mechanical equations of motion[9].

behaviour of the cumulant function will determine the stochastic nature of the dynamical system. This may be understood by an analogy to the concept of random modulation and the narrowing condition mentioned previously. Usual perturbational treatment or more generally the damping-theoretical treatment corresponds to the case where the narrowing condition in some sense is satisfied.[2] It is also clear that Eq. (7.18) for the state vector can be looked from the same point of view.

6) *Relaxation Function*: If a force $F(t)$ acting on a system has the variable A as its conjugate, the linear effect of $F(t)$ on a dynamical variable B of the system, which is assumed to be nearly in equilibrium at a temperature $T=1/k\beta$, is described by an addmittance $\chi_{BA}(\omega)$ or a relaxation function $\Phi_{BA}(t)$. The latter has been shown to be given by[9]

$$\Phi_{BA}(t)=\text{Trace}\left\{\exp(-\beta\mathcal{H})\int_0^\beta A(\lambda)d\lambda B(t)\right\}\Big/\text{Trace }\exp(-\beta\mathcal{H})\ . \tag{7.50}$$

We have assumed here for simplicity

$$\langle A\rangle\equiv\text{Trace}(\exp(-\beta\mathcal{H})A)/\text{Trace}\exp(-\beta\mathcal{H})=0\ . \tag{7.51}$$

In Eq. (7.50) $A(\lambda)$ and $B(t)$ stand for

$$A(\lambda)=\exp(\lambda\mathcal{H})\ A\exp(-\lambda\mathcal{H}) \tag{7.52}$$

and

$$B(t)=\exp(it\mathcal{H}/\hbar)\ B\exp(-it\mathcal{H}/\hbar) \tag{7.53}$$

respectively.

For calculation of the relaxation function (7.50), the following method may be useful. Let us introduce a generating function $Z(a,b,t)$ by the definition,

$$Z(a,b,t)=\text{Trace }\exp-(\beta\mathcal{H}-aA)\exp(bB(t)) \tag{7.54}$$

for which we see that

$$\frac{\partial Z(a,b,t)}{\partial a}\bigg|_{a=b=0}=\text{Trace }\exp(-\beta\mathcal{H})\int_0^\beta A(\lambda)d\lambda$$
$$=\beta\text{ Trace }\exp(-\beta\mathcal{H})A=0 \tag{7.55}$$

and

$$\frac{\partial^2 Z(a,b,t)}{\partial a\partial b}\bigg|_{a=b=0}=\text{Trace }\exp(-\beta\mathcal{H})\int_0^\beta A(\lambda)d\lambda B(t)$$
$$=Z(0,0,0)\ \Phi_{BA}(t)\ . \tag{7.56}$$

Now the Hamiltonian $\mathcal{H}$ of the system is assumed to be divided into the unperturbed part $\mathcal{H}_0$ and the perturbation $\mathcal{H}_1$. Then we may writte

$$Z(a,b,t)$$
$$=\text{Trace}\left\{\exp(-\beta\mathcal{H}_0)\exp_+\left[-\int_0^\beta\{\mathcal{H}_1(\lambda)-aA(\lambda)\}d\lambda\right]\cdot\right.$$
$$\left.\exp_-i\hbar^{-1}\int_0^t\mathcal{H}_1^\times(t')dt'\cdot\exp bB^0(t)\right\} \tag{7.57}$$

by introducing the ordered exponentials

$$\exp_+\left\{-\int_0^\beta X(\lambda)d\lambda\right\}$$
$$=\sum_{n=0}^\infty(-)^n\int_0^\beta d\lambda_1\int_0^{\beta_1}d\lambda_2\cdots\int_0^{\beta_{n-1}}d\lambda_n X(\lambda_1)\cdots X(\lambda_n) \tag{7.58}$$

.and

$$\exp_- \int_0^t Y(t') dt'$$

$$= \sum_{n=0}^{\infty} \int_0^t dt_1 \int_0^{t_1} dt_2 \cdots \int_0^{t_{n-1}} dt_n Y(t_n) Y(t_{n-1}) \cdots Y(t_1) . \qquad (7.59)$$

In Eq. (7.57) $B^0(t)$ represents the unperturbed motion of B, i.e.

$$B^0(t) = \exp(it\mathscr{H}_0/\hbar) B \exp(-it\mathscr{H}_0/\hbar) .$$

The expression (7.57) of $Z(a, b, t)$ is now written as

$$Z(a, b, t)$$
$$= Z_0 \left\langle \exp_P \left\{ -\int_0^\beta (\mathscr{H}_1(\lambda) - aA(\lambda)) d\lambda + \frac{i}{\hbar} \int_0^t \mathscr{H}_1^\times(t') dt' + bB^0(t) \right\} \right\rangle$$
$$= Z_0 \exp K(a, b, t) , \qquad (Z_0 = \text{Trace} \exp(-\beta\mathscr{H}_0)) \qquad (7.60)$$

where the ordering P now prescribes the ordering of $\mathscr{H}_1(\lambda)$, $A(\lambda)$, $\mathscr{H}_1^\times(t)$ and $B^0(t)$ as explicitly indicated in (7.57). Thus $\exp^P$ here is a complex generalization of ordered exponential. By virtue of Eqs. (7.55) and (7.56) we find

$$\frac{\partial^2 Z(a, b, t)}{\partial a \partial b} \bigg|_{a=b=0} = Z_0 e^K \cdot \frac{\partial^2 K}{\partial a \, \partial b} \bigg|_{a=b=0} = Z(0, 0, 0) \cdot \frac{\partial^2 K}{\partial a \, \partial b} \bigg|_{a=b=0}$$

.so that we have

$$\Phi_{BA}(t) = \frac{\partial^2 K}{\partial a \, \partial b} \bigg|_{a=b=0}$$
$$= \left\langle P \exp\left(-\int_0^\beta \mathscr{H}_1(\lambda) d\lambda \right) \cdot \int_0^\beta A(\lambda) d\lambda \cdot \exp\frac{i}{\hbar} \int_0^t \mathscr{H}_1^\times(t') dt' \cdot B^0(t) \right\rangle_c \qquad (7.61)$$

where the ordering operation P works on every product in the expansion of the expression in the cumulant average. Eq. (7.61) has some connection to those methods used by Montroll .and Ward[10] and by Izuyama.[11] Further application of this method will be discussed else-where.

Another method of applying the cumulant expansion for calculation of the relaxation function (7.50) is to write it as

$$\Phi_{BA}(t) = \frac{\text{Trace} \exp(-\beta\mathscr{H}) \int_0^\beta A(\lambda) d\lambda \exp(i\mathscr{H}^\times t/\hbar) B}{\text{Trace} \exp(-\beta\mathscr{H}) \int_0^\beta A(\lambda) d\lambda B}$$

$$= \left\langle \exp_- \frac{i}{\hbar} \int_0^t \mathscr{H}_1^\times(t') dt' \right\rangle$$

$$= \exp\{\psi(t)\} . \qquad (7.62)$$

.Namely a particular kind of average process is introduced here for the ordered exponential $\exp_-\left\{ (i/\hbar) \int_0^t \mathscr{H}_1^\times(t') dt' \right\}$. The cumulant function $\psi(t)$ is simply a c-number, which may be .expanded as

$$\psi(t) = -\frac{i}{\hbar} \int_0^t \langle \mathscr{H}_1^\times(t') \rangle dt' - \frac{1}{\hbar^2} \int_0^t dt_1 \int_0^{t_1} dt_2 \langle \mathscr{H}_1^\times(t_2) \mathscr{H}_1^\times(t_1) \rangle_c + \cdots . \qquad (7.63)$$

The simplest approximation is obtained by terminating the series at the second order, which is allowed if a Gaussian nature can be assumed in the stochastic behaviour of the pertur-

bation $\mathscr{H}_1(t)$. Further, if a certain narrowing condition holds, the cumulant function $\phi(t)$ behaves as

$$\phi(t) \sim -t/\tau_r \qquad (7.64)$$

for practically important regions of t.

The relaxation time τ_r may be approximately estimated by

$$\frac{1}{\tau_r} = -\frac{\int_0^\infty dt \; \text{Trace} \; \exp(-\beta\mathscr{H})\int_0^\beta A(\lambda)d\lambda \cdot \ddot{B}(t)}{\text{Trace} \; \exp(-\beta\mathscr{H})\int_0^\beta A(\lambda)d\lambda B}$$

$$= \frac{\int_0^\infty dt \int_0^\beta d\lambda \; \text{Trace} \; \{\exp(-\beta\mathscr{H})\dot{A}(\lambda)\dot{B}(t)\}}{\int_0^\beta d\lambda \; \text{Trace} \; \{\exp(-\beta\mathscr{H})A(\lambda)B\}} . \qquad (7.65)$$

This approximation corresponds to the crudest Grüneisen formula in the case of electronic conduction in metals[12], but still it is very conveniently applied to various problems.[13]

§ 8. Conclusion

We have so far tried to show how greatly the concepts of cumulants and cumulant functions can be generalized and how these generalizations may be applied to physical problems. Of course, many of such generalizations and applications have been done by various authors for various problems. But such an unified point of view as developed in the present article seems to be very useful to obtain insights into sometimes very much complicated problems which often look quite different from each other. More detailed discussion of this subject is hoped to be published in the Supplement of Theoretical Physics in the near future.

References

1) See for example, H. Cramér: *Mathematical Methods of Statistics* (Princeton University Press, 1946) p. 186.

2) R. Kubo: Stochastic Theory of Line Shape and Relaxation (The Proceedings of the Scottich Universities Summer School at Newbattle Abbey, 1961, to be published).

3) E. Meeron: J. Chem. Phys. **27** (1957) 67, Appendix.

4) See for example, J. E. Mayer and M. G. Mayer: *Statistical Mechanics* (John Wiley and Sons, Inc., 1940) Chapter 13.

5) M. L. Goldberger and E. N. Adams: J. Chem. Phys. **20** (1952) 242.

6) K. A. Brueckner: Theory of Nuclear Structure (The Many Body Problem, Cours donnes a l'école d'été de physique theorique, Les Houches 1958, p. 53).

7) J. Goldstone: Proc. Roy. Soc. **A239** (1957) 267.

8) R. Kubo: Some Aspects of the Statistical-Mechanical Theory of Irreversible Processes (Lectures in Theoretical Physics. Vol. 1. Edited by Brittin and Dunham, Interscience Publishers, 1959. p. 181).

9) R. Kubo: J. Phys. Soc. Japan **12** (1957) 570.

10) E. W. Montroll and J. C. Ward: Physica **25** (1959) 423.

11) T. Izuyama: Progress of Theor. Phys. **25** (1961) 964.

12) H. Nakano: Progress of Theor. Phys. **15** (1956) 77.

13) R. Kubo and K. Tomita: J. Phys. Soc. Japan **9** (1954) 888.