JOURNAL OF
MATHEMATICAL
PHYSICS

VOLUME 1, NUMBER 5 SEPTEMBER–OCTOBER, 1960

On the General Theory of the Approach to Equilibrium. I. Interacting Normal Modes

I. PRIGOGINE AND F. HENIN*
Faculté des Sciences, Université Libre de Bruxelles, Belgium
(Received March 22, 1960)

A general method which permits the derivation of the equations which describe the approach to equilibrium correct to an arbitrary finite order in the coupling constant is presented. This method is applied in the present paper to normal modes interacting through three-phonon processes. In a subsequent paper the method will be applied to interacting particles. The distribution function is first Fourier-analyzed with respect to the angle variables. All Fourier components, except the distribution function of action variables, describe correlations among the normal modes. The formal solution for the Fourier components is studied in the limiting case of a very large number of degrees of freedom $N \to \infty$, and for large times by means of a diagram technique. Each component ρ_{3n} can be split into 2 parts: ρ_{3n}' and ρ_{3n}''; one (ρ') due to "scattering" of the normal modes satisfies diagonal differential equations. The other (ρ_{3n}'') contains the direct interaction between the normal modes involved in the corresponding correlation. It is completely determined by the functions ρ_{3n}'. The study of this set of equations enables us to study the approach to equilibrium.

1. INTRODUCTION

IN two recent papers[1,2] one of us (I. P.) and R. Balescu have studied the approach to equilibrium by a method which involves essentially the following steps:

(a) specification of the class of initial distribution functions to which the method applies (extensive and intensive variables in the thermodynamic sense may be already defined at the initial time);

(b) an expansion of the phase density in a Fourier series (in this "representation" the unperturbed Liouville operator L_0 is diagonal, and δL off-diagonal);

(c) the study of the time evolution of the Fourier coefficients by means of a perturbation method which finds its most concise expression in a diagram technique.

No special assumptions about the Hamiltonian are made; also no assumption equivalent to Van Hove's[3] diagonal singularity condition is required.[4]

In these papers[1,2] we had first considered the situations corresponding to weakly coupled gases or to low concentration. We have then shown that retaining all diagrams which give asymptotically contributions of order $\lambda(\lambda^2 t)^n$, n=arbitrary integer[5] [instead of $(\lambda^2 t)^n$], we obtain an equilibrium distribution correct to order λ. This case, as well as similar ones studied before,[6] led us to the conjecture that if all diagrams up to the order $\lambda^m(\lambda^2 t)^n$ are retained, the final equilibrium distribution function would be correct to order λ^m. This may indeed be expected provided the product $\lambda^2 t$ is finite. The time interval t in which we are interested is of the order of the relaxation time, which may be expected to be proportional to λ^{-2} multiplied by some polynomial in λ which reduces to one for $\lambda \to 0$.

We want therefore to discuss evolution equations for the Fourier coefficients which include all contributions of the form $\lambda^r(\lambda^2 t)^n$, $0 \leqslant r \leqslant m$. The method of "enumeration" of diagrams used in footnote reference 2 becomes rapidly impracticable when one considers the case $m>1$. Therefore a new, much more compact method valid for an arbitrary order λ^m is discussed in this as well as in the next paper of this series.

* Chargé de Recherches du Fonds National Belge de la Recherche Scientifique.

[1] I. Prigogine and R. Balescu, Physica **25**, 281 (1959).

[2] I. Prigogine and R. Balescu, Physica **25**, 302 (1959).

[3] L. Van Hove, Physica **21**, 517 (1955).

[4] The validity of this condition will be discussed by J. Philippot in a forthcoming paper.

[5] λ is the coupling constant in the Hamiltonian ($H=H_0+\lambda V$).

[6] I. Prigogine and F. Henin, Physica **23**, 585 (1957).

Here we shall discuss the case of normal modes interacting through anharmonic forces, while in a subsequent paper we shall discuss the case of molecules interacting through intermolecular forces (the case of "gases"). The method we shall use is the same but there are some minor differences. Indeed in the case of gases, one is concerned with the limiting procedure (N number of particles, Ω volume)

$$N\to\infty,\quad \Omega\to\infty,\quad N/\Omega=\text{concentration, finite.}\quad (1.1)$$

On the contrary, in the case of solids, the ratio N/Ω is fixed by the lattice constant. The limiting procedure then reduces to

$$N\to\infty. \quad (1.2)$$

This is necessary to eliminate surface effects in the evolution of the system and has the important consequence that the frequency spectrum becomes dense. It is only in this limit that one can speak strictly about irreversibility.

There exists also a difference in the nature of the diagrams (see Sec. 6). They are much simpler than for the case of gases. Here there are basically only two diagrams, one corresponding to creation of correlations, the other to destruction. For this reason we have considered first the case of normal modes.

We shall show (see Secs. 9 and 10) that the evolution equations valid to an arbitrary order have a very simple form when expressed in terms of diagrams. We shall split the Fourier coefficients ρ_{3n} corresponding to n 3-phonon correlations into two contributions ρ_{3n}', ρ_{3n}'' according to the nature of the diagrams which give rise to them. We shall then show that ρ_{3n}' satisfies a *diagonal* differential equation, while ρ_{3n}'' is determined by $\rho_{3n'}'$. The existence of such diagonal equations is an enormous simplification.

We had already investigated in an earlier paper the possibility of a preliminary change of variables in order to give to the equations a simpler form.[6] There are many features in common between our earlier approach and the present paper. We had however not taken into account the order in λ of the initial Fourier coefficients. Moreover, we used the diagonal singularity condition. To introduce the necessary changes finally appeared to us to be more difficult than to make a different start using more systematically our diagram technique.

The meaning of the evolution equations is discussed in detail in Sec. 13. We then study the approach to equilibrium in the lowest orders in λ. In the lowest possible approximation corresponding to the weakly coupled case, one obtains asymptotically as expected the equilibrium distribution unperturbed by the anharmonic forces. In the next two approximations, one obtains the equilibrium distribution correct respectively to order λ and λ^2. However, because of the mathematical complexity of some of the operators involved, we have not yet been able, in the case of solids, to discuss the approach to equilibrium to an arbitrary order in λ. From this point of view the situation is simpler in the case of gases. As will be shown in our subsequent paper, there it is possible to discuss completely the approach to equilibrium to an arbitrary order in λ and to establish therefore an H theorem of an extreme generality.

We shall see that this is ultimately caused by the fact that the velocity distribution function for particles interacting through velocity-independent forces is not modified by the interaction, while this is not so for the case of normal modes because the anharmonic forces when expressed in angle-action variables become action dependent.

2. LIOUVILLE OPERATOR

We shall consider three-phonon processes. The extension to higher-order processes is trivial. Using angle-action variables, the Hamiltonian of the system can be written[7] as follows:

$$\begin{aligned}H&=H_0+\lambda V\\&=\sum_k \omega_k J_k+\lambda\sum_{kk'k''}\{V_{kk'k''}\exp[i(\alpha_k+\alpha_{k'}+\alpha_{k''})]\\&\quad+3V_{kk'-k''}\exp[i(\alpha_k+\alpha_{k'}-\alpha_{k''})]+\text{c.c.}\}\\&\quad\times(J_kJ_{k'}J_{k''}/\omega_k\omega_{k'}\omega_{k''})^{\frac{1}{2}},\end{aligned}\quad (2.1)$$

where the summations over wave vectors are over half the Brillouin zone only and where we restrict ourselves to the cubic term in the anharmonicity. The Liouville operator associated with the Hamiltonian (2.1) is

$$L=L_0+\lambda\delta L, \quad (2.2)$$

$$L_0=\sum_k \omega_k\frac{\partial}{\partial\alpha_k}, \quad (2.3)$$

$$\delta L=\sum_l\left\{\frac{\partial V}{\partial J_l}\frac{\partial}{\partial\alpha_l}-\frac{\partial V}{\partial\alpha_l}\frac{\partial}{\partial J_l}\right\}. \quad (2.4)$$

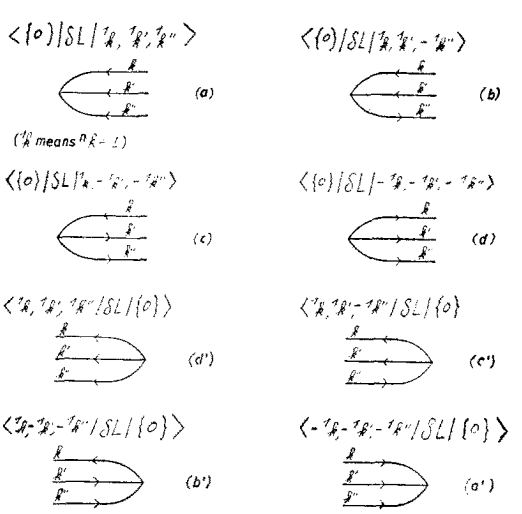


FIG. 1. Basic diagrams.

[7] I. Prigogine and J. Philippot, Physica **23**, 569 (1957).

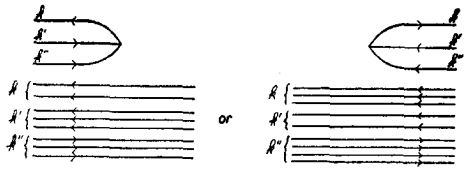


FIG. 2. Example of elementary diagram. Diagrams corresponding to the matrix element $\langle 3_k 2_{k'} - 4_{k''}\{0\}|\delta L|2_k 3_{k'} - 3_{k''}\{0\}\rangle$.

The eigenfunctions $|\{n_k\}\rangle$ of the unperturbed Liouville operator L_0 are

$$L_0|\{n_k\}\rangle = (2\pi)^{-N/2} L_0 \exp(i\textstyle\sum_k n_k\alpha_k) = i(\sum_k n_k\omega_k)|\{n_k\}\rangle, \quad (2.5)$$

where the n_k's are integers. It is easily verified that the only nonvanishing matrix elements $\langle\{n_k\}|\delta L|\{n_k'\}\rangle$ are of the form[7]

$$\langle n_k, n_{k'}, n_{k''}, \{n_l\}|\delta L|n_k\pm 1, n_{k'}\pm 1, n_{k''}\pm 1, \{n_l\}\rangle = iV_{\mp k\mp k'\mp k''}\left[\frac{n_k}{2J_k}+\frac{n_{k'}}{2J_{k'}}+\frac{n_{k''}}{2J_{k''}}\pm\frac{\partial}{\partial J_k}\pm\frac{\partial}{\partial J_{k'}}\pm\frac{\partial}{\partial J_{k''}}\right]\left(\frac{J_k J_{k'} J_{k''}}{\omega_k\omega_{k'}\omega_{k''}}\right)^{\frac{1}{2}} \quad (2.6)$$

with $\pm k\pm k'\pm k''=0$, modulo a reciprocal lattice vector. There is a formal analogy between (2.6) and the matrix elements in the quantum theory in occupation number representation; the k's correspond to momenta and the n_k's to number of particles ($n_k>0$) or holes ($n_k<0$). As in the latter theory, we shall represent an eigenfunction $|\{n_k\}\rangle$ by a set of $\{n_k\}$ lines with arrows toward the left or the right according to whether n_k is positive or negative. The effect of the operator δL in a nonvanishing matrix element is to modify the initial state by destruction or creation of three lines. The basic diagrams are given in Fig. 1. Any nonvanishing matrix element $\langle\{n_k\}|\delta L|\{n_k'\}\rangle$ can be obtained from one of those diagrams by adding either a set of $\{n_k\}$ lines to the diagram abc or d or a set of $\{n_k'\}$ lines to the diagram $a'b'c'$ or d' (see Fig. 2).

In Fig. 2, we have taken into account the fact that only the number of lines with a given wave vector matters. A "positive" (—←) and a "negative" (→—) line cancel each other just as a particle and a hole line do in the quantum theory. In these diagrams, the lines represent angle correlations between the normal modes. The eight basic diagrams of Fig. 1 are of two types: destruction ($abcd$) of a correlation between 3 normal modes or creation ($a'b'c'd'$) of such a correlation. This represents a great simplification if we compare it with the case of gases,[1] where there are six different types of basic diagrams. It might however be interesting to note that not all general diagrams can be interpreted as a "creation" or "destruction" of a correlation. For instance, in the diagram of Fig. 2, we deal with a diagram which represents a continuation of the initial correlation between the normal modes involved in the initial state. However, in the following we shall restrict ourselves to the case of homogeneous systems and the only diagrams which will appear are unambiguously of the "creation" or "destruction" type. Another difference between the problems of solids and gases is related to the fact we deal here with numbers of undiscernible excitations rather than with individual particles. Therefore, connected diagrams do not play any special role. Taking into account the fact that only the total number of lines with a given wave vector matters, any connected diagram can always be written as a nonconnected diagram (see Fig. 3). From this point of view, our diagrams are much more similar to those of the quantum theory in occupation number representation than the diagrams in the case of gases. Except for these differences, the theory for gases and solids will proceed along the same lines. The distribution function is Fourier analyzed with respect to the angle variables. The formal solution for the Fourier components is studied asymptotically; the order of magnitude of a given term depends on the topology of the corresponding diagram and the initial conditions. Those initial conditions are chosen in such a way that extensive or intensive properties of the system in the thermodynamic sense may be defined. Let us consider more closely this last point.

3. EQUILIBRIUM DISTRIBUTION FUNCTION

We consider the canonical equilibrium distribution function and expand it in a power series of λ

$$\rho^{\text{equil}} = \exp[-(H_0+\lambda V)/kT]\Big/\int\cdots\int(dJd\alpha)^N \times\exp[-(H_0+\lambda V)/kT] = \frac{\Pi_k\omega_k}{(kT)^N}\exp[-H_0/kT]\cdot\left\{1-\frac{\lambda V}{kT}+\frac{1}{2}\frac{\lambda^2V^2}{(kT)^2} +\frac{1}{2}\frac{\lambda^2}{(kT)^2}\left[\int\cdots\int(d\alpha)^N V^2 - \int\cdots\int(dJd\alpha)^N V^2\right]+\lambda^3\cdots\right\}, \quad (3.1)$$

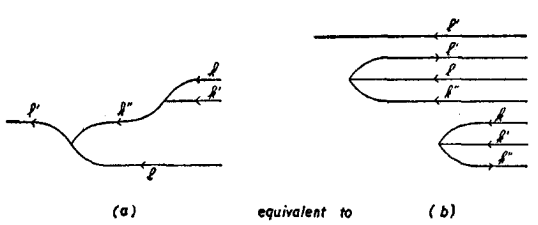


FIG. 3. Example of connected diagrams. Diagram corresponding to the product of matrix elements: $\langle n_{l'}=1|\delta L|n_l=1, n_{k''}=1\rangle \times\langle n_l=1, n_{k''}=1|\delta L|n_l=1, n_k=1, n_{k'}=1\rangle$.

where in V^2, we keep only the terms which are angle dependent. This function has a complicated N dependence. Indeed, each $\sum_{kk'k''}$ contains N^2 terms and each coefficient $V_{kk'k''}=0[1/(N)^{\frac{1}{2}}]$. Therefore, we see that

$$\rho^{\text{equil}}=0\{1+\lambda N^{\frac{1}{2}}+\lambda^2N^3+\lambda^2N+\cdots\}. \quad (3.2)$$

A series expansion in power of λ might therefore be questioned. However, we do not need ρ itself to obtain the macroscopic properties of the system, but we need only reduced distribution functions for a finite number of degrees of freedom. It can be readily verified that those reduced distribution functions are independent of N. As a result we can define extensive and intensive quantities. For instance, one can immediately verify that the mean energy of the system is proportional to N

$$\frac{\bar{E}}{N}=\frac{1}{N}\int\cdots\int(dJd\alpha)^N(H_0+\lambda V)\rho^{\text{equil}} =0(1+\lambda^2+\cdots), \quad (3.3)$$

and that moments of a finite order like $\langle u_{n_1}u_{n_2}u_{n_3}\rangle$, where u_n is the displacement of the nth atom from its equilibrium position, are independent of the size of the system

$$\langle u_{n_1}u_{n_2}u_{n_3}\rangle=[M^{(N)}]^{-\frac{3}{2}}\sum_{kk'k''}\{\exp[i(ka_{n_1}+k'a_{n_2}+k''a_{n_3})] \times\langle(J_kJ_{k'}J_{k''})^{\frac{1}{2}}\exp[i(\alpha_k+\alpha_{k'}+\alpha_{k''})\rangle +\text{similar terms}\}=0(\lambda), \quad (3.4)$$

where $M^{(N)}$ is the total mass of the crystal and a_n the lattice vector for the nth atom. We shall always choose our initial conditions in such a way that extensive and intensive properties may be defined. In other words we only consider such nonequilibrium situations for which $N\to\infty$, $\bar{E}/N=$ independent of N,

$$\langle u_{n_1}u_{n_2}u_{n_3}\rangle=\text{independent of } N, \text{ etc.} \quad (3.5)$$

We shall express this condition in terms of Fourier coefficients in Sec. 5.

4. FOURIER ANALYSIS OF THE DISTRIBUTION FUNCTION

We may Fourier analyze the distribution function $\rho(\{J_k\},\{\alpha_k\},t)$ with respect to the angle variables

$$\rho(\{J_k\},\{\alpha_k\},t)=(2\pi)^{-N/2}\sum_{\{n_k\}}\rho_{\{n_k\}}(\{J_k\},t) \times\exp[i\sum_k n_k(\alpha_k-\omega_kt)]. \quad (4.1)$$

The set of Fourier coefficients $\rho_{\{n_k\}}$ may be called the interaction representation of the distribution function. As in the case of gases, the Fourier coefficients have a simple physical meaning. For instance, $\rho_{\{0\}}$ is the distribution function of the action variables. The other coefficients describe angle correlations between the normal modes.

Homogeneous systems are characterized by the condition that moments like $\langle u_{n_1}u_{n_2}\cdots u_{n_r}\rangle$ are unaffected by a translation

$$\langle u_{n_1}u_{n_2}\cdots u_{n_r}\rangle=\langle u_{n_1+m}u_{n_2+m}\cdots u_{n_r+m}\rangle. \quad (4.2)$$

This implies[8] that the only nonvanishing Fourier coefficients are those for which $\sum_k kn_k=0$. The evolution equations for the Fourier coefficients are readily derived from the Liouville equation[7]:

$$\frac{\partial\rho_{\{n_k\}}}{\partial t}=\lambda\sum_{\{n_k'\}}\exp[i\sum_k(n_k-n_k')\omega_kt] \times\langle\{n_k\}|\delta L|\{n_k'\}\rangle\rho_{\{n_k'\}}(t). \quad (4.3)$$

If we take into account the condition (2.6) for the nonvanishing of the matrix elements of δL, we see that the set of equations (4.3) can be decomposed into completely independent sets of equations. All the Fourier coefficients in a given set are such that they correspond to states with the same value for $\sum_k kn_k$. Therefore all homogeneous Fourier components form an independent set; in other words, if the system is initially homogeneous, it will remain so during its entire evolution.

Because we only kept the cubic term in (2.1) we have a more restrictive condition on the Fourier coefficients which form an independent set. They must correspond to states which are connected by creations or destructions of correlations between three normal modes. For instance,

$$\rho_{\{0\}},\quad \rho_{\pm1_k\pm1_{k'}\pm1_{k''}}\quad\text{with}\quad \pm k\pm k'\pm k''=0,$$
$$\rho_{\pm1_k\pm1_{k'}\pm1_{k''}\pm1_l\pm1_{l'}\pm1_{l''}}\quad\text{with}\quad \pm k\pm k'\pm k''=0,$$

and

$$\pm l\pm l'\pm l''=0,\quad\text{etc.} \quad (4.4)$$

form such a set. Although $\rho_{\pm1_k\pm1_{k'}\pm1_{k''}\pm1_{k'''}}$ with $\pm k\pm k'\pm k''\pm k'''=0$ is also an homogeneous component, it is independent of the previous ones, but this is entirely caused by our neglect of the quartic term in (2.1). For the same reason, the equilibrium distribution (3.1) contain only the coefficients of the series (4.4). For consistency, we shall say that an homogeneous system is a system where the only nonvanishing coefficients are those of the series (4.4). Higher-order terms could be included in (2.1) but this would introduce no new features.

5. INITIAL CONDITIONS—HOMOGENEOUS SYSTEMS

On taking into account (3.1) and (4.1), we see that at equilibrium, the dependence on N and λ of the Fourier coefficients describing a correlation between $3n$ normal modes is

$$\rho_{3n}^{\text{equil}}=(\lambda^n/N^{n/2})\{1+\lambda^2N+\lambda^4N^2+\cdots\}, \quad (5.1)$$

where each term in the bracket is in fact a difference between terms of the given order in N. We have already

[8] R. Brout and I. Prigogine, Physica **22**, 621 (1956).

seen that this type of dependence permits the precise definition of extensive and intensive properties. We shall choose the following initial conditions:

$$\rho_{3n}(0)=0(\lambda^n/N^{n/2}). \tag{5.2}$$

$\rho_{3n}(0)$ can contain higher-order terms as in (5.1). As will be seen from the equations of evolution, the consequence of this is that $\rho_{3n}(t)$ at any time t has the same λ and N dependence as $\rho_{3n}{}^{\text{equil}}$.

With respect to λ, the choice (5.2) is somewhat more arbitrary than the choice for the N dependence. Its main advantage is its persistence during time. However, other situations might be considered. As (5.2) plays a role in determining the order of magnitude of various contributions, other initial conditions will in general lead to different results. The only exception is the equation for the distribution function of action variables for weakly coupled systems which is valid whatever the λ dependence of the Fourier coefficients. In the following we shall be interested in obtaining the evolution equations valid to a given finite order in λ. Due to its N dependence, it has no sense to cut a series like (5.1) at a given power in λ. However, this procedure is perfectly legitimate as soon as reduced distribution functions of a finite number of degrees of freedom are considered. We have given a few examples for the equilibrium functions, but this can be verified at any time with the choice (5.2). Therefore, we shall no longer write the N dependence of our equations. Moreover, we shall write our equations for the complete distribution function, but it must be kept in mind that those equations are valid only to derive the corresponding equations for the reduced distribution function for a finite number of degrees of freedom.[9]

6. DIAGRAMS

The set of equations (4.3) can be formally solved by an iteration procedure

$$\begin{aligned}\rho_{\{n_k\}}(t)=&\rho_{\{n_k\}}(0)+\lambda\sum_{\{n_k'\}}\int_0^t dt_1\exp[i\sum_k(n_k-n_k')\omega_k t_1]\\&\times\langle\{n_k\}|\delta L|\{n_k'\}\rangle\rho_{\{n_k'\}}(0)\\&+\lambda^2\sum_{\{n_k'\}}\sum_{\{n_k''\}}\int_0^t dt_1\int_0^{t_1}dt_2\\&\times\exp[i\sum_k(n_k-n_k')\omega_k t_1]\\&\times\exp[i\sum_k(n_k'-n_k'')\omega_k t_2]\\&\times\langle\{n_k\}|\delta L|\{n_k'\}\rangle\langle\{n_k'\}|\delta L|\{n_k''\}\rangle\\&\times\rho_{\{n_k''\}}(0)+\cdots.\end{aligned} \tag{6.1}$$

FIG. 4. Cycle:
$\langle 0|\delta L|\pm 1_k\pm 1_{k'}\pm 1_{k''}\rangle\langle\pm 1_k\pm 1_{k'}\pm 1_{k''}|\delta L|0\rangle$.

Each term of this equation involves a product of n matrix elements of δL. They can all be represented graphically by a combination of n of the basic matrix elements given in Fig. 1. The n vertices will always be ordered in the chronological order, time running from right to left.

Some interesting combinations of the basic diagrams of Fig. 1 might appear. The simplest case corresponds to two successive inverse transitions and is represented by a cycle (Fig. 4).[10] The other basic combinations of the elementary diagrams of Fig. 1 are:

(a) *diagonal fragment:* Any diagram or part of a diagram such that the initial state is identical with the final state but is different from any of the intermediate states. The cycle is the simplest diagonal fragment. Other examples are given in Fig. 5.

(b) (*nondiagonal*) *destruction fragment:* Any diagram or part of a diagram made of nondiagonal destruction (Fig. 1, *abcd*) transitions (at least one) and possibly diagonal fragments such that no intermediate state is identical with the final state, the latter containing no other lines than the lines which are propagated freely from one end of the diagram to the other. The simplest examples are the diagrams *abcd*, Fig. 1. Other examples are given in Fig. 6.

(c) (*nondiagonal*) *creation fragment:* This is defined in the same way as the destruction fragment, replacing destruction transitions by creation transitions (Fig. 1, $a'b'c'd'$) and final state by initial state. Examples are given in Fig. 7.

(d) *nondiagonal irreducible destruction-creation fragment:* Any diagram made of destruction and creation transitions (at least one of each) and possibly diagonal fragments which cannot be decomposed into a destruction fragment followed by a creation one; in other

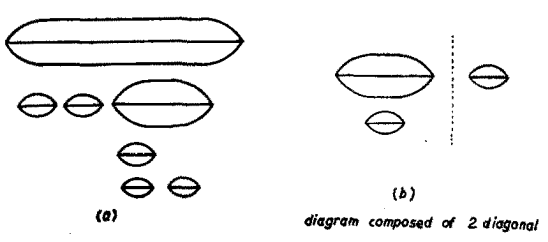


FIG. 5. Examples of diagonal diagrams. [The two lower cycles of diagram (a) must also have a central line.]

[9] The situation is the same as in the case of gases. In the equation for reduced distribution functions corresponding to a finite number of particles, the N factors disappear, while they persist if one consider equations for the whole set of variables. This comes from the simple fact that τ/N will be the order of magnitude of the time between two collisions in the master equation if τ is the time between two collisions of a given molecule.

[10] We shall often drop the arrows and wave vectors of the diagrams; in that case a single diagram will represent a summation over all the diagrams which differ by the wave vectors only; the only exception will be for the lines which remain at the left of the diagram. The wave vectors of those lines are those described by the set $\{n_k\}$ of the Fourier component.

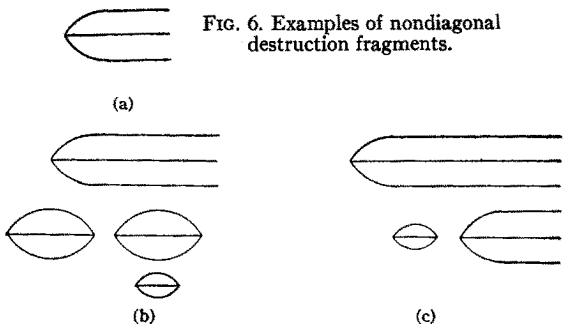


FIG. 6. Examples of nondiagonal destruction fragments.

words, in each of the states, there are at least three lines which come or go to a vertex. Examples are given in Fig. 8.

The most general diagrams are either diagonal or nondiagonal. Diagonal diagrams are made from a succession of diagonal fragments. Nondiagonal diagrams contain at least a destruction or a creation fragment or both. Taking into account the fact that a diagonal fragment added to the right (left) or a destruction (creation) fragment gives rise to a destruction (creation) fragment and that a creation fragment followed by a destruction one gives rise to an irreducible destruction-creation fragment, we are left with four types of nondiagonal diagrams:

(a) nondiagonal destruction diagram: destruction fragment possibly *followed* by a succession of diagonal fragments;

(b) nondiagonal creation diagram: a creation fragment possibly *preceded* by a succession of diagonal fragments;

(c) nondiagonal reducible creation and destruction diagram: destruction fragment *followed* by a creation fragment with or without a succession of diagonal fragments between the the two nondiagonal ones;

(d) nondiagonal irreducible creation-destruction diagram: described in the foregoing.

We obtain in this way a very simple topological classification of the general diagrams. Examples are given in Fig. 9. As we shall see later, the asymptotic contribution of a given diagram is closely related to its topological structure.

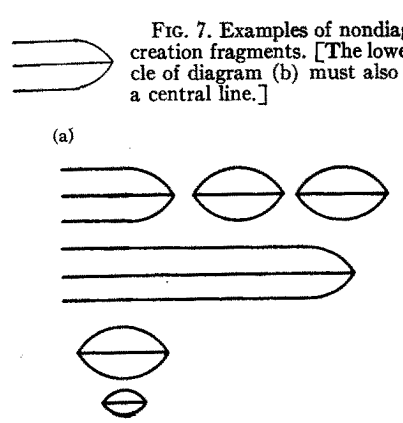


FIG. 7. Examples of nondiagonal creation fragments. [The lower cycle of diagram (b) must also have a central line.]

7. TIME DEPENDENCE OF DIAGRAMS

To each product of n matrix elements [see (6.1)] is associated an integral of the form

$$\int_0^t dt_1 \cdots \int_0^{t_{n-1}} dt_n \exp(-i\alpha_0 t_1) \times\{\prod_{i=1}^{n-1} \exp[i\alpha_i(t_i-t_{i+1})]\} \exp(i\alpha_n t_n). \quad (7.1)$$

From (6.1) we see that

$$\alpha_i=\sum_k{}' n_k{}^{(i)}\omega_k, \quad (7.2)$$

where $|\{n_k{}^{(i)}\}\rangle$ represents the state during the time interval (t_i-t_{i+1}). The $'$ for the $\sum_k$ means that the frequencies corresponding to "free" lines, i.e., to lines which are propagated freely from the initial time to the final time, are not taken into account. Indeed, those lines play no role in the integral; if we consider the integral as written in (6.1), we see that for free lines the difference $(n-n')$ vanishes at any time t_i. In other words, $\alpha_0(\alpha_n)$ is the sum of the frequencies of the lines which have been created (destroyed) somewhere in the diagram, whereas each of the $\alpha_i(1\leqslant i\leqslant n-1)$ is the sum of the frequencies of all the lines except the freely propagated ones which are present during the time interval t_i-t_{i+1}. Examples are given in Fig. 10. For all diagrams, except the *irreducible* destruction-creation diagrams, at least one of the α's is equal to zero; $(m+1)$ of the α's are equal to zero if m is the number of diagonal fragments which can be isolated in the diagram. The expression (7.1) is the time integral which appears in (6.1), and therefore determines the evolution of $\rho_{\{n_k\}}(t)$. However, we are interested in the distribution function rather than in the Fourier components. From (4.1), we thus see that (7.1) has then to be multiplied by a time-dependent function $\exp[-i\sum_k n_k\omega_k t]$ where $|\{n_k\}\rangle$ is the final state. Therefore, rather than (7.1) we have to

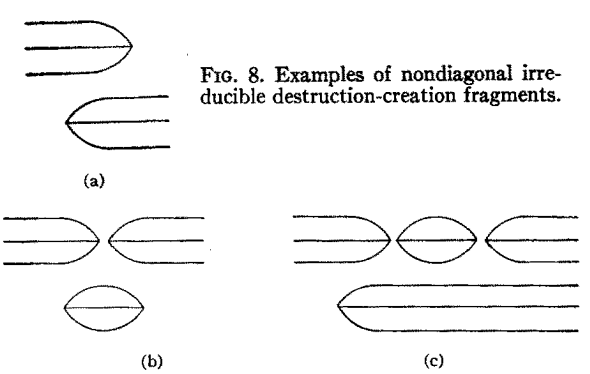


FIG. 8. Examples of nondiagonal irreducible destruction-creation fragments.

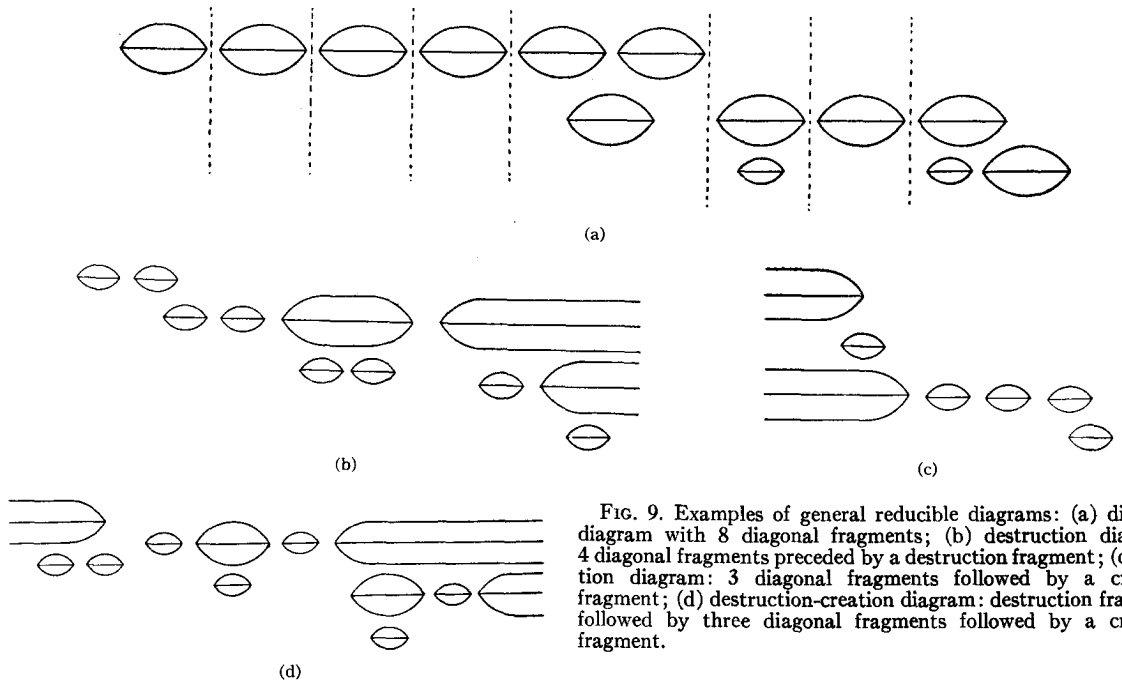


FIG. 9. Examples of general reducible diagrams: (a) diagonal diagram with 8 diagonal fragments; (b) destruction diagram: 4 diagonal fragments preceded by a destruction fragment; (c) creation diagram: 3 diagonal fragments followed by a creation fragment; (d) destruction-creation diagram: destruction fragment followed by three diagonal fragments followed by a creation fragment.

consider

$$I(t)=\exp(i\gamma t)\int_0^t dt_1\cdots\int_0^{t_{n-1}} dt_n \exp[i\alpha_0(t-t_1)] \times\prod_{i=1}^{n-1}\exp[i\alpha_i(t_i-t_{i+1})]\exp(i\alpha_n t_n), \quad (7.3)$$

where γ is the sum of the frequencies of the lines which are propagated freely through the entire diagram. We shall assume that t is sufficiently large to enable us to make an asymptotic evaluation of (7.3). In order to see what this means, let us consider a simple example: the asymptotic contribution of the cycle. We have[7]

$$\lambda^2\sum_{kk'k''}|V_{kk'-k''}|^2\left[\frac{\partial}{\partial J_k}+\frac{\partial}{\partial J_{k'}}-\frac{\partial}{\partial J_{k''}}\right]\left(\frac{J_kJ_{k'}J_{k''}}{\omega_k\omega_{k'}\omega_{k''}}\right) \times\left[\frac{\partial}{\partial J_k}+\frac{\partial}{\partial J_{k'}}-\frac{\partial}{\partial J_{k''}}\right]\int_0^t dt_1\int_0^{t_1} dt_2 \times\exp[i(\omega_k+\omega_{k'}-\omega_{k''})(t_1-t_2)]. \quad (7.4)$$

If we perform the time integrations using as variables

$$\tau=\tau_1-\tau_2 \quad \text{and} \quad T=(\tau_1+\tau_2)/2,$$

we obtain

$$\int d\alpha f(\alpha)\int_0^t d\tau e^{i\alpha\tau}\int_{\tau/2}^{t-\tau/2} dT = \int d\alpha f(\alpha)\int_0^t d\tau e^{i\alpha\tau}(t-\tau) \quad (7.5a)$$

$$= \int d\alpha f(\alpha)\left\{t\frac{e^{i\alpha t}-1}{i\alpha}+i\frac{d}{d\alpha}\frac{e^{i\alpha t}-1}{i\alpha}\right\}, \quad (7.5b)$$

where $\alpha=\omega_k+\omega_{k'}-\omega_{k''}$, and where we have taken into account the fact that for $N\to\infty$, the frequency spectrum becomes continuous and the $\sum_k$ can be transformed into an integral. It is well known that for long times, i.e., for times

$$t\gg(\delta\alpha)^{-1}, \quad (7.6)$$

where $(\delta\alpha)$ is the interval of α over which the function $f(\alpha)$ can be considered as slowly varying, the first term in the rhs of (7.5b) takes the asymptotic form

$$it\int d\alpha f(\alpha)\zeta(\alpha)=it\int d\alpha f(\alpha)\left\{\left(\frac{1}{\alpha}\right)_p-\pi i\delta(\alpha)\right\}. \quad (7.7)$$

The times τ in (7.5a) which give this contribution are of the order of $(\delta\alpha)^{-1}$. If we take this into account as well as (7.6), we may neglect τ in the expression $(t-\tau)$ in the rhs of (7.5a), i.e., we may neglect the second term in the rhs of (7.5b) and the asymptotic value of the inte-

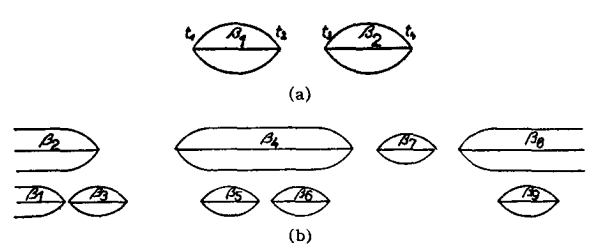


FIG. 10. Examples illustrating (7.1). The β's denote the sum of the frequencies of the corresponding three lines. The α_i's in (7.1) are given by (a) $\alpha_0=0$ $\alpha_1=\beta_1$ $\alpha_2=0$ $\alpha_3=\beta_2$ $\alpha_r=0$

(b) $\alpha_0=\beta_1+\beta_2$ $\alpha_1=\beta_2$ $\alpha_2=\beta_2+\beta_3$ $\alpha_3=\beta_3$ $\alpha_4=0$ $\alpha_5=\beta_4$ $\alpha_6=\beta_4+\beta_5$ $\alpha_7=\beta_4$ $\alpha_8=\beta_4+\beta_6$ $\alpha_9=\beta_4$ $\alpha_{10}=0$ $\alpha_{11}=\beta_7$ $\alpha_{12}=0$ $\alpha_{13}=\beta_8$ $\alpha_{14}=\beta_8+\beta_9$ $\alpha_r=\beta_8$.

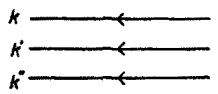


FIG. 11. Example of free propagation diagram.

gral is given by (7.7). As τ is the time interval during which the scattering process occurs, we see that scattering processes which contribute asymptotically are instantaneous events on a macroscopic scale. The condition (7.6) has a simple physical meaning. Indeed, the function $f(\alpha)$ is essentially of the form

$$\sum_{kk'k''} |V_{kk'k''}|^2 \frac{1}{\omega_k\omega_{k'}\omega_{k''}}$$

transformed into an integral. It can be shown (see for instance Peierls[11]) that the coefficient

$$|V_{kk'k''}|^2/\omega_k{}^2\omega_{k'}{}^2\omega_{k''}{}^2$$

is nearly constant. Therefore, the variation of the function $f(\alpha)$ will be mainly determined by the dispersion law which must be used to transform the sum over $kk'k''$ into an integral over α. This function will be slowly varying provided we consider frequency intervals smaller than the Debye frequency ω_D. Therefore, we have

$$t \gg (\delta\alpha)^{-1} \approx 1/\omega_D. \tag{7.8}$$

In other words our asymptotic evaluation will be correct provided we are interested in times much larger than characteristic molecular times of the order of $(\omega_D)^{-1}$. Let us now consider another way to evaluate (7.4) asymptotically which will be more practicable for general integrals of the form (7.3). The integral over time (7.5b) is a regular function of α. If we assume $f(\alpha)$ to be regular on the real axis, we can consider complex variables $(\alpha+i\epsilon)$ and write (7.5b) as

$$\lim_{\epsilon\to 0} \int d\alpha f(\alpha) \left\{ -\frac{t}{i(\alpha+i\epsilon)} + \frac{e^{i(\alpha+i\epsilon)t}-1}{i^2(\alpha+i\epsilon)^2} \right\}. \tag{7.9}$$

If $\epsilon t \gg 1$, we can drop the oscillating term in (7.9).

If we assume that the first complex pole of the function $f(\alpha)$ is at a finite distance[12] from the real axis, we can choose a finite ϵ (of the order of ω_D for instance) and for large t, the condition $\epsilon t \gg 1$ is then satisfied. Then (7.9) becomes

$$\lim_{\epsilon\to 0} \int d\alpha f(\alpha) \left\{ -\frac{t}{i(\alpha+i\epsilon)} + \frac{1}{(\alpha+i\epsilon)^2} \right\}, \tag{7.10}$$

where the integration has to be done along the real axis. These two terms correspond to the two contributions in (7.5a) or (7.5b). Here too, the second term is of the order of the collision time $\omega_D{}^{-1}$ and may be neglected with respect to the first one which is of order t. We obtain therefore the following asymptotic contribution:

$$\lim_{\epsilon\to 0} \int d\alpha f(\alpha) \left\{ -\frac{t}{i(\alpha+i\epsilon)} \right\}. \tag{7.11}$$

This expression is equivalent to (7.7) (see for instance, Heitler[13]). This method of asymptotic evaluation will always be used in the following. In other words, for each sum of frequencies which appear at a vertex, i.e., to which corresponds a Fourier coefficient of the potential $V_{kkk''}$, we shall consider the complex variables $(\alpha\pm i\epsilon)$, the sign being such that we can drop asymptotically the exponentials. Therefore, the only oscillating factor which we shall keep in (7.3) is the factor $e^{i\gamma t}$ which corresponds to the oscillations related to free propagation. For instance, for the diagram of Fig. 11, we have

$$\int d\alpha g(\alpha) e^{i\alpha t}$$

with

$$\alpha = \omega_k + \omega_{k'} + \omega_{k''} \quad \text{and} \quad g(\alpha) = \rho_{1k1k'1k''}(0). \tag{7.12}$$

With this method, it is possible to evaluate asymptotically the integral (7.3). This will be done in Appendix I. Let us first consider the reducible diagrams: the most general diagrams of this class are made of a destruction fragment, followed by a diagonal region containing m diagonal fragments, and then by a creation fragment. Diagonal diagrams, destruction, or creation diagrams can be considered as particular cases of reducible diagrams. For such diagrams we have the following theorem:

Theorem I: Any reducible diagram has an asymptotic contribution proportional to t^m, where m is the number of diagonal fragments in the diagonal region of the diagram.

As examples of application of this theorem, let us consider the diagrams of Fig. 9. For the diagram (a) $m=8$; (b) $m=4$; (c) $m=3$; (d) $m=3$.

For the irreducible destruction-creation diagrams we have theorem II.

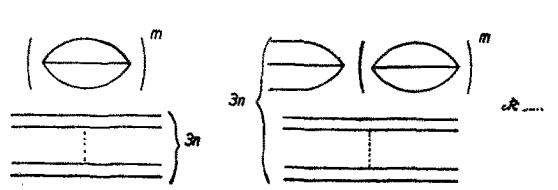


FIG. 12. Lowest-order contributions to $\rho_{3n}'(t)$ [( )^m means a succession of m cycles].

[11] R. Peierls, *Quantum Theory of Solids* (Clarendon Press, Oxford, England, 1955).

[12] For gases, this condition can be studied more precisely. Let us for instance consider a screened Coulomb potential. Then the Fourier coefficients are $V_k = 1/(k^2+K^2)$ and the pole is at $k=iK$. The condition $\epsilon t \gg 1$ can then be interpreted as the fact that the time t must be sufficiently large to allow the particles to travel a distance much larger than the range of the forces.

[13] W. Heitler, *The Quantum Theory of Radiation* (Clarendon Press, Oxford, England, 1954).

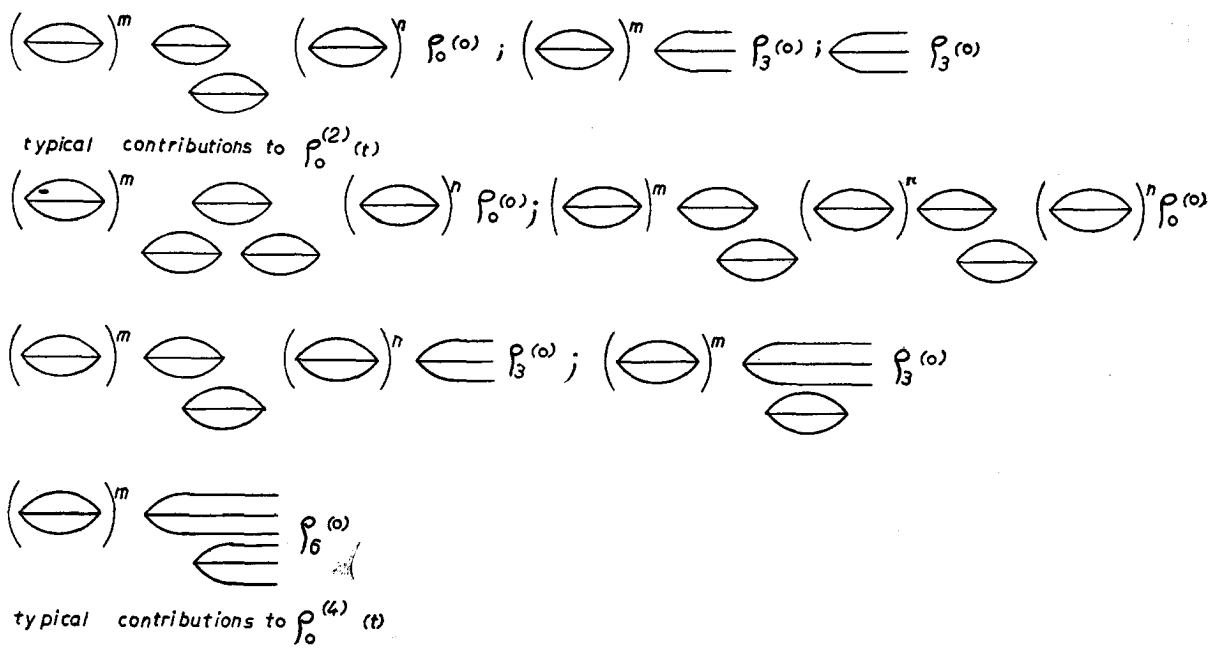


FIG. 13. Contributions to $\rho_0^{(4)}(t)$ obtained from contributions to $\rho_0^{(2)}(t)$.

Theorem II: Any irreducible destruction-creation diagram has a vanishing asymptotic contribution.

8. ASYMPTOTIC CONTRIBUTIONS TO THE FOURIER COMPONENTS

We are now interested in the asymptotic behavior of the Fourier coefficients for homogeneous systems. More precisely, if we call $\rho_{3n}(t)$ the Fourier coefficient describing n correlations between three normal modes, we are interested in the asymptotic behavior of the function

$$\rho_{3n}(t)\exp(-i\omega_{3n}t), \tag{8.1}$$

where ω_{3n} is the sum of the frequencies of the $3n$ normal modes considered. We want to keep in (8.1) all terms up to order λ^{2r}. We have to take into account both the asymptotic contributions of the products of matrix elements and the order of magnitude of the initial Fourier components. From the theorems of Sec. 7, we see that the only diagrams which contribute asymptotically are the reducible diagrams. We can separate those diagrams into two groups: (a) diagonal diagrams and destruction diagrams; (b) creation diagrams and reducible creation-destruction diagrams. Let us also split (8.1) into 2 parts:

$$\rho_{3n}(t)\exp(-i\omega_{3n}t)=\rho_{3n}'(t)\exp(-i\omega_{3n}t)+\rho_{3n}''(t), \tag{8.2}$$

where $\rho_{3n}'(t)\exp(-i\omega_{3n}t)$ contains the asymptotic contributions of all diagrams of group (a) and $\rho_{3n}''(t)$ those of group (b). As all diagrams of group (a), except those which are made by a single nondiagonal destruction fragment, end with a diagonal fragment, $\rho_{3n}'(t)$ might be expected to obey a simple diagonal equation. On the other hand, all contributions to $\rho_{3n}''(t)$ [i.e., of group (b)] end with a nondiagonal fragment, and we may certainly not expect a diagonal equation for those functions.

We may write

$$\rho_{3n}'(t)=\lambda^n\{\rho_{3n}'^{(0)}(t)+\lambda^2\rho_{3n}'^{(2)}(t)+\lambda^4\rho_{3n}'^{(4)}(t)+\cdots+\lambda^{2r}\rho_{3n}'^{(2r)}(t)+\cdots\}, \tag{8.3}$$

where the functions $\rho_{3n}'^{(2i)}(t)$ are defined as the asymptotic contributions of all diagrams of type a, including the order of magnitude of the initial Fourier component in the diagram, whose asymptotic contribution is of the form $\lambda^{2i}(\lambda^2 t)^m$; $m\geq 0$. (8.3) is not a true expansion in power series of λ. Indeed, the initial Fourier components may contain higher-order terms with respect to λ than those given explicitly in (5.2), and the product $(\lambda^2 t)$ itself is λ dependent. The λ dependence in (8.3) may be easily verified. The factor λ^n in front of the rhs of (8.3) is caused by the initial conditions. Indeed, the lowest-order contribution to $\rho_{3n}'(t)$ is that of the diagram of Fig. 12. The fact that the expansion contains only even powers of λ can be seen in the following way. If we consider a diagram contributing to $\rho_{3n}'^{(2i)}(t)$ and want to obtain the next diagrams, we can either replace a diagonal fragment by a higher-order one, or add a cycle in the destruction fragment or finally add a destruction transition in the fragment, but then the initial Fourier component has one more correlation. These three possibilities are illustrated in Fig. 13. Each of them leads to one more uncompensated λ^2 factor.

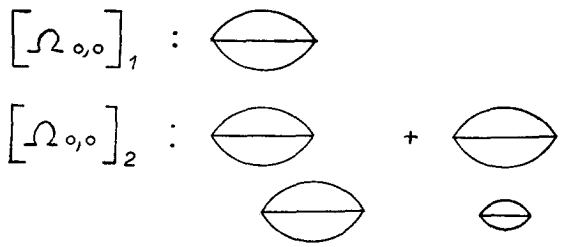


FIG. 14. Simplest diagonal operators.

In the same way, we may write

$$\rho_{3n}''(t)=\lambda^n\{\rho_{3n}''^{(0)}(t)+\lambda^2\rho_{3n}''^{(2)}(t)+\cdots+\lambda^{2r}\rho_{3n}''^{(2r)}(t)+\cdots\},$$
$$\rho_0''(t)=0. \tag{8.4}$$

The last equation comes from the fact that no diagrams involving creations contribute to $\rho_{\{0\}}(t)$.

9. EVOLUTION EQUATIONS FOR $\rho_{3n}'(t)$

We shall now show that the functions $\rho_{3n}'^{(2r)}(t)$ obey the following differential equation:

$$\frac{\partial\rho_{3n}'^{(2r)}}{\partial t}=\lambda^2\sum_{i=1}^{r}[\Omega_{3n,3n}]_{r+1-i}\,\rho_{3n}'^{(2i)}(t). \tag{9.1}$$

The operator in the rhs is the operator corresponding to all diagonal fragments containing $(r+1-i)$ cycles. Examples are given in Fig. 14. The important point in (9.1) is that this equation is diagonal in the index $3n$. It is sufficient to establish this equation for $\rho_0'^{(2r)}$. Indeed, the diagrams which contribute to $\rho_{3n}'^{(2r)}$ differ from those which contribute to $\rho_0'^{(2r)}$ by a set of $3n$ lines which are freely propagated through the diagram. As these lines play no role in the asymptotic evaluation, we can as well add them at the last step. The diagrams contained in $\rho_0'^{(2r)}$ are all diagonal diagrams and all destruction diagrams with asymptotic contributions of the form $\lambda^{2r}(\lambda^2t)^n$. Let us first consider the case $r=0$. Then we have only diagonal diagrams because all destruction diagrams contain at least one uncompensated λ^2 factor. We have to keep all diagonal diagrams whose asymptotic contributions are of the form $(\lambda^2t)^n$. As we get a factor t for each diagonal fragment, we see that the only diagonal fragments we have to consider must be proportional to λ^2, i.e., must be cycles (Fig. 15).

This can be written

$$\rho_0'^{(0)}(t)=\sum_{n=0}^{\infty}\{n \text{ diagonal fragments made by } n \text{ cycles}\}\rho_0(0). \tag{9.2}$$

The asymptotic contribution of the nth term is of the form (see Appendix I)

$$\lambda^{2n}\int_0^t dt_1\cdots\int_0^{t_{2n-1}}dt_{2n}\prod_{j=1}^{n-1}\{\sum_{\alpha_j}\exp[i\alpha_j(t_j-t_{j+1})]\times\langle 0|\delta L_j^2|0\rangle\}$$
$$=\frac{(\lambda^2t)^n}{n!}\{\sum_{\alpha_j}i\zeta(\alpha_j)\langle 0|\delta L_j^2|0\rangle\}^n$$
$$=\lambda^{2n}[\Omega_{0,0}]_1^n\int_0^t dt_1\int_0^{t_1}dt_2\cdots\int_0^{t_{n-1}}dt_n, \tag{9.3}$$

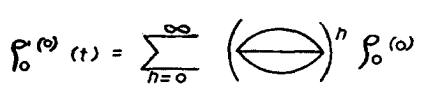


FIG. 15.

where δL_j means that we consider the cycle corresponding to the normal modes involved in α_j, and the sum over α_j means that for each cycle we sum over all possible 3 normal mode processes. The operator $[\Omega_{0,0}]_1$ in the rhs is time independent. We may thus write

$$\rho_0'^{(0)}(t)=\rho_0(0)+\lambda^2[\Omega_{0,0}]_1\int_0^t dt_1\{\sum_{n=0}^{\infty}n \text{ diagonal fragments made by } n \text{ cycles}\}\rho_0(0), \tag{9.4}$$

where in the rhs the contributions in { } are of the same form as in (9.3), but are taken up to time t_1 instead of t. If we take into account (9.2), we obtain an integral equation for $\rho_0'^{(0)}(t)$;

$$\rho_0'^{(0)}(t)=\rho_0(0)+\lambda^2\int_0^t dt_1[\Omega_{0,0}]_1\rho_0'^{(0)}(t_1), \tag{9.5}$$

which leads to the differential equation

$$\frac{\partial\rho_0'^{(0)}}{\partial t}=\lambda^2[\Omega_{0,0}]_1\rho_0'^{(0)}(t). \tag{9.6}$$

We shall now follow exactly the same procedure to obtain (9.1) for $r>0$. In this case, we have (see for instance Fig. 13 for $r=1$ and $r=2$)

$$\rho_0'^{(2r)}(t)=\sum_{m=1}^{\infty}\lambda^{2m}\{m \text{ diagonal fragments formed by } (m+r) \text{ cycles}\}\rho_0(0)$$
$$+\sum_{l=1}^{r}\lambda^{-l}\{\text{destruction fragment formed by } l \text{ destruction transitions and } (r-l) \text{ cycles}\}\rho_{3l}(0)$$
$$+\sum_{k=1}^{\infty}\sum_{n=k}^{r+k-1}\sum_{l=1}^{r-n+k}\lambda^{2k-l}\{k \text{ diagonal fragments formed by } n \text{ cycles}\}\times\{\text{destruction fragment formed by } l \text{ destruction transitions and } (r-n-l+k) \text{ cycles}\}\rho_{3l}(0). \tag{9.7}$$

The asymptotic contribution of all those terms is immediately seen to be proportional to some power of $\lambda^2 t$ if one takes into account theorem I.

To establish an equation for $\rho_0'^{(2r)}(t)$, let us first notice that for times for which the asymptotic time integration holds, the second sum in (9.7) gives simply a constant, independent of time. Therefore it will cancel on being differentiated with respect to time. Let us consider the first term. Let us denote by $[\Omega_{0,0}]_\alpha$ the contribution of a diagonal fragment formed by α cycles. We may transform this term in the same way as we transformed the rhs of (9.2):

$$\sum_{m=1}^{\infty}\lambda^{2m}\{m \text{ diagonal fragments made of } (r+m) \text{ cycles}\}\rho_0(0)$$

$$=\lambda^2[\Omega_{0,0}]_{r+1}\int_0^t dt_1\{\rho_0(0)+\sum_{m=2}^{\infty}[(m-1) \text{ diagonal fragments made of } (m-1) \text{ cycles}]\rho_0(0)\}$$

$$+\lambda^2\sum_{\alpha=1}^{r}[\Omega_{0,0}]_\alpha\int_0^t dt_1\sum_{m=2}^{\infty}\{(m-1) \text{ diagonal fragments made of } (r+m-\alpha) \text{ cycles}\}\rho_0(0), \quad (9.8)$$

where all contributions at the right of the operator $[\Omega_{0,0}]_\alpha$ have to been taken up to time t_1. Taking into account (9.2), (9.8) can be written as

$$\lambda^2[\Omega_{0,0}]_{r+1}\int_0^t dt_1\rho_0^{(0)}(t_1)+\lambda^2\sum_{\alpha=1}^{r}[\Omega_{0,0}]_\alpha\int_0^t dt_1$$

$$\times\sum_{m=1}^{\infty}\{m \text{ diagonal fragments made of } (r+m-\alpha-1) \text{ cycles}\}\rho_0(0). \quad (9.9)$$

We proceed in the same way with the third contribution to (9.7). We again write explicitly the last diagonal fragment

$$\sum_{k=1}^{\infty}\sum_{n=k}^{r+k-1}\sum_{l=1}^{r-n+k}\lambda^{2k-l}\{k \text{ diagonal fragments made of } n \text{ cycles}\}\times\{\text{destruction fragment made of } l \text{ destruction transitions and } (r-n-l+k) \text{ cycles}\}\rho_{3l}(0)$$

$$=\lambda^2\sum_{\alpha=1}^{r}[\Omega_{0,0}]_\alpha\int_0^t dt_1\sum_{l=1}^{r-\alpha+1}\lambda^{-l}\{\text{destruction fragment made of } l \text{ destruction transitions and } (r-\alpha-l+1) \text{ cycles}\}\rho_{3l}(0)$$

$$+\lambda^2\sum_{k=2}^{\infty}\sum_{n=k}^{r+k-1}\sum_{l=1}^{r-n+k}\sum_{\alpha=1}^{n-k+1}[\Omega_{0,0}]_\alpha\lambda^{2k-2-l}$$

$$\times\int_0^t dt_1\{(k-1) \text{ diagonal fragments made of } (n-\alpha) \text{ cycles}\}\times\{\text{destruction fragment made of } l \text{ destruction transitions, and } (r-n-l+k) \text{ cycles}\}\rho_{3l}(0), \quad (9.10)$$

where we have considered separately the term $k=1$. If we replace k by $k+1$ in the second term of (9.10) and change the order of the summations, we obtain

$$\lambda^2\sum_{\alpha=1}^{r}[\Omega_{0,0}]_\alpha\int_0^t dt_1\{\sum_{l=1}^{r-\alpha+1}\lambda^{-l}[\text{destruction fragment made of } l \text{ destruction transitions and } (r-\alpha-l+1) \text{ cycles}]\rho_{3l}(0)$$

$$+\sum_{k=1}^{\infty}\sum_{n=k+\alpha}^{r+k}\sum_{l=1}^{r-n+k+1}\lambda^{-l+2k}[k \text{ diagonal fragments made of } (n-\alpha) \text{ cycles}]\times[\text{destruction fragment made of } l \text{ destruction transitions and } (r-n-l+k+1) \text{ cycles}]\rho_{3l}(0)\}. \quad (9.11)$$

If we replace in (9.7) the 1st and 3rd terms by (9.9) and (9.11) and take into account (9.7) for r replaced by $r-\alpha+1$, we obtain

$$\rho_0'^{(2r)}(t)=\sum_{l=1}^{r}\lambda^{-l}\{\text{destruction fragment made of } l \text{ destruction transitions and } (r-l) \text{ cycles}\}\rho_{3l}(0)$$

$$+\lambda^2\sum_{\alpha=1}^{r+1}[\Omega_{0,0}]_\alpha\int_0^t dt_1\rho_0'^{(2(r-\alpha+1))}(t_1). \quad (9.12)$$

The asymptotic contribution of the 1st term in the rhs being a constant with respect to t (see theorem I), we obtain indeed (9.1) by differentiation with respect to t of (9.12). If we take into account (8.3) we may transform (9.1) into a differential equation for $\rho_{3n}'(t)$ valid

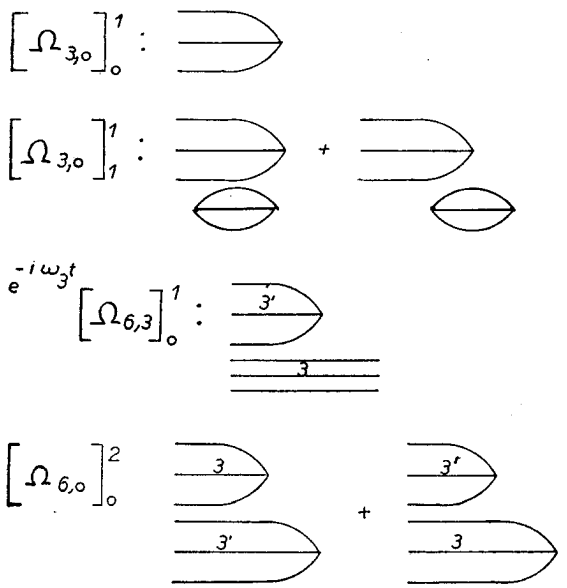


FIG. 16. Simplest nondiagonal operators.

to order λ^{n+2s}. We have

$$\frac{\partial \rho_{3n}'(t)}{\partial t}=\lambda^n \sum_{r=0}^{s} \lambda^{2r} \frac{\partial \rho_{3n}'^{(2r)}}{\partial t}$$

$$=\lambda^{n+2} \sum_{r=0}^{s} \sum_{i=0}^{r} \lambda^{2r}[\Omega_{3n,3n}]_{r+1-i}\, \rho_{3n}'^{(2i)}(t)$$

$$=\lambda^{n+2} \sum_{\alpha=0}^{s} \lambda^{2\alpha}[\Omega_{3n,3n}]_{\alpha+1} \times \sum_{i=0}^{s-\alpha}\{\lambda^{2i} \rho_{3n}'^{(2i)}(t)\}. \tag{9.13}$$

The summation over i can be extended up to $i=s$; this only adds terms of higher order than λ^{n+2s} and the equation is still correct at that order. We then obtain for $\rho_{3n}'(t)$ the very simple "diagonal" differential equation for the time rate of change of $\rho_{3n}'(t)$

$$\frac{1}{\lambda^2}\frac{\partial \rho_{3n}'}{\partial t}=\{\sum_{\alpha=0}^{s} \lambda^{2\alpha}[\Omega_{3n,3n}]_{\alpha+1}\} \rho_{3n}'(t). \tag{9.14}$$

The properties of these equations will be studied in a subsequent paragraph.

10. EVOLUTION OF THE FUNCTIONS $\rho_{3n}''(t)$

We have seen that for $\rho_{3n}'(t)$, the last fragment at the left plays an essential role in determining the evolution equation. This will again be true for the functions $\rho_{3n}''(t)$. As this last fragment is always a creation fragment, we may no longer expect a diagonal equation. Indeed, we shall show that $\rho_{3n}''^{(2r)}(t)$ obeys the following equation:

$$\rho_{3n}''^{(2r)}(t)=\sum_{m=1}^{n} \exp[-i\omega_{3(n-m)}t] \times \sum_{i=0}^{r}[\Omega_{3n,3(n-m)}]_i^m \rho_{3(n-m)}'^{2(r-i)}(t). \tag{10.1}$$

This equation relates $\rho_{3n}''^{(2r)}(t)$ to all the functions $\rho_{3(n-m)}'^{2(r-i)}(t)$. There is a basic difference in the behavior of ρ_{3n}' and of ρ_{3n}''. While the functions ρ_{3n}' satisfy a separate set of equations independent of the ρ_{3n}'', the functions ρ_{3n}'' are completely determined by the values of the ρ_{3n}''s. Moreover, only the functions ρ_{3k}' with $k<n$ and corresponding therefore to a smaller number of correlations enter into the equation for ρ_{3n}''. The operator $[\Omega_{3n,3(n-m)}]_i^m$ corresponds to all the nondiagonal creation fragments made of i cycles and the m nondiagonal creation transitions which are necessary to go from the state $3(n-m)$ to the state $3n$. Examples are given in Fig. 16. The operator Ω is again time independent. The oscillating factor in front corresponds to the free propagation of the $(n-m)$ correlations which were already present initially. It corresponds to the oscillating factor $e^{i\gamma t}$ of (7.3) which is not used in the asymptotic integrations. As in Sec. 9, we again consider first the case $r=0$. We then have only creation diagrams because destructions always introduce at least one uncompensated λ^2 factor. We have

$$\rho_{3l}''^{(0)}(t)=\lambda^{-l} \sum_{m=1}^{l} \sum_{k=0}^{\infty} \lambda^{m+2k}\{\text{creation fragment made of } m \text{ creation transitions}\} \times\{k \text{ cycles}\}\rho_{3(l-m)}(0) \times \exp[-i\omega_{3(l-m)}t]. \tag{10.2}$$

Examples are given in Fig. 17.

Using the notations explained in connection with

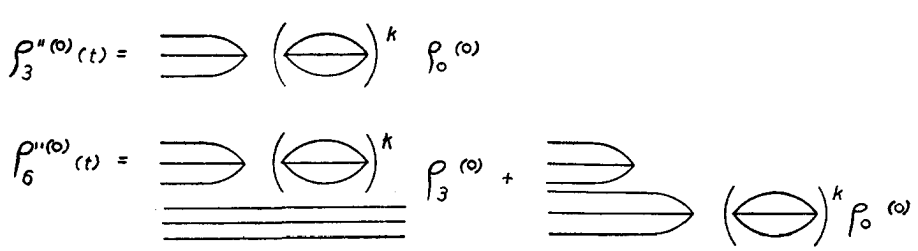


FIG. 17.

(10.1), we can rewrite (10.2) as

$$\rho_{3l}{}''^{(0)}(t)=\lambda^{-l}\sum_{m=1}^{l}\lambda^{m}\exp[-i\omega_{3(l-m)}t][\Omega_{3l,3(l-m)}]_0{}^{m}\times\{\sum_{k=0}^{\infty}\lambda^{2k}[k\text{ cycles}]\rho_{3(l-m)}(0)\}. \quad (10.3)$$

The contributions inside { } have now to be taken up to time t (see Appendix I). If we use (9.2), we obtain

$$\rho_{3l}{}''^{(0)}(t)=\sum_{m=1}^{l}\exp[-i\omega_{3(l-m)}t][\Omega_{3l,3(l-m)}]_0{}^{m}\times\rho_{3(l-m)}{}'^{(0)}(t). \quad (10.4)$$

Let us now consider the case $r>0$ and again follow the same procedure. We have

$$\begin{aligned}\rho_{3l}{}''^{(2r)}(t)=&\sum_{m=1}^{l}\lambda^{-l+m}\exp[-i\omega_{3(l-m)}t]\{\text{creation fragment with } m \text{ creation transitions and } r \text{ cycles}\}\rho_{3(l-m)}(0)\\
&+\sum_{m=1}^{l}\sum_{n=0}^{r}\sum_{s=1}^{\infty}\lambda^{-l+m+2s}\exp[-i\omega_{3(l-m)}t]\times\{\text{creation fragment with } m \text{ creation transitions and } n\\
&\text{cycles}\}\times\{s \text{ diagonal fragments made of } (r+s-n) \text{ cycles}\}\rho_{3(l-m)}(0)\\
&+\sum_{m=1}^{l}\sum_{n=0}^{r-1}\sum_{\mu=1}^{r-n}\lambda^{-l+m-\mu}\exp[-i\omega_{3(l-m)}t]\times\{\text{creation fragment with } m \text{ creation transitions and } n\\
&\text{cycles}\}\times\{\text{destruction fragment with } \mu \text{ destruction transitions and } (r-n-\mu) \text{ cycles}\}\rho_{3(l-m+\mu)}(0)\\
&+\sum_{m=1}^{l}\sum_{n=0}^{r-1}\sum_{\mu=1}^{r-n}\sum_{s=1}^{\infty}\sum_{k=s}^{r+s-n-\mu}\lambda^{-l+m-\mu+2s}\exp[-i\omega_{3(l-m)}t]\times\{\text{creation fragment with } m \text{ creation}\\
&\text{transitions and } n \text{ cycles}\}\times\{s \text{ diagonal fragments made of } k \text{ cycles}\}\times\{\text{destruction fragment}\\
&\text{made of } \mu \text{ destruction transitions and } (r-n-\mu-k+s) \text{ cycles}\}\times\rho_{3(l-m+\mu)}(0). \quad (10.5)\end{aligned}$$

Examples are given in Fig. 18. Let us again write explicitly in (10.5) the operator corresponding to the creation fragment at the left. If we change the order of the summations over μ and k in the last term, we obtain

$$\begin{aligned}\rho_{3l}{}''^{(2r)}(t)=&\sum_{m=1}^{l}\sum_{n=0}^{r-1}\lambda^{-l+m}\exp[-i\omega_{3(l-m)}t][\Omega_{3l,3(l-m)}]_n{}^{m}\times[\sum_{s=1}^{\infty}\lambda^{2s}\\
&\times\{s \text{ diagonal fragments made of } (r+s-n) \text{ cycles}\}\rho_{3(l-m)}(0)\\
&+\sum_{\mu=1}^{r-n}\lambda^{-\mu}\{\text{destruction fragment with } \mu \text{ destruction transitions and } (r-n-\mu) \text{ cycles}\}\rho_{3(l-m+\mu)}(0)\\
&+\sum_{s+1}^{\infty}\sum_{k=s}^{r-n+s-1}\sum_{\mu=1}^{r+s-n-k}\lambda^{-\mu+2s}\{s \text{ diagonal fragments made of } k \text{ cycles}\}\times\{\text{destruction fragment}\\
&\text{made of } \mu \text{ destruction transitions and } (r-n-\mu-k+s) \text{ cycles}\}\rho_{3(l-m+\mu)}(0)]\\
&+\sum_{m=1}^{l}[\Omega_{3l,3(l-m)}]_r{}^{m}\exp[-i\omega_{3(l-m)}t]\{\rho_{3(l-m)}(0)+\sum_{s=1}^{\infty}\lambda^{2s}[s \text{ diagonal fragments made of } s \text{ cycles}]\rho_{3(l-m)}(0),\end{aligned} \quad (10.6)$$

where again all the contributions at the right of the operators Ω have to be taken up to time t. If we take into account (9.7) for r replaced by $(r-n)$ and (9.2), we obtain

$$\rho_{3l}{}''^{(2r)}(t)=\sum_{m=1}^{l}\sum_{n=0}^{r}\exp[-i\omega_{3(l-m)}t][\Omega_{3l,3(l-m)}]_n{}^{m}\rho_{2(l-m)}{}'^{2(r-n)}(t) \quad (10.7)$$

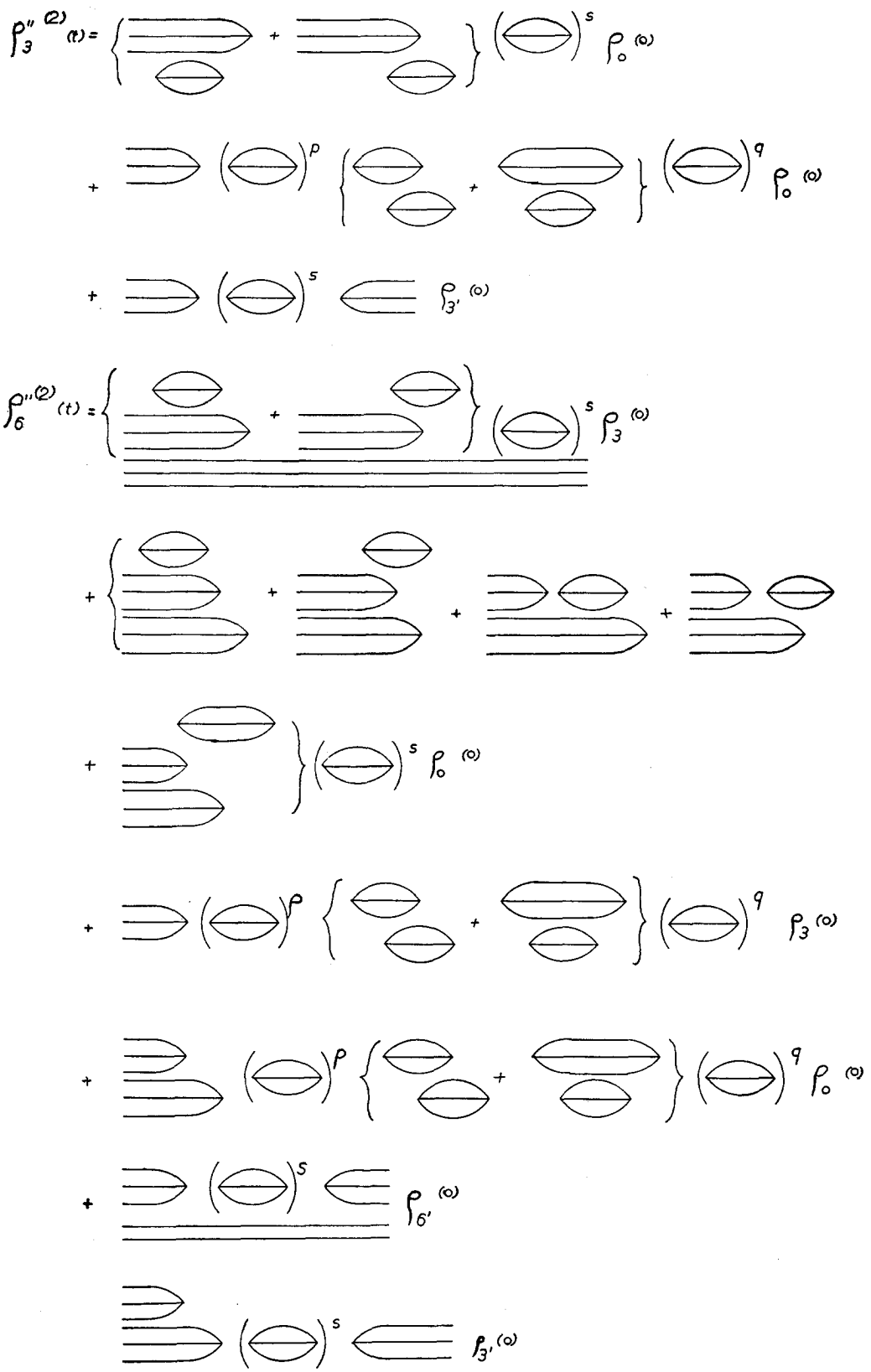


FIG. 18. 6′ has one correlation in common with 6; 3′ is a correlation which does not appear in 6.

in agreement with (10.1). As for $\rho_{3l}'(t)$ we can write an equation for $\rho_{3l}''(t)$ correct up to order λ^{l+2s}:

$$\rho_{3l}''(t)=\lambda^l\sum_{r=0}^{s}\lambda^{2r}\rho_{3n}''^{(2r)}(t)$$

$$=\lambda^l\sum_{m=1}^{l}\exp[-i\omega_{3(l-m)}t]\sum_{i=0}^{s}\lambda^{2i}[\Omega_{3l,3(l-m)}]_i^m\times\sum_{\alpha=0}^{s-i}\lambda^{2\alpha}\rho_{3(l-m)}'^{2\alpha}(t). \tag{10.8}$$

Once more the equation remains correct at order λ^{l+2s} if we extend the summation over α up to $\alpha=s$. Then

$$\rho_{3l}''(t)=\sum_{m=1}^{l}\lambda^m\exp[-i\omega_{3(l-m)}t]\{\sum_{i=0}^{s}\lambda^{2i}[\Omega_{3l,3(l-m)}]_i^m\}\rho_{3(l-m)}'(t). \tag{10.9}$$

This equation relates $\rho_{3l}''(t)$ to the functions $\rho_{3(l-m)}'(t)$. Whereas Eq. (9.14) describes the evolution of the function $\rho_{3n}(t)\exp[-i\omega_{3n}t]$ due to interactions among normal modes not involving the $3n$ normal modes described by that function, the function $\rho_{3n}''(t)$ contains the part of the evolution due to mechanical interactions between those $3n$ normal modes. As can be seen from (10.9), those interactions must occur at the end of the process in order to contribute asymptotically.

11. EXAMPLES OF APPLICATION—EQUATIONS UP TO ORDER λ^2

The result of the two preceding sections can be summarized in the following way. The evolution of the system at a given finite order λ^s is given by

$$\rho_{3n}(t)\exp(-i\omega_{3n}t)=\rho_{3n}'(t)\exp(-i\omega_{3n}t)+\rho_{3n}''(t), \qquad n\leqslant s \tag{11.1}$$

$$\frac{\partial\rho_{3n}'(t)}{\partial t}=\{\sum_{\alpha=0}^{(s-n)/2}\lambda^{2\alpha+2}[\Omega_{3n,3n}]_{\alpha+1}\}\rho_{3n}'(t), \tag{11.2}$$

$$\rho_{3n}''(t)=\sum_{m=1}^{n}\lambda^m\exp[-i\omega_{3(n-m)}t]$$

$$\times\{\sum_{i=0}^{(s-n)/2}\lambda^{2i}[\Omega_{3l,3(l-m)}]_i^m\}\rho_{3(l-m)}'(t), \tag{11.3}$$

$$\rho_0''(t)=0. \tag{11.3'}$$

It is important to remember that those equations are valid only for times such that asymptotic integration is permitted, i.e., much larger than characteristic molecular times (see Sec. 7)

$$t\gg\omega_D^{-1}=\tau_{\text{mol}}. \tag{11.4}$$

Moreover, by using an expansion in terms of $(\lambda^2t)^m\lambda^{2i}$, we introduced a further condition for the validity of these equations. Indeed, these will be valid for times t such that the product λ^2t is finite for λ small; more precisely, we consider times t of the order of the relaxation time τ_{rel}, which is assumed to be of the form

$$\tau_{\text{rel}}\sim\frac{1}{\lambda^2}\{1+\lambda^2+\cdots\}. \tag{11.5}$$

The condition (11.4) of validity of the asymptotic integrations is independent of λ. It is only the fact we use an expansion in terms of $(\lambda^2t)^m\lambda^{2i}$ which introduces, by combination of (11.4) and (11.5), the condition

$$\omega_D\gg\lambda^2/(1-\lambda^2+\cdots). \tag{11.6}$$

The fact that the Eqs. (11.2) and (11.3) are valid to order λ^s also means that all terms of the form $\lambda^{s+m}(\lambda^2t)^n$, $m\geqslant 1$ which may arise in their solutions are meaningless. In other words (11.2) and (11.3) are simply a compact way of writing the sets of Eqs. (9.1) and (10.1).

As an illustration, let us write more explicitly the equations of evolution up to order λ^2. At order λ^0, i.e., for weakly coupled system, we have the simple equation

$$(1/\lambda^2)(\partial\rho_0/\partial t)=[\Omega_{0,0}]_1\,\rho_0(t), \tag{11.7}$$

where the operator $[\Omega_{0,0}]_1$ corresponds to the cycle and is given in[7] [Eq. (2.13)]. This equation is nothing else than the usual master equation for weakly coupled systems. At order λ, we have to add the equations

$$(1/\lambda^2)(\partial\rho_3'/\partial t)=[\Omega_{3,3}]_1\,\rho_3'(t), \tag{11.8}$$

$$\rho_3''=[\Omega_{3,0}]_0{}^1\rho_0'(t), \tag{11.9}$$

where $[\Omega_{3,3}]_1$ corresponds to the cycle to which three freely propagated lines are added, i.e.,

$$[\Omega_{1_k1_{k'}1_{k''},1_k1_{k'}1_{k''}}]_1$$

$$=\sum_{l\,l'\,l''}\{i\zeta(\omega_l+\omega_{l'}+\omega_{l''})\langle 1_k1_{k'}1_{k''}|\delta L|1_k1_{k'}1_{k''}1_l1_{l'}1_{l''}\rangle$$

$$\times\langle 1_k1_{k'}1_{k''}1_l1_{l'}1_{l''}|\delta L|1_k1_{k'}1_{k''}\rangle$$

$$+\text{similar terms with } (1_l1_{l'}-1_{l''}),$$

$$(1_l-1_{l'}-1_{l''}),\text{ and }(-1_l-1_{l'}-1_{l''})\} \tag{11.10}$$

and $[\Omega_{3,0}]_0{}^1$ corresponds to the first diagram of Fig. 16

$$[\Omega_{1_k1_{k'}1_{k''},0}]_0{}^1=i\zeta(\omega_k+\omega_{k'}+\omega_{k''})$$

$$\times\langle 1_k1_{k'}1_{k''}|\delta L|0\rangle. \tag{11.11}$$

Finally, at order λ^2 the equation for ρ_0 itself is modified

$$(1/\lambda^2)(\partial\rho_0/\partial t)=[\Omega_{0,0}]_1\,\rho_0+\lambda^2[\Omega_{0,0}]_2\,\rho_0. \tag{11.12}$$

The operator $[\Omega_{0,0}]_2$ is given in Fig. 14 and is studied in Appendix II. At order λ^2, the functions ρ_3' and ρ_3'' are still given by (11.8) and (11.9). However, we still have to add the following equations in order to have the complete description of the evolution of the system to order λ^2:

$$(1/\lambda^2)(\partial\rho_6'/\partial t)=[\Omega_{6,6}]_1\,\rho_6'(t), \tag{11.13}$$

$$\rho_6''(t)=[\Omega_{6,3}]_0{}^1\rho_3'(t)+[\Omega_{6,3'}]_0{}^1\rho_{3'}'(t) +[\Omega_{6,0}]_0{}^2\rho_0'(t) \quad \text{if } 6=3+3', \tag{11.14}$$

where the operator in (11.13) corresponds to a cycle plus six freely propagated lines and the operators in (11.14) are shown in Fig. 16.

12. APPROACH TO EQUILIBRIUM

Let us first consider the behavior of the functions $\rho_{3n}'(t)$. The lowest-order contribution for each of these functions is given by (9.1) for $r=0$

$$\partial\rho_{3n}'^{(0)}(t)/\partial t=[\Omega_{3n,3n}]_1\,\rho_{3n}'^{(0)}(t). \tag{12.1}$$

The operators $[\Omega_{3n,3n}]_1$ have quite remarkable properties which have been studied previously in a paper by one of us (I.P.) and J. Philippot[7] (where these operators are denoted by $0_{\{3n\}}$). For instance, $[\Omega_{0,0}]_1$ is a self-adjoint operator with negative eigenvalues

$$[\Omega_{0,0}]_1\varphi_m=-\mu_m\varphi_m \quad \mu_m\gtrless 0, \tag{12.2}$$

if φ_m is an eigenfunction of $[\Omega_{0,0}]_1$. To the eigenvalue $\mu_0=0$ correspond the eigenfunctions

$$\varphi_0=f(H_0). \tag{12.3}$$

Now the operators $[\Omega_{3n,3n}]_1$ may be written in the following way:

$$[\Omega_{3n,3n}]_1=[\Omega_{0,0}]_1-M_{3n}{}^2+iF_{3n}, \tag{12.4}$$

where M_{3n} and F_{3n} are not operators but functions of the action variables. The real part of these operators is also a self-adjoint operator with negative eigenvalues

$$\{[\Omega_{0,0}]_1-M_{3n}{}^2\}\,\varphi_m{}^{(3n)}=-\mu_m{}^{(3n)}\varphi_m{}^{(3n)}, \tag{12.5}$$

but this time, the only way to obtain $\mu_m{}^{(3n)}=0$ is to take $\varphi_0{}^{(3n)}=0$. Therefore, for large times, one obtains as a consequence of (12.1)

$$\rho_{\{0\}}{}^{(0)}(t\to\infty)\to f(H_0), \tag{12.6}$$

$$\rho_{3n}'^{(0)}(t\to\infty)\to 0. \tag{12.7}$$

If we use these results, we can study Eqs. (9.1) in the next order. For large times, we can write

$$(1/\lambda^2)(\partial\rho_0'^{(2)}/\partial t)=[\Omega_{0,0}]_1\,\rho_0'^{(2)}+[\Omega_{0,0}]_2 f(H_0), \tag{12.8}$$

whereas for $\rho_{3n}'(t)$, we obtain

$$(1/\lambda^2)[\partial\rho_{3n}'^{(2)}(t)/\partial t]=[\Omega_{3n,3n}]_1\,\rho_{3n}'^{(2)}(t). \tag{12.9}$$

If we use again the properties of the operator in the rhs of (12.9), we obtain

$$\rho_{3n}'^{(2)}(t\to\infty)\to 0. \tag{12.10}$$

It is easily seen that if we go on with this procedure, we shall obtain

$$\rho_{3n}'(t\to\infty)\to 0 \tag{12.11}$$

whatever the order in λ (this order being always finite in order that our equations be valid). Taking into account the physical meaning of $\rho_{3n}'(t)$, we see that for large times the contribution of the scattering of correlations not involved in $\{3n\}$ vanishes.

The behavior of the function $\rho_0'(t)$ for large t is somewhat more involved. One can establish the following property (Appendix II):

$$[\Omega_{0,0}]_2 f(H_0)=-[\Omega_{0,0}]_1\{[\rho_0'^{(2)}]^{\text{equil}}+\text{arbitrary function of } H_0\}. \tag{12.12}$$

If we use this result, the properties of $[\Omega_{0,0}]_1$ and the normalization condition, we can show that Eq. (12.8) leads for large times to the correct equilibrium value of $\rho_0^{(2)}$. We have not yet been able to give a proof of the approach to equilibrium at an arbitrary order in λ. The difficulty lies in the mathematical complexity of the operators involved. Therefore, a generalization of (12.12) is not easy. However, for interacting particles, the proof exists and will be published in a subsequent paper. Let us now consider the behavior of the functions $\rho_{3n}''(t)$. For large times, we can use (12.10) in (10.1)

$$\rho_{3n}''^{(2r)}(t\to\infty)=\sum_{i=0}^{r}[\Omega_{3n,0}]_i{}^n\rho_{\{0\}}{}'^{2(r-i)}(t\to\infty). \tag{12.13}$$

The lowest-order contribution to $\rho_{3n}''(t\to\infty)$ can be shown to be the correct equilibrium expression. This can be easily verified for $n=1$, introducing (12.6) in (11.11)

$$\begin{aligned}\rho_3''^{(0)}&=\lambda[\Omega_{3,0}]_0{}^1 f(H_0)\\ &=\lambda i\zeta(\omega_k+\omega_{k'}+\omega_{k''})\langle 1_k 1_{k'} 1_{k''}|\delta L|0\rangle f(H_0)\\ &=\lambda V_{kk'k''}(J_kJ_{k'}J_{k''}/\omega_k\omega_{k'}\omega_{k''})^{\frac{1}{2}}(\partial f/\partial H_0).\end{aligned} \tag{12.14}$$

This is generalized for any n in the Appendix II. We have not been able to generalize those results for $r>0$, the difficulty being again the mathematical complexity of the operators involved.

13. DISCUSSION—MECHANISM OF IRREVERSIBILITY

An essential feature of the theory, which has already been stressed in the study of gases[1,2] is the fact that the evolution equations at any order λ^r, r finite, can be obtained from a *reduced closed set of equations* rather than from the infinite set of Liouville equations (4.3). Indeed, it can be easily verified for instance that the evolution equations at order λ^0, λ, and λ^2 could have been obtained as the asymptotic solution of the following sets of equations (see Fig. 19). One sees immediately

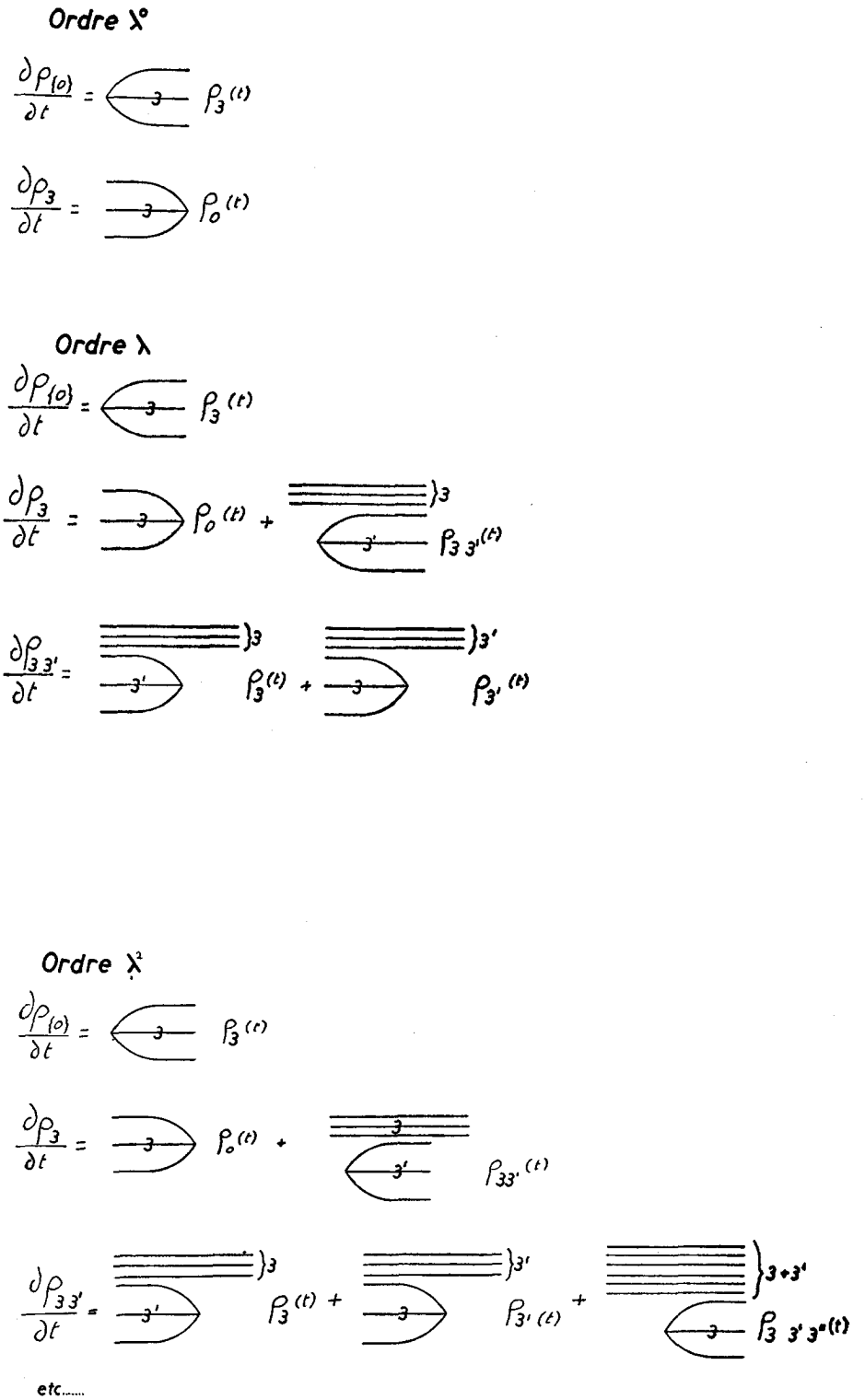


FIG. 19. Reduced sets of equations for the study of the evolution of the system at order λ^0, λ, and λ^2.

that these reduced closed sets of equations are obtained by neglecting in (4.3) all terms involving correlations between more than 3, 6, or 9 normal modes, respectively. In the same way, at order λ^r, the reduced set of equations will contain $(r+2)$ equations never involving diagrams with more than $3(r+1)$ lines. The existence of this reduced closed set of equations implies a well-defined mechanism for irreversibility. In the initial

Liouville equation (4.3), every Fourier component influences the evolution of any other Fourier component. The situation is now completely different. Only a finite number of Fourier components describing correlations between a finite number of normal modes influence each other. These Fourier components still influence the higher correlations, but they are no longer influenced by them. In other words, the reversible behavior is characterized by the fact that the information flows in both directions: from ρ_{3n} to ρ_{3m} $(n<m)$ and vice versa, while in the presence of irreversible processes this is no longer true. This can be represented graphically as follows:

Reversible behavior: $\rho_0 \rightleftharpoons \rho_3 \rightleftharpoons \rho_6 \rightleftharpoons \rho_9 \rightleftharpoons \rho_{12} \rightleftharpoons \cdots$.

Irreversible behavior:

order λ^0: $\rho_0 \rightleftharpoons \rho_3 \rightarrow \rho_6 \rightarrow \rho_9 \rightarrow \cdots$

order λ: $\rho_0 \rightleftharpoons \rho_3 \rightleftharpoons \rho_6 \rightarrow \rho_9 \rightarrow \cdots$

order λ^2: $\rho_0 \rightleftharpoons \rho_3 \rightleftharpoons \rho_6 \rightleftharpoons \rho_9 \rightarrow \cdots$.

It is interesting to consider the existence of such closed sets of equations as a kind of *contraction* or *shortening* in the description of the state of the system which occurs after times long in respect to the duration of an interaction. The existence of such kind of contraction was postulated by Bogolioubov,[14] and discussed among others by Kac and Uhlenbeck.[15]

It is obvious that the closed set of equations is only meaningful for times such that the asymptotic time integration is permitted and that means precisely times long in respect to the duration of an interaction.

The interesting feature is now that for times of the order of the relaxation times a further contraction occurs. This can easily be seen in terms of our decomposition of ρ_{3n} into ρ_{3n}' and ρ_{3n}''. The operators involved in the evolution equation for ρ_{3n}' are diagonal fragments which may be considered as describing *collisions* or *scattering processes* between normal modes. On the other hand, the equations for ρ_{3n}'' contain the direct interactions between the different normal modes necessary to build up the correlations involved in the set $\{3n\}$.

Now we have seen that the ρ_{3n}'''s go to zero in times of the order of the relaxation time. With the disappearance of ρ_{3n}', the system forgets the initial correlations. At the same time new correlations ρ_{3n}'' are built out of ρ_0. Moreover, ρ_0 satisfies a separate diagonal equation given by (9.14) for $n=0$.

Therefore, at this stage a new and more dominant role is played by ρ_0. The correlations become *functionals* of ρ_0. This is again in agreement with the mechanism postulated by Bogolioubov.[14]

In the case studied here this further reduction expresses simply the approach to equilibrium. Indeed let us write the equilibrium distribution in the form

$$\rho^{\text{equil}}(J_1\cdots,\alpha_1\cdots) = \frac{f(H_0+\lambda V)}{\int\cdots\int (dJd\alpha)^N f(H_0+\lambda V)}$$

$$= \left\{1+\lambda V\frac{\partial}{\partial H_0}+\frac{\lambda^2V^2}{2}\frac{\partial^2}{\partial H_0^2}-\frac{\lambda^2}{2}\int\cdots\int (dJd\alpha)^N \times\left[V^2\frac{\partial^2 f(H_0)}{\partial H_0^2}\right]+\lambda^3\cdots\right\} f(H_0) \quad (13.1)$$

with the normalization condition

$$\int\cdots\int (dJ)^N f(H_0)=1. \quad (13.2)$$

Clearly, all ρ_{3n}^{equil}, which are the Fourier coefficients of ρ^{equil}, by (13.1) may be expressed in terms of operations performed on $f(H_0)$, that is, as functionals of $f(H_0)$.

However, this further contraction of the description is by no means trivial in hydrodynamical problems where one may expect that in the same way the correlations will become functionals of the single particle distribution functions depending on the local state of the system.

This problem will be studied in detail in a subsequent paper.

Also a detailed application of this method to the calculation of thermal conductivity of strongly anharmonic solids will be published separately by one of us (F.H.) and L. Blum.

ACKNOWLEDGMENT

We want to thank P. Resibois for important remarks about the asymptotic time integrations.

APPENDIX I. TIME INTEGRALS

Let us evaluate integral (7.3) for various situations. As the factor $e^{i\gamma t}$ plays no role, we shall drop it. Thus we have

$$I(t)=\int_0^t dt_1\cdots\int_0^{t_{n-1}} dt_n \exp[i\alpha_0(t-t_1)] \times\prod_{i=1}^{n-1}\exp[i\alpha_i(t_i-t_{i+1})]\exp(i\alpha_n t_n). \quad (A.1)$$

As we have said in Sec. 7, asymptotically this integral is equal to the sum of all nonoscillating terms where the variables α_i in the denominators are replaced by complex variables $\alpha_i\pm i\epsilon$, and the limit $\epsilon\rightarrow 0$ is taken after integration over α's. Our problem, in order to

[14] N. N. Bogolioubov, J. Phys. U.S.S.R. **10**, 265 (1946).

[15] M. Kac, *Probability and Related Topics in Physical Sciences* (Interscience Publishers, Inc., New York, 1959), p. 132; Appendix I by G. E. Uhlenbeck.

establish theorems I and II is then to find the nonoscillating contributions to (A.1) and their order of magnitude with respect to t. The easiest way to perform an integral like (A.1) is to introduce new variables corresponding to the time intervals (t_i-t_{i+1}). Let us first consider the case where $\alpha_0=0$, i.e., the case of diagonal diagrams or destruction diagrams. Then we choose our new variables in the following way:

$$\begin{aligned} t_i-t_{i+1}&=\tau_i, \quad 1\leqslant i\leqslant n-1 \\ t_n&=\tau_n. \end{aligned} \tag{A.2}$$

The domain of integration is defined by the conditions

$$0\leqslant t_n\leqslant t_{n-1}\cdots\leqslant t_i\leqslant t_{i-1}\cdots\leqslant t_1\leqslant t. \tag{A.3}$$

Equivalent conditions are

$$\begin{aligned} \tau_i&\geqslant 0, \quad 1\leqslant i\leqslant n \\ t-\textstyle\sum_i \tau_i&\geqslant 0. \end{aligned} \tag{A.4}$$

The easiest way to take account of these conditions is to take the integral over each τ_i from 0 to ∞ and multiply the integrand by the Heaviside function $\eta(t-\sum_i \tau_i)$ defined by

$$\begin{aligned} \eta(x)&=1 \quad x>0 \\ &=0 \quad x<0. \end{aligned} \tag{A.5}$$

If we use the representation of the Heaviside function

$$\eta(x)=2\pi i\int_{-\infty-i\epsilon}^{+\infty-i\epsilon} dz\frac{e^{izx}}{z}, \tag{A.6}$$

we obtain for (A.1) in the case $\alpha_0=0$

$$I(t)=2\pi i\int_{-\infty-i\epsilon}^{+\infty-i\epsilon} dz\frac{e^{izt}}{z}\times\prod_{i=1}^{n}\int_0^\infty d\tau_i \times\exp[i(\alpha_i-z)\tau_i]. \tag{A.7}$$

We can easily perform the integrations over the τ_i's

$$I(t)=2\pi i\int_{-\infty-i\epsilon}^{+\infty-i\epsilon} dz\frac{e^{izt}}{z}\prod_{i=1}^{n}\frac{1}{i(z-\alpha_i)}. \tag{A.8}$$

Now if the diagonal region of the diagram contains m diagonal fragments, m of the α_i are equal to zero and (A.8) becomes

$$I(t)=(-i)^m(2\pi i)\int_{-\infty-i\epsilon}^{+\infty-i\epsilon} dz\frac{e^{izt}}{z^{m+1}}\prod_{\substack{\text{all}\\ \alpha_i\neq 0}}\frac{1}{i(z-\alpha_i)}. \tag{A.9}$$

To perform the integral over z, we close the contour by a semicircle at infinity in the upper half-plane and use residue calculus. Then we obtain

$$I(t)=\frac{(-i)^m}{m!}\left\{\frac{\partial^m}{\partial z^m}e^{izt}\prod_{\substack{\text{all}\\ \alpha_i\neq 0}}\frac{1}{i(z-\alpha_i)}\right\}_{z=0}. \tag{A.10}$$

However, only the first contribution, i.e., the contribution proportional to t^m is relevant. Indeed, all the other contributions are at least of order τ/t, where τ is the characteristic molecular time introduced in (11.4), with respect to the first term. The neglect of all oscillating terms corresponds to neglecting terms of that order (see the discussion for the cycle in Sec. 7). We have thus to neglect also constant terms of this order. In other words, the asymptotic contribution of (A.7) will be of the form

$$I(t)=\frac{t^m}{m!}\left\{\prod_{\substack{\text{all}\\ \alpha_i\neq 0}}\frac{1}{i(\alpha_i-z)}\right\}_{z=0}, \tag{A.11}$$

which establishes theorem I.

We may also notice that in this asymptotic evaluation each of the m diagonal fragments as well as the destruction fragment is an instantaneous event on a macroscopic scale. If the last diagonal fragment at the left contains $(2s)$ transitions, the α_i's for $1\leqslant i\leqslant 2s-1$ are different from zero, whereas $\alpha_{2s}=0$; (A.11) may then be written

$$\begin{aligned} I(t)&=\left\{\prod_{i=1}^{2s-1}\frac{1}{i(z-\alpha_i)}\right\}_{z=0}\int_0^t dt_1\int_0^{t_1} dt_2\cdots \\ &\quad\times\int_0^{t_{m-1}} dt_m\left\{\prod_{\substack{\text{all}\\ \alpha_i\neq 0\\ i>2s}}\frac{1}{i(z-\alpha_i)}\right\}_{z=0} \\ &=\left\{\sum_{i=1}^{2s-1}\frac{1}{i(z-\alpha_i)}\right\}_{z=0}\int_0^t dt_1\times\{\text{asymptotic} \\ &\qquad\text{contribution of the remaining part of the} \\ &\qquad\text{diagram taken up to time } t_1\}. \end{aligned} \tag{A.12}$$

This result has been used in (9.8), the factor in front of the rhs of (A.12) corresponding to the ζ functions, etc., which appear in the operator $[\Omega_{3n,3n}]_s$.

Let us now consider the case $\alpha_0\neq 0$, $\alpha_n=0$, i.e., the case of creation diagrams. The natural change of variables in this case is obviously

$$\begin{aligned} t-t_1&=\tau_1 \\ t_i-t_{i+1}&=\tau_{i+1}, \quad 1\leqslant i\leqslant n-1. \end{aligned} \tag{A.13}$$

The domain of integration is defined by

$$\tau_i\geqslant 0, \quad t-\textstyle\sum_i \tau_i\geqslant 0. \tag{A.14}$$

If we introduce again the Heaviside function to take account of the last condition, we obtain

$$I(t)=2\pi i\int_{-\infty-i\epsilon}^{+\infty-i\epsilon} dz\frac{e^{izt}}{z}\prod_{i=0}^{n-1}\int_0^\infty d\tau_{i+1} \times\exp[i(\alpha_i-z)\tau_{i+1}]. \tag{A.15}$$

By following the same procedure as before, it is immediately seen that if m is the number of diagonal fragments preceding the creation fragment, the asymptotic contribution of $I(t)$ is again proportional to t^m. This contribution is

$$I(t)=\frac{t^m}{m!}\left\{\prod_{\substack{\text{all}\\ \alpha_i\neq 0}}\frac{1}{i(z-\alpha_i)}\right\}_{z=0}. \tag{A.16}$$

If the creation fragment at the end of the diagram contains s transitions ($s\geqslant 1$) the α_i's for $1\leqslant i\leqslant s$ are different from zero, whereas $\alpha_{s+1}=0$; (A.16) may then be written

$$I(t)=\prod_{i=1}^{s}\left\{\frac{1}{i(z-\alpha_i)}\right\}_{z=0}\times\left\{\frac{t^m}{m!}\prod_{\substack{\text{all}\\ \alpha_i\neq 0\\ i>s+1}}\frac{1}{i(z-\alpha_i)}\right\}_{z=0}$$

$$=\prod_{i=1}^{s}\left\{\frac{1}{i(z-\alpha_i)}\right\}_{z=0}\times\{\text{asymptotic contribution of the remaining part of the diagram taken up to time } t\}. \tag{A.17}$$

This result has been used in (10.3), the factor in front of the rhs of (A.17) corresponding to the ζ functions, etc., which appear in the operator $[\Omega_{3n,3m}]_s$. The last case of reducible diagram we have to consider is the reducible destruction-creation diagram. In that case α_0 and α_τ are both different from zero, but at least one of the α_i's ($1\leqslant i\leqslant n-1$) is equal to zero. In fact $(m+1)$ of these α_i's are equal to zero if there are m diagonal fragments in between the creation and the destruction fragment. Let us take $\alpha_j=0$. Then, we choose the τ variables in the following way:

$$t-t_1=\tau_1,\quad t_i-t_{i+1}=\tau_{i+1}$$
$$1\leqslant i\leqslant j-1 \quad\text{and}\quad 1+j\leqslant i\leqslant n.$$

Then we have

$$I(t)=2\pi i\int_{-\infty-i\epsilon}^{+\infty-i\epsilon}dz\frac{e^{izt}}{z}\prod_{i=0}^{j-1}\int_0^\infty d\tau_i\exp[i(\alpha_i-z)\tau_i]$$
$$\times\prod_{l=j+1}^{n}\int_0^\infty d\tau_l\exp[i(\alpha_l-z)\tau_l]. \tag{A.18}$$

As m of the remaining α's are equal to zero, it is easily verified that the asymptotic contribution of $I(t)$ is again proportional to t^m, which again is in complete agreement with theorem I of Sec. 7. It can also be verified that (A.17) holds for these diagrams also. We still have to consider the case of the irreducible destruction creation diagrams. Then, none of the α's is equal to zero. If we use for instance, the τ variables as defined by (A.2), we obtain

$$I(t)=2\pi i\int_{-\infty-i\epsilon}^{+\infty-i\epsilon}dz\frac{e^{izt}}{z}e^{i\alpha_0 t}\prod_{i=0}^{n}\int_0^\infty d\tau_i$$
$$\times\exp[i(\alpha_i-\alpha_0-z)\tau_i]$$
$$=2\pi i\int_{-\infty-i\epsilon}^{+\infty-i\epsilon}dz\frac{e^{i(z+\alpha_0)t}}{z}\prod_{i=0}^{n}\frac{1}{i(\alpha_i-\alpha_0-z)}. \tag{A.19}$$

If we take as new variable $z'=z+\alpha_0$ and use residue calculus we immediately see that no term is nonoscillating. Therefore, the contribution vanishes asymptotically as stated in theorem II.

APPENDIX II. PROPERTIES OF SOME OF THE OPERATORS WHICH APPEAR IN THE EQUATIONS OF EVOLUTION

The first property we have to establish is relation (11.13). The operator $[\Omega_{0,0}]_2$ is given in terms of diagrams in Fig. 14. Explicitly, we have

$$t[\Omega_{0,0}]_2=\text{asymptotic limit}\sum_{kk'k''}\sum_{ll'l''}\left[\int_0^t dt_1\exp[-i(\omega_k+\omega_{k'}+\omega_{k''})t_1]\int_0^{t_1}dt_2\exp[-i(\omega_l+\omega_{l'}+\omega_{l''})t_2]\right.$$
$$\times\langle 0|\delta L|1_k1_{k'}1_{k''}\rangle\langle 1_k1_{k'}1_{k''}|\delta L|1_k1_{k'}1_{k''}1_l1_{l'}1_{l''}\rangle\left\{\int_0^{t_2}dt_3\exp[i(\omega_l+\omega_{l'}+\omega_{l''})t_3]\right.$$
$$\times\langle 1_k1_{k'}1_{k''}1_l1_{l'}1_{l''}|\delta L|1_k1_{k'}1_{k''}\rangle\int_0^{t_3}dt_4\exp[i(\omega_k+\omega_{k'}+\omega_{k''})t_4]\langle 1_k1_{k'}1_{k''}|\delta L|0\rangle$$
$$\left.\left.+\text{similar term obtained by permutation of } (kk'k'') \text{ and } (ll'l'')\right\}+\text{similar terms considering } -1_k \text{ instead of } 1_k, \text{ etc.}\right]. \tag{B.1}$$

If we take into account

$$\text{asymptotic limit}\int_0^t dt_1\int_0^{t_1}dt_2\int_0^{t_2}dt_3\int_0^{t_3}dt_4e^{-i\alpha t_1}e^{-i\beta t_2}e^{i\beta t_3}e^{i\alpha t_4}=\lim_{\substack{\epsilon\to 0\\ \epsilon'\to 0}}\frac{t}{[i(\alpha+i\epsilon)]^2i(\alpha+\beta+i\epsilon+i\epsilon')}, \tag{B.2}$$

$$\text{asymptotic limit}\int_0^t dt_1\int_0^{t_1}dt_2\int_0^{t_2}dt_3\int_0^{t_3}dt_4e^{-i\alpha t_1}e^{-i\beta t_2}e^{i\alpha t_3}e^{i\beta t_4}=\lim_{\substack{\epsilon\to 0\\ \epsilon'\to 0}}\frac{t}{[i(\alpha+i\epsilon)]i[\alpha+\beta+i\epsilon+i\epsilon']i(\beta+i\epsilon')}, \tag{B.3}$$

we have

$$[\Omega_{0,0}]_2 f(H_0)=\lim_{\substack{\epsilon\to 0\\ \epsilon'\to 0}}\sum_{kk'k''}\sum_{ll'l''}\Bigg[\frac{1}{i(\omega_k+\omega_{k'}+\omega_{k''}+i\epsilon)}\frac{1}{i(\omega_k+\omega_{k'}+\omega_{k''}+\omega_l+\omega_{l'}+\omega_{l''}+i\epsilon+i\epsilon')}\langle 0|\delta L|1_k1_{k'}1_{k''}\rangle$$

$$\times\langle 1_k1_{k'}1_{k''}|\delta L|1_k1_{k'}1_{k''}1_l1_{l'}1_{l''}\rangle\Bigg\{\frac{1}{i(\omega_k+\omega_{k'}+\omega_{k''}+i\epsilon)}\langle 1_k1_{k'}1_{k''}1_l1_{l'}1_{l''}|\delta L|1_k1_{k'}1_{k''}\rangle$$

$$\times\langle 1_k1_{k'}1_{k''}|\delta L|0\rangle+\frac{1}{i(\omega_l+\omega_{l'}+\omega_{l''}+i\epsilon')}\langle 1_k1_{k'}1_{k''}1_l1_{l'}1_{l''}|\delta L|1_l1_{l'}1_{l''}\rangle$$

$$\times\langle 1_l1_{l'}1_{l''}|\delta L|0\rangle\Bigg\}+\text{similar terms}\Bigg]. \tag{B.4}$$

If we take into account (1.6), we have

$$\lim_{\epsilon\to 0}\frac{1}{i(\omega_k+\omega_{k'}+\omega_{k''}+i\epsilon)}\langle 1_k1_{k'}1_{k''}|\delta L|0\rangle=V_{kk'k''}\left(\frac{J_kJ_{k'}J_{k''}}{\omega_k\omega_{k'}\omega_{k''}}\right)^{\frac{1}{2}}\frac{\partial f}{\partial H_0}, \tag{B.5}$$

and therefore

$$[\Omega_{0,0}]_2 f(H_0)=\lim_{\substack{\epsilon\to 0\\ \epsilon'\to 0}}\sum_{kk'k''}\sum_{ll'l''}\Bigg\{\frac{1}{i(\omega_k+\omega_{k'}+\omega_{k''}+i\epsilon)}\frac{1}{i(\omega_k+\omega_{k'}+\omega_{k''}+\omega_l+\omega_{l'}+\omega_{l''}+i\epsilon+i\epsilon')}\langle 0|\delta L|1_k1_{k'}1_{k''}\rangle$$

$$\times\langle 1_k1_{k'}1_{k''}|\delta L|1_k1_{k'}1_{k''}1_l1_{l'}1_{l''}\rangle\Bigg[\langle 1_k1_{k'}1_{k''}1_l1_{l'}1_{l''}|\delta L|1_k1_{k'}1_{k''}\rangle V_{kk'k''}\left(\frac{J_kJ_{k'}J_{k''}}{\omega_k\omega_{k'}\omega_{k''}}\right)^{\frac{1}{2}}$$

$$+\langle 1_k1_{k'}1_{k''}1_l1_{l'}1_{l''}|\delta L|1_l1_{l'}1_{l''}\rangle V_{ll'l''}\left(\frac{J_lJ_{l'}J_{l''}}{\omega_l\omega_{l'}\omega_{l''}}\right)^{\frac{1}{2}}\Bigg]\frac{\partial f}{\partial H_0}+\text{similar terms}\Bigg\}. \tag{B.6}$$

If we use again (1.6) and take into account that we might have $k=l$, etc.

$$\langle 1_k1_{k'}1_{k''}1_l1_{l'}1_{l''}|\delta L|1_k1_{k'}1_{k''}\rangle\left(\frac{J_kJ_{k'}J_{k''}}{\omega_k\omega_{k'}\omega_{k''}}\right)^{\frac{1}{2}}\frac{\partial f(H_0)}{\partial H_0}$$

$$=-iV_{ll'l''}\left(\frac{J_lJ_{l'}J_{l''}}{\omega_l\omega_{l'}\omega_{l''}}\right)^{\frac{1}{2}}\Bigg[\frac{\delta_{kl}+\delta_{k'l}+\delta_{k''l}}{2J_l}+\frac{\delta_{kl'}+\delta_{k'l'}+\delta_{k''l'}}{2J_{l'}}+\frac{\delta_{kl''}+\delta_{k'l''}+\delta_{k''l''}}{2J_{l''}}$$

$$-\frac{\partial}{\partial J_l}-\frac{\partial}{\partial J_{l'}}-\frac{\partial}{\partial J_{l''}}\Bigg]\left(\frac{J_kJ_{k'}J_{k''}}{\omega_k\omega_{k'}\omega_{k''}}\right)^{\frac{1}{2}}\frac{\partial f}{\partial H_0}$$

$$=iV_{ll'l''}\left(\frac{J_kJ_{k'}J_{k''}J_lJ_{l'}J_{l''}}{\omega_k\omega_{k'}\omega_{k''}\omega_l\omega_{l'}\omega_{l''}}\right)^{\frac{1}{2}}(\omega_l+\omega_{l'}+\omega_{l''})\frac{\partial^2 f}{\partial H_0^2}, \tag{B.7}$$

we obtain

$$[\Omega_{0,0}]_2 f(H_0)=\lim_{\epsilon\to 0}\sum_{kk'k''}\sum_{ll'l''}\Bigg[V_{kk'k''}V_{ll'l''}\frac{1}{i(\omega_k+\omega_{k'}+\omega_{k''}+i\epsilon)}\langle 0|\delta L|1_k1_{k'}1_{k''}\rangle\langle 1_k1_{k'}1_{k''}|\delta L|1_k1_{k'}1_{k''}1_l1_{l'}1_{l''}\rangle$$

$$\times\left(\frac{J_kJ_{k'}J_{k''}J_lJ_{l'}J_{l''}}{\omega_k\omega_{k'}\omega_{k''}\omega_l\omega_{l'}\omega_{l''}}\right)^{\frac{1}{2}}\frac{\partial^2 f}{\partial H_0^2}+\text{similar terms}\Bigg]. \tag{B.8}$$

If we consider simultaneously the term written explicitly in (B.8) and the corresponding term where $(1_l1_{l'}1_{l''})$ is replaced by $(-1_l-1_{l'}-1_{l''})$, we obtain

$$[\Omega_{0,0}]_2 f(H_0)=\lim_{\epsilon\to 0}\sum_{kk'k''}\sum_{ll'l''}\Bigg[iV_{kk'k''}|V_{ll'l''}|^2\frac{1}{i(\omega_k+\omega_{k'}+\omega_{k''}+i\epsilon)}\Bigg\{\frac{\delta_{kl}+\delta_{k'l}+\delta_{k''l}}{J_l}+\frac{\delta_{kl'}+\delta_{k'l'}+\delta_{k''l'}}{J_{l'}}$$

$$+\frac{\delta_{kl''}+\delta_{k'l''}+\delta_{k''l''}}{J_{l''}}\Bigg\}\left(\frac{J_lJ_{l'}J_{l''}}{\omega_l\omega_{l'}\omega_{l''}}\right)\left(\frac{J_kJ_{k'}J_{k''}}{\omega_k\omega_{k'}\omega_{k''}}\right)^{\frac{1}{2}}\frac{\partial^2 f}{\partial H_0^2}+\text{similar terms}\Bigg]. \tag{B.9}$$

On taking into account

$$\left[\frac{\delta_{kl}}{J_l}+\frac{\delta_{kl'}}{J_{l'}}+\frac{\delta_{kl''}}{J_{l''}}\right]J_lJ_{l'}J_{l''}=\frac{\partial}{\partial J_k}(J_lJ_{l'}J_{l''}), \tag{B.10}$$

(B.9) can be written

$$[\Omega_{0,0}]_2 f(H_0)=\sum_{kk'k''}\left\{\zeta(\omega_k+\omega_{k'}+\omega_{k''})V_{kk'k''}\langle 0|\delta L|1_k1_{k'}1_{k''}\rangle\left(\frac{J_kJ_{k'}J_{k''}}{\omega_k\omega_{k'}\omega_{k''}}\right)^{\frac{1}{2}}\frac{\partial^2 f}{\partial H_0^2}\left[\frac{\partial}{\partial J_k}+\frac{\partial}{\partial J_{k'}}+\frac{\partial}{\partial J_{k''}}\right]\right.$$

$$\left.\times\sum_{ll'l''}|V_{ll'l''}|^2\left(\frac{J_lJ_{l'}J_{l''}}{\omega_l\omega_{l'}\omega_{l''}}\right)+\text{similar terms}\right\}. \tag{B.11}$$

If all terms are written explicitly, we have

$$[\Omega_{0,0}]_2 f(H_0)=-i\sum_{kk'k''}\left\{V_{kk'-k''}\delta(\omega_k+\omega_{k'}-\omega_{k''})\langle 0|\delta L|1_k1_{k'}-1_{k''}\rangle\left(\frac{J_kJ_{k'}J_{k''}}{\omega_k\omega_{k'}\omega_{k''}}\right)^{\frac{1}{2}}\frac{\partial^2 f}{\partial H_0^2}\left\{\frac{\partial}{\partial J_k}+\frac{\partial}{\partial J_{k'}}-\frac{\partial}{\partial J_{k''}}\right\}\right.$$

$$\left.\times\sum_{ll'l''}[|V_{ll'l''}|^2+3|V_{ll'-l''}|^2]\left(\frac{J_lJ_{l'}J_{l''}}{\omega_l\omega_{l'}\omega_{l''}}\right)+\text{c.c.}\right\}. \tag{B.12}$$

On taking into account the property

$$\alpha\delta(\alpha)=0 \tag{B.13}$$

and (1.6), this can be written

$$[\dot{\Omega}_{0,0}]_2 f(H_0)=-\sum_{kk'k''}\{\delta(\omega_k+\omega_{k'}-\omega_{k''})\langle 0|\delta L|1_k1_{k'}-1_{k''}\rangle\langle 1_k1_{k'}-1_{k''}|\delta L|0\rangle+\text{c.c.}\}$$

$$\times\left\{\sum_{ll'l''}[|V_{ll'l''}|^2+3|V_{ll'-l''}|^2]\left(\frac{J_lJ_{l'}J_{l''}}{\omega_l\omega_{l'}\omega_{l''}}\right)\frac{\partial^2 f}{\partial H_0^2}\right\}, \tag{B.14}$$

which establishes (11.13).

The last property we have to establish is the generalization of (11.12) for $n>1$

$$\begin{aligned}\rho_{3n}{}^{\prime\prime(0)}(t\to\infty)&=[\Omega_{3n,0}]_0{}^n f(H_0)\\&=[\rho_{3n}{}^{\prime\prime(0)}]^{\text{equil}}.\end{aligned} \tag{B.15}$$

If we denote by $3_1, 3_2, \cdots, 3_n$ the n correlations involved in ρ_{3n}, we have to show

$$[\Omega_{3_1\cdots 3_n,0}]_0{}^n f(H_0)=V_{3_1}\cdots V_{3_n}\left(\frac{J_{3_1}\cdots J_{3_n}}{\omega_{3_1}\cdots\omega_{3_n}}\right)^{\frac{1}{2}}\frac{\partial^n f(H_0)}{\partial H_0{}^n}, \tag{B.16}$$

where the operator in the lhs contains all possible permutations of the n creation transitions and where in the rhs if $3_1=1_k1_{k'}1_{k''}$, then $V_{3_1}=V_{kk'k''}$, $(J_{3_1}/\omega_{3_1})=(J_kJ_{k'}J_{k''}/\omega_k\omega_{k'}\omega_{k''})$. We shall establish (B.16) by recurrence. First of all, let us consider the structure of the terms in the operator in the lhs. We have

$$\sum_{\substack{\text{all possible}\\ \text{permutations}\\ \text{of }\{3_1\cdots 3_n\}}}\text{asymptotic limit}\int_0^t dt_1\cdots\int_0^{t_{n-1}}dt_n\exp[-i(\omega_{3_1}+\omega_{3_2}+\cdots+\omega_{3_n})(t-t_1)$$

$$\times\exp[-i(\omega_{3_1}+\omega_{3_2}+\cdots+\omega_{3_{n-1}})(t_1-t_2)]\times\cdots\times\exp[-i(\omega_{3_1}+\omega_{3_2}+\cdots+\omega_{3_{n-j}})(t_j-t_{j+1})]\times\cdots$$

$$\times\exp[-i\omega_{3_1}(t_{n-1}-t_n)]\langle 3_1\cdots 3_n|\delta L|3_1\cdots 3_{n-1}\rangle\langle 3_1\cdots 3_{n-1}|\delta L|3_1\cdots 3_{n-2}\rangle\times\cdots\times\langle 3_1|\delta L|0\rangle$$

$$=\lim_{\substack{\epsilon_1\to 0\\ \vdots\\ \epsilon_n\to 0}}\sum_{\substack{\text{all possible}\\ \text{permutations}\\ \text{of }\{3_1\cdots 3_n\}}}\frac{1}{i(\omega_{3_1}+\omega_{3_2}+\cdots+\omega_{3_n}-i\epsilon_1-i\epsilon_2-\cdots-i\epsilon_n)}\times\frac{1}{i(\omega_{3_1}+\omega_{3_2}+\cdots+\omega_{3_{n-1}}-i\epsilon_1\cdots-i\epsilon_{n-1})}\times\cdots$$

$$\times\frac{1}{i(\omega_{3_1}-i\epsilon_1)}\langle 3_1\cdots 3_n|\delta L|3_1\cdots 3_{n-1}\rangle\times\langle 3_1\cdots 3_{n-1}|\delta L|3_1\cdots 3_{n-2}\rangle\cdots\langle 3_1|\delta L|0\rangle. \tag{B.17}$$

This can be written as

$$\lim_{\epsilon\to 0}\sum_{j=1}^{n}\frac{1}{i(\omega_{3_1}+\omega_{3_2}+\cdots+\omega_{3_n}-i\epsilon)}\langle 3_1\cdots 3_n|\delta L|3_1\cdots 3_{j-1}3_{j+1}\cdots 3_n\rangle[\Omega_{3_1\cdots 3_{j-1}3_{j+1}\cdots 3_n,0}]_0{}^{n-1}. \tag{B.18}$$

Therefore, if we assume (B.16) to be true for $n-1$, for n we have

$$[\Omega_{3_1\cdots 3_n,0}]_0{}^n f(H_0)=\lim_{\epsilon\to 0}\sum_{j=1}^{n}\frac{1}{i(\omega_{3_1}+\omega_{3_2}+\cdots+\omega_{3_n}-i\epsilon)}\langle 3_1\cdots 3_n|\delta L|3_1\cdots 3_{j-1}3_{j+1}\cdots 3_n\rangle$$
$$\times V_{3_1}V_{3_2}\cdots V_{3_{j-1}}V_{3_{j+1}}\cdots V_{3_n}\left(\frac{J_{3_1}J_{3_2}\cdots J_{3_{j-1}}J_{3_{j+1}}\cdots J_{3_n}}{\omega_{3_1}\omega_{3_2}\cdots\omega_{3_{j-1}}\omega_{3_{j+1}}\cdots\omega_{3_n}}\right)^{\frac{1}{2}}\frac{\partial^{n-1}f(H_0)}{\partial {H_0}^{n-1}}. \quad \text{(B.19)}$$

If we take into account that

$$\langle 3_1\cdots 3_n|\delta L|3_1\cdots 3_{j-1}3_{j+1}\cdots 3_n\rangle\left(\frac{J_{3_1}\cdots J_{3_{j-1}}J_{3_{j+1}}\cdots J_{3_n}}{\omega_{3_1}\cdots\omega_{3_{j-1}}\omega_{3_{j+1}}\cdots\omega_{3_n}}\right)^{\frac{1}{2}}\frac{\partial^{n-1}f(H_0)}{\partial {H_0}^{n-1}}$$
$$=iV_{3_j}\left(\frac{J_{3_1}\cdots J_{3_{j-1}}J_{3_j}J_{3_{j+1}}\cdots J_{3_n}}{\omega_{3_1}\cdots\omega_{3_{j-1}}\omega_{3_j}\omega_{3_{j+1}}\cdots\omega_{3_n}}\right)^{\frac{1}{2}}\left\{\frac{\partial}{\partial J_{3_j}}\frac{\partial^{n-1}f}{\partial {H_0}^{n-1}}\right\}$$
$$=iV_{3_j}\left(\frac{J_{3_1}\cdots J_{3_n}}{\omega_{3_1}\cdots\omega_{3_n}}\right)^{\frac{1}{2}}\omega_{3_j}\frac{\partial^n f(H_0)}{\partial {H_0}^n}, \quad \text{(B.20)}$$

(B.16) is easily established.

Liouville Equation and the Resolvent Formalism

S. TEITLER AND R. F. WALLIS
U. S. Naval Research Laboratory, Washington, D. C.
(Received May 2, 1960)

With the use of the Prigogine and Balescu representation for the description of relaxation starting from the Liouville equation, we apply Van Hove's resolvent formalism to obtain higher-order contributions for classical weakly coupled homogeneous gases. These results confirm Van Hove's observation that, for appropriate representations, persistent (long time) effects are determined by diagonal matrix elements of operators consisting of products of perturbative operators separated by diagonal operators. The mechanism of relaxation is also discussed.

INTRODUCTION

THE important contributions of Prigogine and co-workers using the Liouville equation as the starting point in the investigation of various physical problems is well exemplified by the work of Prigogine and Balescu[1] (hereafter referred to as P-B), dealing with relaxation in classical gases. In general, they deal with a system of N particles and volume $8\pi^3\Omega$ in the limit

$$N\to\infty,\quad \Omega\to\infty,\quad N/\Omega=8\pi^3c=\text{finite constant},\tag{1.1}$$

where c is the average concentration or number density. The Hamiltonian for the system may be written

$$H=H_0+V=\sum_j H_j+\lambda\sum_{j<r}V_{jr},\tag{1.2}$$

where H_j is dependent on the individual particle coordinates and momenta. The potential V is here written as a sum of binary interactions but may more generally be made up of higher-order interactions.

$$\frac{\partial f_N}{\partial t}+\sum_m\left(\frac{\partial H}{\partial\mathbf{p}_m}\cdot\frac{\partial}{\partial\mathbf{q}_m}-\frac{\partial H}{\partial\mathbf{q}_m}\cdot\frac{\partial}{\partial\mathbf{p}_m}\right)f_N=0,\tag{1.3}$$

where f_N, the N particle distribution function in phase space, is normalized to unity:

$$\int(d\mathbf{q})^N(d\mathbf{p})^N f_N=1.\tag{1.4}$$

The unperturbed Hamiltonian H_0 is either transformed to, or originally expressed in, a representation in which all the canonical coordinates are cyclic and V depends generally on the momenta and periodically on the cyclic coordinates. Then Eq. (1.3) may be written

$$i\frac{\partial f_N}{\partial t}=-i\sum_m\frac{\partial H_0}{\partial\mathbf{p}_m}\cdot\frac{\partial}{\partial\mathbf{q}_m}f_N+i\lambda\sum_m\left[\frac{\partial V}{\partial\mathbf{q}_m}\cdot\frac{\partial}{\partial\mathbf{p}_m}-\frac{\partial V}{\partial\mathbf{p}_m}\cdot\frac{\partial}{\partial\mathbf{q}_m}\right]f_N.\tag{1.5}$$

We define the operators

$$L^0=-i\sum_m\frac{\partial H_0}{\partial\mathbf{p}_m}\cdot\frac{\partial}{\partial\mathbf{q}_m},\qquad \lambda\delta L=i\lambda\sum_m\left[\frac{\partial V}{\partial\mathbf{q}_m}\frac{\partial}{\partial\mathbf{p}_m}-\frac{\partial V}{\partial\mathbf{p}_m}\frac{\partial}{\partial\mathbf{q}_m}\right],\tag{1.6}$$

so that

$$i(\partial f_N/\partial t)=(L^0+\lambda\delta L)f_N.\tag{1.7}$$

We follow P-B and Fourier analyze f_N. We order this Fourier expansion according to the number of nonvanishing wave vectors and explicitly write all volume dependencies so that we may assume the coefficients do not depend on N or Ω separately in the asymptotic limit (1.1):

$$\begin{aligned}f_N=(8\pi^3\Omega)^{-N}\{&\rho_0+\Omega^{-1}\sum_{\mathbf{k}}{}'\sum_j\rho_{\mathbf{k}}{}^j\exp(i\mathbf{k}\cdot\mathbf{q}_j)\\&+\Omega^{-2}\sum_{\mathbf{k}_1}{}'\sum_{\mathbf{k}_2}{}'\sum_{j_1<j_2}\rho_{\mathbf{k}_1\mathbf{k}_2}{}^{j_1j_2}\\&\times\exp(i\mathbf{k}_1\cdot\mathbf{q}_{j_1}+i\mathbf{k}_2\cdot\mathbf{q}_{j_2})+\cdots\\&+\Omega^{-s}\sum_{\mathbf{k}_1}{}'\cdots\sum_{\mathbf{k}_s}{}'\sum_{j_1<j_2<\cdots<j_s}\rho_{\mathbf{k}_1\cdots\mathbf{k}_s}{}^{j_1\cdots j_s}\\&\times\exp(i\mathbf{k}_1\cdot\mathbf{q}_{j_1}+\cdots+i\mathbf{k}_s\cdot\mathbf{q}_{j_s})\end{aligned}\tag{1.8}$$

where the prime on the $\mathbf{k}$ sums signifies the omission of the value $\mathbf{k}=0$ and $\Omega^{-1}\sum_{\mathbf{k}}{}'\to\int'd\mathbf{k}$ in the limit $\Omega\to\infty$. For further details and many pertinent comments, see P-B. However, we point out here that ρ_0 is the momentum distribution function and the other coefficients are related to correlations of the particles.

In the present paper we depart from P-B most significantly in our means of calculating the time evolution of the ρ's. While P-B use a time iterative technique, we use a form of the resolvent technique given by Van Hove.[2] The compactness of this latter technique allows us to investigate higher-order contributions with relative ease. We are then able to classify these higher-order contributions showing that dissipative and cloud effects may, in the appropriate representation and for certain long time limits, be determined by diagonal matrix elements of operators consisting of products of perturbations separated by

[1] I. Prigogine and R. Balescu, Physica 25, 281, 302 (1959).

[2] L. Van Hove, Physica 21, 517, 901 (1955); 22, 343 (1956); 23, 441 (1957).

diagonal operators which may be singular. This is not to say that the evolution of the system has no nondiagonal determination which in some cases may yield long time effects, but only that, in the appropriate representation and for certain long time limits the diagonal matrix elements determine the persistent (long time) behavior. This distinction between diagonal and nondiagonal matrix elements of the product operators was already stated in the masterful papers of Van Hove.

In the last section we make some comments concerning the mechanism of relaxation.

ITERATION OF THE RESOLVENT

Let us write

$$f_N(t)=U(t)f_N(0), \tag{2.1}$$

so that

$$\begin{aligned}\rho_{\mathbf{k}_1\cdots\mathbf{k}_s}{}^{j_1\cdots j_s}=&\langle k_1\cdots k_s|\,U(t)\,|0\rangle\rho_0(0)+\cdots\\&+\sum_{\mathbf{k}_1'}\cdots\sum_{\mathbf{k}_r'}\sum_{j_1'<j_2'<\cdots<j_r'}\\&\times\langle\mathbf{k}_1\cdots\mathbf{k}_s|\,U(t)\,|\mathbf{k}_1'\cdots\mathbf{k}_r'\rangle\\&\times\rho_{\mathbf{k}_1\cdots\mathbf{k}_r'}{}^{j_1'\cdots j_r'}+\cdots,\end{aligned} \tag{2.2}$$

where

$$\begin{aligned}&\langle\mathbf{k}_1\cdots\mathbf{k}_s|\,U(t)\,|\mathbf{k}_1'\cdots\mathbf{k}_r'\rangle\\&=(8\pi^3\Omega)^{-N}\int(d\mathbf{q})^N\exp(-i\sum_{l=1}^{s}\mathbf{k}_l\cdot\mathbf{q}_{jl})U(t)\\&\times\exp(i\sum_{l=1}^{r}\mathbf{k}_l'\cdot\mathbf{q}_{j'l'}).\end{aligned} \tag{2.3}$$

From Eq. (1.7), the unitary operator U obeys the equation

$$i(dU/dt)=(L^0+\lambda\delta L)U \tag{2.4}$$

with the initial condition $U(0)=1$. Then following Hugenholtz,[3] we may write

$$U(t)=\frac{i}{2\pi}\oint dzR(z)e^{-izt}, \tag{2.5}$$

where the path of integration is any positive contour completely enclosing the real axis. Here

$$R(z)=(L^0+\lambda\delta L-z)^{-1}. \tag{2.6}$$

If we denote $(L^0-z)^{-1}$ by $D_z{}^{(0)}$, we may iterate (assuming here once and for all that all such iterations converge) to obtain

$$\begin{aligned}R(z)=&D_z{}^{(0)}-\lambda D_z{}^{(0)}\delta LD_z{}^{(0)}\\&+\lambda^2D_z{}^{(0)}\delta LD_z{}^{(0)}\delta LD_z{}^{(0)}-\cdots.\end{aligned} \tag{2.7}$$

The diagonal part of $R(z)$ is written as

$$\{R(z)\}_D=D_z{}^{(0)}+\lambda D_z{}^{(0)}\sum D_z{}^{(0)}, \tag{2.8}$$

where

$$\begin{aligned}\sum=\{&-\delta L+\lambda\delta LD_z{}^{(0)}\delta L\\&-\lambda^2\delta LD_z{}^{(0)}\delta LD_z{}^{(0)}\delta L+\cdots\}_D.\end{aligned} \tag{2.9}$$

Let us define the diagonal operator

$$\begin{aligned}G_z=\{&-\delta L+\lambda\delta LD_z{}^{(0)}\delta L\\&-\lambda^2\delta LD_z{}^{(0)}\delta LD_z{}^{(0)}\delta L+\cdots\}_{SD},\end{aligned} \tag{2.10}$$

where SD means simply diagonal in the sense there are no intermediate states equal to the final (or initial) state. Terms which have intermediate states equal to initial (or final) state do arise in $\sum$. It is clear that the sum of all possible diagonal terms, with or without intermediate states equal to the initial or final states, is

$$\begin{aligned}\sum&=G_z+\lambda G_zD_z{}^{(0)}G_z+\lambda^2G_zD_z{}^{(0)}G_zD_z{}^{(0)}G_z+\cdots\\&=G_z[1-\lambda G_zD_z{}^{(0)}]^{-1}=G_zD_z(D_z{}^{(0)})^{-1},\end{aligned} \tag{2.11}$$

where

$$D_z=[L^0-z-\lambda G_z]^{-1}. \tag{2.12}$$

It should be noted that $\sum$ includes all terms which are made up of products of diagonal fragments with their initial and final states equal to the initial (and final) state of the complete $\{R(z)\}_D$. Since our interest will be in an expansion of $R(z)$ in powers of λ, we shall not consider separately those diagonal fragments formed by two equal intermediate states which are *not* equal to the initial (or final) state of the complete $\{R(z)\}_D$.

We want the nondiagonal contribution to $R(z)$ to be expressed explicity in simply nondiagonal terms, i.e., no intermediate state should equal either the initial or final state of the complete nondiagonal $\{R(z)\}_{ND}$. To do this, we regroup terms so that all possible diagonal fragments connecting either the initial or final states with an intermediate state are included in diagonal operators separating perturbative terms in the product. It follows then that the total resolvent is

$$\begin{aligned}R(z)=D_z-\lambda D_z[&\delta L-\lambda\delta LD_z\delta L\\&+\lambda^2\delta LD_z\delta LD_z\delta L-\cdots]_{SND}D_z,\end{aligned} \tag{2.13}$$

where SND means simply nondiagonal and the contributions from all the diagonal fragments connecting an intermediate state with either the initial or final state are included in the D_z.[4]

The role of G_z in causing either cloud or dissipative effects is much as described by Van Hove. In the Van Hove formulation, cloud effects corresponding to G_z real when z approaches the real axis are easily treated since they act only to shift residues (eigenvalue renormalization). The integral over z to obtain the diagonal part of $U(t)$ follows immediately from the Cauchy integral formula. When G_z is complex or

[3] N. M. Hugenholtz, *The Many Body Problem* (John Wiley & Sons, Inc., New York, 1959), p. 1 ff.

[4] The expression for the simply nondiagonal part of the resolvent given in Eq. (2.13) counts certain terms more than once. For most potentials of interest, however, the error introduced is negligible.

imaginary as z approaches the real axis, there are no poles apparent and the residues on the real axis in the unperturbed problem do not occur. This is the dissipative case.

To treat this latter case we follow Van Hove[5] and note that $G_{z^*}=G_z^\dagger$ since L is Hermitian. Thus if we set $z=\zeta+i\eta$, we may write for G_z close to the real axis

$$G(\zeta\pm i0)=K(\zeta)\pm iJ(\zeta) \tag{2.14}$$

with $K(\zeta)$ and $J(\zeta)$ real and the notation $\pm i0$ means $\pm i|\eta|$ when $\eta\to 0$. We see then that G_z is discontinuous across the real axis whenever $J(\zeta)$ is nonvanishing. We write

$$D_\alpha(\zeta\pm i0)=[L_\alpha{}^0-\zeta\mp i0-\lambda K_\alpha(\zeta)\mp i\lambda J_\alpha(\zeta)]^{-1}, \tag{2.15}$$

where

$$D_\alpha=\langle\alpha|D|\alpha\rangle. \tag{2.16}$$

We note that $D_\alpha(\zeta\pm i0)$ would have a pole when

$$L_\alpha{}^0-\zeta-\lambda K_\alpha(\zeta)=0 \tag{2.17}$$

if $J_\alpha(\zeta)$ vanished. If we denote by l_α the value of ζ which satisfies Eq. (2.17), we may write for $D_\alpha(\zeta\pm i0)$ in the neighborhood of l_α

$$D_\alpha(\zeta\pm i0)=N_\alpha(\pm)[l_\alpha-\zeta\mp i0\mp\lambda iN_\alpha(\pm)J_\alpha(l_\alpha)]^{-1}, \tag{2.18}$$

where

$$N_\alpha(\pm)=[1+\lambda K_\alpha{}'(l_\alpha)\pm\lambda iJ_\alpha{}'(l_\alpha)+\cdots]^{-1}. \tag{2.19}$$

For ζ far from l_α we may also approximate $D_\alpha(\zeta\pm i0)$ by Eq. (2.18) since the difference $l_\alpha-\zeta$ then dominates its behavior if we assume that $\lambda K_\alpha{}'(l_\alpha)$ and $\lambda J_\alpha{}'(l_\alpha)$ are at most of order λ. Then to a given order in an iteration procedure we may represent $D_\alpha(\zeta\pm i0)$ by its form near its maximal point, i.e., at the point where (2.17) holds.

We may now write, using Eq. (2.5) and setting $x\pm=\zeta\pm i0$,

$$U_\alpha(t)=\frac{i}{2\pi}\int_{+\infty}^{-\infty}\frac{N_\alpha(+)e^{-ix_+t}dx_+}{l_\alpha-x_+-i\lambda N_\alpha(+)J_\alpha(l_\alpha)}+\frac{i}{2\pi}\int_{-\infty}^{+\infty}\frac{N_\alpha(-)e^{-ix_-t}dx_-}{l_\alpha-x_-+i\lambda N_\alpha(-)J_\alpha(l_\alpha)}. \tag{2.20}$$

If we make use of a formula given by Lighthill,[6] we obtain for positive time,

$$U_\alpha(t)=N_\alpha(\pm)\exp(-itl_\alpha)\times\exp[-\lambda|N_\alpha(\pm)J_\alpha(l_\alpha)|t], \tag{2.21}$$

where $N(+)$ applies when $J_\alpha(l_\alpha)$ is negative and $N(-)$ applies when $J_\alpha(l_\alpha)$ is positive.

[5] See also L. Van Hove, "Lecture notes," University of Washington, Seattle, Washington, 1958.

[6] M. J. Lighthill, *Introduction to Fourier Analysis and Generalized Functions* (Cambridge University Press, New York, 1958), p. 44.

We now must see how this formalism works. What we do is evaluate $G(\zeta\pm i0)$ to some order in the iteration procedure. We are thereby able to identify $K(\zeta)$ and $J(\zeta)$, which we then use as in the foregoing general formulation. We consider then the problem so well treated by P-B; i.e., the classical homogeneous weakly coupled gas.

CLASSICAL WEAKLY COUPLED HOMOGENEOUS GAS

We take over the P-B formulation and notation for this problem. Thus

$$H=\sum_j(p_j^2/2m)+\lambda\sum_{j<n}V_{jn}(|\mathbf{q}_j-\mathbf{q}_n|), \tag{3.1}$$

where λ measures the strength of interaction, λV_{jn} is the interaction potential between particles j and n which is assumed to depend only on the relative distance between the jth and nth particles, and we shall take mass equal unity for convenience. We expand V_{jn} in a Fourier series

$$V_{jn}=\Omega^{-1}\sum_{\mathbf{l}}V_l\exp[i\mathbf{l}\cdot(\mathbf{q}_j-\mathbf{q}_n)], \tag{3.2}$$

where V_l depends only on the absolute value of $\mathbf{l}$.

P-B specify the initial conditions such that the Fourier coefficients initially have the same order of magnitude as in equilibrium

$$\begin{aligned}&\rho_0(0)\sim 1;\ \rho_{\mathbf{k}_1\mathbf{k}_2}{}^{\alpha\beta}\sim\lambda \quad \text{for} \quad \mathbf{k}_1=-\mathbf{k}_2;\\ &\rho_{\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3}{}^{\alpha\beta\gamma}\sim\lambda^2 \quad \text{for} \quad \mathbf{k}_1+\mathbf{k}_2+\mathbf{k}_3=0, \text{ etc.}\\ &\rho_{\mathbf{k}_1\mathbf{k}_2\cdots}{}^{\alpha\beta\cdots}=0 \quad \text{for} \quad \mathbf{k}_1+\mathbf{k}_2+\cdots\neq 0.\end{aligned} \tag{3.3}$$

The vanishing of the $\rho_{\mathbf{k}_1\mathbf{k}_2\cdots}{}^{\alpha\beta\cdots}$ for $\mathbf{k}_1+\mathbf{k}_2+\cdots\neq 0$ is a manifestation of the fact that δL only couples the completely homogeneous states (for which all wave vectors vanish) to quasi-homogeneous states where the sum of the wave vectors present vanishes.

Let us now consider the time evolution of the ρ's to lowest order. We note that δL derived from the potential V_{jn}, Eq. (3.2), has vanishing diagonal matrix elements in the $\mathbf{k}-$ representation. Thus, to lowest order,

$$G(\zeta\pm i0;\lambda)=[\lambda\delta L(L^0-\zeta\mp i0)^{-1}\delta L]_{SD}, \tag{3.4}$$

where the order in λ is explicitly indicated. For ρ_0 we are concerned with

$$\begin{aligned}&\langle 0|\delta L(L^0-\zeta\mp i0)^{-1}\delta L|0\rangle\\ &=\Omega^{-1}\sum_{j<n}{\int}'d\mathbf{l}|V_l|^2\mathbf{l}\cdot\mathbf{D}_{jn}\\ &\times\left[\pm i\pi\delta(\mathbf{l}\cdot\mathbf{g}_{jn}-\zeta)+P\left(\frac{1}{\mathbf{l}\cdot\mathbf{g}_{jn}-\zeta}\right)\right]\mathbf{l}\cdot\mathbf{D}_{nj},\end{aligned} \tag{3.5}$$

where P just means principal value, $\mathbf{D}_{jn}=[(\partial/\partial\mathbf{p}_j)+(\partial/\partial\mathbf{p}_n)]$, and $\mathbf{g}_{jn}=\mathbf{p}_j-\mathbf{p}_n$. Equation (3.5) corresponds

to the cycle (or λ^2 chain) diagram in Fig. 1(a). Then

$$K_0(\zeta;\lambda)=\lambda\Omega^{-1}\sum_{j<n}\int' d\mathbf{l}\,|V_l|^2\mathbf{l}\cdot\mathbf{D}_{jn}\times P\left(\frac{1}{\mathbf{l}\cdot\mathbf{g}_{jn}-\zeta}\right)\mathbf{l}\cdot\mathbf{D}_{jn}\equiv\lambda M_2(\zeta) \tag{3.6}$$

$$J_0(\zeta;\lambda)=\lambda\Omega^{-1}\sum_{j<n}\int' d\mathbf{l}\,|V_l|^2\mathbf{l}\cdot\mathbf{D}_{jn}\pi\delta\times(\mathbf{l}\cdot\mathbf{g}_{jn}-\zeta)\mathbf{l}\cdot\mathbf{D}_{jn}\equiv\lambda O_2(\zeta). \tag{3.7}$$

We assume the derivatives of $K_0(\zeta;\lambda)$ and $J_0(\zeta;\lambda)$ are of the same order of magnitude as $K_0(\zeta;\lambda)$ and $J_0(\zeta;\lambda)$ themselves. This means that we assume K_0 and J_0 are not only small, but also slowly varying. Now iterating Eq. (2.17) we find that $l_0=0$ is the maximal point for $D_0(\zeta\pm i0;\lambda^2)$. Thus, for the moment discounting any nondiagonal contribution and keeping only terms in λ^2, we find for ρ_0, using Eq. (2.2),

$$\rho_0(\lambda^2 t)\simeq[1+\lambda^2M_2'(0)\pm\lambda^2O_2'(0)]\times\exp[-\lambda^2tO_2(0)]\rho_0(0). \tag{3.8}$$

In the approximation $\lambda^2\to 0$, $t\to\infty$ such that $\lambda^2 t\to$ finite constant, which we call the $\lambda^2 t$ limit (see also Van Hove), Eq. (3.8) reduces to the P-B result

$$[\rho_0(\lambda^2t)]_{\text{P-B}}=\exp[-\lambda^2tO_2(0)]\rho_0(0). \tag{3.9}$$

We note that the compactness of the resolvent formalism allows us to consider only one diagram to obtain Eq. (3.9) rather than the very large number needed for all the $(\lambda^2 t)^n$ contributions necessary in the P-B iterative procedure.

Let us carry our results to next highest order. We are then interested in

$$G(\zeta\pm i0;\lambda^2)=[\lambda\delta L(L^0-\zeta\mp i0)^{-1}\delta L]_{SD}-[\lambda^2\delta L(L^0-\zeta\mp i0)^{-1}\delta L\times(L^0-\zeta\mp i0)^{-1}\delta L]_{SD}. \tag{3.10}$$

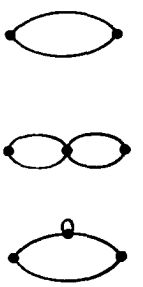


FIG. 1. (a) The cycle or λ^2 chain diagram; (b) λ^3 chain diagram; (c) λ^3 ring diagram.

The first term on the right-hand side of Eq. (3.10) was treated previously. For ρ_0, there are two types of diagrams that correspond to the second term. These are the λ^3 chain, $C_0(\lambda^2)$, shown in Fig. 1(b) and the λ^3 ring, $R_0(\lambda^2)$, shown in Fig. 1(c), so that

$$\lambda^2\langle 0|\delta L(L^0-\zeta\mp i0)^{-1}\delta L(L^0-\zeta\mp i0)^{-1}\delta L|0\rangle=C_0(\zeta\pm i0;\lambda^2)+R_0(\zeta\pm i0;\lambda^2). \tag{3.11}$$

Then

$$C_0(\zeta\pm i0;\lambda^2)=\lambda^2\Omega^{-1}\int' d\mathbf{l}\int' d\mathbf{l}'\sum_{j<n}\times V(l)V(|\mathbf{l}-\mathbf{l}'|)V(l')\mathbf{l}\cdot\mathbf{D}_{jn}\times\{(\mathbf{l}\cdot\mathbf{g}_{jn}-\zeta\mp i0)^{-1}(\mathbf{l}-\mathbf{l}')\cdot\mathbf{D}_{jn}\times(\mathbf{l}'\cdot\mathbf{g}_{jn}-\zeta\mp i0)^{-1}\mathbf{l}'\cdot\mathbf{D}_{jn}\} \tag{3.12}$$

$$R_0(\zeta\pm i0;\lambda^2)=\lambda^2\Omega^{-2}\int' d\mathbf{l}\sum_{j<n}\sum_{j'\neq n}|V(l)|^3\mathbf{l}\cdot\mathbf{D}_{nj'}\times\{(\mathbf{l}\cdot\mathbf{g}_{nj'}-\zeta\mp i0)^{-1}\mathbf{l}\cdot\mathbf{D}_{j'j}\times(\mathbf{l}\cdot\mathbf{g}_{j'j}-\zeta\mp i0)^{-1}\mathbf{l}\cdot\mathbf{D}_{jn}\}. \tag{3.13}$$

For $\zeta=0$ only terms made up of products of a δ function and a principal value make nonvanishing contributions. Thus the iteration of Eq. (2.17) again yields $l_0=0$ for the maximal point of $\langle 0|D(\zeta\pm i0;\lambda^3)|0\rangle$, and we have

$$C_0(0\pm i0;\lambda^2)=\pm i\pi\lambda^2\Omega^{-1}\int' d\mathbf{l}\int' d\mathbf{l}'\sum_{j<n}V(l)V(|\mathbf{l}-\mathbf{l}'|)V(l')\left\{\mathbf{l}\cdot\mathbf{D}_{jn}\delta(\mathbf{l}\cdot\mathbf{g}_{jn})(\mathbf{l}-\mathbf{l}')\cdot\mathbf{D}_{jn}P\left(\frac{1}{\mathbf{l}'\cdot\mathbf{g}_{jn}}\right)\mathbf{l}'\cdot\mathbf{D}_{jn}\right.$$
$$\left.+\mathbf{l}\cdot\mathbf{D}_{jn}P\left(\frac{1}{\mathbf{l}\cdot\mathbf{g}_{jn}}\right)(\mathbf{l}-\mathbf{l}')\cdot\mathbf{D}_{jn}\delta(\mathbf{l}'\cdot\mathbf{g}_{jn})\mathbf{l}'\cdot\mathbf{D}_{jn}\right\}; \tag{3.14}$$

$$R_0(0\pm i0;\lambda^2)=\pm i\pi\lambda^2\Omega^{-2}\int' d\mathbf{l}\sum_{j<n}\sum_{j'\neq n}|V(l)|^3\mathbf{l}\cdot\mathbf{D}_{nj'}\left\{\delta(\mathbf{l}\cdot\mathbf{g}_{nj'})\mathbf{l}\cdot\mathbf{D}_{j'j}P\left(\frac{1}{\mathbf{l}\cdot\mathbf{g}_{j'j}}\right)\mathbf{l}\cdot\mathbf{D}_{jn}\right.$$
$$\left.+P\left(\frac{1}{\mathbf{l}\cdot\mathbf{g}_{nj'}}\right)\mathbf{l}\cdot\mathbf{D}_{j'j}\delta(\mathbf{l}\cdot\mathbf{g}_{j'j})\mathbf{l}\cdot\mathbf{D}_{jn}\right\}. \tag{3.15}$$

By defining the operators

$$O_{3C}(0)=(i/\lambda^2)C_0(0+i0;\lambda^2)$$
$$O_{3R}(0)=(i/\lambda^2)R_0(0+i0;\lambda^2), \tag{3.16}$$

and again discounting any nondiagonal contributions, we have for ρ_0 in the $\lambda^2 t$ limit corrected to first order in λ

$$\rho_0(\lambda^3t)\simeq\exp[-(\lambda^2O_2+\lambda^3O_{3C}+\lambda^3O_{3R})t]. \tag{3.17}$$

In contrast to the corresponding P-B result which

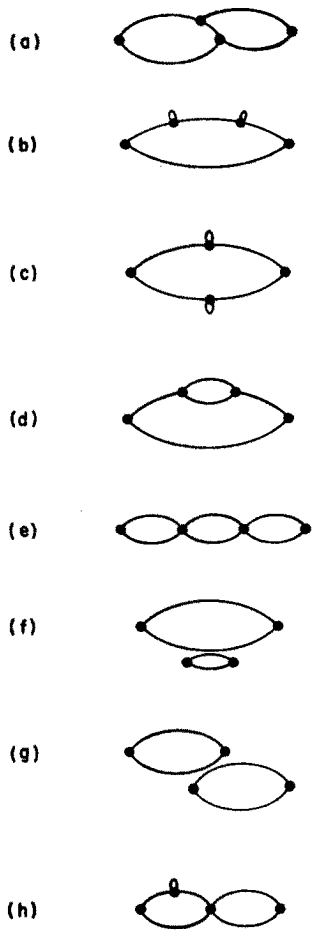


FIG. 2. Diagrams for the λ^4 simply diagonal contributions to $\langle 0|G(\zeta \pm i0)|0\rangle$.

includes only $\lambda(\lambda^2 t)^n$ terms, Eq. (3.17) includes contributions from all possible diagrams which are made up of cycles, λ^3 chains, and λ^3 rings. In the time iterative procedure, this means we would have certain terms of the form $\lambda^j(\lambda^2 t)^n$. The resolvent technique enables us to pick out and classify those $\lambda^j(\lambda^2 t)^n$ terms arising from diagonal fragments in an orderly way.

Let us consider the diagonal contributions for ρ_0 to next highest order. In this case, things get much more complicated but there is no difficulty in principle. Aside from lower-order terms already contributing to the exponential part as in Eq. (3.16), we have the λ^4 *SD* terms (see their diagrams in Fig. 2). Now, however, the $\lambda^2 t$ limit is corrected to order λ^2 so that we may no longer neglect the contributions from the derivatives of G in the long time limit.

Let us turn now to the nondiagonal resolvent contributions. The first such contribution to the evolution of $\rho_0(t)$ is

$$\rho_0^{nd}(t)=(\lambda i/2\pi)\oint dz e^{-izt}\langle 0|D_z|0\rangle\Omega^{-2}\int' d\mathbf{l}\sum_{j<n}\times\{V_{\mathbf{l}}\mathbf{l}\cdot\mathbf{D}_{jn}\langle -\mathbf{l}\mathbf{l}|D_z|-\mathbf{l}\mathbf{l}\rangle\rho_{-1\mathbf{l}}{}^{jn}(0)\}. \quad (3.18)$$

We define a quantity Q by using Eq. (3.18) and setting

$$\rho_0^{nd}(t)=(i/2\pi)\oint dz e^{-izt}\langle 0|D_z|0\rangle[\lambda^2 Q\rho_0(0)], \quad (3.19)$$

where the extra factor of λ arises because of the initial conditions, Eq. (3.3). With the use of Eq. (3.16), we may write[7]

$$\rho_0(t)\simeq\exp[-(\lambda^2 O_2+\lambda^3 O_{3C}+\lambda^3 O_{3R})t]\times[1+\lambda^2 Q]\rho_0(0). \quad (3.20)$$

We note that in the $\lambda^2 t$ limit, even corrected by $\lambda^3 t$ terms, the nondiagonal term makes no contribution. However, in an iterative expansion, the nondiagonal term does provide terms of the form $\lambda^2(\lambda^2 t)^n$, etc., and the nature of its contribution is not clearly separated from similar, more important diagonal terms. Also, Eq. (3.19) points out the importance of the initial conditions in determining the long time behavior. Thus, if the nondiagonal contribution is of the order λ instead of λ^2, we would not be consistent in keeping $\lambda^3 t$ terms in the exponential. Another important point is the dependence on representation. Suppose to obtain an adequate description of the relaxation we had to keep many higher-order terms in the exponential. Then for practically any reasonable initial conditions, the nondiagonal terms would contribute to persistent (long time) behavior. This, however, may be remedied by transforming to an appropriate representation. Consider the following transformation:

$$e^{-S}Re^{+S}=\{R\}_D+\{R\}_{ND}+[\{R\}_D,S]+[\{R\}_{ND},S]+\tfrac{1}{2}([\{R\}_D,S],S)+\cdots. \quad (3.21)$$

If we choose S such that

$$\{R\}_{ND}+[\{R\}_D,S]=0,$$

we see that the original nondiagonal terms make a contribution to higher-order diagonal terms and the new nondiagonal terms are of a higher order than the original nondiagonal terms and therefore more negligible. We may continue to do this until only diagonal terms (in an appropriate representation) determine the persistent behavior so that an equation such as Eq. (3.17) accurately describes the long time relaxation of a transformed ρ_0. However, the nondiagonal terms clearly make a contribution to the evolution of this transformed ρ_0 for shorter times. This contribution may be reversible or irreversible depending on the nature of the Q's.

MECHANISM OF RELAXATION

It is not immediately apparent that a function such as ρ_0 should obey an irreversible equation since the Liouville equation from which such an equation is

[7] Actually, to this order in λ, we should also include λ^2 terms which arise in $N(\pm)$ [see Eq. (2.21)].

derived is reversible. However, equations such as (3.17) are indeed irreversible and it is of interest to attempt to find the origin of this irreversible behavior.

From Eq. (3.17), it is apparent that for dissipative effects to occur in the $\lambda^2 t$ limit, it is of primary importance that the imaginary parts of the $G(\zeta \pm i0)$'s are nonvanishing. The nonvanishing of the imaginary parts of the appropriate $G(\zeta \pm i0)$'s is related to the choice of representation and in our example to the possibility of degeneracy of the eigenvalues of the unperturbed operator. This, in turn, is related to the need for a large system with a continuous spectrum. It was Van Hove who first emphasized the importance of the properties of the perturbation when he noted that persistent (long time) effects are determined by the diagonal matrix elements of operators composed of products of perturbation operators separated by diagonal operators which may have singularities.

ACKNOWLEDGMENTS

We thank Professor A. A. Maradudin of this laboratory and the University of Maryland for several interesting discussions. One of us (S.T.) also thanks Dr. J. Philippot for some interesting discussions and acknowledges the stimulating effect of a series of lectures given by Professor I. Prigogine at the National Bureau of Standards in December, 1959.

Differential-Operator Approximations to the Linear Boltzmann Equation*

ARMAND SIEGEL† ‡
University of Michigan, Ann Arbor, Michigan, and Boston University, Boston, Massachusetts
(Received April 19, 1960)

A measure of deviation from equilibrium of an ensemble of particles is proposed, which is physically appropriate and of especially simple form when expressed in terms of the expansion coefficients of the ensemble distribution function with respect to the system of orthogonal polynomials obtained by using the equilibrium distribution function as weight function. The linear Boltzmann operator can then be expanded in a series of terms which, under certain circumstances, may be regarded as of successively diminishing magnitude in their effect on the rate of approach to equilibrium. This expansion of the operator is different from the expansion due to Kramers (later discussed by Moyal) in derivate moments, commonly used in approximate stochastic treatments of irreversible processes. With the aid of a theorem on definite operators, it is possible to break off the series at any point and thereby obtain a correspondingly accurate approximation to the linear Boltzmann operator, whose temporal solutions tend to the correct equilibrium distribution function. The first approximation is the Fokker-Planck operator, exactly. The next approximation would be the appropriate operator to use when the stochastic variable begins to deviate appreciably from a linear dissipation law, etc. The method is applied to the "Rayleigh process" (ensemble of particles in a rarefied gas medium, the medium itself being in internal equilibrium), and the second approximation to the linear Boltzmann operator for this case is explicitly derived. A possible form for the second approximation in more general processes, suggested by this, is also given.

I. INTRODUCTION AND BASIC THEOREMS

1. Introduction

THE central theoretical tools in the study of time-varying thermal fluctuations have long been the Fokker-Planck equation and its *alter ego* the Langevin equation.[1] The use of these powerful mathematical devices has conferred on the subject a considerable degree of logical cohesion, but they limit its scope to phenomena obeying a linear friction, or dissipation, law (the terminology is defined in footnote 1). From the point of view of experiment this limitation is of no consequence at present, because there are as yet no temporal observations outside the linear friction range. But from the point of view of theory, the extension of our understanding to the nonlinear range appears desirable, because the dominating purpose of theory in this field is to bridge the gap between the fundamental theoretical postulates of kinetic theory and the phenomenological formalism, namely, thermodynamics (*sensu* antonym of thermostatics). The Fokker-Planck-Langevin formalism does make contact with thermodynamics[2,3] on the one side of this gap. But the fundamental theory is certainly nonlinear, hence "nonlinearization" of the Fokker-Planck-Langevin formalism is necessary to further the linkage.

A number of papers and reports on this subject have appeared in recent years.[4–9] Some of these make more or less tentative assumptions regarding the fundamental statistical equations governing the nonlinear systems, and go on to obtain solutions of these equations. Others emphasize only the problem of deriving and justifying the fundamental statistical equations. The present paper is addressed to this latter problem.

The thinking that underlies the present work is as follows: The Fokker-Planck equation may be rigorously derived in the case of the random walk in velocity space.[10] The random walk, as a random impact process, may be regarded as a simplified version of the *Rayleigh process*.[11] The Rayleigh process, which is defined and

* This work supported in part by the U. S. Air Force through the Air Force Office of Scientific Research, Air Research and Development Command.

† John Simon Guggenheim Memorial Foundation Fellow, 1957–58.

‡ Present address: Boston University.

[1] The Fokker-Planck equation for the temporal evolution of the probability density function $P(\xi,t)$ of a scalar variable ξ reads

$$\frac{\partial P(\xi,t)}{\partial t}=-\frac{\partial}{\partial \xi}a\xi P(\xi,t)+\frac{b}{2}\frac{\partial^2}{\partial \xi^2}P(\xi,t).$$

Here $a\xi$ is (if the equation is applicable) minus the ensemble average rate of change of ξ due to "friction" or dissipative effects in general; i.e., $\langle\dot{\xi}\rangle=-a\xi$. For a particle undergoing Brownian motion, $\langle\dot{\xi}\rangle$ is literally due to friction, being attributable to viscosity; more generally, ξ may be any thermodynamic observable in its range of linear dissipation, according to the theories referred to in footnotes 2 and 3. The constant b is (again, if the equation is applicable) a measure of the amplitude of thermal fluctuations, or "noise."

In the mathematically equivalent Langevin formalism, $\xi(t)$ is a random function of time satisfying the Langevin equation

$$\dot{\xi}+a\xi=(b)^{\frac{1}{2}}\epsilon(t),$$

where $\epsilon(t)$ is the "ideal random function" normalized so that

$$\left\langle\left[\int_0^t\epsilon(t)dt\right]^2\right\rangle=t.$$

$\xi(t)$ will then be found to have a probability density satisfying the Fokker-Planck equation as just given.

Introductory treatments of these matters will be found in the well-known review articles by S. Chandrasekhar [Revs. Modern Phys. **15**, 1 (1943)] and by M. C. Wang and G. E. Uhlenbeck [Revs. Modern Phys. **17**, 323 (1945)].

[2] N. Hashitsume, Progr. Theoret. Phys. (Kyoto) **8**, 461 (1952).
[3] L. Onsager and S. Machlup, Phys. Rev. **91**, 1505 (1953).
[4] D. K. C. MacDonald, Phys. Rev. **108**, 541 (1957).
[5] N. G. van Kampen, Phys. Rev. **110**, 319 (1958).
[6] R. O. Davies, Physica **24**, 1055 (1958).
[7] C. T. J. Alkemade, Physica **24**, 1029 (1958).
[8] M. Lax, Revs. Modern Phys. **32**, 25 (1960).
[9] N. G. van Kampen (unpublished report, 1959).
[10] See S. Chandrasekhar or M. C. Wang and G. E. Uhlenbeck, cited in footnote 1.
[11] Lord Rayleigh, *Scientific Papers* (Cambridge University Press, New York, 1902), Vol. 3, p. 273; discussed by C. S. Wang Chang and G. E. Uhlenbeck, *Kinetic Theory of a Gas in Alternating Outside Force Fields*, Engineering Research Institute Report 2457-3-T (University of Michigan, Ann Arbor, Michigan, 1956).

discussed in Sec. II and Appendix A of this paper, is a process which is itself simple enough for easy, explicit mathematical description, yet real enough to embody some basic features of thermal fluctuation phenomena. In the simplified version referred to in the foregoing, the friction dependence is linear, but in the exact formulation it is definitely nonlinear. Thus it might be possible to derive from it *a* counterpart, if not *the* counterpart, of the Fokker-Planck equation for the nonlinear friction region.

The probability density of the random variable in a Rayleigh process obeys a linear Boltzmann equation, the operator of which contains an explicit expansion parameter. When the operator is appropriately expanded in terms of this parameter (this is done in Sec. II, where it is shown that this expansion is different from the customarily employed Kramers or Moyal expansion), the condition that the probability density tends to the known equilibrium form can be applied to the problem of approximating this series. The first approximation is, as it must be, the Fokker-Planck equation. The second approximation is a sixth-order differential operator of precisely defined form containing two independent physical parameters (in addition to that of the first approximation), and an arbitrary parameter which does not affect any experimentally measurable results. The method is, moreover, a general one and yields approximations of arbitrary order. Thus there is a regular sequence of approximations linking the Fokker-Planck and linear Boltzmann equations.

The paper is organized as follows: Since the result may have validity for processes other than the Rayleigh process, the subsections of Sec. I which follow this introductory section discuss the general case of a linear Boltzmann-operator expansion having the necessary properties, and state and prove the theorems for the construction of satisfactory approximations from this expansion. In Sec. II the Rayleigh process is described, certain necessary expressions are derived from it, and the general theorem is applied. In Sec. III certain generalizations suggested by the Rayleigh process analysis are discussed. The equation generalizing the Fokker-Planck equation to cubic friction is then given for a hypothetical process which is mathematically similar to the Rayleigh process but which does not possess an explicit expansion parameter, or for which the parameter is unknown.

2. Précis of Method for Construction of Successive Approximations to the Linear Boltzmann Operator

Consider the linear Boltzmann equation for an ensemble of particles moving in one dimension. Assume no force field, and that the particles have already attained a spatially uniform density. The distribution function will then depend only on velocity and time: We write $P(V,t)$ for the probability density of velocity V, normalized to unity, as a function of time. The linear Boltzmann equation will then be

$$\partial P/\partial t = BP, \tag{1}$$

where B is a linear integral operator. The function BP is given, more explicitly, in terms of a kernel $B(V,V')$ as

$$\int B(V,V')P(V')dV'. \tag{2}$$

The eigenvalues of B must all be negative, except for a nondegenerate zero eigenvalue which has the Maxwell-Boltzmann distribution function

$$F(V) = (2\pi V_R^2)^{-\frac{1}{2}} \exp\left(-\frac{V^2}{2V_R^2}\right) \tag{3}$$

(V_R = root-mean-square value of V in the equilibrium distribution) (4)

as its eigenfunction. This "equilibrium requirement" ensures that an arbitrary initial distribution will always decay into the Maxwell-Boltzmann distribution. It may be equivalently formulated by saying that B must be *negative semidefinite*, in the sense that

$$\int_{-\infty}^{\infty} Y(V)B(V,V')Y(V')F(V')dVdV' \leqq 0 \tag{5}$$

for any polynomial $Y(V)$, with the equality sign holding only for Y = constant.

In the following discussion we take B to be in a Hermitian matrix representation. The vector corresponding to $F(V)$ is then the (unique) eigenvector of B for eigenvalue zero. The negative semidefiniteness requirement will take the form

$$\sum_{r,s} u_r B_{rs} u_s \leqq 0 \tag{6}$$

for all normalizable vectors (u_m), with the equality holding when (u_m) corresponds to $F(V)$. (Details of such a representation will be given in Sec. I.3.)

Suppose B to depend on some parameter λ with respect to which it may be expanded in a convergent series:

$$B = c(\lambda) \sum_{m=0}^{\infty} \lambda^m b_m, \tag{7}$$

where $c(\lambda)$ is a positive c-number function of λ and b_m is a matrix independent of λ. The existence of such an expansion suggests the possibility of approximating B, for small values of λ, by terminating the series at some finite value of m. In so doing, however, it will be important to retain the negative semidefiniteness property in the approximate operator: Lack of this property will imply the existence of at least one eigenvector of B which grows, instead of decaying, expo-

nentially with time; if present in the initial distribution, in however small an admixture, this mode or modes will grow in amplitude indefinitely large with time, hence the Maxwell-Boltzmann distribution will never be reached.

It is not possible to prove from the negative semidefiniteness of B that an operator obtained by terminating the series (7) is negative semidefinite. In fact, we shall find that the model to be discussed below furnishes a counterexample to such a supposition. However, it may be possible to retain the negative semidefiniteness property by a simple construction. This is based on the following factorization theorem[12]: *A positive semidefinite Hermitian matrix can always be written as the product of some (suitably chosen) matrix Q and its adjoint $Q^\dagger$.* Let us put, for the sum in Eq. (7),

$$\mathbf{B}=\sum_{m=0}^{\infty}\lambda^m b_m. \tag{8}$$

Then, applying this theorem to $\mathbf{B}$, which is negative semidefinite like B,

$$\mathbf{B}=-QQ^\dagger. \tag{9}$$

[N.B.: Q is determined only to within a unitary postmultiple; if U is unitary, $QU(QU)^\dagger=QU\cdot U^{-1}Q^\dagger=QQ^\dagger$.] Suppose now that Q also can be expanded in terms of λ:

$$Q=\sum_{l=0}^{\infty}\lambda^l q_l. \tag{10}$$

For the operator obtained by taking the first m_0 terms of Q, write

$$Q^{(m_0)}=\sum_{l=0}^{m_0}\lambda^l q_l. \tag{11}$$

Then

$$\mathbf{B}=-Q^{(m_0)}Q^{(m_0)\dagger}+o(\lambda^{m_0}) \tag{12}$$

where $o(\lambda^{m_0})$ is an operator of order higher than λ^{m_0}; i.e., $\mathbf{B}$ may be approximated to order λ^{m_0} by $-Q^{(m_0)}Q^{(m_0)\dagger}$, a form analogous to the exact factorization $QQ^\dagger$.

We now assert that $-Q^{(m_0)}Q^{(m_0)\dagger}$ is a negative semidefinite matrix. That it is at least negative definite follows from its very form, since for any vector $\mathfrak{u}$, $\mathfrak{u}\cdot Q^{(m_0)}Q^{(m_0)\dagger}\mathfrak{u}=(Q^{(m_0)\dagger}\mathfrak{u})\cdot(Q^{(m_0)\dagger}\mathfrak{u})\geqq 0$. This being so, it will moreover be negative semidefinite in the sense desired if F [Eq. (3)] is an eigenvector, with eigenvalue zero. This can be shown to be true, as follows: $Q^\dagger F=0$ since by hypothesis [remark following Eq. (5)] $(F,QQ^\dagger F)=(Q^\dagger F,Q^\dagger F)=0$. But if $Q^\dagger$ annihilates F, and if the series of Eq. (10) represents Q over some nonzero range of λ, then the individual $q_l^\dagger$ must also annihilate F; since, putting $q_l^\dagger F=F_l$, we must have $\sum\lambda^l F_l=0$ over this range of λ, which means that the individual F_l must vanish; whence $Q^{(m_0)\dagger}$ annihilates F, and F is an eigenvector of $Q^{(m_0)}Q^{(m_0)\dagger}$ for eigenvalue zero, q.e.d.

We use a subscript m_0 to denote a negative semidefinite approximation to $\mathbf{B}$ of order λ^{m_0} as constructed in the foregoing, i.e.,

$$\mathbf{B}_{m_0}=-Q^{(m_0)}Q^{(m_0)\dagger}. \tag{13}$$

It should be noted that $\mathbf{B}_{m_0}$ could, unlike $\mathbf{B}$, be degenerate with respect to the zero eigenvalue, so far as the present proof goes. If this should happen, it would of course be quite unsuitable, since the time-asymptotic distribution, in contradiction to the H theorem, would not be uniquely the Maxwell-Boltzmann distribution, but would depend on the initial distribution. Until more is known, the success of the method sketched above in yielding an approximation suitable in this respect cannot be guaranteed in advance; individual cases to which it is applied will have to be inspected after the event for satisfaction of this criterion.

Construction of $Q^{(m_0)}$

If we substitute the series expressions for Q and for $\mathbf{B}$ into Eq. (9), and equate coefficients of like powers of λ, we obtain an infinite set of equations

$$-\tfrac{1}{2}\sum_{j=0}^{m}(q_j q_{m-j}^\dagger+q_{m-j}q_j^\dagger)=b_m;\quad m=0,1,2,\cdots. \tag{14}$$

The first m_0+1 equations of this set involve only the first $(m_0+1)q$'s, $q_0, q_1\cdots q_{m_0}$. Suppose we have a solution to these first m_0+1 equations, say $q_0', q_1', \cdots q_{m_0}'$; primes are used to allow for the possibility that these may not agree with the first m_0 terms of Q itself (i.e., of the solution to the infinite set of equations), even after allowing for the possibility of an arbitrary common unitary postmultiple. From these we can construct an operator

$$\mathbf{B}_{m_0}'=-Q'^{(m_0)}Q'^{(m_0)\dagger}, \tag{15}$$

where

$$Q'^{(m_0)}=\sum_{l=0}^{m_0}\lambda^l q_l'. \tag{16}$$

$\mathbf{B}_{m_0}'$, like $\mathbf{B}_{m_0}$, agrees with $\mathbf{B}$ to order λ^{m_0}. In this way, if Eqs. (14) can be solved for $m=0, 1, 2\cdots m_0$, we have successive approximations up to m_0th order in λ to $\mathbf{B}$.

3. Estimation of the Degree of Deviation from Equilibrium

The essential nature of B is to drive its operand $P(V,t)$ toward the equilibrium function $F(V)$. Thus when it is expanded, the increasing smallness of its successive terms should be with respect to their effectiveness in this sense. With this in mind, we adopt the

[12] F. D. Murnaghan, *The Theory of Group Representations* (The Johns Hopkins Press, Baltimore, 1938), p. 20.

following as a measure of the degree of deviation of $P(V,t)$ from the equilibrium distribution $F(V)$:

$$\chi_0^2=\int\frac{[P(V,t)-F(V)]^2}{F(V)}dV. \tag{17}$$

It will be noted that this is the same, to within a constant factor, as Pearson's noted χ^2 of statistical theory,[13] with $F(V)$ the hypothetical and $P(V,t)$ the sampling distribution, and with an infinitely fine subdivision of the range of V. However, since there are a number of possible measures of "goodness of fit" of a distribution, mere coincidence with one of these which happens to be famous is not sufficient reason for its adoption; it is necessary to demonstrate the suitability of the choice (17). Our reasons for adopting it are the following:

(1) It emphasizes deviations from equilibrium according to the magnitude of the V values involved: Since $F(V)$ will be essentially localized in the region of equilibrium values of V, it is increasingly small for increasingly large deviations of V from its equilibrium range, and with $F(V)$ in the denominator of the integrand of χ_0^2 these large deviations are increasingly heavily weighted. This is appropriate in a study of the approach to equilibrium, because a given amount of probability added to or taken away from the equilibrium distribution curve in the neighborhood of some V value becomes increasingly important in its effects with increasing deviation from the rms value of V.

(2) It is precisely adapted to formulation in terms of a matrix representation, and therefore to the utilization of the theorem of Sec. I.2: given a set of polynomials $p_r(V)$ orthogonalized with respect to $F(V)$ as weight function (in particular, the Hermite polynomials),

$$\int p_r(V)p_s(V)F(V)dV=(N_r/N_0)\delta_{rs}, \tag{18}$$

where N_r is a normalization constant. Expand $P(V,t)$ in terms of normalized functions $(N_0/N_r)^{\frac{1}{2}}p_r(V)F(V)$:

$$P(V,t)=N_0^{\frac{1}{2}}\sum_{r=0}^{\infty}a_r(t)N_r^{-\frac{1}{2}}P_r(V)F(V). \tag{19}$$

Then

$$\chi_0^2=N_0\int[\sum{}'a_r(t)N_r^{-\frac{1}{2}}P_r(V)]^2F(V)dV=\sum{}'a_r^2, \tag{20}$$

where the prime on the summation sign denotes omission of the term $r=0$. Noting that the a_r are the components of the Hilbert-space vector $P(V,t)$, we see that χ_0^2 is now represented by the squared length of the part of $P(V,t)$ orthogonal to the equilibrium function. Thus, when B is put into a matrix representation having the above as basis functions, its tendency to promote equilibrium will be measured in the simplest possible way—by its effect on the components of its vector operand.

We are now in a position to relate the approach to equilibrium to the series expansion of B. The matrix elements of B are

$$B_{rs}=N_0(N_rN_s)^{-\frac{1}{2}}\int\int p_r(V)B(V,V')p_s(V')dVdV'. \tag{21}$$

[Because of the detailed balancing condition, $B(V,V')F(V')=B(V',V)F(V)$, these are Hermitian.] If Eq. (1) is taken in matrix form, and B expanded according to Eq. (7), we find for the rate of decrease of χ_0^2,

$$\frac{d\chi_0^2}{dt}=2c(\lambda)\sum_{m=0}^{\infty}\lambda^m\sum_{r,s}{}'a_r(b_m)_{rs}a_s. \tag{22}$$

Thus if $\lambda\ll1$, the successive matrices $b_m(m=0, 1, 2, \cdots)$ make rapidly decreasing contributions to the trend to equilibrium, at least if the sums which are the coefficients of the λ^m do not increase markedly with m. Thus we have achieved the interpretation of the expansion (7) which was sought at the beginning of this section.

4. Summary of Method

Successive approximations to the Boltzmann operator with respect to its tendency to promote equilibrium may be constructed then along the following lines: (1) B must be put into a matrix representation, using as basis the orthogonal [with respect to $F(V)$ as weighting function] functions $p_r(V)$; (2) the matrix B is then expanded (if possible) in terms of a small parameter λ; (3) the procedure of Sec. I.2 is then used to obtain a negative semidefinite approximation of the desired order in λ; (4) the coefficients of successive powers of λ in the expansion of $d\chi_0^2/dt$ must not increase too rapidly.

We do not discuss requirement (4) in the present paper, but merely apply the method outlined assuming it to be satisfied. We shall show that the zeroth approximation thus obtained is exactly the Fokker-Planck equation, which is evidence that this procedure is justified. Before leaving this subject, however, it might be mentioned that from the general behavior of the successive matrices b_m in the special case to be studied in Sec. II of this paper, it appears [see Eq. (76)] that they satisfy requirement (4) to the extent that $P(V,t)$ approximates to the equilibrium function; whence higher terms of the expansion become important, despite the decreasing values of λ^m, with highly disequilibrated ensembles.

[13] See, e.g., H. Cramér, *Mathematical Methods of Statistics* (Princeton University Press, Princeton, New Jersey, 1945), Chap. 30.

5. Considerations Related to the H Theorem

Brinkman[14] has suggested the application, in the study of the linear Boltzmann process, of the criterion

$$\dot{F} \leqq 0, \tag{23}$$

where

$$F = U - TS, \tag{24}$$

U and S being the internal energy and entropy of the set of particles whose distribution function obeys the postulated linear Boltzmann equation, and T the temperature of the medium whose molecules generate the random motion of the particles. Since Brinkman did not consider the linear Boltzmann equation as such, it is of interest to see how his criterion fits into the present work.

First, to settle on a simple nomenclature, let us agree henceforth to restrict the term "particle" to the individuals whose random motion is being studied, and the term "molecule" to the constituent individuals of the medium. Assume both particles and molecules to be uniformly distributed in space. If $f(v,t)$ is the distribution function of molecules with respect to their velocity variate v, and $P(V,t)$ that of particles with respect to their velocity variate V, then the H function per unit volume of the combined systems is

$$H(t) = \int f(v,t) \log f(v,t) dv + \int P(V,t) \log P(V,t) dV. \tag{25}$$

The application of the *Stosszahlansatz* to collisions of *all* kinds will lead to the usual nonlinear coupled Boltzmann equations for the two distributions, and thereby to the H theorem:

$$\dot{H} = \int \dot{f} \log f dv + \int \dot{P} \log P dV \leqq 0. \tag{26}$$

Now let us specialize to the conditions under which the linear Boltzmann equation holds for the particles—to the "linear Boltzmann regime," as we shall call it. $f(v,t)$ is then negligibly different from the equilibrium Maxwell-Boltzmann distribution. This does not mean that $\dot{f}$ is necessarily to be neglected, but merely that its effect relative to f is. As for $\dot{f}$, it has a contribution due to collisions of molecules with one another, and one due to collisions with the particles. Under the linear Boltzmann regime, the former can be made arbitrarily small independently of the latter—for example, by making the medium sufficiently rarefied while compensating for this by increasing the size of the particles; thus we assume it to be negligible.

For the second term in (26) we shall have

$$\int \dot{P} \log P dV = -\dot{S}/k. \tag{27}$$

As for the first term, the f contribution due to collisions between molecules and particles cannot be neglected, unlike the other contribution, since these are the collisions responsible for the process itself. But since the molecules are in equilibrium,

$$\log f = -\tfrac{1}{2} m v^2/kT + \text{const}, \tag{28}$$

and

$$\int \dot{f} \log f dv = -\frac{\dot{U}(\text{medium})}{kT} = +\frac{\dot{U}}{kT}. \tag{29}$$

We will thus have from (26), with $H = -S/k$,

$$\dot{F} \leqq 0, \tag{30}$$

showing that Brinkman's criterion is necessarily satisfied under the linear Boltzmann regime, i.e., $\dot{F}$ will decrease when $P(V,t)$ satisfies the linear Boltzmann equation.

We shall return to this matter at the end of Sec. II.5.

II. APPLICATION TO THE RAYLEIGH MODEL

1. Linear Boltzmann Operator for the Rayleigh Model

As an illustration of the method described above, we apply it to a simplified case which is still interesting from the kinetic theory point of view; that of the random velocity of a particle suspended in a rarefied gas in internal equilibrium. In order to simplify the mathematical analysis, we introduce certain artifices which are more or less familiar in this classical problem. We study the random motion of *particles* (as defined in Sec. I.5) of mass M. Their random motion results from collisions with the *molecules* of the rarefied gas, of mass m. The molecules have uniform spatial density ρ and a Maxwell-Boltzmann distribution with respect to the molecular velocity $\mathbf{v}$,

$$(2\pi v_R^2)^{-\frac{3}{2}} \exp\left(-\frac{\mathbf{v}^2}{2v_R^2}\right), \tag{31}$$

where

$$v_R^2 = \text{rms value of any component of } \mathbf{v} = kT/m. \tag{32}$$

The mass ratio m/M will at a later stage be assumed less than one, but not necessarily very small. We assume the gas sufficiently rarefied, and the particle concentration low enough, for the initial Maxwell-Boltzmann distribution of the gas molecules not to change appreciably in time, no matter what the initial velocity distribution of the particles. The particle concentration is also to be so low that collisions of particles with one another occur with negligible frequency, i.e., the random velocity changes of the particles are entirely caused by collisions with the molecules.

For further simplicity we take the particles to be infinitely thin disks, each constrained to move only in a direction perpendicular to its plane. This constraint, while artificial, is of a purely passive nature and does not dynamically affect the spontaneous statistical fluc-

[14] H. C. Brinkman, Physica 23, 82 (1957).

tuations responsible for the random process studied; in particular, we shall obtain in lowest approximation exactly the classical Browian motion.

With the foregoing constraint, the distributed variable of the disk distribution is the velocity component along the line of the allowed motion, which we may call V. However, we shall use instead of V the dimensionless

$$y=V/V_R \tag{33}$$

[V_R is defined in Eq. (4); here $V_R=(kT/M)^{\frac{1}{2}}$].

The linear Boltzmann equation of the foregoing process, which is obeyed by $P(y,t)$, the probability density function in y, is

$$\frac{\partial P(y',t)}{\partial t}=\int B(y',y'')P(y'',t)dy'', \tag{34}$$

where the linear Boltzmann operator $B(y',y'')$ has the form

$$B(y',y'')=C(y'|y'')-\delta(y'-y'')\int C(y|y')dy. \tag{35}$$

It is shown in Appendix A that

$$C(y''|y')=V_R(\mu/2\pi)^{\frac{1}{2}}\left(\frac{1+\mu}{2\mu}\right)^2|y''-y'|$$
$$\times\exp\left\{-\frac{1}{8\mu}[(1+\mu)y''-(1-\mu)y']^2\right\}. \tag{36}$$

$C(y''|y')$ is the transition probability density-in-y'' for a particle having initial velocity y'.

2. Expansion in Kramers Series

The operator B can be expanded in Kramers series[15–19]

$$B=\sum_{n=1}^{\infty}\frac{1}{n!}\left(-\frac{\partial}{\partial y'}\right)^n\alpha_n(y'), \tag{37}$$

in terms of the "derivate moments" α_n (Moyal's[18] terminology)

$$\alpha_n(y')=\int_{-\infty}^{\infty}(y''-y')^nC(y''|y')dy''. \tag{38}$$

(B is now an infinite-order differential operator, not an integral operator.)

Some simplification results if $\alpha_n(y')$ is replaced by the function

$$A_n(y')=\frac{1}{V_R\mu^{\frac{1}{2}}n!}\left(\frac{1+\mu}{\mu}\right)^n\alpha_n(y') \tag{39}$$

and if in the integral for $\alpha_n(y')$ the transformation

$$x=\frac{1+u}{2\mu}(y''-y'),\quad y=y' \tag{40}$$

is introduced (x is then the new variable of integration, and y the new argument of α_n). On substituting (36) into the integral for α_n, we then find

$$A_n(y)=\frac{2^n}{(2\pi)^{\frac{1}{2}}n!}\int\exp\left[-\frac{\mu}{2}(x+y)^2\right]x^n|x|dx, \tag{41}$$

and the expansion (37) becomes

$$B=V_R\mu^{\frac{1}{2}}\sum_{n=1}^{\infty}\left(\frac{\mu}{1+\mu}\right)^n\left(-\frac{\partial}{\partial y}\right)^nA_n(y). \tag{42}$$

3. Discussion of the Kramers Expansion

The simple appearance of the Kramers expansion is somewhat misleading from the point of view of the study of the approach to equilibrium, since it is not clear whether or how the successive terms represent decreasing contributions to the equilibrium-seeking tendency. In fact, we shall find that the expansion of B according to the matrix method of Sec. I.3, and which is adapted to the especially simple criterion of deviation from equilibrium there introduced, *is quite different from the Kramers expansion.* Thus the Kramers expansion plays no fundamental role in this work; however, we shall find it useful in deriving actual expressions for matrices.

The foregoing assertion is foreshadowed by the result of the traditional "random walk" analysis of the velocity-space progress of a particle subject to successive independent random impacts, which is a valid approximation in the present kind of system in the limit of vanishingly small μ. But it is certainly not clear how one could, at least in any offhand way, apply this limiting process to Eq. (42) directly. The "random walk" derivation, to be sure, does use the Kramers expansion, but it uses limiting approximations for the derivate moments from the outset, and does not make explicit use of μ.

On the other hand, Wang Chang and Uhlenbeck[11] obtained the Fokker-Planck equation directly from the linear Boltzmann equation by combining with the assumption of vanishing mass ratio the assumption that the velocity variable of the particles never gets much larger than the rms value it would have in equilibrium. [In our case this would amount to assuming $P(y,t)$ negligible for y much larger than μ.] It can be shown that these combined requirements are equivalent to taking the first two terms only in the expansion (42), and simultaneously approximating A_1 and A_2 by their lowest-order terms in y, these being of first and zeroth

[15] H. A. Kramers, Physica **7**, 284 (1940).
[16] S. Chandrasekhar, Revs. Modern Phys. **15**, 1 (1943).
[17] M. C. Wang and G. E. Uhlenbeck, Revs. Modern Phys. **17**, 323 (1945).
[18] J. E. Moyal, J. Roy. Stat. Soc. (London) **B11**, 150 (1949).
[19] J. Keilson and J. E. Storer, Quart. Appl. Math. **10**, 243 (1952).

order, respectively[20]; and this same result will be seen to follow quite simply from our analysis. The work that will be presented here extends the procedure of Wang Chang and Uhlenbeck in that successive approximations, rather than a single limiting approximation, can be obtained.

3. Symmetrization of B; Basis Functions of the Matrix Representation

The kernel of Eq. (34) can be symmetrized by the transformation

$$\bar{B}(y',y'')=\exp\left(\frac{y'^2}{4}\right)B(y',y'')\exp\left(-\frac{y''^2}{4}\right); \quad (43)$$

with the accompanying transformation of the distribution function,

$$\bar{P}(y')=\exp\left(\frac{y'^2}{4}\right)P(y'), \quad (44)$$

the linear Boltzmann equation (34) is unchanged in form:

$$\frac{\partial\bar{P}(y',t)}{\partial t}=\int\bar{B}(y',y'')\bar{P}(y'',t)dy''. \quad (45)$$

Being symmetric, $\bar{B}$ is Hermitian with respect to an unweighted inner product, thus for any two functions $\varphi(y)$, $\psi(y)$, we henceforth define

$$(\varphi,\psi)=\int_{-\infty}^{\infty}\varphi^*(y)\psi(y)dy. \quad (46)$$

This definition of the inner product is of heuristic convenience for the geometrical interpretation, and facilitates the use of standard (at least to mathematical physicists) definitions of the Hermite functions, which we shall use extensively.

The corresponding transformation of B as a *differential* operator [Eq. (42)] is

$$\bar{B}=\exp\left(\frac{y^2}{4}\right)B\exp\left(-\frac{y^2}{4}\right)=V_R\mu^{\frac{1}{2}}\sum_{n=1}^{\infty}\left(\frac{\mu}{1+\mu}\right)^n \times\exp\left(\frac{y^2}{4}\right)\left(-\frac{\partial}{\partial y}\right)^n\exp\left(-\frac{y^2}{4}\right)A_n(y). \quad (47)$$

Since the equilibrium function of the linear Boltzmann operator of the process we are considering is $(2\pi)^{-\frac{1}{2}}\exp(-y^2/2)$, we use as the basis for our matrix representation, for the reasons given in Sec. I.3, the Hermite polynomials that are orthogonal in the sense of Eq. (18) with respect to this as weighting function. But with the definition of inner product we are now using, Eq. (46), the basis consists rather of the Hermite *functions*

$$h_r(y)=\bar{C}^r\exp\left(-\frac{y^2}{4}\right), \quad (48)$$

where $\bar{C}$ is the "creation operator"

$$\bar{C}=\exp\left(\frac{y^2}{4}\right)\left(-\frac{d}{dy}\right)\exp\left(-\frac{y^2}{4}\right). \quad (49)$$

These satisfy

$$(h_r,h_s)=N_r\delta_{rs}, \quad (50)$$

where

$$N_r=(2\pi)^{\frac{1}{2}}r! \quad (51)$$

The Hermite *polynomials* are

$$H_r(y)=\exp\left(\frac{y^2}{4}\right)h_r(y). \quad (52)$$

They are the counterparts of the polynomials $p_r(V)$ of Sec. I.3. If in that section we take N_0 (which is arbitrary at that stage)

$$N_0=(2\pi)^{\frac{1}{2}}, \quad (53)$$

let y correspond directly to V, and put

$$F(y)=(2\pi)^{-\frac{1}{2}}\exp(-y^2/2), \quad (54)$$

then N_0 times formula (18) corresponds exactly to (50).

The matrix elements of B as defined by (21) now become matrix elements of B with respect to the new definition of inner product:

$$B_{rs}(\text{sense of Sec. I.3})=\left(\frac{h_r}{N_r^{\frac{1}{2}}},B\frac{h_s}{N_s^{\frac{1}{2}}}\right) =\bar{B}_{rs}(\text{sense of this section}). \quad (55)$$

Henceforth matrix elements are to be understood as defined according to the *second* of the foregoing equalities.

For χ_0^2, Sec. I.3, we now have

$$\chi_0^2=N_0\int(\bar{P}-\bar{F})^2dy=N_0(\bar{P}-\bar{F},\bar{P}-\bar{F}). \quad (56)$$

4. Matrix Expansion of $\bar{B}$

$\bar{B}$ can now be written in terms of creation operators

$$\bar{B}=V_R\mu^{\frac{1}{2}}\sum_{n=1}^{\infty}\left(\frac{\mu}{1+\mu}\right)^n\bar{C}^nA_n(y). \quad (57)$$

We are now ready to derive the matrix expansion of $\bar{B}$ analogous to Eq. (7). The crucial step for this derivation is to expand the A_n in Hermite polynomials

$$A_n(y)=\sum_{k=0}^{\infty}A_{nk}H_k(y), \quad (58)$$

where the A_{nk} are constant coefficients. It is just this

[20] As pointed out by H. A. Kramers (footnote 15), the derivate moments α_n are even or odd functions as n is an even or odd number.

device that will be seen to make possible a simple expansion of the matrix of $\bar{B}$.

It is shown in Appendix B that the expansion coefficients A_{nk} have the form

$$A_{nk}=(1+\mu)^{-\frac{1}{2}}\left(\frac{\mu}{1+\mu}\right)^{(k-n-2)/2}a_{nk}, \tag{59}$$

where a_{nk} is independent of μ. It should be noted that a_{nk} is nonvanishing only when n and k have the same parity.[20] This gives

$$\bar{B}=\left(\frac{\mu}{1+\mu}\right)^{\frac{1}{2}}V_R\sum_{n=1}^{\infty}\sum_{(k)}\left(\frac{\mu}{1+\mu}\right)^{(n+k-2)/2}a_{nk}\bar{C}^nH_k(y), \tag{60}$$

in which the sum denoted by (k) is over all positive k values having the same parity as n; or, by transforming the indices of summation

$$\bar{B}=\left(\frac{\mu}{1+\mu}\right)^{\frac{1}{2}}V_R\sum_{m=0}^{\infty}\left(\frac{\mu}{1+\mu}\right)^{m}\sum_{p=-m}^{m+1}a_{m+p+1,m-p+1}\times\bar{C}^{m+p+1}H_{m-p+1}(y). \tag{61}$$

The matrix form of $\bar{B}$ is obtained almost immediately from the foregoing. We define the matrix element of $\bar{C}$ as the inner product between *normalized* Hermite functions:

$$C_{rs}=\left(\frac{h_r}{N_r^{\frac{1}{2}}},\bar{C}\frac{h_s}{N_s^{\frac{1}{2}}}\right)=r^{\frac{1}{2}}\delta_{r,s+1}. \tag{62}$$

In order to ascertain the properties of $H_r(y)$ as an operator, it is convenient to introduce the destruction operator

$$\bar{D}=y/2+d/dy, \tag{63}$$

whose effect on the Hermite functions is given by

$$\bar{D}h_r(y)=rh_{r-1}(y), \tag{64}$$

and whose matrix element with respect to normalized Hermite functions is

$$D_{rs}=\left(\frac{h_r}{N_r^{\frac{1}{2}}},\bar{D}\frac{h_s}{N_s^{\frac{1}{2}}}\right)=s^{\frac{1}{2}}\delta_{r,s-1}. \tag{65}$$

Since

$$\bar{C}=y/2-d/dy, \tag{66}$$

we have

$$y=\bar{C}+\bar{D}. \tag{67}$$

The matrix characterization of $H_r(y)$ then follows immediately from its functional form and from Eqs. (62) and (65), if one substitutes for y using Eq. (67)

$$H_r(y)=H_r(\bar{C}+\bar{D}). \tag{68}$$

The important thing about H_r is that its matrix elements with respect to normalized Hermite functions are *independent of* μ.

The upshot is that Eq. (61) may just as correctly stand for the matrix equation giving the matrix expansion of $\bar{B}$, as for a differential operator equation; and in this matrix expansion the coefficients a_{nk} and the matrices $\bar{C}^n$ and H_k are independent of μ. Thus Eq. (61) corresponds to the desired expansion (7), provided

$$\lambda=\mu/(1+\mu). \tag{69}$$

5. Fokker-Planck Equation and the Next Approximation

We write $\bar{B}^{(m_0)}$ for the approximation to $\bar{B}$ obtained by terminating the sum in Eq. (61) at $m=m_0$. Then from the expressions derived in Appendix B we find, for $m_0=0$,

$$\bar{B}^{(0)}=8V_R\left(\frac{\mu}{2\pi(1+\mu)}\right)^{\frac{1}{2}}(-\bar{C}y+\bar{C}^2)$$

$$=-8V_R\left(\frac{\mu}{2\pi(1+\mu)}\right)^{\frac{1}{2}}\bar{C}\bar{D}. \tag{70}$$

This stage corresponds to the case $m_0=0$, $\mathbf{B}_0=b_0=q_0q_0^{\dagger}$, of Eq. (14), and the approximate operator is exactly factorizable. Being in the form of a negative numerical factor times $-\bar{C}\bar{C}^{\dagger}=-\bar{C}\bar{D}$, it very transparently exhibits the negative semidefinite property. To get the operator which operates on the true probability density function, we invert the transformation (43) on $\bar{B}^{(0)}$, $\bar{C}$, and $\bar{D}$. We find

$$C=\exp\left(-\frac{y^2}{4}\right)\bar{C}\exp\left(\frac{y^2}{4}\right)=-\frac{d}{dy}, \tag{71}$$

and

$$D=\exp\left(-\frac{y^2}{4}\right)\bar{D}\exp\left(\frac{y^2}{4}\right)=y+\frac{d}{dy}. \tag{72}$$

Thus

$$B^{(0)}=8V_R\left(\frac{\mu}{2\pi(1+\mu)}\right)^{\frac{1}{2}}\cdot\frac{d}{dy}\left(y+\frac{d}{dy}\right), \tag{73}$$

which is the Fokker-Planck operator, as promised. Note the convenience of the form (70), from which it can be seen by inspection that the eigenfunctions of $\bar{B}^{(0)}$ are the $h_r(y)$, and that the eigenvalues (also those of $B^{(0)}$) are $-8nV_R[\mu/2\pi(1+\mu)]^{\frac{1}{2}}$.

$\bar{B}^{(1)}$, the next approximation to $\bar{B}$ which would be obtained by an uncritical inspection of Eq. (61), involves the addition to $\bar{B}^{(0)}$ of the operator

$$V_r\left(\frac{\mu}{1+\mu}\right)^{\frac{1}{2}}[a_{13}\bar{C}H_3+a_{22}\bar{C}^2H_2+a_{31}\bar{C}^3H_1+a_{40}\bar{C}^4H_0]. \tag{74}$$

This must be simplified. From Eq. (68) and the definitions of the polynomials H_r, it may be written as a function of creation and destruction operators. In the resulting expression all $\bar{D}$ operators may be moved to

the right of all $\bar{C}$ operators with the aid of the commutation relation, and the numerical values of the a_{nk} may then be substituted. When this is done, we find

$$\bar{B}^{(1)}=-8V_R\left(\frac{\mu}{2\pi(1+\mu)}\right)^{\frac{1}{2}} \times\left[\bar{C}\bar{D}+\frac{1}{6}\frac{\mu}{1+\mu}(\bar{C}^3\bar{D}-6\bar{C}^2\bar{D}^2+\bar{C}\bar{D}^3)\right]. \quad (75)$$

It is evident by inspection that this expression is Hermitian and has $h_0(y)$ as eigenfunction for eigenvalue zero, as it should. But it is not negative definite, since

$$(h_n,\bar{B}^{(1)}h_n)=-8V_RN_n\left(\frac{\mu}{2\pi(1+\mu)}\right)^{\frac{1}{2}} \times\left[n-\frac{\mu}{1+\mu}n(n-1)\right], \quad (76)$$

which becomes positive for sufficiently large n. Thus the procedure of Sec. I.2 must be used.

The construction of a negative semidefinite operator from $\bar{B}^{(1)}$ as given in Sec. I.2 amounts to "completing the square" of the expression in brackets, Eq. (75), as follows: Conditions on three constants α, β, and γ are found such that if

$$q=\alpha\bar{C}^3+\beta\bar{C}^2\bar{D}+\gamma\bar{C}\bar{D}^2, \quad (77)$$

then

$$\left(\bar{C}+\frac{1}{6}\frac{\mu}{1+\mu}q\right)\left(\bar{C}+\frac{1}{6}\frac{\mu}{1+\mu}q\right)^{\dagger} \quad (78)$$

agrees with the operator in brackets in Eq. (75) to terms of order $\mu/(1+\mu)$. The conditions are found to be

$$\alpha+\gamma=1, \qquad \beta=-3. \quad (79)$$

We thus have arrived at the following operator:

$$\bar{B}_1=-8V_R\left(\frac{\mu}{2\pi(1+\mu)}\right)^{\frac{1}{2}} \times\left\{\bar{C}+\frac{1}{6}\frac{\mu}{1+\mu}[\alpha\bar{C}^3-3\bar{C}^2\bar{D}+(1-\alpha)\bar{C}\bar{D}^2]\right\} \cdot\left\{\bar{D}+\frac{1}{6}\frac{\mu}{1+\mu}[\alpha\bar{D}^3-3\bar{C}\bar{D}^2+(1-\alpha)\bar{C}^2\bar{D}]\right\}. \quad (80)$$

This operator, multiplied by $[(1+\mu)/\mu]^{\frac{1}{2}}$, agrees with $\bar{B}$, multiplied by the same quantity, to terms of order $\mu/(1+\mu)$. It is Hermitian, and negative semidefinite with the transformed [according to Eq. (44)] Maxwell-Boltzmann distribution function as its stable stationary distribution. As a differential operator it is of sixth-order, and therefore is probably easier to handle in the form (80) than in strict differential-operator form. It does not appear likely that any simpler operator can furnish an equivalent approximation.

$\bar{B}_1$ is somewhat arbitrary in that the constant α is arbitrary; only the two Eqs. (79) determine the three constants α, β, and γ. However, the arbitrariness is in a term of higher order than that to which the operator is accurate; when (80) is multiplied out we must, of course, get

$$\bar{B}_1=-8V_R\left(\frac{\mu}{2\pi(1+\mu)}\right)^{\frac{1}{2}} \times\left\{\bar{C}\bar{D}+\frac{1}{6}\frac{\mu}{1+\mu}[\bar{C}^3\bar{D}-6\bar{C}^2\bar{D}^2+\bar{C}\bar{D}^3] +\frac{1}{36}\left(\frac{\mu}{1+\mu}\right)^2[\alpha\bar{C}^3-3\bar{C}^2\bar{D}+(1-\alpha)\bar{C}\bar{D}^2] \times[\alpha\bar{D}^3-3\bar{C}\bar{D}^2+(1-\alpha)\bar{C}^2\bar{D}]\right\}, \quad (81)$$

in which the term inside the braces with coefficient $\mu/(1+\mu)$, which is the first correction term, is independent of α. However, the higher-order term in braces, which has coefficient $[\mu/(1+\mu)]^2$, will not, except by coincidence, agree with the term of like order in the exact operator $\bar{B}$, no matter what the value of α, since the former is in general only part of the term in the exact operator. $\bar{B}_1$ is, however, not meant to be accurate to this order (this higher term in $\bar{B}_1$ would not be accurate even if there were no arbitrariness), and computations should not be carried beyond terms which are determined by the $\mu/(1+\mu)$ term in braces in (81). Then the arbitrariness due to the indeterminateness of α will play no part in the results.

It should not be concluded, from the fact that the $[\mu/(1+\mu)]^2$ term in braces in (81) is to be disregarded where it affects computational results, that it can be dispensed with. By ensuring negative semidefiniteness it prevents runaway solutions; without it, probability modes $h_r(y)$ of very large r value will grow indefinitely in amplitude. It is to be expected that, when used in the proper range of deviations from equilibrium, $\bar{B}_1$ will yield nonarbitrary results. The term which contains the arbitrariness must be included to prevent the intrusion, into solutions of the approximate equation, of spurious effects.

The arbitrariness due to α does not affect the possibility of constructing a sequence of approximations to $\bar{B}$. As constructed according to the prescription in Sec. I.2, $[(1+\mu)/\mu]^{\frac{1}{2}}\bar{B}_N$ will always be correct to order $[\mu/(1+\mu)]^N$. Thus in $\bar{B}_2$ the error due to α will be made good, although a new error will be introduced in a term of higher order.

The results obtained will now be considered in relation to Brinkman's assertion that the condition (23) requires

that all α_n vanish for $n>2$ (our α_n = Brinkman's μ_n). In the first place, the fact that F decreases when $P(V,t)$ obeys the linear Boltzmann equation, as proved in Sec. I.5, makes it impossible from our point of view to agree with this conclusion, since the linear Boltzmann operator in general has nonvanishing derivate moments of all orders. However, in any case, what we have sought is an approximation to an operator which in its exact form does satisfy the requirement. If this operator can be expanded in powers of $\mu/(1+\mu)$, then it is clear that successive approximations to its effect on $P(V,t)$ can be obtained by breaking the series off at successively higher terms, and that these approximations might be useful even if they did not satisfy some of the requirements the exact operator satisfies.

III. MORE GENERAL SYSTEMS

The above work can be generalized to other linear Boltzmann operators in the following two ways: by leaving the a_{nk} general, and by suppressing explicit reference to the expansion parameter $\mu/(1+\mu)$. In the following sections we take up these two modes of generalization successively.

1. Case of General a_{nk}

The a_{nk} are not mutually independent. Let us write the expansion of $\bar{B}$ in the form

$$\bar{B}=V_R\left(\frac{\mu}{1+\mu}\right)^{\frac{1}{2}}\sum_{m=0}^{\infty}\left(\frac{\mu}{1+\mu}\right)^m\bar{b}_m. \tag{82}$$

Each $\bar{b}_m$ in (82) must end in a destruction operator in order that $h_0(y)$ be a stable equilibrium solution. In the case $m=0$ we have, from Eq. (61),

$$\begin{aligned}\bar{b}_0&=[a_{11}\bar{C}y+a_{20}\bar{C}^2]\\&=[a_{11}\bar{C}(\bar{C}+\bar{D})+a_{20}\bar{C}^2].\end{aligned} \tag{83}$$

The stable equilibrium condition here requires that the coefficient of $\bar{C}^2$ vanish, or

$$a_{20}=-a_{11}. \tag{84}$$

This is, of course, just the classic relation between viscosity and diffusion coefficient discovered by Einstein.[21]

A similar relation can be found in the next order: After bringing $\bar{D}$ operators to the right in all terms, we have

$$\begin{aligned}\bar{b}_1=&(a_{13}+a_{22}+a_{31}+a_{40})\bar{C}^4+(3a_{13}+2a_{22}+a_{31})\bar{C}^3\bar{D}\\&+(3a_{13}+a_{22})\bar{C}^2\bar{D}^2+a_{13}\bar{C}\bar{D}^3.\end{aligned} \tag{85}$$

The condition that this annihilate $h_0(y)$ is that the coefficient of $\bar{C}^4$ vanish:

$$a_{13}+a_{22}+a_{31}+a_{40}=0. \tag{86}$$

In this case another restriction on the a_{nk} must be satisfied too, in order that $\bar{b}_1$ be Hermitian. Namely, the coefficients of $\bar{C}^3\bar{D}$ and of $\bar{C}\bar{D}^3$ must be equal:

$$3a_{13}+2a_{22}+a_{31}=a_{13}. \tag{87}$$

With Eqs. (84), (86), and (87) we can eliminate a_{20}, a_{31}, and a_{40}. When this is done, we obtain

$$\begin{aligned}B^{(1)}=V_R\left(\frac{\mu}{1+\mu}\right)^{\frac{1}{2}}\Big\{&a_{11}\bar{C}\bar{D}+\frac{\mu}{1+\mu}[a_{13}\bar{C}^3\bar{D}\\&+(3a_{13}+a_{22})\bar{C}^2\bar{D}^2+a_{13}\bar{C}\bar{D}^3]+o\left(\frac{\mu}{1+\mu}\right)\Big\}.\end{aligned} \tag{88}$$

In the Rayleigh process the a_{nk} are known, and in fact the relations derived above can be verified for the expressions given in Appendix B. However, Eq. (88) may also be applied in the following way. Suppose thermal fluctuations are to be studied beyond the range where linear friction applies, in some system whose linear Boltzmann operator is unknown, but in which the following hypotheses may be justifiable: (a) The (unknown) linear Boltzmann operator is expansible in terms of some parameter analogous to $\mu/(1+\mu)$, and (b) the successive derivate moments are expansible in Hermite polynomials. This amounts to saying that the random process involved is mathematically of the same type as the Rayleigh process. Equation (88) or its equivalent then tells us that in order to study the random process in a consistent way with inclusion of the V^3 term in the friction[22] (i.e., in the first derivate moment), it is sufficeint to know just the coefficient of this term and that of the V^2 term in the noise, i.e., in the second derivate moment; the remaining relevant coefficients a_{31} and a_{40}, which appear in the third and fourth derivate moments, being determined by the former two.

As in Sec. II.5, a negative semidefinite operator agreeing with $\bar{B}^{(1)}$ to order $[\mu/(1+\mu)]^{\frac{1}{2}}$ can be constructed by completing the operator absolute-square. The result is

$$\begin{aligned}\bar{B}_1=\left(\frac{\mu}{1+\mu}\right)^{\frac{1}{2}}V_R\Big\{&a_{11}^{\frac{1}{2}}\bar{C}+\frac{\mu}{1+\mu}a_{11}^{-\frac{1}{2}}\\&\times[\alpha\bar{C}^3+\tfrac{1}{2}(3a_{13}+a_{22})\bar{C}^2\bar{D}+(a_{13}-\alpha)\bar{C}\bar{D}^2]\Big\}\\&\times\Big\{a_{11}^{\frac{1}{2}}\bar{D}+\frac{\mu}{1+\mu}a_{11}^{-\frac{1}{2}}[\alpha\bar{D}^3+\tfrac{1}{2}(3a_{13}+a_{22})\bar{C}\bar{D}^2\\&+(a_{13}-\alpha)\bar{C}^2\bar{D}]\Big\}.\end{aligned} \tag{89}$$

[21] A. Einstein, Ann. Physik 17, 549 (1905).

[22] Carried out to its second term, the expansion of A_1 is

$$\begin{aligned}A_1(y)&=\frac{(1+\mu)^{\frac{1}{2}}}{\mu}\left\{a_{11}H_1+\frac{\mu}{1+\mu}a_{13}H_3+\cdots\right\}\\&=\frac{(1+\mu)^{\frac{1}{2}}}{\mu}\left\{a_{11}y+\frac{\mu}{1+\mu}a_{13}(y^3-3y)+\cdots\right\}.\end{aligned}$$

If $\mu\ll1$, the H_3 contribution will not come in until $y^3\sim1/\mu$, and then the $-3y$ term will be negligible compared to y^3. Similarly with all higher H_r contributions. If they contribute significantly at all, the highest power in them will dominate. Thus when $\mu\ll1$ the Hermite expansion will not be appreciably different from a power series. And this will, of course, hold for the expansion of any derivate moment.

2. Suppression of the Expansion Parameter

Let us put

$$k_1=\left(\frac{\mu}{1+\mu}\right)^{1/4} a_{11}^{\frac{1}{2}}$$

$$k_2=\left(\frac{\mu}{1+\mu}\right)^{5/4} a_{11}^{-\frac{1}{2}}\alpha$$

$$k_3=\left(\frac{\mu}{1+\mu}\right)^{5/4} \tfrac{1}{2}a_{11}^{-\frac{1}{2}}(3a_{13}+a_{22}) \tag{90}$$

$$k_4=\left(\frac{\mu}{1+\mu}\right)^{5/4} a_{11}^{-\frac{1}{2}}(a_{13}-\alpha).$$

Then

$$\begin{aligned}\bar{B}_1&=V_R[k_1\bar{C}+k_2\bar{C}^3+k_3\bar{C}^2\bar{D}+k_4\bar{C}\bar{D}^2]\\&=[k_1\bar{D}+k_2\bar{D}^3+k_3\bar{C}\bar{D}^2+k_4\bar{C}^2\bar{D}].\end{aligned} \tag{91}$$

This form of $\bar{B}_1$ would be usable, if valid, for a system not possessing an expansion parameter, or for which this parameter was unknown. In the absence of as yet unsuspected restrictions, the four k's of Eqs. (90) are mutually independent; they certainly are so for the Rayleigh model, since α, a_{11}, a_{13}, and a_{22} are independent.

The form of the operator of Eq. (91) with arbitrary k's is sufficient for negative semidefiniteness. Let us try to define the conditions under which it is also *necessarily* the next negative semidefinite approximation after the Fokker-Planck operator. In terms of α_{nk}, the kth Hermite coefficient of $\alpha_n(y)$, Eq. (61) reads

$$\bar{B}=\sum_{m=0}^{\infty}\sum_{p=0}^{m+1}\frac{\alpha_{m+p+1,m-p+1}}{(m+p+1)!}\bar{C}^{m+p+1}H_{m-p+1}(y). \tag{92}$$

This is a perfectly general formal expression for any linear Boltzmann operator, since it may be derived without any further assumptions from Eq. (37). If the equilibrium distribution of y is Gaussian, $\exp(-y^2/2)$, sufficient conditions on $\alpha_{nk}/n!$ in order that $\bar{B}$ be Hermitian and promote stable equilibrium are the same as Eq. (84) for the $\alpha_{nk}/n!$ with $n+k=2$, and the same as Eqs. (86) and (87) for those with $n+k=4$. To prove these conditions necessary as well, a variable expansion parameter analogous to $\mu/(1+\mu)$ is needed, in order to make possible the device of setting the coefficient of each power of the parameter equal to zero. However, in the spirit of a phenomenological approach it may be justifiable to hypothesize the existence of such a parameter, when definite knowledge about a given system is not available. Assuming the hypothetical parameter to be small as well—as would be reasonable for any macroscopic variable which fluctuates due to molecular impulses or contributions—Eq. (91) would then be the most general form for the indicated approximation.

ACKNOWLEDGMENTS

This work is the outgrowth of a problem suggested to the author by Professor George E. Uhlenbeck, whose friendly and stimulating advice has moreover been invaluable throughout its course. The author wishes also to thank the John Simon Guggenheim Memorial Foundation for the grant of a fellowship for the period 1957–58, during which roughly the first half of the work was done, while on leave from Boston University.

APPENDIX A

Derivation of the Transition Probability for the Rayleigh Model

We here evaluate the function $C(y''|y')$ of Eq. (36) in the text. To do so we first work in terms of the ordinary velocity V. Let

$C(V''|V')=$ probability density-in-V'' per unit time that a particle with given velocity V' undergoes a collision which changes its velocity to V'' ("transition probability" from V' to V''). (A1)

Let v stand for the component of velocity of a molecule in the direction of the constrained motion of the particles (note that v therefore does *not* stand for the *speed* of the molecules). Given a particle with initial velocity V, and assuming the distribution (31) for the vector velocity v, the probability density-in-v for a collision of the particle with a molecule having velocity component (in the foregoing sense) v, per unit time, is

$$A\rho f_0(v)|V-v|, \tag{A2}$$

if A is the area of the disk of the particle, ρ the spatial density of molecules, and $f_0(v)$ the one-dimensional Maxwell-Boltzmann function

$$f_0(v)=(2\pi v_R^2)^{-\frac{1}{2}}\exp\left(-\frac{v^2}{2v_R^2}\right). \tag{A3}$$

The fraction of all disks which are knocked out of the infinitesimal range dV at V by molecules in dv at v is, per unit time,

$$A\rho f_0(v)|V-v|P(V)dvdV, \tag{A4}$$

where $P(V)$ is the velocity probability density of particles.

The coefficient of $dvdV$ in (A4) is the probability density-in-v-and-V, per unit time, of the process described. However, for use in Eq. (35) we require the transition probability between two values of V, namely, V' and V'' (ultimately, y' and y''). These may be related to v and V by the dynamics of the collision, as follows: Let

$$V'=V. \tag{A5}$$

Assume that molecules are reflected specularly from the disks; then V', V'' are related to the variables V, v by Eq. (A5) and by

$$V''-V'=2\frac{\mu}{1+\mu}(v-V) \tag{A6}$$

where

$$\mu=m/M. \tag{A7}$$

To obtain a transition probability for $V'\rightarrow V''$ from the expression (A4) one must (a) write this expression in terms of V' and V''; (b) make it a probability density in V' and V'' by multiplying it by $\partial(v,V)/\partial(V',V'')=(1+\mu)/2\mu$, and (c) divide by $P(V')$ to obtain a probability *conditional* in V'. In this way one obtains

$$C(V''|V')=A\rho\left(\frac{1+\mu}{2\mu}\right)^2 f_0\left(\frac{1+\mu}{2\mu}V''-\frac{1-\mu}{2\mu}V'\right)\cdot|V''-V'|. \tag{A8}$$

An expression formally the same as this would be obtained if the gas were linear instead of three-dimensional, with the particles on a line with the gas molecules, and if every encounter between a particle and a molecule resulted in a collision. The latter is the model originally introduced by Rayleigh,[11] and the present one is mathematically equivalent to it.

Now transform to the variable y [Eq. (33)]. Writing $C(y''|y')$ for the transition probability per unit time for the event $y'\rightarrow y''$, which is a probability density in y'', we shall have

$$C(y''|y')=C(V''|V')dV''/dy'' \tag{A9}$$

(in this equation the C's stand for transition probabilities with respect to the arguments in their respective parentheses; since the arguments are different random variables on the two sides of the equation, the C's on the two sides are not meant to be the same functions). Put

$$A\rho=1, \tag{A10}$$

since this combination of constants plays no further part in the analysis. Equation (A9) applied to Eq. (A8) then gives Eq. (36) of the main text.

APPENDIX B

We here evaluate the Hermite expansion coefficients of $A_n(y)$. From Eq. (58) and the normalization constant of the $H_k(y)$,

$$A_{nk}=\frac{1}{(2\pi)^{\frac{1}{2}}k!}\int\exp\left(-\frac{y^2}{2}\right)H_k(y)A_n(y)dy. \tag{B1}$$

We utilize the generating function of the Hermite polynomials,

$$\exp\left(-\frac{z^2}{2}+zy\right)=\sum_{m=0}^{\infty}z^r\frac{H_r(y)}{r!}, \tag{B2}$$

whence A_{nk} is the coefficient of z^k in the power series expansion of

$$I(z)=(2\pi)^{-\frac{1}{2}}\int\exp\left(-\frac{y^2}{2}-\frac{z^2}{2}+zy\right)A_n(y)dy$$
$$=\frac{2^n}{2\pi n!}\int dy\exp\left(-\frac{y^2}{2}-\frac{z^2}{2}+zy\right)\times\int dx\exp\left[-\frac{\mu}{2}(x+y)^2\right]x^n|x|dx. \tag{B3}$$

Inverting the order of integration, it is possible to integrate immediately over y, using the formula

$$\int_{-\infty}^{\infty}\exp(-ay^2-by)dy=\left(\frac{\pi}{a}\right)^{\frac{1}{2}}\exp\left(\frac{b^2}{4a}\right). \tag{B4}$$

This gives

$$I(z)=\frac{2^n}{[2\pi(1+\mu)]^{\frac{1}{2}}n!}\int\exp\left(-\frac{\mu z^2}{2(1+\mu)}-\frac{\mu zx}{1+\mu}\right)\exp\left(-\frac{\mu}{2(1+\mu)}x^2\right)x^n|x|dx. \tag{B5}$$

But here we recognize, in the first exponential in the integrand, the generating function according to formula (B2) of the functions

$$H_r\left(-\left[\frac{\mu}{1+\mu}\right]^{\frac{1}{2}}x\right).$$

From this it follows that

$$A_{nk}=\frac{(-1)^k2^n}{[2\pi(1+\mu)]^{\frac{1}{2}}n!k!}\left(\frac{\mu}{1+\mu}\right)^{k/2}\times\int\exp\left(-\frac{\mu}{2(1+\mu)}x^2\right)H_k\times\left(\left[\frac{\mu}{1+\mu}\right]^{\frac{1}{2}}x\right)x^n|x|dx. \tag{B6}$$

Changing to $[\mu/(1+\mu)]^{\frac{1}{2}}x$ as variable of integration,

$$A_{nk}=\frac{(-1)^k2^n}{[2\pi(1+\mu)]^{\frac{1}{2}}n!k!}\left(\frac{\mu}{1+\mu}\right)^{(k-n-2)/2}\times\int\exp\left(-\frac{x^2}{2}\right)H_k(x)x^n|x|dx. \tag{B7}$$

The integral can be simplified as follows:

$$\int_{-\infty}^{\infty} \exp\left(-\frac{x^2}{2}\right) H_k(x) x^n |x| dx$$

$$= \int_0^{\infty} [(-1)^{n+k+2}+1] H_k(x) \exp\left(-\frac{x^2}{2}\right) x^{n+1} dx$$

$$= 2\int_0^{\infty} H_k(x) \exp\left(-\frac{x^2}{2}\right) x^{n+1} dx, \quad \text{if } n+k \text{ even}$$

$$= 0 \qquad \text{if } n+k \text{ odd.} \quad \text{(B8)}$$

The dependence on parity of $n+k$ agrees with the fact that $A_n(y)$ is even or odd according to the parity of n (cf. footnote 20).

The integral in (B8) is

$$J = \int_0^{\infty} H_k(x) \exp\left(-\frac{x^2}{2}\right) x^{n+1} dx$$

$$= \int_0^{\infty} x^{n+1} \left(-\frac{d}{dx}\right)^k \exp\left(-\frac{x^2}{2}\right) dx. \quad \text{(B9)}$$

If $k \leqq n+1$ we integrate by parts k times to get

$$J = (n+1)n \cdots (n-k+2) \int_0^{\infty} x^{n-k+1} \exp\left(-\frac{x^2}{2}\right) dx$$

$$= \frac{(n+1)!}{(n-k+1)!!} \quad (k \leqq n+1), \quad \text{(B10)}$$

where the double factorial $N!!$ is defined by

$$N!! = N(N-2)(N-4)\cdots 3\cdot 1 \quad (N \text{ odd})$$
$$= N(N-2)(N-4)\cdots 4\cdot 2 \quad (N \text{ even}), \quad \text{(B11)}$$

and, by convention, $0!! = (-1)!! = 1$.

If $k > n+1$ we integrate by parts $n+1$ times:

$$J = (n+1)! \left[-\left(-\frac{d}{dx}\right)^{k-n-2} \exp\left(-\frac{x^2}{2}\right)\right]\Bigg|_0^{\infty}$$

$$= (n+1)! H_{k-n-2}(0)$$

$$= (-1)^{(k-n-2)/2} (n+1)! (k-n-3)!!, \quad \text{(B12)}$$

the last form being obtained by adapting, to our definition of Hermite polynomials, formula (13.15), Sec. 10, of *Higher Transcendental Functions* [edited by A. Erdélyi (McGraw-Hill Book Company, Inc., New York, 1953), Vol. II].

If we combine the foregoing results, we get

$$A_{nk} = (1+\mu)^{-\frac{1}{2}} \left(\frac{\mu}{1+\mu}\right)^{(k-n-2)/2} a_{nk}, \quad \text{(B13)}$$

with

$$a_{nk} = (-1)^k \frac{2^{n+1}}{(2\pi)^{\frac{1}{2}}} \frac{n+1}{k!(n-k+1)!!} \quad (k \leqq n+1) \quad \text{(B14)}$$

$$= (-1)^{(n+k+2)/2} \frac{2^{n+1}}{(2\pi)^{\frac{1}{2}}} \frac{(n+1)(k-n-3)!!}{k!}$$
$$(k > n+1). \quad \text{(B15)}$$

Topological Derivation of the Mayer Density Series for the Pressure of an Imperfect Gas

MELVILLE S. GREEN*
Hughes Aircraft Company, Culver City, California
(Received April 8, 1960)

A new derivation of Mayer's classical density expansion for the pressure of an imperfect gas based on a classification of cluster graphs according to topological criteria is presented. The classification is a generalization of the classification of simple trees into trees with centers and trees with bicenters.

RECENTLY, a number of significant advances in the theory of many-particle systems have been made by the method of summation of infinite subclasses of terms in series defined by diagrams. Perhaps the classic example of such series are Mayer's cluster expansion[1] for thermodynamic properties and distribution functions in nonquantum equilibrium statistical mechanics. Since Mayer's graphs are simpler and more purely combinatorial objects than the Feynman-type graphs[2] which occur in quantum perturbation theory, it is illuminating to go back and look at the classical version of some of the combinatorial problems of the quantum theory of many-body systems. In the course of just such an investigation the author has come upon a new derivation of Mayer's density series for the pressure which is the classical analog of a rearrangement of the quantum perturbation expansion for the grand partition function of a bose gas.[3] This derivation is expounded here for its own sake as well as for the light it throws on the quantum problem.

Of the many derivations of Mayer's density series,[4] the present one is most closely related to Salpeter's, but unlike his, it gives directly an expansion for the pressure rather than the chemical potential. Like Salpeter's, the present derivation does not require the evaluation of a complex numerical combinatorial factor but relies rather on certain topological properties of cluster graphs. The essential part of the present derivation is a theorem, to be demonstrated below, which shows that each reducible cluster can be considered to be built upon a foundation which is either an irreducible cluster or a point.

Mayer has given the following expression for the logarithm of the classical grand partition function of a monatomic gas or, what is the same thing, PV/kT:

$$\frac{PV}{kT}=V\left[z+\sum_{n=2}^{\infty}\frac{z^n}{n!}\sum_{\substack{\text{all clusters of}\\ n\text{ particles}}}\int\prod f_{ij}d(i)\cdots d(n)\right], \quad (1)$$

where z is the "active" number density or activity,

$$f_{ij}=\exp-u(|\mathbf{x}_i-\mathbf{x}_j|)/kT-1,$$

$u(|\mathbf{x}_i-\mathbf{x}_j|)$ is the intermolecular potential between molecule i and molecule j, P is the pressure, V is the volume, T is the temperature of the gas, and k is Boltzmann's constant. $(1)\cdots(n)$ is an abbreviation for $\mathbf{x}_i\cdots\mathbf{x}_n$, $d(1)\cdots d(n)$ is an abbreviation for $d\mathbf{x}_1\cdots d\mathbf{x}_n$.

The product of Πf_{ij} contains a number of factors f_{ij} with indices ij taken from the set $1\cdots n$. Every term on the right in (1) corresponds to a diagram consisting of points labeled $1\cdots n$, connected pairwise by a number of arcs. For every arc joining points, say $k1$, in a particular diagram, a factor f_{k1} appears in the corresponding terms of the series. In mathematical terminology a diagram constructed in this way is called a graph with labeled points. The second summation on the right is over all products f_{ij} whose corresponding graph is a cluster, i.e., a connected graph in which no pair of points joined by more than one arc.

In what follows, we shall also need the activity series for the number of particles:

$$N=PV=V\left[z+\sum_{n=1}\frac{z^{n+1}}{n!}\times\sum_{\substack{\text{all clusters of}\\ n+1\text{ particles}}}\int\prod f_{ij}d(2)\cdots d(n+1)\right]. \quad (2)$$

We consider now two particular clusters for the purpose of illustrating a number of definitions and simple theorems about graphs. (We use the terminology of Ford and Uhlenbeck.[5]) Points such as 10 in Fig. 1(a) are called articulation points. The removal of 10 together with all arcs joined to it divides the graph into two disjoint parts. This is not true, for instance, of 12, which is therefore not an articulation point. A connected graph without articulation points is called a star. [Thus 10–12 and 4, 5 are stars.] An arbitrary cluster is divided by its articulation points into a collection of stars. Thus the articulation points 2, 3, 4, 6, 8 in Fig. 1(b) divide it into the stars (12), (23), (346), (45), (687), (89). Conversely, any cluster can be considered as a collection of stars joined together at articulation

* Permanent Address: National Bureau of Standards, Washington 25, D. C.

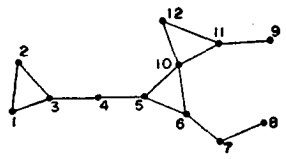
[1] J. E. Mayer and M. G. Mayer *Statistical Mechanics* (John Wiley & Sons, Inc., New York, 1940).

[2] N. M. Hugenholtz, Physica **23**, 48 (1957); T. D. Lee and C. N. Yang, Phys. Rev. **113**, 1165 (1959).

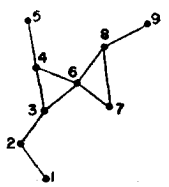
[3] M. S. Green, Phys. Rev. Letters **1**, 409 (1958).

[4] J. E. Mayer, J. Chem. Phys. **5**, 67 (1937); J. E. Mayer and P. F. Ackermann, *ibid.* **5**, 74 (1937); B. Kahn and G. E. Uhlenbeck, Physica **5**, 399 (1938); M. Born and K. Fuchs, Proc. Roy. Soc. (London) **A166**, 391 (1938); K. Husimi, J. Chem. Phys. **18**, 682 (1950); E. A. Salpeter, Ann. Phys. **5**, 183 (1958).

[5] G. W. Ford and G. E. Uhlenbeck, Proc. Natl. Acad. Sci. U. S. **42**, 122 (1956).



(a)



(b)

FIG. 1. Two representative clusters.

points, and from this point of view the cluster is called a star tree. A star tree with one point singled out is called a rooted star tree, and the singled-out point is called the root. A star tree may consist of a number of branches disjoint except for the root. If in Fig. 1(b) the point 3 is designated as the root, we obtain a rooted star tree with two branches, (123) and (3456789).

A cluster may be considered to be made up in a number of different ways of stars, rooted star trees, and branches joined at articulation points. For each such way, the corresponding integral can be written as a product of integrals referring to the component stars, rooted star trees, and branches. This is a consequence of the fact that f_{ij} depends only on the distance $|\mathbf{x}_i-\mathbf{x}_j|$, which is, indeed, what makes graph-theoretic concepts important in the theory of cluster series. Thus, the integral corresponding to Fig. 1(b) may be written

$$\int f_{12}f_{23}f_{34}f_{36}f_{46}f_{45}f_{67}f_{68}f_{87}f_{89}d(1)\cdots d(9)$$
$$=\int f_{12}d(1)\times\int f_{23}d(2)\times\int f_{45}d(5)\times\int f_{89}d(9)$$
$$\times\int f_{34}f_{36}f_{46}d(3)d(4)\times\int f_{67}f_{68}f_{87}d(7)d(8). \quad (3)$$

Alternatively, this same integral may be considered to be the product of an integral referring to one of the stars in Fig. 1(b) and integrals referring to each one of the rooted star trees which are attached to this star through articulation points.

It is perhaps obvious from the principle illustrated by (3) that it should be possible to express (1) and (2) in terms of integrals referring only to stars. It is not so obvious, however, that these equations attain their simplest form when expressed as a power series in the density. As is well known, the result of reexpressing (1) in terms of integrals referring to stars (irreducible cluster integrals in Mayer's terminology) is the density series for the pressure, which may be written in the form

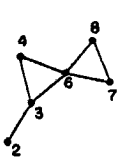
$$PV/kT=V(\rho+S-\rho\partial S/\partial\rho), \quad (4)$$

where ρ is the average number density and

$$VS=\sum_{n=2}^{\infty}\frac{\rho^n}{n!}\sum_{\substack{\text{all stars of}\\ n\text{ points}}}\int\prod f_{ij}d(1)\cdots d(n). \quad (5)$$

The appearance of a power series in ρ rather than z in which the coefficients are determined by stars rather than by arbitrary clusters becomes less surprising when it is noted that the series in (2) for ρ is essentially a sum over rooted star trees and that any cluster may be considered to be a star with a number of rooted star trees attached to it. The difficulty with this last point of view, however, is that in general a cluster can be represented as a star with a number of rooted star trees attached in several different ways. We show now that all clusters can be divided into two classes. For every member of the first class there is a certain star which has a special right to be the basis upon which the cluster is built by attaching rooted star trees. For every member of the second class there is a special point from which the cluster is built by adding branches.

Consider again Fig. 1(b). There are three stars (89), (45), and (12), which are attached to the remainder of the graph by only one (articulation) point. Let us remove these stars and obtain Fig. 2(a). This graph has two stars, (23) and (678), attached by only one point. Removing these, we obtain the star (436) in Fig. 2(b). Figure 1(b) can thus be considered to be constructed from the fundamental star (436) by the attachment of the rooted star trees (321), (45), and (6879). We will call the process of successive removal of all stars attached at only one point the reduction process. It is essential that all stars containing a single articulation point be completed before the new graph is examined and the process is applied again. The reduction process applied to the graph in Fig. 1(a) yields the graph in Fig. 3(a) after two steps, while the final step yields simply the point 5. This graph may thus be considered to consist of the two branches $(1\cdots5)$ and $(5\cdots12)$ joined together at the point 5.



(a)

(b)

FIG. 2. The reduction of Fig. 1(b); (a) the first step; (b) the second step.

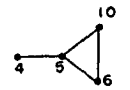
That the reduction process can be applied to any reducible cluster and will result in either a star or a point after a finite number of steps follows from the fact that if a graph has an articulation point, there is at least one star attached to the remainder of the graph at only one point. If the result of the process is a star, we will call the star the fundamental star and if a point, the fundamental point. Any graph can thus be considered to be built of a fundamental star with a number of rooted trees attached or of a number of branches joined at a fundamental point.

The question immediately arises: What is it that distinguishes the fundamental star or fundamental point from all other stars or points of a graph? This question can be answered by introducing the notion of order of a rooted star tree. The order of a star tree is the number of steps in the reduction process required to reduce it to a point. Now if a graph has a fundamental star, the order of no one of the *attached rooted star trees* is greater than the order of all of the others, and if a graph has a fundamental point, the order of no one of the *attached branches* is greater than the order of all the others. Suppose, for instance, that a graph with a fundamental star has an attached rooted star tree of the order n greater than the order of all the others. After $n-1$ steps in the reduction process, all the other rooted star trees will have been reduced to points, while the given one will have been reduced to one or more stars attached to the fundamental star at one point. If this is so, however, the reduction process correctly applied one more step should have resulted in a point, which contradicts the hypothesis that the graph has a fundamental star. On the other hand, if a graph with a fundamental point has an attached branch of largest order, at the $(n-1)$th step, all other attached branches except the maximum will have been reduced to a point, while the maximum will have been reduced to a star. Again this contradicts the hypothesis that the graph has a fundamental point since, if the reduction process had been correctly applied, it would have yielded a star at the $(n-1)$th step. It is easy to see that if a graph has a star with the foregoing property, no other star has the foregoing property, nor does it have a point with the foregoing property; and if a graph has a point with the foregoing property, it has no other point with the foregoing property, nor does it have a star with the foregoing property. Thus we have the following theorem:

All clusters can be divided into two disjoint classes, those with a unique fundamental point, and those with a unique fundamental star. If a graph has a fundamental point, no branch attached to this point has an order greater than the rest. If a graph has a fundamental star, no rooted star tree attached to this star has an order greater than the rest.[6]

[6] The notions of fundamental star and fundamental point of a star tree are generalizations of the notion of center and bicenter of a simple tree. See D. König, *Theorie der Endlichen und Unendlichen Graphen* (Chelsea Publishing Company, New York, 1950), Chap. 5. A similar, but not identical, generalization has been made by F. Harary and R. Z. Norman [Ann. Math. 58, 134 (1953)].

FIG. 3. The next-to-last step in the reduction of Fig. 1(a).



With this theorem we are in a position to reexpress (1) in terms of star integrals. Consider a particular term on the right in (1) corresponding to a labeled cluster which has a fundamental star σ and attached rooted star trees $\tau_1, \tau_2, \cdots \tau_r$. Let m be the number of points in σ and $m_1 \cdots m_r$, the number of points besides the root in $\tau_1 \cdots \tau_r$. The total number of points is $n=m+m_1+\cdots m_r$. There will be a number of other clusters differing only in labeling which yield the same contribution. We compute the contribution of all clusters identical to the given one, except for labeling. We suppose that in the given cluster the fundamental star is labeled with the integers $1 \cdots m$, while $\tau_1 \cdots \tau_r$ are labeled with integers taken from the sets $(m_1) \cdots (m_r)$. The operation of summing over all labelings of the given graph can be written as the sum over all labelings which do not interchange labels among σ and $\tau_1 \cdots \tau_r$ times the number of ways of choosing sets of $m, m_1 \cdots m_r$ integers out of n integers $(=n!/m_1! \cdots m_r!)$. Thus we have for the contribution of the given cluster and all its relabelings

$$\frac{1}{m!m_1! \cdots m_r} \sum_{\substack{\text{all labelings of } \sigma \\ \text{with integers } 1 \cdots n}} \int f_{ij} d(1) \cdots d(n) \times \prod_{s=1}^{r} \sum_{\substack{\text{all labelings of } \tau_s \\ \text{with integers taken from } (m_s)}} \int \prod f_{ij} d[m_s], \quad (6)$$

where $d[m_s]$ is the product of volume elements for all molecules of the sets $[m_s]$.

We now sum over all clusters which have the same fundamental star as the given cluster. All such clusters can be built up by attaching various collections of trees to the points of the star. The trees, however, may not be added independently. We must be sure that there are at least two of maximum order, otherwise the given star will not be the fundamental star of the cluster.

In order to find the total contribution to (1) from clusters with fundamental stars, it is most convenient first to compute the sum without regard to restrictions in adding chains and then to subtract the superfluous contributions. The sum without regard to restrictions is

$$VS=\sum_{m=2}^{\infty} \frac{\rho^m}{m!} \sum_{\text{all stars}} \int \prod f_{ij} d(1) \cdots d(n), \quad (7)$$

$$\rho=z+\sum_{n=1} \frac{z^{n+1}}{n!} \times \sum_{\substack{\text{all rooted star} \\ \text{trees with root 1} \\ \text{from points } 2 \cdots n+1}} \int \prod f_{ij} d(2) \cdots d(n+1). \quad (8)$$

Any point of the star σ may have either no rooted tree attached, corresponding to the first term on the right in (8), or an arbitrary rooted tree, corresponding to one of the second group of terms. Since we allow trees to be attached independently, the sum of products is the product of the sums. Since there is a one-to-one correspondence between labeled rooted star trees with root 1 formed from $n+1$ points and labeled clusters formed from $n+1$ points, (8) is simply another form of (2), and ρ is, in fact, the number density.

The contributions to be subtracted are those from clusters formed from σ (which is, of course, no longer the fundamental star of the cluster) by attaching a number of rooted star trees, one of which has greater order than the rest. If s is the (articulation) point at which the tree of largest order τ_s is attached to σ, the remaining trees $\tau_i (i \neq s)$ together with σ, form a branch β whose order is almost equal to that of τ_s. The contributions to be subtracted may thus also be characterized as those from clusters formed by joining a branch of order p to a rooted star tree of order q with $p \leqslant q$.

We turn now to the contributions to (1) from clusters with fundamental points. These are formed by attaching a number of branches at a point, taking care that at least two branches have order greater than the rest. Again, it is convenient first to compute the sum without restrictions, then to subtract the superfluous contributions. It is easily seen that the unrestricted contribution together with the first term z in (1) is simply ρ as given by (8). The contribution to be subtracted is that from clusters formed by joining one or more branches $\beta_1 \cdots \beta_s$ to a common root such that one, say β_s, has order p greater than the rest. The remaining branches β_i $(i \neq s)$ form a tree of order q with $p > q$. The case in which β_s is the only branch must also be included. If we call a point a tree of order zero, the above superfluous contributions may be described as those arising by joining a rooted star tree of order q and a branch of order p with $p > q \geqslant 0$.

We see now that the two collections of superfluous contributions complement each other. Together they comprise contributions from all graphs formed by joining a branch and a tree to a common root with no restrictions on the order of the branch or the rooted tree. It may be remarked that in general a given cluster appears several times in this collection. A cluster appears, in fact, as many times as the number of ways it can be constructed by attaching a tree and a branch at a common root.

We consider a particular labeled cluster formed by joining a particular branch β with a root 1 and remaining points labeled with integers $2 \cdots n+1$, and a particular tree τ with a root 1 and remaining points labeled with integers $n+2, \cdots, n+m+1$. The contribution of this cluster together with all distinct clusters derived by permutation of labels among points within the rooted star tree and branch may be written

$$V \sum_{\substack{\text{all labelings of } \beta \\ \text{with integers } 2 \cdots n+1}} \int \prod f_{ij} d(2) \cdots d(n+1)$$

$$\times \sum_{\substack{\text{all labelings of } \tau \\ \text{with integers} \\ n+2 \cdots n+m+1}} \int \prod f_{ij} d(n+2) \cdots d(n+m+1). \quad (9)$$

On multiplying by the number of ways of choosing n integers to label the points of the branch, m integers to label the points of the rooted star tree and one integer to label the common root, and summing over all branches and rooted star trees, we obtain for the total superfluous contribution

$$V \sum_{n=1}^{\infty} \sum_{\substack{\text{all} \\ \text{branches}}} \frac{z^n}{n!} \int \prod f_{ij} d(2) \cdots d(n+1)$$

$$\times \left(z + \sum_{m=1}^{\infty} \sum_{\substack{\text{all rooted star trees} \\ \text{with root 1 and other} \\ \text{points } n+2 \cdots n+m+1}} \frac{z^{m+1}}{m!} \right.$$

$$\left. \times \int \prod f_{ij} d(n+2) \cdots d(n+m+1) \right). \quad (10)$$

(The term z in the parentheses corresponds to "rooted star trees of order zero.")

We immediately note that, by (8), the second factor in (10) is simply ρ. We can eliminate z from the first factor by using the representation of a branch implicitly used in the foregoing: every branch can be represented in one and only one way as a rooted star (i.e., a star with a marked-out point) with a number of rooted star trees attached at points other than the root. With the aid of this representation, the first factor in (10) can be represented as a sum over rooted stars:

$$\sum_{n=1}^{\infty} \sum_{\substack{\text{all rooted stars with} \\ \text{root 1 and points } 2 \cdots n+1}} \frac{\rho^n}{n!}$$

$$\times \int \prod f_{ij} d(2) \cdots d(n+1). \quad (11)$$

Finally, we note that (11) is simply $V\rho \partial S / \partial \rho$ and that the total quantity to be subtracted is $V\rho \partial S / \partial \rho$. Thus we have (4) for PV/kT.

Zero-Point Energy of an Electron Lattice*

ROSEMARY A. COLDWELL-HORSFALL
Department of Physics, University of Maryland, College Park, Maryland

AND

ALEXEI A. MARADUDIN†
Institute for Fluid Dynamics and Applied Mathematics, University of Maryland, College Park, Maryland

(Received May 27, 1960)

At very low densities an electron gas in a compensating uniform background of positive charge crystallizes into a bcc lattice for which the correlation energy per electron is $(-1.792/r_s)$ ry. At higher densities the first correction to this result arises from the zero-point energy of the electrons, which can be expanded in terms of the even moments of the frequency spectrum. We have computed the first five nonvanishing moments and have estimated the contribution to the zero-point energy from the remaining moments using their known asymptotic behavior. This procedure leads to the value $(2.638/r_s^2)$ ry for the zero-point energy per electron. The low temperature specific heat per electron is found to be $56.21kr_s^{9/2}\cdot(kT)^3$ ry. The range of r_s values for which these results should be valid is discussed on the basis of Lindemann's melting formula.

I. INTRODUCTION

IN 1934, Wigner, in a study of the effect of electron interactions on the energy levels of the electrons in a metal[1] stated that as the density of an electron gas in a compensating uniform background of positive charge tends to zero, the Coulomb interactions dominate the kinetic energy of the electrons and they arrange themselves in the configuration of lowest potential energy, probably a bcc lattice. In 1938, he presented a quantitative treatment of this problem.[2] He calculated the energy of a lattice of electrons in a uniform background of positive charge on the assumption that the potential at any electron is that due to a uniform sphere of positive charge surrounding it. The radius of this sphere r_s measured in units of the Bohr radius, was chosen to be such that the volume occupied by one electron is $\frac{4}{3}\pi r_s^3$. This simple electrostatic calculation yields an energy per electron which is proportional to r_s^{-1}. A careful calculation by Fuchs[3] of the static energies of monovalent ions arrayed in the three primitive cubic-lattice structures in a compensating uniform background of charge, showed that for a given number density of ions the bcc structure has indeed the lowest energy, with the fcc and simple cubic structures having higher energies in that order. When he took into account the oscillation of the electrons about their lattice points, Wigner obtained a correction to the static energy which is proportional to $r_s^{-\frac{3}{2}}$. The value of this energy given in the text of Wigner's paper is $3r_s^{-\frac{3}{2}}$ ry, which is that obtained in the Einstein approximation in which each electron is assumed to vibrate independently of all the others. However, in a footnote to this paper Wigner mentions that when one takes into account the coupling between the displacement components of the electrons, as in the theories of Debye,[4] and of Born and von-Karman,[5] the value $2.7r_s^{-\frac{3}{2}}$ ry is obtained. Since Wigner's 1938 paper no detailed calculation of the zero-point energy of an electron lattice has been published. At the time of writing, however, we have received a report of the work of Carr[6] in which the value $2.66r_s^{-\frac{3}{2}}$ ry is given for the harmonic contribution to the zero-point energy. Carr obtained the zero-point energy of a bcc lattice of electrons by numerical integration, having first calculated the normal mode frequencies of the electron vibrations at 512 points in the first Brillouin zone.

Here we apply the moment-trace method[7] to the calculation of the zero-point energy of a bcc lattice of electrons in a uniform background of positive charge. In Sec. II we express the zero-point energy as an expansion in all the even moments of the frequency distribution function of the normal modes of vibration, and show how the asymptotic behavior of the moments may be used to sum all terms past the few that can be determined explicitly.

In Sec. III we obtain the dynamical matrix of the lattice from the equations of motion of the electrons. We evaluate the first five even moments of the frequency distribution function in Sec. IV, and we use these results to obtain an estimate of the contribution to the zero-point energy from the remaining moments. In Sec. V we discuss the range of values of r_s for which our result may be expected to be valid, using essentially Lindemann's melting-point formula as a criterion for the stability of the electron lattice.

* This research was supported by the U. S. Air Force through the Air Force Office of Scientific Research, Air Research and Development Command.

† Present address: Westinghouse Research Laboratories, Beulah Road, Churchill Borough, Pittsburgh 35, Pennsylvania.

[1] E. P. Wigner, Phys. Rev. **46**, 1002 (1934).

[2] E. P. Wigner, Trans. Faraday Soc. **34**, 678 (1938).

[3] K. Fuchs, Proc. Roy. Soc. (London) **151**, 585 (1935).

[4] P. Debye, Ann. Physik **39**, 789 (1912).

[5] M. Born and Th. von-Karman, Physik. Z. **13**, 297 (1912); **14**, 15 (1913).

[6] W. J. Carr, Jr., Scientific Paper 6-40601-1-P2, Westinghouse Research Laboratories (December 31, 1959); and private communication.

[7] E. W. Montroll, J. Chem. Phys. **10**, 218 (1942).

II. ZERO-POINT ENERGY

We begin this section by introducing the distribution function $g(\omega)$ for the normal mode frequencies of our electron lattice. $g(\omega)$ is defined in such a way that $g(\omega)d\omega$ is the *fraction* of normal modes with frequencies in the interval $(\omega, \omega+d\omega)$ in the limit as $d\omega \to 0$. The moments μ_n of this distribution function are defined by

$$\mu_n = \int_0^{\omega_p} \omega^n g(\omega) d\omega, \tag{2.1}$$

and the fact that ω_p is the maximum normal mode frequency follows readily from Kohn's sum rule.

The squares of the normal mode frequencies are the eigenvalues of a secular determinant D. It is well known that the trace of the nth power of D is equal to the sum of the nth powers of its eigenvalues, or

$$\mathrm{Tr}D^n = \sum_j \sum_{\mathbf{k}} \omega_j^{2n}(\mathbf{k}), \tag{2.2}$$

where j labels the branches of the frequency spectrum and $\mathbf{k}$ labels the frequencies in each branch. If we divide both sides of this equation by $3N$, where N is the total number of electrons, since there are $3N$ normal mode frequencies, we have

$$\frac{1}{3N} \sum_j \sum_{\mathbf{k}} \omega_j^{2n}(\mathbf{k}) \equiv \mu_{2n} = \frac{1}{3N} \mathrm{Tr}D^n, \tag{2.3}$$

$$E_0 = -\frac{\hbar}{2} \sum_{j=1}^{3} \sum_{\mathbf{k}} \omega_j(\mathbf{k}). \tag{2.4}$$

With the use of the distribution function $g(\omega)$, Eq. (2.4) can be written as

$$E_0 = \frac{3N\hbar}{2} \int_0^{\omega_p} \omega g(\omega) d\omega. \tag{2.5}$$

If we define a dimensionless distribution function $f(x)$ by

$$f(x) = \omega_p g(\omega_p x), \tag{2.6}$$

the zero-point energy per electron can be written as

$$\frac{E_0}{N} = \frac{3\hbar\omega_p}{2} \int_0^1 x f(x) dx. \tag{2.7}$$

In terms of $f(x)$ we can define dimensionless moments u_n by

$$u_n = \frac{\mu_n}{\omega_p{}^n} = \int_0^1 x^n f(x) dx, \tag{2.8}$$

and we see that

$$E_0/N = (3\hbar\omega_p/2) u_1. \tag{2.9}$$

Since the moment-trace method gives only the even moments u_{2n}, we have to express u_1 in terms of the even moments. This is done by writing[8,9]

$$u_1 = \int_0^1 [1-(1-x^2)]^{\frac{1}{2}} f(x) dx$$

$$= \sum_{n=0}^{\infty} (-1)^n \binom{\frac{1}{2}}{n} \int_0^1 (1-x^2)^n f(x) dx. \tag{2.10}$$

We now define new dimensionless moments v_{2n} by

$$v_{2n} = \int_0^1 (1-x^2)^n f(x) dx = \sum_{j=0}^{n} (-1)^j \binom{n}{j} u_{2j}. \tag{2.11}$$

In terms of these moments, u_1 becomes

$$u_1 = v_0 - \tfrac{1}{2}v_2 - \tfrac{1}{8}v_4 - \tfrac{1}{16}v_6 - (5/128)v_8 - \cdots$$

$$= \sum_{n=0}^{\infty} (-1)^n \binom{\frac{1}{2}}{n} v_{2n}. \tag{2.12}$$

Since the v_{2n} are all positive we see that breaking off the expansion (2.12) at any term yields an upper bound for u_1.

In practice, we can calculate only the first few moments u_{2n}, and hence only the first few moments v_{2n}. We must therefore estimate the contribution to E_0/N from the sum

$$\sum_{n=k}^{\infty} (-1)^n \binom{\frac{1}{2}}{n} v_{2n}, \tag{2.13}$$

when all the moments up to v_{2k-2} are known explicitly.

In general, the low-frequency expansion of the frequency distribution function $f(x)$ has the form

$$f(x) = c_2 x^2 + c_4 x^4 + c_6 x^6 + \cdots. \tag{2.14}$$

It has then been shown[8] that the asymptotic behavior of the v_{2n} in the limit of large n is given by

$$v_{2n} \sim \frac{c_2}{2} \frac{\Gamma(\frac{3}{2})\Gamma(n+1)}{\Gamma(n+\frac{5}{2})} + \frac{c_4}{2} \frac{\Gamma(\frac{5}{2})\Gamma(n+1)}{\Gamma(n+\frac{7}{2})} + \frac{c_6}{2} \frac{\Gamma(\frac{7}{2})\Gamma(n+1)}{\Gamma(n+9/2)} + \cdots. \tag{2.15}$$

If this result is substituted into Eq. (2.13), the summations over n can be carried out in closed form, and

[8] C. Domb, A. A. Maradudin, E. W. Montroll, and G. H. Weiss, Phys. Rev. **115**, 24 (1959).
[9] C. Domb and L. Salter, Phil. Mag. **43**, 1083 (1952).

we obtain

$$\sum_{n=k}^{\infty}(-1)^n\binom{\frac{1}{2}}{n}v_{2n}\sim-\frac{c_2}{4(2k-1)(2k+1)}-\frac{c_4}{2(2k-1)(2k+1)(2k+3)}-\frac{15}{8}\frac{c_6}{(2k-1)(2k+1)(2k+3)(2k+5)}-\cdots. \quad (2.16)$$

Our final expression for the zero-point energy per electron becomes

$$\frac{E_0}{N}=\frac{3\hbar\omega_p}{2}\left\{\sum_{n=0}^{k-1}(-1)^n\binom{\frac{1}{2}}{n}v_{2n}-\frac{c_2}{4(2k-1)(2k+1)}-\frac{c_4}{2(2k-1)(2k+1)(2k+3)}-\frac{15}{8}\frac{c_6}{(2k-1)(2k+1)(2k+3)(2k+5)}-\cdots\right\}. \quad (2.17)$$

III. DYNAMICAL MATRIX

We consider a lattice of electrons in a compensating uniform background of positive charge. The equilibrium position of the lth electron relative to an origin at some lattice point is given by the vector

$$\mathbf{r}^l=l_1\mathbf{a}_1+l_2\mathbf{a}_2+l_3\mathbf{a}_3,$$

where the $\{l_i\}$ are integers and $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_3$ are the primitive translation vectors of our lattice.

For small displacements of the electrons about their equilibrium positions, the total potential energy of this system can be expanded as

$$\Phi=\Phi_0-\tfrac{1}{2}\sum_{\substack{ll'\\xy}}\phi_{xy}{}^{ll'}u_x{}^l u_y{}^{l'}+\cdots, \quad (3.1)$$

where $u_x{}^l$ is the x component of displacement of the lth electron, and for central force interactions

$$\phi_{xy}{}^{ll'}=[\partial^2\phi(r)/\partial x\partial y]|_{r=r^{ll'}}, \quad (3.2)$$

with

$$\mathbf{r}^{ll'}=\mathbf{r}^l-\mathbf{r}^{l'}.$$

Furthermore, since the net force on any electron must vanish in a uniform translation of the lattice, we have the additional condition that

$$\phi_{xy}{}^{ll}=-\sum_{l'}{}'\phi_{xy}{}^{ll'}\delta_{xy}, \quad (3.3)$$

where the prime on the summation excludes $l=l'$. In the present problem

$$\phi(r)=\frac{e^2}{r}-ne^2\int\frac{d^3\mathbf{r}'}{|\mathbf{r}-\mathbf{r}'|}, \quad (3.4)$$

$$=\phi^C(r)+\phi^B(r),$$

where n is the number density of electrons and equals $2/a_0{}^3$ for a bcc lattice in terms of the lattice parameter a_0, and we find

$$(\phi^C)_{xy}{}^{ll'}=\frac{e^2}{(r^{ll'})^5}[3x^{ll'}y^{ll'}-(r^{ll'})^2\delta_{xy}], \quad l\neq l' \quad (3.5a)$$

$$(\phi^B)_{xy}{}^{ll'}=-\tfrac{4}{3}\pi ne^2\delta_{xy}\delta_{ll'}. \quad (3.5b)$$

In obtaining Eq. (3.5b) we have used the fact that

$$\nabla^2(1/r)=-4\pi\delta(r).$$

The equations of motion for an electron acting under the potential given in Eq. (3.1) are

$$-\frac{\partial\Phi}{\partial u_x{}^l}=M\ddot{u}_x{}^l=\sum_{\substack{l'\\y}}\phi_{xy}{}^{ll'}u_y{}^{l'}, \quad (x=x,y,z). \quad (3.6)$$

If we assume solutions of the form

$$u_x{}^l(\mathbf{k})=u_x{}^{(0)}\exp[i\omega t-i\mathbf{k}\cdot\mathbf{r}^l], \quad (3.7)$$

Eq. (3.6) becomes

$$-\omega^2u_x{}^{(0)}=\sum_y\left\{\frac{1}{M}\sum\phi_{xy}{}^{ll'}\exp[i\mathbf{k}\cdot(\mathbf{r}^l-\mathbf{r}^{l'})]\right\}u_y{}^0,$$

$$=-\sum_y D_{xy}(\mathbf{k})u_y{}^0, \quad (3.8)$$

where the $D_{xy}(\mathbf{k})$ are obtained from Eqs. (3.8) and (3.5) as

$$D_{xy}(\mathbf{k})=\frac{4\pi ne^2}{3M}\delta_{xy}+\frac{e^2}{M}\sum_{l'}{}'\times\frac{[3x^{ll'}y^{ll'}-(r^{ll'})^2\delta_{xy}][\delta_{xy}-\exp(i\mathbf{k}\cdot\mathbf{r}^{ll'})]}{(r^{ll'})^5}. \quad (3.9)$$

We assume that the displacement components $u_x{}^l$ satisfy the Born-von Karman cyclic boundary condition, and this implies that the allowed values of the wave vector $\mathbf{k}$ are uniformly distributed throughout the first Brillouin zone with a density $\Omega/(2\pi)^3$, where Ω is the normalization volume for our lattice. It is convenient to define new matrix elements $C_{xy}(\mathbf{k})$ such that

$$C_{xy}(\mathbf{k})=D_{xy}(\mathbf{k})-\tfrac{1}{3}\omega_p{}^2\delta_{xy}, \quad (3.10)$$

where ω_p is the classical plasma frequency,

$$\omega_p{}^2=(4\pi ne^2)/M.$$

The condition that the set of Eqs. (3.8) has a nontrivial solution is that the determinant of the coefficients vanishes. In the present notation this condition becomes

$$|C-I(\omega^2-\tfrac{1}{3}\omega_p{}^2)|\equiv|C-I\lambda|=0. \quad (3.11)$$

It is immediately obvious that $\mathrm{Tr}C=0$ for all $\mathbf{k}$, by Laplace's equation, so we obtain the sum rule previously derived by Kohn,[10]

$$\sum_{j=1}^{3}\omega_j^2(\mathbf{k})=\omega_p^2. \tag{3.12}$$

IV. CALCULATION OF THE MOMENTS OF THE FREQUENCY DISTRIBUTION FUNCTION

In Sec. II we saw that the zero-point energy of our lattice can be expanded in terms of the even dimensionless moments of the frequency distribution function, and that these moments are related to the traces of powers of the dynamical matrix $\mathbf{D}$ by

$$u_{2n}=\frac{1}{3N\omega_p^{2n}}\mathrm{Tr}\mathbf{D}^n=\frac{1}{3N\omega_p^{2n}}\sum_{j=1}^{3}\sum_{\mathbf{k}}[D^n(\mathbf{k})]_{jj}. \tag{4.1}$$

In practice, however, it is more convenient to work with the matrix $\mathbf{C}(\mathbf{k})$ defined by Eq. (3.10). We must thus relate the moments u_{2n} to the traces of powers of $\mathbf{C}(\mathbf{k})$. This we do as follows. We write

$$\mathrm{Tr}C^n=\sum_{j=1}^{3}\sum_{\mathbf{k}}\lambda_j^n(\mathbf{k}), \tag{4.2}$$

where

$$\lambda_j(\mathbf{k})=\omega_j^2(\mathbf{k})-\tfrac{1}{3}\omega_p^2. \tag{4.3}$$

The determinantal equation (3.11) can be expanded in the form

$$\lambda^3-a\lambda^2+b\lambda-c=0, \tag{4.4}$$

where the coefficients a, b, c are functions of $\mathbf{k}$ which are given explicitly by

$$\begin{aligned}a&=C_{11}+C_{22}+C_{33},\\ b&=C_{11}C_{22}+C_{22}C_{33}+C_{33}C_{11}-C_{12}C_{21}\\ &\qquad -C_{23}C_{32}-C_{31}C_{13},\\ c&=C_{11}C_{22}C_{33}+2C_{12}C_{23}C_{31}-C_{11}C_{23}C_{32}\\ &\qquad -C_{22}C_{31}C_{13}-C_{33}C_{12}C_{21}.\end{aligned} \tag{4.5}$$

We now introduce a function $S_n(\mathbf{k})$ by the relation

$$S_n(\mathbf{k})=\lambda_1^n(\mathbf{k})+\lambda_2^n(\mathbf{k})+\lambda_3^n(\mathbf{k}), \tag{4.6}$$

where the λ_j are the three roots of the secular equation for a given value of $\mathbf{k}$. With the aid of known relations[11] between the sums of powers of the roots of an algebraic equation and the coefficients in the equation, and Eq. (4.4), we find that the functions S_n satisfy the following recurrence relation:

$$S_n=aS_{n-1}-bS_{n-2}+cS_{n-3}.$$

Thus the first five S_n are

$$\begin{aligned}S_0&=3,\\ S_1&=a=0,\\ S_2&=a^2-2b=-2b,\\ S_3&=a^3-3ab+3c=3c,\\ S_4&=a^4-4a^2b+4ac+2b^2=2b^2,\end{aligned} \tag{4.7}$$

where we have used the fact that a vanishes as was indicated at the end of the preceding section. It is the simplification of the expressions for the S_n expressed by the second column in Eq. (4.7) which prompted the use of the matrix $\mathbf{C}(\mathbf{k})$. With the aid of Eqs. (4.3) and (4.6), we can write $S_n(k)$ as

$$\begin{aligned}S_n(\mathbf{k})&=\sum_{j=1}^{3}(\omega_j^2-\tfrac{1}{3}\omega_p^2)^n\\ &=\sum_i\Big\{\omega_j^{2n}-n\omega_j^{2n-2}\tfrac{1}{3}\omega_p^2+\frac{n(n-1)}{2!}\omega_j^{2n-4}\frac{1}{9}\omega_p^4\\ &\quad-\frac{n(n-1)(n-2)}{3!}\omega_j^{2n-6}\frac{1}{27}\omega_p^6\\ &\quad+\frac{n(n-1)(n-2)(n-3)}{4!}\omega_j^{2n-8}\frac{1}{81}\omega_p^8\\ &\quad-\frac{n(n-1)(n-2)(n-3)(n-4)}{5!}\\ &\qquad\times\omega_j^{2n-10}\frac{1}{243}\omega_p^{10}+\cdots\Big\},\end{aligned}$$

and if we now take the average value of S_n over all wave vectors $\mathbf{k}$, we obtain

$$\begin{aligned}S_n=\langle S_n\rangle&=\frac{1}{3N}\sum_{\mathbf{k}}\frac{S_n(\mathbf{k})}{\omega_p^{2n}}\\ &=u_{2n}-\frac{n}{3}u_{2n-2}+\frac{n(n-1)}{2!}\frac{1}{3^2}u_{2n-4}\\ &\quad-\frac{n(n-1)(n-2)}{3!}\frac{1}{3^3}u_{2n-6}\\ &\quad+\frac{n(n-1)(n-2)(n-3)}{4!}\frac{1}{3^4}u_{2n-8}\\ &\quad-\frac{n(n-1)(n-2)(n-3)(n-4)}{5!}\frac{1}{3^5}u_{2n-10}+\cdots.\end{aligned} \tag{4.8}$$

[10] W. Kohn (unpublished work); see R. Brout, Phys. Rev. **113**, 43 (1959), footnote reference 4.

[11] E. P. Adams, *Smithsonian Mathematical Formulae and Tables of Elliptic Functions* (The Smithsonian Institution, Washington, D. C., 1922).

Therefore we can express the even moments u_{2n} in terms of the S_n by simple inversion of the Eqs. (4.8); thus,

$$\begin{aligned} u_0&=S_0,\\ u_2&=\tfrac{1}{3}S_0+S_1,\\ u_4&=\tfrac{1}{9}S_0+\tfrac{2}{3}S_1+S_2,\\ u_6&=\frac{1}{27}S_0+\frac{3}{9}S_1+\frac{3}{3}S_2+S_3,\\ u_8&=\frac{1}{81}S_0+\frac{4}{27}S_1+\frac{6}{9}S_2+\frac{4}{3}S_3+S_4, \end{aligned} \tag{4.9}$$

and so on.

The first nontrivial summation is that for S_2. From Eqs. (4.8), (4.9), and (3.9), we have

$$\begin{aligned} S_2&=-\frac{2}{3N\omega_p{}^4}\sum_{\mathbf{k}}[3C_{11}C_{22}-3C_{12}C_{21}]\\ &=-\frac{2\cdot 3\cdot e^4}{3N\omega_p{}^4M^2}\sum_{\mathbf{k}}\sum_{\mathbf{r}_l,\mathbf{r}_m}{}'\\ &\quad\times\frac{[(3x_l^2-r_l^2)(3y_m^2-r_m^2)-9x_ly_lx_my_m]}{r_l{}^5r_m{}^5}\\ &\quad\times\exp(i\mathbf{k}\cdot\mathbf{r}^l+\mathbf{r}^m). \end{aligned} \tag{4.10}$$

We first perform the summation over all wave vectors and use the result that

$$\sum_{\mathbf{k}}\exp(i\mathbf{k}\cdot\mathbf{r})=N\delta_{\mathbf{r},0}; \tag{4.11}$$

thus the number of summations over the lattice vectors is reduced by unity, and we obtain the simple sum

$$S_2=-\frac{2e^4}{\omega_p{}^4M^2}\sum_{\mathbf{r}_l}{}'\left(-\frac{1}{r_l{}^6}\right). \tag{4.12}$$

This sum is tabulated by Misra[12] and Ingham and Jones[13] and has the value $0.45383(2/a_0)^6$ for the bcc lattice; therefore,

$$S_2=0.091965. \tag{4.13}$$

The succeeding S_n become increasingly difficult to evaluate: S_3 involves a double sum over all lattice vectors and S_4 a triple sum, after the summation over wave vectors has been carried out (see Appendix B).

The values we obtain for the S_n are

$$S_3=7.3542\times10^{-3},\quad S_4=1.6630\times10^{-3}. \tag{4.14}$$

The values of the u_{2n} and the v_{2n} defined in Eqs. (4.9) and (2.11), respectively, become

$$\begin{aligned} u_0&=1.000000 & v_0&=1.000000\\ u_2&=0.333333 & v_2&=0.666667\\ u_4&=0.203076 & v_4&=0.536409\\ u_6&=0.136356 & v_6&=0.472872\\ u_8&=0.109816 & v_8&=0.449514. \end{aligned} \tag{4.15}$$

We substitute these values of the v_{2n} into Eq. (2.12) and obtain the contribution to the zero-point energy per electron from the first five moments:

$$E_0/N=\tfrac{3}{2}\hbar\omega_p(0.552503). \tag{4.16}$$

The contribution to E_0/N from the remaining terms in the expansion is given by Eq. (2.16), viz.,

$$\begin{aligned} \frac{\Delta E_0}{N}=\frac{3\hbar\omega_p}{2}\Bigg\{&-\frac{1}{4}\frac{c_2}{(2k-1)(2k+1)}\\ &-\frac{1}{2}\frac{c_4}{(2k-1)(2k+1)(2k+3)}\\ &-\frac{15}{8}\frac{c_6}{(2k-1)(2k+1)(2k+3)(2k+5)}\Bigg\}. \end{aligned} \tag{4.17}$$

Values of c_2 and c_4 were obtained by substituting the values given in Eq. (4.15) for v_6 and v_8 into Eq. (2.15), and we find that

$$c_2=27.892,\quad c_4=-68.136. \tag{4.18}$$

The extrapolated value of the zero-point energy is

$$E_0/N=\tfrac{3}{2}\hbar\omega_p(0.5085)=2.642/r_s{}^{\frac{1}{2}}\text{ ry}. \tag{4.19}$$

A more accurate value of c_2 is calculated in Appendix C:

$$c_2=29.984. \tag{4.20}$$

With this result and the values of v_6 and v_8 given by Eq. (4.15), we find for the coefficients c_4 and c_6 the values

$$c_4=-86.264,\quad c_6=27.192. \tag{4.21}$$

The extrapolated value of E_0/N obtained using these c_{2n} is

$$E_0/N=\tfrac{3}{2}\hbar\omega_p(0.5077)=2.638/r_s{}^{\frac{1}{2}}\text{ ry}. \tag{4.22}$$

These results are in good agreement with Carr's result, $2.66/r_s{}^{\frac{1}{2}}$ ry.

V. STABILITY OF THE ELECTRON LATTICE

The success of Lindemann's melting formula[14] when applied to ordinary solids led us to use the formula to find a criterion for the stability of the electron lattice. Lindemann stated that melting occurs when the mean amplitude of vibration of a particle about its lattice position is greater than some fraction δ of the inter-

[12] R. D. Misra, Proc. Cambridge Phil. Soc. **36**, 173 (1940).
[13] A. E. Ingham and J. E. Jones, Proc. Roy. Soc. (London) **107**, 636 (1925).
[14] F. A. Lindemann, Physik. Z. **11**, 609 (1910).

particle spacing. The mean amplitude of vibration is determined as follows.

We expand the α component of the displacement of the lth electron in terms of the normal coordinates of the lattice

$$u_\alpha{}^l=\frac{1}{(NM)^{\frac{1}{2}}}\sum_{\mathbf{k},j}Q_j(\mathbf{k})e_\alpha{}^j(\mathbf{k})\exp(i\mathbf{k}\cdot\mathbf{r}^l), \tag{5.1}$$

where the $\{e_\alpha{}^j(\mathbf{k})\}$ are the components of the eigenvectors of the matrix $[D_{xy}(\mathbf{k})]$. Since every lattice point is equivalent to every other one when the cyclic boundary condition is assumed, and since the displacement components are real, we have the identity

$$\langle(u^l)^2\rangle=\frac{1}{N}\sum_{\alpha,l}u_\alpha{}^l(u_\alpha{}^l)^*$$

$$=\sum_\alpha\frac{1}{NM}\sum_{\mathbf{k},j,j'}Q_j(\mathbf{k})Q_{j'}{}^*(\mathbf{k})e_\alpha{}^{*j'}(\mathbf{k})e_\alpha{}^j(\mathbf{k}), \tag{5.2}$$

so that

$$\langle u^2\rangle=\frac{1}{NM}\sum_{\mathbf{k},j}Q_j(\mathbf{k})Q_j{}^*(\mathbf{k}). \tag{5.3}$$

In evaluating the thermal average of $\langle u^2\rangle$, we thus require the thermal average of $Q_j(\mathbf{k})Q_j{}^*(\mathbf{k})$. This is the well-known expression

$$\langle Q_j(\mathbf{k})Q_j{}^*(\mathbf{k})\rangle_{\text{th}}=[\hbar/2\omega_j(\mathbf{k})]\coth[\hbar\omega_j(\mathbf{k})/2kT], \tag{5.4}$$

and the mean square displacement at temperature T is therefore given by

$$\langle u^2\rangle_{\text{av}}=\frac{1}{NM}\frac{\hbar}{2}\sum_{\mathbf{k},j}\frac{1}{\omega_j(\mathbf{k})}\coth\frac{\hbar\omega_j(\mathbf{k})}{2kT}. \tag{5.5}$$

In terms of the frequency distribution function this result becomes

$$\langle u^2\rangle_{\text{av}}=\frac{3\hbar}{2M}\int_0^{\omega_p}\frac{g(\omega)}{\omega}\coth\left(\frac{\hbar\omega}{2kT}\right)d\omega. \tag{5.6}$$

In the limit as $T\to 0$, we obtain

$$\langle u^2\rangle_{\text{av}}=\frac{3\hbar}{2M}\int_0^{\omega_p}\frac{g(\omega)}{\omega}d\omega$$

$$=\frac{3\hbar}{2M\omega_p}u_{-1} \tag{5.7a}$$

$$=\frac{3\hbar}{2M\omega_p}\left[v_0+\frac{1}{2}v_2+\frac{3}{8}v_4+\frac{5}{16}v_6+\frac{35}{128}v_8+\cdots\right]. \tag{5.7b}$$

Lindemann's criterion can be stated as

$$\langle u^2\rangle_{\text{av critical}}=\delta^2r_0{}^2, \tag{5.8}$$

where r_0 is the mean particle spacing. For the bcc lattice r_0 is $\sqrt{3}a_0/2$. On using the values of v_{2n} given in Eq. (4.15), we obtain the critical value of r_s,

$$\begin{aligned} r_s&= 0.4054\,\delta^{-4}\\ &= 6.4857 \quad \text{for } \delta=\tfrac{1}{2},\\ &=103.771 \quad \text{for } \delta=\tfrac{1}{4}. \end{aligned} \tag{5.9}$$

Correction terms to the expansion in Eq. (5.7) in terms of c_2, c_4, $\cdots$ can be obtained in just the same way as was used in Sec. II. However, in view of the arbitrariness in the choice of δ, a higher degree of accuracy than is represented by Eq. (5.7) seemed pointless.

The value $r_s\sim 20$ given by Nozières and Pines[15] for $\delta=\frac{1}{4}$ is calculated on the assumption that only longitudinal plasmons can contribute to the lattice vibrations, and certainly is an overestimate of the stability of the lattice, as they expect. However, the real problem in this calculation is the choice of a reasonable value for δ.

Finally, the fact that the electrons in a compensating uniform background of positive charge crystallize into a lattice at some critical value of the number density suggests that in the absence of detailed knowledge regarding the nature of this phase transition, the extrapolation of the low density result for the correlation energy to the region of metallic densities, at least by any simple procedure, is likely to be in error.

APPENDIX A

In order to make this paper self-contained, we present an "elementary" calculation of the potential energy of an electron lattice for each of the three primitive cubic-lattice structures. It is elementary in that Ewald's generalized theta-function transformation[16] is not employed.

For each of the lattice structures the potential energy of interaction of a single electron with all other electrons and with the uniform background of positive charge is given by

$$U=U_1+U_2, \tag{A1}$$

where

$$U_1=e^2{\sum_l}'\frac{1}{(x_l{}^2+y_l{}^2+z_l{}^2)^{\frac{1}{2}}}, \tag{A2}$$

$$U_2=-ne^2\int\int\int_{(\Omega)}\frac{dxdydz}{(x^2+y^2+z^2)^{\frac{1}{2}}}. \tag{A3}$$

The prime on the summation in Eq. (A2) means that the point $x_l=y_l=z_l=0$ is excluded. n is the number density of electrons, and Ω is the normalization volume for our lattices.

We consider the sum U_1 first. The vector $\mathbf{x}_l=(x_l,y_l,z_l)$ takes the following forms for the three primitive cubic

[15] P. Nozières and D. Pines, Phys. Rev. **111**, 442 (1958).
[16] P. P. Ewald, Ann. Physik **64**, 253 (1921).

lattices:

$$\begin{aligned}&\text{sc:}\quad \mathbf{x}_l=a_0(l_1,l_2,l_3) \qquad l_1,\ l_2,\ l_3 \text{ unrestricted integers,}\\ &\text{fcc:}\quad \mathbf{x}_l=(a_0/2)(l_1,l_2,l_3) \qquad l_1+l_2+l_3=\text{even integers,}\\ &\text{bcc:}\quad \mathbf{x}_l=(a_0/2)(l_1,l_2,l_3) \quad l_1,\ l_2,\ l_3 \text{ all even or all odd.}\end{aligned} \tag{A4}$$

With the aid of the relation

$$\frac{1}{z^k}=\frac{1}{\Gamma(k)}\int_0^\infty t^{k-1}e^{-zt}dt, \tag{A5}$$

we rewrite Eq. (A2) as

$$U_1=\frac{\zeta}{\pi^{\frac{1}{2}}}\frac{e^2}{a_0}\sum_{l_1l_2l_3}{}'\int_0^\infty t^{-\frac{1}{2}}\exp[-(l_1^2+l_2^2+l_3^2)t]dt, \tag{A6}$$

where ζ is 1 for the sc lattice and 2 for the fcc and bcc lattices. We now break up the range of integration into two parts $(0,\epsilon)$ and (ϵ,∞), where the choice of ϵ will be deferred to a later point in the calculations. Thus we define

$$U_{11}=\frac{\zeta}{\pi^{\frac{1}{2}}}\frac{e^2}{a_0}\sum_{l_1l_2l_3}{}'\int^\infty t^{-\frac{1}{2}}\exp[-(l_1^2+l_2^2+l_3^2)t]dt, \tag{A7}$$

$$U_{12}=\frac{\zeta}{\pi^{\frac{1}{2}}}\frac{e^2}{a_0}\sum_{l_1l_2l_3}{}'\int_0^\epsilon t^{-\frac{1}{2}}\exp[-(l_1^2+l_2^2+l_3^2)t]dt. \tag{A8}$$

In terms of the auxiliary integrals[12]

$$\phi_m(x)=\int_1^\infty t^m e^{-xt}dt, \tag{A9}$$

Eq. (A7) can be written as

$$U_{11}=\zeta\left(\frac{\epsilon}{\pi}\right)^{\frac{1}{2}}\frac{e^2}{a_0}\sum_{l_1l_2l_3}{}'\phi_{-\frac{1}{2}}[\epsilon(l_1^2+l_2^2+l_3^2)]. \tag{A10}$$

In the sum U_{12} we interchange the order of summation and integration and remove the restriction on the sum to obtain

$$\begin{aligned}U_{12}&=\frac{\zeta}{\pi^{\frac{1}{2}}}\frac{e^2}{a_0}\int_0^\epsilon t^{-\frac{1}{2}}\{\sum_{l_1l_2l_3}\exp[-(l_1^2+l_2^2+l_3^2)t]-1\}dt\\ &=\frac{\zeta}{\pi^{\frac{1}{2}}}\frac{e^2}{a_0}\int_0^\epsilon t^{-\frac{1}{2}}\sigma(t)dt-2\zeta\left(\frac{\epsilon}{\pi}\right)^{\frac{1}{2}}\frac{e^2}{a_0},\end{aligned} \tag{A11}$$

where we have put

$$\sigma(t)=\sum_{l_1l_2l_3}\exp[-(l_1^2+l_2^2+l_3^2)t]. \tag{A12}$$

For the three lattice types, $\sigma(t)$ takes the following forms[12]:

$$\text{sc:}\quad \sigma(t)=\sum_{l_1l_2l_3}\exp[-(l_1^2+l_2^2+l_3^2)t]=\{\sum_{l=-\infty}^{\infty}\exp(-l^2t)\}^3, \tag{A13a}$$

$$\text{fcc:}\quad \sigma(t)=\sum_{\substack{l_1l_2l_3\\ l_1+l_2+l_3 \text{ even}}}\exp[-(l_1^2+l_2^2+l_3^2)t]=\{\sum_{l=-\infty}^{\infty}\exp(-4l^2t)\}^3+3\sum_{l=-\infty}^{\infty}\exp(-4l^2t)\{\sum_{l=-\infty}^{\infty}\exp[-4(l-\tfrac{1}{2})^2t]\}^2, \tag{A13b}$$

$$\text{bcc:}\quad \sigma(t)=\sum_{\substack{l_1l_2l_3\\ \text{all even, all odd}}}\exp[-(l_1^2+l_2^2+l_3^2)t]=\{\sum_{l=-\infty}^{\infty}\exp(-4l^2t)\}^3+\{\sum_{l=-\infty}^{\infty}\exp[-4(l-\tfrac{1}{2})^2t]\}^3. \tag{A13c}$$

However, we employ the following transformations[17]:

$$\sum_{l=-\infty}^{\infty}\exp(-l^2t)=\left(\frac{\pi}{t}\right)^{\frac{1}{2}}\sum_{l=-\infty}^{\infty}\exp\left(-\frac{\pi^2}{t}l^2\right), \tag{A14a}$$

$$\sum_{l=-\infty}^{\infty}\exp[-4(l-\tfrac{1}{2})^2t]=\frac{1}{2}\left(\frac{\pi}{t}\right)^{\frac{1}{2}}\sum_{l=-\infty}^{\infty}(-1)^l\exp\left(-\frac{\pi^2}{4t}l^2\right), \tag{A14b}$$

and the auxiliary integrals $\phi_n(x)$ to rewrite U_{12} as

$$\text{sc:}\quad U_{12}=-2\left(\frac{\epsilon}{\pi}\right)^{\frac{1}{2}}\frac{e^2}{a_0}+\frac{\pi}{\epsilon}\frac{e^2}{a_0}\sum_{l_1l_2l_3}{}'\phi_0\left[\frac{\pi^2}{\epsilon}(l_1^2+l_2^2+l_3^2)\right]+\pi\frac{e^2}{a_0}\int_0^\epsilon\frac{dt}{t^2}, \tag{A15a}$$

$$\text{fcc:}\quad U_{12}=-4\left(\frac{\epsilon}{\pi}\right)^{\frac{1}{2}}\frac{e^2}{a_0}+2\frac{\pi}{\epsilon}\frac{e^2}{a_0}\left[4\phi_0\left(\frac{3\pi^2}{4\epsilon}\right)+3\phi_0\left(\frac{\pi^2}{\epsilon}\right)+6\phi_0\left(\frac{2\pi^2}{\epsilon}\right)+\cdots\right]+\pi\frac{e^2}{a_0}\int_0^\epsilon\frac{dt}{t^2}, \tag{A15b}$$

$$\text{bcc:}\quad U_{12}=-4\left(\frac{\epsilon}{\pi}\right)^{\frac{1}{2}}\frac{e^2}{a_0}+\frac{3\pi}{\epsilon}\frac{e^2}{a_0}\left[2\phi_0\left(\frac{\pi^2}{2\epsilon}\right)+\phi_0\left(\frac{\pi^2}{\epsilon}\right)+4\phi_0\left(\frac{3\pi^2}{\epsilon}\right)+2\phi_0\left(\frac{2\pi^2}{\epsilon}\right)+\cdots\right]+\frac{1}{2}\pi\frac{e^2}{a_0}\int_0^\epsilon\frac{dt}{t^2}. \tag{A15c}$$

[17] E. Whittaker and G. N. Watson, *Modern Analysis* (Cambridge University Press, New York, 1952), 4th ed., p. 474.

The parameter ϵ is now chosen so as to ensure equal rates of convergence in the two sums U_{11} and U_{12}. Since for large x

$$\phi_m(x)\sim(e^{-x}/x),$$

in the present case the following choices were made:

$$\epsilon_{sc}=\pi;\quad \epsilon_{fcc}=\pi/2;\quad \epsilon_{bcc}=\pi/2. \tag{A16}$$

We now turn to the energy U_2. The number density n in each of the three cases is

$$\begin{aligned}&\text{sc:}\quad n=1/a_0^3,\\ &\text{fcc:}\quad n=4/a_0^3,\\ &\text{bcc:}\quad n=2/a_0^3.\end{aligned} \tag{A17}$$

We must evaluate U_2 in the same representation that we used for U_1 in order that the divergence which occurs in the evaluation of U_{12}, Eq. (A15), be canceled by a corresponding divergence in U_1, as is required by the charge neutrality of our system. To this end we make a change of variables,

$$x_j=(a_0/\zeta)u_j, \tag{A18}$$

and rewrite Eq. (A.3) as

$$\begin{aligned}U_2&=-ne^2\left(\frac{a_0}{\zeta}\right)^2\int\int\int\frac{dudvdw}{(u^2+v^2+w^2)^{\frac{1}{2}}}\\&=-\frac{ne^2}{\pi^{\frac{1}{2}}}\left(\frac{a_0}{\zeta}\right)^2\int\int\int\int_0^\infty t^{-\frac{1}{2}}\\&\quad\times\exp[-(u^2+v^2+w^2)t]dtdudvdw.\end{aligned}$$

Having made this transformation we now extend the limits of the integrals over u, v, w to $\pm\infty$. We interchange the order of integration and find

$$U_2=-ne^2\pi\left(\frac{a_0}{\zeta}\right)^2\int_0^\infty\frac{dt}{t^2}. \tag{A19}$$

Combining U_{11}, U_{12}, and U_2, and substituting the values of ζ and ϵ which are given by Eq. (A16), we obtain

$$U_{sc}=\frac{e^2}{a_0}[6\phi_{-\frac{1}{2}}(\pi)+12\phi_{-\frac{1}{2}}(2\pi)+8\phi_{-\frac{1}{2}}(3\pi)+6\phi_{-\frac{1}{2}}(4\pi)+\cdots]$$
$$+\frac{e^2}{a_0}\left[6\phi_0(\pi)+12\phi_0(2\pi)+8\phi_0(3\pi)+6\phi_0(4\pi)+\cdots\right]-3\frac{e^2}{a_0}, \tag{A20a}$$

$$U_{fcc}=\sqrt{2}\frac{e^2}{a_0}[12\phi_{-\frac{1}{2}}(\pi)+6\phi_{-\frac{1}{2}}(2\pi)+24\phi_{-\frac{1}{2}}(3\pi)+12\phi_{-\frac{1}{2}}(4\pi)+\cdots]$$
$$+4\frac{e^2}{a_0}[4\phi_0(3\pi/2)+3\phi_0(2\pi)+6\phi_0(4\pi)+\cdots]-2(1+\sqrt{2})\frac{e^2}{a_0}, \tag{A20b}$$

$$U_{bcc}=\sqrt{2}\frac{e^2}{a_0}[8\phi_{-\frac{1}{2}}(3\pi/2)+6\phi_{-\frac{1}{2}}(2\pi)+12\phi_{-\frac{1}{2}}(4\pi)+\cdots]$$
$$+6\frac{e^2}{a_0}[2\phi_0(\pi)+\phi_0(2\pi)+4\phi_0(3\pi)+2\phi_0(4\pi)+\cdots]-(2\sqrt{2}+1)\frac{e^2}{a_0}. \tag{A20c}$$

Tables of $\phi_n(x)$ have been prepared by Misra[12] and by Born and Misra.[18] With the aid of these the sums are readily evaluated, and we finally obtain

$$\begin{aligned}U_{sc}&=-(2e^2/a_0)(1.4186488)\\&=-2(1.7601188)(1/r_s)\text{ ry},\\U_{fcc}&=-(2e^2/a_0)(2.2924378)\\&=-2(1.791753)(1/r_s)\text{ ry},\\U_{bcc}&=-(2e^2/a_0)(1.819620)\\&=-2(1.791860)(1/r_s)\text{ ry}.\end{aligned} \tag{A21}$$

These values represent the interaction energy of each electron with its surroundings. The potential energy per electron is half of this interaction energy.

[18] M. Born and R. D. Misra, Proc. Cambridge Phil. Soc. **36**, 466 (1940).

APPENDIX B

From Eqs. (4.5), (4.7), (4.8), and (3.9), we have

$$\begin{aligned}S_3&=\frac{1}{3N}\sum_{\mathbf{k}}\frac{3c}{\omega_p^6}\\&=\frac{1}{N\omega_p^6}\sum_{\mathbf{k}}[C_{11}C_{22}C_{33}+2C_{12}C_{23}C_{31}-3C_{11}C_{23}^2]\\&=-\frac{e^6}{NM^3\omega_p^6}\frac{N}{\sum_{r_l,r_m,r_n}}'[(3x_l^2-r_l^2)(3y_m^2-r_m^2)(3z_n^2-r_n^2)\\&\quad+2\cdot3^3x_ly_lz_my_mz_nx_n-3^3(3x_l^2-r_l^2)y_mz_my_nz_n]\\&\quad\times\frac{1}{r_l^5r_m^5r_n^5}\delta_{r_l+r_m+r_n,0}\\&=(e^6/M^3\omega_p^6)(2/a_0)^9\sum\\&=\sum/\pi^3,\end{aligned} \tag{B1}$$

where $\sum$ is the sum in Eq. (B1) evaluated for the values of $\mathbf{n}_l$ from $(3)^{\frac{1}{2}}$ to $(27)^{\frac{1}{2}}$, and for $\mathbf{n}_m$ from $(3)^{\frac{1}{2}}$ to $(20)^{\frac{1}{2}}$; $\mathbf{n}_j$ is a vector with integer components which are all even or all odd [see Appendix A, Eq. (A4)]. We obtain the value

$$\sum = 228.02\times 10^{-3}.$$

Hence

$$S_3 = 7.3542\times 10^{-3}. \tag{B2}$$

Similarly, we have

$$S_4 = \frac{2}{3N\omega_p{}^8}\sum_{\mathbf{k}} b^2, \tag{B3}$$

where

$$\begin{aligned}\sum_{\mathbf{k}} b^2 = 3\left(\frac{e^2}{M}\right)^4 N \sum_{\mathbf{r}_l,\mathbf{r}_m,\mathbf{r}_n,\mathbf{r}_p}{}' \{&(3x_l{}^2-r_l{}^2)(3x_m{}^2-r_m{}^2)\\ &\times(3y_n{}^2-r_n{}^2)(3y_p{}^2-r_p{}^2)+2(3x_l{}^2-r_l{}^2)\\ &\times(3y_m{}^2-r_m{}^2)(3y_n{}^2-r_n{}^2)(3z_p{}^2-r_p{}^2)\\ &+3^2[3^2x_ly_lx_my_m-2(3x_l{}^2-r_l{}^2)(3y_m{}^2-r_m{}^2)]\\ &\times[x_ny_nx_py_p+y_nz_ny_pz_p+z_nx_nz_px_p]\}\\ &\times\delta_{\mathbf{r}_l+\mathbf{r}_m+\mathbf{r}_n+\mathbf{r}_p,\,0}/r_l{}^5r_m{}^5r_n{}^5r_p{}^5\\ =3(e^2/M)^4N(2/a_0)^{12}\cdot\sum,&\end{aligned}$$

where, in this case,

$$\sum = \sum_{n_l{}^2=3}^{8}\sum_{n_m{}^2=3}^{8}\sum_{n_p{}^2=3}^{8} = 80.998\times 10^{-3}.$$

Therefore, we obtain

$$\begin{aligned}S_4 &= (2/\pi^4)\sum\\ &= 1.663\times 10^{-3}.\end{aligned} \tag{B4}$$

APPENDIX C

We have remarked in Sec. II that it is a general result for three-dimensional lattices that the low-frequency expansion of the frequency distribution function $f(x)$ has the form

$$f(x)\sim c_2x^2+c_4x^4+c_6x^6+\cdots. \tag{C1}$$

Here we obtain an accurate value of the coefficient c_2 for our model.

We begin by writing down the small $\mathbf{k}$ expansions of the elements $D_{xy}(\mathbf{k})$ of the 3×3 dynamical matrix defined by Eq. (3.9)[19]:

$$\begin{aligned}D_{xy}(\mathbf{k}) &= \tfrac{1}{3}\omega_p{}^2\delta_{xy}-\frac{e^2}{M}\sum_l{}'\frac{3x_ly_l-r_l{}^2\delta_{xy}}{r_l{}^5}e^{i\mathbf{k}\cdot\mathbf{r}_l},\\ D_{xx}(\mathbf{k}) &\underset{|\mathbf{k}|\to 0}{\longrightarrow} \omega_p{}^2\frac{k_x{}^2}{k^2}-\omega_p{}^2\frac{a_0{}^2k^2}{16\pi}\left[0.849+0.637\frac{k_x{}^2}{k^2}-1.591\frac{k_y{}^2+k_z{}^2}{k^2}\right],\\ D_{xy}(\mathbf{k}) &\underset{|\mathbf{k}|\to 0}{\longrightarrow} \omega_p{}^2\frac{k_xk_y}{k^2}-\omega_p{}^2\frac{a_0{}^2k^2}{16\pi}\left[0.942\frac{k_xk_y}{k^2}\right].\end{aligned} \tag{C2}$$

The secular equation has been solved along six directions in $\mathbf{k}$ space, viz., the [100], [110], [111], [210], [211], [221] directions. The results are as follows ($x=\omega/\omega_p$, $\gamma=a_0{}^2/16\pi$):

$$\begin{aligned}[100]:\quad & x^2=\gamma(0.742)k^2 \quad \text{(twice)}\\ &\quad =1-\gamma(1.486)k^2;\end{aligned} \tag{C3a}$$

$$\begin{aligned}[110]:\quad & x^2=\gamma(0.742)k^2\\ &\quad =\gamma(0.099)k^2\\ &\quad =1-\gamma(0.843)k^2;\end{aligned} \tag{C3b}$$

$$\begin{aligned}[111]:\quad & x^2=\gamma(0.313)k^2 \quad \text{(twice)}\\ &\quad =1-\gamma(0.629)k^2;\end{aligned} \tag{C3c}$$

$$\begin{aligned}[210]:\quad & x^2=\gamma(0.742)k^2\\ &\quad =\gamma(0.363)k^2\\ &\quad =1-\gamma(1.107)k^2;\end{aligned} \tag{C3d}$$

$$\begin{aligned}[211]:\quad & x^2=\gamma(0.528)k^2\\ &\quad =\gamma(0.351)k^2\\ &\quad =1-\gamma(0.880)k^2;\end{aligned} \tag{C3e}$$

$$\begin{aligned}[221]:\quad & x^2=\gamma(0.528)k^2\\ &\quad =\gamma(0.170)k^2\\ &\quad =1-\gamma(0.700)k^2.\end{aligned} \tag{C3f}$$

Only the two acoustic branches contribute to the low-frequency end of the frequency spectrum. The dispersion relations for these two branches can clearly be written as

$$\omega^2=\omega_p{}^2\frac{a_0{}^2}{16\pi}C_j(\theta,\phi)k^2 \quad (j=1,2). \tag{C4}$$

It is well known[20] that the distribution function for the squares of the normal mode frequencies in a given branch is given by

$$G_j(\omega^2)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\exp(-i\alpha\omega^2)f_j(\alpha)d\alpha, \tag{C5}$$

[19] M. H. Cohen and F. Keffer, Phys. Rev. **99**, 1128 (1955).
[20] A. A. Maradudin and G. H. Weiss, Nuovo cimento **15**, 408 (1960).

where

$$f_j(\alpha)=E\{\exp(i\alpha\omega_j^2)\}$$
$$=\frac{1}{3N}\sum_{\mathbf{k}} e^{i\alpha\omega_j^2(\mathbf{k})}$$
$$=\frac{1}{3N}\frac{\Omega}{(2\pi)^3}\int\int\int e^{i\alpha\omega_j^2(\mathbf{k})}d^3\mathbf{k}. \quad \text{(C6)}$$

The integration in Eq. (C6) is carried out throughout the first Brillouin zone. The normalization volume Ω is given by

$$\Omega=N(a_0^3/2). \quad \text{(C7)}$$

The small ω^2 behavior of $G_j(\omega^2)$ is determined by the large $|\alpha|$ behavior of $f_j(\alpha)$. If we substitute Eq. (C4) into Eq. (C6), we can extend the integration throughout all space with little error, since there is only one minimum in each Brillouin zone, to find the leading term in the large $|\alpha|$ expansion of $f_j(\alpha)$:

$$f_j(\alpha)\sim\frac{a_0^3}{6(2\pi)^3}\int_0^\infty\int_0^\pi\int_0^{2\pi}\exp\left[i\alpha\omega_p^2\frac{a_0^2}{16\pi}C_j(\theta,\phi)k^2\right]\times k^2\sin\theta dk d\theta d\phi$$
$$=-\frac{\sqrt{2}I_j}{6\pi\omega_p^3}\frac{1-i\,\mathrm{sgn}\alpha}{|\alpha|^{\frac{3}{2}}}, \quad \text{(C8)}$$

where

$$I_j=\int_0^\pi\int_0^{2\pi}\frac{\sin\theta d\theta d\phi}{[C_j(\theta,\phi)]^{\frac{3}{2}}}. \quad \text{(C9)}$$

If we substitute Eq. (C8) back into Eq. (C5), we obtain[21]

$$G_j(\omega^2)\sim\frac{2}{3\pi^{\frac{3}{2}}\omega_p^3}I_j(\omega^2)^{\frac{1}{2}}, \quad \text{(C10)}$$

and since $g_j(\omega)=2\omega G_j(\omega^2)$, we have

$$g_j(\omega)=(4I_j/3\pi^{\frac{3}{2}}\omega_p^3)\omega^2 \quad \omega^2\to 0. \quad \text{(C11)}$$

By introducing $f_j(x)=\omega_p g(\omega_p x)$, we finally obtain

$$f_j(x)\sim(4I_j/3\pi^{\frac{3}{2}})x^2. \quad \text{(C12)}$$

The coefficient c_2 is therefore given by

$$c_2=(4/3\pi^{\frac{3}{2}})(I_1+I_2). \quad \text{(C13)}$$

[21] M. J. Lighthill, *Fourier Analysis and Generalized Functions* (Cambridge University Press, New York, 1958), p. 43.

To evaluate I_1 and I_2 we have used Houston's[22] method as developed by Betts, Bhatia, and Wyman.[23] They obtain the result that if we have an integral of the form

$$I=\int_0^\pi\int_0^{2\pi}F(\theta,\phi)\sin\theta d\theta d\phi, \quad \text{(C14)}$$

where $F(\theta,\phi)$ has cubic symmetry and is known only along the six special directions, [100], [110], [111], [210], [211], [221], which we denote by A, B, C, D, E, and F, respectively, the most accurate value of I obtainable from these data is

$$I=\frac{4\pi}{1081080}\{117603F_A+76544F_B+17496F_C$$
$$+381250F_D+311040F_E+177147F_F\}. \quad \text{(C15)}$$

With the aid of this result and Eqs. (C3) and (C9), we evaluated (I_1+I_2) with the result that

$$I_1+I_2=125.22, \quad \text{(C16)}$$

so that

$$c_2=29.984. \quad \text{(C17)}$$

With a value for c_2 we can obtain the low-temperature specific heat of the electron lattice. The specific heat per electron can be written as

$$\frac{C_v}{N}=3k\int_0^{\omega_p}\frac{(\hbar\omega/2kT)^2}{\sinh^2(\hbar\omega/2kT)}g(\omega)d\omega,$$

so that

$$C_v/N\sim c_2k(4\pi^4/5)(T/\Theta)^3 \quad \text{as } T\to 0, \quad \text{(C18)}$$

where we have put $\Theta=\hbar\omega_p/k$. If we substitute the values of c_2 and the fundamental constants into Eq. (C18) we obtain finally in the low-temperature limit

$$C_v/N=56.21k{r_s}^{9/2}(kT)_{\mathrm{ry}}^3. \quad \text{(C19)}$$

ACKNOWLEDGMENT

We would like to thank Dr. W. J. Carr, Jr., for sending us his results prior to their publication, and for several helpful discussions.

[22] W. V. Houston, Revs. Modern Phys. **20**, 161 (1948).
[23] D. D. Betts, A. B. Bhatia, and M. Wyman, Phys. Rev. **104**, 37 (1956).

Generalization of the Lagrange Expansion with Application to Physical Problems*

P. A. STURROCK
Microwave Laboratory, W. W. Hansen Laboratories of Physics, Stanford University, Stanford, California
(Received May 2, 1960)

Certain problems of fluid dynamics are conveniently discussed in the Lagrangian description by means of the displacement-vector function. The Eulerian variables describing the motion may be obtained from the displacement-vector function by a generalization of the Lagrange expansion which is here established. Two examples of the application of this expansion are given; the derivation of the Fokker-Planck equation, and the description of the surface conditions of a rippling electron stream by the device of equivalent surface charges.

1. INTRODUCTION

IN problems concerned with the dynamics of distributions of particles, such as the theories of electron tubes and of electron-ion plasmas[1] and some aspects of the theory of neutral gases, it is convenient to study perturbations of the particle distribution by means of a displacement vector rather than by the associated variation of the Eulerian variables $\rho(x,t)$, etc. Thus we assume that in the perturbed state of the system, the particle which was at point x_r at time t is now to be found at point $x_r+\xi_r(x,t)$ at the same time t. We here assume that the particles constitute a single stream, but one may take account of multistream systems by introducing a further parameter into $\xi_r(x,t)$.

The advantage of this formalism is that it readily lends itself to the application of particle dynamics in Lagrangian or Hamiltonian form.[2] We note that in this picture the single vector $\xi_r(x,t)$ replaces the scalar $\rho(x,t)$, the particle density, and the current density $j_r(x,t)$, which are related by the continuity equation. However, although the displacement vector is convenient for dynamical calculations, it is necessary to have some method for determining the Eulerian variables from the displacement-vector function. This communication deals with this problem.

We begin by considering a one-dimensional problem. We are given the perturbation of a linear array of particles by means of the displacement-vector function, and wish to evaluate the density in the perturbed state. We find the answer to this problem to be given by the familiar Lagrange expansion.[3]

We next consider the more general question posed by relaxing the restriction that the system be one-dimensional, and by seeking formulas for other Eulerian quantities such as the current density. These questions are answered by similar expansions, which it is natural to regard as generalizations of the original expansion, although this generalization is not possible in the form in which the expansion is usually given.

We consider two simple examples of the application of this expansion: the derivation of the Fokker-Planck equation[4] and the problem of representing perturbations of a finite electron beam by surface charges and surface currents.[5] We show in an appendix how our formulas may be derived by the methods of gas dynamics.

2. ONE-DIMENSIONAL PROBLEM

The familiar statement of the Lagrange expansion is the following: Let x and x_0 be related by the equation

$$x=x_0+\lambda\xi(x_0); \tag{2.1}$$

then x_0 is implicitly defined as a function of x. Any function of x_0, $F(x_0)$, should therefore be expressible in terms of λ and the functions $\xi(x)$, $F(x)$. The appropriate relation is

$$F(x_0(x))=F(x)+\sum_{\nu=0}^{\infty}(-)^{\nu}\frac{\lambda^{\nu}}{\nu!}\frac{d^{\nu}}{dx^{\nu}}\left\{\frac{dF(x)}{dx}\xi^{\nu}(x)\right\}. \tag{2.2}$$

Let us now consider the problem of the perturbation of a linear continuous distribution of particles, initially of density $\rho(x)$. We suppose that under the perturbation the particle initially at x_0 is transferred to the point x as given by (2.1). We wish to determine the density of the perturbed system $\tilde{\rho}(x)$.

For this purpose, it is convenient to introduce the cumulative measure function defined as the number or "weight" of particles to the left of a given point

$$F(x)=\int_{-\infty}^{x}\rho(x')dx', \quad \tilde{F}(x)=\int_{-\infty}^{x}\tilde{\rho}(x')dx'. \tag{2.3}$$

We now see that

$$\tilde{F}(x)\equiv F[x_0(x)], \tag{2.4}$$

if we assume, here and throughout, that λ is not so large that "crossover" occurs. We also note that

$$\rho(x)=dF(x)/dx, \quad \tilde{\rho}(x)=d\tilde{F}(x)/dx. \tag{2.5}$$

The relation we seek may now be obtained by dif-

* The research reported in this paper has been sponsored by the Electronics Research Directorate of the Air Force Cambridge Research Center, Air Research and Development Command.

[1] P. A. Sturrock, Phys. Rev. 117, 1426 (1960).

[2] P. A. Sturrock, Ann. Phys. 4, 306 (1958).

[3] E. T. Whittaker and G. N. Watson, *Modern Analysis* (Cambridge University Press, New York, 1952), 4th ed., p. 132.

[4] S. Chandrasekhar, Revs. Modern Phys. 15, 31 (1943).

[5] E. L. Chu, J. Appl. Phys. 31, 381 (1960).

ferentiating (2.2),

$$\tilde{\rho}(x)=\sum_{\nu=1}^{\infty}(-)^{\nu}\frac{\lambda^{\nu}}{\nu!}\frac{d^{\nu}}{dx^{\nu}}\{\rho(x)\xi^{\nu}(x)\} \tag{2.6}$$

or, more explicitly,

$$\tilde{\rho}(x)=\rho(x)-\lambda(d/dx)[\rho(x)\xi(x)] +\tfrac{1}{2}\lambda^2(d^2/dx^2)[\rho(x)\xi^2(x)]-\cdots. \tag{2.7}$$

Equation (2.6) is equivalent to the Lagrange expansion (2.2), but we shall see that (2.6) can be generalized to n dimensions, whereas (2.2) apparently cannot.

3. GENERALIZATIONS OF THE LAGRANGE EXPANSION

Now reconsider the problem of evaluating the density of a distribution of particles consequent upon displacement of these particles, but consider space of a higher number of dimensions by replacing the single variable x by a vector with coordinates x_r where, for definiteness, we assume that r takes the values of 1, 2, 3. The summation convention will be used.

We introduce the Fourier transform of the density function $\rho(x)$ defined by

$$\rho(x)=\int d^3x e^{ik_r x_r}\rho_F(k), \tag{3.1}$$

$$\rho_F(k)=\left(\frac{1}{2\pi}\right)^3\int d^3x e^{-ik_r x_r}\rho(x). \tag{3.2}$$

We now consider Eq. (3.2) as it applies to the perturbed state, and note that the integration over $\tilde{\rho}(x)d^3x$ is simply a counting of particles, and so may equally well be replaced by integration over $\rho(x_0)d^3x_0$. Hence, we see from (2.1) and (3.2) that

$$\tilde{\rho}_F(k)=\left(\frac{1}{2\pi}\right)^3\int d^3x_0\rho(x_0) \times\exp\{-ik_r[x_{0,r}+\lambda\xi_r(x_0)]\}. \tag{3.3}$$

We now insert (3.3) into (3.1), expanding one of the exponentials, to obtain

$$\tilde{\rho}(x)=\left(\frac{1}{2\pi}\right)^3\int d^3x'\rho(x')\int d^3k \times\exp[ik_r(x_r-x_r')]\sum_{\nu=0}^{\infty}(-i)^{\nu}\frac{\lambda^{\nu}}{\nu!}\{k_r\xi_r(x')\}^{\nu}. \tag{3.4}$$

The k integration may be expressed in terms of δ functions as follows:

$$\tilde{\rho}(x)=\int d^3x'\rho(x')\left[\delta^3(x-x')-\lambda\xi_r(x')\frac{\partial\delta^3(x-x')}{\partial x_r} +\frac{\lambda^2}{2!}\xi_r(x')\xi_s(x')\frac{\partial^2\delta^3(x-x')}{\partial x_r\partial x_s}-\cdots\right], \tag{3.5}$$

which leads finally to the expansion

$$\tilde{\rho}(x)=\rho(x)-\lambda\frac{\partial}{\partial x_r}\{\rho(x)\xi_r(x)\} +\frac{\lambda^2}{2!}\frac{\lambda^2}{\partial x_r\partial x_s}\{\rho(x)\xi_r(x)\xi_s(x)\}-\cdots. \tag{3.6}$$

This represents a generalization of the Lagrange expansion.

Since the density of the perturbed system may be expressed by the expansion (3.6), we should expect that other Eulerian quantities, such as current density, energy density, etc., may be expressed in a similar fashion. Let us consider the quantity $\phi(x,t)$, which may be a vector or tensor, and suppose that it is expressible as

$$\phi=\rho\psi(v_r), \tag{3.7}$$

in which $\rho(x,t)$ and $v_r(x,t)$ are known for the unperturbed system. On perturbation, the particle which was at x_r at time t is now at $x_r+\xi_r(x,t)$ at time t (here and henceforth we dispense with the parameter λ), and has velocity $v_r+(\partial\xi_r/\partial t)+v_s(\partial\xi_r/\partial x_s)$. Hence the Fourier transform of ϕ is given by the expression formed from (3.3) by replacing ρ by $\rho\psi$, the argument of ψ being the perturbed velocity. Hence we obtain, in place of (3.6), the expansion

$$\tilde{\phi}(x,t)=\rho(x)\psi\left(v_r+\frac{\partial\xi_r}{\partial t}+v_s\frac{\partial\xi_r}{\partial x_s}\right) -\frac{\partial}{\partial x_u}\left\{\rho(x)\psi\left(v_r+\frac{\partial\xi_r}{\partial t}+v_s\frac{\partial\xi_r}{\partial x_s}\right)\xi_u(x)\right\}+\cdots. \tag{3.8}$$

As a particular example, consider the current density j_r, for which $\psi(v_r)=v_r$. The appropriate expansion is seen to be

$$\tilde{j}_r(x,t)=\rho(x)\left(v_r+\frac{\partial\xi_r}{\partial t}+v_s\frac{\partial\xi_s}{\partial x_s}\right) -\frac{\partial}{\partial x_u}\left\{\rho(x)\left(v_r+\frac{\partial\xi_r}{\partial t}+v_s\frac{\partial\xi_r}{\partial x_s}\right)\xi_u(x)\right\}+\cdots. \tag{3.9}$$

4. FOKKER-PLANCK EQUATION[6]

Consider an assembly of particles, with a range of "thermal" velocities, acted upon by "external" forces and by random microforces such as those which give rise to interparticle small-angle collisions. We represent the distribution by the function $f(x,v,t)$ in the usual way. We now suppose that at time $t+\Delta t$ the particle which at time t had position x_r and velocity v_r is found to have position $x_r+\Delta x_r$ and velocity $v_r+\Delta v_r$. Then

we see from (3.6) that

$$f(x,v,t+\Delta t)=f-\frac{\partial}{\partial x_r}(f\Delta x_r)-\frac{\partial}{\partial v_r}(f\Delta v_r)$$
$$+\frac{1}{2}\frac{\partial^2}{\partial x_r\partial x_s}(f\Delta x_r\Delta x_s)+\frac{\partial^2}{\partial x_r\partial v_s}(f\Delta x_r\Delta v_s)$$
$$+\frac{1}{2}\frac{\partial^2}{\partial v_r\partial v_s}(f\Delta v_r\Delta v_s)-\cdots, \quad (4.1)$$

wherein all quantities on the right-hand side are evaluated at time t.

We now suppose that Δt may be chosen small enough for the change in the macroscopic quantities such as the external forces and f itself to be small, yet large enough for a large number of small-angle collisions to take place in this period. Then we may write

$$f(x,v,t+\Delta t)-f(x,v,t)=(\partial f/\partial t)\cdot\Delta t, \quad \Delta x_r=v_r\Delta t. \quad (4.2)$$

The change in the velocity may be written as the sum of two contributions, $\Delta_F v_r$ caused by external forces and $\Delta_C v_r$ caused by collisions. Then

$$\Delta_F v_r=(F_r/m)\Delta t, \quad (4.3)$$

according to our assumptions concerning Δt, where $F_r(x,v,t)$ is the external force. In dealing with the collision term, we consider averages over a large number of collisions. Then both $\Delta_C v_r$ and $\Delta_C v_r\Delta_C v_s$ will give contributions linear in Δt, whereas higher-order products will tend to zero more rapidly than Δt as Δt is diminished, subject always to the restriction that Δt is long compared with the collision interval. In this sense, we may write

$$\left\langle\frac{\Delta_C v_r}{\Delta t}\right\rangle\rightarrow\left\langle\frac{\partial v_r}{\partial t}\right\rangle_C, \quad \left\langle\frac{\Delta_C v_r\Delta_C v_s}{\Delta t}\right\rangle\rightarrow\left\langle\frac{\partial v_r\partial v_s}{\partial t}\right\rangle_C,$$
$$\text{as} \quad \text{“}\Delta t\rightarrow 0\text{.”} \quad (4.4)$$

The corresponding limit of higher-order products is zero.

On combining the foregoing equations, we find that (4.1) may be written as

$$\frac{\partial f}{\partial t}+v_r\frac{\partial f}{\partial x_r}+\frac{\partial}{\partial v_r}\left(\frac{F_r}{m}f\right)$$
$$=-\frac{\partial}{\partial v_r}\left(\left\langle\frac{\partial v_r}{\partial t}\right\rangle_C f\right)+\frac{1}{2}\frac{\partial^2}{\partial v_r\partial v_s}\left(\left\langle\frac{\partial v_r\partial v_s}{\partial t}\right\rangle_C f\right), \quad (4.5)$$

where we adopt the conventional arrangement of grouping collisions on the right-hand side. Note that the acceleration term on the left-hand side is in the form appropriate to velocity-dependent forces.

5. SURFACE CHARGE AND DIPOLE LAYERS OF RIPPLED BEAMS

A problem which has received some attention of late and which may be conveniently handled by means of the Lagrange expansion is that of determining the equivalent surface charges to ascribe to a rippling beam with a sharp boundary.[6] We shall discuss only the charge density, but the current density and other quantities could be evaluated similarly. The problem which we have to solve is that of handling the discontinuity represented by the edge of the beam. For this purpose, it is convenient to introduce the step function defined by

$$\theta(\alpha)=0, \quad \alpha<0,$$
$$=1, \quad \alpha>0, \quad (5.1)$$

and the function $N(x)$ which denotes the normal distance from the point x to the surface of the beam, counted positive if the point is outside the beam and negative if the point is inside the beam. We may now write the unperturbed density $\rho(x)$ of the beam as

$$\rho(x)=\rho_0(x)\theta[-N(x)], \quad (5.2)$$

wherein $\rho_0(x)$ may be taken to be a well-behaved function, the discontinuity being represented by the function $\theta[-N(x)]$ which is unity for interior points and zero for exterior points. We note that

$$\partial N/\partial x_r=n_r, \quad (5.3)$$

where n_r is the outward normal vector. Since $n_r^2=1$, we also see that

$$n_s(\partial n_s/\partial x_r)=0 \quad \text{or} \quad n_s(\partial n_r/\partial x_s)=0, \quad (5.4)$$

so that $\partial n_r/\partial x_s$ may be evaluated from knowledge of the normal-vector field at the boundary surface itself.

We may now apply the Lagrange expansion (3.6). Upon noting that

$$(\partial/\partial x_r)\theta(-N)=-n_r\delta(N), \quad (5.5)$$

and

$$(\partial^2/\partial x_r\partial x_s)\theta(-N)=-(\partial n_r/\partial x_s)\delta(N)-n_r n_s\delta'(N), \quad (5.6)$$

we see that

$$\tilde{\rho}(x)=\tilde{\rho}_V(x)\theta(-N)+\tilde{\sigma}(x)\delta(N)-\tilde{\pi}(x)\delta'(N)+\cdots, \quad (5.7)$$

where

$$\tilde{\rho}_V(x)=\rho_0-(\partial/\partial x_r)(\rho_0\xi_r)$$
$$+\tfrac{1}{2}(\partial^2/\partial x_r\partial x_s)(\rho_0\xi_r\xi_s)+\cdots, \quad (5.8)$$

$$\tilde{\sigma}(x)=\rho_0 n_r\xi_r-n_r(\partial/\partial x_s)(\rho_0\xi_r\xi_s)$$
$$-\tfrac{1}{2}(\partial n_r/\partial x_s)\rho_0\xi_r\xi_s+\cdots, \quad (5.9)$$

and

$$\tilde{\pi}(x)=\tfrac{1}{2}\rho_0 n_r n_s\xi_r\xi_s+\cdots, \quad (5.10)$$

to second order in the displacement. Formulas (5.9) and (5.10) are to be evaluated at the bounding surface. We see from the form of (5.7) that $\tilde{\rho}_V(x)$ represents the

perturbed space-charge density inside the beam, $\tilde{\sigma}(x)$ represents the surface-charge density, and $\tilde{\pi}(x)$ represents the surface-dipole layer.

6. DISCUSSION

The generalization of the Lagrange expansion which was established in Sec. 3 is useful in relating the Lagrangian and Eulerian descriptions of the motion of a fluid or assembly of particles. The transition from Lagrangian to Eulerian variables may be desired for the purpose of interpreting mathematical results. It may also be necessary if an Eulerian description is required for one part of the problem, whereas a Lagrangian description is convenient for another part. Problems concerning electron beams and plasmas are of this category, since it is convenient to discuss the motion of particles by Lagrangian variables whereas the equations of the electromagnetic field require the Eulerian formulation. In this connection, however, we may note that the complete problem may be described by an action principle which makes it unnecessary to introduce the Eulerian description of the particle motion.[2]

The conditions for validity of our generalized forms of the Lagrange expansion have not been investigated. It may be possible to examine the series for convergence in particular examples. However, we have seen in Sec. 5 that the expansion is informative even when it is not convergent. Such applications of the expansion are subject to the usual restrictions applied to the calculus of the Dirac delta function.

APPENDIX. DERIVATION OF THE LAGRANGE EXPANSION FROM GAS KINETICS

The Lagrange expansion (3.6), which we derived by Fourier-transform theory, may also be established by arguments of gas kinetics. In place of (2.1), let us now write

$$x_r \rightarrow \bar{x}_r(x,t) = x_r + t\xi_r(x), \tag{A.1}$$

where we shall interpret t as time. Under this perturbation, the initial density $\rho(x)$ becomes $\tilde{\rho}(x,t)$, which we shall expand as

$$\tilde{\rho} = \sum_{\nu=0}^{\infty} \frac{1}{\nu!} t^\nu \tilde{\rho}_\nu. \tag{A.2}$$

By Taylor's theorem,

$$\tilde{\rho}_\nu = \partial^\nu \tilde{\rho}/\partial t^\nu \quad (t=0). \tag{A.3}$$

We see from (A.1) that the velocity of any particle of the fluid is constant, being given by $\xi_r(x)$ evaluated for the position occupied by the particle at $t=0$. Hence we obtain the sequence of conservation relations

$$(\partial/\partial t)(\tilde{\rho}\psi) + (\partial/\partial x_r)(\tilde{\rho}\psi v_r) = 0, \tag{A.4}$$

where $\psi = 1, v_s, v_s v_t, v_s v_t v_u, \cdots$. From this equation, we may readily prove by induction that

$$\partial^\nu \tilde{\rho}/\partial t^\nu = (-)^\nu (\partial^\nu/\partial x_r \partial x_s \cdots)(\tilde{\rho} v_r v_s \cdots). \tag{A.5}$$

On evaluating (A.5) at $t=0$, we see that

$$\tilde{\rho}_\nu = (-)^\nu (\partial^\nu/\partial x_r \partial x_s \cdots)(\rho \xi_r \xi_s \cdots); \tag{A.6}$$

(A.2) and (A.6) yield once more the Lagrange expansion (3.6).

Normal Form of Antiunitary Operators

EUGENE P. WIGNER
Palmer Physical Laboratory, Princeton University, Princeton, New Jersey
(Received April 25, 1960)

Antiunitary operators are characterized in a manner similar to the characterization of unitary operators by their characteristic vectors and characteristic values. It is shown that a complete orthonormal set of vectors can be defined, some of which are invariant under the antiunitary operator. The rest of the vectors, which are always even in number, form pairs in such a way that the antiunitary operator transforms each member of a pair into a multiple of the other member of the same pair [Eq. (11)]. The extent to which the vectors of the orthonormal set are determined by the antiunitary operator is ascertained and the number of free parameters in the various cases of degeneracy found.

1.

ANTIUNITARY operators[1] play a significant role in the theory of the invariance of quantum mechanical equations. The symmetry operators which involve the operation of time-inversion are antiunitary. The antiunitary operators are antilinear, i.e., if φ and ψ are two vectors of the complex Hilbert space in which the antiunitary operator A is defined and if a and b are two complex numbers,

$$A(a\varphi+b\psi)=a^*A\varphi+b^*A\psi. \tag{1}$$

The asterisk denotes the conjugate complex. Furthermore, A changes the scalar product into its conjugate complex

$$(A\varphi,A\psi)=(\varphi,\psi)^*=(\psi,\varphi). \tag{2}$$

Actually, (1) follows from (2) so that the latter equation can serve as the definition of the antiunitary nature of A. However, unless the Hilbert space has only a finite number of dimensions, it is also necessary to specify that A has an inverse. This is also antiunitary.

If A is antiunitary, A^2 defined by

$$A^2\psi=A(A\psi) \tag{3}$$

is unitary. This follows directly from the defining Eqs. (1) and (2), and it is also clear that A^2 has an inverse if A does.

If Av_k is given for all the members of a complete orthonormal set of vectors $v_1,v_2,\cdots$, its antilinear property defines it for all vectors $v=\Sigma a_kv_k$:

$$Av=A(\sum a_kv_k)=\sum a_k^*Av_k. \tag{4}$$

Hence, the normal form of A will be obtained by specifying a complete set of orthonormal vectors v_k for which Av_k has a particularly simple form. These vectors are the analogs of the characteristic vectors for unitary operators and will be, indeed, characteristic vectors of A^2. However, this property does not define them completely.

If $v_1,v_2,\cdots$ form a complete orthonormal set, $Av_1,Av_2,\cdots$ also form such a set. The orthonormal nature of the latter set follows directly from (2), the completeness from the existence of the inverse of A. If w were orthogonal to all Av_k, then $A^{-1}w$ would be orthogonal to all v_k.

We mention further for the sake of completeness, that if K is the operation of complex conjugation so that, in a particular coordinate system,

$$K\psi=\psi^*;\quad K^2=1, \tag{5}$$

AK is unitary and it follows that every antiunitary operator can be written in the form

$$A=UK, \tag{6}$$

where U is unitary. It follows from (6) that

$$A^2=UKUK=UU^*K^2=UU^*, \tag{7}$$

where U^* is the conjugate complex of U in the coordinate system in which (5) is valid. Since UU^* is equivalent to its conjugate complex

$$UU^*=UU^*UU^{-1}=U(UU^*)^*U^{-1}, \tag{8}$$

its characteristic values are either real or pairwise conjugate complex. It follows that the square of an antiunitary operator is equivalent to a rotation. The last four equations will not be used explicitly.

2.

It will be assumed that the spectrum of $A^2=\Lambda$ is discrete. The complications which arise if Λ has a continuous spectrum are not serious, but their elimination is cumbersome. Let us consider then a characteristic vector of Λ:

$$\Lambda v=A^2v=\omega v. \tag{9}$$

Since Λ is unitary, $|\omega|=1$. It then follows that Av is also a characteristic vector of Λ,

$$\Lambda Av=A^2Av=AA^2v=A\omega v=\omega^*Av, \tag{10}$$

[1] Some of the results of the present article can be obtained on the basis of theorems derived by E. Cartan in his *Leçons sur la Géométrie Projective Complexe* (Gauthier-Villars, Paris, 1931). I am much indebted to Professor S. Bochner for drawing my attention to the very profound investigations contained in this treatise, which deals with general linear and antilinear transformations. However, the direct derivations, given in the text of the present paper, are hardly longer than the reinterpretation and amplification of Cartan's results (see particularly pp. 124–137) would have been.

and belongs to the characteristic value ω^*. Unless $\omega=1$ or $\omega=-1$, $\omega\neq\omega^*$ and Av is orthogonal to v. Hence, if we choose an arbitrary orthonormal base, $v_{\omega 1}, v_{\omega 2}, \cdots$ among the characteristic vectors of ω, we can define, if ω is complex,

$$v_{\omega^*,k}=\omega^{\frac{1}{2}}Av_{\omega,k} \quad \text{or} \quad Av_{\omega,k}=(\omega^{\frac{1}{2}})^*v_{\omega^*k} \tag{11}$$

and the $v_{\omega^*,k}$ will form a full base of orthonormal characteristic vectors to ω^*. The sign of the square root in (11) is best fixed in such a way that the imaginary part of $\omega^{\frac{1}{2}}$ shall have the same sign as the imaginary part of ω. Then $(\omega^*)^{\frac{1}{2}}=(\omega^{\frac{1}{2}})^*$. The purpose of the $\omega^{\frac{1}{2}}$ factor will become evident at once.

Application of A to both sides of (11) gives

$$Av_{\omega^*,k}=(\omega^{\frac{1}{2}})^*A^2v_{\omega,k}=(\omega^{\frac{1}{2}})^*\omega v_{\omega,k}=\omega^{\frac{1}{2}}v_{\omega,k}, \tag{12}$$

so that the choice of the characteristic vectors to ω^* made in (11) renders this equation valid also if ω is replaced by ω^*. The $v_{\omega,k}$ may be called characteristic vectors of A also.[2] However, in contrast to the unitary case, the characteristic vectors of A to ω also define the characteristic vectors of A to ω^* if we want (11) to hold. If one recalls that Λ is equivalent to a rotation it is not surprising that a certain amount of simplification results if a relation exists between the characteristic vectors of ω and of ω^*. In the case of a rotation one would set $v_{\omega^*,k}=v_{\omega,k}{}^*$.

Let us consider now a characteristic vector v to the characteristic value 1:

$$\Lambda v=A^2v=v. \tag{13}$$

It then follows from (10) that Av is also a characteristic vector to the characteristic value 1 and so is, unless it vanishes, $v_{11}=c(v+Av)$; c is a real normalization constant. It follows from (13) that

$$Av_{11}=Ac(v+Av)=c(Av+v)=v_{11}, \tag{14}$$

so that v_{11} is invariant under A. If $v=-Av$ we choose $v_{11}=iv$ and have again

$$Av_{11}=Aiv=-iAv=iv=v_{11}. \tag{15}$$

Next we consider another characteristic vector $v'=\Lambda v'$ which is orthogonal to v_{11}:

$$(v_{11},v')=0. \tag{16}$$

Because of (2) and (14),

$$(v_{11},Av')=(A^2v',Av_{11})=(\Lambda v',v_{11})=(v',v_{11})=0, \tag{17}$$

Av' will also be orthogonal to v_{11}. We can write therefore $v_{12}=c(v'+Av')$ or, if this vanishes, $v_{12}=iv'$, and this will still be orthogonal to v_{11} and also invariant under A. Proceeding in the same way, a full orthonormal base $v_{11}, v_{12}, \cdots$ of characteristic vectors of Λ to the characteristic value 1 can be found which are invariant under A,

$$Av_{1k}=v_{1k}. \tag{18}$$

The vectors which satisfy (18) can be called the invariant vectors of A. The procedure just used to ensure (18) is similar to the separation of real and imaginary parts of a number.

Let us finally consider a characteristic vector of Λ to -1:

$$\Lambda v_{-11}=A^2v_{-11}=-v_{-11}. \tag{19}$$

In this case again, because of (10), Av_{-11} is also a characteristic value to -1. Furthermore, Av_{-11} is orthogonal to v_{-11} because of (2) and (19):

$$(v_{-11},Av_{-11})=(A^2v_{-11},Av_{-11})=-(v_{-11},Av_{-11}). \tag{20}$$

Hence we can write

$$v_{-1^*1}=iAv_{-11} \quad v_{-11}=i^*Av_{-1^*1}=-iAv_{-1^*1}. \tag{21}$$

If Λ has further linearly independent characteristic vectors to -1, a normalized v_{-12} can be found which is orthogonal to both v_{-11} and v_{-1^*1}. Furthermore, the same will be true of $v_{-1^*2}=iAv_{-12}$. Thus, for instance,

$$\begin{aligned}(v_{-1^*2},v_{-11})&=(iAv_{-12},v_{-11})=-i(Av_{-11},A^2v_{-12})\\&=-i(-iv_{-1^*1},-v_{-12})=0.\end{aligned} \tag{22}$$

Hence, proceeding in the same way, one can find a full orthonormal base of characteristic vectors of Λ to -1,

$$\Lambda v_{-1k}=-v_{-1k} \quad \Lambda v_{-1^*k}=-v_{-1^*k}, \tag{23}$$

for which

$$v_{-1^*k}=iAv_{-1k} \quad v_{-1k}=-iAv_{-1^*k}=i^*A_{-1^*k} \tag{24}$$

holds. These equations are formally identical with the Eqs. (11) for complex characteristic values if one considers -1 to be two conjugate complex characteristic values -1 and -1^* of Λ, which happen to coincide. The v_{-1k} belong to the characteristic value -1, the v_{-1^*k} to the characteristic value -1^*. Equation (24) becomes a special case of (11) if one sets $(-1)^{\frac{1}{2}}=i$; $(-1^*)^{\frac{1}{2}}=i^*=-i$.

3.

On summarizing the preceding results, we can characterize an antiunitary operator by two sets of vectors, which jointly form a complete orthonormal set, together with the characteristic values $\omega_1, \omega_1{}^*, \omega_2, \omega_2{}^*, \cdots$ belonging to the second set. These characteristic values are pairwise conjugate complex, of modulus 1, but are not equal to 1. The first set of vectors are invariant under the antiunitary operator, i.e., (18) applies to them; (11) is valid for the members of the second set. The ω may also be equal to -1, but this characteristic value always occurs in pairs and one member of the pair is denoted by -1, the other by -1^*.

It will be shown now that any two sets of vectors v_{1k} and $v_{\omega k}$ which jointly form a complete orthonormal

[2] The two vectors $v_{\omega k}$ and $v_{\omega^* k}$ form a plane in our Hilbert space. The line which corresponds to this plane in Cartan's projective space is the invariant line of the passage cited in footnote reference 1.

set, together with the corresponding ω, give an antiunitary operator by means of (4), (11), and (18). In other words, the sets v_{1k} and $v_{\omega k}$ are not subject to any further conditions except that there are just as many vectors bearing the index ω as there are with the index ω^*. The number of vectors in the first set is arbitrary and so are the values of ω except that $\omega \neq 1$, $|\omega| = 1$ and they occur in conjugate complex pairs.

In order to prove the preceding assertion we consider two vectors φ and ψ and expand them in terms of the orthonormal set

$$\varphi = \sum_k a_k v_{1k} + \sum_{\omega k} b_{\omega k} v_{\omega k}$$
$$\psi = \sum_k c_k v_{1k} + \sum_{\omega k} d_{\omega k} v_{\omega k}. \tag{25}$$

$A\varphi$ and $A\psi$ are then given by

$$A\varphi = \sum_k a_k^* v_{1k} + \sum_{\omega k} b_{\omega k}^* (\omega^{\frac{1}{2}})^* v_{\omega^* k}$$
$$A\psi = \sum_k c_k^* v_{1k} + \sum_{\omega k} d_{\omega k}^* (\omega^{\frac{1}{2}})^* v_{\omega^* k}. \tag{26}$$

Both conditions (1) and (2) of the antiunitary nature of A can be verified to be consequences of (26) and the orthonormality of the v_{1k}, $v_{\omega k}$, provided that

$$(\omega^*)^{\frac{1}{2}} = (\omega^{\frac{1}{2}})^* \quad |\omega| = 1. \tag{27}$$

For $\omega = -1$, this last condition is spelled out explicitly in (23). As was mentioned before, (27) can most simply be assured for complex ω by using that sign for $\omega^{\frac{1}{2}}$ for which the signs of the imaginary parts of ω and of $\omega^{\frac{1}{2}}$ are the same.

4.

Evidently, the two sets v_{1k}, $v_{\omega k}$ and the corresponding ω completely determine A. Conversely, A determines the number of vectors contained in the set v_{1k}—this is the multiplicity of the characteristic value 1 of A^2—and the value of the ω and their multiplicities. However, the vectors v are not completely determined by A and the present section will be devoted to the determination of the freedom that remains in the choice of these vectors.

Let us denote two other orthonormal sets which characterize the same antilinear operator by w_{1k} and $w_{\omega k}$. Since the w_{1k} form a base for the characteristic functions to the characteristic value 1 of $\Lambda = A^2$, they are connected with the v_{1k} by a nonsingular transformation

$$w_{1k} = \sum r_{kl} v_{1l}. \tag{28}$$

In fact, it follows from the orthonormality of the v_{1k} and of the w_{1k} that r is unitary. This is, however, not the only condition on r: If the vectors w_{1k} are to be invariant under A, i.e., if they satisfy (18),

$$Aw_{1k} = \sum r_{kl}^* A v_{1l} = \sum r_{kl}^* v_{1l} = w_{1k}, \tag{29}$$

the r_{kl} must be real. Hence, two different invariant sets of vectors of the same antiunitary operator are related to each other by a rotation

$$r = r^* \quad rr' = rr^\dagger = 1. \tag{30}$$

The prime denotes the transpose, the dagger the Hermitian adjoint.

For complex ω, the sets $w_{\omega k}$ and $v_{\omega k}$ span the same linear manifold. Hence, we have

$$w_{\omega k} = \sum u_{kl}^{(\omega)} v_{\omega l}, \tag{31}$$

and it again follows from the orthonormality of the w_ω and v_ω that $u^{(\omega)}$ is unitary. By calculating $Aw_{\omega k}$ again, we find

$$Aw_{\omega k} = \sum u_{kl}^{(\omega)*} A v_{\omega l} = \sum u_{kl}^{(\omega)*} (\omega^*)^{\frac{1}{2}} v_{\omega^* l}, \tag{32}$$

so that if we want $Aw_{\omega k} = (\omega^*)^{\frac{1}{2}} w_{\omega^* k}$ to remain valid, we must have

$$u^{(\omega^*)} = u^{(\omega)*}, \tag{33}$$

i.e., the unitary transformations which belong to conjugate complex characteristic value are conjugate complex.

The preceding argument does not apply if $\omega = -1$. It is indeed clear that in this case the w_{-1k} may be linear combinations of the v_{-1k} and of the $v_{-1^* k}$ because all these belong to the characteristic value -1 of Λ. Hence we set

$$w_{-1k} = \sum s_{kl} v_{-1l} + \sum t_{kl} v_{-1^* l}. \tag{34}$$

The condition (24) that $w_{-1^* k} = iAw_{-1k}$ now reads

$$\begin{aligned} w_{-1^* k} &= iA\left(\sum s_{kl} v_{-1l} + \sum t_{kl} v_{-1^* l}\right) \\ &= \sum s_{kl}^* iA v_{-1l} + \sum t_{kl}^* iA v_{-1^* l} \\ &= \sum -t_{kl}^* v_{-1l} + \sum s_{kl}^* v_{-1^* l}. \end{aligned} \tag{35}$$

Hence, the sets of vectors w_{-1} and w_{-1^*} are obtained from the sets v_{-1}, v_{-1^*} by the transformation

$$S = \begin{Vmatrix} s & t \\ -t^* & s^* \end{Vmatrix}. \tag{36}$$

This will guarantee that (24) is valid for the w_{-1}, w_{-1^*} if it is valid for the v_{-1}, v_{-1^*} because the second set of Eqs. (24) can be obtained from the first set by applying A to these. However, in order to make the w_{-1}, w_{-1^*} an orthonormal set, the S of (36) must be unitary. The conditions for this are obtained by setting $SS^\dagger = 1$ or, in terms of the submatrices s and t,

$$ss^\dagger + tt^\dagger = 1 \quad st' = ts'. \tag{37}$$

It is easy to see that if the conditions (37) are satisfied, S becomes a simplectic matrix, i.e., it leaves the form

$$F = \begin{Vmatrix} 0 & 1 \\ -1 & 0 \end{Vmatrix} \tag{38}$$

invariant in the sense that

$$SFS'=F. \tag{39}$$

It follows that the sets w_{-1k}, w_{-1^*k} are obtained from the sets v_{-1k}, v_{-1^*k} by a unitary simplectic transformation.

The calculation of the last paragraph shows that the role of vectors $v_{\omega k}$, $v_{\omega^* k}$ for $\omega=-1$ is quite different from the role of the vectors $v_{\omega k}$, $v_{\omega^* k}$ for complex ω. The fact that the same Eq. (24) holds for $\omega=-1$ and for complex ω is somewhat accidental.

It may be well to note at this point that the equation

$$Aw=\nu w \tag{40}$$

with complex ν does not imply that ν is one of the ω. In fact, (40) holds with $w=(\nu^{\frac{1}{2}})^* v_{1k}$ and an arbitrary ν.

5.

Lastly, we shall determine the number of free parameters in an antiunitary transformation which can be characterized by l invariant vectors; $2m$ vectors with the characteristic value -1; 2ρ different complex characteristic values with positive imaginary parts and their complex conjugates with multiplicities $c_1, c_2, \cdots, c_\rho$. These are then also the multiplicities of the corresponding conjugate complex characteristic values. Hence,

$$l+2m+2c_1+2c_2+\cdots+2c_\rho=n, \tag{41}$$

where n is the number of dimensions of the underlying Hilbert space which will be assumed to be finite dimensional in the present section.

The number of free parameters will be calculated by adding the free parameters necessary to characterize the complete orthonormal set v_{1k}, $v_{\omega k}$ and the ω, and subtracting the number of parameters contained in the transformations which alter the v but leave A unchanged. These were determined in the preceding section.

A complete orthonormal set in n dimensions can be characterized by $2n-1+(2n-3)+\cdots+3+1=n^2$ parameters. The number of free parameters in the ω is just ρ. Hence, $n^2+\rho$ parameters are necessary to characterize the v and the ω.

A rotation in the l-dimensional space of the v_{1k} does not change A. The number of parameters of such a rotation is $\frac{1}{2}l(l-1)$. Similarly, a $2m$-dimensional unitary simplectic transformation remains free for the vectors $v_{-11}, \cdots, v_{-1m}, v_{-1^*1}, \cdots, v_{-1^*m}$. The number of parameters of such a transformation is $m(2m+1)$. Finally, an arbitrary unitary transformation of the vectors $v_{\omega 1}, v_{\omega 2}, \cdots$ leaves A also unchanged if the conjugate complex transformation is applied to the vectors $v_{\omega^* 1}, v_{\omega^* 2}, \cdots$. The number of parameters in such a transformation is just the square of the corresponding c. Hence, the total number of free parameters in the antiunitary transformation is

$$p=n^2+\rho-\tfrac{1}{2}l(l-1)-m(2m+1)-\sum_1^\rho c_r^2$$

$$=n^2-\tfrac{1}{2}l(l-1)-m(2m+1)-\sum_1^\rho (c_r^2-1). \tag{42}$$

For even n, the number of parameters is just n^2 if all the characteristic values are complex and simple. Two invariant vectors decrease the number of parameters by 1, two characteristic values -1 by 3, if a complex characteristic value is doubly degenerate (the same then holds for the conjugate complex characteristic value) the number of parameters is also decreased by three.

The number of free parameters is also n^2 if n is odd and there are $n-1$ simple complex characteristic values and one invariant vector. Multiplicities among the complex characteristic values and the presence of a characteristic value -1 (which is always at least double) reduce the number of free parameters as in the case of even n.

The fact that the number of parameters is n^2 in the general case could have been inferred from the possibility of representing an antiunitary transformation in the form (6), i.e., as the product of a unitary transformation and complex conjugation. The number of free parameters in an n-dimensional unitary transformation is just n^2. The decrease in the number of free parameters (by 3) caused by the presence of a single pair of characteristic values -1 is remarkable.

6.

The preceding results will now be formulated in the language of projection operators and thus extended to the case in which there is a continuous spectrum. However, the proofs, which are rather obvious, will be omitted.

Consider again the unitary operator $\Lambda=A^2$. If 1 and -1 belong to the point spectrum of Λ, denote the corresponding projection operators by E_1 and E_{-1}. The projection operator which belongs to an interval J of the unit circle in the complex plane will be denoted by E_J. All these projection operators are self-adjoint, commute with Λ and with each other; the product of two of them is equal to the projection operator which corresponds to the intersection of the domains to which the two factors correspond. Furthermore,

$$\Lambda E_1=E_1 \quad \Lambda E_{-1}=-E_{-1} \quad \lim \Lambda E_J=\omega E_J, \tag{43}$$

where the lim in the last equation indicates that J is an infinitely narrow interval around ω. We define the antiunitary operators

$$A_1=AE_1 \quad A_{-1}=AE_{-1} \quad A_J=AE_J, \tag{44}$$

then

$$A=A_1+A_{-1}+\lim \sum A_J. \tag{45}$$

The lim again indicates that the intervals J are infinitely narrow; they cover all the unit circle with the exceptions of the points 1 and -1. The intervals J will be assumed to lie either entirely in the upper half-plane, or entirely in the lower half-plane. The interval J^* will be the conjugate complex of the interval J.

It is good to recall, for the rest of this discussion, that A^{-1} is also an antiunitary operator and is, in fact, given by

$$A^{-1}=\Lambda^{-1}A=A\Lambda^{-1}. \tag{46}$$

A transforms every projection operator into the projection operator which corresponds to the conjugate complex domain

$$AE_1A^{-1}=E_1 \quad AE_{-1}A^{-1}=E_{-1} \quad AE_JA^{-1}=E_{J^*}. \tag{47}$$

These equations can be given a variety of forms by combining them with (43) and (44). The most interesting of these forms gives the projection operators in terms of the A_J. Thus

$$A_1^2=AE_1AE_1=E_1A^2E_1=E_1\Lambda E_1=E_1^2=E_1. \tag{48}$$

Similarly,

$$A_{-1}^2=-E_{-1} \quad \lim A_{J^*}A_J=\omega E_J. \tag{49}$$

Whereas, if J and L do not overlap,

$$A_{L^*}A_J=0. \tag{50}$$

These equations form a substitute for the equations involving the characteristic vectors v of A. As an example, we show that v_{-1k} and Av_{-1k} are orthogonal or, in the present language, that $E_{-1}\varphi$ and $AE_{-1}\varphi$ are orthogonal for any φ

$$\begin{aligned}(E_{-1}\varphi,AE_{-1}\varphi)&=(A^2E_{-1}\varphi,AE_{-1}\varphi)\\&=(-E_{-1}\varphi,AE_{-1}\varphi)=0.\end{aligned} \tag{51}$$

The second form follows from the antiunitary nature of A, the third from (43).

Phenomenological Distinction between Unitary and Antiunitary Symmetry Operators

EUGENE P. WIGNER
Palmer Physical Laboratory, Princeton University, Princeton, New Jersey
(Received April 25, 1960)

It is well known that one always can find as many orthogonal states (i.e., states between which the transition probability is zero) as the Hilbert space has dimensions which are invariant under a given unitary transformation. The corresponding vectors are characteristic vectors of the unitary operator. In contrast, most antiunitary operators leave not more than one state invariant. However, if there are two orthogonal invariant states, a consideration of the states for which the transition probability is ½ into both invariant states surely provides a distinction. In the antiunitary case, one of these states is also invariant, another one is transformed into an orthogonal state, the rest are in between. In the unitary case, the transition probability between original state and transformed state is the same for all states for which the transition probability is ½ into two orthogonal states. This provides a "directly observable" distinction between unitary and antiunitary transformations.

1.

THE invariance transformations of quantum mechanics are transformations in a complex Hilbert space which leave the absolute value of the scalar product of any two vectors φ, ψ invariant:

$$|(\varphi,\psi)| = |(T\varphi,T\psi)|. \tag{1}$$

The reason for the invariance of (1) is that this absolute value (or, rather, its square) is, according to the usual physical interpretation of quantum mechanics, operationally meaningful: It is the transition probability between the two states characterized by the two vectors φ and ψ. It is well known that the transformations T which satisfy (1) fall into one of two categories: They can be unitary or antiunitary.

The last statement should be formulated more precisely. The physical state does not determine the state vector φ completely, all multiples of φ (the whole "ray of φ") describe the same state. It is customary and useful to normalize the state vectors, i.e., to choose a vector from the ray of φ which is normalized,

$$(\varphi,\varphi)=1. \tag{2}$$

Even then, a phase factor (i.e., a factor of modulus 1) remains free in φ. The same applies to the transformed state, the ray of $T\varphi$. Note that (1) remains valid if φ, ψ, $T\varphi$, and $T\psi$ are multiplied by arbitrary phase factors. The precise formulation of the statement at the end of the preceding paragraph stipulates the possibility of a certain choice of the state vector from the ray of the transformed state, which corresponds to any choice of a state vector φ from the ray of the original state. The choice referred to therefore replaces the physically given ray correspondence by a vector correspondence. The theorem states that the choice can be made either in such a way that for any two state vectors φ, ψ and any two numbers a and b,

$$T(a\varphi+b\psi)=aT\varphi+bT\psi \quad \text{and} \quad (\varphi,\psi)=(T\varphi,T\psi) \tag{3a}$$

(unitary case), or in such a way that

$$T(a\varphi+b\psi)=a^*T\varphi+b^*T\psi \quad \text{and} \quad (\varphi,\psi)=(T\psi,T\varphi). \tag{3b}$$

In the latter case T (and also the ray correspondence from which it derives) is called antiunitary.

It is easy to show, by abstract reasoning, that no transformation can fall into both classes, i.e., that if the choice of $T\varphi$ from its ray can be made in such a way that (3a) becomes valid, it cannot be made in such a way that (3b) becomes valid, and conversely.

It follows that it must be possible to ascertain the unitary or antiunitary nature of a transformation by considering only transition probabilities, i.e., absolute values of scalar products. However, it may be of some interest to spell this out in detail, and it is the aim of the present article to do this. Use will be made, for this purpose, of the normal form of antiunitary operators obtained in the preceding article.[1] Similarly, the unitary operators will be described by their characteristic values and the corresponding characteristic vectors.

We note for further reference that if φ and ψ are normalized state vectors, they represent the same state if $\varphi=\omega\psi$ with an arbitrary ω of modulus 1. In this case $|(\varphi,\psi)|=1$. On the other hand, if

$$(\varphi,\psi)=0, \tag{4}$$

φ and ψ represent states which have, at least in some respect, opposite properties. If (4) holds for two states, we call them orthogonal: The transition probability between them is zero. Orthogonality is, therefore, an operationally verifiable relation between two states. The number of the mutually orthogonal states is equal to the dimension of the underlying Hilbert space. We shall assume henceforth that this is larger than 1.

[1] E. P. Wigner, J. Math. Phys. 1, 409 (1960), preceding article.

2.

Let us consider, first, the states which are left unchanged by the transformation. For these,

$$T\varphi=\omega\varphi, \tag{5}$$

and we shall call them invariant states. If T is unitary, φ is one of its characteristic vectors. It follows that there is at least one set of mutually orthogonal states which contains as many members as *any* set of mutually orthogonal states contains. In particular, unless there are at least two orthogonal invariant states, the symmetry operator cannot be unitary and must be, therefore, antiunitary.

If the spectrum of the unitary operator is simple, the invariant states are isolated, but if it has a characteristic value of multiplicity l, the corresponding invariant states form a continuous manifold with $2l-2$ parameters. The characteristic functions have l complex or $2l$ real parameters but the normalization condition subjects these to one real equation and one real parameter, characterizing the phase factor, is physically meaningless.

Let us consider now an antiunitary transformation. We decompose the state vector into the invariant and characteristic vectors of the operator T

$$\varphi=\sum_k a_k v_{1k}+\sum_{\omega k} b_{\omega k} v_{\omega k}. \tag{6}$$

If $T\varphi=\omega'\varphi$ (i.e., φ is an invariant state), we can consider $\varphi'=\omega'^{\frac{1}{2}}\varphi$ and have

$$T\varphi'=(\omega'^{\frac{1}{2}})^*T\varphi=(\omega'^{\frac{1}{2}})^*\omega\varphi=(\omega'^{\frac{1}{2}})^*\varphi=\varphi'. \tag{7}$$

Since φ and $\omega'^{\frac{1}{2}}\varphi$ represent the same state, it suffices, in the antiunitary case, to find those state vectors for which

$$T\varphi=\varphi. \tag{8}$$

Since, for the φ of (6)

$$T\varphi=\sum a_k^* v_{1k}+\sum b_{\omega k}^*(\omega^{\frac{1}{2}})^* v_{\omega^* k}$$

$$=\sum a_k^* v_{1k}+\sum b_{\omega^* k}^*\omega^{\frac{1}{2}} v_{\omega k}, \tag{9}$$

φ will satisfy (8) if

$$a_k=a_k^* \tag{10}$$

and

$$\omega^{\frac{1}{2}} b_{\omega^* k}^*=b_{\omega k}. \tag{11}$$

Since (11) must hold for all ω and all k, ω can be replaced therein by ω^* to give

$$(\omega^*)^{\frac{1}{2}} b_{\omega k}^*=b_{\omega^* k} \quad \text{or} \quad b_{\omega^* k}^*=\omega^{\frac{1}{2}} b_{\omega k}. \tag{12}$$

Insertion of the latter expression into (11) gives

$$\omega b_{\omega k}=b_{\omega k} \quad \text{or} \quad b_{\omega k}=0 \tag{13}$$

since $\omega\neq 1$. It follows that if the symmetry operator is antiunitary, the invariant states can be described by invariant vectors, i.e., by real linear conbinations of the v_{1k}.

It follows that the structure of the set of invariant states is, in general, very different for unitary and antiunitary operators. An antiunitary operator may have no invariant state, or it may have only one. However, no matter how many invariant states it has, they form a continuous manifold each member of which can be changed continuously into any other. If the antiunitary operator has n orthogonal invariant vectors, the manifold of invariant states is an $n-1$ dimensional continuous manifold. Its state vector can be described by n real coefficients the sum of the squares of which is 1. There is no arbitrary phase factor in this case because (8) already determines the phase factor.

3.

The number and topological properties of the invariant states actually permit a phenomenological distinction between unitary and antiunitary transformations. The following distinction is, however, more direct.

Consider two orthogonal invariant states, φ_1 and φ_2. If there are no such, the transformation is surely antiunitary. Next, consider the states for which the transition probability is $\frac{1}{2}$ into both φ_1 and φ_2. The state vectors of these states are

$$\psi_\alpha=2^{-\frac{1}{2}}(\varphi_1+e^{i\alpha}\varphi_2). \tag{14}$$

Consider finally the transition probability of these states into the states $T\psi_\alpha$,

$$P(\alpha)=|(\psi_\alpha,T\psi_\alpha)|^2. \tag{15}$$

If T is unitary,

$$T\psi_\alpha=2^{-\frac{1}{2}}(\omega_1\varphi_1+e^{i\alpha}\omega_2\varphi_2), \tag{16}$$

and the transition probability becomes

$$P(\alpha)=\tfrac{1}{4}|(\varphi_1+e^{i\alpha}\varphi_2,\ \omega_1\varphi_1+e^{i\alpha}\omega_2\varphi_2)|^2=\tfrac{1}{4}|\omega_1+\omega_2|^2$$

$$=\tfrac{1}{2}(1+\mathrm{Re}\omega_1\omega_2^*). \tag{17}$$

It is independent of α, i.e., the same for all the states ψ_α with transition probabilities $\frac{1}{2}$ into φ_1 and φ_2.

Let us assume next that T is antiunitary. In this case

$$T\psi_\alpha=2^{-\frac{1}{2}}(\varphi_1+e^{-i\alpha}\varphi_2), \tag{18}$$

and the transition probability becomes

$$P(\alpha)=\tfrac{1}{4}|(\varphi_1+e^{i\alpha}\varphi_2,\ \varphi_1+e^{-i\alpha}\varphi_2)|^2=\tfrac{1}{4}|1+e^{-2i\alpha}|^2$$

$$=\tfrac{1}{2}(1+\cos 2\alpha). \tag{19}$$

It varies, for the states in question, between 0 and 1.

The preceding argument can yet be greatly generalized. In the unitary case, if the transition probabilities into n orthogonal invariant states $\varphi_1, \varphi_2, \cdots, \varphi_n$ are prescribed to be $r_1^2, r_2^2, \cdots, r_n^2$ so

that

$$\sum r_k^2=1, \tag{20}$$

the general form of the state vector is

$$\psi_{\alpha_1\cdots\alpha_n}=\sum r_k e^{i\alpha k}\varphi_k, \tag{21}$$

with arbitrary $\alpha_1, \cdots, \alpha_n$. Nevertheless, the transition probability from any of these $\psi_{\alpha_1\cdots\alpha_n}$ into the corresponding $T\psi_{\alpha_1\cdots\alpha_n}$ is the same, namely,

$$\begin{aligned}|(\psi_{\alpha_1\cdots\alpha_n}, T\psi_{\alpha_1\cdots\alpha_n})|^2 &= |(\sum r_k e^{i\alpha k}\varphi_k, \sum r_k e^{i\alpha k}\omega_k\varphi_k)|^2 \\ &= |\sum r_k^2\omega_k|^2. \end{aligned}\tag{22}$$

In the antiunitary case, on the other hand,

$$T\psi_{\alpha_1\cdots\alpha_n}=\sum r_k e^{-i\alpha k}\varphi_k, \tag{23}$$

and the transition probability becomes

$$|(\psi_{\alpha_1\cdots\alpha_n}, T\psi_{\alpha_1\cdots\alpha_n})|^2=|\sum r_k^2 e^{-2i\alpha k}|^2. \tag{24}$$

Unless one of the r_k^2 is larger than $\frac{1}{2}$, this still will assume every value between 0 and 1 for suitably chosen α.

The striking difference in the relation of the original and transformed states to each other shows particularly clearly how definite the relations in question are in either case.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 1, NUMBER 5 SEPTEMBER-OCTOBER, 1960

Real Representations of Coordinate Rotations

U. FANO
National Bureau of Standards, Washington, D. C.
(Received April 8, 1960)

Since irreducible tensorial sets that represent observables are of integral degree, their transformations under coordinate rotations have real representations. Real representations, with rows and columns classified by eigenvalues of the commuting operators J_z^2 and $\exp(i\pi J_y)$, are given explicitly, so that complex functions of rotation angles need not be used. The addition of angular momenta is worked out for sets in the real representation such as the sets of real orbital wave functions. Applications to the theory of angular distributions are discussed.

1. INTRODUCTION

THE quantum theory of the coupling and recoupling of angular momenta constitutes an extension of vector algebra which has applications in macroscopic as well as in atomic and nuclear physics. The elements of this algebra have been called "irreducible tensorial sets" in a recent publication.[1] (Eigenstates of an atomic system with a given angular momentum quantum number j and different magnetic quantum numbers m constitute a typical example of irreducible tensorial set of $2j+1$ elements and of "degree" j.) In the same reference it was emphasized that calculations with tensorial sets can be carried to their final analytical result treating each set as a unit, as a vector is usually treated; that is, without considering explicitly the individual elements of the sets or their transformation law under coordinate rotations. Set elements and their transformations become relevant when a final result, for example an angular distribution of nuclear radiation, has to be evaluated numerically. Thus, even though irreducible sets can be expressed in different alternative representations, related to one another by unitary matrices, the choice of a specific representation may become relevant only at the last stage of a calculation.

In accordance with current practice of quantum physics, FR[1] considered primarily a "standard" representation of irreducible sets in which each set element is an eigenstate of coordinate rotations about the z axis, and therefore gets merely multiplied by $\exp(im\varphi)$ under such rotations. This standardization causes the transformation law—which is an irreducible representation of the rotation group—to be complex. Complex transformations are unavoidable when one deals with the wave functions of Fermi particles, but it should be noticed that no wave functions appear in the final results of many calculations with tensorial sets. Final results involve probabilities rather than probability amplitudes and often take the form of invariant products of tensorial sets of *integral degree*, which represent observables, such as, e.g., the response characteristics of a polarimeter. Tensorial sets of integral degree and their transformations under coordinate rotations which relate, e.g., the orientations of different parts of an experimental arrangement can be expressed in *real form*. This possibility should enable physicists to apply the results of analytical calculations without any use of complex numbers.

It may therefore be worthwhile to describe the characteristics of real irreducible tensorial sets and the explicit form of their transformation law. Examples of irreducible real sets of degree l are the set of real components of the electric 2^l-pole moment of a system of charges and the set of $2l+1$ spherical harmonics $\mathcal{P}_{lm}(\theta)\cos m\varphi$, $\mathcal{P}_{lm}(\theta)\sin m\varphi$. It is perhaps a curious historical accident that the properties of these sets do not seem to have been described systematically thus far.

Irreducible sets in "standard" form have a simple structure, in that each of their elements is classified by an eigenvalue of the infinitesimal rotation operator J_z. It is the diagonalization of this operator that causes the transformation under rotations about the z axis, represented by the operator $\exp(i\varphi J_z)$, to be complex. To analyze the structure of real sets with real transformations, we shall have to diagonalize two different operators rather than a single one, so that two quantum numbers will be required to classify each element of a real set. As a compensation, the new scheme of classification emphasizes reflection symmetries that tend to be overlooked in the standard formulation and reduces the matrices of rotations about the y axis to pairs of separate submatrices.

Our first task will consist of providing a suitable classification for the elements of a real set (Sec. 2). As a next step, the unitary matrix Q will be constructed that serves to transform standard (or contrastandard) sets of integral degree into a "real-standard" form, or vice versa (Sec. 3). This matrix will then be applied to the transformation law of irreducible sets under coordinate rotations (i.e., to the irreducible representations of the rotation group) to convert them from their usual standard form into a real-standard form (Sec. 4). Finally, it will be shown in Sec. 5 how irreducible products of irreducible real sets are constructed by means of Wigner coefficients transformed by the matrix Q. (This procedure enables one, e.g., to carry out the addition of orbital momenta of atomic particles starting from real

[1] U. Fano and G. Racah, *Irreducible Tensorial Sets* (Academic Press, Inc., New York, 1959); this book will be referred to as FR.

wave functions.) Some applications will be discussed in Sec. 6.

The material thus developed may be regarded as a complement to Chaps. 5–7 of FR. The treatment of multiple coupling and recoupling of irreducible sets in the following chapters of FR applies equally to irreducible sets in different representations. The results should be particularly relevant to applications of Chap. 19 of FR.

2. QUANTUM NUMBERS

Since the elements of an irreducible tensorial set experience a linear transformation (called "r-transformation" in FR, Chap. 2) when the system of space coordinates rotates, there is a one-to-one correspondence between the set elements and the rows or columns of the r-transformation matrices. It is convenient to study first the structure of these matrices, and in particular of the related Hermitian matrices J_x, J_y, and J_z pertaining to infinitesimal rotations, and then to classify the tensorial set elements accordingly (FR, Chap. 3 and Appendix B). Owing to the commutation laws $J_xJ_y-J_yJ_x=iJ_z$, etc., any two of these matrices determine the third one; in particular, if two of the matrices are imaginary the third one is imaginary too. We shall focus on J_y and J_z since a general rotation is expressed in terms of Euler angles by the product of operators $\exp(i\psi J_z)\exp(i\theta J_y)\exp(i\varphi J_z)$.

As pointed out in the foregoing, the matrix J_z cannot be diagonal and different from zero in a real representation, because the r-transformation matrix $\exp(i\varphi J_z)$ would then be complex. Indeed, all three matrices J_x, J_y, J_z must be imaginary in a real representation (FR, Appendix C). Now, if J_z is imaginary, J_z^2 will be real. Therefore J_z^2 can be real in a real representation. We shall consider in this paper a real-standard representation in which J_z^2 is diagonal and in which its eigenvalues are indicated by m^2. When one deals with irreducible sets of integral degree l and with J_z diagonal, the eigenvalues of J_z run from $-l$ to l. Therefore, J_z^2 has l doubly degenerate eigenvalues m^2, where m is an integer between 1 and l, and the nondegenerate eigenvalue $\mathrm{m}^2=0$. Accordingly, we shall use m as a "quantum number," i.e., as an index to label the elements of a real-standard irreducible set. A single set element will be labeled with $\mathrm{m}=0$ and l pairs of elements with each of the other values of m. Thereby, the element with $\mathrm{m}=0$ is completely identified; this element is also an eigenvector of J_z, corresponding to $J_z=0$, and remains invariant under coordinate rotations about the z axis.

An additional index is required to distinguish the set elements of each pair with $\mathrm{m}\neq 0$. This index must be an eigenvalue of an operator that commutes with J_z^2. One such operator is $\exp(i\pi J_y)$, which represents a rotation of 180° about the y axis, and therefore changes J_z into $-J_z$. Since $[\exp(i\pi J_y)]^2=1$, for a set of integral degree, $\exp(i\pi J_y)$ has only the two eigenvalues ± 1. That is, $\exp(i\pi J_y)$ represents a parity operation.

The single set element with $\mathrm{m}=0$ has the same parity with respect to $\exp(i\pi J_y)$ as the element with $J_z=0$ of a standard set of the same order. This parity is $(-1)^l$ (see FR, p. 23). We shall classify the elements of a real standard set with $\mathrm{m}\neq 0$ according to their parity *relative* to that of the element with $\mathrm{m}=0$. That is, we represent the eigenvalues of $\exp(i\pi J_y)$ in the form $(-1)^lP$, where $P=\pm 1$ serves as a parity quantum number.

A real-standard set of degree l will be indicated by a boldface Roman letter with a superior index, e.g., $\mathbf{a}^{(l)}$, and its individual elements by the corresponding nonboldface letter with subscripts m and P,

$$\mathrm{a}^{(l)}{}_{\mathrm{m}P}, \tag{1}$$

where $P=\pm 1$ for $\mathrm{m}\neq 0$ and $P=1$ for $\mathrm{m}=0$. The index P may be indicated simply by $+$ or $-$, if convenient, and may be omitted for $\mathrm{m}=0$.

The coordinate rotation represented by $\exp(i\pi J_y)$ is related to coordinate reflections. The operation I of coordinate inversion (i.e., reflection at the origin) combined with $\exp(i\pi J_y)$ yields a reflection on the xz plane,

$$R_{xz}=I\exp(i\pi J_y). \tag{2}$$

A tensorial set of degree l constructed as the product of l polar vectors has the parity $I=(-1)^l$. The elements of this set have then the parity $R_{xz}=(-1)^l(-1)^lP=P$. In general an element of a real-standard set has parity $\pm P$ with respect to R_{xz} according to whether $I=\pm(-1)^l$. When the lower sign obtains, as it does for the sets of components of axial vectors, the set might be said to be of "pseudodegree" l, in analogy with the fact that axial vectors are called pseudovectors.

Notice that the operators J_z^2 and $\exp(i\pi J_y)$ commute, respectively, with the operators $\exp(i\varphi J_z)$ and $\exp(i\theta J_y)$, which represent rotations about the z and y axes. Therefore m and P are, respectively, "good" quantum numbers under these separate operations. This is important because any rotation is usually resolved into a sequence of Euler angle rotations about these two axes. That is, the matrix $\exp(i\varphi J_z)$ reduces to l submatrices corresponding to the eigenvalues of J_z^2 and the matrix $\exp(i\theta J_y)$ to two submatrices corresponding to the two eigenvalues of $\exp(i\pi J_y)$.

3. MATRIX $Q^{(l)}$

In this section we construct the matrix $Q^{(l)}$ that transforms a standard set of degree l into a real-standard one as indicated by the formula

$$\mathbf{a}^{(l)}=Q^{(l)}\mathfrak{a}^{(l)}. \tag{3}$$

A contrastandard set (see FR, pp. 22–23) is transformed according to the corresponding formula

$$\mathfrak{b}^{(l)}=Q^{(l)*}\mathfrak{b}^{[l]}=\mathfrak{b}^{[l]}Q^{(l)-1}. \tag{3'}$$

Each row of $Q^{(l)}$ is labeled by a pair of values of the two quantum numbers m, P, except that P is automatically $+1$ and may be omitted when $\mathrm{m}=0$. Each

column could be labeled by a value of the single quantum number m which is an eigenvalue of J_z and identifies an element of a standard or contrastandard set. However, since J_z^2 is also diagonal when J_z is, we may set

$$m=S\mathrm{m}, \quad S=\pm 1, \tag{4}$$

where m^2 is an eigenvalue of J_z^2. Each column of $Q^{(l)}$ will then also be labeled with a pair of values of the two quantum numbers m, S, except that S is automatically $+1$ and may be omitted when $\mathrm{m}=0$.

Since the elements of both standard and real-standard sets are eigenvectors of J_z^2, the matrix $Q^{(l)}$ commutes with J_z^2 and accordingly reduces to submatrices, one for each value of m. The single elements with $m=\mathrm{m}=0$ of a standard set is only multiplied by a phase factor N_l when the set is transformed by $Q^{(l)}$. It will be seen that the whole matrix depends on l only through this factor, so that we have

$$(\mathrm{m}'P|Q^{(l)}|\mathrm{m}S)=N_l\,(P|S)^{(\mathrm{m})}\,\delta_{\mathrm{m}'\mathrm{m}}, \tag{5}$$

where $(P|S)^{(\mathrm{m})}$ indicates a unitary 2×2 matrix for $\mathrm{m}\neq 0$, but is just 1 for $\mathrm{m}=0$, and where $|N_l|^2=1$.

To complete the calculation of the matrix $Q^{(l)}$ it suffices to specify a set of spherical harmonics $\mathbf{Y}^{(l)}(\theta,\varphi)$ that shall be regarded as real-standard and observe how it relates to the contrastandard $\mathfrak{Y}^{[l]}(\theta,\varphi)$ described in FR, p. 25. There is some arbitrariness in the choice of the relative sign of the different elements of $\mathbf{Y}^{(l)}(\theta,\varphi)$. Our choice will be such that the three elements of $\mathbf{Y}^{(l)}(\theta,\varphi)$ coincide with the Cartesian components of a unit vector in the direction (θ,φ). Indicating the two values of P simply by $+$ or $-$, we set

$$\mathrm{Y}^{(l)}{}_{\mathrm{m}+}(\theta,\varphi)=\pi^{-\frac{1}{2}}\mathcal{P}_{l\mathrm{m}}(\theta)\,\mathrm{cosm}\varphi, \tag{6a}$$

$$\mathrm{Y}^{(l)}{}_{\mathrm{m}-}(\theta,\varphi)=\pi^{-\frac{1}{2}}\mathcal{P}_{l\mathrm{m}}(\theta)\,\mathrm{sinm}\varphi, \tag{6b}$$

$$\mathrm{Y}^{(l)}{}_{0}(\theta,\varphi)=(2\pi)^{-\frac{1}{2}}\mathcal{P}_{l0}(\theta) =[(2l+1)/4\pi]^{\frac{1}{2}}\,P_l(\cos\theta), \tag{6c}$$

where the $\mathcal{P}_{l\mathrm{m}}$ are the associated Legendre polynomials normalized according to Bethe.[2] Since the contrastandard harmonics of FR are

$$\mathfrak{Y}^{[l]}{}_{m}=i^l(-1)^m(2\pi)^{-\frac{1}{2}}\mathcal{P}_{lm}(\theta)\,\exp(im\varphi)$$

[where the factor $(-1)^m$ stems from the Condon-Shortley normalization], and since (3′) implies that $\mathfrak{Y}^{[l]}=\mathbf{Y}^{(l)}Q^{(l)}$, the matrix $Q^{(l)}$ is easily seen to be given by (5) with

$$N_l=i^l, \quad (P|S)^{(\mathrm{m})}=\sqrt{\tfrac{1}{2}}\begin{vmatrix}(-1)^{\mathrm{m}} & 1\\ (-1)^{\mathrm{m}}i & -i\end{vmatrix} \quad (\text{for } \mathrm{m}\neq 0). \tag{7}$$

4. r-TRANSFORMATIONS

The real-standard r-transformation matrix of degree l is given by $Q^{(l)}\mathfrak{D}^{(l)}Q^{(l)-1}$, where $\mathfrak{D}^{(l)}$ is the standard matrix on page 22 of FR and $Q^{(l)}$ is given by (5) and (7).

[2] H. A. Bethe and E. E. Salpeter, *Encyclopedia of Physics* (Springer-Verlag, Berlin, Germany, 1957), Vol. 35I, p. 431.

Any r-transformation matrix expressed as a function of Euler angles factors out (see, e.g., FR, p. 140) in the form

$$D(\psi,\theta,\varphi)=e^{i\psi J_z}e^{i\theta J_y}e^{i\varphi J_z}. \tag{8}$$

The first and last factors of this expression are trivially simple only in the representations where J_z is diagonal. Therefore we shall transform each factor of (8) separately.

As a preliminary it may be worthwhile to write the real-standard forms of J_z and J_y. From Eq. (5.1) of FR and from (5) and (7) of the foregoing, we find

$$\begin{aligned}(\mathrm{m}'P'|J_z|\mathrm{m}P)&=\textstyle\sum_{S'S}(P'|S')^{(\mathrm{m}')}(\mathrm{m}'S'|J_z|\mathrm{m}S)(S|P)^{(\mathrm{m})}\\&=\textstyle\sum_{S}(P'|S)^{(\mathrm{m})}S\mathrm{m}(S|P)^{(\mathrm{m})}\delta_{\mathrm{m}'\mathrm{m}}\\&=i\mathrm{m}P\delta_{\mathrm{m}'\mathrm{m}}\delta_{-P'P}, \quad \text{for} \quad \mathrm{m}\neq 0,\end{aligned} \tag{9}$$

$$\begin{aligned}(\mathrm{m}+1,P'|J_y|\mathrm{m}P)&=\textstyle\sum_{S'S}(P'|S')^{(\mathrm{m}+1)}(\mathrm{m}+1,S'|J_y|\mathrm{m}S)(S|P)^{(\mathrm{m})}\\&=\textstyle\sum(P'|S)^{(\mathrm{m}+1)}(-iS)\frac{1}{2}[l(l+1)-\mathrm{m}(\mathrm{m}+1)]^{\frac{1}{2}}(S|P)^{(\mathrm{m})}\\&=i\tfrac{1}{2}[l(l+1)-\mathrm{m}(\mathrm{m}+1)]^{\frac{1}{2}}\delta_{P'P}, \quad \text{for} \quad \mathrm{m}\neq 0,\end{aligned} \tag{10a}$$

$$(1P|J_y|0)=\textstyle\sum(P|S)^{(1)}(1S|J_y|0)=i[\frac{1}{2}l(l+1)]^{\frac{1}{2}}\delta_{P1}. \tag{10b}$$

All other elements of these matrices vanish or are Hermitian conjugates of those given here.[3]

It follows readily from (9) that a coordinate rotation by an angle ψ about the z axis is represented by the real-standard r-transformation matrix

$$(\mathrm{m}'P'|e^{i\psi J_z}|\mathrm{m}P) =\delta_{\mathrm{m}'\mathrm{m}}[\delta_{P'P}\,\mathrm{cosm}\psi-\delta_{-P'P}P\,\mathrm{sinm}\psi]. \tag{11}$$

Therefore this rotation transforms a real-standard set $\mathbf{a}^{(l)}$ into a set $\mathbf{a}^{(l)\prime}$ whose elements are[4]

$$\mathrm{a}^{(l)}{}_{\mathrm{m}P}{}'=\mathrm{cosm}\psi\;\mathrm{a}^{(l)}{}_{\mathrm{m}P}+P\,\mathrm{sinm}\psi\;\mathrm{a}^{(l)}{}_{\mathrm{m}-P}. \tag{A}$$

The real-standard form of the matrix $\exp(i\theta J_y)$ will be obtained directly by transformation of the standard form whose elements are the functions $\mathfrak{d}^{(l)}{}_{m'm}(\theta)$ discussed in FR, Appendixes D–F. The $\mathfrak{d}$'s with positive or

[3] The magnitudes of the matrix elements (9) and (10) could have been obtained directly, i.e., without resorting to transformation of the standard matrices, from the characterization of real-standard sets in Sec. 2 utilizing the commutation rules of the J matrices in analogy with the treatment in FR, Appendix B. The choice of signs in (9) and (10) involves an arbitrary convention equivalent to that performed in (6). Therefore the values (9) and (10) of the real-standard matrix elements could have been taken as a point of departure instead of (6), and the values of the matrix elements $(P|S)^{(\mathrm{m})}$ could have then been obtained by solving (9) and (10). The value of N_l can be determined, to within an arbitrary sign which is otherwise fixed by comparing (6) with (5.18) of FR, from Eq. (C.11) of FR which requires that $(\mathrm{m}'S'|\tilde{Q}^{(l)}Q^{(l)}|\mathrm{m}S) =(\mathrm{m}'S'|\mathfrak{U}^{(l)}|\mathrm{m}S)=(-1)^{l-\mathrm{m}}\delta_{\mathrm{m}'\mathrm{m}}\delta_{-S'S}$.

[4] When this transformation is applied to the set $\mathbf{Y}^{(l)}(\theta,\varphi)$, it should be kept in mind that the coordinate rotation by ψ about the z axis charges the variable φ of these harmonics into $\varphi-\psi$ rather than into $\varphi+\psi$. This remark applies also to rotations about the y axis.

TABLE I. The matrices $d^{(l)}_{mm'P}(\theta)$ for $l=1, 2, 3$.

$l=1 \quad P=1$

m \ m′	1	0
1	$\cos\theta$	$-\sin\theta$
0	$\sin\theta$	$\cos\theta$

$l=1 \quad P=-1$

m \ m′	1
1	1

$l=2 \quad P=1$

m \ m′	2	1	0
2	$\frac{1}{2}(\cos^2\theta+1)$	$-\cos\theta\sin\theta$	$\frac{1}{2}\sqrt{3}\sin^2\theta$
1	$\cos\theta\sin\theta$	$2\cos^2\theta-1$	$-\sqrt{3}\cos\theta\sin\theta$
0	$\frac{1}{2}\sqrt{3}\sin^2\theta$	$\sqrt{3}\cos\theta\sin\theta$	$\frac{1}{2}(3\cos^2\theta-1)$

$l=2 \quad P=-1$

m \ m′	2	1
2	$\cos\theta$	$-\sin\theta$
1	$\sin\theta$	$\cos\theta$

$l=3 \quad P=1$

m \ m′	3	2	1	0
3	$\frac{1}{4}(\cos^2\theta+3)\cos\theta$	$-\sqrt{\frac{3}{8}}(\cos^2\theta+1)\sin\theta$	$\frac{1}{4}\sqrt{15}\cos\theta\sin^2\theta$	$-\sqrt{\frac{5}{8}}\sin^3\theta$
2	$\sqrt{\frac{3}{8}}(\cos^2\theta+1)\sin\theta$	$\frac{1}{2}(3\cos^2\theta-1)\cos\theta$	$-\sqrt{\frac{5}{8}}(3\cos^2\theta-1)\sin\theta$	$\frac{1}{2}\sqrt{15}\cos\theta\sin^2\theta$
1	$\frac{1}{4}\sqrt{15}\cos\theta\sin^2\theta$	$\sqrt{\frac{5}{8}}(3\cos^2\theta-1)\sin\theta$	$\frac{1}{4}(15\cos^2\theta-1)\cos\theta$	$-\sqrt{\frac{3}{8}}(5\cos^2\theta-1)\sin\theta$
0	$\sqrt{\frac{5}{8}}\sin^3\theta$	$\frac{1}{2}\sqrt{15}\cos\theta\sin^2\theta$	$\sqrt{\frac{3}{8}}(5\cos^2\theta-1)\sin\theta$	$\frac{1}{2}(5\cos^2\theta-3)\cos\theta$

$l=3 \quad P=-1$

m \ m′	3	2	1
3	$\frac{1}{4}(3\cos^2\theta+1)$	$-\sqrt{\frac{3}{2}}\cos\theta\sin\theta$	$\frac{1}{4}\sqrt{15}\sin^2\theta$
2	$\sqrt{\frac{3}{2}}\cos\theta\sin\theta$	$2\cos^2\theta-1$	$-\sqrt{\frac{5}{2}}\cos\theta\sin\theta$
1	$\frac{1}{4}\sqrt{15}\sin^2\theta$	$\sqrt{\frac{5}{2}}\cos\theta\sin\theta$	$\frac{1}{4}(5\cos^2\theta-1)$

negative indices (m',m) can be expressed in terms of those with positive indices only, evaluated at θ and $\pi-\theta$, utilizing the symmetry properties represented by Eqs. (D.24, 25) of FR. We have, for m′, m$\neq$0,

$$(m'S'|e^{i\theta J_y}|mS) = \mathfrak{d}^{(l)}_{S'\cdot m',S\cdot m}(\theta)=\delta_{S'S}S^{m'-m}\mathfrak{d}^{(l)}_{m'm}(\theta) + (-1)^{l-m}\delta_{-S'S}S^{m'-m}\mathfrak{d}^{(l)}_{m'm}(\pi-\theta), \quad (12a)$$

and, for m′=0,

$$(0|e^{i\theta J_y}|mS)=\mathfrak{d}^{(l)}_{0,S\cdot m}(\theta)=S^m\,\mathfrak{d}^{(l)}_{0,m}(\theta). \quad (12b)$$

Transformation with the matrix $Q^{(l)}$ yields now

$$(m'P'|e^{i\theta J_y}|mP) = \delta_{P'P}\,d^{(l)}_{m'mP}(\theta) = \delta_{P'P}\{(-1)^{m'-m}\mathfrak{d}^{(l)}_{m'm}(\theta)+(-1)^l P\,\mathfrak{d}^{(l)}_{m'm}(\pi-\theta)\}, \quad \text{for } m', m\neq 0, \quad (13a)$$

$$(0|e^{i\theta J_y}|mP) = \delta_{P1}\,d^{(l)}_{0m+}(\theta)=\delta_{P1}(-1)^m\sqrt{2}\,\mathfrak{d}^{(l)}_{0m}(\theta) = \delta_{P1}\left(\frac{4}{2l+1}\right)^{\frac{1}{2}}\mathcal{P}_{lm}(\theta), \quad \text{for } m\neq 0, \quad (13b)$$

and finally

$$(0|e^{i\theta J_y}|0)=d^{(l)}_{00}(\theta)=\mathfrak{d}^{(l)}_{00}(\theta)=P_l(\cos\theta). \quad (13c)$$

These equations serve as definitions of the new functions $d^{(l)}_{m'mP}(\theta)$, in which the value of P will be indicated by a + or − sign. According to (13), a coordinate rotation by an angle θ about the y axis transforms a real-standard set $\mathbf{a}^{(l)}$ into a set $\mathbf{a}^{(l)\prime}$ whose elements are[4]

$$a^{(l)}_{m'P}{}' = \textstyle\sum_m d^{(l)}_{m'mP}(\theta)\,a^{(l)}_{mP}. \quad (B)$$

We have thus obtained separate transformation formulas for the set elements with different values of P, as anticipated in Sec. 2.

As a result of the factoring represented by (8), any real-standard r-transformation can be carried out by repeated application of Eqs. (A) and (B). Whereas (A) involves only sines and cosines, (B) involves the functions $d^{(l)}_{m'mP}(\theta)$, which are defined by (13) in terms of the $\mathfrak{d}^{(l)}_{m'm}(\theta)$ and can be calculated using the algebraic expression for the $\mathfrak{d}^{(l)}$ on page 143 of FR. Table I gives the $d^{(l)}_{m'mP}$ in terms of sines and cosines for $l=1$, 2, and 3.[5] Notice that the $d^{(l)}$ have, like the $\mathfrak{d}^{(l)}$, the symmetry property[6]

$$d^{(l)}_{mm'}(\theta)=(-1)^{m-m'}d^{(l)}_{m'm}(\theta). \quad (14)$$

[5] The functions $d^{(l)}_{mm'P}$ are not normalized independently of each other, because the normalization property of r-transformations, given by Eq. (A.32) of FR, concerns only the complete matrix $D^{(l)}(\psi,\theta,\varphi)$. The resulting normalization of the $d^{(l)}$ is $\sum_P\int_0^\pi[d^{(l)}_{m'mP}(\theta)]^2\sin\theta\,d\theta=8/(2l+1)$, for m, m′$\neq$0, where the factor 8 makes allowance for the fact that the mean square of $\cos m\psi$ or $\sin m\psi$ is $\frac{1}{2}$.

[6] This property and the properties (D.23-25) of FR rest on simple geometrical considerations. Since the matrices $\mathbf{d}^{(l)}$ and $\mathfrak{d}^{(l)}$ are real, their transposition amounts to an inversion, i.e., to replacing $e^{i\theta J_y}$ by $e^{-i\theta J_y}$. The negative rotation $e^{-i\theta J_y}$ is in turn equivalent to the sequence of positive rotations $e^{i\pi J_z}e^{i\theta J_y}e^{i\pi J_z}$, where $e^{i\pi J_z}$ is diagonal and equal to $(-1)^m$ and $(-1)^m$, respectively, in the two representations. The remaining symmetries, including the symmetry of (13) under the replacement of θ by $\pi-\theta$, follow from the fact that $e^{-i\theta J_y}=e^{i\pi J_y}e^{i(\pi-\theta)J_y}$.

5. IRREDUCIBLE PRODUCTS

The concept of irreducible product of two irreducible tensorial sets (FR, Chap. 7) is, in essence, independent of the representation in which the set elements are specified. Accordingly, it applies to real-standard as well as to standard sets. However, the procedure for the actual construction of the elements of irreducible products has been described in FR and in other references only for standard or contrastandard sets. In order to construct irreducible products of real-standard sets one can convert these sets to their standard form, by means of the transformation Q^{-1}, take their product in accordance with FR, and then convert it back to real-standard form. We define, then, the irreducible product of degree l of two real-standard sets $\mathbf{a}^{(l_1)}$ and $\mathbf{b}^{(l_2)}$ by

$$[\mathbf{a}^{(l_1)}\times\mathbf{b}^{(l_2)}]^{(l)}=Q^{(l)}[\mathfrak{a}^{(l_1)}\times\mathfrak{b}^{(l_2)}]^{(l)}$$
$$=Q^{(l)}[Q^{(l_1)-1}\mathbf{a}^{(l_1)}\times Q^{(l_2)-1}\mathbf{b}^{(l_2)}]^{(l)}. \quad (15)$$

The construction of irreducible products of two standard sets, as described in FR, consists of reducing the elements of the reducible direct product set $\mathfrak{a}^{(l_1)}{}_{m_1}\mathfrak{b}^{(l_2)}{}_{m_2}$ by means of a unitary transformation matrix M which diagonalizes the operator $\mathbf{J}^2=(\mathbf{J}_1+\mathbf{J}_2)^2$. The elements of this matrix M are the Wigner coefficients $(l_1l_2lm|l_1m_1,l_2m_2)$. To construct the irreducible products (15) of real-standard sets directly we have to calculate the transformed matrix $Q^{(l)}MQ^{(l_1)-1}Q^{(l_2)-1}$. This matrix will reduce the direct product set $\mathrm{a}^{(l_1)}{}_{\mathrm{m}_1P_1}\mathrm{b}^{(l_2)}{}_{\mathrm{m}_2P_2}$.

We set, then,

$$[\mathbf{a}^{(l_1)}\times\mathbf{b}^{(l_2)}]^{(l)}{}_{\mathrm{m}P}$$
$$=\sum{}_{\mathrm{m}_1P_1\mathrm{m}_2P_2}(l_1l_2l\ \mathrm{m}\ P|l_1\mathrm{m}_1P_1,l_2\mathrm{m}_2P_2)$$
$$\times\mathrm{a}^{(l_1)}{}_{\mathrm{m}_1P_1}\,\mathrm{b}^{(l_2)}{}_{\mathrm{m}_2P_2}, \quad (C)$$

where

$$(l_1l_2l\ \mathrm{m}\ P|l_1\mathrm{m}_1P_1,l_2\mathrm{m}_2P_2)$$
$$=N_l\sum{}_{SS_1S_2}(P|S)^{(\mathrm{m})}\ (l_1l_2l\ S\mathrm{m}|l_1S_1\mathrm{m}_1,l_2S_2\mathrm{m}_2)$$
$$\times(S_1|P_1)^{(\mathrm{m}_1)}\ (S_2|P_2)^{(\mathrm{m}_2)}\ N_{l_1}{}^{-1}N_{l_2}{}^{-1}. \quad (16)$$

The transformation coefficient on the right of (16) is a Wigner coefficient in which $S\mathrm{m}=\pm\mathrm{m}$, etc. The result of the calculation indicated by (16) takes a somewhat different form depending on which among m, m_1, and m_2 is largest and on whether any of these numbers vanishes. It is

$$(l_1l_2l\mathrm{m}P|l_1\mathrm{m}_1P_1,l_2\mathrm{m}_2P_2)$$
$$=\frac{1}{\sqrt{2}}\cos\left[\frac{\pi}{2}(l_1+l_2-l+\delta_{-P_11}+\delta_{-P_21}-\delta_{-P1})\right]$$
$$\times(l_1l_2l\mathrm{m}|l_1\mathrm{m}_1,l_2\mathrm{m}_2),\quad\text{for}\quad \mathrm{m}>\mathrm{m}_1,\mathrm{m}_2, \quad (16a)$$

$$=\frac{1}{\sqrt{2}}\cos\left[\frac{\pi}{2}(l_1+l_2-l+\delta_{-P_11}-\delta_{-P_21}-\delta_{-P1}+2\mathrm{m}_2)\right]$$
$$\times(l_1l_2l\mathrm{m}|l_1\mathrm{m}_1,l_2-\mathrm{m}_2),\quad\text{for}\quad \mathrm{m}_1>\mathrm{m},\mathrm{m}_2, \quad (16b)$$

$$=\frac{1}{\sqrt{2}}\cos\left[\frac{\pi}{2}(l_1+l_2-l-\delta_{-P_11}+\delta_{-P_21}-\delta_{-P1}+2\mathrm{m}_1)\right]$$
$$\times(l_1l_2l\mathrm{m}|l_1-\mathrm{m}_1,l_2\mathrm{m}_2)\quad\text{for}\quad \mathrm{m}_2>\mathrm{m},\mathrm{m}_1, \quad (16c)$$

$$=\cos\left[\frac{\pi}{2}(l_1+l_2-l+\delta_{-P_11}-\delta_{-P1})\right]$$
$$\times(l_1l_2l\mathrm{m}|l_1\mathrm{m}_2,l_20),\quad\text{for}\quad \mathrm{m}_2=0, \quad (16d)$$

$$=\cos\left[\frac{\pi}{2}(l_1+l_2-l+\delta_{-P_21}-\delta_{-P1})\right]$$
$$\times(l_1l_2l\mathrm{m}|l_10,l_2\mathrm{m}_2),\quad\text{for}\quad \mathrm{m}_1=0, \quad (16e)$$

$$=\cos\left[\frac{\pi}{2}(l_1+l_2-l+\delta_{-P_11}-\delta_{-P_21}+2\mathrm{m}_2)\right]$$
$$\times(l_1l_2l0|l_1\mathrm{m}_1,l_2-\mathrm{m}_2),\quad\text{for}\quad \mathrm{m}=0, \quad (16f)$$

where the transformation coefficients on the right side are the ordinary Wigner coefficients.[7]

The cosine factor in Eqs. (16) equals ±1 or 0, and represents the effect of the selection rule that requires the transformation matrix elements to vanish unless they commute with $\exp(i\pi J_y)$; that is, unless

$$(-1)^lP=(-1)^{l_1+l_2}P_1P_2. \quad (17)$$

For given values of m, m_1, and m_2 other than zero and for a given P, this selection rule allows two alternative pairs of values of P_1, P_2; the normalization factor $1/\sqrt{2}$ in (16a–c) makes allowance for this alternative. No such alternative arises when one of the m vanishes, since the corresponding P is automatically 1.

The real-standard transformation matrix (16) enables one to reduce the set of components of any ordinary tensor of degree n, by the procedure indicated on page 41 of FR, without having to resort to an intermediate complex transformation. As an example we give here the result of the reduction of the set of nine components $T_{ik}(i,k=x,y,z)$ of a 2nd-degree tensor. This reduction parallels the construction of irreducible real-standard products $[\mathbf{a}^{(1)}\times\mathbf{b}^{(1)}]^{(l)}$, with $l=0$, 1, 2, of the sets of components of two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$. Specifically, the component T_{xy} of a tensor corresponds to the product a_xb_y of vector components; it can therefore be indicated by $\mathrm{T}_{1+,1-}$ just as a_x and b_y are indicated by $\mathrm{a}^{(1)}{}_{1+}$ and $\mathrm{b}^{(1)}{}_{1-}$, respectively, according to the conventions of Sec. 2. The reduction of the tensor components proceeds then through the formula analogous to (C):

$$\mathrm{T}^{(l)}{}_{\mathrm{m}P}=\sum{}_{\mathrm{m}_1P_1\mathrm{m}_2P_2}(11l\mathrm{m}P|1\mathrm{m}_1P_1,1\mathrm{m}_2P_2)$$
$$\times T_{\mathrm{m}_1P_1,\mathrm{m}_2P_2}, \quad (18)$$

[7] See, e.g., FR, p. 36, for definitions, notations, and symmetries, regarding these coefficients. Numerical tables are given by A. Simon, Oak Ridge National Laboratory Rept. 1718 (Special), and M. Rotenberg *et al.*, *The 3-j and 6-j Symbols* (Technology Press, Cambridge, Massachusetts, 1959).

where $T^{(l)}{}_{mP}$ is a real-standard component of an irreducible tensor $\mathbf{T}^{(l)}$ of degree l contained in the initial tensor with components T_{ik}. Entering in (18) the expressions (16) with the numerical values of the Wigner coefficients, one finds

$$\begin{aligned}
&T^{(2)}{}_{2+}=2^{-\frac{1}{2}}(T_{xx}-T_{yy}), \quad T^{(2)}{}_{2-}=2^{-\frac{1}{2}}(T_{xy}+T_{yx}),\\
&T^{(2)}{}_{1+}=2^{-\frac{1}{2}}(T_{xz}+T_{zx}), \quad T^{(2)}{}_{1-}=2^{-\frac{1}{2}}(T_{yz}+T_{zy}),\\
&\qquad T^{(2)}{}_{0}=6^{-\frac{1}{2}}(2T_{zz}-T_{xx}-T_{yy}),\\
&T^{(1)}{}_{1+}=2^{-\frac{1}{2}}(T_{zy}-T_{yz}), \quad T^{(1)}{}_{1-}=2^{-\frac{1}{2}}(T_{xz}-T_{zx}),\\
&\qquad T^{(1)}{}_{0}=2^{-\frac{1}{2}}(T_{yz}-T_{xy}),\\
&\qquad T^{(0)}{}_{0}=3^{-\frac{1}{2}}(T_{xx}+T_{yy}+T_{zz}).
\end{aligned} \tag{19}$$

The elements of the products of vector components $[\mathbf{a}^{(1)}\times\mathbf{b}^{(2)}]^{(l)}$ are obtained by replacing in (19) T_{xx} by a_xb_x, etc. The numerical coefficients in (19) arise from the fact that the coefficients (16) constitute a unitary transformation. Notice, however, that the set of degree one $[\mathbf{a}^{(1)}\times\mathbf{b}^{(1)}]^{(1)}$ has a sign opposite to that of the components of the vector product $\boldsymbol{a}\times\boldsymbol{b}$.[8] The reduction of the sets of components of nth degree ordinary tensors, with $n>2$, proceeds by repeated application of (18) to successive pairs of tensor indices, in analogy with the construction of irreducible multiple products of n irreducible sets of degree 1 in Chap. 8 of FR.

6. APPLICATIONS

It was emphasized in FR, Chaps. 6 and 19, that functions of the mutual orientation of two systems are often conveniently expanded into invariant products of pairs of tensorial sets pertaining to the two systems. Each set is conveniently defined in a frame of reference attached to the system to which it pertains, and each product contains accordingly an r-transformation matrix that relates the orientations of the two systems. Thus the potential energy of a system of electric charges in a potential field $V(r,\theta,\varphi)$ is expressed by Eq. (6.15) of FR in the form

$$U=\sum_l \mathfrak{M}^{[l]}\ \mathfrak{D}^{(l)}(\psi,\theta,\varphi)\ \mathfrak{A}^{(l)}, \tag{20}$$

where $\mathfrak{M}^{[l]}$ is a contrastandard set of components of the 2^l-pole moment of the charges, $\mathfrak{A}^{(l)}$ a set of coefficients of the expansion of V into contrastandard spherical harmonics $[4\pi/(2l+1)]^{\frac{1}{2}}\mathfrak{Y}^{[l]}(\theta,\varphi)$, and $\mathfrak{D}^{(l)}$ pertains to the coordinate rotation from axes (xyz) pertaining to the field to axes (x',y',z') pertaining to the system of charges. In a real-standard representation, V would be expanded into harmonics $[4\pi/(2l+1)]^{\frac{1}{2}}\mathbf{Y}^{(l)}(\theta,\varphi)$ with coefficients $\mathbf{A}^{(l)}$. If the two coordinate systems have one common coordinate axis $(y=y')$, (20) is then replaced by the explicit formula

$$\begin{aligned}
U=\sum_{lm'm}[&M^{(l)}{}_{m'+}\,d^{(l)}{}_{m'm+}(\theta)\,A^{(l)}{}_{m+}\\
&+M^{(l)}{}_{m'-}d^{(l)}{}_{m'm-}(\theta)\,A^{(l)}{}_{m-}],
\end{aligned} \tag{21}$$

in which only (B) has been utilized. Relaxing the requirement $y=y'$ often enables one to choose y and y' so that the planes xz and $x'z'$ be symmetry planes of the field and charges respectively, i.e., so that $A^{(l)}{}_{m-}$ and $M^{(l)}{}_{m'-}$ vanish. In this event, combined application of (A) and (B) yields

$$\begin{aligned}
U=\sum_{lm'm} M^{(l)}{}_{m'+}[&\cos m'\psi\ d^{(l)}{}_{m'm+}(\theta)\cos m\varphi\\
&-\sin m'\psi\ d^{(l)}{}_{m'm-}(\theta)\sin m\varphi]\,A^{(l)}{}_{m+},
\end{aligned} \tag{22}$$

which brings out explicitly the dependence of the interaction energy on the mutual orientation of the symmetry planes of field and charges.

The mean values of quantum mechanical operators indicated in Eq. (18.27) of FR and throughout the following Chap. 19 have the same geometrical structure as the interaction energy U considered in the foregoing and can, therefore, be expressed in real-standard forms analogous to (21) or (22). Notice, however, that the standard sets of matrix elements of an operator F, defined by (18.1) of FR, do not become automatically real when transformed to a real-standard representation by the matrix $Q^{(l)}$. Only sets that are self-conjugate (FR, Chap. 4) in an arbitrary representation become real in the real-standard representation. As shown on page 102 of FR, standard sets of matrix elements $[(\alpha j|F|\alpha' j')]^{(k)}$ of a Hermitian operator F that belong to a diagonal submatrix, with $(\alpha' j')=(\alpha j)$, are self-conjugate or anti-self-conjugate depending on whether k is even or odd. The transformation $Q^{(l)}$, with $l=k$, will accordingly make these sets real or pure imaginary, respectively. All these sets can be made real by adding to their definition a renormalization factor i^k. On the other hand, sets $[(\alpha j|F|\alpha' j')]^{(k)}$ with $(\alpha' j')\neq(\alpha j)$ are conjugate or anticonjugate to the sets $[(\alpha' j'|F|\alpha j)]^{(k)}$ that belong on the opposite side of the diagonal, depending on the parity of $j-j'-k$. Therefore, the real-standard sets $Q^{(k)}[(\alpha j|F|\alpha' j')]^{(k)}$ and $Q^{(k)}[(\alpha' j'|F|\alpha j)]^{(k)}$ are, in essence, complex conjugate. Their real and imaginary parts may be separated out as new sets of matrix elements that are actually real and are symmetric or antisymmetric with respect to interchange of (αj) and $(\alpha' j')$. Phase normalization conventions are involved in this separation. No suggestion is made on these conventions while their suitability has not been tested.[9]

As a more specific application, we shall consider here

[8] This fact, pointed out on page 39 of FR, stems from the convention on the sign normalization of standard sets embodied in (5.17) and (5.13) of FR; it would not have occurred if i^l had been replaced by i^{-l} in (5.17). It appears that the convention (5.17) of FR, the convention (G.5) of FR regarding the sign of Wigner coefficients, and the convention about right-handed coordinate systems embodied in the definition of vector product form an inconsistent triad. I am indebted to G. Racah for clarification of this point.

[9] The various alternative irreducible sets of matrix elements may be regarded as coefficients of the expansion of an operator into alternative orthogonal systems of operators, from the point of view of U. Fano [Revs. Modern Phys. **29**, 74 (1957), Sec. 6]. See Eq. (18.4) of FR, where the orthogonal operators are $[|\alpha j\}\times\tilde{\mathfrak{U}}^{(j')}\{\alpha' j'|]^{(k)}$.

the scattering of light or γ rays by an atomic system through an electric dipole interaction process. The general formula for this type of process, namely, (19.11) of FR, represents the response of an analyzer of the scattered radiation as a linear combination of invariant products of the type

$$[(E1|\rho|E1)]^{(k)*}\mathfrak{D}^{(k)-1}(\psi,\theta,\varphi)[(E1|F|E1)]^{(k)}, \quad (23)$$

where (ψ,θ,φ) are Euler angles relating the orientations of the analyzer and of a polarizer of the incident beam. In (23), F indicates the operator that represents the analyzer response, ρ is the density matrix of the incident radiation as prepared by a polarizer, and all the sets of quantum numbers (αj) have been replaced by $E1$ to specify electric dipole radiation. Since $j=1$, the degree k of the products (23), can be 0, 1, and 2. Geometry alone determines the relative values of (23) for the three values of k. The nature of the scatterer determines the coefficients with which these products are combined in the final results; the coefficients are equal when the scatterer is isotropic.

To work out the real-standard form of (23), we replace the standard sets of matrix elements by real-standard ones

$$\begin{aligned}\mathbf{C}^{(k)}&=Q^{(k)}[(E1|\rho|E1)]^{(k)},\\ \mathbf{F}^{(k)}&=Q^{(k)}[(E1|F|E1)]^{(k)}.\end{aligned} \quad (24)$$

These two sets can be expressed, to within irrelevant normalization factors, in terms of the electric field components of the radiation that emerges from the polarizer or is accepted by the analyzer. We indicate these components by (a_x,a_y) and $(b_{x'},b_{y'})$, respectively, for the incident and detected radiation; the primes refer to coordinate axes attached to the analyzer. The direct products of these components and of their complex conjugates are the components $C_{ik}=a_ia_k^*$, $F_{ik}=b_ib_k^*$ of intensity tensors. The elements of the real-standard sets (24) are obtained from the C_{ik} and F_{ik}, respectively, by the reduction formulas (19). Notice that (a) the complex phase of the field components a_i and b_i has the essential function of representing the phase relationships in the polarization; (b) the intensity parameters C_{ik}, F_{ik} represent observables in any event, whereas the field components a_i, b_i are not observable under conditions of partial polarization; (c) the intensity parameters can be expressed in terms of the Stokes parameters

$$\begin{aligned}I_0&=a_xa_x^*+a_ya_y^*=C_{xx}+C_{yy},\\ I_1&=a_xa_x^*-a_ya_y^*=C_{xx}-C_{yy},\\ I_2&=a_xa_y^*+a_ya_x^*=C_{xy}+C_{yx},\\ I_3&=i(a_xa_y^*-a_ya_x^*)=i(C_{xy}-C_{yx}),\end{aligned}$$

which describe the intensity and polarization of the incident radiation, and in terms of the corresponding efficiency parameters

$$E_0=b_{x'}b_{x'}^*+b_{y'}b_{y'}^*=F_{x'x'}+F_{y'y'},\ E_1,\ E_2,\ E_3,$$

which characterize the analyzer response. Equation (19) yields

$$\begin{aligned}\mathbf{C}^{(2)}{}_{2+}&=\sqrt{\tfrac{1}{2}}(a_xa_x^*-a_ya_y^*)=\sqrt{\tfrac{1}{2}}I_1,\\ \mathbf{C}^{(2)}{}_{2-}&=\sqrt{\tfrac{1}{2}}(a_xa_y^*+a_ya_x^*)=\sqrt{\tfrac{1}{2}}I_2,\\ \mathbf{C}^{(2)}{}_{1+}&=\mathbf{C}^{(2)}{}_{1-}=\mathbf{C}^{(1)}{}_{1+}=\mathbf{C}^{(1)}{}_{1-}=0,\\ \mathbf{C}^{(2)}{}_{0}&=\sqrt{\tfrac{1}{6}}(-a_xa_x^*-a_ya_y^*)=-\sqrt{\tfrac{1}{6}}I_0,\\ \mathbf{C}^{(1)}{}_{0}&=\sqrt{\tfrac{1}{2}}(a_ya_x^*-a_xa_y^*)=i\sqrt{\tfrac{1}{2}}I_3,\\ \mathbf{C}^{(0)}{}_{0}&=\sqrt{\tfrac{1}{3}}(a_xa_x^*+a_ya_y^*)=\sqrt{\tfrac{1}{3}}I_0.\end{aligned} \quad (25)$$

The elements of $\mathbf{F}^{(k)}$ are given by analogous formulas.

The products (23) for $k=0, 1, 2$ will now be expressed in terms of the Stokes parameters, which represent observables, and in the same real-standard form as (21), i.e., assuming that the axes y and y' coincide. On utilizing the matrices given in Table I, we find[10]

$$\mathbf{C}^{(0)*}\mathbf{D}^{(0)-1}\mathbf{F}^{(0)}=\tfrac{1}{3}I_0E_0, \quad (26a)$$

$$\mathbf{C}^{(1)*}\mathbf{D}^{(1)-1}\mathbf{F}^{(1)}=\tfrac{1}{2}I_3E_3\cos\theta, \quad (26b)$$

$$\begin{aligned}\mathbf{C}^{(2)*}\mathbf{D}^{(2)-1}\mathbf{F}^{(2)}=[&I_0E_0\tfrac{1}{12}(3\cos^2\theta-1)-(I_0E_1+I_1E_0)\tfrac{1}{4}\sin^2\theta\\ &+I_1E_1\tfrac{1}{4}(1+\cos^2\theta)]+\tfrac{1}{2}I_2E_2\cos\theta.\end{aligned} \quad (26c)$$

The sum of these three products yields the complete intensity and polarization formula for scattering by an isotropic center. In the absence of polarization effects (i.e., when only I_0 and E_0 differ from zero), this formula reduces to the familiar angular distribution law $\frac{1}{4}I_0E_0(1+\cos^2\theta)$. However, different results are obtained when (26a–c) are added with different weight factors, as they generally must be for anisotropic scatterers such as spinning atoms.

The performance of polarizers and analyzers is characterized, with regard to linear polarization, by the maximum values P_l and Q_l that are achieved by I_1/I_0 and E_1/E_0 when the x and x' axes lie, respectively, in the direction of linear polarization characteristic for the polarizer and analyzer; I_2 and E_2 vanish when the axes are so oriented. Calling ψ the angle from this x axis to the plane of scattering and φ the angle from the plane of scattering to the x' axis, we must then substitute in (26c)[10]

$$\begin{aligned}I_1&=I_0P_l\cos2\psi, \quad I_2=I_0P_l\sin2\psi,\\ E_1&=E_0Q_l\cos2\varphi, \quad E_2=-E_0Q_l\sin2\varphi.\end{aligned} \quad (27)$$

This substitution brings (26c) to the form equivalent to (22).

[10] The -1 exponent on the $\mathbf{D}$ matrix, equivalent to transposition, stems from the fact that the Euler angles pertain to rotation from the polarizer to the analyzer axis, rather than to the opposite rotation.

Algebraic Characterization of the TCP Operation

A. GROSSMANN*
Physics Department, Brandeis University, Waltham, Massachusetts
(Received May 2, 1960)

If an "operation" $\mathfrak{M}$ is real and preserves Lorentz covariant relations, then $\mathfrak{M}$ is either a proper Lorentz operation or the product of a proper Lorentz operation and the TCP operation.

I. INTRODUCTION

THE TCP theorem[1-5] has been stated and proved within the framework of quantized field theories. It is shown in this paper that the TCP operation emerges, quite unambiguously, from considerations which involve only the proper orthochronous Lorentz group L. The idea is, roughly, to define an "operation" as something which acts simultaneously on all Lorentz covariant quantities, and to ask for the operations which do not destroy Lorentz covariant relations. It turns out that, apart from Lorentz transformations, only the TCP operation has this property. (Theorem 1, Sec. III.)

These results can be viewed as an application of the duality theory of non-Abelian groups.[6-8] The reading of this paper, however, does not require any knowledge of duality theory.

II. OPERATIONS

Let A be the group of unimodulary 2×2 matrices (the universal covering group of L). Denote by $u(m,n)$ quantities that transform according to the $(m+1)\times(n+1)$-dimensional irreducible representation of A (see Appendix). A quantity $u(m,n)$ is a spinor[9] symmetric in m undotted and n dotted indices.

The TCP operation as defined by Pauli[2] transforms $u(m,n)$ into

$$(-1)^n u(m,n) \quad \text{if} \quad m+n \text{ is even}, \tag{1a}$$

$$i(-1)^n u(m,n) \quad \text{if} \quad m+n \text{ is odd}. \tag{1b}$$

Let us define, in general, an *operation* $\mathfrak{M}$ as an arbitrarily given law by which, for $m, n=0, 1, 2, \cdots$, the $(m+1)(n+1)$ independent components of $u(m,n)$ are linearly transformed among themselves. This can be written as

$$u(m,n) \to M^{(m,n)}u(m,n)$$
$$(m, n=0,1,2,\cdots),$$

where $M^{(m,n)}$ is a linear operator in the $(m+1)(n+1)$-dimensional space of components of $u(m,n)$.

The *product* $\mathfrak{M}_1\mathfrak{M}_2$ of the operations $\mathfrak{M}_1$ and $\mathfrak{M}_2$ is defined by

$$u(m,n) \to M_1{}^{(m,n)}M_2{}^{(m,n)}u(m,n)$$
$$(m, n=0,1,2,\cdots).$$

Let now $u(D)$ be a quantity which transforms according to an arbitrary (not necessarily irreducible) representation D of A. In order to define the transformation $M(D)$ which $u(D)$ undergoes under $\mathfrak{M}$, we require:

(A) If $a \to D'(a)(a\epsilon A)$ and $a \to \gamma D'(a)\gamma^{-1}$ are any two equivalent representations of A, then

$$M(\gamma D'\gamma^{-1}) = \gamma M(D')\gamma^{-1}.$$

(B) If $D=D'+D''$ is the direct sum of the representations D' and D'', then

$$M(D) = M(D') + M(D''),$$

where $+$ denotes a direct sum of operators.

Every unimodulary 2×2 matrix a defines[10] an operation

$$u(m,n) \to D^{(m,n)}(a)u(m,n)$$
$$(m, n=0,1,2,\cdots).$$

A proper orthochronous Lorentz transformation $l=\Lambda(a,a)$ defines[10] not one operation but an infinite number of them. They are all of the form

$$u(m,n) \to D^{(m,n)}(a)u(m,n) \qquad (m+n \text{ even}), \tag{2a}$$

$$u(m,n) \to \pm D^{(m,n)}(a)u(m,n) \qquad (m+n \text{ odd}), \tag{2b}$$
$$(m, n=0,1,2,\cdots),$$

and are distinguished from each other by the choice of signs in (2b). This choice is arbitrary, since, under the Lorentz transformation $l=\Lambda(a,a)=\Lambda(-a,-a)$, the quantity $u(m,n)$ can equally well be thought to transform into $D^{(m,n)}(a)u(m,n)$ as into $D^{(m,n)}(-a)u(m,n) = (-1)^{m+n}D^{(m,n)}(a)u(m,n)$.

* On leave of absence from the Institute R. Boskovic, Zagreb, Yugoslavia.

[1] G. Lüders, Kgl. Danske Videnskab. Selskab. Mat.-fys. Medd. 28, 5 (1954); Ann. Phys. 2, 1 (1957).

[2] W. Pauli, *Niels Bohr and the Development of Physics* (Pergamon Press, New York, 1955), p. 30.

[3] J. Schwinger, Phys. Rev. 82, 914 (1951); Proc. Natl. Acad. Sci. U. S. 44, 223, 617 (1958).

[4] R. Jost, Helv. Physica Acta 30, 409 (1957).

[5] F. J. Dyson, Phys. Rev. 110, 579 (1958).

[6] C. Chevalley, *Theory of Lie Groups* (Princeton University Press, Princeton, New Jersey, 1946), Chap. VI (for compact groups).

[7] Harish-Chandra, Ann. Math. 51, 299 (1950).

[8] G. Hochschild and G. D. Mostow, Ann. Math 60, 495 (1957).

[9] See e.g., B. L. van der Waerden, *Die Gruppentheoretische Methode in der Quantenmechanik* (Springer-Verlag, Berlin, 1932), Sec. 20.

[10] For the definition of $\Lambda(a,b)$, $D^{(m,n)}(a,b)$, and $D^{(m,n)}(a)$, see the Appendix.

Similarly, a proper complex Lorentz transformation[10] $\Lambda=\Lambda(a,b)$ defines all the operations

$$u(m,n)\rightarrow D^{(m,n)}(a,b)u(m,n) \qquad (m+n \text{ even}),$$
$$u(m,n)\rightarrow \pm D^{(m,n)}(a,b)u(m,n) \qquad (m+n \text{ odd}),$$

with arbitrary choice of signs in the second line.

If $\mathfrak{L}$ is one of the operations defined by a Lorentz transformation, we shall say that $\mathfrak{L}$ is a *Lorentz operation* (proper complex, etc., as the case may be).

It will often be unnecessary to distinguish between the operations that are defined by one and the same Lorentz transformation. This leads to the following definition:

Let $\mathfrak{M}_1$ and $\mathfrak{M}_2$ be two operations. We shall write

$$\mathfrak{M}_1\equiv\mathfrak{M}_2$$

if

$$M_1{}^{(m,n)}=M_2{}^{(m,n)} \qquad (m+n \text{ even}),$$
$$M_1{}^{(m,n)}=\pm M_2{}^{(m,n)} \qquad (m+n \text{ odd}).$$

The *complex conjugate*, $\mathfrak{M}^*$, of an operation $\mathfrak{M}$ is defined by

$$u(m,n)\rightarrow (M^{(m,n)})^*u(m,n)$$
$$(m,n=0,1,2,\cdots),$$

where $(M^{(m,n)})^*$ is the complex conjugate of the matrix[11] $M^{(m,n)}$.

If $\mathfrak{M}\equiv\mathfrak{M}^*$, we shall say that the operation $\mathfrak{M}$ is *real*.

The product of two real operations is real. The operation $\mathcal{T}$, defined by (1), is real.

Let a and b be two unimodulary 2×2 matrices. Consider the operation $\mathfrak{L}(a,b)$ defined as follows: All undotted indices of quantities $u(m,n)$ are transformed by a, and all dotted indices by b^* (the complex conjugate of b). From the foregoing definitions and the results of the Appendix, it follows that $\mathfrak{L}(a,b)$ is a proper complex Lorentz operation.

In particular, $\mathfrak{L}(a,a)$ is a proper orthochronous and $\mathfrak{L}(a,-a)$ a proper antiorthochronous real Lorentz operation.

III. PRESERVATION OF LORENTZ COVARIANT RELATIONS

In this section we shall study the operations which preserve relations covariant under L. It will be important to distinguish covariance under L from covariance under A.

Covariant relations between spinors are obtained[9] by multiplication and contraction; furthermore, certain reality conditions are preserved by L.

Multiplication: A relation of the form

$$u(m_1,n_1)v(m_2,n_2)=w(m_1+m_2,\ n_1+n_2) \qquad (3)$$

is preserved by A. Under L, however, (3) is transformed into

$$u'(m_1,n_1)v'(m_2,n_2)=w'(m_1+m_2,\ n_1+n_2)$$
if both m_1+n_1 and m_2+n_2 are even, and (4a)

$$u'(m_1,n_1)v'(m_2,n_2)=\pm w'(m_1+n_1,\ m_2+n_2),$$
if at least one of the numbers m_1+n_1, m_2+n_2 is odd. (4b)

This is so because we are free to choose either determination of a when transforming each of the spinors u, v, w.

Contraction: Suppose that $v(m-2p,\ n-2q)$ has been obtained from $u(m,n)$ by contraction of p pairs of undotted indices and q pairs of dotted indices. Then $v'(m-2p,\ n-2q)$ (the transform of v by a proper orthochronous Lorentz transformation) can be obtained from $u'(m,n)$ by the same contractions. There is no ambiguity of sign here, because only pairs of indices are involved. Both indices of a pair belong to the same quantity and, consequently, transform according to the same determination of a.

Reality conditions: Consider now a relation of the form $u(m,n)=v^*(n,m)$ where the star denotes complex conjugation. Under $a\epsilon A$, the quantity $u(m,n)$ is transformed into $u'(m,n)=D^{(m,n)}(a)u(m,n)$, and $v(n,m)$ into $v'(n,m)=D^{(n,m)}(a)v(n,m)=D^{(m,n)}(a)^*v(n,m)$. So $v^*(n,m)$ is transformed into

$$v^*(n,m)'=D^{(m,n)}(a)v^*(n,m)=D^{(m,n)}(a)u(m,n)=u'(m,n).$$

It follows that under a proper orthochronous Lorentz transformation,

$$[v^*(n,m)]'=u'(m,n) \qquad (m+n \text{ even})$$
$$[v^*(n,m)]'=\pm u'(m,n) \qquad (m+n \text{ odd}).$$

These considerations show how proper orthochronous Lorentz transformations affect covariant spinor relations. The problem now is to find all operations $\mathfrak{M}$ which are exactly "as good" as proper orthochronous Lorentz transformation in preserving covariant relations.

Let $u'(m,n)=M^{(m,n)}u(m,n)$, $v'(m,n)=M^{(m,n)}v(m,n)$, etc., be quantities transformed by $\mathfrak{M}$.

Problem: Study the set of all operations $\mathfrak{M}$ which satisfy the following conditions:

(C_1) If $u(m_1,n_1)v(m_2,n_2)=w(m_1+m_2,\ n_1+n_2)$, then $u'(m_1,n_1)v'(m_2,n_2)=\pm w'(m_1+n_1,\ m_2+n_2)$.

(C_2) Furthermore, if both m_1+n_1 and m_2+n_2 are even, then $u'(m_1,n_1)v'(m_2,n_2)=w'(m_1+m_2,\ n_1+n_2)$.

(D) If $v(m-2p,\ n-2q)$ is obtained from $u(m,n)$ by contraction of p pairs of undotted and q pairs of dotted indices, then $v'(m-2p,\ n-2q)$ is obtained from $u'(m,n)$ by contraction of the same pairs of indices.

[11] That is, the matrix corresponding to the operator $M^{(m,n)}$ in the basis in which the action of A is represented by the matrices $D^{(m,n)}(a)$.

(E) If $u(m,n)=v^*(n,m)$, then

$$u'(m,n)=[v^*(n,m)]' \quad \text{for} \quad m+n \text{ even}$$
$$u'(m,n)=\pm[v^*(n,m)]' \quad \text{for} \quad m+n \text{ odd}.$$

The rest of this section is devoted to the solution of the above problem.

According to Sec. II, the condition (E) can be expressed as $\mathfrak{M}\equiv\mathfrak{M}^*$ or the requirement that $\mathfrak{M}$ be real.

If two operations satisfy (C) and (D), then their product also satisfies (C) and (D).

If an operation $\mathfrak{M}$ satisfies (C), (D), and (E), so does its inverse $\mathfrak{M}^{-1}$. Every proper orthochronous Lorentz operation satisfies (C), (D), and (E), and so does the operation (1). So our operations form a group G which contains all proper orthochronous Lorentz operations and the operation (1). We shall prove that G is precisely the group generated by all proper real (not necessarily orthochronous) Lorentz operations, and the TCP operation (1).

In the lemmas that follow we shall always assume that the operation $\mathfrak{M}$ satisfies (C) and (D). This assumption will not be repeated. However, whenever the reality assumption (E) is made, it will be stated explicitly.

Lemma 1: The number $M^{(0,0)}$ is equal to one.

Indeed, for any m, n such that $m+n$ is even, we have, by (C_2), $M^{(0,0)}M^{(m,n)}=M^{(m,n)}$.

Lemma 2: The effect of $\mathfrak{M}$ on a quantity $u(m,n)$ can be obtained as follows: The undotted indices of $u(m,n)$ are transformed by $M^{(0,1)}$ and the dotted by $M^{(1,0)}$. In addition, the sign of $u(m,n)$ may get changed.

Proof: By (C), the lemma is certainly true if $u(m,n)$ is a product of quantities $u(0,1)$ and $u(1,0)$. It remains true for an arbitrary combination of such products; that is, for every quantity $u(m,n)$.

Lemma 3: The determinant of the 2×2 matrix $M^{(0,1)}$ is either $+1$ or -1.

Proof: For the sake of shortness, write M instead of $M^{(0,1)}$. Let $u_{k\dot{s}}$ be a spinor with two undotted indices. By Lemma 2, $u_{k\dot{s}}$ is transformed, under $\mathfrak{M}$, into

$$u_{k\dot{s}}'=\pm M_k{}^p M_{\dot{s}}{}^{\dot{q}} u_{p\dot{q}}.$$

Because of (D), the contraction $(u')_k{}^k$ has to be equal to the contraction $u_p{}^p$. This gives

$$(u')_k{}^k=\pm M_k{}^p M^{k\dot{q}} u_{p\dot{q}}=u_p{}^p=\epsilon^{p\dot{q}}u_{p\dot{q}},$$

where

$$\epsilon=\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}$$

is the matrix that raises spinor indices.[9] Since $u_{p\dot{q}}$ is arbitrary, we have

$$\pm M_k{}^p M^{k\dot{q}}=\epsilon^{p\dot{q}}.$$

A simple calculation shows that this is equivalent to $\det M=\pm1$.

Lemma 4: $\det M^{(1,0)}=\pm1$.

The proof is the same as that of Lemma 3.

Lemma 5: If $\det M^{(0,1)}=\det M^{(1,0)}=1$, then $\mathfrak{M}$ is a proper complex Lorentz operation. More precisely: either

$$\mathfrak{M}\equiv\mathfrak{L}(M^{(0,1)},(M^{(1,0)})^*), \tag{5a}$$

or

$$\mathfrak{M}\equiv\mathfrak{L}(M^{(0,1)},-(M^{(1,0)})^*), \tag{5b}$$

where $\mathfrak{L}(M^{(0,1)},M^{(1,0)})$ is the proper complex Lorentz operation defined at the end of Sec. II.

Proof: Consider the effect of $\mathfrak{M}$ on quantities $u(m,n)$. For $m+n$ odd, our assertion is true by Lemma 2. For $m+n$ even, it is necessary to verify that the action of $\mathfrak{M}$ coincides, without discrepancies in sign, with the action of one of the operations (5). Notice that every $u(m,n)(m+n$ even) can be obtained, by contractions, from products of quantities $u(1,1)$. By (C_2) and (D), then, the matrices $M^{(m,n)}(m+n$ even) are fully determined—without ambiguity in sign—by $M^{(1,1)}$. Suppose first that $M^{(1,1)}u(1,1)$ is obtained from $u(1,1)$ by transformation of spinor indices, without the additional change of sign. Then $\mathfrak{M}\equiv\mathfrak{L}(M^{(0,1)},(M^{(1,0)})^*)$. If the change of sign does occur, then

$$\mathfrak{M}\equiv\mathfrak{L}(M^{(0,1)},-(M^{(1,0)})^*)\equiv\mathfrak{L}(-M^{(0,1)},(M^{(1,0)})^*).$$

Lemma 6: If $\det M^{(0,1)}=\det M^{(1,0)}=-1$, then $\mathfrak{M}\equiv\mathcal{T}\mathfrak{L}$ where $\mathcal{T}$ is the TCP operation (1) and $\mathfrak{L}$ is a proper complex Lorentz operation.

Proof: The operation $\mathcal{T}^{-1}\mathfrak{M}$ satisfies the assumption of Lemma 5, and is, consequently, a proper complex Lorentz operation.

Lemma 7: If $\mathfrak{M}$ is real and $\det M^{(0,1)}=1$, then $\mathfrak{M}$ is a real proper Lorentz operation (not necessarily orthochronous).

Proof: Since $\mathfrak{M}$ is real, we have $M^{(1,0)}=\pm(M^{(0,1)})^*$, and $\det M^{(1,0)}=\det M^{(0,1)}=1$. So, by Lemma 5, either $\mathfrak{M}\equiv\mathfrak{L}(M^{(0,1)},M^{(0,1)})$ (proper orthochronous) or $\mathfrak{M}\equiv\mathfrak{L}(M^{(0,1)},-M^{(0,1)})$ (proper antiorthochronous).

Lemma 8: If $\mathfrak{M}$ is real and $\det M^{(0,1)}=-1$, then $\mathfrak{M}\equiv\mathcal{T}\mathfrak{L}$, where $\mathcal{T}$ is the TCP operation (1) and $\mathfrak{L}$ is a real proper Lorentz operation (not necessarily orthochronous).

Proof: Lemma 7 is applicable to the operation $\mathcal{T}^{-1}\mathfrak{M}$.

Combining the Lemmas 3, 7, and 8, we obtain the following theorem:

Theorem 1: If an operation $\mathfrak{M}$ is real and preserves Lorentz covariant relations [that is, satisfies (C), (D), and (E)], then $\mathfrak{M}$ is either a real proper Lorentz operation or the product of a real proper Lorentz operation and of the TCP operation (1).

ACKNOWLEDGMENTS

The author acknowledges with gratitude financial support from Harvard University in the years 1955–56 and 1956–57, and many fruitful discussions on the sub-

ject of this paper with both mathematicians and physicists—in particular with Dr. R. Marr.

APPENDIX

This Appendix deals with real and complex Lorentz transformations, their relation to unimodulary, unitary and orthogonal transformations, and their finite-dimensional representations. The results on orthogonal groups are not needed in this paper but are given for future reference.

The proper complex Lorentz transformations Λ are defined by

$$\Lambda_T g \Lambda = g,$$

and

$$\det\Lambda = 1.$$

The metric tensor g is

$$g = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & +1 \end{pmatrix}.$$

The matrix s, defined as

$$s = \begin{pmatrix} i & 0 & 0 & 0 \\ 0 & i & 0 & 0 \\ 0 & 0 & i & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

satisfies $s^2 = g$.

Denote by σ the unitary matrix

$$\sigma = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & i & -i & 0 \\ 1 & 0 & 0 & -1 \\ 1 & 0 & 0 & 1 \end{pmatrix}.$$

Given any pair a, b, of 2×2 matrices, we may consider the 4×4 matrix $\Lambda(a,b)$ defined by

$$\Lambda(a,b) = \sigma(a\times b^*)\sigma^{-1},$$

and the matrix $R(a,b)$ defined as

$$R(a,b) = s\Lambda(a,b)s^{-1}.$$

Here $\times$ denotes the Kronecker product, and b^* is the complex conjugate of b. From the simplest properties of the Kronecker product, it follows that

(a) $\Lambda(a',b')\Lambda(a'',b'') = \Lambda(a'a'',b'b'')$

(b) $R(a',b')R(a'',b'') = R(a'a'',b'b'')$.

Also,

(c) The complex conjugate of $\Lambda(a,b)$ is $\Lambda(b,a)$.

This can be verified by noticing that $(a\times b^*)^* = (a^*\times b) = P(b\times a^*)P^{-1}$, where

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

and that $\sigma P = \sigma^*$.

(d) The complex conjugate of $R(a,b)$ is $gR(b,a)g$.

(e) If we establish a one-to-one correspondence between the four-vector $z = (z_1,z_2,z_3,z_0)$ and the matrix

$$\hat{z} = \begin{pmatrix} z_0+z_3 & z_1-iz_2 \\ z_1+iz_2 & z_0-z_3 \end{pmatrix},$$

then the transformation $z \longrightarrow \Lambda(a,b)z$ corresponds to the transformation $\hat{z} \longrightarrow a\hat{z}b^+$.

This is verified by straightforward computation.

The foregoing statements are valid for arbitrary a and b. We shall now impose restrictions on these matrices.

(f) If $\det a = \det b = 1$, then $\Lambda(a,b)$ is a proper complex Lorentz transformation and $R(a,b)$ is a proper complex orthogonal transformation.

It is sufficient to prove that $\Lambda(a,b)$ is a proper complex Lorentz transformation; the assertion on $R(a,b)$ follows then from the definition of R. By the theorem on determinants of Kronecker products, $\det\Lambda(a,b) = 1$. To prove that $\Lambda(a,b)$ is a Lorentz transformation, we notice that, for 2×2 matrices, the condition $\det a = 1$ is equivalent to

$$a_T = \epsilon a^{-1}\epsilon^{-1},$$

where

$$\epsilon = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

and a_T is the transposed of a. Furthermore, the equation $\sigma_T g \sigma = \epsilon\times\epsilon$ holds. Then

$$\begin{aligned}[\Lambda(a,b)]_T g \Lambda(a,b) &= \sigma_T^{-1}(a\times b^*)_T \sigma_T g \sigma (a\times b^*)\sigma^{-1} \\ &= \sigma_T^{-1}(a_T\times b_T^*)(\epsilon\times\epsilon)(a\times b^*)\sigma^{-1} \\ &= \sigma_T^{-1}(\epsilon\times\epsilon)[a^{-1}\times(b^*)^{-1}](\epsilon\times\epsilon)^{-1}(\epsilon\times\epsilon)(a\times b^*)\sigma^{-1} \\ &= \sigma_T^{-1}(\epsilon\times\epsilon)\sigma^{-1} = g,\end{aligned}$$

which shows that Λ is a Lorentz transformation.

(g) It can be shown by a continuity argument that every proper complex Lorentz transformation can be written in the form $\Lambda(a,b)$ with suitably chosen a and b. Similarly, every proper complex orthogonal transformation can be written as $R(a,b)$ for some a, b.

(h) If $\det a = 1$, then $\Lambda(a,a)$ is a proper orthochronous real Lorentz transformation.

From (c) and (f) we know that $\Lambda(a,a)$ is a proper real Lorentz transformation. Since every $\Lambda(a,a)$ can be continuously connected to unity, it must be orthochronous.

Consequently, if $\det a = 1$, then $R(a,a)$ varies over a subgroup of the group of proper complex orthogonal

transformations, isomorphic to the group of proper orthochronous real Lorentz transformations. The matrices $R(a,a)$ are, in general, not real.

(h′) If $\det a=1$, then $\Lambda(a,-a)=-\Lambda(a,a)$ is a proper real Lorentz transformation changing the sense of time.

(i) If $\det a=\det b=1$, and both a and b are unitary ($aa^{+}=bb^{+}=1$), then $R(a,b)$ is a real proper orthogonal transformation in four dimensions.

We know that $R(a,b)$ is a proper orthogonal transformation whenever $\det a=\det b=1$. We have only to show that $R(a,b)$ is real, or, by (d), that $R(a,b)=gR(b,a)g$. Now, since $g\sigma=\sigma P(\epsilon\times\epsilon)$, we have, for all unimodulary a and b,

$$\begin{aligned} gR(b,a)g^{-1}&=sg\sigma(b\times a^{*})\sigma^{-1}g^{-1}s^{-1}\\ &=s\sigma P(\epsilon\times\epsilon)(b\times a^{*})(\epsilon\times\epsilon)^{-1}P^{-1}\sigma^{-1}s^{-1}\\ &=s\sigma P[b_T{}^{-1}\times(a^{+})^{-1}]P^{-1}\sigma^{-1}s^{-1}\\ &=s\sigma[(a^{+})^{-1}\times b_T{}^{-1}]\sigma^{-1}s^{-1}=R[(a^{+})^{-1},(b^{+})^{-1}] \end{aligned}$$

which, for unitary a and b is equal to $R(a,b)$.

Consequently, if $\det a=\det b=1$ and both a and b are unitary, then $\Lambda(a,b)$ varies over a subgroup of the group of proper complex Lorentz transformations, isomorphic to the group of proper real orthogonal transformations in four dimensions. These matrices $\Lambda(a,b)$ are, in general, not real.

(j) If $\det a=1$ and $a^{+}a=1$, then $\Lambda(a,a)=R(a,a)$ varies over the group of real proper rotations in three dimensions.

Let now a and b be again two arbitrary unimodulary matrices. Consider a quantity $u(m,n)$ symmetric in m undotted and in n dotted indices. Transform the undotted indices by a and the dotted indices by b^{*}. This induces a linear transformation in the space of the $(m+1)(n+1)$ independent components of $u(m,n)$. The matrix of this transformation is denoted by $D^{(m,n)}(a,b)$. Finally, $D^{(m,n)}(a)$ is defined by $D^{(m,n)}(a)=D^{(m,n)}(a,a)$. The representation $a\rightarrow D^{(m,n)}(a)$ is irreducible. For every a, the matrix $D^{(n,m)}(a)$ is complex conjugate to the matrix $D^{(m,n)}(a)$.

Singularities and Discontinuities of Feynman Amplitudes*

R. E. CUTKOSKY†
Carnegie Institute of Technology, Pittsburgh, Pennsylvania
(Received March 31, 1960)

The Landau singularities of the amplitude calculated from an arbitrary Feynman graph are considered. It is shown that the discontinuity across a branch cut starting from any Landau singularity is obtained by replacing Feynman propagators by delta functions for those lines which appear in the Landau diagram. The general formula is a simple generalization of the unitarity condition. The discontinuity is then considered as an analytic function of the momenta and masses; it is shown that its singularities are a subclass of the singularities of the original amplitude which corresponds to Landau diagrams with additional lines. The general results are illustrated by application to some single loop graphs. In particular, the general formula gives an immediate calculation of the Mandelstam spectral function for fourth-order scattering. Singularities not of the Landau type are discussed and illustrated by the third-order vertex part.

I. INTRODUCTION

KARPLUS, Sommerfield, and Wichman[1] and Landau[2] have emphasized the importance of examining the analyticity of the amplitudes corresponding to Feynman graphs, and have discussed some simple graphs in detail. Landau has also given a criterion for determining the position of certain singularities of the amplitude for an arbitrary graph. In this paper we shall derive a formula for the discontinuity across a cut starting from any one of Landau's branch points, and shall determine where this discontinuity is singular. The result is a very natural generalization of the well-known expression, given by the unitarity condition, for the discontinuity across a cut starting from any physical threshold. The general result is extremely useful for analyzing spectral representations. For example, it leads immediately to an explicit expression for the Mandelstam spectral function for the fourth-order scattering amplitude.[3]

Before proceeding with the calculation, let us recapitulate Landau's discussion. He considers the amplitude

$$F=\int B\prod(d^4k)A_1^{-1}\cdots A_N^{-1} \tag{1}$$

(where $A_i=M_i^2-q_i^2$ and B is an arbitrary polynomial) corresponding to a graph with N internal lines and n independent loops. In (1) and the following we adhere closely to Landau's notation. The q_i are linear combinations of the k_i and the external momenta p_i. On its principal branch F has no singularities for sufficiently small, real p_i^2; if the M_i^2 are positive, we may take the p_i^2 to be positive without passing a singularity, and begin the investigation with real p_{i4} and imaginary $\mathbf{p}_i$. We denote by z_a the independent invariants formed from the p_i.

Now introduce the Feynman parametrization

$$F=(N-1)!\int\prod(d\alpha)\prod(d^4k)BD^{-N}\delta(1-\tilde{\alpha}), \tag{2}$$

where $D=\sum_{i=1}^N\alpha_iA_i$ and $\tilde{\alpha}=\sum\alpha_i$. Let $\varphi=\max_k(D)$ (the maximization is carried out with real k_{i4} and imaginary $\mathbf{k}_i$). According to Landau, if $\min\varphi>0$, F is nonsingular, where the minimum is taken with respect to nonnegative α's satisfying $\tilde{\alpha}=1$. As the p_i^2 are increased, the first singularity of F occurs when $\min\varphi\to 0$. This, Landau shows, means that for each i

$$\alpha_iA_i=0, \tag{3}$$

and for each closed loop

$$\sum\alpha_iq_i=0, \tag{4}$$

where the sum is extended over all the lines in the loop; moreover, (4) must be satisfied with nonnegative α's. Landau pointed out that a singularity exists when (4) is satisfied with arbitrary α_i, but did not give an explicit proof of this; as this point is important to our subsequent discussion we show that this follows from an analytic continuation in the internal masses, and the continuity theorem for singularity surfaces.[4]

The following remarks are contained implicitly in Landau's paper.

Let D_m be obtained from D by setting the $\alpha_i=0$ for $i>m$, and let $\varphi_m=\max_k(D_m)$. If for some α_i, $\max_k(D_m)$ occurs for $q_i^2=M_i^2(i\leqslant m)$, then for any other nonnegative $\alpha_i(i\leqslant m)$, $\varphi_m\geqslant 0$. Now, we may choose the M_i^2 for $i>m$ so large that φ_m is the minimum of φ for nonnegative α's. For any $\alpha_i>0(i\leqslant m)$ and $p_i^2>0$ we determine q_i which satisfy (4) (this is just the maximization problem) and *define* for $i\leqslant m$ masses M_i by the equation $q_i^2=M_i^2$. Hence masses exist such that any

* Supported in part by the U. S. Atomic Energy Commission and by the Alfred P. Sloan Foundation, Inc.
† Alfred P. Sloan Foundation Fellow.
[1] R. Karplus, C. M. Sommerfield, and E. H. Wichman, Phys. Rev. 111, 1187 (1958); 114, 375 (1959).
[2] L. D. Landau, Nuclear Phys. 13, 181 (1959). *Note added in proof.* Results similar to Landau's were also obtained by J. C. Taylor [Phys. Rev. 117, 261 (1960)], which the author received after submission of this paper.
[3] S. Mandelstam, Phys. Rev. 115, 1742 (1959).
[4] H. Behnke and P. Thullen, *Theorie der Functionen Mehrerer Komplexer Veränderlichen* (Springer-Verlag, Berlin, 1934), p. 49.

"Landau diagram" corresponds to the first singularity. There are two cases to be considered, which can easily be distinguished upon inspection of the Landau diagram. If some of the masses obtained by the procedure described are constant, or satisfy a relation independent of the α's and the z's, then we do not in general have either a solution to (4) or a singularity. Otherwise, as the α's and z's are varied these masses take on all possible values, in which case it follows from the continuity theorem that for any internal masses there is always a singularity when Eqs. (3) and (4) are satisfied, although this singularity might not appear on the principal sheet of the Riemann surface.

In order to discuss the analytic continuations of F, we eliminate the delta function from (2), by replacing the α_i by $\lambda\alpha_i$, multiplying by a suitable entire function of λ (say $e^{-\lambda}$) and integrating over λ. This gives the equation

$$F=(N-1)!\int\prod(d\alpha)\prod(d^4k)BD^{-N}\tilde{\alpha}^{-1}\exp(-\tilde{\alpha}^{-1}). \quad (5)$$

In (5), the α_i vary independently over any suitable contours from 0 to ∞. We may use this equation to interpret Landau's conditions in the complex region. We use an idea introduced by Hadamard,[5] which has already been exploited in a similar problem by Eden.[6] If we first integrate over the $k_{i\mu}$, we obtain an integrand which is singular when φ vanishes, where in the general case φ is an extremum of D. The singularities of F occur when some of the α_i are fixed at the lower limit of integration, while the contours over which the remaining α_i are integrated are trapped between coalescing singularities. In other words, φ must have a double zero with respect to each of the free variables, which leads directly to Landau's conditions (3) and (4). It is also necessary that for z_a in the neighborhood of a singularity of F, the contours actually pass between these nearly coalescent zeros. We know that this occurs when we consider the first singularity; we obtain an illustration of the continuity theorem if we note that when the M_i are varied, if the z_a are simultaneously varied so as to keep the zeros in a nearly coalescent configuration, the contours must remain entrapped.

Since the integrand in (5) is always singular when $\tilde{\alpha}=0$, if D vanishes for $\tilde{\alpha}=0$ the condition of a double zero with respect to the free variables is relaxed. In this case we might have a singularity even if conditions (3) and (4) do not hold, although such a singularity could never appear on the principal sheet. We shall show, in Sec. III, that an "anomalous" singularity of this type actually occurs in the third-order vertex.

[5] J. Hadamard, Acta Math. **22**, 55 (1898).

[6] R. J. Eden, Proc. Roy. Soc. (London) **A210**, 388 (1952). *Note added in proof.* After submission of this paper, the author received two papers containing a similar discussion of the complex singularities: J. C. Polkinghorne and G. R. Screaton, Nuovo cimento **15**, 289 (1960); J. Tarski, J. Math. Phys. **1**, 154 (1960).

II. DISCONTINUITIES OF FEYNMAN AMPLITUDES

A. Calculation of the Discontinuity

We shall prove the following theorem: Let F denote the amplitude defined by Eq. (1), and let F_m denote the discontinuity of F across a branch cut starting from a singularity defined by Landau's conditions (3) and (4) in which $A_i=0$ for $i\leqslant m$; then

$$F_m=(2\pi i)^m \times\int\frac{B\prod(d^4k)\delta_p(q_1^2-M_1^2)\cdots\delta_p(q_m^2-M_m^2)}{A_{m+1}\cdots A_N}. \quad (6)$$

(The notation implies a particular ordering of the lines.) The subscript p on the delta functions means that only the contribution of the "proper" root of $q_i^2=M_i^2$ is to be taken. Equation (6) is a simple generalization of well-known results, and follows directly from the Hadamard-Eden analysis.

Consider the contracted Feynman graph obtained by fusing the vertices connected by the lines $i>m$. Let ν be the number of independent loops in this contracted graph. We can choose the k_j so that the $q_i(i\leqslant m)$ depend only on those k_j for which $j\leqslant\nu$. If the $m\times 4\nu$ matrix

$$J_{i,j\mu}=\partial q_i^2/\partial k_{j\mu}$$

is of rank m, we may choose as integration variables $\xi_i=q_i^2$ for $i\leqslant m$, and $4\nu-m$ additional variables. The q_i^2 are the squared distances between certain points in momentum space, and the ξ_i for $m<i\leqslant 4\nu$ may be interpreted as related angle variables. We shall discuss later the circumstance that $J_{i,j\mu}$ has a rank smaller than m for all $k_{j\mu}$. If the rank is too small only when the $k_{j\mu}$ satisfy particular relations, these exceptional points may in general be avoided by appropriate indentations of the $k_{j\mu}$ contours. We therefore obtain

$$F=\int_{a_1}^{b_1}dq_1^2\cdots\int_{a_m}^{b_m}dq_m^2 \times\int\frac{\prod_{m<i\leq 4\nu}(d\xi_i)\prod_{j>\nu}(d^4k_j)}{JA_1\cdots A_N}, \quad (7)$$

where

$$J=\det(\partial\xi_i/\partial k_{j\mu}).$$

The limits of integration (a_j,b_j) for the q_j^2 integration are the extrema of q_j^2 for fixed $q_i^2(i<j)$. This leads to the equations (for each loop of the contracted graph)

$$\sum_{(i\leq j)}\beta_i q_{i\mu}=0, \quad (8)$$

where the β_i are Lagrange multipliers. From (8) for $j=m$ we see that Landau's conditions (3), (4) imply that when a singularity develops, the point where the $A_i=0$ for $i\leqslant m$ lies on the boundary of the region of integration. Equation (8) also shows that the rank of $J_{i,j\mu}$ is always too small on the boundary of the integration region, but this gives no difficulty. In certain

cases each set of the q_i^2 corresponds to two points in momentum space; in these cases we interpret the q_m^2 integration as being taken over the closed contour which encloses the two points a_m and b_m where J is singular.[7]

For brevity we denote by z a point in the (many-sheeted) space of the invariants. Let z_0 denote any point on the singularity surface in question which does not also lie on some other singularity surface.

We first suppose that all the integrations in (7) have been performed, except that over q_1^2. Then we write

$$F=\int_{a_1}^{b_1} dq_1^2(M_1^2-q_1^2)^{-1}F_{(1)}(q_1^2). \tag{9}$$

Now, by hypothesis, (A) F is singular when $z\to z_0$, and (B) F would not be singular at z_0 if the factor $(M_1^2-q_1^2)^{-1}$ were absent or if the mass M_1 were changed. Therefore, the contour of the q_1^2 integration must pass between the pole $q_1^2=M_1^2$ and a singularity of $F_{(1)}(q_1^2)$ at $q_1^2=Q^2$, where $Q^2\to M_1^2$ when $z\to z_0$. We may replace this contour by one on the other side of the pole $q_1^2-M_1^2$ and a very small circle enclosing this pole, where the contour which avoids the pole gives a contribution to F which is regular in the neighborhood of z_0. The singular part of F is therefore

$$F_s=\pm 2\pi i F_{(1)}(M_1^2). \tag{10}$$

The argument given is not sufficient to determine the sign.

After applying the foregoing argument in succession to the variables $q_2^2\cdots q_{m-1}^2$, we obtain

$$F_s=\int_{a_m}^{b_m} dq_m^2(M_m^2-q_m^2)^{-1}F_{(m)}(q_m^2). \tag{11}$$

In (11), a_m and b_m are the limits calculated with $q_1^2=M_i^2$ for $i<m$. When $z\to z_0$, it follows from (8) that one of these limits coincides with the point $q_m^2=M_m^2$. It is obvious that the discontinuity across a branch cut starting from z_0 is $2\pi i F_{(m)}(M_m^2)$. When the q_m^2 integration is taken over a contour enclosing the points a_m and b_m, the two branches of F_s are determined by whether the pole $q_m^2=M_m^2$ lies inside this contour or not, so we obtain the same result.

We now *define* the sign of F_m by analytic continuation from the case where the masses are such that the singularity in question is the first encountered as the z_a are continued through real values from the singularity-free region, and z is a real point just beyond this singularity. It was shown in the Introduction that it is possible to do this. We define the discontinuity $F_m(z)$ to be the difference between $F(z)$ as calculated by giving the masses small *negative* imaginary parts and that calculated with small *positive* imaginary parts; that is,

$$F_m(z)=F_{-i\epsilon}(z)-F_{+i\epsilon}(z). \tag{12}$$

Now consider the q_m^2 integration: Equation (12) implies that the discontinuity is given by a clockwise contour around the pole. But the same result must hold for all q_i^2. This proves Eq. (6) for the case that the rank of the matrix $\partial q_i^2/\partial k_{j\mu}$ is equal to m, except that in transforming back to the $k_{i\mu}$ we must be careful to keep only the contribution from the proper root of $q_i^2=M_i^2$.

There are two cases in which the rank of $J_{i,j\mu}$ is too small; either this happens only for z which satisfy some particular relation, which restricts these z to lie on some surface, or else it occurs identically, for all z. In the first case, (6) is valid for all nonexceptional z, but the discontinuity might be singular when $J_{i,j\mu}$ is singular. If the rank is always too small, as when $m>4\nu$, we consider the singularity obtained by eliminating a sufficient number of lines (say for $m'<i\leqslant m$) that the rank of the reduced matrix $\partial q_i^2/\partial k_{i\mu}$ is m'. The singularity of the larger matrices implies that the eliminated q_i^2 can be expressed in terms of the q_i^2 for $i\leqslant m'$. Hence when we evaluate the discontinuity $F_{m'}$ by Eq. (6), we find that $F_{m'}$ has not a branch point but a pole when one of the eliminated A_i vanishes. These exceptional cases will be illustrated in Sec. III.

B. Singularities of the Discontinuity Function

We may think of $F_m(z)$ as the difference between the values of $F(z)$ on two different sheets, so the singularity surfaces of $F_m(z)$ will be contained among those of $F(z)$. We discuss these singularities by introducing $N-m$ Feynman parameters $\alpha_i(i>m)$ and repeating Landau's calculation. When we integrate over the $k_{i\mu}$, we obtain a singularity for those values of the α_i for which

$$\varphi=\text{Extremum}_k(\textstyle\sum_{i>m}\alpha_i A_i)$$

vanishes. However, the variables $k_{i\mu}$ are not all independent, because they satisfy the constraints $A_i=0$ for $i\leqslant m$. These constraints are introduced into the extremization by using m Lagrange multipliers, which we also call $\alpha_i(i\leqslant m)$. This leads to the equation $\sum\alpha_i q_{i\mu}=0$, which is identical to (4). The integration over the Feynman parameters is singular when some of them are zero, and φ is a vanishing extremum with respect to the rest. This leads to Eq. (3) for $i>$m. We are not allowed to omit any of the conditions $A_i=0$ for $i\leqslant m$, so the singularities of $F(z)$ which are also singularities of $F_m(z)$ correspond to Landau diagrams in which lines have been added to the Landau diagram which defined the original singularity. The other singularities of F necessarily appear on both sheets and cancel when we calculate the difference. As we have pointed out before, there is also a possibility of non-Landauian singularities.

[7] For some graphs with more than one loop, several of the q_i^2 integrations need to be interpreted in this way.

Let us denote by $F_{m,m'-m}(z)$ the discontinuity of $F_m(z)$ across a branch cut starting from the branch point for which $A_i=0$ for $m<i\leqslant m'$. We calculate $F_{m,m'-m}$ by the same method used to calculate F_m; we use the q_i^2 as variables for $i\leqslant m'$. It is clear that all the steps in the proof (except for determination of the sign) are identical. Moreover, we find that

$$F_{m,m'-m}(z)\equiv F_{m'}(z). \tag{13}$$

[We use Eq. (13) to define the sign of $F_{m,m'-m}$.] It may be noted that it can be proved independently (by extending the argument in the Introduction) that the singularity of F which corresponds to $A_i=0$ for $i\leqslant m'$ only appears on one of two adjacent sheets connected by the branch point corresponding to $A_i=0$ for $i\leqslant m$.

C. Unitarity Condition

Consider two graphs, each with m outgoing lines, and with r and s incoming lines, respectively. Let F and G denote the corresponding amplitudes. The unitarity of the S matrix implies that these two graphs give a contribution to the imaginary part of the T matrix (for r outgoing and s incoming particles) which is, apart from numerical factors and with neglect of the spins of the particles,

$$\mathcal{T}_{rs(m)}=\int d\tau_m F^* G, \tag{14}$$

where $d\tau_m$ is the volume element in the phase space of m particles. Let $\mathbf{q}_i$ and W_i denote the momenta and energies of these m particles. As a consequence of momentum conservation, the $\mathbf{q}_i$ depend linearly on $m-1$ integration variables $\mathbf{k}_i$. With a covariant normalization of states, we have

$$d\tau_m=\frac{d^3k_1\cdots d^3k_{m-1}}{(2W_1)\cdots(2W_m)}\delta(\sum W_i-E), \tag{15}$$

where E is the total energy. We may introduce $m-1$ new integration variables k_{i4} and write (15) as follows:

$$d\tau_m=d^4k_1\cdots d^4k_{m-1}\delta_p(q_1^2-M_1^2)\cdots\delta_p(q_m^2-M_m^2). \tag{16}$$

In (16) the q_{i4} are the same functions of the k_{i4} as the $\mathbf{q}_i$ are of the $\mathbf{k}_i$. The subscript p means that only the "proper" root of $q_i^2=M_i^2$, that for which q_{i4} is positive, is to be considered when the integrations are carried out.

Equation (14) is first obtained for real momenta. To continue it to the complex region we introduce the explicit forms of G and F, with the notation that q_i is the momentum of any internal line, and k_i is any integration variable. Then (14) becomes

$$\mathcal{T}_{rs(m)}=\int\frac{\prod(d^4k)B\delta_p(q_1^2-M_1^2)\cdots\delta_p(q_m^2-M_m^2)}{A_{m+1}\cdots A_N}, \tag{17}$$

where $A_i=M_i^2+i\epsilon-q_i^2$ for lines belonging to the graph F, and $A_i=M_i^2-i\epsilon-q_i^2$ for lines belonging to the graph G.

Equation (17) is just a special case of the general discontinuity formula (6) for the graph obtained by joining the graphs F and G by the m common lines. In (17) the analytic continuations have been defined in a particular way (by the $\pm i\epsilon$ rule), while in (6) the masses may be considered to be arbitrary. The discussion in Sec. II.B of the location of the singularities of $F_m(z)$ applies without modification to $\mathcal{T}_{rs(m)}$.

The correspondence between the unitarity condition (17) and the general discontinuity formula (6) suggests that the general discontinuity may be looked on as a pseudounitarity condition. The particles, instead of being divided into the two groups of "initial" and "final" particles, may be divided into three or more groups.

III. ILLUSTRATIONS

In this section we illustrate the results derived in Sec. II by applying them to the three graphs shown in Fig. 1.

A. Fourth-Order Scattering

The singularities correspond to the vanishing of the following combinations of the A_i: (13), (24), (12), (23), (34), (41), (123), (134), (124), (234), and (1234). The ordinary threshold is the (13) singularity. The corresponding discontinuity is obtained by replacing A_1^{-1} and A_3^{-1} by $2\pi i\delta_p(A_1)$ and $2\pi i\delta_p(A_3)$. The discussion in Sec. II.B shows that this discontinuity has only the singularities (13), (123), (134), and (1234). The Mandelstam spectral function[3] is, apart from a factor of four, the discontinuity of this discontinuity function across the (1234) singularity, which is

$$F_4=\int d^4k\delta_p(q_1^2-M_1^2)\cdots\delta_p(q_4^2-M_4^2). \tag{18}$$

Reverting to the variables used in the proof of (6),

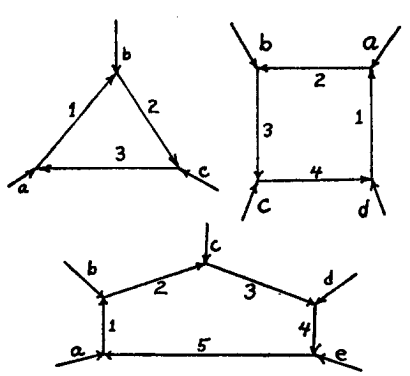


FIG. 1. Feynman graphs considered in Sec. III.

we have

$$F=\int J^{-1}dq_1^2\cdots dq_4^2\delta(q_1^2-M_1^2)\cdots\delta(q_4^2-M_4^2) \quad (19)$$
$$=J^{-1},$$

where $J=\det\partial q_i^2/\partial k_\mu=2^4\det q_{i\mu}$ is evaluated for $q_i^2=M_i^2$. The result of Mandelstam[3] and Kibble[8] is obtained from (18) by noting that $[\det q_{i\mu}]^2=\det q_i q_j$.

The reader will recognize $\det q_{i\mu}$ as the volume of the four-dimensional parallelepiped constructed with the q_i as edges. The vectors q_i have lengths M_i, and they have such directions that when drawn from a common vertex Q, their ends are vertices of the tetrahedron constructed from the external momenta (see Fig. 2). Complex vectors are to be used in drawing the figure, when necessary. This figure (a simplex) is one corner of the parallelepiped; its volume V is 1/4! times the volume of the parallelepiped. Hence $J=2^4 4!V$.

Landau's condition for the location of the (1234) singularity is that the point Q should lie in the hyperplane of the tetrahedron. In this case $V=0$. It should

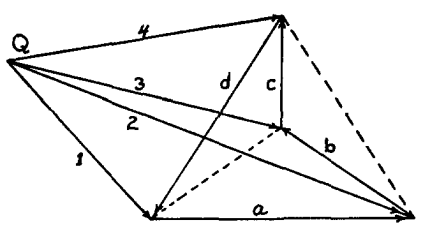


FIG. 2. The Mandelstam spectral function is the reciprocal of the volume of this figure.

be noted that the transformation from the k_μ to the q_i^2 is singular when the tetrahedron degenerates to a planar figure. But $4V$ is the product of the volume of the tetrahedron and the altitude of the point Q from the hyperplane of the tetrahedron, and when the volume of the tetrahedron vanishes, the altitude, for *fixed* lengths of the q_i, becomes infinite in such a way that V^{-1} is analytic.

B. Third-Order Vertex

The discontinuity across the (123) branch cut is

$$F_3=\frac{1}{2\pi}\int d^4k\delta_p(q_1^2-M_1^2)\delta_p(q_2^2-M_2^2)\delta_p(q_3^2-M_3^2). \quad (20)$$

Consider the point Q whose squared distances from the vertices of the triangle (p_a, p_b, p_c) are q_i^2 (see Fig. 3). The locus of Q in four-dimensional space is a circle whose radius κ is the altitude of Q from the plane of the triangle. Transforming to new variables, we have

$$d^4k=\kappa d\varphi d^3k=\kappa d\varphi\prod(dq_i^2)J_3^{-1}, \quad (21)$$

[8] T. W. B. Kibble, Phys. Rev. **117**, 1159 (1960).

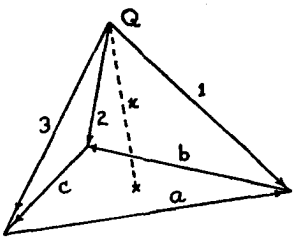


FIG. 3. Geometrical construction associated with the third-order vertex.

where $J_3=8\det q_{ia}$ is a 3×3 determinant. Hence we obtain

$$F_3=\kappa J_3^{-1}. \quad (22)$$

Now $\det q_{ia}$ is 3! times the volume of the tetrahedron in Fig. 3, which in turn is $\frac{1}{3}\kappa\mathfrak{A}$, where $\mathfrak{A}$ is the area of the triangle. Therefore,

$$F_3=2^{-4}\mathfrak{A}^{-1}$$
$$=\tfrac{1}{4}\{p_a^4+p_b^4+p_c^4-2p_a^2p_b^2-2p_a^2p_c^2-2p_b^2p_c^2\}^{-\frac{1}{2}}. \quad (23)$$

We see that F_3, and therefore also F on at least one sheet, is singular when $\mathfrak{A}=0$. In this example, a singularity of the matrix $\partial q_i^2/\partial k_\mu$ actually is associated with a singularity of F. The singularity can be shown to correspond, in terms of the Feynman parametrization discussed in the Introduction, to the case $\alpha_1+\alpha_2+\alpha_3=0$.

C. Example of Redundant Lines

Consider the graph shown in Fig. 1 which has five lines in one loop. Landau's procedure shows there is a singularity when all five $A_i=0$, but this is not a branch point. The discontinuity across the (1234) branch cut is shown by the method of Sec. III.A to be

$$F_4=J^{-1}(q_5^2-M_5^2)^{-1}, \quad (24)$$

where J and q_5^2 are functions of the external momenta and of $M_1,\cdots,M_4$. When the external momenta are such that $q_5^2=M_5^2$, F_4 has a pole. Since F_4 is the difference between values of the amplitude F on two adjacent sheets, and since the (12345) singularity only appears on one of them, F also has a pole. The location of the pole corresponds to the possibility of drawing the Landau diagram with four-dimensional vectors; the nonexistence of a branch cut corresponds to the impossibility of buckling the diagram into an extra dimension.

ACKNOWLEDGMENTS

This work was all done at the Carnegie Institute of Technology, but the author's introduction to this subject and acquaintance with Landau's paper began while he was a summer visitor at the Brookhaven National Laboratory. The author accordingly wishes to thank Dr. S. A. Goudsmit, Dr. G. C. Wick, and the other members of the Brookhaven National Laboratory staff for their hospitality, and also wishes to thank the A. P. Sloan Foundation for financial support. The author is grateful to Professor Z. Nehari for discussion of Hadamard's method and its applications.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 1, NUMBER 5 SEPTEMBER-OCTOBER, 1960

Consistency of the Canonical Reduction of General Relativity

R. ARNOWITT*
Department of Physics, Syracuse University, Syracuse, New York, and Brandeis University, Waltham, Massachusetts

S. DESER*
Department of Physics, Brandeis University, Waltham, Massachusetts

AND

C. W. MISNER†
Palmer Physical Laboratory, Princeton University, Princeton, New Jersey, and Brandeis University, Waltham, Massachusetts

(Received April 22, 1960)

The question of consistency of the canonical reduction of general relativity (obtained by eliminating constraints and also imposing coordinate conditions in the action or generator) is examined. It is shown that the equations of motion obtained from this "reduced" formalism agree with the original Einstein equations. Agreement is also established for the generators of space-time translations. In order to establish consistency, it is necessary to discard certain well-defined divergence terms in the original Lagrangian. These would otherwise appear as nondivergences in the reduced Lagrangian, incorrectly altering the equations.

1. INTRODUCTION

As is well known, general relativity differs from all other field theories in that its "stress tensor" vanishes (to within a divergence) as a consequence of the constraint equations of the theory. This singular behavior arises due to the invariance of relativity under general coordinate transformations. Thus the generator of translations with respect to the (arbitrary) time coordinate would indeed be expected to vanish since any apparent "time" translation could be removed by a mere relabeling. Real motion is to be expressed in terms of physically meaningful time and space coordinates which must be functionals of the metric. Therefore, in order to isolate the dynamics of the system (i.e., to obtain the correct nonvanishing stress tensor), one must specify the dynamical variables as functions of those variables being used as coordinates. Such a procedure is equivalent to imposing coordinate conditions; it also involves elimination of redundant variables by means of the constraint equations.

This program has been previously carried out,[1] and results in a canonical form for the theory, that is, in a "reduced" Lagrangian of the form

$$\mathfrak{L}=\sum_{A=1}^{2}\pi_A\partial\phi^A/\partial t-\mathfrak{IC}[\pi_A,\phi^A], \tag{1.1}$$

where ϕ^A,π_A are two independent canonically conjugate pairs of field variables and $\mathfrak{IC}$ is the nonvanishing Hamiltonian density. The functions ϕ^A and π_A, as well as those variables being used as the space-time coordinates, were expressed explicitly in terms of the metric field $g_{\mu\nu}$ and its first time derivatives. The form of Eq. (1.1) automatically ensured the internal consistency of the canonical formalism in that the Poisson bracket (P.B.) equations of motion are identical with the Hamilton-Lagrange equations of motion obtained by varying Eq. (1.1). Correspondingly, the canonical momentum $\mathbf{P}^C=-\int d^3r\Sigma\pi_A\nabla\phi^A$ arising in the reduced formalism of Eq. (1.1), correctly generates spatial translations there:

$$[f(\pi_A,\phi^A),\mathbf{P}^C]=\nabla f. \tag{1.2}$$

The primary consistency check of the canonical formalism, however, lies in the demonstration of agreement between it and the original Einstein equations. In the process of eliminating redundant variables some subtleties arise, which it is the purpose of this note to examine. In particular, we shall verify that indeed this "external" consistency does hold and that the expressions given for energy-momentum correctly generate space-time translations in the chosen coordinate frame. This result is valid also when matter is coupled to the gravitational field (see Appendix).

In the canonical reduction to Eq. (1.1) from the Einstein-Lagrangian, a complication arises not present in analogous situations for particle mechanics or simple field theories: a term which, in the original Lagrangian (or "Hamiltonian") is a pure divergence, may cease to be a divergence upon elimination of the redundant variables and hence may contribute to the equations of motion obtained from the reduced Lagrangian (or Hamiltonian). Consequently, a Lagrangian obtained by substituting solutions from the constraint equations into the original one may give incorrect equations of motion. For example, suppose a theory contained several dynamical variables ϕ_A and a redundant variable C with the constraint

* Supported in part by the National Science Foundation and Air Force Office of Scientific Research under contract.

† Alfred P. Sloan Research Fellow.

[1] In text the following papers will be referred to as I through V: (I) R. Arnowitt and S. Deser, Phys. Rev. **113**, 745 (1959); (II) R. Arnowitt, S. Deser, and C. Misner, *ibid.* **116**, 1322 (1959); (III) **117**, 1595 (1960); (IIIa) present paper; (IV) **118**, 1100 (1960); (IVa) Nuovo cimento (to be published); (IVb) Phys. Rev. (to be published); (IVc) *ibid.* (to be published); (V) *ibid.* (to be published).

equation

$$\nabla^2 C=\phi_1{}^2. \tag{1.3}$$

A term $\nabla\cdot\nabla C$ in the original Lagrangian makes no contribution to the equations of motion there, while in the reduced Lagrangian (where the constraints are eliminated), this term would appear as $\phi_1{}^2$ and *would* contribute. Further, the addition of a divergence may correspond merely to the change of position of a derivative in the original Lagrangian. Whether or not such terms should be included in the original Lagrangian is not directly determined by the original field equations. Yet, the decision to keep or drop these divergences can strongly modify the resulting field equations of the reduced system.

In Lagrangians of the form obtained by parameterizing standard field theories (which form includes general relativity as an "already parameterized" case), there is a unique specification of what divergence terms are to be retained. This requirement leads precisely to a form for the field Lagrangian, which is the natural generalization of the particle case. In the next section it will be shown that upon elimination of redundant variables at this point, the correct equations of motion may be obtained from the reduced Lagrangian. In Sec. 3, the consistency of the spatial generators will be demonstrated. We shall show that the expression previously obtained (in III) for the momentum density (by inserting constraints and coordinate conditions into the original generator), differs from the canonical one merely by a divergence in the canonical variables. Finally, some of our results will be used to comment (in Sec. 4) on other techniques of dealing with general relativity.

2. REDUCTION OF THE LAGRANGIAN

It was shown in III that the Lagrangian of general relativity, $\mathfrak{L}_{GR}=(-{}^4g)^{\frac{1}{2}}\,{}^4R$, could be recast into the form[2]

$$\mathfrak{L}_{GR}=-g_{ij}\partial\pi^{ij}/\partial t-N_\mu R^\mu -2[g^{\frac{1}{2}}N_{0|}{}^{i}+N_j(\pi^{ij}-\tfrac{1}{2}g^{ij}\pi)]_{,i}, \tag{2.1}$$

where

$$R^0\equiv -g^{\frac{1}{2}}[{}^3R+g^{-1}(\tfrac{1}{2}\pi^2-\pi_{ij}\pi^{ij})] \tag{2.2a}$$

and

$$R^i\equiv -2\pi^{ij}{}_{|j} \tag{2.2b}$$

are linear combinations of $G^0{}_\mu\equiv{}^4R^0{}_\mu-\frac{1}{2}\delta^0{}_\mu\,{}^4R$. Here

$$N_0\equiv(-g^{00})^{-\frac{1}{2}},\quad N_i\equiv{}^4g_{0i},\quad \pi\equiv\pi^{ij}g_{ij},$$
$$\text{and}\qquad \pi^{ij}\equiv(-{}^4g)^{\frac{1}{2}}[{}^4\Gamma^0{}_{mn}-g_{mn}\,{}^4\Gamma^0{}_{pq}g^{pq}]g^{mi}g^{nj}. \tag{2.3}$$

[2] We use units such that $\kappa\equiv 16\pi\gamma c^{-4}=1=c$, where γ is the Newtonian gravitational constant. Latin indices run from 1 to 3, Greek from 0 to 3, and $x^0=t$. All tensors and covariant operations are *three*-dimensional unless specified; ${}^3R$ is the curvature scalar of g_{ij} (not ${}^4g_{\mu\nu}$) and $g^{ij}(\neq{}^4g^{ij})$ is the matrix inverse to g_{ij}. The vertical bar "|" indicates covariant differentiation with respect to g_{ij}, and N_0 is treated as a three-scalar. Ordinary differentiation is denoted by a comma or the symbol ∂_μ.

This Lagrangian is of the general form[3]

$$\mathfrak{L}=\sum_{I=1}^{N+4}\pi_I\partial\phi^I/\partial t-N_\mu R^\mu(\phi^J,\pi_J) +[\mathfrak{F}^i(\pi_J,\phi^J,N_\mu)]_{,i}, \tag{2.4}$$

which form is also found when matter is coupled to the gravitational field (see IV and V). The same structure arises upon parameterization of standard field theories (e.g., the scalar meson example in III), and is the straightforward generalization of the parameterized particle Lagrangian[4]

$$L=\sum_{i=1}^{M}p_iq^{i\prime}+p_{M+1}q^{M+1\prime}-NR,$$

where the constraint R may be taken as $R\equiv p_{M+1}+\mathfrak{IC}(p_i,q_i)$; a prime denotes a derivative with respect to the (arbitrary) parameter τ which replaces the time in this formulation. Note that the divergence term of $\mathfrak{L}$ in Eq. (2.4) is determined uniquely by the requirement that the R^μ not be functions of the Lagrange multipliers N_μ. For example, gradients appearing in R^μ cannot be moved by means of integrations by parts, which would give rise to gradients of N_μ outside the total divergence term; this is not allowed in the standard form of Eq. (2.4).

In the remainder of this section, we shall compute the equations of motion obtained by varying the action

$$I=\int d^4x\mathfrak{L}, \tag{2.5}$$

and we shall insert into these the solutions of the constraint equations and the coordinate conditions. (The term $\mathfrak{F}^i{}_{,i}$ in $\mathfrak{L}$ does not contribute in this analysis.) We will then verify that these equations agree with those obtained by varying the reduced Lagrangian $\mathfrak{L}_R$; here $\mathfrak{L}_R$ is that Lagrangian obtained by substituting constraints and coordinate conditions into $\mathfrak{L}$ *with the divergence* $\mathfrak{F}^i{}_{,i}$ *discarded.*

The constraint equations $R^\mu=0$ result from varying N_μ. They state that four of the momenta π_I are not independent canonical variables, corresponding to the fact that four of the ϕ^I are not really field variables, but rather physical space-time coordinates. We assume that ϕ^I and π_I are so chosen that the coordinate conditions

$$\phi^{N+\mu}=x^\mu \tag{2.6}$$

[3] Since constraint equations, by definition, contain no time derivatives, their solutions never eliminate them. Thus a total time derivative in the Lagrangian is harmless and $\pi\partial_0\phi$ is equivalent to $-\phi\partial_0\pi$ there, either before or after the elimination of redundant variables. For further discussion of the relation of total time derivatives to canonical transformations, see IVa.

[4] See, for example, C. Lanczos, *The Variational Principles of Mechanics* (University of Toronto Press, Toronto, Canada, 1949), or II.

are physically acceptable.[5] With the convention that the index A runs from 1 to N, we write the solutions of the constraint equations

$$\delta I/\delta N_\mu = R^\mu(\pi_{N+\mu},\phi^A,\pi_A,\phi^{N+\mu})=0, \tag{2.7}$$

in the form

$$\pi_{N+\mu}=\mathcal{T}^0{}_\mu[\phi^A,\pi_A,\phi^{N+\mu}]. \tag{2.8}$$

Since R^μ may involve spatial derivatives, we use the functional notation $\mathcal{T}^0{}_\mu[\phi^A,\pi_A,\phi^{N+\mu}]$ to take into account the possible appearance of operators such as $(1/\nabla^2)$ in Eqs. (2.8). Indeed, just this circumstance (having no simple analog in particle mechanics) gives rise to the "divergence" problem under discussion here.

Varying ϕ^A and π_A in the action integral yields the equations of motion

$$\partial\pi_A(x)/\partial t=-\int d^4y N_\mu(y)[\delta R^\mu(y)/\delta\phi^A(x)] \tag{2.9a}$$

$$\partial\phi^A(x)/\partial t=\int d^4y N_\mu(y)[\delta R^\mu(y)/\delta\pi_A(x)]. \tag{2.9b}$$

Since the R^μ contain no time derivatives, $\delta R^\mu(y)/\delta\phi^A(x)$ will contain $\delta(y^0-x^0)$ as a factor; however, spatial derivatives of ϕ^A in R^μ produce derivatives of delta functions in $\delta R^\mu/\delta\phi^A$, and thus spatial derivatives of N_μ may appear when the integral in Eq. (2.9a) [or Eq. (2.9b)] is evaluated. The equations

$$\partial\phi^{N+\mu}(x)/\partial t=\int d^4y N_\alpha(y)[\delta R^\alpha(y)/\delta\pi_{N+\mu}(x)] \tag{2.10}$$

serve to determine the Lagrange multipliers N_μ when the functions $\phi^{N+\mu}$ are specified by the coordinate conditions.

In order to substitute the solution of the constraints, Eq. (2.8), into the dynamic equations (2.9) we expand R^μ in a functional Taylor series[6] about the point $\pi_{N+\mu}=\mathcal{T}^0{}_\mu$ (indicated in the following by $\pi=\mathcal{T}$). The zeroth-order term vanishes and we have

$$R^\mu(y)=\int d^4z[\pi_{N+\alpha}(z)-\mathcal{T}^0{}_\alpha(z)]\times[\delta R^\mu(y)/\delta\pi_{N+\alpha}(z)]_{\pi=\mathcal{T}}+\cdots. \tag{2.11}$$

This allows us to compute

$$\delta R^\mu(y)/\delta\phi^A(x)=-\int d^4z[\delta\mathcal{T}^0{}_\alpha(z)/\delta\phi^A(x)]\times[\delta R^\mu(y)/\delta\pi_{N+\alpha}(z)]_{\pi=\mathcal{T}}+\cdots \tag{2.12a}$$

and

$$\delta R^\mu(y)/\delta\pi_A(x)=-\int d^4z[\delta\mathcal{T}^0{}_\alpha(z)/\delta\pi_A(x)]\times[\delta R^\mu(y)/\delta\pi_{N+\alpha}(z)]_{\pi=\mathcal{T}}+\cdots, \tag{2.12b}$$

where the terms represented by $\cdots$ contain $[\pi_{N+\alpha}-\mathcal{T}^0{}_\alpha]$ as a factor and so vanish when the constraints are inserted. We now substitute Eqs. (2.12) into (2.9) and eliminate $\pi_{N+\alpha}$ by using Eqs. (2.8). The coordinate conditions (2.6) reduce Eq. (2.10) to $\delta^\mu{}_0=\int N_\alpha\delta R^\alpha/\delta\pi_{N+\mu}$. By using these results, the dynamic equations then become

$$\partial\pi_A(x)/\partial t=\int d^4z[\delta\mathcal{T}^0{}_0(z)/\delta\phi^A(x)] \tag{2.13a}$$

$$\partial\phi^A(x)/\partial t=-\int d^4z[\delta\mathcal{T}^0{}_0(z)/\delta\pi_A(x)]. \tag{2.13b}$$

These equations, equivalent to the original set [Eqs. (2.9) and (2.7)] under the coordinate condition $\partial_0\phi^{N+\mu}=\delta^\mu{}_0$ are easily seen to be just the Hamiltonian equations obtained from the Lagrangian

$$\mathfrak{L}_R=\sum_{A=1}^{N}\pi_A\partial\phi^A/\partial t-\mathfrak{IC}, \tag{2.14a}$$

where

$$\mathfrak{IC}=-\mathcal{T}^0{}_0[\phi^A,\pi_A]. \tag{2.14b}$$

Note that only the differential statement of the coordinate conditions, $\partial_0\phi^{N+\mu}=\delta^\mu{}_0$, was needed in deriving the preceding result. In Eq. (2.14b) we have assumed that these coordinate conditions are chosen in such a way that $\mathcal{T}^0{}_0$ has no explicit x^μ dependence in order that a set of standard conservations laws hold.[7]

For general relativity, one must rearrange the $g_{ij}\partial_t\pi^{ij}$ part of the Lagrangian of Eq. (2.1) in order to explicitly apply the general methods discussed above. This is accomplished by making an orthogonal decomposition on π^{ij} and g_{ij} similar to the one used in III. We write

$$g^{ij}=\delta_{ij}+h_{ij}=\delta_{ij}+h_{ij}{}^{TT}+\tfrac{1}{2}[\delta_{ij}h^T-(1/\nabla^2)h^T{}_{,ij}]+h_{i,j}+h_{j,i} \tag{2.15a}$$

$$\pi^{ij}=\pi^{ijTT}+\tfrac{1}{2}[\delta_{ij}\nabla^2\bar{\pi}^T-\bar{\pi}^T{}_{,ij}]+\pi^i{}_{,j}+\pi^j{}_{,i}. \tag{2.15b}$$

Note that $\nabla^2\bar{\pi}^T$ used here is just π^T of III and h_{ij} approaches zero asymptotically. The Lagrangian (2.1) now becomes

$$\mathfrak{L}=\pi^{ijTT}\partial_t h_{ij}{}^{TT}+\nabla^2 h^T\partial_0(-\tfrac{1}{2}\bar{\pi}^T)-2\pi^{ij}{}_{,j}\partial_0 h_i+N_\mu R^\mu+\mathfrak{F}^{\prime i}{}_{,i}. \tag{2.16}$$

[5] For general relativity, the main requirement is that these conditions are consistent with an asymptotically flat metric at spatial infinity.

[6] V. Volterra, *Theory of Functionals* (Blackie and Son Ltd., London, 1930), p. 25.

[7] That this can in fact be achieved for general relativity was shown in III.

In obtaining Eq. (2.16), various spatial divergences have been included in $\mathfrak{F}^i{}_{,j}$ and a total time derivative neglected.[3] By choosing

$$\phi^{N+0}\equiv-\tfrac{1}{2}\bar{\pi}^T,\quad \phi^{N+i}=x^i+h_i, \tag{2.17}$$

one has

$$\mathfrak{L}=\pi^{ij TT}\partial_t h_{ij}{}^{TT}-(-\nabla^2 h^T)\partial_t\phi^{N+0}-2\pi^{ij}{}_{,j}\partial_t\phi^{N+i}+N_\mu R^\mu+\mathfrak{F}^i{}_{,j}. \tag{2.18}$$

The formal derivation now follows with the association

$$\pi_A\equiv\pi^{ij TT},\qquad \phi^A\equiv h_{ij}{}^{TT}$$
$$\pi_{N+0}\equiv\nabla^2 h^T,\quad \pi_{N+i}\equiv-2\pi^{ij}{}_{,j}, \tag{2.19}$$

since $\mathfrak{L}$ is now in the form of Eq. (2.4). As was discussed in III, the constraint equations $R^\mu=0$ can indeed be solved for $\pi_{N+\mu}$ and are independent of x^μ when the coordinate conditions $\phi^{N+\mu}=x^\mu$ are imposed.

The divergence $\mathfrak{F}^i{}_{,j}$ in Eq. (2.4) played no role whatsoever, as a divergence never affects a variational derivative. *However*, $\mathfrak{L}_R$ of Eq. (2.14) will not necessarily be the reduced Lagrangian obtained by inserting coordinate conditions (2.6) and constraints (2.8) into the original Lagrangian of Eq. (2.4) unless $\mathfrak{F}^i{}_{,j}$ is dropped, as was illustrated by the simple example in Sec. I. (If some particular $\mathfrak{F}^i{}_{,j}$ happens to remain a divergence after $\phi^{N+\mu}$, $\pi_{N+\mu}$ and N_μ have been eliminated, the reduced Lagrangian will be equivalent to $\mathfrak{L}_R$, but in general the correct way to obtain $\mathfrak{L}_R$ is to drop the $\mathfrak{F}^i{}_{,j}$ term before eliminating the redundant variables.) In relativity it is very easy to obtain equations such as $\partial\phi^A/\partial t=0=\partial\pi_A/\partial t$ by including a wrong divergence term in $\mathfrak{L}$ while substituting constraints. In fact, it is easy to obtain such equations for variables ϕ^A and π_A, which are not even constants of motion in the linearized approximation.

The result obtained in this section is not an unexpected one. If one drops the term $\mathfrak{F}^i{}_{,j}$ in Eq. (2.4) (which does not affect the equations of motion obtained from $\mathfrak{L}$), $\mathfrak{L}$ becomes just the field analog of the parameterized particle Lagrangian. It is well known[4] in the particle mechanics case that one can then impose the constraint and coordinate conditions without error.

3. REDUCTION OF THE GENERATOR AND FIELD MOMENTUM

We consider here the generator associated with the Lagrangian of Eq. (2.4). As was discussed in III, the process of reducing the theory to canonical form can be carried out directly in the generator. We will show explicitly that this reduction is identical to the previous one. In addition, it will enable us to display the field momentum density automatically. We will see that this differs from the field momentum density derived from the reduced Lagrangian, Eq. (2.14), by a divergence of canonical variables (a "true divergence"); this will check the consistency of the spatial displacement generators.

The generator arises from the Lagrangian of Eq. (2.4) (with the $\mathfrak{F}^i{}_{,j}$ now discarded) as the surface integral term in the endpoint variation of the action. It has the form

$$G=\int d^3r\Big[\sum_{I=1}^{N+4}\pi_I\delta\phi^I-\mathcal{T}^0{}_\mu{}'\delta x^\mu\Big]. \tag{3.1}$$

The stress tensor $\mathcal{T}^0{}_\mu{}'$ vanishes (to within a divergence) as a consequence of the constraint equations, which again reflects the general covariance of the theory. Thus, the canonical tensor which arises when one makes arbitrary coordinate variations δx^μ in Eq. (2.5) (but no associated Lorentz transformations) is given by

$$\mathcal{T}^0{}_0{}'=N_\mu R^\mu \tag{3.2}$$

and

$$\mathcal{T}^0{}_k{}'=-\sum_{I=1}^{N+4}\pi_I\phi^I{}_{,k}. \tag{3.3}$$

The constraint equations $R^\mu=0$ obviously account for the vanishing of $\mathcal{T}^0{}_0{}'$ and we will see that $\mathcal{T}^0{}_k{}'$ also vanishes (to within a divergence). The generator therefore becomes

$$G=\int d^3r\Big[\sum_{A=1}^{N}\pi_A\delta\phi^A+\sum_{\mu=0}^{3}\pi_{N+\mu}\delta\phi^{N+\mu}\Big]. \tag{3.4}$$

On inserting solutions (2.8) of the constraint equations and the coordinate conditions (2.6), one has

$$G=\int d^3r\Big[\sum_{A=1}^{N}\pi_A\delta\phi^A+\mathcal{T}^0{}_\mu\delta x^\mu\Big], \tag{3.5}$$

which is the standard canonical form for a field theory generator with N canonical pairs of variables ϕ^A,π_A and a generator of space-time translations $\int\mathcal{T}^0{}_\mu\delta x^\mu d^3r$. The Hamiltonian $\mathcal{H}=-\int d^3r\,\mathcal{T}^0{}_0$ is identical to the one obtained in Sec. 2, thus showing the correctness of the time-translation part. From the reduced Lagrangian $\mathfrak{L}_R$ of Eq. (2.14), one knows that the correct generator of space translations is the canonical field momentum

$$P_k=\int d^3r\,\mathcal{T}^{C0}{}_k\equiv-\int d^3r\sum_A\pi_A\phi^A{}_{,k}. \tag{3.6}$$

{By Eq. (3.6) and the fundamental P.B. relations, one has $[f,P_k]=\partial_k f$}. From Eq. (3.3), one has

$$\mathcal{T}^0{}_k{}'=\mathcal{T}^{C0}{}_k-\sum_{\mu=0}^{3}\pi_{N+\mu}\phi^{N+\mu}{}_{,k}. \tag{3.7}$$

On inserting the solutions of the constraint equations (2.8) and the coordinate conditions (2.6), one obtains

$$\mathcal{T}^0{}_k{}'=\mathcal{T}^{C0}{}_k-\mathcal{T}^0{}_k. \tag{3.8}$$

Thus, the vanishing of $\mathcal{T}^0{}_k{}'$ coincides with the

consistency requirement for the spatial-translation generators.[8]

We now show that $\mathcal{T}^0{}_k$ differs from $\mathcal{T}^{C0}{}_k$ by at most a divergence of the canonical variables. We limit ourselves to the case of general relativity. Here $\phi^A \equiv h_{ij}{}^{TT}$, $\pi_A \equiv \pi^{ijTT}$, $\mathcal{T}^0{}_0 = \nabla^2 h^T$, and $\mathcal{T}^0{}_i = -2\pi^{ij}{}_{,j}$ (notation is as in Sec. 2), while $\mathcal{T}^{C0}{}_k = -\pi^{ijTT} h_{ij}{}^{TT}{}_{,k}$. The constraint equation which determines $\mathcal{T}^0{}_i$ in terms of the canonical variables reads

$$-2\pi^{ij}{}_{|j} = 0 \tag{3.9}$$

and can be written as

$$-2\pi_i{}^j{}_{,j} \equiv -2(g_{ik}\pi^{kj})_{,j} = -\pi^{lj} g_{lj,i}. \tag{3.10}$$

On inserting the orthogonal decomposition for π^{lj}, g_{lj} in Eq. (3.10), one has

$$-2\pi_i{}^j{}_{,j} = -\pi^{ljTT} h_{lj}{}^{TT}{}_{,i} - [2\pi^l (h_{lj}{}^{TT}{}_{,i} + h_{lj}{}^T{}_{,i}) - \tfrac{1}{2}\pi^{ljTT}(1/\nabla^2) h^T{}_{,li}]_{,j}. \tag{3.11}$$

It is now necessary to show that the $[\ ]_{,j}$ term of Eq. (3.11) is indeed a true divergence in the sense that $\int [\ ]_{,j} d^3r$ vanishes for arbitrary values of the canonical variables. These latter must vanish rapidly for the system to be bounded (see IVb). When π^l and g^T are expressed in terms of the canonical variables, they are seen to decrease as $1/r$ at infinity. Thus all the terms in the bracket vanish faster than $1/r^2$, and therefore $-2\pi_i{}^j{}_{,j}$ differs from $\mathcal{T}^{C0}{}_i$ by a true divergence. Further, since

$$g_{ij} = \delta_{ij} + h_{ij}{}^{TT} + h_{ij}{}^T,$$

we find

$$-2\pi_i{}^j{}_{,j} = -2\pi^{ij}{}_{,j} + 2[\pi^{lj}(h_{il}{}^{TT} + h_{ij}{}^T)]_{,j}, \tag{3.12}$$

where the bracket in Eq. (3.12) is also a true divergence. Hence

$$P_i = \int d^3r\, \mathcal{T}^{C0}{}_i = \int \mathcal{T}^0{}_i d^3r \equiv -2\int d^3r \pi^{ij}{}_{,j}. \tag{3.13}$$

This establishes explicitly that the term $\int d^3r \mathcal{T}^0{}_i \delta x^i$ arising in Eq. (3.5) correctly generates spatial translations. The result holds also when matter is coupled to the gravitational field, as shown in the Appendix.

Equivalence between $\mathcal{T}^0{}_k$ and $\mathcal{T}^{C0}{}_k$ is valid for a parameterized Lorentz covariant theory as well. If the parameterization is carried out by rewriting the Lagrangian in a generally covariant form, but with $g_{\mu\nu} = \eta_{\alpha\beta}(\partial\phi^{N+\alpha}/\partial x^\mu)(\partial\phi^{N+\beta}/\partial x^\nu)$ (where $\eta_{\alpha\beta}$ is the Lorentz metric), then one finds for the $\mathcal{T}^0{}_k$ of Eq. (3.5) the standard symmetric stress tensor of the original theory.[9] As is well known, this differs by a divergence from $\mathcal{T}^{C0}{}_k$.

[8] Conversely, the physical requirement that $T^0{}_i'$ must vanish since it is the generator of translations with respect to the parameters (on which the theory does not depend), would then lead to the equivalence of $T^0{}_i$ and $T^{C0}{}_i$.

[9] It might be noted that the alternate type of parameterization carried out for the scalar meson field in III, would lead to the canonical tensor for $\mathcal{T}^0{}_k$.

4. DISCUSSION

In the preceding sections we have seen that in a Lagrangian of the form

$$\mathcal{L} = \sum_{I=1}^{N+\mu} \pi_I \partial\phi^I/\partial t - N_\mu R^\mu(\pi_J, \phi^J), \tag{4.1}$$

one may insert the solutions of the constraint equations $\pi_{N+\mu} = \mathcal{T}^0{}_\mu$ and the canonical coordinate conditions $\partial_t \phi^{N+\mu} = \delta_0{}^\mu$ to obtain a reduced Lagrangian

$$\mathcal{L}_R = \sum_{A=1}^{N} \pi_A \partial\phi^A/\partial t + \mathcal{T}^0{}_0[\pi_A, \phi^A] \tag{4.2}$$

whose equations of motion are equivalent to those of Eq. (4.1). Consequently, given a Lagrangian that differs from Eq. (4.1) by a total 3-divergence, the consistent reduction method requires that this divergence be neglected before eliminating the constraints. In general relativity, then, the last term in Eq. (2.16) should be omitted (as was done in III). Thus, the canonical equations of motion in III are correctly the Einstein equations. Indeed, if one had included the divergence which actually appears in Eq. (3.1), the energy obtained from the reduced Hamiltonian for the Schwarzschild solution would have become $\frac{3}{2}m$ rather than m. Similarly, this Hamiltonian would give rise to wrong equations of motion even in the linearized approximation.

It was also seen that the generator associated with the Lagrangian of Eq. (4.1) gives rise, when constraints are eliminated and coordinate conditions imposed, to the generator obtained from the reduced Lagrangian of Eq. (4.2). Thus the consistency of the spatial translation generators is guaranteed.

Recently, Dirac[10] has suggested an entirely different procedure for obtaining a nonvanishing Hamiltonian in general relativity. His analysis is performed within a generalized Hamiltonian formalism[11] and does not make direct use of the fact that general relativity is a parameterized theory when presented in generally covariant form. The method involves writing the vanishing Hamiltonian $N_\mu R^\mu$ as

$$N_\mu R^\mu = (N_0 - 1)R^0 + N_i R^i + R^0, \tag{4.3}$$

and dropping a particular divergence in the last term. Thus, before the redundant variables are eliminated, the new Hamiltonian density is weakly a divergence. Next, the redundant variables are eliminated by means of the constraint equations, and one arrives at a reduced Hamiltonian which is not a divergence in terms of the remaining variables. The situation here is just of the type discussed in Eq. (1.3). Such a procedure seems to us to be logically incomplete. While with the

[10] P. A. M. Dirac, Phys. Rev. **114**, 924 (1959); Proc. Roy. Soc. (London) **A246**, 333 (1958); Phys. Rev. Letters **2**, 368 (1959).

[11] P. A. M. Dirac, Proc. Roy. Soc. (London) **A246**, 326 (1958).

particular choice of divergence that Dirac makes, the reduced Hamiltonian gives correct equations of motion in the linearized approximation, no general proof is given that the full theory is correct in this respect. It is of interest to note however, that Dirac's choice does lead to the correct *numerical* value of the energy.[12]

APPENDIX

Here we extend, for the case of coupling, the proof given in text that the field energy momentum of the reduced generator correctly generates space-time translations. For the momentum, our derivation consists, as in Sec. 3, in showing that this momentum density differs from the canonical one by a true divergence. In V, it is shown that when the Maxwell field and point charges are coupled to the gravitational field, the coefficient of δx^i in the total reduced generator is still $-2\pi^{ij}{}_{,j}$. The effect of the matter enters through the constraint equations used to solve for $-2\pi^{ij}{}_{,j}$ in terms of the canonical variables of all the fields. Thus, Eq. (3.9) now reads

$$-2\pi^{ij}{}_{|j}=g^{ik}\mathcal{T}_M{}^0{}_k, \tag{A1}$$

where the matter momentum density is

$$\mathcal{T}_M{}^0{}_k=B_{kj}\mathcal{E}^j+[p_k(t)-eA_k{}^T]\delta^3[\mathbf{r}-\mathbf{r}(t)]. \tag{A2}$$

In Eq. (A2), $B_{kj}\equiv A_j{}^T{}_{,k}-A_k{}^T{}_{,j}$ is the magnetic field, $\mathcal{E}^j\equiv\mathfrak{F}^{0j}$ is the electric field density and $A_k{}^T$ is the transverse part of the vector potential. The quantities $r^k(t)$ and $p_k(t)$ are the canonical variables of the particle. Hence $\mathcal{T}_M{}^0{}_k$ is independent of the metric and has the same form as the symmetric stress-tensor's momentum density in flat space. Consequently, it differs from the canonical momentum density,

$$\mathcal{T}_M^{C0}{}_k=\mathcal{E}^{jT}A_j{}^T{}_{,k}+p_k(t)\delta^3[\mathbf{r}-\mathbf{r}(t)]$$

only by a divergence. The proof of this makes use of the fact that $\mathcal{E}=\mathcal{E}^T+\nabla(1/\nabla^2)e\delta^3$, where ∇^2 is the flat space Laplacian operator.

From Eq. (A1), one obtains the extended form of Eq. (3.11):

$$-2\pi_i{}^j{}_{,j}=-\pi^{lmTT}g_{lm}{}^{TT}{}_{,i}+\mathcal{T}_M{}^0{}_i+\mathfrak{D}_{ij,j}, \tag{A3}$$

where $\mathfrak{D}_{ij,j}$ is the divergence in Eq. (3.11). Hence $-2\pi_i{}^j{}_{,j}$ differs by a divergence from the *total* canonical momentum density,

$$\mathcal{T}^{C0}{}_i=-\pi^{lmTT}g_{lm}{}^{TT}{}_{,i}+\mathcal{T}_M^{C0}{}_i. \tag{A4}$$

Finally, as was shown in Eq. (3.12), $-2\pi_i{}^j{}_{,j}$ differs from $-2\pi^{ij}{}_{,j}$ by a divergence.

That $-\nabla^2 g^T$ correctly gives rise to time translations in the coupled case, follows immediately from the results of Sec. 2. In V, it was shown that the coefficient of $\partial_t(-\frac{1}{2}\bar{\pi}^T)$ in the Lagrangian of Eq. (2.18) was unaltered by the presence of matter. Therefore, the discussion of Sec. 2 is completely unchanged by coupling.

[12] Since Dirac's Hamiltonian differs from zero by a divergence, its numerical value for a computation of the energy is given by this divergence. Thus

$$E=-\int d^3x[g^{-\frac{1}{2}}(gg^{ij})_{,j}]_{,i}=-\oint g^{-\frac{1}{2}}(gg^{ij})_{,j}dS_i.$$

On introducing the orthogonal decomposition of the metric [Eq. (2.15)], we may, in the surface integral at spatial infinity, neglect all terms beyond the linear one since $g_{ij}\to\delta_{ij}$ at spatial infinity. This gives $E=-\oint[g^T{}_{,i}+(g_{j,i}-g_{i,j})_{,j}]dS_i$, but the second term vanishes by Gauss' theorem, leaving

$$E=-\oint g^T{}_{,i}dS_i,$$

the coordinate independent value obtained in III and IV. This is also equal, as Dirac has noted, to the value obtained from the surface integral form of the Einstein pseudotensor. (This discussion assumes $g_{\mu\nu,\alpha}\sim 1/r^2$ at spatial infinity; see IVc for a more general treatment.)

Spacetime of the Elementary Particles*

R. FINKELSTEIN†
Department of Physics, University of California, Los Angeles, California
(Received February 10, 1960)

The possibility is examined that physical space is characterized by a torsion, or an asymmetric connection, which is determined by the matter field. There exists a space with uniform torsion with the same metrical properties as conventional microspace; it is isotropic and homogeneous with a very large radius of curvature ($R \cong 10^{28}$ cm). The momentum operators form a group and for practical purposes commute. The torsion defines at every point two kinds of parallel transfer or two screw motions of opposite helicity. There are, consequently, two kinds of spinor field associated with the space; they are distinguished by opposite coupling to the torsion. Viewed from within the Lorentz group the torsion produces an axial vector interaction. To interpret the given mathematical model, it is suggested that there exists a universal axial vector coupling between fermions represented by the spinor fields and bosons associated with the torsion; and that this interaction manifests itself macroscopically as a torsion of space, in the same general way that gravitational interactions correspond to a curvature of space. This general assumption leads to cosmological models characterized by relations connecting the average density of matter and the strength of the assumed interaction. For the observed average density of matter in the known universe ($\sim 10^{-30}$ g/cm^3) the proposed axial vector coupling turns out, for a space of uniform torsion, to be of the order of the strong interactions.

1. INTRODUCTION

QUANTUM electrodynamics does not exclude small departures from a Euclidean spacetime in situations where it has been tested and large departures at smaller distances. The failure of parity conservation, the new conservation laws, which are foreign to the Lorentz group, and the old difficulties of field theory suggest that the usual assumptions about microspace should be reexamined.

The vacuum correlation functions of quantum field theory (e.g., $\langle\psi(x)\psi(x')\rangle$) presumably characterize empty space; nevertheless, they are calculated under the assumption that configuration space is Euclidean. This procedure appears to ignore Mach's principle in the sense that configuration space is postulated to be independent of the matter field. At the same time the matter field does define, in terms of such correlation functions, a mathematical structure to be associated with space, but which does not, according to the usual theory, determine the geometry of that space.

Here we investigate, in a preliminary way, a generalized spacetime in which the structure of the continuum and certain quantum expectation functions are codetermined in accord with Mach's principle. In particular, the possibility is examined that physical space is characterized by a torsion, which is determined by the matter field. With respect to metrical properties, the proposed space agrees with conventional microspace; it is isotropic and homogeneous, with a very large radius of curvature ($R \cong 10^{28}$ cm). The particular model considered does not admit space inversions. It turns out that the proposed theory may describe, if it has any physical content, a universal axial vector coupling.

Attempts to generalize the theory of gravitation have concentrated on the role of the electromagnetic field. Here, on the contrary, we work in the neutral approximation suggested by elementary particle theory: all particles are regarded as neutral and the electromagnetic field vanishes identically. One might hope to get in this way a unified theory of the Fermi and gravitational interactions, which have the common properties of weakness and universality. There is nothing, however, in this approach, which obviously excludes the strong interactions and in fact the vanishing of the electromagnetic field is the condition for the exact isospin symmetry which characterizes the strong couplings.

The mathematical basis for our work was given by Cartan and Schouten[1]; an adaptation of their work for the purposes of this paper will now be presented.

2. COORDINATE AND ENNUPLE TRANSFORMATIONS

We consider a four-dimensional continuum in which the coordinates x label the abstract points. It is postulated that the physical laws do not depend on how this labelling is done, i.e., that the physical laws are covariant under the general coordinate transformation

$$x'^{\mu} = f(x^{\mu}). \tag{2.1}$$

Equation (2.1) induces the tensor transformation of a contravariant vector

$$\lambda'^{\mu} = \lambda^{\alpha} \partial x'^{\mu} / \partial x^{\alpha} \tag{2.1a}$$

and corresponding transformations for other tensors.

In addition to the global coordinate transformations, it is also necessary to consider certain local transformations (of ennuples). An ennuple is defined to be a set of n linearly independent vectors,[2] $\lambda^{\mu}{}_{s}$, where the latin index indicates the vector; then

$$\lambda \equiv |\lambda^{\mu}{}_{s}| \neq 0.$$

* This work was supported by the National Science Foundation.
† J. S. Guggenheim Fellow, 1959–1960.

[1] E. Cartan and J. A. Schouten, Akad. van Wetens., Amsterdam, Proc. 29, 803 (1926).
[2] Although ennuple refers to arbitrary n, here it will always mean $n=4$ or $n=3$.

An ennuple transformation is defined as follows:

$$\lambda'^{\mu}{}_{s}=\lambda^{\mu}{}_{m}a^{m}{}_{s}, \tag{2.2}$$

$$\lambda'=\lambda a. \tag{2.2a}$$

The determinant λ is therefore a pseudoscalar. Ennuples are important in any generally covariant spinor theory because they are needed to define spin transformations and the Lorentz limit of the general theory. Moreover, the structure of a space which admits distant parallelism, such as we shall discuss, may be exactly specified by a parallel ennuple field.

Since λ does not vanish, it is possible to find a covariant vector $\lambda^{t}{}_{\alpha}$, such that

$$\lambda^{\alpha}{}_{s}\lambda^{t}{}_{\alpha}=\delta^{t}{}_{s}, \tag{2.3a}$$

$$\lambda^{\alpha}{}_{s}\lambda^{s}{}_{\beta}=\delta^{\alpha}{}_{\beta}. \tag{2.3b}$$

The two ennuples $\lambda^{\alpha}{}_{s}$ and $\lambda^{s}{}_{\alpha}$ may be called reciprocal. They are not contravariant and covariant components of the same vector unless a metric is so defined that

$$\lambda^{s}{}_{\alpha}=g_{\alpha\beta}\lambda^{\beta}{}_{s}.$$

Then

$$\sum_{s}\lambda^{s}{}_{\alpha}\lambda^{s}{}_{\beta}=g_{\alpha\beta}.$$

However, consideration of the nature of the metric will be deferred until the connection has been discussed.

3. CONNECTION

Associate with a given vector λ^{μ} at x^{μ} another vector $\lambda^{\mu}+\delta\lambda^{\mu}$ at a nearby point according to either of the two equations

$$\delta\lambda^{\mu}=-L^{\mu}{}_{\alpha\beta}\delta x^{\alpha}\lambda^{\beta}, \tag{A $-$}$$

$$\delta\lambda^{\mu}=-L^{\mu}{}_{\alpha\beta}\lambda^{\alpha}\delta x^{\beta}. \tag{A $+$}$$

$\delta\lambda^{\mu}$ is not a vector in general and $L^{\mu}{}_{\alpha\beta}$ not a tensor. Let the symmetric and antisymmetric parts of L be

$$\Gamma^{\mu}{}_{\alpha\beta}=\tfrac{1}{2}(L^{\mu}{}_{\alpha\beta}+L^{\mu}{}_{\beta\alpha}), \tag{3.1a}$$

$$\Omega^{\mu}{}_{\alpha\beta}=\tfrac{1}{2}(L^{\mu}{}_{\alpha\beta}-L^{\mu}{}_{\beta\alpha}), \tag{3.1b}$$

where $\Omega^{\mu}{}_{\alpha\beta}$, which is called the torsion, is a tensor and its contraction is a vector.

4. ABSOLUTE PARALLELISM

By integrating (A $\pm$) along a particular path, we may associate two λ's at the end of the path with the initial vector. However, the final vectors will depend on the path in general.

Consider (A $+$) first. In order that the final vector be independent of the path, the following equation must be integrable:

$$\lambda^{\mu}{}_{|\beta^{+}}\equiv\partial\lambda^{\mu}/\partial x^{\beta}+\lambda^{\alpha}L^{\mu}{}_{\alpha\beta}=0. \tag{4.1a}$$

The conditions of integrability are

$$L^{\alpha}{}_{\beta\gamma\delta}=0, \tag{4.1b}$$

where $L^{\alpha}{}_{\beta\gamma\delta}$ is the curvature tensor [see Eq. (7.1)]. We shall assume that the conditions (4.1b) are satisfied, and that it is therefore possible to associate with a vector at P another vector at any other point P' in a way which depends only on the points and is independent of the path joining them. (A $+$) is said to define an infinitesimal parallel displacement and vectors related by the integration of (4.1a) are said to be (+) parallel. The condition (4.1b) is the condition of (+)-absolute parallelism according to which finitely separated vectors may be said without ambiguity to be parallel.

When this is the case the manifold may be specified by an ennuple field as well as by a connection, for we now have

$$\partial\lambda^{\mu}{}_{i}/\partial x^{\beta}+\lambda^{\alpha}{}_{i}L^{\mu}{}_{\alpha\beta}=0.$$

By (2.3b),

$$L^{\mu}{}_{\alpha\beta}=-\lambda^{i}{}_{\alpha}\partial\lambda^{\mu}{}_{i}/\partial x^{\beta}=(\partial\lambda^{i}{}_{\alpha}/\partial x^{\beta})\lambda^{\mu}{}_{i}. \tag{4.2a}$$

The antisymmetric part of the connection is

$$\Omega^{\mu}{}_{\alpha\beta}=\tfrac{1}{2}\lambda^{\mu}{}_{i}(\partial_{\alpha}\lambda^{i}{}_{\beta}-\partial_{\beta}\lambda^{i}{}_{\alpha}) \tag{4.2b}$$

or

$$=\tfrac{1}{2}(\lambda^{i}{}_{\alpha}\partial_{\beta}-\lambda^{i}{}_{\beta}\partial_{\alpha})\lambda^{\mu}{}_{i}, \tag{4.2c}$$

where

$$\partial_{\beta}=\partial/\partial x^{\beta}.$$

The following invariants will also prove to be important:

$$\tfrac{1}{2}\omega^{i}{}_{jk}=\Omega^{\mu}{}_{\alpha\beta}\lambda^{i}{}_{\mu}\lambda^{\alpha}{}_{j}\lambda^{\beta}{}_{k}. \tag{4.3a}$$

The inverse relation expresses the torsion in terms of these invariants and the components of the ennuple

$$\Omega^{\mu}{}_{\alpha\beta}=\tfrac{1}{2}\omega^{i}{}_{jk}\lambda^{\mu}{}_{i}\lambda^{j}{}_{\alpha}\lambda^{k}{}_{\beta}. \tag{4.3b}$$

Expressed directly in terms of the ennuple, these invariants are

$$\omega^{i}{}_{jk}=\lambda^{\alpha}{}_{j}\lambda^{\beta}{}_{k}(\partial_{\alpha}\lambda^{i}{}_{\beta}-\partial_{\beta}\lambda^{i}{}_{\alpha}). \tag{4.3c}$$

5. DISPLACEMENT OPERATORS

Because of their relation to the momentum, displacement operators are of fundamental physical significance. They are now defined as follows:

$$X_{i}=\lambda^{\mu}{}_{i}\partial_{\mu}, \tag{5.1}$$

where X_{i} means the gradient operator in the direction of λ_{i}. The commutator of two displacement operators is

$$(X_{j},X_{k})=c^{i}{}_{jk}X_{i}, \tag{5.2a}$$

where

$$c^{i}{}_{jk}=-\omega^{i}{}_{jk}. \tag{5.2b}$$

The displacement operators do not commute unless the torsion vanishes. They do form a a group, however, if the structure constants $c^{i}{}_{jk}$ are constants which satisfy the Jacobi relations; we return to this point later.

6. TORSION

In the usual theory of gravitation the connection is symmetric and the torsion therefore vanishes. In the

kind of manifold now under consideration the torsion does not vanish, but for simplicity we restrict ourselves to spaces which are homogeneous with respect to it. Since absolute parallelism has been defined with respect to the (+) connection, uniform torsion will be defined with respect to the same connection; and the consequences for the (−) connection will be examined later. Therefore the condition of uniform torsion reads

$$\Omega^{\mu}{}_{\alpha\beta|\gamma^{+}}=0. \tag{6.1}$$

The corresponding condition

$$\Gamma^{\mu}{}_{\alpha\beta|\gamma^{+}}=0$$

is not a tensor equation in general, and (6.1) is therefore the simplest differential restriction of the connection. The left side of (6.1) means the (+)-covariant derivative with respect to $L^{\mu}{}_{\alpha\beta}$ and is an abbreviation for

$$\Omega^{\mu}{}_{\alpha\beta|\gamma^{+}}=\partial_{\gamma}\Omega^{\mu}{}_{\alpha\beta}+\Omega^{\sigma}{}_{\alpha\beta}L^{\mu}{}_{\sigma\gamma}-\Omega^{\mu}{}_{\sigma\beta}L^{\sigma}{}_{\alpha\gamma}-\Omega^{\mu}{}_{\alpha\sigma}L^{\sigma}{}_{\beta\gamma} \tag{6.2a}$$

$$=\Omega^{\mu}{}_{\alpha\beta,\gamma}+\Omega^{\sigma}{}_{\alpha\beta}\Omega^{\mu}{}_{\sigma\gamma}-\Omega^{\mu}{}_{\sigma\beta}\Omega^{\sigma}{}_{\alpha\gamma}-\Omega^{\mu}{}_{\alpha\sigma}\Omega^{\sigma}{}_{\beta\gamma}, \tag{6.2b}$$

where $\Omega^{\mu}{}_{\alpha\beta,\gamma}$ is the covariant derivative with respect to $\Gamma^{\mu}{}_{\alpha\beta}$. The procedure followed in imposing (6.1) is very special. In a more complete theory permitting variable torsion, Eq. (6.1) would represent only a special class of solutions.

7. CURVATURE TENSORS

In general one may define the curvature tensors[3]

$$L^{\mu}{}_{\alpha\beta\gamma}=(1-P_{\beta\gamma})\{\partial_{\beta}L^{\mu}{}_{\alpha\gamma}+L^{\sigma}{}_{\alpha\gamma}L^{\mu}{}_{\sigma\beta}\}, \tag{7.1}$$

$$B^{\mu}{}_{\alpha\beta\gamma}=(1-P_{\beta\gamma})\{\partial_{\beta}\Gamma^{\mu}{}_{\alpha\gamma}+\Gamma^{\sigma}{}_{\alpha\gamma}\Gamma^{\mu}{}_{\sigma\beta}\}, \tag{7.2}$$

$$\Omega^{\mu}{}_{\alpha\beta\gamma}=L^{\mu}{}_{\alpha\beta\gamma}-B^{\mu}{}_{\alpha\beta\gamma}$$
$$=(1-P_{\beta\gamma})\{\Omega^{\mu}{}_{\alpha\gamma|\beta^{+}}+\Omega^{\mu}{}_{\sigma\gamma}\Omega^{\sigma}{}_{\alpha\beta}+\Omega^{\mu}{}_{\sigma\alpha}\Omega^{\sigma}{}_{\beta\gamma}\}, \tag{7.3}$$

where $1-P_{\beta\gamma}$ indicates antisymmetrization with respect to $(\beta\gamma)$. Because of the $(\beta\gamma)$ antisymmetry there are only two contractions of B and Ω.

The manifold under consideration is defined by the conditions (4.1b) and (6.1). The $B^{\mu}{}_{\alpha\beta\gamma}$ satisfy the following symmetry conditions:

$$B^{\mu}{}_{\alpha\beta\gamma}+B^{\mu}{}_{\beta\gamma\alpha}+B^{\mu}{}_{\gamma\alpha\beta}=0. \tag{7.4a}$$

From (4.1b) and (7.4a) it follows that

$$\Omega^{\mu}{}_{\alpha\beta\gamma}+\Omega^{\mu}{}_{\beta\gamma\alpha}+\Omega^{\mu}{}_{\gamma\alpha\beta}=0. \tag{7.4b}$$

From (6.1) and (7.3)

$$\Omega^{\mu}{}_{\sigma\alpha}\Omega^{\sigma}{}_{\beta\gamma}+\Omega^{\mu}{}_{\sigma\beta}\Omega^{\sigma}{}_{\gamma\alpha}+\Omega^{\mu}{}_{\sigma\gamma}\Omega^{\sigma}{}_{\alpha\beta}\equiv 0, \tag{7.5a}$$

and

$$\Omega^{\mu}{}_{\alpha\beta\gamma}=\Omega^{\mu}{}_{\sigma\alpha}\Omega^{\sigma}{}_{\beta\gamma}, \tag{7.5b}$$

$$B^{\mu}{}_{\alpha\beta\gamma}=\Omega^{\mu}{}_{\alpha\sigma}\Omega^{\sigma}{}_{\beta\gamma}. \tag{7.6a}$$

There is only one nonvanishing contraction of $B^{\mu}{}_{\alpha\beta\gamma}$, namely,

$$B_{\alpha\beta}=\Omega^{\mu}{}_{\alpha\sigma}\Omega^{\sigma}{}_{\beta\gamma}, \tag{7.6b}$$

since $B_{\alpha\beta}$ is symmetric.

The preceding results depend on the assumptions of (+)-absolute parallelism (4.1b) and (+)-uniform torsion (6.1).

8. (−) CONNECTION

Denote the (±) connections by $L^{\mu}{}_{\alpha\beta}(\pm)$. Then by definition

$$L^{\mu}{}_{\alpha\beta}(-)=L^{\mu}{}_{\beta\alpha}(+), \tag{8.1a}$$

$$\Gamma^{\mu}{}_{\alpha\beta}(-)=\Gamma^{\mu}{}_{\alpha\beta}(1), \tag{8.1b}$$

$$\Omega^{\mu}{}_{\alpha\beta}(-)=-\Omega^{\mu}{}_{\alpha\beta}(+). \tag{8.1c}$$

By definition of (+) and (−) derivatives (6.2b),

$$\tfrac{1}{2}(\Omega^{\mu}{}_{\alpha\gamma|\beta^{+}}-\Omega^{\mu}{}_{\alpha\gamma|\beta^{-}})=\Omega^{\mu}{}_{\alpha\gamma|\beta^{+}}-\Omega^{\mu}{}_{\alpha\gamma,\beta}$$
$$=\Omega^{\sigma}{}_{\alpha\gamma}\Omega^{\mu}{}_{\sigma\beta}+\Omega^{\sigma}{}_{\beta\alpha}\Omega^{\mu}{}_{\sigma\gamma}+\Omega^{\sigma}{}_{\gamma\beta}\Omega^{\mu}{}_{\sigma\alpha}, \tag{8.2}$$

and

$$\Omega^{\mu}{}_{\alpha\beta|\gamma^{+}}=\Omega^{\mu}{}_{\alpha\beta|\gamma^{-}}=\Omega^{\mu}{}_{\alpha\beta,\gamma}=0 \tag{8.2a}$$

according to (7.5a).

Hence the given space has uniform torsion with respect to both (+) and (−) connections.

Since B depends only on Γ, it is unchanged by the transformation (8.1)

$$B^{\mu}{}_{\alpha\beta\gamma}(+)=B^{\mu}{}_{\alpha\beta\gamma}(-). \tag{8.3}$$

By Eq. (7.3)

$$\Omega^{\mu}{}_{\alpha\beta\gamma}(+)-\Omega^{\mu}{}_{\alpha\beta\gamma}(-)=(1-P_{\beta\gamma})[\Omega^{\mu}{}_{\alpha\gamma|\beta^{+}}+\Omega^{\mu}{}_{\alpha\gamma|\beta^{-}}].$$

By (8.2a)

$$\Omega^{\mu}{}_{\alpha\beta\gamma}(+)=\Omega^{\mu}{}_{\alpha\beta\gamma}(-). \tag{8.4}$$

By (8.3) and (8.4)

$$L^{\mu}{}_{\alpha\beta\gamma}(-)=L^{\mu}{}_{\alpha\beta\gamma}(+)=0.$$

Hence, if the (+) curvature vanishes, the (−) curvature does also. Therefore the given space admits absolute parallelism with respect to both (+) and (−) connections.

9. DISPLACEMENT GROUP

The displacement operators X_k are the symbols of a continuous group provided that the coefficients $c^{i}{}_{jk}$ in Eq. (5.2) satisfy the following conditions:

$$c^{i}{}_{jk}=-c^{i}{}_{kj}, \tag{9.1}$$

$$c^{p}{}_{ij}c^{s}{}_{kp}+c^{p}{}_{jk}c^{s}{}_{ip}+c^{p}{}_{ki}c^{s}{}_{jp}=0, \tag{9.2}$$

$$\partial c^{i}{}_{jk}/\partial x^{\gamma}=0. \tag{9.3}$$

By definitions (4.3a) and (5.2b) $c^{i}{}_{jk}$ is antisymmetric in k and j because $\Omega^{\mu}{}_{\alpha\beta}$ is antisymmetric in α and β. The Jacobi relations (9.2) are satisfied in virtue of (7.5a). That the $c^{i}{}_{jk}$ are constants may be seen as follows. We have

$$\omega^{i}{}_{jk|\gamma^{+}}=0$$

[3] L. P. Eisenhart, *Non-Riemannian Geometry* (American Mathematical Society, New York, 1927).

by (4.1a), (4.3a), and (8.2). But $\omega^i{}_{jk}$ is an invariant. Hence

$$\partial\omega^i{}_{jk}/\partial x^\gamma=\omega^i{}_{jk|\gamma}{}^+=0.$$

Therefore, the X_k do define a group.

The Cartan group metric may be defined in the usual way,

$$\bar{g}_{mn}=c^r{}_{ms}c^s{}_{nr}. \tag{9.4}$$

Cartan's criterion that the group be semisimple is that

$$\bar{g}=|\bar{g}_{mn}|\neq 0. \tag{9.5}$$

In the Appendix it is shown that

$$\bar{g}=0 \tag{9.6}$$

for a four-parameter group. Hence the group of displacement operators is not semisimple.

From Eqs. (4.3b), (7.6b), and (9.4), we have

$$B_{\alpha\beta}=\tfrac{1}{4}\bar{g}_{mn}\lambda^m{}_\alpha\lambda^n{}_\beta. \tag{9.7}$$

From (9.6) and (9.7) it follows that the determinant of $B_{\alpha\beta}$ vanishes:

$$|B_{\alpha\beta}|=(\tfrac{1}{4})^4\lambda^2\bar{g}=0. \tag{9.8}$$

10. METRIC

The manifold so far discussed has no metric properties. To connect with either the quantum theory of fields or the classical theory of gravitation a metric must appear at some stage. In the nonsymmetric theories of Einstein and Schrödinger, the metric, like the connection, is asymmetric; but the asymmetry in the metric is there interpreted in terms of an electromagnetic field, which vanishes in our neutral approximation. We are here going to assume that the metric is symmetric:

$$g_{\alpha\beta}=g_{\beta\alpha}. \tag{10.1}$$

In addition it will be assumed that the space is homogeneous with respect to the metric in the same sense that it is with respect to the torsion [Eq. (8.2a)], that is,

$$g_{\alpha\beta|\gamma}{}^+=g_{\alpha\beta|\gamma}{}^-=0 \tag{10.2}$$

or

$$(\partial g_{\alpha\beta}/\partial x^\gamma)+g_{\alpha\sigma}L^\sigma{}_{\beta\gamma}+g_{\sigma\beta}L^\sigma{}_{\alpha\gamma}=0 \tag{10.2a}$$

and

$$(\partial g_{\alpha\beta}/\partial x^\gamma+g_{\alpha\sigma}L^\sigma{}_{\gamma\beta}+g_{\sigma\beta}L^\sigma{}_{\gamma\alpha}=0. \tag{10.2b}$$

If we add (10.2a) and (10.2b) we obtain

$$(\partial g_{\alpha\beta}/\partial x^\gamma)+g_{\alpha\sigma}\Gamma^\sigma{}_{\gamma\beta}+g_{\sigma\beta}\Gamma^\sigma{}_{\gamma\alpha}=0 \tag{10.3}$$

or

$$g_{\alpha\beta,\gamma}=0. \tag{10.3a}$$

Hence,

$$g_{\alpha\beta|\gamma}{}^+=g_{\alpha\beta|\gamma}{}^-=g_{\alpha\beta,\gamma}=0 \tag{10.4}$$

in correspondence with (8.2a). The geometric meaning of (10.4) is that angle is unaltered in (+)-, (−)-, and (0)-parallel displacement. The postulate (10.2) is especially important since it permits the introduction of two spinor fields everywhere.

Our other assumption about the metric (10.1) permits us to solve (10.3) simply, with the result

$$\Gamma^\mu{}_{\alpha\beta}=\left\{\begin{matrix}\mu\\ \alpha\beta\end{matrix}\right\}, \tag{10.5}$$

where { } is the Christoffel symbol. It is therefore no longer necessary to distinguish between covariant derivatives with respect to Γ and { }.

If (10.2b) is subtracted from (10.2a) the result is

$$g_{\alpha\sigma}\Omega^\sigma{}_{\beta\gamma}+g_{\sigma\beta}\Omega^\sigma{}_{\alpha\gamma}=0 \tag{10.6}$$

or

$$\Omega_{\alpha\beta\gamma}=-\Omega_{\beta\alpha\gamma}, \tag{10.7}$$

where

$$\Omega_{\alpha\beta\gamma}=g_{\alpha\sigma}\Omega^\sigma{}_{\beta\gamma}. \tag{10.7a}$$

The torsion now determines the Ricci tensor by (10.5) and (7.6b) according to the equation

$$R_{\alpha\beta}=B_{\alpha\beta}=\Omega^\lambda{}_{\alpha\mu}\Omega^\mu{}_{\beta\lambda}. \tag{10.8}$$

It also follows that

$$R_{\alpha\beta,\gamma}=0. \tag{10.9}$$

In addition there is still the algebraic restriction (9.8) or

$$|R_{\alpha\beta}|=0. \tag{10.10}$$

Finally,

$$\Omega^\alpha{}_{\alpha\beta}=g^{\alpha\gamma}\Omega_{\gamma\alpha\beta}=0. \tag{10.11}$$

The independent tensor fields defined in the space are $\delta^\mu{}_\alpha$, $\epsilon_{\alpha\beta\gamma\delta}$, $g_{\alpha\beta}$, and $\Omega^\mu{}_{\alpha\beta}$. From these may be derived $\delta^\alpha{}_\mu\Omega^\mu{}_{\alpha\beta}=\Omega^\alpha{}_{\alpha\beta}$, which vanishes by (10.11); $g_{\alpha\mu}\Omega^\mu{}_{\beta\gamma}=\Omega_{\alpha\beta\gamma}$, which is completely antisymmetric; and

$$\varphi^\mu=(1/3!)\epsilon^{\mu\alpha\beta\gamma}\Omega_{\alpha\beta\gamma}, \tag{10.12}$$

which is an axial vector satisfying the equation

$$\varphi^\mu{}_{,\lambda}=0. \tag{10.13}$$

The general solution of (10.9) is, therefore,

$$R_{\alpha\beta}=ag_{\alpha\beta}+bg\varphi_\alpha\varphi_\beta, \tag{10.14}$$

subject to the algebraic condition (10.10).

There are 10 equations to be solved for the 10 $g_{\alpha\beta}$ in terms of the φ_μ. The special type of space under consideration is therefore entirely characterized by the single axial vector field φ_μ. The same conclusion may be reached by writing $R_{\alpha\beta}$ in terms of the 24 $\Omega^\mu{}_{\alpha\beta}$ which are subject to the 16 Jacobi relations (7.5a) and the four conditions (10.11). It may also be remarked that in general one cannot begin by assigning the metric, because the 10 equations for φ_μ may then be inconsistent.

11. PARALLEL DISPLACEMENT OF ORTHOGONAL AXES

Three kinds of parallel displacement have been defined, namely,

$$\delta_0\lambda^\mu{}_i=-\Gamma^\mu{}_{\alpha\beta}\lambda^\alpha{}_i\delta x^\beta, \tag{11.0}$$

$$\delta_\pm\lambda^\mu{}_i=\delta_0\lambda^\mu{}_i\mp\Omega^\mu{}_{\alpha\beta}\lambda^\alpha{}_i\delta x^\beta, \tag{11.$\pm$}$$

$(\pm)$ are integrable and define two ennuple fields everywhere. On the other hand, the condition for the integrability of (0) is

$$R^{\mu}{}_{\alpha\beta\gamma}=0,$$

which in general is not satisfied.

According to (10.4) angle is preserved by all three kinds of parallel displacement. Hence, an orthogonal ennuple, defined at any point P, remains orthogonal after displacement to any other P' according to each of the three equations (0), (+), and (−). However, if the displacement is made according to (0), the resulting ennuple at P' depends on the path.

Let $\lambda^{\mu}{}_{i}$ be a set of linearly independent vectors at P. We may in the usual way construct from $\lambda^{\mu}{}_{i}$ and the given metric at P an orthonormal set $a^{\mu}{}_{i}$, say. Then

$$g_{\mu\nu}a^{\mu}(i)a^{\nu}(j)=\delta(i,j) \tag{11.1a}$$

or

$$a^{\mu}(i)a_{\mu}(j)=\delta(i,j), \tag{11.1b}$$

where

$$a_{\mu}(j)=g_{\mu\nu}a^{\nu}(j). \tag{11.1c}$$

According to (11.1b), $a^{\mu}(i)$ and $a_{\mu}(i)$ are reciprocal matrices. They are also reciprocal in the reverse direction, i.e.,

$$\textstyle\sum_{k} a^{\mu}(k)a_{\lambda}(k)=\delta^{\mu}{}_{\lambda} \tag{11.2a}$$

or

$$\textstyle\sum_{k} a^{\mu}(k)a^{\lambda}(k)=g^{\mu\lambda}. \tag{11.2b}$$

The previously introduced vectors $\lambda^{\mu}{}_{i}$ and $\lambda^{i}{}_{\mu}$ now are replaced by $a^{\mu}(i)$ and $a_{\mu}(i)$, which may be regarded as contravariant and covariant components of the same vector.

The orthogonal ennuple at P displaced according to (0), $(\pm)$ leads to three others at P', namely,

$$a^{\mu}{}_{0}(i,P') \quad \text{and} \quad a^{\mu}{}_{\pm}(i,P').$$

These three orthogonal ennuples at P' are then related by rotations. The formulas will be somewhat simplified if the metric at P' is chosen Cartesian so that covariant and contravariant components are equal. Then one has for infinitesimal rotations, corresponding to infinitesimal displacements

$$\delta^{*}{}_{\pm}a(i)\equiv a_{\pm}(i,P')-a_{0}(i,P')=\textstyle\sum_{k} a(k,P')\delta_{\pm}O(k,i), \tag{11.3}$$

where $a(i)$ means the vector with components $a^{\mu}(i)=a_{\mu}(i)$. According to the formula for parallel displacement (11. $\pm$), we have

$$\textstyle\sum_{k} a^{\mu}(k)\delta_{\pm}O(k,i)=\mp\Omega^{\mu}{}_{\alpha\beta}a^{\alpha}(i)\delta x^{\beta},$$

and therefore

$$\delta_{\pm}O(k,i)=\mp\Omega^{\mu}{}_{\alpha\beta}a^{\alpha}(i)a_{\mu}(k)\delta x^{\beta}. \tag{11.4a}$$

Since the metric at P' is assumed to be Cartesian, the preceding equation may be written in the following way:

$$\delta_{\pm}O(k,i)=\mp\sum_{\mu,\alpha,\beta}\Omega_{\mu\alpha\beta}a_{\mu}(k)a_{\alpha}(i)\delta x^{\beta}. \tag{11.4b}$$

Since $\Omega_{\mu\alpha\beta}=-\Omega_{\alpha\mu\beta}$, we also have

$$\delta_{\pm}O(k,i)=-\delta_{\pm}O(i,k),$$

which is necessary for an infinitesimal rotation.

Equation (11. $\pm$) becomes

$$\delta_{\pm}a(i)=\delta_{0}a(i)+\textstyle\sum_{k} a(k)\delta_{\pm}O(k,i). \tag{11.5a}$$

The integral form of Eq. (11.5a) may be written as follows:

$$\Delta_{\pm}a(i)=\Delta_{0}a(i)+\textstyle\sum_{k} a(k)O_{\pm}(k,i), \tag{11.5b}$$

where $\Delta_{0}a(i)$ means a path-dependent parallel displacement and $O_{\pm}(k,i)$ means a rotation at the final point of such magnitude that $\Delta_{\pm}a(i)$ is path independent. Here, $O_{+}=-O_{-}$.

In the way just discussed one may associate with the given space two orthogonal ennuple fields, which are (+) and (−) parallel. These define at every point the connection according to (4.2a) and the metric according to (11.2b). We may finally check that all covariant derivatives of $g_{\alpha\beta}$ vanish, when they are defined according to (11.2b):

$$\begin{aligned} g_{\alpha\beta|\gamma}&=\textstyle\sum_{i}[a_{\alpha|\gamma}(i)a_{\beta}(i)+a_{\alpha}(i)a_{\beta|\gamma}(i)]\\ &=0, \end{aligned}$$

since

$$a_{\alpha|\gamma}(i)=0.$$

We also check

$$\begin{aligned} g_{\alpha\beta,\lambda}&=\textstyle\sum_{i}[a_{\alpha,\lambda}(i)a_{\beta}(i)+a_{\alpha}(i)a_{\beta,\lambda}(i)]\\ &=\pm\textstyle\sum_{i}[\Omega^{\sigma}{}_{\alpha\lambda}a_{\sigma}(i)a_{\beta}(i)+\Omega^{\sigma}{}_{\beta\lambda}a_{\sigma}(i)a_{\alpha}(i)]\\ &=\pm(\Omega^{\sigma}{}_{\alpha\lambda}g_{\sigma\beta}+\Omega^{\sigma}{}_{\beta\lambda}g_{\sigma\alpha})=0. \end{aligned}$$

12. SPINORS

To introduce spinors it is necessary to define an orthogonal ennuple field everywhere. In general the field so defined is altogether unrelated to the structure of the space. When the geometry permits absolute parallelism, however, as here, the situation is much simpler since in that case the connection is completely specified by a parallel ennuple field [Eq. (4.2a)].

Except for the fundamental simplification just noted, the introduction of spinors proceeds in the usual way. Corresponding to the (+) and (−) parallelism, however, there are two kinds of ennuple field $a^{\mu}{}_{\pm}(i)$ and two kinds of spinor $\psi_{\pm}$. Define

$$\gamma^{\mu}{}_{\pm}=\textstyle\sum a^{\mu}{}_{\pm}(i)\gamma_{\pm}(i), \tag{12.1}$$

where the $\gamma_{\pm}(i)$ are the usual constant Dirac matrices

$$\{\gamma(i),\gamma(k)\}_{+}=2\delta(i,k). \tag{12.2a}$$

[The distinction $(\pm)$ will not be indicated explicitly in general.] Then

$$\{\gamma^{\mu},\gamma^{\nu}\}_{+}=2\textstyle\sum_{i} a^{\mu}(i)a^{\nu}(i)=2g^{\mu\nu}. \tag{12.2b}$$

The γ^{μ} are vectors under coordinate transformations. Next, define the fields $\psi(x)$ and $\bar{\psi}(x)$, scalars under

coordinate transformations and spinors under ennuple rotations. Under the rotation $\delta O(i,k)$,

$$\delta\psi=[\tfrac{1}{2}\delta O(i,k)\gamma(i)\gamma(k)]\psi, \tag{12.3a}$$

$$\delta\bar{\psi}=-\bar{\psi}[\tfrac{1}{2}\delta O(i,k)\gamma(i)\gamma(k)]. \tag{12.3b}$$

13. PARALLEL DISPLACEMENT OF SPINORS

Define the infinitesimal parallel transfer of a spinor by the equations

$$\delta\psi=-\Delta\psi, \tag{13.1a}$$

$$\delta\bar{\psi}=\bar{\psi}\Delta, \tag{13.1b}$$

where $\Delta=\Delta_\mu\delta x^\mu$ is the spin representation of a rotation associated with δx^μ. This rotation is usually determined by comparing the orientation of an ennuple moved by (0) transfer with that of the reference ennuple. Here the reference fields are defined by the structure of the space, and they are of two types, corresponding to (+) and (−) parallelism. Consequently, there are also two ways in which a spinor may be parallel displaced; these will be distinguished by $\delta_\pm$ and $\Delta_\pm$. Equation (13.1) will now be completed by the specification of Δ as follows:

$$\Delta_\pm=\mp\tfrac{1}{4}\Omega_{\mu\alpha\beta}\gamma^\mu\gamma^\alpha\delta x^\beta. \tag{13.1c}$$

The Δ defined by (13.1) is, by (12.3), the spin representation of the rotation (11.4a). Corresponding to (13.1) there must be similar relations for the $\gamma^\mu{}_\pm$, which are defined by (12.1). Since the law of parallel transfer for the $a^\mu{}_\pm(i)$ is already fixed it is only necessary to assign the corresponding law for either $\gamma^\mu{}_\pm$, or $\gamma_\pm(i)$. It is possible to adopt the following law:

$$\delta\gamma^\mu{}_\pm=-(\Gamma^\mu{}_{\alpha\beta}\pm\Omega^\mu{}_{\alpha\beta})\gamma^\alpha{}_\pm\delta x^\beta. \tag{13.2}$$

Since both (+) and (−) derivatives of $g_{\alpha\beta}$ vanish, according to (10.4), Eq. (13.2) preserves the commutation rules as required:

$$\tfrac{1}{2}\delta(\gamma^\mu\gamma^\nu+\gamma^\nu\gamma^\mu)=\delta g^{\mu\nu}. \tag{13.3}$$

From (13.1) it follows at once that

$$\delta(\bar{\psi}\psi)=0. \tag{13.4}$$

The corresponding equation for a vector may be checked as follows: Consider

$$A^\lambda=\bar{\psi}_\pm\gamma^\lambda{}_\pm\psi_\pm, \tag{13.5}$$

then

$$\begin{aligned}\delta A^\lambda&=\bar{\psi}(\Delta\gamma^\lambda-\gamma^\lambda\Delta)\psi+\bar{\psi}(\delta\gamma^\lambda)\psi\\&=\bar{\psi}(\Delta,\gamma^\lambda)\psi-(\Gamma^\lambda{}_{\alpha\beta}\pm\Omega^\lambda{}_{\alpha\beta})A^\alpha\delta x^\beta.\end{aligned}$$

By (13.1c),

$$(\Delta,\gamma^\lambda)=\mp\tfrac{1}{4}\Omega_{\mu\alpha\beta}(\gamma^\mu\gamma^\alpha,\gamma^\lambda)\delta x^\beta.$$

But

$$(\gamma^\mu\gamma^\alpha,\gamma^\lambda)=2(\gamma^\mu g^{\alpha\lambda}-\gamma^\alpha g^{\mu\lambda}).$$

Therefore,

$$(\Delta,\gamma^\lambda)=\pm\Omega^\lambda{}_{\mu\beta}\gamma^\mu\delta x^\beta$$

and

$$\delta A^\lambda=-\Gamma^\lambda{}_{\mu\beta}A^\mu\delta x^\beta. \tag{13.6}$$

Therefore, if the displacement of a spinor is defined by (13.1), then the resulting equations for the vector $(\bar{\psi}_\pm\gamma^\lambda{}_\pm\psi_\pm)$ come out correctly for (0) transfer, if (13.2) is adopted for $\delta\gamma^\mu{}_\pm$.

The corresponding covariant derivatives are

$$\psi_{|\beta}=\partial_\beta\psi\pm\tfrac{1}{4}\Omega_{\mu\alpha\beta}\gamma^\mu\gamma^\alpha\psi, \tag{13.7a}$$

$$\bar{\psi}_{|\beta}=\partial_\beta\bar{\psi}\mp\bar{\psi}(\tfrac{1}{4}\Omega_{\mu\alpha\beta}\gamma^\mu\gamma^\alpha). \tag{13.7b}$$

The two signs correspond to the (+)- and (−)-parallel ennuple fields. The invariant derivative is

$$\begin{aligned}\gamma^\beta\psi_{|\beta}&=\gamma^\beta\partial_\beta\psi\pm\tfrac{1}{4}\Omega_{\mu\alpha\beta}\gamma^\beta\gamma^\mu\gamma^\alpha\psi\\&=(\gamma^\beta\partial_\beta\pm\tfrac{1}{4}\Omega_{\mu\alpha\beta}\gamma^\mu\gamma^\alpha\gamma^\beta)\psi.\end{aligned} \tag{13.8}$$

Define

$$\gamma^5\equiv(1/4!)\epsilon_{\lambda\mu\rho\tau}\gamma^\lambda\gamma^\mu\gamma^\rho\gamma^\tau, \tag{13.9a}$$

$$\gamma^{\tau5}\equiv\tfrac{1}{2}(\gamma^\tau,\gamma^5), \tag{13.9b}$$

then

$$\gamma^{\tau5}=(1/3!)g^{\tau\lambda}\epsilon_{\lambda\mu\rho\sigma}\gamma^\mu\gamma^\rho\gamma^\sigma, \tag{13.9c}$$

and (13.8) may be written

$$\gamma^\mu\psi_{|\mu}=[\gamma^\mu\partial_\mu\pm\tfrac{3}{2}\gamma^{\mu5}\varphi_\mu]\psi, \tag{13.10}$$

where φ_μ is given by (10.12). The formal analog of the Dirac equation is

$$[\gamma^\mu\partial_\mu\pm\tfrac{3}{2}\gamma^{\mu5}\varphi_\mu]\psi_\pm=m\psi_\pm, \tag{13.11a}$$

and the corresponding massless equation is

$$[\gamma^\mu\partial_\mu\pm\tfrac{3}{2}\gamma^{\mu5}\varphi_\mu]\psi_\pm=0. \tag{13.11b}$$

As emphasized before, the ennuple field ordinarily introduced to define the parallel displacement of spinors is unrelated to the affine connection; instead, it defines a spin connection. Consequently the usual generalization of Dirac's equation depends only partly on the affine connection. Here, on the other hand, the interaction term is determined entirely, and simply, by the torsion.

14. SUMMARY

The space under consideration has been defined by the following properties:

(a) distant parallelism: $L^\mu{}_{\alpha\beta\gamma}(\pm)=0$;
(b) uniform torsion: $\Omega^\mu{}_{\alpha\beta|\gamma^+}=\Omega^\mu{}_{\alpha\beta|\gamma^-}=0$;
(c) symmetric metric: $g_{\alpha\beta}=g_{\beta\alpha}$;
(d) distant congruence: $g_{\alpha\beta|\gamma^+}=g_{\alpha\beta|\gamma^-}=0$.

By (d) is meant the congruence at different points of two configurations which are related by parallel transfer. There are two kinds of distant parallelism and congruence, and they are formally related by a transposition of the connection. In any physical interpretation transposition symmetry must play a fundamental role.

15. COSMOLOGICAL MODEL

In this section we consider the simplest nontrivial example of a spacetime which satisfies the general con-

ditions which have been set. This model was first studied by Robertson[4] in a different connection.

Consider the line element

$$ds^2=dt^2-F(r,t)(dx^2+dy^2+dz^2), \tag{15.1}$$

where

$$F(r,t)=[1+\tfrac{1}{4}K(t)r^2]^{-2}=G(t)[1+\tfrac{1}{4}K_0r^2]^{-2} \tag{15.1a}$$

and

$$r^2=x^2+y^2+z^2. \tag{15.1b}$$

Here t means the time, and surfaces of constant t define hypersurfaces of constant curvature $K(t)$. These three-dimensional sections are isotropic and homogeneous.

We shall further specialize by assuming that $K(t)$ is independent of the time. This specialization corresponds exactly to the static Einstein model when the connection is symmetric; however, since we are postulating an asymmetric connection, it is necessary also to specify the torsion, and in such a way as to be consistent with the given metric.

We shall see that the following eight functions are consistent and define a continuum with the metrical properties of an Einstein space.

$$\begin{aligned} g_{ii}&=-F(r), \quad \Omega_{ij4}=0 \\ g_{44}&=1, \qquad \Omega_{ijk}=\omega\epsilon_{ijk}. \end{aligned} \tag{15.2}$$

The Christoffel symbols are

$$\Gamma^4{}_{\alpha\beta}=\Gamma^\alpha{}_{4\beta}=0,$$

$$\Gamma^i{}_{jk}=-\tfrac{1}{2}KF^{\frac{1}{2}}(\delta_{ik}x^j+\delta_{ij}x^k-\delta_{jk}x^i), \quad i,\ j,\ k\neq 4. \tag{15.3a}$$

The asymmetric part of the connection is

$$\Omega^4{}_{ij}=\Omega^i{}_{4j}=0, \quad \Omega^i{}_{jk}=-F^{-1}\Omega_{ijk}=-F^{-1}\omega\epsilon_{ijk}. \tag{15.3b}$$

The Ricci tensor calculated directly from the metric is

$$R_{ij}=-2KF\delta_{ij}, \quad i,\ j\neq 4 \tag{15.4a}$$

$$R_{\alpha 4}=0, \quad \alpha=1,\cdots,4. \tag{15.4b}$$

The following relations, which must hold according to (10.9) and (10.10), may now be checked:

$$R_{\alpha\beta,\gamma}=0, \tag{15.5}$$

$$|R_{\alpha\beta}|=0. \tag{15.6}$$

On the other hand, the Ricci tensor is directly related to $\Omega^\alpha{}_{\beta\gamma}$

$$\begin{aligned} R_{ij}&=\Omega^m{}_{in}\Omega^n{}_{jm} \\ &=-(\omega/F)^2\Sigma_{m,n}\epsilon_{imn}\epsilon_{jmn}, \\ &=-2(\omega/F)^2\delta_{ij}, \end{aligned} \tag{15.7a}$$

and

$$R_{\alpha 4}=0, \quad \alpha=1,\cdots,4. \tag{15.7b}$$

One sees that (15.7) and (15.4) are consistent if

$$\omega^2=KF^3 \tag{15.8a}$$

or

$$\omega^2=-Kg. \tag{15.8b}$$

[4] H. P. Robertson, Ann. Math. 33, 496 (1932).

By (15.3a) and (15.3b), we have

$$\begin{aligned} &L^4{}_{\alpha\beta}=L^\alpha{}_{4\beta}=L^\alpha{}_{\beta 4}=0, \\ &L^i{}_{jk}=-(KF)^{\frac{1}{2}}\epsilon_{ijk}-(K/2)F^{\frac{1}{2}}(\delta_{ik}x^j+\delta_{ij}x^k-\delta_{jk}x^i). \end{aligned}$$

The fundamental $(\pm)$-ennuple fields satisfy the equations

$$\lambda^\mu{}_{i,\alpha}\pm\Omega^\mu{}_{\beta\alpha}\lambda^\beta{}_i=0. \tag{15.9}$$

If we choose $\lambda^\mu{}_4=\delta^\mu{}_4$, these equations are satisfied in a trivial way. In order that the other three vectors be orthogonal to $\lambda^\mu{}_4$ choose

$$\lambda^4{}_i=0.$$

Then the Eqs. (15.9) simplify:

$$\partial\lambda^\mu{}_i/\partial x^4=0 \tag{15.9a}$$

and

$$\lambda^\mu{}_{i,\alpha}\pm\Omega^\mu{}_{\beta\alpha}\lambda^\beta{}_i=0, \quad i,\alpha,\beta,\mu=1,2,3. \tag{15.9b}$$

The set (15.9b) has to be solved for three orthogonal fields $\lambda^\mu{}_i$.

16. ENNUPLE FIELD AND THE DISPLACEMENT GROUP

The solutions of (15.9b) may be obtained by the following geometric method. The equation of the hypersphere (t=const) may be written in a simpler form by introducing the four variables y^α such that

$$y^k=F(r)x^k, \tag{16.1a}$$

$$y^4=K^{-\frac{1}{2}}(2F^{\frac{1}{2}}-1)=K^{-\frac{1}{2}}(1-\tfrac{1}{4}Kr^2)/(1+\tfrac{1}{4}Kr^2). \tag{16.1b}$$

Then

$$\sum_1^4(y^\alpha)^2=1/K. \tag{16.1c}$$

This sphere is invariant under the four-dimensional orthogonal group with the six generators

$$Y_{\alpha\beta}=y^\alpha\partial/\partial y^\beta-y^\beta\partial/\partial y^\alpha. \tag{16.2}$$

The group O_4 has two invariant subgroups with generators

$$\text{(a)} \quad Y_{34}+Y_{12} \quad Y_{14}+Y_{23} \quad Y_{24}+Y_{31}, \tag{16.3a}$$

$$\text{(b)} \quad Y_{34}-Y_{12} \quad Y_{14}-Y_{23} \quad Y_{24}-Y_{31}. \tag{16.3b}$$

Each of these two subgroups is a three-dimensional rotation group.

The symbol Y_{12} corresponds to a rotation in the 12 plane; Y_{34} corresponds to a rotation in the 34 plane, and therefore to a translation in the 3 direction. The two invariant subgroups (a) and (b) therefore describe screw motions in the direction of the three spatial axes; (a) and (b) differ in the helicity of the screw displacement. We shall now see that these two screw displacement groups (a) and (b) are exactly the two displacement groups, based on the (+) and (−) connection.

In terms of the original coordinates (x,y,z,t), which may be more directly interpreted, the generators of the screw motions become[4]

$$X_i=\lambda^\mu{}_i\partial/\partial x^\mu, \quad i,\mu=1,2,3 \tag{16.4a}$$

where

$$\lambda^\mu{}_i(\pm)=K^{\frac{1}{2}}(\eta^\mu{}_i\pm\xi^\mu{}_i), \tag{16.4b}$$

$$\xi^\mu{}_i=\epsilon_{i\mu\nu}x^\nu, \tag{16.4c}$$

$$\eta^\mu{}_i=K^{-\frac{1}{2}}[\delta^\mu{}_i(2-F^{-\frac{1}{2}})+\tfrac{1}{2}Kx^\mu x^i], \tag{16.4d}$$

and

$$\textstyle\sum_i \lambda^\alpha{}_i\lambda^\beta{}_i=F^{-1}\delta^{\alpha\beta}. \tag{16.4e}$$

It may now be verified directly by substitution that the ennuples (16.4b) satisfy the differential equations (15.4) and therefore may be obtained by (+)- and (−)-parallel displacement.

The ennuple invariants are

$$\omega^i{}_{jk}=2\Omega^\alpha{}_{\beta\gamma}\lambda^i{}_\alpha\lambda^\beta{}_j\lambda^\gamma{}_k. \tag{16.5}$$

By (16.4e),

$$\lambda^i{}_\alpha=F\lambda^\alpha{}_i. \tag{16.6}$$

Hence by (15.3)

$$\omega^i{}_{jk}=-2\omega\hat{\lambda}\epsilon_{ijk},$$

where

$$\hat{\lambda}=\epsilon_{rst}\lambda^r{}_1\lambda^s{}_2\lambda^t{}_3, \quad r,s,t=1,2,3, \tag{16.7a}$$

but

$$\hat{\lambda}(-g)^{\frac{1}{2}}=1, \tag{16.7b}$$

so that

$$\omega^i{}_{jk}=2K^{\frac{1}{2}}\epsilon_{ijk}. \tag{16.8}$$

The commutation rules for the two displacement groups are now

$$(X_4,X_k)=0, \tag{16.9a}$$

$$(X_k,X_l)=c^m{}_{kl}X_m, \quad k,l\neq 4 \tag{16.9b}$$

where

$$c^m{}_{kl}=\pm\omega^m{}_{kl}=\pm 2K^{\frac{1}{2}}\epsilon_{klm}.$$

The commutator of two spacelike displacements is proportional to $K^{\frac{1}{2}}$ or inversely proportional to the radius of curvature. The usual commuting momentum operators may be regarded as approximations to the X_k from which they differ by a small rotation, which is again inversely proportional to the radius of the space.

The group metric is

$$\begin{aligned}\bar{g}_{mn}&=c^p{}_{mq}c^q{}_{np}\\&=-8K\delta_{mn}, \quad m,n=1,2,3\end{aligned} \tag{16.10}$$

$$\bar{g}_{\alpha 4}=0, \quad \alpha=1,\cdots,4$$

and by (9.7)

$$\begin{aligned}R_{kl}&=\tfrac{1}{4}\bar{g}_{mn}\lambda^m{}_\kappa\lambda^n{}_l\\&=-2K\textstyle\sum_m\lambda^m{}_k\lambda^m{}_l\\&=-2KF\delta_{kl}, \quad k,l=1,2,3\\R_{\alpha 4}&=0 \qquad\qquad \alpha=1,\cdots,4\end{aligned}$$

by (16.4e) and (16.6). This checks with (15.4).

17. GROUP OF THE SPACE

The three-dimensional sections ($t=$const) are carried into themselves by the G_6 which has already been discussed. However, as was already pointed out by Robertson,[4] the three-dimensional sections do not admit improper rotations unless $K=0$; but if $K=0$, the model is empty since then the curvature and torsion both vanish. Therefore the proposed model of spacetime is not reflection invariant. On the other hand, the complete four-dimensional space does admit time reversal.

The reflection $x_\kappa\rightarrow -x_k$ induces the transformation $n\pm\xi\rightarrow\eta\mp\xi$. If this transformation leaves invariant the structure of both ennuple fields, and therefore carries the space into itself, it must be equivalent to an ennuple transformation with constant coefficients. This is clearly not the case, however, since ξ itself depends on position. A reflection therefore changes the structure of the space. Our tensor equations are to be understood as follows: the space is defined in a special coordinate system by Eq. (15.2) and in all other coordinate systems connected with the special one by a transformation with positive Jacobian.

18. GROUP OF THE SPINOR EQUATION

Equation (13.11) is generally covariant by derivation. Both terms of the operator are separately invariant with respect to general coordinate transformations. The spinor is invariant with respect to general coordinate transformations and therefore so is the equation. With respect to ennuple rotations, ψ is a spinor, and covariance may be shown in the usual way. (The permitted rotations must be independent of position to preserve the parallelism of the ennuple field.)

Although the differential equation is formally invariant under coordinate reflections, these transformations are not permitted since they do not take the space into itself. Reflection now corresponds not to an alternative description of the same physical space, but to a different physical space. The group of (13.11) therefore does not include improper coordinate transformations.

However, there are no similar restrictions about using either right- or left-handed ennuples. Reversing an ennuple changes the signs of $X_i(i=1,2,3)$ while ∂_μ and φ_μ are invariant. It follows that the equation is invariant if

$$\psi'(x)=\gamma(4)\psi(x). \tag{18.1a}$$

The equation is therefore invariant under ennuple reflections. Hence in general we take ψ to be a spin basis for the improper Lorentz group.

In order to compare (13.11) with the usual Dirac equation one has to consider regions of space which are small enough to be regarded as approximately flat. For such a small region (13.11) admits coordinate reflections, even though it does not admit coordinate reflections in the large. To see this, consider the neighbor-

hood of the point $x=0$ (which is not a special point since the space is homogeneous). There $\lambda(x)$ differs from $\lambda(0)$ by terms of order x/R; it is these terms which change sign and prevent the space from going into itself under a reflection. If these terms of order x/R are neglected, however, then the neighborhood of the origin does admit reflections. Since the differential equation is formally invariant in any case, one may now say that to this approximation the differential equation also admits coordinate reflections. In this same approximation the displacement operators commute.

To complete the correspondence with Dirac's equation one notes that the ennuple ordinarily used to define spinors also defines the coordinate system. The conventional reflection therefore means simultaneous coordinate and ennuple reflection. The equation is consequently invariant under conventional reflection if ψ changes according to the familiar equation

$$\psi'(x)=\gamma(4)\psi(-x). \tag{18.1b}$$

Consider next the case of zero mass. It is now possible, though not necessary, to work with two component functions ψ_+ and ψ_- restricted by the conditions

$$\gamma_5\psi_+=\psi_+, \tag{18.2a}$$

$$\gamma_5\psi_-=-\psi_-. \tag{18.2b}$$

The differential equations then become

$$[\gamma^\mu\partial u+\tfrac{3}{4}\gamma^\mu(1+\gamma^5)\varphi_\mu]\psi_+=0, \tag{18.3a}$$

$$[\gamma^\mu\partial\mu+\tfrac{3}{4}\gamma^\mu(1-\gamma^5)\varphi_\mu]\psi_-=0. \tag{18.3b}$$

The violation of parity in (18.3) is associated with the vanishing of the rest mass and has nothing to do with the torsion.

19. MACH'S PRINCIPLE

According to (10.14) the structure of the given space is entirely specified by φ^μ.[5] By Mach's principle φ^μ in turn should be determined by the matter field. However, that is not yet possible because our field equations are incomplete. We intend to continue the systematic discussion elsewhere but for the present let us tentatively suppose that the torsion is determined by the matter field in the simplest possible manner, namely,

$$\varphi^\mu=r^2(\bar\psi\gamma^5\gamma^\mu\psi), \tag{19.1a}$$

where r is a constant which must have the dimensions of a length. Then (13.11) becomes

$$[\gamma^\mu\partial_\mu+\tfrac{3}{2}r^2(\bar\psi\gamma^5\gamma^\mu\psi)\gamma_5\gamma_\mu]\psi=m\psi. \tag{19.1b}$$

[5] According to gravitational theory the corresponding equation is

$$R_{\mu\nu}-\tfrac{1}{2}Rg_{\mu\nu}+\Lambda g_{\mu\nu}=-kT_{\mu\nu}.$$

This equation is consistent with (10.14) and restricts the energy momentum tensor to the form

$$T_{\mu\nu}=dg_{\mu\nu}+e\varphi_\mu\varphi_\nu.$$

Equation (19.1b) is the nonlinear spinor equation. It has been pointed out by Pauli and Heisenberg[6] that, in the case $m=0$, this equation admits the Gursey group, which is isomorphic to the neutrino and isospin groups.

Equation (19.1) is meant to be illustrative only. Another possibility is

$$(\Box-\lambda^{-2})\varphi^\mu=\gamma(\bar\psi\gamma^5\gamma^\mu\psi), \tag{19.2}$$

where λ is a length and γ a dimensionless coupling constant. Equations (19.1) and (19.2) correspond in the language of the quantized theory to axial vector interactions of the Fermi and Yukawa type. From the point of view given here, equations such as (19.1) and (19.2) would have to be obtained from the generally covariant theory with variable torsion.

20. QUANTIZED THEORY

If φ and ψ are regarded as quantized fields then of course quanta will be associated with the torsion as well as with the matter field, and in that case the assumption of constant torsion cannot be made. The previous remarks about reflection invariance in the small also no longer apply: one expects, however, that it is possible to construct a quantized field theory which does not conserve parity.

Since we do not have a complete set of field equations, little can be said about the quantized theory at this point. However, Eqs. (13.11) and (19.2), or (19.1b), are complete and may be regarded as the equations of motion of a quantized theory which is suggested by the geometry. In order to interpret (19.1) as an ordinary Fermi coupling between quantized fields, the conventional coupling constant must be related to the length appearing here as follows:

$$g/\hbar c=r_f^2. \tag{20.1}$$

With $g=1.4\times10^{-49}$ erg cm^3, one has $r_f=2.1\times10^{-16}$ cm.

Although quantum mechanical effects are generally beyond the scope of this paper, it is of interest to discuss the following situation, which is not entirely classical, in order to interpret our results. Consider a closed space of uniform torsion uniformly and diffusely filled with a gas of identical fermions which are described by the quantized field ψ obeying Eq. (13.11a). For ψ write

$$\psi=\textstyle\sum_k u_k a_k, \tag{20.2}$$

where the a_k are the annihilation operators and u_k is a free particle solution satisfying the classical equation

$$[\gamma^\mu\partial_\mu+\tfrac{3}{2}\gamma^5\gamma^\mu\varphi_\mu]u_k=mu_k \tag{20.2a}$$

with

$$\varphi^\mu=(0,0,0,\omega). \tag{20.2b}$$

[6] W. Heisenberg, *1958 International Conference on High Energy Physics at CERN* (CERN Scientific Information Service, Geneva, 1958).

The u_k are very nearly plane waves, but closed on themselves. Assume that all energies are so low that the number of free bosons present is negligible and that the only effect of the boson field is to mediate collisions between the fermions. Then we may make the classical approximations, corresponding to (19.1b) and (19.2), respectively,

$$\varphi^\mu = r^2\langle\bar{\psi}\gamma^5\gamma^\mu\psi\rangle \tag{20.3a}$$

$$\varphi^\mu = \gamma\int[e^{-r/\lambda}/r]\langle\bar{\psi}\gamma^5\gamma^\mu\psi\rangle d\mathbf{r}, \tag{20.3b}$$

where $\langle\cdots\rangle$ means expectation value for a given quantum mechanical state of the whole gas. In this approximation one may say that the structure of space (which is entirely defined by φ^μ) and the expectation function $\langle\bar{\psi}\gamma^5\gamma^\mu\psi\rangle$ of the matter field are codetermined.

Ignoring velocity dependence, we crudely approximate the magnitude of φ in the two cases as follows:

$$\varphi \cong r^2\rho, \tag{20.4a}$$

$$\varphi \cong 4\pi\gamma\lambda^2\rho, \tag{20.4b}$$

where ρ is the average effective particle density in space. On the other hand, according to (20.2b) and (15.8b), the magnitude of φ is

$$\omega = K^{\frac{1}{2}}(-g)^{\frac{1}{4}} = 1/R, \tag{20.5}$$

where R is the radius of curvature and we have put $(-g)^{\frac{1}{4}}=1$. For the given model of spacetime, the total three-dimensional volume (V) and mass (M) are

$$V = 2\pi^2R^3, \tag{20.6a}$$

$$M = 4\pi^2R/\kappa, \tag{20.6b}$$

where κ is the gravitational constant: $\kappa = 2\times10^{-27}$ cm/g. The matter density is

$$\rho_m = M/V = 2/\kappa R^2. \tag{20.6c}$$

We have assumed that all particles are alike and neutral. Let the total number of particles be N and the mass of a single particle be M_N. Then the particle density is

$$\rho = N/V = \rho_m/M_N, \tag{20.7}$$

and by (20.5) and (20.6c) it is possible to write (20.4a) in either of the equivalent forms

$$N = 2\pi^2(R/r)^2, \tag{20.8a}$$

$$r = [\tfrac{1}{2}R\kappa M_N]^{\frac{1}{2}}. \tag{20.8b}$$

The corresponding formulas based on (20.4b) may be obtained by replacing r by $(4\pi\gamma)^{\frac{1}{2}}\lambda$. These results relate the radius of the space to microscopic constants. Although the argument leading to (20.8) is meant to be only roughly illustrative, it is clear that relations like (20.8) do exist for the given model.

21. PHYSICAL INTERPRETATION

The model of physical space under consideration has the properties (14a)–(14d), and in addition is static, homogeneous, and isotropic; it is therefore very similar to the usually accepted microspace. It is nevertheless essentially different from the configuration space of quantum field theory in having an asymmetric connection or torsion. The existence of the torsion prevents the space from having reflection invariance. Viewed from within the Lorentz subgroup, the torsion transforms as an axial vector; roughly speaking then, one may say that the proposed model modifies the usual equations of elementary particle theory by the introduction of a fundamental, axial vector coupling.

If an ennuple is (+) or (−) displaced from P to P', the new ennuple at P' may be regarded as obtained by first making a (0) displacement and then a rotation at P', where the amount of the rotation is determined by the torsion and the sign is opposite in the two cases. Regarding the (0) displacement as the analog of an ordinary translation, one may say that the torsion defines at every point two screw motions of opposite helicity. One may say also that the torsion field distinguishes between two kinds of spinor field ($\psi_\pm$) according to Eq. (13.11), where ψ_+ and ψ_- are associated with parallel transfers of opposite helicity.

If one tries to relate the massless equation to the neutrino then (18.3) appears to be a better choice than (13.11b) since the only interactions of neutrinos with other matter appear to be weak; with this choice the parity violating interaction term in (18.3) would be identified with the universal weak coupling. According to such a viewpoint φ^μ is, in the classical theory, related to the torsion of space and in the quantized theory it is related to the weak couplings.

Such an interpretation is not necessarily inconsistent but is, in fact, similar to the usual view of the quantized gravitational field. That is, the distribution of matter is believed to produce a curvature of space which may be specified by the metric field ($g_{\alpha\beta}$). We shall now postulate that the matter distribution also produces a torsion (φ^μ) of space. In the geometric analysis of this paper, φ^μ must be regarded as a classical field just as $g_{\alpha\beta}$ represents the classical gravitational field. By this one means that φ^μ and $g_{\alpha\beta}$ are to be interpreted as quantum expectation values. At the classical level they are both universal because of their common geometrical interpretation.

In order to establish a meaningful connection with elementary particle theory, the classical axial vector field should, according to the usual view of such matters, be interpreted in terms of quantum field operators. Because so little is known there are several possibilities.

First, it is not clear to what extent contributions to

the torsion could arise from the strong as well as the weak couplings. The weak interactions are now believed to be axial vector (or vector) with equal parity conserving and parity violating parts. However, it is quite possible that the strong interactions are also axial vector (possibly with a small parity violating part). [In this connection it is important that Eq. (19.1b), with $m=0$, admits the Gursey group, which describes baryon as well as lepton symmetries.] If it does turn out that the strong interactions are partly or wholly axial vector, then the torsion should be determined mainly by them.

There is next the formal question whether the φ operator should be regarded as a boson field or as a bilinear product of fermion fields. The former view corresponds to assuming an intermediate boson. According to the latter possibility, the basic interaction is a four fermion coupling.

The last question which we shall mention concerns the possible existence of diagonal terms in the Fermi couplings, i.e., terms of the form

$$(\bar{\psi}_i\gamma_5\gamma_\mu\psi_j)(\bar{\psi}_k\gamma_5\gamma_\mu\psi_l),$$

where $k=l$ and $i=j$. We shall assume that there are such terms and, in fact, that the macroscopic torsion is related to the microscopic theory as follows:

$$\varphi_\mu=(g/\hbar c)\sum_i\langle\bar{\psi}_i\gamma_5\gamma_\mu\psi_i\rangle=(g/\hbar c)\langle\cdots\rangle \quad (21.1)$$

summed over all fields interacting with the given field.

This last equation should be interpreted in the same way as (20.3a) and (20.3b) of the previous section. There we considered a gas of identical neutral fermions filling a closed space and determining the torsion of that space according to Eqs. (20.3), which are essentially the same as (21.1). We shall now ask in general terms how such a space checks as a cosmological model if the fermions are imagined to be nucleons and their φ coupling is regarded as a point or very short range interaction. In sufficiently energetic collisions new particles will be produced because of this coupling and the cross sections for such processes will be determined by the lengths r_0 and $\gamma\lambda$. However, under cosmological conditions, nuclear collisions occur at low energy and are adequately described by a classical nuclear potential. In the approximation adopted, it may be supposed that these collisions are just sufficient to maintain the consistency of the very slightly curved $(1/R)$ plane waves.

The observed density of matter in space is $\rho_m\cong10^{-30}$ g/cm^3. That corresponds by Eq. (20.6c) to $R=10^{28}$ cm. If this value of R is put into (20.8b) one obtains $r=10^{-12}$ cm. That may be compared with $r_f=10^{-16}$ cm, which characterizes the weak couplings. If this estimate is taken seriously, then it follows that the torsion estimated from the observed average density of matter cannot be attributed to the weak couplings. On the other hand, if the pion coupling is written as a pair interaction, as in the theory of the composite pion,[7] then the value of the Fermi length is $r=10^{-13}$ cm (since the corresponding coupling constants differ by 10^6). Therefore if the proposed geometry and the estimate based on it have any physical content, it would appear that the torsion of space is associated mainly with the strong interactions. These results, if they are not fortuitous,[8] suggest that a field theory based on an asymmetric connection is of great interest in connection with both the strong and the weak interactions.

These final remarks, beginning in Sec. 19, have outrun the systematic analysis, which is based throughout on the assumption of uniform torsion. This assumption permits an exact geometric analysis but prevents us from writing a complete set of field equations. In its present form, the physical model is subject to the restriction that it describe only one neutral fermion field—except for the distinction between $\psi(+)$ and $\psi(-)$. Finally, no physical interpretation of the fundamental duality between $\psi(+)$ and $\psi(-)$, i.e., of the transposition symmetry, has been given in this paper.[9]

ACKNOWLEDGMENTS

I wish to thank William Ramsay for the remark which is proved in the Appendix and also to thank him and Dr. Herbert Fried for helpful discussion.

[7] This may be estimated by writing

$$g_F(\bar{p}_<\cdots n_>)(\bar{n}_>\cdots p_>)=g\hbar c(2\pi/\epsilon_\pi)^{\frac{1}{2}}\exp(ip_\pi x/\hbar)(\bar{n}_>\cdots p_>).$$

The left (right) sides refer to a Fermi (Yukawa) description of $\pi^-+p\to n$. Here $(\bar{p}_<\cdots n_>)$ is the composite pion and $<(>)$ means negative (positive) energy. We factor

$$\bar{p}_<\cdots n_>=\exp(ip_\pi x/\hbar)\mu(0),$$

where p_π is the momentum of the composite pion and $\mu(0)$ is the amplitude of the internal wave function at zero separation. Let us put

$$(4\pi/3)(\hbar/Mc)^3|\mu(0)|^2=1,$$

where $\hbar/Mc$ represents, very roughly, the "radius" of the pion. Then

$$g_F\cong(4\pi/3)^{\frac{1}{2}}(M/2m_\pi)^{\frac{1}{2}}(4\pi g^2/\hbar c)^{\frac{1}{2}}(Mc^2)(\hbar/Mc)^3.$$

With $g^2/\hbar c=15$, one obtains $g_F\cong10^{-43}$ erg cm^3.

[8] It would be meaningless to refine the numerical estimates at this stage. However, if the general point of view of this paper can be maintained, then in a more complete theory the following improvements, among others, should be made: (a) the astronomical facts should not be approximated by the static Einstein model, but instead by a more realistic expanding model, (b) the density $\rho=10^{-30}$ should be replaced by an effective density which takes into account the velocity dependence implied by γ^5. This effective density would be negligible for intergalactic hydrogen and somewhat reduced for nucleons in nuclei.

[9] The distinction between $\psi(+)$ and $\psi(-)$ may depend on the theory of the electromagnetic field. It is possible to introduce the electromagnetic field in such a way that transposition symmetry corresponds to charge symmetry.

APPENDIX

Define

$$c_{mnp}=\bar{g}_{mt}c^{t}{}_{np}.$$

The c_{mnp} are completely antisymmetric because of the Jacobi relations.

Construct

$$\theta=(1/3!)\epsilon^{mnps}c_{mnp}X_s;$$

then

$$(\theta,X_r)=(1/3!)\epsilon^{mnps}c_{mnp}c^{t}{}_{sr}X_t.$$

Assume $\bar{g}\neq 0$. Then $\bar{g}^{mn}$ may be constructed. Assume $\bar{g}^{mn}$ diagonal. Then

$$\begin{aligned}(\theta,X_1)&=[c_{134}c^{t}{}_{21}+c_{214}c^{t}{}_{31}+c_{123}c^{t}{}_{41}]X_t\\&=B^tX_t\end{aligned}$$

and, if $t=2$, for example,

$$\begin{aligned}B^2&=c_{134}c^{2}{}_{21}+c_{214}c^{2}{}_{31}+c_{123}c^{2}{}_{41}\\&=\bar{g}^{22}[c_{214}c_{231}+c_{123}c_{241}]=0.\end{aligned}$$

Hence, if $\bar{g}\neq 0$, it follows that

$$(\theta,X_t)=0.$$

In a new representation such that $X_1{}'=\theta$ one has

$$c^{r}{}_{1s}{}'=0.$$

It follows that

$$\bar{g}'=\bar{g}=0.$$

Therefore, the assumption that $\bar{g}\neq 0$ is inconsistent.

Errata

Double Commutator in Quantum Field Theory

[J. Math. Phys. 1, 231 (1960)]

R. F. STREATER
CERN, Geneva, Switzerland

Equations (11) and (12) should read:

$$\bar{D}(p,q)=0 \quad \text{unless} \quad \begin{cases} p^2 \geqslant m_1^2,\ q^2 \geqslant m_3^2,\ p\cdot q>0 & \text{or} \\ p^2 \geqslant m_1'^2,\ (p-q)^2 \geqslant m_2^2,\ p\cdot(p-q)>0 \end{cases} \tag{11}$$

$$\bar{D}(p,q)\neq 0 \tag{12}$$

in this spectrum.

That is, the conditions $p\cdot q>0$, $p\cdot(p-q)>0$ to Eq. (11) have been added.

Analytic Properties of Radial Wave Functions

[J. Math. Phys. 1, 319 (1960)]

ROGER G. NEWTON
Indiana University, Bloomington, Indiana

Page 321, line before Eq. (2.15): Read "(2.14)" for "(2.12)." Page 330, line 8 below Eq. (5.1′): Insert "if" between "$|k_0|$" and "is." Page 334, line 5 below Eq. (6.2): Read

$$\mathfrak{G}_l(k;r,r')=\begin{cases} \varphi_l(k,r)a(k,r'), & r<r', \\ f_l(-k,r)b(k,r'), & r>r'. \end{cases}$$

Page 339, Eq. (9.7) should read:

$$F_J{}^T(k)\equiv k^L W[F_J(k,r),\Phi_J(k,r)]. \tag{9.7}$$

Equation (9.9′) should read:

$$S_J(k)=[F_J(-k)]^{-1}F_J(k). \tag{9.9′}$$

Page 344, line 6 from bottom, right-hand column: Insert "upper half of the" before "complex."