JOURNAL OF
MATHEMATICAL
PHYSICS

VOLUME 2, NUMBER 3 MAY–JUNE, 1961

On the Connection of Spin and Commutation Relations between Different Fields

HUZIHIRO ARAKI
Department of Nuclear Engineering, Kyoto University, Yosida, Kyoto, Japan
(Received October 14, 1960)

The connection of spin and commutation relations for different fields is studied. The normal locality is defined as the property that two Boson fields as well as a Boson field and a Fermion field commute, while two Fermion fields anticommute with each other at a spacelike distance. A regular locality is defined as any combination of commutativity and anticommutativity between various pairs of fields at spacelike distance, where the kinematically related fields are assumed to obey the same type of commutation relations. The normal and regular weak locality is defined in a corresponding way. It is proved on the basis of the Lorentz invariance and spectrum conditions that any regular locality is equivalent to the normal locality plus a set of even-oddness conservation laws. It is further shown, under the assumption of the normal weak locality between pairs of the same field, that any regular weak locality is equivalent to the normal weak locality plus a set of even-oddness conservation laws.

1. INTRODUCTION

AS is well known, the connection of spin and statistics has been proved first by Pauli[1] for free fields. Later, proofs based on conventional basic postulates of the relativistic quantum field theory have been given by Luders and Zumino[2] and by Burgoyne[3] for interacting fields. The theorem proved by the latter authors states that the wrong connection of spin and statistics can never be postulated for the relativistic quantum field theory where the wrong connection means the anticommutation relation between $\varphi_\alpha(x)$ and $\varphi_\beta(y)^*$ for a Boson field φ and the commutation relation between $\psi_\alpha(x)$ and $\psi_\beta(y)^*$ for a Fermion field ψ, for the spacelike separation of x and y.

The connection of spin and statistics in a wide sense, namely, commutation relations between different fields, has been studied by various authors.[4–6] The normal connection of spin and statistics is defined as the case where two Boson fields, or a Boson and a Fermion field, commute while two Fermion fields anticommute for the spacelike separation of their arguments. Luders[5] and Kinoshita[6] have shown on the basis of the Hamiltonian formalism that the general case can be reduced to the normal case by a series of the Klein transformations[7] which do not affect any physical content of the theory.

The purpose of this paper is to give a different proof of the connection of spin and statistics *in a wide sense* which is based solely on the conventional basic postulates of the relativistic quantum field theory.

To state our results, let us define normal and regular localities. The normal locality is what has been called the normal connection of spin and statistics above. The regular locality is any specific combination of commutation and anticommutation relations between various pairs of fields for the spacelike separation of their arguments where the kinematically related fields (namely, various components of covariant fields as well as their hermitian conjugates) are assumed to obey the commutation or anticommutation relation simultaneously.

We also introduce similar definitions for the weak

[1] W. Pauli, Phys. Rev. **58**, 716 (1940). *Note added in proof.* The first proof of spin and statistics for free fields appear in the paper of M. Fierz, Helv. Phys. Acta **12**, 3 (1939). The author is indebted to Professor R. Jost for his remark on this point.

[2] G. Luders and B. Zumino, Phys. Rev. **110**, 1450 (1958).

[3] N. Burgoyne, Nuovo cimento **8**, 607 (1958).

[4] K. Nishijima, Progr. Theoret. Phys. (Kyoto) **5**, 187 (1950); S. Oneda and H. Umezawa, *ibid.* **9**, 685 (1953); T. Kinoshita, Phys. Rev. **96**, 199 (1954); H. Umezawa, J. Podolanski, and S. Oneda, Proc. Phys. Soc. (London) **A68**, 503 (1955); R. Spitzer, Phys. Rev. **105**, 1919 (1957).

[5] G. Luders, Z. Naturforsch. **13a**, 254 (1958).

[6] T. Kinoshita, Phys. Rev. **110**, 978 (1958).

[7] O. Klein, J. Phys. (U.S.S.R.) **9**, 1 (1938).

locality.[8] A set of fields is said to have the weak locality if their vacuum expectation values satisfy

$$(\Psi_0, \varphi_1(x_1)\cdots\varphi_n(x_n)\Psi_0) = (-1)^{\eta(1\cdots n)}(\Psi_0, \varphi_n(x_n)\cdots\varphi_1(x_1)\Psi_0) \quad (1.1)$$

for any real regularity points $x_1\cdots x_n$. If η is the same as the one derived from the normal locality, the weak locality is said to be normal, and if η is the same as the one derived from a regular locality, the weak locality is said to be regular.

Our result is theorem 1.

Theorem 1. In a relativistic quantum-field theory satisfying (1) the invariance under the inhomogeneous restricted Lorentz group, and (2) spectrum conditions, namely, the existence of the unique vacuum, the positivity of the energy, and the existence of the lowest nonzero mass, any regular locality is equivalent to the normal locality plus a set of even-oddness conservation laws. With the additional assumption of the normal weak locality between the *same* field, any regular weak locality is equivalent to the normal weak locality plus the same set of even-oddness conservation laws.

It should be stressed that the connection of spin and statistics between *different* fields can be proved even with *weak* locality.

2. EVEN-ODDNESS CONSERVATION LAWS

A theory is said to have an even-oddness conservation law for a set α of fields[9] if all the Wightman functions containing an odd number of fields belonging to α vanish. In such a case, the two spaces H_0 and H_e generated from the vacuum by polynomials of fields which are of odd or even degrees, respectively, with regard to fields in α, are orthogonal to each other and the both spaces are obviously invariant under the inhomogeneous restricted Lorentz group. Hence the operator $p(\alpha)$, defined to be 1 on H_e and -1 on H_0, is well defined[10] and commutes with the representation $U(a,\Lambda)$ of the inhomogeneous restricted Lorentz group.

Suppose we define new fields φ' by multiplying $p(\alpha)$ to all the fields φ belonging to a certain set β,[9]

$$\begin{aligned} \varphi' &= p(\alpha)\varphi \quad \text{if } \varphi\in\beta, \\ \varphi' &= \varphi \quad\quad\ \text{if } \varphi\notin\beta. \end{aligned} \quad (2.1)$$

Then, the φ' are again Lorentz covariant fields and the commutation and anticommutation relations are interchanged for the localities between a field in $\alpha-\beta$[11] and a field in β, as well as those between a field in $\alpha\cap\beta$ and a field in $\beta-\alpha$, while the rest of the localities are unchanged. The Wightman functions for the new fields φ differ at most from those for old φ by sign, and the physical content of the theory such as the S-matrix elements is obviously the same for φ' and for φ.

As we shall see in Sec. 4, if we have sufficiently many even-oddness conservation laws in the theory, the transformation of the type (2.1)[7] enables us to reduce any regular locality to the normal locality. Hence, it is important to derive sufficiently many even-oddness conservation laws from a regular locality. This can be accomplished by the following theorem.[12]

Theorem 2. If a theory satisfies (1) and (2) of theorem 1, if C_1 and C_2 are products of field operators, and if C_1 and $U(\lambda a,1)C_2U(\lambda a,1)^{-1}$ anticommute for a spacelike vector a and sufficiently large λ, then the vacuum expectation value of *either* C_1 *or* C_2 vanishes.

The proof of this theorem is based on the following lemma.

Lemma 1. In a theory satisfying (1) and (2) of theorem 1,

$$\begin{aligned} \lim_{\lambda\to\infty} &(\Psi_0, \varphi_1(x_1)\cdots\varphi_k(x_k) \\ &\times\varphi_{k+1}(x_{k+1}-\lambda a)\cdots\varphi_n(x_n-\lambda a)\Psi_0) \\ &= (\Psi_0, \varphi_1(x_1)\cdots\varphi_k(x_k)\Psi_0) \\ &\quad\times(\Psi_0, \varphi_{k+1}(x_{k+1})\cdots\varphi_n(x_n)\Psi_0), \end{aligned} \quad (2.2)$$

where a is a spacelike vector.

This lemma has been proved in an averaged sense for an arbitrary configuration of x's in theorem 3 of footnote 13, and the exponential approach to the limit is provable in a manner similar to the proof of theorem 1 of footnote 13 if $x_1-x_2, \cdots x_k-x_{k+1}, a, x_{k+1}-x_{k+2}\cdots x_{n-1}-x_n$ form a Jost point.

By this lemma we have

$$\lim_{\lambda\to\infty} (\Psi_0, C_1U(\lambda a,1)C_2U(\lambda a,1)^{-1}\Psi_0) = (\Psi_0,C_1\Psi_0)(\Psi_0,C_2\Psi_0), \quad (2.3a)$$

$$\lim_{\lambda\to\infty} (\Psi_0, U(\lambda a,1)C_2U(\lambda a,1)^{-1}C_1\Psi_0) = (\Psi_0,C_2\Psi_0)(\Psi_0,C_1\Psi_0), \quad (2.3b)$$

and hence, we have theorem 2.

A similar theorem holds for weak locality.[14]

Theorem 3. If a theory satisfies (1) and (2) of theorem 1 plus a regular weak locality, and if $\eta(1\cdots n)+\eta(1\cdots k)+\eta(k+1\cdots n)$ is odd, then either

$$(\Psi_0,\varphi_1(x_1)\cdots\varphi_k(x_k)\Psi_0) \text{ or } (\Psi_0,\varphi_{k+1}(x_{k+1})\cdots\varphi_n(x_n)\Psi_0)$$

vanishes identically.

The proof of this theorem is similar to the previous theorem, namely, if $x_1-x_2\cdots x_k-x_{k+1}, a, x_{k+1}-x_{k+2}\cdots$

[8] R. Jost, Helv. Phys. Acta **30**, 409 (1957).

[9] The kinematically related fields are assumed to be *simultaneously* in the set or not in the set.

[10] The vacuum state is assumed to be cyclic with respect to the set of all the fields.

[11] $\alpha-\beta$ is the set of elements belonging to α, but not belonging to β.

[12] H. Araki, J. Math. Phys. **2**, 163 (1961). See Appendix B.

[13] H. Araki, Ann. Phys. **11**, 260 (1960).

[14] The author is indebted to Professor R. Jost for pointing out the applicability of our proof to the case of the weak locality.

$x_{n-1}-x_n$ form a Jost point, we have

$$\lim_{\lambda\to\infty}(\Psi_0,\ \varphi_1(x_1)\cdots\varphi_k(x_k)\times\varphi_{k+1}(x_{k+1}-\lambda a)\cdots\varphi_n(x_n-\lambda a)\Psi_0)$$
$$=(\Psi_0,\ \varphi_1(x_1)\cdots\varphi_k(x_k)\Psi_0)\times(\Psi_0,\ \varphi_{k+1}(x_{k+1})\cdots\varphi_n(x_n)\Psi_0), \quad (2.4a)$$

$$\lim_{\lambda\to\infty}(\Psi_0,\ \varphi_n(x_n-\lambda a)\cdots\varphi_{k+1}(x_{k+1}-\lambda a)\times\varphi_k(x_k)\cdots\varphi_1(x_1)\Psi_0)$$
$$=(\Psi_0,\ \varphi_n(x_n)\cdots\varphi_{k+1}(x_{k+1})\Psi_0)\times(\Psi_0,\ \varphi_k(x_k)\cdots\varphi_1(x_1)\Psi_0). \quad (2.4b)$$

Hence, if $\eta(1\cdots n)+\eta(1\cdots k)+\eta(k+1\cdots n)$ is odd, we have, from (1.1),

$$(\Psi_0,\ \varphi_1(x_1)\cdots\varphi_k(x_k)\Psi_0)\times(\Psi_0,\ \varphi_{k+1}(x_{k+1})\cdots\varphi_n(x_n)\Psi_0)=0. \quad (2.5)$$

Thus, one of the two factors vanishes at least in a real neighborhood of its real regularity point. By the well-known argument of the analytic continuation, we have theorem 3.

3. ILLUSTRATIVE EXAMPLE

As a simple example, we consider a theory of two scalar fields φ and ψ which anticommute with each other at the spacelike separation of their arguments. Let C_1 be a product of $m_1\varphi$'s and $n_1\psi$'s, and C_2 be a product of $m_2\varphi$'s and $n_2\psi$'s. By theorem 2, if $m_1n_2+m_2n_1$ is odd, then

$$(\Psi_0,C_1\Psi_0)(\Psi_0,C_2\Psi_0)=0. \quad (3.1)$$

Now at least one of the following four possibilities always holds for the vacuum expectation value w of $m\varphi$'s and $n\psi$'s:

(1) all the w with odd m vanish,
(2) all the w with odd n vanish,
(3) at least one w with odd m and odd n does not vanish,
(4) at least one w with odd m and even n *and* at least one w with even m and odd n do not vanish.

For case (1), we have the conservation law of the even-oddness $p(\varphi)$ of φ. If we define $\psi'=p(\varphi)\psi$, then ψ' is a scalar field and ψ' and φ commute with each other. The pair of fields φ and ψ' is physically equivalent to the pair φ and ψ and satisfies the normal locality. For case (2), we have the conservation law of the even-oddness $p(\psi)$ of ψ, and $\varphi'=p(\psi)\varphi$ commutes with ψ. For case (3), by taking the nonvanishing w with odd m and odd n as $(\Psi_0, C_1\Psi_0)$ in (3.1), we see that all the w with even m and odd n, as well as those with odd m and even n, vanish. Hence, we have the conservation law of the even-oddness p of φ and ψ, and $\varphi'=p\varphi$ commutes with ψ. For case (4), by taking the non-vanishing w with odd m and even n as $(\Psi_0, C_1\Psi_0)$ in (3.1). we see that (2) holds. [In a similar manner (1) also holds.] Thus, we see that theorem 1 holds for this simple example.

4. GENERAL CASE

For the discussion of the general case, it is advantageous to use the vector space S over the prime field with characteristic 2.[15,16] We number different fields, giving the same number to kinematically related fields. To each Wightman function w, we assign a vector $s(w)$ in S whose ith component is 1 or 0 according to whether the number of ith field contained in the Wightman function w is odd or even.

For any set α of fields, we denote by $t(\alpha)$ the vector in S whose ith component is 1 or 0 according to whether or not the ith field belongs to α. Then, a given theory has an even-oddness conservation law for the set α if and only if

$$(t(\alpha),s(w))=1 \quad \text{implies} \quad w=0. \quad (4.1)$$

As is well known, for the set f of all the Fermions, (4.1) holds.

To each regular locality we assign a matrix (ϵ_{ij}) whose matrix element ϵ_{ij} is 1 if ith field and jth field anticommute and is 0 if they commute with each other. Note that kinematically related fields are assumed to obey the same locality. We also assign the same matrix (ϵ_{ij}) to the corresponding regular weak locality. By definition

$$\epsilon_{ij}=\epsilon_{ji}. \quad (4.2)$$

By the theorem proved in footnotes 2 and 3 (or by assumption in the case of the weak locality),

$$\epsilon_{ii}=t(f)_i. \quad (4.3)$$

We define a functional γ over S (not necessarily linear) by

$$\gamma(s)=0 \quad \text{if } s(w)=s \text{ implies } w=0;\quad =1 \text{ otherwise.} \quad (4.4)$$

Then, (4.1) can be restated as

$$\gamma(s)[t(\alpha),s]=0 \quad (4.5)$$

for all s if and only if we have the even-oddness conservation law for the set α.

If we define a linear transformation ϵ of the space S by

$$(\epsilon s)_i=\textstyle\sum_j \epsilon_{ij}s_j, \quad (4.6)$$

[15] The prime field with the characteristic 2 consists of two elements, 0 and $1.\ 0+0=1+1=0\cdot0=0\cdot1=1\cdot0=0.\ 0+1=1\cdot1=1$. A vector in S consists of an n-tuple of elements of this field: $s=(s_1\cdots s_n)$. The sum of two vectors is defined by the component-wise sum, and the inner product is defined (in a fixed basis) by

$$(s,t)=\sum_{i=1}^{n} s_it_i.$$

[16] See appendix of footnote 5.

then theorems 2 and 3 state that for any s and t,

$$\gamma(t)(t,\epsilon s)\gamma(s)=0. \tag{4.7}$$

For the normal locality, ϵ is given by

$$\epsilon_{ij}{}^{(n)}=t(f)_i t(f)_j. \tag{4.8}$$

The transformation (2.1) of fields changes ϵ into

$$\epsilon_{ij}+t(\alpha)_i t(\beta)_j+t(\beta)_i t(\alpha)_j. \tag{4.9}$$

To prove theorem 1, we have to show that any ϵ satisfying (4.2), (4.3), and (4.7) can be reduced to $\epsilon^{(n)}$ by a series of transformations of the type (4.9) where α satisfies (4.5). Hence, it suffices to prove the following lemma.

Lemma 2. Any ϵ satisfying (4.2), (4.3), and (4.7) can be written in the form

$$\epsilon_{ij}=\sum_k (t(\alpha_k)_i s^{(k)}{}_j+s^{(k)}{}_i t(\alpha_k)_j)+\epsilon_{ij}{}^{(n)} \tag{4.10}$$

where each α_k satisfies (4.5).

For the proof, we define $\sigma=\epsilon-\epsilon^{(n)}$. We then restate (4.2) and (4.3) in a basis-independent way as

$$(s,\sigma s)=0, \quad (s,\sigma t)=(t,\sigma s). \tag{4.11}$$

We now define a subset Γ of S by

$$\Gamma=\{s\in S;\ \gamma(s)=1\}. \tag{4.12}$$

Then (4.7) can be restated as

$$\sigma\Gamma\perp\Gamma. \tag{4.13}$$

If we denote by $\bar{\Gamma}$ the linear subset of S generated by Γ, then $\Gamma^\perp=\bar{\Gamma}^\perp$ is a linear subset of S, and hence, (4.13) is equivalent to

$$\sigma\bar{\Gamma}\perp\bar{\Gamma}. \tag{4.14}$$

Furthermore, we have an even-oddness conservation law for the set α if and only if

$$t(\alpha)\perp\bar{\Gamma}. \tag{4.15}$$

Let $e^{(1)}\cdots e^{(m)}$ be a basis of $\bar{\Gamma}$ and let $e^{(1)}\cdots e^{(n)}$ be a basis of S containing the above basis of $\bar{\Gamma}$. Let $\delta^{(1)}\cdots\delta^{(n)}$ be linear functionals over S defined by

$$\delta^{(i)}e^{(k)}=\delta_{ik}. \tag{4.16}$$

As one can easily show, any linear functional δ over S can be expressed as

$$\delta s=(d,s),$$

with a suitable element d in S. Let $d^{(i)}$ be such an element in S corresponding to $\delta^{(i)}$.

Let us denote the matrix elements of σ in this new coordinate system by

$$\sigma_{ik}{}'=(e^{(i)},\sigma e^{(k)}). \tag{4.17}$$

Now, for any elements s and t in S,

$$(t,\sigma s)=\sum_{i,k}(t,d^{(i)})\sigma_{ik}{}'(d^{(k)},s). \tag{4.18}$$

By (4.11) and (4.17),

$$\sigma_{ii}{}'=0, \quad \sigma_{ik}{}'=\sigma_{ki}{}'. \tag{4.19}$$

By (4.14),

$$\sigma_{ik}{}'=0 \quad \text{if } i,k<m. \tag{4.20}$$

Hence, we have

$$\sigma=\sum_{k>m}|d^{(k)}\rangle\langle s^{(k)}|+|s^{(k)}\rangle\langle d^{(k)}|, \tag{4.21}$$

where

$$s^{(k)}=\sum_{i<k}\sigma_{ik}{}'d^{(i)}. \tag{4.22}$$

Since $d^{(k)}$ for $k>m$ are orthogonal to $\bar{\Gamma}$, (4.21) is the same as (4.10) [cf. (4.15)].

Free Quantized Lorentzian Fields

J. SCHWARTZ
Institute of Mathematical Sciences, New York University, New York, New York
(Received March 11, 1960)

A systematic classification is made of all local Lorentz-invariant quantized field theories which are free in the sense that $[\psi(x),\psi(x')]$ is a c number. It is shown that none exist except those given in the classical paper of Fierz, and that, in fact, certain of the fields listed by Fierz are redundant.

1. INTRODUCTION

THE aim of the present paper is to classify the set of all possible (proper) Lorentz invariant quantum field theories of free fields, these theories being specified by a set of axiomatic requirements to be described. More specifically, it will be shown that the set of fields occurring in any such theory may be decomposed into sets of fields describing particles of various positive masses and of zero mass. These various sets of fields will then be completely analyzed.

The problem considered in this paper has been considered from slightly different points of view by quite a number of authors. The classical paper of Fierz,[1] gives an account of the particular fields to be noted in the present paper, without remarking that, for positive mass at any rate, the fields of the transformation $\psi_{\lambda_1\cdots\lambda_m}$ are complete and canonical, or noting that for positive mass these fields with the transformation $\psi_{\lambda_1\cdots\lambda_n\mu_1\cdots\mu_m}$ are superfluous. (Here we have employed a notation to be explained in detail in what follows.) Gel'fand and Yaglom[2] give a general, though largely unquantized, treatment of fields under the rather special assumption that the fields satisfy first-order equations of a generalized Dirac form. This work of Gel'fand and Yaglom is related to earlier works of Bhabha,[3] Harish-Chandra,[4] Moses,[5] Umezawa and Visconti,[6] Rarita and Schwinger,[7] and Lamont.[8]

2. LORENTZ GROUP AND ITS REPRESENTATIONS

Résumé

As a necessary preliminary to the detailed statement both of problems and results, we must begin with a discussion reviewing the most important properties of the Lorentz group and its irreducible representations. All this material, comprising the whole of the present section, is classical, and is derived in detail by Van der Waerden.[9] Nevertheless, it is well to review this material, albeit hurriedly, both for the convenience of the reader, and in order to establish certain details of notation to be used subsequently.

The Lorentz metric in 4-space time E^4 is

$$\begin{aligned} x^2 &= x_0^2-x_1^2-x_2^2-x_3^2 \quad x=(x_0,x_1,x_2,x_3) \\ &= t^2-x^2-y^2-z^2 \qquad\quad =(t,x,y,z)=(t,\mathbf{x}) \\ &= t^2-\mathbf{x}^2. \end{aligned} \tag{2.1}$$

A linear inhomogeneous transformation which preserves this metric will be called an extended Lorentz transformation; the collection of all such transformations is the extended Lorentz group $\mathcal{E}$. The subgroup of homogeneous transformations is called the Lorentz group and denoted by $\mathcal{L}$. If $e\in\mathcal{E}$, there exist a unique $l\in\mathcal{L}$ and $a\in E^4$ such that $ex=lx+a$; $l=l(e)$ is called the rotation part of e, $a=a(e)$ the translation part of e. The map $l\to l(e)$ is evidently a homomorphism. The translations $x\to x+a$ form an abelian subgroup $\mathcal{T}$ of $\mathcal{E}$; evidently $\mathcal{E}$ is homeomorphic to $\mathcal{L}\oplus\mathcal{T}$, that is, $\mathcal{E}$ is $\mathcal{L}\oplus E^4$ topologically.

Connectedness

$\mathcal{L}$ and consequently $\mathcal{E}$ has four connected components, since $\det(l)=\pm1$, and $\operatorname{sgn}[l(1,0,0,0)]_0=\pm1$ are both continuous on $\mathcal{L}$. Elements of the 4 components are $x\to x$ (proper component); $(t,\mathbf{x})\to(t,-\mathbf{x})$ (parity); $(t,\mathbf{x})\to(-t,\mathbf{x})$ (time reversal); $x\to -x(TP)$. The proper components of $\mathcal{L}$ and $\mathcal{E}$, that is, the connected subgroups of $\mathcal{L}$ and $\mathcal{E}$ consisting of all transformations which may be connected to the identity transformations by a continuous curve, will be denoted by $\mathcal{L}_p$ and $\mathcal{E}_p$, respectively. Note that $e\in\mathcal{E}_p$ if and only if $l(e)\in\mathcal{L}_p$.

It should be emphasized that all the results to be deduced below will follow from $\mathcal{E}_p$ invariance.

A vector $x\in E^4$ is timelike if $x^2>0$, spacelike if $x^2<0$, and null if $x^2=0$. The indecomposable invariant subsets of E^4 under $\mathcal{L}_p$ are evidently the set consisting of the origin alone, and the three families of sets

$$\begin{aligned} &\Sigma_m=\{x\,|\,x^2=-m^2\};\quad m>0\text{: These are the hyperboloids of spacelike vectors.} \\ &\Sigma_m{}^+=\{x\,|\,x^2=m^2,\ x_0>0\};\quad m\geq 0\text{: These are the hyperboloids of timelike forward vectors.} \\ &\Sigma_m{}^-=\{x\,|\,x^2=m^2,\ x_0<0\},\quad m\leq 0\text{: These are the hyperboloids of timelike backward vectors.} \end{aligned} \tag{2.2}$$

[1] Marcus Fierz, Helv. phys. Acta **12**, 3 (1938).
[2] I. M. Gel'fand and A. M. Yaglom, J. Exptl. Theoret. Phys. (U.S.S.R.) **18**, 703 (1948a).
[3] H. J. Bhabha, Revs. Modern Phys. **17**, 200 (1945).
[4] Harish-Chandra, Phys. Rev. **71**, 793 (1947).
[5] H. Moses, Nuovo cimento Suppl. **1**, 1 (1958).
[6] H. Umezawa and A. Visconti, Nuclear Phys. **7**, 348 (1956).
[7] P. Rarita and J. Schwinger, Phys. Rev. **60**, 61 (1941).
[8] J. Lamont (unpublished).
[9] B. Van der Waerden, *Die Gruppentheoretische Methoden in Quantenmechanik* (Verlag Julius Springer, Berlin, Germany, 1932).

The integral $\int f(x)dx$ of a function on 4-space can be decomposed as an integral of $\mathfrak{L}_p$-invariant surface integrals:

$$\int f(x)dx=\int_0^\infty dm\Big\{\int_{\Sigma_m} f(x)d\sigma+\int_{\Sigma_{m^+}} f(x)d\sigma+\int_{\Sigma_{m^-}} f(x)d\sigma\Big\}. \quad (2.3)$$

A corresponding decomposition of distributions may be made using the Fourier transform. If f is a (tempered) distribution, and ϕ its Fourier transform, so that we have (symbolically)

$$f(x)=\int e^{i(k,x)}\phi(k)dk,$$

we put

$$f_{(m)}(x)=\int_{\Sigma_m} e^{i(k,x)}\phi(k)dk; \quad (2.4a)$$

and

$$f^{(m)}(x)=\int_{\Sigma_{m^+}\cup\Sigma_{m^-}} e^{i(k,x)}\phi(k)dk. \quad (2.4b)$$

Then

$$f=\int (f_{(m)}+f^{(m)})dm. \quad (2.5)$$

All of these symbolic operations may readily be carried out rigorously when suitably restrictive hypotheses are made on the distribution f.

In our subsequent work with the maps $f\to f^{(m)}$ and $f\to f_{(m)}$, we shall always assume (at least implicitly) that the carrier of the Fourier transform of the distribution f is the union of a *finite* collection of the Lorentz-invariant hyperboloids Σ_m, $\Sigma_m{}^+$, and $\Sigma_m{}^-$. Under this hypothesis, formula (2.5) may be written in the simpler form

$$f=\sum_{i=1}^{N} f_{(m_i)}+\sum_{j=1}^{M} f^{(m_j)}. \quad (2.6)$$

The distributions of the form $f_{(m)}$ and $f^{(m)}$ appearing in Eq. (2.6) have Fourier transforms carried by the hyperboloids Σ_m and $\Sigma_m{}^+\cup\Sigma_m{}^-$, respectively. It follows, according to the theory of distributions, that $f_{(m)}$ and $f^{(m)}$ satisfy partial differential equations

$$(\Box+m^2)^k f_{(m)}(x)=0, \quad \text{and} \quad (\Box-m^2)^k f^{(m)}=0, \quad (2.7)$$

provided only that the exponent k is sufficiently large. Here,

$$\Box=\nabla^2-(\partial^2/\partial t^2)$$

denotes the d'Alembertian operator.

It follows easily from Eq. (2.7) that, under the hypothesis that the carrier of the Fourier transform of the distribution f is the union of a finite collection of the Lorentz-invariant hyperboloids Σ_m, $\Sigma_m{}^+$, and $\Sigma_m{}^-$, we may write $f_{(m)}$ and $f^{(m)}$ in the form

$$\begin{aligned} f_{(m)}&=P_{(m)}(\Box)f \\ f^{(m)}&=P^{(m)}(\Box)f, \end{aligned} \quad (2.8)$$

where $P_{(m)}(\Box)$ and $P^{(m)}(\Box)$ are suitably chosen finite polynomials (with constant coefficients) in the d'Alembertian operator $\Box$. This shows that if $f(x)=0$ for x in a region, then $f_{(m)}(x)$ and $f^{(m)}(x)$ vanish in the same region.

We note the following properties of the maps $f\to f_{(m)}$ and $f\to f^{(m)}$ as a formal lemma, leaving the elementary proofs to the reader.

Lemma. *The maps $f\to f_{(m)}$ and $f\to f^{(m)}$ are linear; real in the sense that if $g=f^{(m)}$, then $\bar{g}=\bar{f}^{(m)}$; and Lorentz invariant; so that $f_{(m)}(Lx)=[f(Lx)]_{(m)}$.*

Next we wish to reiterate some well-known topological properties of the Lorentz group, especially those having to do with the fundamental group of this group. The group $\mathfrak{L}_p$ is doubly connected, i.e., doubly covered by its simply connected covering group. Call this covering group U. Then the natural mapping of the covering group U onto the group $\mathfrak{L}_p$ maps precisely two points onto each point of $\mathfrak{L}$; each of these points has a neighborhood mapped in 1−1 way, so that the covering of $\mathfrak{L}_p$ by U is by smooth sheets without branch points (as in Riemann surface theory).

Thus $\mathfrak{L}_p$ admits continuous double-valued functions, but not continuous n-valued functions for $n>2$.

More explicitly: Let U be the group of 2×2 complex matrices of determinant 1. Let S be the set of all self-adjoint 2×2 matrices. Let U act on S by $a\to uau^*=h(u)a$; then U is mapped *homomorphically* into a set of linear transformations of S, preserving the determinant of $a\in S$. Write $a\in S$ as

$$a=\begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix}; \quad (2.9)$$

and put

$$a_{12}=x+iy, \quad a_{11}=t+z, \quad a_{22}=t-z \quad \text{for} \quad (t,x,y,z)\in E^4.$$

Then $\det(a)=(t+z)(t-z)-(x+iy)(x-iy)=t^2-x^2-y^2-z^2$; so that if we map each 2×2 unimodular matrix u onto the map $h(u)$ of E^4 defined by the two formulas

$$u\begin{pmatrix} t+z & x+iy \\ x-iy & t-t \end{pmatrix}u^*=\begin{pmatrix} t'+z' & x'+iy' \\ x'-iy' & t'-z' \end{pmatrix} \quad (2.10a)$$

and

$$h(u)(t,x,y,z)=(t',x',y',z'), \quad (2.10b)$$

then h is a homomorphism of U into the Lorentz group. Since U is connected, h must map U into the connected component $\mathfrak{L}_p$ of the Lorentz group. Thus, since U and $\mathfrak{L}_p$ are both six-dimensional, U must be mapped onto $\mathfrak{L}_p$. If $h(u)=1$, i.e., $uau^*=a$ for all $a\in S$, we have $u=\pm 1$. Thus $h(u)=h(v)$ if and only

if $u=\pm v$. Moreover, it clearly follows that inverse map $\sigma(l)$ is defined up to a $\pm$ sign; this inverse homomorphism is called the *spin representation* of $\mathfrak{L}_p$. The homomorphism $e\to\sigma[l(e)]$ is *spin representation* of $\mathfrak{E}_p$.

This particular representation of the group $\mathfrak{E}_p$ exemplifies the general notion of linear representation.

Definition. If G is a group, a representation of G is a homomorphism of G into a group of linear transformations in a linear vector or Hilbert space V; i.e., a set of mappings $v\to L(g)v$ such that $L(g_1g_2)v=L(g_1)L(g_2)v$. [We permit $L(g)$ to be defined up to sign.]

Two representations $L(g)$ in V and $L'(g)$ in V' are called *equivalent* if there exists a $1-1$ map M of V on V' such that $ML(g)M^{-1}=L(g)$. We write then $L\sim L'$.

If $L(g)$ is a representation of G in V and $L'(g)$ is a representation in V', then the representation $L''(g)$ in $V\oplus V'$ defined by $L''(g)(v\oplus v')=[L(g)v]\oplus[L'(g)v']$ is called the *direct sum* of L and L', and we write $L''=L\oplus L'$.

A representation is *irreducible* if it is not the direct sum of two other representations.

Theorem. Every finite-dimensional representation L of the proper Lorentz group $\mathfrak{L}_p$ or the Lorentz group $\mathfrak{L}$ may be written uniquely as a direct sum of irreducible representations.

The irreducible representations which appear in this decomposition are said to appear in L.

What are the irreducible representations? We quote the well-known answer.

Let $t_{\lambda_1\cdots\lambda_m\mu_1\cdots\mu_n}$ denote an arbitrary tensor defined for λ, $\mu=1, 2$, and symmetric in λ's and μ's separately; we may on occasion write this as $t(\lambda_1\cdots\lambda_m;\mu_1\cdots\mu_n)$. The collection of all such tensors evidently forms a linear space. For each $u=(u_\lambda{}^{\lambda'})$ in U, put $u_\lambda{}^{\lambda'}=u(\lambda';\lambda)$, let $\bar{u}(\lambda';\lambda)$ be the complex conjugate of $u(\lambda';\lambda)$, and put

$$[\tilde{L}(u)t](\lambda_1\cdots\lambda_m;\mu_1\cdots\mu_n)=t(\lambda_1'\cdots\lambda_m';\mu_1'\cdots\mu_n')\times u(\lambda_1';\lambda_1)\cdots\bar{u}(\mu_1';\mu_1)\cdots\bar{u}(\mu_n';\mu_n). \quad (2.9)$$

This evidently defines a linear representation of U; and if for each $l\in\mathfrak{L}_p$ we put $L(l)=\tilde{L}[\sigma(l)]$, we define a linear representation of $\mathfrak{L}_p$, which we will denote by the symbol $(m)\otimes(\bar{n})$.

Theorem. The representations $(m)\otimes(\bar{n})$ of $\mathfrak{L}_p$ are all irreducible, and any irreducible representation of $\mathfrak{L}_p$ is equivalent to one of them. If $m+n$ is even, the representation is single valued. If $m+n$ is odd, the representation is double valued.

We speak in these two cases of even or odd representations, or of "even spin" or "odd spin."

If $L(g)$ is a representation of G in an m-dimensional space V, then by introducing a basis in V we may think of $L(g)$ as being given by any set of matrices $R_\beta{}^\alpha(g)$, $1\le\alpha,\beta\le m$, such that $R_\beta{}^{\alpha'}(g)R_{\alpha'}{}^\alpha(g')=R_\beta{}^\alpha(gg')$. This enables us to introduce the two useful notions of *complex conjugate* and *Kronecker product* representations:

If the representation L is given by the matrices $R_\beta{}^\alpha(g)$, then $\bar{L}$ is the representation given by the matrices $\bar{R}_\beta{}^\alpha(g)$.

If the representations L_1 and L_2 are given by the matrices $R_{1,\beta_1}{}^{\alpha_1}(g)$ and $R_{2,\beta_2}{}^{\alpha_2}(g)$, then $L_1\otimes L_2$ is the representation given by the matrices $R_{1,\beta_1}{}^{\alpha_1}(g)R_{2,\beta_2}{}^{\alpha_2}(g)$. Having such a collection of matrices in mind, we may speak of an index α running between 1 and m as an index belonging to the representation L. These compositions between representations have various useful properties:

(i) $L\otimes(L_1\oplus L_2)\sim(L\otimes L_1)\oplus(L\otimes L_2)$

(ii) $L_1\otimes(L_2\otimes L_3)\sim(L_1\otimes L_2)\otimes L_3$

(iii) If $M=L\otimes L'$, then $\bar{M}=\bar{L}\otimes\bar{L}'$; a similar formula holding for the complex conjugate of the direct sum of two representations.

(iv) $L_1\otimes L_2\sim L_2\otimes L_1$.

We have $[(m)\otimes(\bar{0})]\otimes[(0)\otimes(\bar{n})]=(m)\times(\bar{n})$. Thus if we abbreviate by writing $(m)=(m)\times(\bar{0})$ and $(\bar{n})=(0)\times(\bar{n})$, we obtain a new notation consistent with our old notation.

The fundamental result on the product of two representations is the

Clesbsch-Gordan formula:

$$(m)\otimes(n)=\sum_{k=0}^{\min(m,n)}(m+n-2k).$$

By using this formula and the algebraic rules for direct sums and Kronecker products, any Kronecker product of the irreducible representations $(m)\times(\bar{n})$ of $\mathfrak{L}_p$ can be written as a direct sum of irreducible representations.

The "natural" irreducible representation $L_{\text{nat}}: x\to lx$ of $\mathfrak{L}_p$ as a group of transformations of the vectors $x\in E^4$ is equivalent to one of the complete set of irreducible representations $(m)\times(\bar{n})$. To which? As is well known, it is equivalent to $(1)\otimes(\bar{1})$. Thus, expressing the equivalence in terms of explicit matrices, there must exist matrices $\sigma_j{}^{\lambda\mu}$, $(\lambda,\mu=1, 2, j=1,\cdots,4,)$ such that if $u=\sigma(l)$, we have $\sigma_{j'}{}^{\lambda\mu}l_j{}^{j'}=\sigma_j{}^{\lambda'\mu'}u_{\lambda'}{}^\lambda u_{\mu'}{}^\mu$. These matrices are simply the matrices which give the (uniquely determined) map M of vectors V onto tensors $t=[t_{\lambda\mu}]$ such that $ML_{\text{nat}}=L_{(1)\times(\bar{1})}M$.

This set of four 2×2 matrices, whose abstract significance appears in the last formulas, is known as the set of 2×2 Pauli spin matrices.

To deduce the specific form of the 2×2 Pauli spin matrices, we may reason as follows. The representation $(1)\times(\bar{1})$ of U is the representation of U in the set $[t]=[t_{\lambda\mu}]$ of all 2×2 matrices determined by the formula

$$[L(u)t]_{\lambda\mu}=t_{\lambda'\mu'}u_\lambda{}^{\lambda'}\bar{u}_\mu{}^{\mu'}, \quad (2.11)$$

i.e., by the formula

$$L(u)=utu^*. \quad (2.12)$$

What we require then is a one-to-one linear mapping $t \leftrightarrow V$ of 2×2 Hermitian matrices t onto vectors V in E^4 such that

$$t \leftrightarrow V \text{ implies } \sigma(l)t\sigma(l)^* \leftrightarrow lV. \tag{2.13}$$

It is evident from the definition of the map σ [cf. formulas (2.10a) and (2.10b)] that this map is the map

$$V \leftrightarrow \begin{pmatrix} V_0+V_3 & V_1+iV_2 \\ V_1-iV_2 & V_0-V_3 \end{pmatrix}, \tag{2.14}$$

so that the 2×2 Pauli matrices are the matrices corresponding under the correspondence (2.14) to the vectors (1,0,0,0), (0,1,0,0), (0,0,1,0), and (0,0,0,1); that is, are the matrices

$$\begin{aligned} \{\sigma_0{}^{\lambda\mu}\} &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \{\sigma_1{}^{\lambda\mu}\} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \\ \{\sigma_2{}^{\lambda\mu}\} &= \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}, \quad \{\sigma_3{}^{\lambda\mu}\} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \end{aligned} \tag{2.15}$$

By $\partial^{\lambda\mu}$ we shall denote the first-order matrix of partial derivatives given by the formula

$$\partial^{\lambda\mu} = \sigma_j{}^{\lambda\mu}(\partial/\partial x j) \quad \text{(summation convention)}; \tag{2.16}$$

so that $\partial^{\lambda\mu}$ is the matrix

$$\begin{pmatrix} \dfrac{\partial}{\partial x^0}+\dfrac{\partial}{\partial x^3} & \dfrac{\partial}{\partial x^1}+i\dfrac{\partial}{\partial x^2} \\ \dfrac{\partial}{\partial x^1}-i\dfrac{\partial}{\partial x^2} & \dfrac{\partial}{\partial x^0}-\dfrac{\partial}{\partial x^3} \end{pmatrix} \tag{2.17}$$

according to formula (2.15). Where typographically necessary, we shall write this operator as $\partial(\lambda,\mu)$. Note that the transformation $x \to lx$ of the variables $x=(x_0,x_1,x_2,x_3)$ induces the corresponding spin-transformation $\partial^{\lambda\mu} \to \partial^{\lambda_1\mu_1}u_{\lambda_1}{}^{\lambda}\bar{u}_{\mu_1}{}^{\mu}$ of the "partial derivative" matrix $\partial^{\lambda\mu}$; here $u=\sigma(l)$. It is also convenient to introduce the spin-invariant "spinor-metric" matrices

$$\epsilon_{\lambda_1\lambda_2}, \quad \epsilon^{\lambda_1\lambda_2}, \quad \epsilon_{\mu_1\mu_2}, \quad \epsilon^{\mu_1\mu_2}$$

all given by the matrix

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{2.18}$$

Where typographically necessary, we shall write these matrices as $\epsilon(\lambda_1,\lambda_2)$, etc. (The "invariance" of these forms appears in such formulas as

$$\epsilon_{\lambda_1'\lambda_2'}u_{\lambda_1}{}^{\lambda_1'}u_{\lambda_2}{}^{\lambda_2'} = u_{\lambda_1}{}^{1}u_{\lambda_2}{}^{2} - u_{\lambda_1}{}^{2}u_{\lambda_2}{}^{1} = \epsilon_{\lambda_1\lambda_2},$$

valid for every 2×2 unimodular matrix $u_{\lambda}{}^{\lambda'}$.) In terms of these "metric forms" we shall "raise" and "lower" spinor indices λ and μ in the manner customary in tensor analysis. We note especially the formula

$$\begin{aligned} \epsilon_{\lambda\lambda'}\partial^{\lambda\mu}\partial^{\lambda'\mu'} &= \epsilon_{\mu\mu'}\Box, \\ \epsilon_{\mu\mu'}\partial^{\lambda\mu}\partial^{\lambda'\mu'} &= \epsilon^{\lambda\lambda'}\Box, \end{aligned} \tag{2.19}$$

which follow immediately on taking the determinant of the matrix (2.17) and which will be useful later.

By $\epsilon_{\lambda_1\cdots\lambda_n\lambda_1'\cdots\lambda_n'}$, we will denote the symmetrized sum

$$\epsilon_{\lambda_1\cdots\lambda_n\lambda_1'\cdots\lambda_n'} = \frac{1}{n!}\sum \epsilon(\lambda_1\lambda i_1')\epsilon(\lambda_2\lambda i_2')\cdots\epsilon(\lambda_n\lambda i_n') \tag{2.20}$$

taken over all permutations $(i_1\cdots i_n)$ of $(1\cdots n)$; the symbol $\epsilon_{\mu_1\cdots\mu_n\mu_1'\cdots\mu_n'}$ will be defined in an exactly corresponding way. By $\partial_{\lambda_1\cdots\lambda_n\mu_1\cdots\mu_n}$ we will denote the similar symmetrized sum

$$\partial_{\lambda_1\cdots\lambda_n\mu_1\cdots\mu_n} = \frac{1}{n!}\sum \partial(\lambda_1,\mu i_1)\partial(\lambda_2,\mu i_2)\cdots\partial(\lambda_n,\mu i_n). \tag{2.21}$$

Another useful fact about the representations $(m)\times(\bar{n})$ is the following: The representations of the type $(n)\times(\bar{n})$ (that is, those representations equivalent to their own complex-conjugates) may be written in another way. The representation $(n)\times(\bar{n})$ is the representation of $\mathfrak{L}_p$ by tensors $a^{i_1\cdots i_n}$ symmetric in $j_1\cdots j_n$ and satisfying $a^{i_1i_2\cdots i_n}G_{j_1j_2}=0$ (where $G_{j_1j_2}$ is the Lorentz metric form); these tensors are to transform according to the law

$$a^{i_1\cdots i_n} \to a^{i_1'\cdots i_n'}l_{j_1}{}^{i_1'}\cdots l_{j_n}{}^{i_n'},$$

$l\in\mathfrak{L}_p$ denoting a generic Lorentz transformation.

3. AXIOMATIC DEFINITION OF FREE FIELD THEORIES

After this hurried summary of some more or less familiar mathematical preliminaries, we introduce the mathematical axioms which are to define the physical theories forming the objects of our study. We shall take a relativistically invariant local quantum field theory to be defined by

(a) A set of linear operators $\psi_\alpha(x)$ (field strength operators) in a Hilbert space H, defined for all $x\in E^4$ and all α in a finite set of indices [the operators $\psi_\alpha(x)$ are actually not functions but distributions in their dependence on x; this technicality like many others will be ignored in what follows]. For $v\in H$, $[v,\psi_\alpha(x)v]$ signifies physically the expectation value of the α component of the field strength at the point x.

For notational convenience, we shall suppose that the fields $\psi_\alpha(x)^*$ are included among the fields $\psi_\alpha(x)$, so that for each α there exists an $\tilde{\alpha}=\beta$ such that

$$\psi_\alpha(x)^* = \psi_\beta(x).$$

(b) A finite-dimensional linear representation $R_\alpha{}^\beta(l)$ of the group $\mathfrak{L}_p$, and a unitary representation $U(e)$ of $\mathcal{E}_p$ in H, such that

$$R_\alpha{}^\beta[l(e)]\psi_\beta(e^{-1}x)=U(e)^*\psi_\alpha(x)U(e).$$

This axiom expresses the extended Lorentz invariance of our theories; that is, the homogeniety of space-time, in the usual way. We have here made the customary assumption that $R_\alpha{}^\beta$ depends only on the $\mathfrak{L}$ part of e; i.e., that translations do not call for change of field components.

According to a theorem stated previously, the representation $R_\alpha{}^\beta$ may be decomposed into the direct sum of irreducible representations; we will assume that this has been done, so that the *indices α are divided into various classes, and $R_\alpha{}^\beta=0$ if α and β belong to distinct classes.* Note that we do not assume that the representation $R_\alpha{}^\beta$ is irreducible itself, but only that it has been written as the sum of irreducible representations in the manner just stated.

(c) By Stone's theorem, the four-parameter group $U(x\rightarrow x+a)$ of unitary operators may be represented in terms of a set of four self-adjoint infinitesimal generators through the formula

$$U(x\rightarrow x+a)=\exp[i(a_0P^0-\cdots-a_4P^4)].$$

The set P^μ of self-adjoint operators is called the four-vector of momentum operators, and in particular P^0 is called the energy operator. We require (for physical reasons) that the energy operator be semidefinite, i.e., that $P^0\geq 0$. It follows at once by Lorentz invariance that we have $P^\mu x_\mu\geq 0$ for every positive timelike vector x.

Next we have an axiom which expresses at once the customary "law of causality" as it would apply not only to free but also to coupled fields, and expresses at the same time the "freeness" of the fields.

(d) For each pair α, β of indices there exists a sign $\sigma(\alpha,\beta)=\pm 1$ and a c number (scalar) valued function $f_{\alpha\beta}(x,x')$ such that

$$[\psi_\alpha(x),\psi_\beta(x')]_{\sigma(\alpha,\beta)}=f_{\alpha\beta}(x,x')$$

and such that $f_{\alpha\beta}(x,x')=0$ if $x-x'$ is a spacelike vector. It is also assumed that $\sigma(\alpha,\beta)=\sigma(\alpha',\beta)=\sigma(\alpha,\beta)$ if the indices α and α' belong to the same class and the indices β and β' belong to the same class.

Here we have written $[A,B]_{\pm 1}=[A,B]_\pm=AB\pm BA$, and shall write $[A,B]$ when one of $[A,B]_\pm$ is meant but it is desired not to specify which in particular.

The assumption that $[\psi_\alpha(x),\psi_\beta(x')]=0$ if $x-x'$ is spacelike is the ordinary "causality" assumption which applies as well to "coupled field" as to "free field" theories. The additional assumption contained in axiom (d) is the assumption that the function $[\psi_\alpha(x),\psi_\beta(x')]$ is a *scalar*-valued function. This very restrictive assumption defines the fields as *free* fields, from the point of view of the present paper.

In addition to the four explicitly stated "algebraic" axioms (a)–(d), it is well to note certain additional assumptions which will be made implicitly in the following analysis. In the first place, we suppose the operator-valued distributions $\psi_\alpha(x)$ to be *tempered* distributions. This assumption justifies the uses of the Fourier transform to be made in what follows. The carrier of the Fourier transform of the distributions $\psi_\alpha(x)$ is, by the Lorentz invariance of our field theory, a closed *Lorentz-invariant* set. We shall assume in what follows that this set is the union of a *finite* collection of the Lorentz-invariant hyperboloids Σ_m, $\Sigma_m{}^+$, and $\Sigma_m{}^-$. This corresponds to the assumption that the "mass spectrum" of the theory under consideration is *discrete.*

It is well to note that our aim is to uncover the algebraic structure of theories satisfying axioms (a)–(d), and that in consequence our reasoning will be formal and algebraic, rather than fully rigorous. We shall emphasize the algebraic side of the problem, and let the technical side fend for itself. There should, however, be no great difficulty in supplying missing technical axioms and deductions, so as to perfect this side of our arguments.

Having stated our four principal axioms we are in a position to begin our analysis.

4. DEFINITION OF EQUIVALENT THEORIES. STATEMENT OF RESULTS

It will be shown in the present paper that theories satisfying axioms (a)–(d) of the previous section can, after systematic utilization of simple algebraic transformations, be written in a uniquely determined canonical form. The following definitions, fundamental to the subsequent analysis, are intended to specify the precise nature of the allowed transformations.

Definition. Two free-field theories satisfying axioms (a)–(d), transforming according to representations $R_\alpha{}^{\alpha'}$ and $P_\beta{}^{\beta'}$ of the proper Lorentz group and with field operators ψ_α and ϕ_β will be called *equivalent* if

(i) ψ_α can be written as a Lorentz-invariant linear combination of the partial derivatives of ϕ_β with constant coefficients

$$\psi_\alpha=\Sigma_k c_{j_1\cdots j_k}{}^\beta{}_{,\alpha}\partial_{j_1}\cdots\partial_{j_k}\phi_\beta,$$

the coefficients c satisfying the scalar transformation law

$$c_{j_1\cdots j_k}{}^\beta{}_{,\alpha}=c_{j_1'\cdots j_k'}{}^{\beta'}{}_{,\alpha'}P_{\beta'}{}^\beta(l)R_\alpha{}^{\alpha'}(l)l_{j_1}{}^{i_1'}\cdots l_{j_k}{}^{i_k'}.$$

(ii) ϕ_β can be written in the same way in terms of ψ_α.

Were it not for the occurence of fields corresponding to particles of mass zero, then, allowing only transformations of the type described in the preceding definition, we could show that every field theory satisfying the "free-field" axioms described in the preceding section could be written in a drastically reduced canonical form. In treating fields "of mass zero" however, we need a looser definition than that given previously, which we state as follows.

Definition. Two free-field theories satisfying axioms (a)–(d), transforming according to representations $R_\alpha{}^{\alpha'}$ and $P_\beta{}^{\beta'}$ of the proper Lorentz group and with field operators ψ_α and φ_β are called semiequivalent if there exists a theory of the same sort with field operators η_γ such that

(i) the theory with operators ψ_α is equivalent to the theory with operators η_γ;

(ii) the fields φ_β can all be obtained by repeated application of the four-dimensional gradient operator to the fields η_γ.

The sole difference between "equivalence" and "semiequivalence" of $\{\psi_\gamma\}$ and $\{\varphi_\beta\}$ is that in the latter case, formulas expressing ψ_γ in terms φ_β must involve integrations as well as differentiations. In momentum space, this is a matter of small consequence, and the physical content (structure of spectrum, etc.) of two semiequivalent theories is identical.

Allowing transformations of the type just described, we shall show in the subsequent sections of the present paper that every field theory satisfying the axioms of Sec. 3 can be written in a canonical form. More precisely, we shall establish the following theorem.

Theorem. Let there be given a Lorentz-invariant quantized field theory of free fields, with field operators Ψ_α satisfying the axioms of Sec. 3, especially axioms (a)–(d). Then a second semiequivalent theory with field operators ψ_α satisfying these same axioms may be found, such that in addition:

(1) Every field component ψ_α satisfies a Klein-Gordon equation $\Box\psi_\alpha=m^2\psi_\alpha$, where $m=m(\alpha)\geq 0$.

(2) If $m(\alpha)\neq m(\beta)$, $[\psi_\alpha(x), \psi_\beta(x')]=0$.

(3) The field components ψ_α consist of a family of fields $\psi_{\lambda_1\cdots\lambda_n}{}^{(j)}$, $n=n(j)$, of the indicated Lorentz transformation, together with the Hermitian conjugates of these fields.

(4) The field components $\{\psi_\alpha\}=\{\psi_{\lambda_1\cdots\lambda_n}{}^{(j)}\}$ with $m(\alpha)=m>0$ all satisfy the reality condition.

$$\psi_{\mu_1\cdots\mu_n}{}^{*(j)}=m^{-n}\partial_{\mu_1}{}^{\lambda_1}\cdots\partial_{\mu_n}{}^{\lambda_n}\psi_{\lambda_1\cdots\lambda_n}{}^{(j)}.$$

Moreover, these same field components have commutation relations given by the formulas

(5) $[\psi_{\lambda_1\cdots\lambda_n}{}^{(j)}(x),\ \psi_{\lambda_1'\cdots\lambda_{n'}}{}^{(j')}(x')]=0$ if $j\neq j'$;

$$[\psi_{\lambda_1\cdots\lambda_n}{}^{(j)}(x), \psi_{\lambda_1'\cdots\lambda_{n'}}{}^{(j)}(x')]=i^{n-1}\epsilon_{\lambda_1\cdots\lambda_n,\,\lambda_1'\cdots\lambda_{n'}}D^{[m(\alpha)]}(x-x')$$

with $\sigma(\alpha,\tilde{\alpha})=\sigma(\alpha,\alpha)$ necessarily equal to $(-1)^{n+1}$.

(6) The field components $\{\psi_\alpha\}=\{\psi^{(j)}\}$ with $m(\alpha)=0$ transforming according to the representation $(0)\times(\bar{0})$ of the Lorentz group satisfy the reality condition $\psi^{(j)}=\psi^{(j)*}$, and have the commutation relations

$$[\psi^{(j)}(x), \psi^{(k)}(x')]_-\equiv 0 \quad \text{if} \quad j\neq k$$
$$[\psi^{(j)}(x), \psi^{(j)}(x')]_-=-iD^{(0)}(x-x').$$

(7) The field components $\{\psi_\alpha\}=\{\psi_{\lambda_1\cdots\lambda_n}{}^{(j)}\}$ with $m(\alpha)=0$ and $n>0$ all satisfy the first-order equation (transversality condition)

$$\partial_\mu{}^{\lambda_1}\psi_{\lambda_1\cdots\lambda_n}{}^{(j)}(x)\equiv 0,$$

and have the commutation relations

$$[\psi_{\lambda_1\cdots\lambda_n}{}^{(j)}(x), \psi_{\lambda_1'\cdots\lambda_{n'}}{}^{(k)}(x')]\equiv 0;$$
$$[\psi_{\lambda_1\cdots\lambda_n}{}^{(j)}(x), \psi_{\mu_1'\cdots\mu_{n'}}{}^{(k)*}(x')]\equiv 0$$

if $j\neq k$;

$$[\psi_{\lambda_1\cdots\lambda_n}{}^{(j)}(x), \psi_{\mu_1\cdots\mu_n}{}^{(j)*}(x')]=i^{n+1}\partial_{\lambda_1\cdots\lambda_n\mu_1\cdots\mu_n}D^{(0)}(x-x'),$$

with $\sigma(\alpha,\tilde{\alpha})=\sigma(\alpha,\alpha)$ necessarily equal to $(-1)^{n+1}$.

(8) All the field components ψ_α are linearly independent.

Thus, the "invariants" describing a theory written in canonical form are simply

(a) a set of masses (mass-spectrum);

(b) for each mass, a spin (spin-spectrum);

(c) for each pair of mass-spin combinations, a sign σ equal to ± 1 (specification of commutation or anticommutation relations).

Two theories for which (a)–(c) turn out to be the same when the theories are written in canonical form are semiequivalent. Since masses, spins, and the signs σ of (c) can be arbitrarily prescribed, we have before us a systematic and unredundant classification of all free-field theories.

5. SEPARATION OF MASSES

The general theorem of the preceding section will be proved by a succession of reductions; that is, by exhibiting a sequence of equivalent theories, each of which has more of the formal properties listed in this theorem than did the previous. In the present section, we take a first essential step, by showing that each field theory satisfying the axioms of Sec. 3 is equivalent to a theory in which each field component satisfies a wave or Klein-Gordon equation.

It is clear from Lorentz invariance that the function $f_{\alpha\beta}$ of axiom (d) depends only on $x-x'$. From the expression for $f^{(m)}$ given in formula (2.8) we have

$$[\psi_\alpha{}^{(m)}(x), \psi_\beta(x')]=f_{\alpha\beta}{}^{(m)}(x-x'). \tag{5.1}$$

Thus

$$[\psi_\alpha{}^{(m)}(x), \psi_\beta{}^{(m)}(x')]=f_{\alpha\beta}{}^{(m)}(x-x'). \tag{5.2}$$

while

$$[\psi_\alpha{}^{(m)}(x), \psi_\beta{}^{(m')}(x')]=0 \quad \text{if} \quad m\neq m'. \tag{5.3}$$

Since the mapping $f\to f^{(m)}$ is real, it follows that $(\psi_\alpha{}^{(m)})^*=\psi_\alpha{}^{(m)}$. Thus the collection $(\psi_\alpha)^{(m)}$, $(\psi_\alpha)_{(m)}$ of fields satisfies axioms (a)–(d). According to formula (2.6), the field ψ_α can be reconstructed as a linear combination from the fields $\psi_\alpha{}^{(m)}$ and $(\psi_\alpha)_{(m)}$. By formula (2.8), the fields $\psi_\alpha{}^{(m)}$ and $(\psi_\alpha)_{(m)}$ can be

written as Lorentz-invariant linear combinations with constant coefficients of the partial derivatives of ψ_α. Thus, the theory defined by the system of fields $(\psi_\alpha)^{(m)}$ and $(\psi_\alpha)_{(m)}$ is equivalent to the theory defined by the system of fields ψ_α. [It should be noted that in drawing this conclusion we have implicitly assumed that for each of our original field components α, $\psi_\alpha{}^{(m)}=0=(\psi_\alpha)_{(m)}$ except for a finite collection of numbers m. As was pointed out at the end of Sec. 3, this assumption must be regarded as an axiom supplementing axioms (a)–(d).]

Thus, upon passing from our original theory to an equivalent theory, we may assume without loss of generality that for each index α there exists a real number m and a positive integer k such that either

$$(\Box\psi_\alpha-m^2)^k\psi_\alpha=0, \quad m\geq 0 \tag{5.4}$$

or

$$(\Box\psi_\alpha+m^2)^k\psi_\alpha=0, \quad m>0; \tag{5.5}$$

the number m, the integer k and the case (5.4) or (5.5) depending only on the class of the index α.

Our next aim is to show that the integer k in (5.4) and (5.5) can be taken equal to 1, so that every field component satisfies an equation of Klein-Gordon type. To prove this, we require the following lemma.

Lemma. Suppose that a field-operator component $\psi_\alpha(x)$ satisfies

$$[\psi_\alpha(x),\psi_\beta(x')]_{\sigma(\alpha,\beta)}\equiv 0,$$

where $\beta=\tilde{\alpha}$. Then it follows that $\psi_\alpha(x)\equiv 0$.

Proof. First consider the case $\sigma(\alpha,\tilde{\alpha})=+1$. In this case, our hypothesis gives

$$\psi_\alpha(x)\psi_\alpha{}^*(x)+\psi_\alpha{}^*(x)\psi_\alpha(x)\equiv 0.$$

Since both terms in the preceding equation are positive-definite operators, it follows at once that $\psi_\alpha{}^*(x)\psi_\alpha(x)\equiv 0$, and hence that $\psi_\alpha(x)\equiv 0$.

Next consider the case $\sigma(\alpha,\tilde{\alpha})=-1$. Let $\psi_\alpha(p)$ denote the Fourier-transformed field

$$\psi_\alpha(p)=(2\pi)^{-2}\int e^{ix\cdot p}\psi_\alpha(x)d^4x.$$

It follows at once from Lorentz invariance [cf. the statements of axioms (b) and (c), Sec. 3] that

$$\exp(ia_\mu P^\mu)\psi_\alpha(p)\exp(-ia_\mu P^\mu)=\psi_\alpha(p)\exp(ia_\mu p^\mu).$$

Thus, differentiating with respect to a, we have

$$P^\mu\psi_\alpha(p)-\psi_\alpha(p)P^\mu=p^\mu\psi_\alpha(p). \tag{5.6}$$

Thus, if v_q is a vector in Hilbert space such that $P^\mu v_q=q^\mu v_q$, it follows that

$$P^\mu[\psi_\alpha(p)v_q]=(q^\mu+p^\mu)[\psi_\alpha(p)v_q]. \tag{5.7}$$

It follows in exactly the same way that if v_q is a vector in Hilbert space such that $P^\mu v_q=q^\mu v_q$, then

$$P^\mu[\psi_\alpha{}^*(p)v_q]=(q^\mu-p^\mu)[\psi_\alpha(p)v_q]. \tag{5.8}$$

Now, by axiom (c), it follows that the eigenvalues $[q^\mu]=[q^0,q^1,q^2,q^3]$ of the operators P^μ satisfy $q^0\geq 0$. It follows by Lorentz invariance that these eigenvalues $[q^0,q^1,q^2,q^3]$ always define a *positive timelike or null vector*. Thus, if p^μ points out of the positive light cone, it follows from (5.7) that $[\psi_\alpha(p)]^n v_q=0$ for n sufficiently large. On the other hand, since our hypothesis gives

$$\psi_\alpha(p)\psi_\alpha{}^*(p')-\psi_\alpha{}^*(p')\psi_\alpha(p)\equiv 0,$$

in the present case $\sigma(\alpha,\tilde{\alpha})=-1$, the operator $\psi_\alpha(p)$ is *normal*. Thus, it follows that $\psi_\alpha(p)v_q=0$ whenever p points out of the positive light cone. Similarly, $\psi_\alpha{}^*(p)v_q=0$ whenever p points out of the negative light cone. Since, for a normal operator N, $Nv=0$ if and only if $N^*v=0$, it follows that $\psi_\alpha(p)v_q=0$ whenever p points out of either the positive or the negative light cone. Since the eigenvectors v_q of the operators P^μ form a complete set, it follows that $\psi_\alpha(p)=0$ whenever $p\neq 0$. Thus $\psi_\alpha(x)\equiv 0$. Q.E.D.

By using this lemma, we may prove that if a field-component operator $\psi_\alpha(x)$ satisfies $(\Box-m^2)^k\psi_\alpha(x)\equiv 0$ for some positive integer k, then it satisfies $(\Box-m^2)\times\psi_\alpha(x)\equiv 0$. The proof is as follows. Consider the expression

$$[\psi_\alpha(x),\psi_\beta(x')]_{\sigma(\alpha,\beta)}.$$

By Lorentz invariance, this function of the variables x, x' depends only on $x-x'$, so that we may write

$$[\psi_\alpha(x),\psi_\alpha{}^*(x')]=f_{\alpha\beta}(x-x').$$

Suppose first that $k=2v$ is even. On applying the operator $(\Box-m^2)^k$ to $\psi_\alpha(x)$, it follows that $(\Box-m^2)^k\times f_{\alpha\beta}(x)\equiv 0$. On applying the operator $(\Box-m^2)^v$ to $\psi_\alpha(x)$ and $(\Box-m^2)^v$ to $\psi_\alpha{}^*(x')$, it follows that

$$\begin{aligned}[(\Box-m^2)^v\psi_\alpha(x),\ (\Box-m^2)^v\psi_\alpha{}^*(x')] &= (\Box_x-m^2)^v(\Box_{x'}-m^2)^v f_{\alpha\beta}(x-x')\\ &= (\Box-m^2)^{2v}f_{\alpha\beta}(x-x')\\ &=0.\end{aligned}$$

Thus, by the preceding lemma, $(\Box-m^2)^v\psi_\alpha(x)=0$.

What this means is that if the field component $\psi_\alpha(x)$ satisfies the equation $(\Box-m^2)^k\psi_\alpha(x)\equiv 0$, with k even, then it also satisfies the equation $(\Box-m^2)^{k/2}\psi_\alpha(x)\equiv 0$. It is an evident consequence of this that if the field component $\psi_\alpha(x)$ satisfies the equation $(\Box-m^2)^k\psi_\alpha(x)=0$ (k being either odd or even), then it also satisfies the equation $(\Box-m^2)^{k'}\psi_\alpha(x)\equiv 0$, where $k'=[(k+1)/2]$ is the greatest integer in $(k+1)/2$. Since $[(k+1)/2]<k$ unless $k=1$, it follows that if the field component $\psi_\alpha(x)$ satisfies the equation $(\Box-m^2)^k\psi_\alpha(x)=0$, it also satisfies the equation $(\Box-m^2)\psi_\alpha(x)=0$.

We may show in exactly the same way that if the field component $\psi_\alpha(x)$ satisfies the equation $(\Box+m^2)^k\times\psi_\alpha(x)=0$, it also satisfies the equation $(\Box+m^2)\psi_\alpha(x)=0$.

This concludes the proof of the principal assertion made previously.

The following lemma summarizes the facts established in the present section.

Lemma. Let there be given a Lorentz-invariant quantized field theory of free fields, with field operators satisfying the axioms of Sec. 3, especially axioms (a)–(d). Then there exists a second equivalent theory, with field operators ψ_α, satisfying these same axioms, and such that in addition every field operator ψ_α satisfies either

$$(\Box-m^2)\psi_\alpha(x)\equiv 0 \tag{5.9}$$

or

$$(\Box+m^2)\psi_\alpha(x)\equiv 0, \tag{5.10}$$

$m=m(\alpha)$ being some suitable real number.

Our next aim is to show that case (5.10) is impossible, and to find a more detailed canonical form for field components satisfying (5.9).

6. IMPOSSIBILITY OF IMAGINARY MASSES

In the present section we shall show that if a field component $\psi_\alpha(x)$ satisfies the equation $(\Box+m^2)\psi_\alpha(x)=0$ with $m>0$, then $\psi_\alpha(x)\equiv 0$.

This is most conveniently done as follows. We have observed repeatedly that if $\beta=\tilde{\alpha}$ the bracket $[\psi_\alpha(x), \psi_\beta(x')]$ has the form

$$[\psi_\alpha(x), \psi_\alpha^*(x')]=f_{\alpha\beta}(x-x').$$

On introducing the Fourier-transformed field

$$\psi_\alpha(p)=(2\pi)^{-2}\int e^{ix\cdot p}\psi_\alpha(x)d^4x, \tag{6.1}$$

so that

$$\psi_\alpha^*(p)=(2\pi)^{-2}\int e^{-ix\cdot p}\psi_\alpha^*(x)dx, \tag{6.2}$$

it follows that

$$\begin{aligned}
&[\psi_\alpha(p), \psi_\alpha^*(p')]\\
&=(2\pi)^{-4}\int e^{ix\cdot p}e^{-ix'\cdot p'}[\psi_\alpha(x), \psi_\alpha^*(x')]d^4xd^4x'\\
&=(2\pi)^{-4}\int e^{ix\cdot p}e^{-ix'\cdot p'}f_{\alpha\beta}(x-x')d^4xd^4x'\\
&=(2\pi)^{-4}\int e^{i(x-x')\cdot p}e^{ix'\cdot(p-p')}f_{\alpha\beta}(x-x')d^4xd^4x'\\
&=f_{\alpha\beta}(p)\delta(p-p').
\end{aligned} \tag{6.3}$$

Note that since $(\Box+m^2)\psi_\alpha(x)=0$, $\psi_\alpha(p)$ vanishes except for p on the (spacelike) hyperboloid $\{p\,|\,p^2+m^2=0\}$. Since $[\psi_\alpha(p), \psi_\alpha^*(p')]$ is Hermitian, the function $f_{\alpha\beta}(p)$ is real. If $\psi_\alpha(x)\neq 0$, we can choose p such that $\psi_\alpha(p)\neq 0$.

First consider the case $\sigma(\alpha,\tilde{\alpha})=+1$. In this case, since $\psi_\alpha(p)\psi_\alpha^*(p)+\psi_\alpha(p)^*\psi_\alpha(p)$ is a positive definite operator, we have $f_{\alpha\beta}(p)>0$.

By Lorentz invariance [cf. the statements of axioms (b) and (c), Sec. 3] we have

$$\exp(ia_\mu P^\mu)\psi_\alpha(p)\exp(-ia_\mu P^\mu)=\psi_\alpha(p)\exp(ia_\mu p^\mu). \tag{6.4}$$

Thus differentiating with respect to a we have

$$P^\mu\psi_\alpha(p)-\psi_\alpha(p)P^\mu=p^\mu\psi_\alpha(p). \tag{6.5}$$

Thus, if v_q is a vector in Hilbert space such that $P^\mu v_q=q^\mu v_q$, it follows that

$$P^\mu[\psi_\alpha(p)v_q]=(q^\mu+p^\mu)[\psi_\alpha(p)v_q]. \tag{6.6}$$

Since by axiom (c) the eigenvalues q^μ of the operators P^μ satisfy $q^0\geq 0$, it follows that $\psi(p)v_q=0$ if $(q+p)^0<0$. Similarly, $\psi(p)^*v_q=0$ if $(q-p)^0<0$. By Lorentz invariance, $\psi_\alpha(p)v_q=0$ if the vector $q+p$ can be mapped by a proper Lorentz transformation $l\in\mathfrak{L}_p$ onto a four-vector r such that $r^0<0$. In particular, $\psi_\alpha(p)v_q=0$ if $p+q$ is spacelike. Similarly, $\psi_\alpha(p)^*v_q=0$ if $p-q$ is spacelike. For any q, there exists a p on the spacelike hyperboloid $p_\mu p^\mu=-m^2$ such that *both* $p+q$ and $p-q$ are spacelike. Thus, there exists a p such that $\psi_\alpha(p)v_q=\psi_\alpha(p)^*v_q=0$. Since then $0=[\psi_\alpha(p)\psi_\alpha^*(p)+\psi_\alpha^*(p)\times\psi_\alpha(p)]v_q=f_{\alpha\beta}(p)\delta(0)v_q$ and $f_{\alpha\beta}(p)>0$ has been shown previously, we have a contradiction. This contradiction proves the desired result in case $\sigma(\alpha,\tilde{\alpha})=+1$.

Next consider the case $\sigma(\alpha,\tilde{\alpha})=-1$, and again let $\beta=\tilde{\alpha}$. Here we will consider the cases $f_{\alpha\beta}(p)=0$, $f_{\alpha\beta}(p)>0$, and $f_{\alpha\beta}(p)<0$ separately.

If $f_{\alpha\beta}(p)>0$, we reason as follows. If v is a vector such that $\psi_\alpha^*(p)v=0$, we have

$$\begin{aligned}
\|\psi_\alpha^*(p)v\|^2&=[\psi_\alpha(p)\psi_\alpha(p)^*v,v]\\
&=[\psi_\alpha^*(p)\psi_\alpha(p)v,v]+f_{\alpha\beta}(p)\delta(0)(v,v)\\
&=\|\psi_\alpha(p)v\|^2+\delta(0)f_{\alpha\beta}(p)\|v\|^2,
\end{aligned}$$

so that $v=0$.

Now, the vector $p=[p_0,p_1,p_2,p_3]$ lies on the spacelike hyperboloid $p^2=-m^2$. It follows [cf. formula (6.6)] that if v_q is a nonzero vector in Hilbert space such that $P^\mu v_q=q^\mu v_q$, then

$$P^\mu[\psi_\alpha^*(p)v_q]=(q^\mu-p^\mu)[\psi_\alpha(p)v_q]. \tag{6.7}$$

Now, by axiom (c), it follows that the eigenvalues $[q^\mu]=[q^0,q^1,q^2,q^3]$ of the operators P^μ satisfy $q^0\geq 0$. Therefore, by Lorentz invariance, these eigenvalues define a positive timelike or null vector. Since the vector p points out of the light cone, it follows that $[\psi_\alpha^*(p)]^k v_q=0$ for k sufficiently large. By the foregoing $v_q=0$. This contradiction shows that $f_{\alpha\beta}(p)>0$ is impossible. Similarly, $f_{\alpha\beta}(p)<0$ may be shown to be impossible. Hence $f_{\alpha\beta}(p)\equiv 0$. But then, by the first lemma proved in the preceding section, $\psi_\alpha(x)\equiv 0$.

This proves that if $(\Box+m^2)\psi_\alpha(x)=0$, and $m>0$, then $\psi_\alpha(x)=0$.

Alternative (5.10) of the final lemma stated in Sec. 5 is consequently impossible; so that in our subsequent analysis we have only to consider sets of

fields satisfying either the equation

$$\Box\psi_\alpha(x)=0 \tag{6.7}$$

or

$$(\Box-m^2)\psi_\alpha(x)=0, \quad \text{where} \quad m=m(\alpha)>0. \tag{6.8}$$

Field components satisfying (6.8) will be called *fields of positive mass*; field components satisfying (6.7) will be called *fields of zero mass*.

Our next aim is to find a detailed canonical form for our fields. We shall first analyze the somewhat simpler case (6.8) of fields of positive mass, and then analyze the fields of zero mass. But before any of this is done, we shall need some results on the sign $\sigma(\alpha,\gamma)$ (related to the theorem of Pauli on spin and statistics). The next section is devoted to establishing these results.

7. LEMMAS ON THE SIGN $\sigma(\alpha,\gamma)$

It has been observed that the scalar-valued function $[\psi_\alpha(x),\psi_\beta(x')]$ depends only on the difference $x-x'$. We may consequently write

$$[\psi_\alpha(x),\psi_\gamma(x')]=f_{\alpha\gamma}(x-x').$$

If we pass to the Fourier-transformed field,

$$\psi_\alpha(p)=(2\pi)^{-2}\int e^{ix\cdot p}\psi_\alpha(x)d^4x, \tag{7.1}$$

so that, writing $\beta=\tilde{\alpha}$, we have

$$\psi_\alpha^*(p)=(2\pi)^{-2}\int e^{-ix\cdot p}\psi_\alpha^*(x)dx$$

$$=(2\pi)^{-2}\int e^{-ix\cdot p}\psi_\beta(x)d^4x, \tag{7.2}$$

it follows [cf. formula (6.3)] that, writing $\nu=\tilde{\beta}$, we have

$$\begin{aligned}[\psi_\alpha(p),\psi_\gamma^*(p')]&=f_{\alpha\gamma}(p)\delta(p-p')\\ [\psi_\alpha(p),\psi_\gamma(p')]&=f_{\alpha\gamma}(p)\delta(p+p')\\ [\psi_\alpha(p),\psi_\alpha^*(p)]&=f_{\alpha\beta}(p)\delta(p-p'),\end{aligned} \tag{7.3}$$

where

$$f_{\alpha\gamma}(p)=(2\pi)^{-2}\int e^{ix\cdot p}f_{\alpha\gamma}(x)d^4x.$$

Our main aim in the present section is to prove that $[\psi_\alpha(x),\psi_\gamma(x')]\not\equiv 0$, then $\sigma(\alpha,\eta)=\sigma(\gamma,\eta)$ for every field component $\psi_\eta(x)$. To do this, we shall first establish the following.

Lemma. If $[\psi_\alpha(x),\ \psi_\nu(x')]\not\equiv 0$, where $\nu=\tilde{\gamma}$, then $\sigma(\alpha,\tilde{\gamma})=\sigma(\alpha,\tilde{\alpha})$.

Proof. On passing to the Fourier-transformed fields, we have (7.3).

First consider the case $\sigma(\alpha,\tilde{\alpha})=+1$. Then clearly $\psi_\alpha(p)\psi_\alpha^*(p)+\psi_\alpha^*(p)\psi_\alpha(p)>0$, so that by (7.3), if $\beta=\tilde{\alpha}$, we have $f_{\alpha\beta}(p)>0$ for each p. Moreover, by (7.3) $\psi_\alpha(p)^2=0$. Choose some particular p such that $f_{\alpha\nu}(p)\neq 0$. Then, putting $A=\psi_\alpha(p)/[f_{\alpha\beta}(p)\delta(0)]^{\frac{1}{2}}$, we have $A^2=0$, $AA^*+A^*A=I$. Hence, by a well-known result, we may write Hilbert space as an orthogonal direct sum $H\oplus H$ of two replicas of the same space H, in such a way that when this is done A is represented by the matrix

$$A=\begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix}. \tag{7.4}$$

Suppose that $\sigma(\alpha,\tilde{\gamma})=-1$. Then, putting $B=\psi_\gamma(p)$, we have $[A,B]_-=0$ by (7.3), and $[A,B^*]_-=\tau I$, where $\tau=[\delta(0)]^{\frac{1}{2}}f_{\alpha\nu}(p)/[f_{\alpha\beta}(p)]^{\frac{1}{2}}$. Since $[A,B]_-=0$, it follows that B is represented by a matrix

$$B=\begin{pmatrix}a & 0\\ b & a\end{pmatrix}. \tag{7.5}$$

Then $[A,B^*]_-$ is represented by the matrix

$$\begin{pmatrix}-b & 0\\ 0 & +b\end{pmatrix}, \tag{7.6}$$

and hence can only be equal to τI if $\tau=0$. This shows that our assumptions imply $f_{\alpha\nu}(p)=0$; and this contradiction proves the present lemma in case $\sigma(\alpha,\tilde{\alpha})=+1$.

Next consider the case $\sigma(\alpha,\tilde{\alpha})=-1$. Choose some particular p such that $f_{\alpha\nu}(p)\neq 0$, where as before $\nu=\tilde{\gamma}$. Here we shall consider the two subcases $f_{\alpha\beta}(p)=0$ and $f_{\alpha\beta}(p)\neq 0$ separately. First suppose that $f_{\alpha\beta}(p)\neq 0$. Suppose for the sake of definiteness that $f_{\alpha\beta}(p)>0$; the proof in the case $f_{\alpha\beta}(p)<0$ is similar. Then, putting $A=\psi_\alpha(p)/[f_{\alpha\beta}(p)\delta(0)]^{\frac{1}{2}}$, we have $AA^*-A^*A=I$. Hence, by a well-known result our Hilbert space may be represented as an infinite direct sum $H\oplus H\oplus H\oplus\cdots$ of a sequence of replicas of the same space H, in such a way that when this is done A is represented by the matrix

$$A=\begin{bmatrix}0 & 0 & 0 & \cdots\\ 1 & 0 & 0 & \cdots\\ 0 & \sqrt{2} & 0 & \cdots\\ 0 & 0 & \sqrt{3} & \cdots\\ \cdots & \cdots & \cdots & \cdots\end{bmatrix}. \tag{7.7}$$

Let J_1 and K denote the operators defined by the matrices

$$J=\begin{bmatrix}1 & 0 & 0 & 0 & \cdots\\ 0 & -1 & 0 & 0 & \cdots\\ 0 & 0 & 1 & 0 & \cdots\\ 0 & 0 & 0 & -1 & \cdots\\ \cdots & \cdots & \cdots & \cdots & \cdots\end{bmatrix}; \tag{7.8}$$

$$K=\begin{bmatrix}(0!)^{\frac{1}{2}} & 0 & 0 & \cdots\\ 0 & (1!)^{\frac{1}{2}} & 0 & \cdots\\ 0 & 0 & (2!)^{\frac{1}{2}} & \cdots\\ \cdots & \cdots & \cdots & \cdots\end{bmatrix}. \tag{7.9}$$

Then clearly $A=KA_1K^{-1}$, where the operator A_1 is

defined by the matrix

$$A_1=\begin{pmatrix} 0 & 0 & 0 & \cdots \\ 1 & 0 & 0 & \cdots \\ 0 & 1 & 0 & \cdots \\ 0 & 0 & 1 & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{pmatrix}. \tag{7.10}$$

Now, on operator C_1 commuting with A_1 is readily seen by examination of its matrix to have the form

$$C_1=\begin{pmatrix} c_0 & 0 & 0 & 0 & \cdots \\ c_1 & c_0 & 0 & 0 & \cdots \\ c_2 & c_1 & c_0 & 0 & \cdots \\ c_3 & c_2 & c_1 & c_0 & \cdots \\ \cdots & & & & \end{pmatrix}, \tag{7.11}$$

i.e., to have the form $C_1=c_0+c_1A_1+c^2A_1{}^2+\cdots$, where c_0, c_1, c_2, $\cdots$ etc., are linear transformations in H. It follows immediately from the formula $A=KA_1K^{-1}$ that an operator C_k commuting with A has the form $C=c_0I+c_1A+c_2A^2+\cdots$, where c_0, c_1, c_2, $\cdots$ etc., are linear transformations in H. It is clear that $A_1J=-JA_1$. Hence, since $J=KJK^{-1}$, $AJ=-JA$. Consequently, an operator anticommuting with A has the form $c_0J+c_1JA+c_2JA^2+\cdots$.

Suppose now that $\sigma(\alpha,\gamma)=+1$. Then putting $B=\psi_\gamma(p)$, we have $[A,B]_+=0$ by (7.3) and $[A,B]_+{}^*=\tau I$, where $\tau=[\delta(0)]^{\frac{1}{2}}f_{\alpha\nu}(p)/[f_{\alpha\alpha}(p)]^{\frac{1}{2}}$; note that we continue to write $\nu=\tilde{\gamma}$. Thus B anticommutes with A and has the form $B=c_0J+c_1JA+c_2JA^2+\cdots$. Hence

$$[A,B^*]_+=\sum_{k=0}^{\infty} c_k[A,(A^*)^kJ^*]_+.$$

Since $J=J^*$, this may be written as

$$[A,B^*]_+=\sum_{k=0}^{\infty} c_k[A(A^*)^kJ+(A^*)^kJA]$$

$$=\sum_{k=0}^{\infty} c_k[A,(A^*)^k]_-J$$

$$=\sum_{k=0}^{\infty} kc_k(A^*)^{k-1}J.$$

It is clear from the equation $(A^*)^j=K^{-1}(A_1{}^*)^jK$ that none of the matrices in this last sum but the matrix C_1J have nonzero diagonal elements. And then from the form of the matrix J, it is clear that this sum cannot be a multiple of the identity matrix unless it is identically zero. Thus, if $[A,B^*]_+=\tau I$, it follows that $\tau=0$. This shows that our assumptions imply $f_{\alpha\nu}(p)=0$; and this contradiction proves the present lemma if $\sigma(\alpha,\tilde{\alpha})=-1$ and $f_{\alpha\beta}(p)\neq 0$, where $\beta=\tilde{\alpha}$.

If $\sigma(\alpha,\beta)=-1$ and $f_{\alpha\beta}(p)=0$, where $\beta=\tilde{\alpha}$, it follows from (7.3) that the operator $\psi_\alpha(p)$ is normal. Let $A=\psi_\alpha(p)$, and, supposing that $\sigma(\alpha,\gamma)=+1$, put $B=\psi_\gamma(p)$, so that $B^*=\psi_\nu(p)$, where $\nu=\tilde{\gamma}$. Then by (7.3) we have $[A,B]_+=0$ and $[A,B^*]_+=\tau I$, where $\tau=\delta(0)f_{\alpha\nu}(p)$. The operator A is normal. Let v_λ be an eigenvector of A, belonging to the eigenvalue λ. Then clearly $A(Bv_\lambda)=-BAv_\lambda=-\lambda(Bv_\lambda)$. Hence Bv_λ is an eigenvector of A belonging to the eigenvalue $-\lambda$. This shows that, if $E(e)$ denotes the spectral projection of the normal operator A corresponding to the subset e of the plane of λ, then $BE(e)=E(-e)B$. Since the projections $E(e)$ and $E(-e)$ are self-adjoint, it follows that $B^*E(e)=E(-e)B^*$. Since $A=\int\lambda E(d\lambda)$ by the spectral resolution theorem, it follows that $B^*A=-AB^*$. Thus $[B^*,A]_+=0$, so that our assumptions imply $f_{\alpha\nu}(p)=0$. This contradiction proves our lemma in the final case $\sigma(\alpha,\tilde{\alpha})=-1$, $f_{\alpha\beta}(p)=0$. Q.E.D.

Corollary. If $[\psi_\alpha(x),\psi_\gamma(x')]_{\sigma(\alpha,\gamma)}\not\equiv 0$, then $\sigma(\alpha,\gamma)=\sigma(\alpha,\tilde{\alpha})=\sigma(\gamma,\tilde{\gamma})$.

Proof. The first equation follows if we put $\tilde{\gamma}$ for γ in the preceding lemma. Since $\sigma(\alpha,\gamma)=\sigma(\gamma,\alpha)$, the second equation follows by symmetry. Q.E.D.

We may now generalize the preceding corollary, to put our result on the sign $\sigma(\alpha,\gamma)$ into its most complete form. This is done in the following lemma. The reader should note that in stating the lemma and in its proof we write $\beta=\tilde{\alpha}$.

Lemma. If $[\psi_\alpha(x),\psi_\nu(x')]_{\sigma(\alpha,\gamma)}\not\equiv 0$, where $\nu=\tilde{\gamma}$, then $\sigma(\alpha,\gamma)=\sigma(\alpha,\tilde{\alpha})=\sigma(\gamma,\tilde{\gamma})=\sigma(\alpha,\eta)=\sigma(\beta,\eta)$ for every field component $\psi_\alpha(x)$ which is not identically zero.

Proof. Suppose in contradiction of what we wish to establish that ψ_η is a field component such that $\sigma(\alpha,\eta)\neq\sigma(\gamma,\eta)$. For the sake of definiteness we may assume that $\sigma(\alpha,\eta)\neq\sigma(\alpha,\tilde{\alpha})$, $\sigma(\gamma,\eta)=\sigma(\gamma,\tilde{\gamma})$. First consider the case $\sigma(\alpha,\tilde{\alpha})=+1$. Then, by the above lemma and corollary, $[\psi_\eta(x),\psi_\alpha(x')]\equiv 0$ and $[\psi_\eta(x),\psi_\beta(x)]\equiv 0$, $\psi_\eta(x)$ commutes with $\psi_\alpha(x)$ and with $\psi_\alpha{}^*(x)$. (We have as before, written $\beta=\tilde{\alpha}$.) Again by the previous lemma and corollary, $\sigma(\alpha,\gamma)=+1$. Put $A=\psi_\alpha(p)/[f_{\alpha\beta}(p)\times\delta(0)]^{\frac{1}{2}}$, [cf. formula (31) and note that as in the second sentence of the proof of the preceding lemma, $f_{\alpha\beta}(p)>0$]. Put $B=\psi_\gamma(p)$, $C=\psi_\eta(p')$, where p is chosen so that $AB^*+B^*A=\tau I$ with τ different from zero. Then it follows exactly as in the proof of the preceding lemma that we may write Hilbert space as an orthogonal direct sum $H\oplus H$ of two replicas of the same space H in such a way that when this is done A, B, and C are represented by matrices

$$A=\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

$$B=\begin{pmatrix} a & 0 \\ b & -a \end{pmatrix}$$

$$C=\begin{pmatrix} c & 0 \\ 0 & c \end{pmatrix}.$$

Since $AB^*+B^*A=\tau I$, we have $b=\tau I$. Since $\sigma(\gamma,\eta)$

$=+1$, $BC+CB$ is a multiple of the identity. But

$$BC+CB=\begin{pmatrix}(ac+ca) & 0\\ 2\tau c & -(ac+ca)\end{pmatrix}$$

can only be a multiple of the identity if $c=0$. Thus $\psi_\eta(p')=0$ for all p', so $\psi_\eta(x)=0$. This contradiction proves our assertion in case $\sigma(\alpha,\tilde{\alpha})=+1$.

Next consider the case $\sigma(\alpha,\tilde{\alpha})=-1$. Then, by the preceding lemma, ψ_η anticommutes with $\psi_\alpha(x)$ and $\psi_\alpha{}^*(x)$. Again by the same lemma, $\sigma(\alpha,\gamma)=-1$. Choose $\bar{\tau}$ so that $\psi_\alpha(p)\psi_\gamma{}^*(p)-\psi_\gamma(p)^*\psi_\alpha(p)=\bar{\tau}I$ with $\bar{\tau}\neq 0$. We shall consider the two cases $f_{\alpha\beta}(p)\neq 0$ and $f_{\alpha\beta}(p)=0$ separately.

If $f_{\alpha\beta}(p)\neq 0$, we may suppose for the sake of definiteness that $f_{\alpha\beta}(p)>0$; the proof in the case $f_{\alpha\beta}(p)<0$ is similar. Put $A=\psi_\alpha(p)/[f_{\alpha\beta}(p)\delta(0)]^{\frac{1}{2}}$, $B=\psi_\gamma(p)$, $C=\psi_\eta\times(p')$. It follows exactly as in the proof of the preceding lemma that we may write Hilbert space as an orthogonal direct sum $H\oplus H\oplus H\oplus\cdots$ of a sequence of replicas of the same space H in such a way that when this is done, A is represented by the matrix (7.7), and B and C have the form $B=b_0I+b_1A+b_2A^2+\cdots$, $C=c_0J+c_1JA+c_2JA^2+\cdots$; b_0, b_1, $\cdots$, c_0, c_1, $\cdots$, etc., being linear mappings of H into itself. Since $0=CA^*+A^*C=c_1J+2c_2JA+3c_3JA^2+\cdots$, it follows that $c_1=c_2=c_3=\cdots=0$, so that $C=c_0J$. Since $AB^*-B^*A=\tau I$, we have $\tau I=b_1I+2b_2A+3b_3A^2+\cdots$; thus $B=b_0I+\tau A$. Hence, $BC-CB=2\tau c_0JA$ can only be a multiple of the identity if $c_0=0$. Thus $\psi_\eta(p')=0$ for all p', so $\psi_\eta(x)=0$. This contradiction proves our assertion in case $\sigma(\alpha,\tilde{\alpha})=-1$, so that our assertion is proved in case $f_{\alpha\beta}(p)\neq 0$.

In case $\sigma(\alpha,\tilde{\alpha})=-1$, and $f_{\alpha\beta}(p)=0$, where, as before, $\beta=\tilde{\alpha}$, the operator $\psi_\alpha(p)$ is normal. We then can show, exactly as in the final paragraph of the proof of the preceding lemma, that $f_{\alpha\nu}(p)=0$, where, as before, $\nu=\tilde{\gamma}$, contrary to our assumption; and this proves our lemma in the sole remaining case. Thus, our lemma is generally proved. QED.

In stating the axioms of Sec. 3, we required that the indices α should be divided into various classes, each of which transforms according to a given irreducible representation of the Lorentz group, and that $\sigma(\alpha,\beta)=\sigma(\alpha',\beta')$ if α, α' and β, β' belong to the same class. Now that we have proved the preceding lemma, it is seen that if we put two β, β' indices into the same *family* whenever $\sigma(\alpha,\beta)=\sigma(\alpha,\beta')$ for all β', then $[\psi_\beta(x),\psi_{\beta'}(x')]\equiv 0$ if β and β' belong to distinct families.

On summarizing the work of this section and the previous sections, we see that we have proved the following lemma:

Lemma. Let there be given a Lorentz-invariant quantized field theory of free fields, with field operators satisfying the axioms of Sec. 3. Then there exists a second equivalent theory with field operators ψ_α, satisfying these same axioms, and such that in addition:

(i) Every field operator ψ_α satisfies $(\Box-m^2)\psi_\alpha(x)\equiv 0$, $m=m(\alpha)$ being some suitable real number.

(ii) The indices α may be subdivided into *families*, which may in turn be subdivided into *classes*, such that

(a) $\sigma(\alpha,\beta)$ depends only on the families of α and β;

(b) α and $\tilde{\alpha}$ always belong to the same family;

(c) Each $[\psi_\alpha(x),\psi_\beta(x')]\equiv 0$ unless α and β belong to the same family;

(d) Each class of indices transforms according to some certain irreducible representation of the proper homogeneous Lorentz group.

8. DETAILED ANALYSIS OF THE FIELDS OF POSITIVE MASS

We are now in a position to complete our analysis. In the present section we study field components of positive mass; field components of zero mass will be analyzed in a subsequent section. We consider a field theory satisfying the axioms of Sec. 3, and having the additional properties noted in the final lemma of the preceding section. By Lorentz-invariance, the scalar-valued function $[\psi_\alpha(x),\psi_\beta(x')]_{\sigma(\alpha,\beta)}$ must have the form

$$[\psi_\alpha(x),\psi_\beta(x')]=f_{\alpha\beta}(x-x'). \tag{8.1}$$

If $[\Box-m(\alpha)^2]\psi_\alpha(x)\equiv 0$ and $[\Box-m(\beta)^2]\psi_\beta(x)\equiv 0$, then clearly

$$[\Box-m(\alpha)^2]f_{\alpha\beta}(x)\equiv 0\equiv[\Box-m(\beta)^2]f_{\alpha\beta}(x).$$

Thus $[\psi_\alpha(x),\psi_\beta(x')]\equiv 0$ unless $m(\alpha)=m(\beta)$. We may consequently suppose without loss of generality that all the fields ψ_α whose indices α lie in a given one of the *families* of the final lemma of the preceding section satisfy $(\Box-m^2)\psi_\alpha(x)=0$ with a value of m depending only on the family to which α belongs.

In the present section, we analyze a given family for which $m>0$, and, unless the contrary is explicitly specified, consider only indices α, β, etc., which lie in this particular family.

We have then $(\Box-m^2)f_{\alpha\beta}(x)\equiv 0$.

Thus, the function $f_{\alpha\beta}(x)$ is determined by its value and the value of its normal derivative on any plane $t=$ constant. By our axioms, $f_{\alpha\beta}(x)=0$ if x is spacelike. Thus $f_{\alpha\beta}(\mathbf{x},0)=(\partial/\partial t)f_{\alpha\beta}(\mathbf{x},0)=0$ for $\mathbf{x}\neq 0$. By a well-known theorem in the theory of distributions,

$$\begin{aligned} f_{\alpha\beta}(\mathbf{x},0)&=\Lambda_{\alpha\beta}\left(\frac{\partial}{\partial x}\right)\delta(\mathbf{x});\\ \frac{\partial}{\partial t}f_{\alpha\beta}(\mathbf{x},0)&=\tilde{\Lambda}_{\alpha\beta}\left(\frac{\partial}{\partial x}\right)\delta(\mathbf{x}),\end{aligned} \tag{8.2}$$

where $\Lambda_{\alpha\beta}$ and $\tilde{\Lambda}_{\alpha\beta}$ are partial differential operators of finite order with constant coefficients in the *space*

derivatives $\partial/\partial x$. Since $\Box f_{\alpha\beta}=m^2 f_{\alpha\beta}$, it follows that

$$f_{\alpha\beta}(x)=\Lambda_{\alpha\beta}\left(\frac{\partial}{\partial x}\right)D_1^{(m)}(x)+\tilde{\Lambda}_{\alpha\beta}\left(\frac{\partial}{\partial x}\right)D^{(m)}(x), \quad (8.3)$$

where $D_1^{(m)}$ and $D^{(m)}$ are solutions of the equation $D=m^2D$ with initial condition

$$D_1^{(m)}(\mathbf{x},0)=\delta(\mathbf{x}); \quad \frac{\partial}{\partial t}D_1^{(m)}(\mathbf{x},0)=0$$
$$D^{(m)}(\mathbf{x},0)=0; \quad \frac{\partial}{\partial t}D^{(m)}(\mathbf{x},0)=\delta(\mathbf{x}). \quad (8.4)$$

Since evidently $D_1^{(m)}=(\partial/\partial t)D^{(m)}$, it follows that $f_{\alpha\beta}$ may be written in the form

$$f_{\alpha\beta}(x)=(a_{\alpha\beta}+a_{\alpha\beta}{}^{j}\partial_j+a_{\alpha\beta}{}^{j_1j_2}\partial_{j_1}\partial_{j_2}+\cdots)D^{(m)}(x), \quad (8.5)$$

where the matrices of coefficients $a^{j_1\cdots j_k}$ may be taken to be symmetric in $j_1\cdots j_k$, and, since $D^{(m)}=m^2D^{(m)}$, to satisfy $a_{\alpha\beta}{}^{j_1j_2\cdots j_k}G_{j_1j_2}=0$. Here, $G_{j_1j_2}$ is the Lorentz-metric form $G_{j_1j_2}=0$ for $j_1\neq j_2$, $G_{0,0}=1$, $G_{ii}=-1$, $i=1, 2, 3$.

Let $\{\psi_\alpha\}$ be a class [cf. the axioms of Sec. 3, especially axiom (b), second paragraph] of components of our quantized fields, the index α belonging to the irreducible representation $(m)\times(\bar{n})$ of the Lorentz group. For explicitness, we shall write the class of indices $\{\alpha\}$ in the form $\{\alpha\}=\{\lambda_1\cdots\lambda_k\mu_1\cdots\mu_j\}$, $\lambda, \mu=1, 2$; and in the same way write the field components ψ_α as $\psi_{\lambda_1\cdots\lambda_k\mu_1\cdots\mu_j}$. When we adopt this notation, the transformation matrix $R_\beta{}^\alpha(1)$ takes on the standard form $u_{\lambda_1}{}^{\lambda_1'}\cdots u_{\lambda_k}{}^{\lambda_k'}\times\bar{u}_{\mu_1}{}^{\mu_1'}\cdots\bar{u}_{\mu_j}{}^{\mu_j'}$, where u is the spin matrix $\sigma(l)$ representing the Lorentz transformation l. Then consider the set of fields obtained by partial differentiation:

$$\chi_{\lambda,\lambda_1\cdots\lambda_k,\mu_2\cdots\mu_j}=\epsilon^{\mu,\mu}\partial_{\lambda\mu}\psi_{\lambda_1\cdots\lambda_k\mu_1\cdots\mu_j}. \quad (8.6)$$

According to the Clebsch-Gordan formula, these fields {which transform according to the representation $(1)\otimes[(k)\otimes(j_1)]$, where $j_1=j-1$} may be written in a Lorentz-invariant way as a linear combination with constant coefficients of fields

$$\psi_{\lambda_1\cdots\lambda_{k+1}\mu_1\cdots\mu_{j-1}} \quad \text{and} \quad \psi_{\lambda_1\cdots\lambda_{k-1}\mu_1\cdots\mu_{j-1}}$$

which transform according to the representations $(k+1)\otimes(\bar{j_1})$ and $(k-1)\otimes(\bar{j_1})$, respectively. (Note that "Lorentz invariant," in the preceding sentence, means that the constant coefficients spoken of there remain unchanged if all the fields $\chi_{\lambda,\lambda_1\cdots\lambda_k,\mu_2\cdots\mu_j}$; $\psi_{\lambda_1\cdots\lambda_{k+1}\mu_1\cdots\mu_{j-1}}$; and $\psi_{\lambda_1\cdots\lambda_{k-1}\mu_1\cdots\mu_{j-1}}$ are simultaneously transformed in the manner which their subscripts indicate.) Moreover, according to formula (2.19) we have

$$\psi_{\lambda_1\cdots\lambda_k\mu_1\cdots\mu_j}=m^{-2}\Box\psi_{\lambda_1\cdots\lambda_k\mu_1\cdots\mu_2}$$
$$=\epsilon^{\lambda\lambda'}\epsilon^{\mu\mu'}\partial_{\lambda'\mu_1}\partial_{\lambda\mu}\psi_{\lambda_1\cdots\lambda_k\mu_1\cdots\mu_j} \quad (8.7)$$

so that $\psi_{\lambda_1\cdots\lambda_k\mu_1\cdots\mu_j}$ can be written as a Lorentz-invariant linear combination with constant coefficients of partial derivatives of the fields $\psi_{\lambda_1\cdots\lambda_{k+1}\mu_1\cdots\mu_{j-1}}$ and $\psi_{\lambda_1\cdots\lambda_{k-1}\mu_1\cdots\mu_{j-1}}$. Proceeding inductively in this way, we find that there exists fields $\psi_{\lambda_1\cdots\lambda_q}$ and $\psi_{\mu_1\cdots\mu_n}$ belonging, respectively, to the representations $(k)\otimes(\bar{0})$ and $(0)\otimes(\bar{j})$ of the proper Lorentz group $\mathfrak{L}_p$, which are invariant linear combinations with constant coefficients of the partial derivatives of the field components $\psi_{\lambda_1\cdots\lambda_k\mu_1\cdots\mu_j}$, such that the field components $\psi_{\lambda_1\cdots\lambda_k\mu_1\cdots\mu_j}$ are in turn invariant linear combinations with constant coefficients of the partial derivatives of the fields $\psi_{\lambda_1\cdots\lambda_q}$ and $\psi_{\mu_1\cdots\mu_n}$. Hence it follows that our original quantized free-field theory is equivalent to one in which all the field operators ψ_α which satisfy an equation $\Box\psi_\alpha=m^2\psi_\alpha$, $m>0$, transform according to a representation of the Lorentz group having either the form $(k)\times(\bar{0})$ or the complex-conjugate form $(0)\times(\bar{j})$.

The mapping

$$\psi_{\lambda_1\cdots\lambda_k}\longrightarrow(\mathcal{J}\psi)_{\lambda_1\cdots\lambda_k}=\partial_{\lambda_1}{}^{\mu_1}\partial_{\lambda_2}{}^{\mu_2}\partial_{\lambda_k}{}^{\mu_k}\psi_{\mu_1\cdots\mu_k}^* \quad (8.8)$$

is an antilinear mapping whose square is

$$\psi_{\lambda_1\cdots\lambda_k}\longrightarrow\partial_{\lambda_1\mu_1}\partial_{\lambda_2\mu_2}\cdots\partial_{\lambda_k\mu_k}\partial^{\nu_1\mu_1}\partial^{\nu_2\mu_2}\partial^{\nu_k\mu_k}\psi_{\nu_1\cdots\nu_k}$$
$$=m^k\psi_{\lambda_1\cdots\lambda_k} \quad (8.9)$$

by formula (2.19). Thus, if we put $J=m^{-k}\mathcal{J}$, J is an antilinear mapping whose square is $+1$. Thus the field ψ may be written as

$$\psi=1/2(\psi+J\psi)+1/2i(\psi-J\psi): \quad (8.10)$$

the fields $\psi_1=1/2(\psi+J\psi)$ and $\psi_2=1/2i(\psi-J\psi)$ satisfy $J\psi_1=\psi_1$; $J\psi_2=\psi_2$. It is again clear that passing to an equivalent free-field theory we may replace the field ψ by the fields ψ_1 and ψ_2, so that without loss of generality we may suppose that all the field operators $\psi_{\lambda_1\cdots\lambda_k}$ of our theory which satisfy an equation

$$\Box\psi_{\lambda_1\cdots\lambda_k}=m^2\psi_{\lambda_1\cdots\lambda_k}, \quad m>0,$$

satisfy the reality condition $J\psi=\psi$, i.e., satisfy

$$\psi_{\mu_1\cdots\mu_k}{}^*=m^{-k}\partial_{\mu_1}{}^{\lambda_1}\partial_{\mu_2}{}^{\lambda_2}\partial_{\mu_k}{}^{\lambda_k}\psi_{\lambda_1\cdots\lambda_k}. \quad (8.11)$$

It follows in the same way that we may suppose without loss of generality that all the field operators $\psi_{\mu_1\cdots\mu_j}$ of our theory which satisfy an equation $\Box\psi_{\mu_1\cdots\mu_j}=m^2\times\psi_{\mu_1\cdots\mu_j}$, $m>0$, satisfy the reality condition

$$\psi_{\lambda_1\cdots\lambda_j}{}^*=m^{-j}\partial_{\lambda_1}{}^{\mu_1}\partial_{\lambda_2}{}^{\mu_2}\cdots\partial_{\lambda_j}{}^{\mu_j}\psi_{\mu_1\cdots\mu_j}. \quad (8.12)$$

In virtue of the reality conditions (8.11) and (8.12), it is clear that if we simply drop the field components $\psi_{\mu_1\cdots\mu_j}$ transforming according to representations $(0)\times(\bar{j})$, and retain only those field-components $\psi_{\lambda_1\cdots\lambda_k}$ transforming according to representations $(k)\times(\bar{0})$, and their adjoints, we obtain a theory equivalent to the theory with which we started.

To summarize: Upon passing to an equivalent theory, the field operators of a quantized free-field theory satisfying axioms (a)–(d) may without loss of generality

be taken to satisfy axioms (a)–(d), to have the properties described in the final lemma of the preceding section, and in addition to have the following properties:

(i) Every field component ψ_α satisfies the equation

$$\Box\psi_\alpha = m^2\psi_\alpha, \quad m = m(\alpha) \geq 0. \tag{8.13}$$

(ii) The field components ψ_α with $m(\alpha)>0$ are all fields $\psi_{\lambda_1\cdots\lambda_n}$ belonging to representations $(n)\times(\bar{0})$ of the proper Lorentz group $\mathfrak{L}_p$, and the adjoints of such fields.

(iii) The field components $\{\psi_\alpha\}$ with $m(\alpha)>0$ all satisfy the reality conditions

$$\psi_{\mu_1\cdots\mu_n}{}^* = m^{-n}\partial_{\mu_1}{}^{\lambda_1}\partial_{\mu_2}{}^{\lambda_2}\cdots\partial_{\mu_n}{}^{\lambda_n}\psi_{\lambda_1\cdots\lambda_n}$$

and

$$\psi_{\lambda_1\cdots\lambda_n} = m^{-n}\partial_{\lambda_1}{}^{\mu_1}\partial_{\lambda_2}{}^{\mu_2}\cdots\partial_{\lambda_n}{}^{\mu_n}\psi_{\mu_1\cdots\mu_n}{}^*.$$

(iv) The various sets $\{\psi_{\lambda_1\cdots\lambda_n}\}$ of field components ψ_α with $m(\alpha)>0$ satisfy no relation of linear dependence.

According to the Clebsch-Gordan formula, the representation $(k)\otimes(k')\otimes(n)\otimes(\bar{n})$ contains a scalar [i.e., contains the representation $(0)\times(\bar{0})$] if and only if $n=0$ and $k_1=k_2$. Now, we have shown previously that the expression

$$[\psi_{\lambda_1\cdots\lambda_k}(x),\psi_{\lambda_1'\cdots\lambda_{k'}'}(x')]$$

must have the form

$$[\psi_{\lambda_1\cdots\lambda_k}(x),\psi_{\lambda_1'\cdots\lambda_{k'}'}(x')] = (a_{\lambda_1\cdots\lambda_k,\lambda_1'\cdots\lambda_{k'}'} + a_{\lambda_1\cdots\lambda_k,\lambda_1'\cdots\lambda_{k'}'}{}^{j_1}\partial_{j_1} + \cdots + a_{\lambda_1\cdots\lambda_k,\lambda_1'\cdots\lambda_{k'}'}{}^{j_1\cdots j_n}\partial_{j_1}\cdots\partial_{j_n} + \cdots) \times D^{(m)}(x-x').$$

Since the function $D^{(m)}$ satisfies the equation $\Box D^{(m)} = m^2 D^{(m)}$, it is clear that we may without loss of generality suppose the coefficients a to (be symmetric in their indices j and) satisfy

$$a_{\lambda_1\cdots\lambda_k,\lambda_1'\cdots\lambda_{k'}'}{}^{j_1\cdots j_n}G_{j_1j_2} = 0,$$

$G_{j_1j_2}$ being the Lorentz metric. Thus the coefficients $a_{\lambda_1\cdots\lambda_k,\lambda_1'\cdots\lambda_{k'}'}{}^{j_1\cdots j_n}$ transform according to the representation $(k)\otimes(k')\otimes(n)\otimes(\bar{n})$ (cf. the final paragraph of Sec. 2). If $k\neq k'$, we conclude that all the coefficients a are zero. If $k=k'$, we conclude similarly from the uniqueness of the invariant contained in the representation that

$$[\psi_{\lambda_1\cdots\lambda_k}(x),\psi_{\lambda_1'\cdots\lambda_k'}(x')] = c\epsilon_{\lambda_1\cdots\lambda_k;\,\lambda_1'\cdots\lambda_k'} \times D^{(m)}(x-x'), \tag{8.14}$$

where c is some complex constant, and where

$$\epsilon_{\lambda_1\cdots\lambda_k;\,\lambda_1'\cdots\lambda_k'} = \frac{1}{k!}\sum \epsilon(\lambda_1,\lambda i_1')\epsilon(\lambda_2,\lambda i_2')\cdots\epsilon(\lambda_k,\lambda i_k'), \tag{8.15}$$

the summation being extended over all permutations $(i_1\cdots i_n)$ of $(1\cdots n)$. Thus (14) holds if $k=k'$, while in the contrary case we must have, even more simply,

$$[\psi_{\lambda_1\cdots\lambda_k}(x),\psi_{\lambda_1'\cdots\lambda_{k'}'}(x')] \equiv 0, \quad k\neq k'. \tag{8.16}$$

To investigate the constant c in Eq. (14), it is convenient to make some explicit calculations of the usual sort. By $d\sigma$ we shall denote the Lorentz-invariant element

$$d\sigma = m\,dp_1\,dp_2\,dp_3/(m^2+p_1{}^2+p_2{}^2+p_3{}^2)^{\frac{1}{2}} \tag{8.17}$$

on the two-sheeted hyperboloid h_m: $\{p^\mu \mid p_\mu p^\mu = m^2\}$, $m\geq 0$. In terms of our earlier notation for the two branches of this hyperboloid, $h_m = \sum_m{}^+ \cup \sum_m{}^-$. The invariant function $D^{(m)}(x)$ may be represented by

$$D^{(m)}(x) = \int_{h_m} D^{(m)}(p)e^{ix\cdot p}d\sigma, \tag{8.18}$$

where $D(p)$ is an odd function of p given on the branch $\sum_m{}^+$ by

$$D^{(m)}(p) = \frac{1}{i}\frac{1}{2m}\frac{1}{(2\pi)^3}. \tag{8.19}$$

Thus,

$$D^{(m)}(x-x') = \int_{h_m}\int_{h_m} D^{(m)}(p)\delta(-p,p')e^{ix\cdot p}e^{ix'\cdot p'}d\sigma d\sigma', \tag{8.20}$$

where $\delta(p,p')$ is the Lorentz-invariant δ function defined on the Cartesian product of h_m with itself by the formula

$$\int_{h_m}\int_{h_m}\phi(p)\phi'(p')\delta(p,p')d\sigma d\sigma' = \int_{h_m}\phi(p)\phi'(p)d\sigma. \tag{8.21}$$

By $\partial_{\lambda_1\cdots\lambda_n\mu_1\cdots\mu_n}$ and $p_{\lambda_1\cdots\lambda_n\mu_1\cdots\mu_n}$ we shall denote

$$\partial_{\lambda_1\cdots\lambda_n\mu_1\cdots\mu_n} = \frac{1}{n!}\sum \partial(\lambda_1,\mu i_1)\partial(\lambda_2,\mu i_2)\cdots\partial(\lambda_n,\mu i_n) \tag{8.22}$$

and

$$p_{\lambda_1\cdots\lambda_n\mu_1\cdots\mu_n} = \frac{1}{n!}\sum p(\lambda_1,\mu i_1)p(\lambda_2,\mu i_2)\cdots p(\lambda_n,\mu i_n), \tag{8.23}$$

where $\partial(\lambda,\mu)=\partial_{\lambda\mu}$, $p(\lambda,\mu)=p_{\lambda\mu}=p_j\sigma_{\lambda\mu}{}^j$, and where the summation is in each case taken over all permutations $i_1\cdots i_n$ of the integers $1\cdots n$. If $\psi_{\lambda_1\cdots\lambda_n}$ is a field satisfying the differential equation $\Box\psi_{\lambda_1\cdots\lambda_n} = m^2\psi_{\lambda_1\cdots\lambda_n}$, we may evidently write

$$\psi_{\lambda_1\cdots\lambda_n}(x) = \int_{h_m} e^{ix\cdot p}\psi_{\lambda_1\cdots\lambda_n}(p)d\sigma. \tag{8.24}$$

If ψ also satisfies the commutation relations

$$\begin{aligned}&[\psi_{\lambda_1\cdots\lambda_n}(x),\psi_{\mu_1\cdots\mu_n}{}^*(x')]\\&\qquad=a\partial_{\lambda_1\cdots\lambda_n\mu_1\cdots\mu_n}D^{(m)}(x-x')\\&[\psi_{\lambda_1\cdots\lambda_n}(x),\psi_{\lambda_1'\cdots\lambda_n'}(x')]\\&\qquad=b\epsilon_{\lambda_1\cdots\lambda_n\lambda_1'\cdots\lambda_n'}D^{(m)}(x-x')\end{aligned}\tag{8.25}$$

(the differential operator being applied to the variable x'), it evidently follows that

$$\begin{aligned}&[\psi_{\lambda_1\cdots\lambda_n}(p),\psi_{\mu_1\cdots\mu_n}{}^*(p')]\\&\qquad=ai^{n-1}(-1)^n p_{\lambda_1\cdots\lambda_n,\mu_1\cdots\mu_n}D(p)\delta(p,p')\\&[\psi_{\lambda_1\cdots\lambda_n}(p),\psi_{\lambda_1'\cdots\lambda_n'}(p')]\\&\qquad=b\epsilon_{\lambda_1\cdots\lambda_n\lambda_1'\cdots\lambda_n'}D(p)\delta(-p,p');\end{aligned}\tag{8.26}$$

for each p, $p'\epsilon h_m$. The matrix $p_{\lambda_1\cdots\lambda_n,\mu_1\cdots\mu_n}=p_{LM}$ is evidently positive-definite if $p_0>0$. Thus since $AA^*\pm A^*A$ is Hermitian, $(-1)^n ai^n$ is real.

If the brackets in formula (8.25) are anticommutation brackets, and the field is not identically zero, then if we put $p=p'$ with $p_0>0$ in (26), it clearly follows that $(-1)^n ai^{n-1}$ is positive, while if we put $p=p'=-q$ with $q_0>0$ in (8.26), it clearly follows that $(-1)^n ai^{n-1}\times(-1)^n(-1)$ is positive. From the positivity of both these quantities, it follows at once than n is odd, and that $a=-a'i^{n-1}$, where $a'>0$.

Next consider the case in which the brackets in formula (8.25) are commutator brackets. Since we assume all our fields linearly independent, it follows at once from the first lemma of Sec. 5 that $a\neq 0$. Put

$$A=|a(-1)^n i^{n-1}p_{1\cdots1,1\cdots1}D(p)\delta(p,p)|^{\frac{1}{4}}\psi_{1\cdots1}(p).\tag{8.27}$$

Then, if

$$\begin{aligned}&(-1)^n ai^{n-1}>0\\&\qquad\text{and } p^0>0,\text{ then } AA^*-A^*A=I\\&(-1)^n ai^{n-1}<0\\&\qquad\text{and } p^0>0,\text{ then } AA^*-A^*A=-I\\&(-1)(-1)^n(-1)^n ai^{n-1}>0\\&\qquad\text{and } p^0<0,\text{ then } AA^*-A^*A=I\\&(-1)(-1)^n(-1)^n ai^{n-1}<0\\&\qquad\text{and } p^0<0,\text{ then } AA^*-A^*A=-I.\end{aligned}\tag{8.28}$$

Now, the Fourier-transformed operators $\psi_{\lambda_1\cdots\lambda_n}(p)$ of our theory satisfy

$$[P^\mu,\psi_{\lambda_1\cdots\lambda_n}(p)]_-=P^\mu\psi_{\lambda_1\cdots\lambda_n}(p)\tag{8.29}$$

$$[P^\mu,\psi_{\mu_1\cdots\mu_n}{}^*(p)]_-=-P^\mu\psi_{\mu_1\cdots\mu_n}(p),\tag{8.30}$$

P^μ denoting the 4-vector of energy-momentum operators, as has been shown previously [cf. formula (5.6)]. Thus

$$\begin{aligned}&[P^0,A]=p^0A\\&[P^0,A^*]=-p^0A^*.\end{aligned}\tag{8.31}$$

Hence, if x is a vector such that $P^0x=\lambda x$, then since the spectrum of P^0 is entirely non-negative, we have $A^nx=0$ for sufficiently large n if p^0 is negative and $(A^*)^nx=0$ for sufficiently large n if p^0 is positive. But, in the second case of (8.28) it follows that $|A^*x|^2=|Ax|^2+|x|^2$ for each x, so that A^* has no zeroes; and in the third case of (8.29) it follows similarly that A has no zeroes. These cases are consequently excluded, and it follows that $(-1)^n ai^{n-1}>0$, $(-1)ai^{n-1}<0$, so that $(-1)^{n-1}$ is positive, n is even, and $a=ia'$ where a' is real and of the same sign as i^n. Thus, we have proved the following.

Lemma. In case the brackets in formula (25) are commutator brackets, then either $a=0$, or n is even and $a=ia'$, where a' is real and of the same sign as i^n.

Note that this lemma and the corresponding statement for anticommutator brackets are together equivalent to the well-known theorem of Pauli in spin and statistics.

By (ii), each of our classes of field components may be taken to consist of a set of fields $\psi_{\lambda_1\cdots\lambda_n}{}^{(j)}$ with a fixed n or of the class of adjoints of such class of fields; as n and j vary, j going from $1\cdots R_n$ for fixed n, the various separate classes of fields are enumerated.

Our next aim is to examine a family of indices having some fixed value of n more closely. By (14), we have

$$\begin{aligned}&[\psi_{\lambda_1\cdots\lambda_n}{}^{(j)}(x),\psi_{\lambda_1'\cdots\lambda_n'}{}^{(k)}(x')]\\&\qquad=c_{jk}\epsilon_{\lambda_1\cdots\lambda_n,\lambda_1'\cdots\lambda_n'}D^{(m)}(x-x').\end{aligned}\tag{8.32}$$

Let $\{\alpha\}$ be such a family, and suppose that $\{\psi_\alpha\}=\{\psi_{\lambda_1\cdots\lambda_n}{}^{(j)}\}$ is the corresponding family of field operators, the various classes making up the family being enumerated as the index j varies.

Then $\sigma(\alpha,\alpha)=\sigma$ for all α in the given class of indices; where $\sigma=+1$ if n is odd, and $\sigma=-1$ if n is even by what has just been proved. Moreover, by the final lemma of Sec. 7, $\sigma(\alpha,\beta)=\sigma$ for all α, β in the family. Since

$$\epsilon_{\lambda_1\cdots\lambda_n,\lambda_1'\cdots\lambda_n'}=(-1)^n\epsilon_{\lambda_1'\cdots\lambda_n',\lambda_1\cdots\lambda_n},\tag{8.33}$$

and $D^\Delta(x)$ is an odd function of x, it follows at once from (8.25) that

$$c_{jk}=(-1)^{n+1}\sigma c_{kj}.\tag{8.34}$$

Thus, since $\sigma(-1)^{n+1}=1$, $c_{jk}=c_{kj}$. On taking the adjoint of (8.32), we have

$$\begin{aligned}&[\psi_{\mu_1'\cdots\mu_n'}{}^{*(k)}(x'),\psi_{\mu_1\cdots\mu_n}{}^{*(j)}(x')]\\&\qquad=\bar c_{jk}\epsilon_{\mu_1\cdots\mu_n\mu_1'\cdots\mu_n'}D^{(m)}(x-x').\end{aligned}\tag{8.35}$$

Then, using the reality condition of (iii), it follows that

$$\begin{aligned}&\bar c_{jk}\epsilon_{\mu_1\cdots\mu_n\mu_1'\cdots\mu_n'}D^{(m)}(x-x')\\&\quad=\partial_{\mu_1}{}^{\lambda_1'}\partial_{\mu_2}{}^{\lambda_2'}\cdots\partial_{\mu_n}{}^{\lambda_n'}\partial_{\mu_1}{}^{\lambda_1}\cdots\partial_{\mu_n}{}^{\lambda_n}m^{-2n}c_{kj}\\&\qquad\times\epsilon_{\lambda_1'\cdots\lambda_n',\lambda_1\cdots\lambda_n}D^{(m)}(x'-x)\\&\quad=m^{-2n}\Box^n c_{kj}(-1)^n(-1)D^{(m)}(x-x')\epsilon_{\mu_1\cdots\mu_n\mu_1'\cdots\mu_n'}\\&\quad=c_{kj}(-1)^{n+1}\epsilon_{\mu_1\cdots\mu_n\mu_1'\cdots\mu_n'}D^{(m)}(x-x').\end{aligned}$$

Consequently,

$$c_{jk}=(-1)^{n+1}c_{jk}.\tag{8.36}$$

Hence, if $\sigma=+1$, c_{jk} is real and symmetric.

If $\sigma=-1$, so that n is even, then it follows in exactly the same way that c_{jk} is imaginary and symmetric.

Finally then, since c_{jk} has been shown to be symmetric, real in case $\sigma=+1$, and imaginary in case $\sigma=-1$, it follows that there exists a real orthogonal matrix $O_k{}^j$ such that OcO^{-1} is diagonal. On replacing the system of fields $\psi_{\lambda_1\cdots\lambda_n}{}^{(j)}$ by the equivalent system of fields $O_k{}^j\psi_{\lambda_1\cdots\lambda_n}{}^{(k)}$, it follows that we may assume without loss of generality that $c_{jk}=0$ for $j\neq k$. Then, replacing each field $\psi_{\lambda_1\cdots\lambda_n}{}^{(j)}$ by a real multiple of itself, we may assume without loss of generality that $|c_{jj}|=1$. By what has been proved previously, we have

$$c_{jj}=i^{n+1}. \tag{8.37}$$

Thus, summarizing: Upon passing to an equivalent theory, the field operators of a quantized free-field theory satisfying axioms (a)–(d) may without loss of generality to satisfy axioms (a)–(d), and (i)–(iv) of the present section. In addition, we may suppose that $[\psi_\alpha,\psi_\beta]=0$ if $m(\alpha)\neq m(\beta)$ or if $m(\alpha)=m(\beta)>0$ while α and β lie in *distinct classes* of indices, while if $\{\psi_{\lambda_1\cdots\lambda_n}\}=\{\psi_\alpha\}$ is any given class of fields with $m(\alpha)>0$, we may suppose that $\sigma(\alpha,\tilde{\alpha})=(-1)^{n+1}$, and

$$[\psi_{\lambda_1\cdots\lambda_n}(x),\psi_{\lambda_1'\cdots\lambda_n'}(x')]=i^{n+1}\epsilon_{\lambda_1\cdots\lambda_n,\lambda_1'\cdots\lambda_n'}D^{(m)}(x-x').$$

9. SOME SPECIAL EXAMPLES OF FIELDS OF POSITIVE MASS

The best-known examples, aside from the trivial case of a scalar field satisfying the Klein-Gordon equation, are those (Majorana and Dirac cases) related to the Dirac equation. In the present notation, the Dirac four-spinor field is a pair $(\psi_\lambda,\tilde{\psi}_{\dot{\lambda}})$ of classes of fields, transforming according to the representations (1) and $(\bar{1})$, respectively. The Dirac equation may be written

$$\begin{aligned}\partial_\mu{}^\lambda\psi_\lambda&=m\tilde{\psi}_{\dot{\mu}}\\ \partial_\lambda{}^\mu\tilde{\psi}_\mu&=m\psi_\mu.\end{aligned} \tag{9.1}$$

The commutation relations are

$$[\psi_\lambda,\tilde{\psi}_\mu]_+=[\psi_\lambda,\psi_{\lambda'}]_+=[\psi_\mu,\psi_{\mu'}]_+\equiv 0 \tag{9.2}$$

and

$$\begin{aligned}&[\psi_\lambda,\psi_\mu{}^*]_+\equiv-\partial_{\lambda\mu}D^{(m)}\equiv[\tilde{\psi}_\lambda{}^*,\tilde{\psi}_\mu]_+;\\ &[\psi_\lambda,\tilde{\psi}_{\lambda'}{}^*]_+\equiv-m\epsilon_{\lambda\lambda'}D^{(m)};\\ &[\tilde{\psi}_\mu,\psi_{\mu'}{}^*]_+\equiv-m\epsilon_{\mu\mu'}D^{(m)}.\end{aligned} \tag{9.3}$$

Equation (9.1) makes it evident that the field $\tilde{\psi}$ is merely a "gradient" of ψ, and can be dropped. Thus, the Dirac theory can as well be written as a two-spinor theory, with the equations

$$(\Box-m^2)\psi_\lambda=0 \tag{9.4}$$

$$[\psi_\lambda,\psi_{\lambda'}]_+=0 \tag{9.5}$$

$$[\psi_\lambda,\psi_\mu{}^*]=-\partial_{\lambda\mu}D^{(m)}. \tag{9.6}$$

The analysis in the previous sections shows that the Dirac field ψ_λ of (9.4)–(9.6) allows a "decomposition into real and imaginary part"

$$\begin{aligned}\psi_\lambda{}^{(R)}&=\tfrac{1}{2}(\psi_\lambda-m^{-1}\partial_\lambda{}^\mu\psi_\mu{}^*);\\ \psi_\lambda{}^{(I)}&=(1/2i)(\psi_\lambda-m^{-1}\partial_\lambda{}^\mu\psi_\mu{}^*)\end{aligned} \tag{9.7}$$

$$\psi_\lambda=\psi_\lambda{}^{(R)}+i\psi_\lambda{}^{(I)}. \tag{9.8}$$

The commutation relations (9.5) and (9.6) are then those uniquely prescribed by symmetry under "change conjugation"

$$\psi_\lambda{}^{(R)}+i\psi_\lambda{}^{(I)}\to\psi_\lambda{}^{(R)}-i\psi_\lambda{}^{(I)},\quad\text{i.e.,}\quad\psi_\lambda\to m^{-1}\partial_\lambda{}^\mu\psi_\mu{}^*.$$

The reality condition $\tilde{\psi}_\lambda{}^*=\psi_\lambda$ which may be imposed upon the components of the Dirac four-spinor, can be written in two-component notation as

$$\psi_\lambda=m^{-1}\partial_\lambda{}^\mu\psi_\mu{}^*, \tag{9.9}$$

which explains why we have chosen to call the general equation of this form a "reality condition." Thus, the Majorana theory, which we obtain from the Dirac theory by imposing the reality condition (9.9), is, from the point of view of the previous section, the simplest theory of positive mass with spin, satisfying the equations

$$\Box\psi_\lambda=m^2\psi_\lambda \tag{9.10}$$

$$\psi_\lambda=m^{-1}\partial_\lambda{}^\mu\psi_\mu{}^* \tag{9.11}$$

$$[\psi_\lambda,\psi_{\lambda'}]_+=-\epsilon_{\lambda\lambda'}D^{(m)}. \tag{9.12}$$

The next higher spin case is the theory of a field $\psi_{\lambda\lambda'}$ satisfying

$$\Box\psi_{\lambda\lambda'}=m^2\psi_{\lambda\lambda'} \tag{9.13}$$

$$\psi_{\lambda\lambda'}=m^{-2}\partial_\lambda{}^\mu\partial_{\lambda'}{}^{\mu'}\psi_{\mu\mu'}{}^* \tag{9.14}$$

$$[\psi_{\lambda\lambda'}(x),\psi_{\lambda''\lambda'''}(x')]_-=-i\epsilon_{\lambda\lambda',\lambda''\lambda'''}D^{(m)}(x-x'). \tag{9.15}$$

If we put

$$A_j=\sigma_j{}^{\lambda\mu}\partial_\mu{}^{\lambda'}\psi_{\lambda\lambda'}, \tag{9.16}$$

so that A is a four-vector, then

$$\Box A_j=m^2A_j \tag{9.17}$$

$$\partial^jA_j=\partial^{\lambda\mu}\partial_\mu{}^\lambda\psi_{\lambda\lambda'}=\Box\epsilon^{\lambda\lambda'}\psi_{\lambda\lambda'}=0 \tag{9.18}$$

$$\begin{aligned}A_j{}^*&=(\sigma_j{}^{\lambda\mu}\partial_\mu{}^{\lambda'}\psi_{\lambda\lambda'})^*=(\sigma_j{}^{\lambda\mu}\sigma_k{}^{\lambda'\mu'}\epsilon_{\mu\mu'}\partial^k\psi_{\lambda\lambda'})^*\\ &=\sigma_j{}^{\mu\lambda}\sigma_k{}^{\mu'\lambda'}\epsilon_{\mu\mu'}\partial^k\psi_{\lambda\lambda'}{}^*\\ &=m^{-2}\sigma_j{}^{\mu\lambda}\partial_\mu{}^{\lambda'}\partial_{\lambda'}{}^{\mu'}\partial_{\lambda'}{}^{\mu''}\psi_{\mu'\mu''}\\ &=m^{-2}\sigma_j{}^{\mu\lambda}\Box\partial_{\lambda'}{}^{\mu'}\psi_{\mu'\mu}\\ &=\sigma_j{}^{\mu\lambda}\partial_\lambda{}^{\mu'}\psi_{\mu'\mu}=A_j.\end{aligned} \tag{9.19}$$

$$\begin{aligned}&[A_j(x),A_k(x')]\\ &=-i\sigma_j{}^{\lambda\mu}\sigma_k{}^{\lambda'\mu'}\partial_\mu{}^{\lambda''}\partial_{\mu'}{}^{\lambda'''}\epsilon_{\lambda\lambda'',\lambda'\lambda'''}D^{(m)}(x-x')\\ &=-i\sigma^{\lambda\mu}\sigma^{\lambda'\mu'}(\epsilon_{\mu\mu'}\epsilon_{\lambda\lambda'}m^2-\sigma_j{}^{\lambda\mu}\sigma_k{}^{\lambda'\mu'}\partial_{\lambda'\mu}\partial_{\lambda\mu'})D^{(m)}(x-x')\\ &=-i\gamma(m^2G-\partial_i\partial_j)D^{(m)}(x-x').\end{aligned} \tag{9.20}$$

We do not bother to evaluate the numerical constant γ. Thus, the theory described by (9.13)–(9.15) is simply the ordinary theory of the "real vector meson of

positive mass" satisfying the "divergence condition" $\partial^j A_j = 0$. It should be noted that the "reality condition" (9.14) is equivalent to the condition $A_j = A_j^*$.

10. EXPLICIT ANALYSIS OF FIELDS OF MASS ZERO

In the present section we will complete our analysis of the field components of zero mass. The general structure of these fields will be most apparent if we examine a very simple special case first: the case of a field ψ_λ transforming according to the representation (10.1) and satisfying the zero-mass equation $\Box\psi_\lambda = 0$.

It follows from the equation $\Box\psi_\lambda = 0$ and from Lorentz invariance (compare the corresponding arguments at the beginning of Sec. 8) that the bracket

$$[\psi_\lambda(x), \psi_\mu^*(x')]_+ \tag{10.1}$$

must be given by an expression

$$[\psi_\lambda(x), \psi_\mu^*(x')]_+ = c\partial_{\lambda\mu} D^{(0)}(x-x'). \tag{10.2}$$

Consequently, if we put $\chi_\mu(x) = \partial_\mu{}^\lambda \psi_\lambda(x)$, it follows that

$$[\chi_\mu(x), \chi_\lambda^*(x')]_+ = -c\partial_{\mu\lambda}\Box D^{(0)}(x-x') = 0. \tag{10.3}$$

On putting $x = x'$ and $\mu = \lambda$, it follows immediately that $\chi_\mu(x) \equiv 0$. That is, in the zero-mass case, the second-order equation $\Box\psi_\lambda(x) = 0$, together with our other axioms, implies the first-order equation

$$\partial_\mu{}^\lambda \psi_\lambda(x) = 0!$$

We shall see in the present section that this phenomenon is perfectly general.

If we note that the initial arguments of Sec. 8 apply as well to fields satisfying the equation $\Box\psi_\alpha = 0$ as to fields satisfying the equation $\Box\psi_\alpha = m^2\psi_\alpha$, it follows at once that if $\{\psi_\alpha\}$ is a given *family* of fields satisfying $\Box\psi_\alpha = 0$, we have

$$[\psi_\alpha(x), \psi_\beta(x')]_{\sigma(a,\beta)} = (a_{\alpha\beta} + a_{\alpha\beta}{}^j\partial_j + a_{\alpha\beta}{}^{j_1 j_2}\partial_{j_1}\partial_{j_2} + \cdots) D^{(0)}(x-x'), \tag{10.4}$$

where $D^{(0)}(x)$ is the solution of the equations

$$\Box D^{(0)} = 0; \quad D^{(0)}(\mathbf{x}, 0) = 0; \quad (\partial/\partial t) D^{(0)}(\mathbf{x}, 0) = \delta(\mathbf{x}). \tag{10.5}$$

As in Sec. 8, the matrices of coefficients $a^{j_1 \cdots j_k}$ may be taken to be symmetric in $j_1 \cdots j_k$ and to satisfy $a^{j_1 j_2 \cdots j_k} G_{j_1 j_2} = 0$, $G_{j_1 j_2}$ denoting the Lorentz metric.

Consider more specifically a class of field components $\psi_{\lambda_1 \cdots \lambda_a \mu_1 \cdots \mu_b}$ transforming according to the representation of the Lorentz group indicated by the indices displayed. If we note that by the Clebsch-Gordan formula the representation $(a) \times (\bar{b}) \times (\bar{a}) \times (b) \times (n) \times (\bar{n})$ contains exactly one invariant quantity for $a+b \geq n \geq |a-b|$, and exactly zero invariant quantities for n not in this range, it follows at once that for the field components under consideration, we may write formula (10.4) more specifically as

$$[\psi_{\lambda_1 \cdots \lambda_a \mu_1 \cdots \mu_b}, \psi_{\mu_1' \cdots \mu_a' \lambda_1' \cdots \lambda_b'}^*] = \sum_{n=0}^{\min(a,b)} \sum c_n \epsilon_{\lambda_1 \lambda_1'} \cdots \epsilon_{\lambda_n \lambda_n'} \epsilon_{\mu_1 \mu_1'} \cdots \epsilon_{\mu_n \mu_n'} \times \partial_{\lambda_{n+1} \cdots \lambda_b' \mu_{n+1} \cdots \mu_a'} D^{(0)}; \tag{10.6}$$

here, every subscript λ_j, λ_j', μ_j, or μ_j' which does not occur attached to one of the symbols "ϵ" is to be attached to the symbol "∂"; the inner symbol sym indicates "symmetrization" by summation over all permutations of the subscripts $\lambda_1' \cdots \lambda_b'$ and all permutations of the subscripts $\mu_1' \cdots \mu_a'$.

We shall now show that $c_n = 0$ for all $n > 0$. Suppose, in fact, that the nonzero coefficient c in (10.6) with largest subscript is c_k, and that $k > 0$. If $b = 0$, then $\min(a,b) = 0$, and we have nothing to prove. Hence, we may also assume without loss of generality that $b > 0$. This being the case, we form the field

$$\chi_{\nu_1 \cdots \nu_k \eta \lambda_{k+1} \cdots \lambda_a \mu_2 \cdots \mu_b} = \partial_{\nu_1}{}^{\lambda_1} \cdots \partial_{\nu_k}{}^{\lambda_k} \partial_\eta{}^{\mu_1} \psi_{\lambda_1 \cdots \lambda_a \mu_1 \cdots \mu_b}. \tag{10.7}$$

The commutation relations satisfied by this field follow at once by differentiation of (10.6). We find

$$[\chi_{\nu_1 \cdots \nu_k \eta \lambda_{k+1} \cdots \lambda_a \mu_2 \cdots \mu_b}, \chi_{\eta_1 \cdots \eta_k \nu \mu_{k+1}' \cdots \mu_a' \lambda_2' \cdots \lambda_b'}^*] = \sum_{n=0}^{k} \sum c_n \partial_{\nu_1 \lambda_1'} \partial_{\nu_2 \lambda_2'} \cdots \partial_{\nu_k \lambda_k'} \partial_{\eta_1 \mu_1} \partial_{\eta_2 \mu_2} \cdots \partial_{\eta_k \mu_k} \times \partial_\eta{}^{\mu_1} \partial_\nu{}^{\lambda_1'} \partial_{\nu_{n+1}}{}^{\lambda_{n+1}} \cdots \partial_{\nu_k}{}^{\lambda_k} \partial_{\eta_{n+1}}{}^{\mu_{n+1}'} \cdots \partial_{\eta_k}{}^{\mu_k'} \partial_{\lambda_{n+1} \cdots \lambda_b' \mu_{n+1} \cdots \mu_a'} D^{(0)}. \tag{10.8}$$

The inner sum $\sum$ here denoting a sum over permutations of the indices as in (10.6). It is apparent from (10.7) that each term on the right-hand side contains a factor having the structure

$$\partial_{\nu_1 \lambda_1'} \partial^{\lambda_1'} \partial_{\eta_1 \mu_1} \partial^{\mu_1} D^{(0)} = \epsilon_{\nu_1 \nu} \epsilon_{\eta_1 \eta} \Box\Box D^{(0)};$$

since $\Box D^{(0)} = 0$, the right-hand side of (10.8) is zero. If the bracket in (10.8) is $[\quad]_-$, it follows from the first lemma of Sec. 5 that

$$\chi_{\nu_1 \cdots \nu_k \eta \lambda_{k+1} \cdots \lambda_a \mu_2 \cdots \mu_b} = 0; \tag{10.9}$$

if the bracket in (10.8) is $[\quad]_+$, then the same conclusion follows in an even more elementary way, since $AA^* + A^*A = 0$ evidently implies that $A = 0$. Thus, we learn in any case that

$$\partial_{\nu_1}{}^{\lambda_1} \cdots \partial_{\nu_k}{}^{\lambda_k} \partial_\eta{}^{\mu_1} \psi_{\lambda_1 \cdots \lambda_a \mu_1 \cdots \mu_b} = 0. \tag{10.10}$$

Then, differentiating (10.6) and using (10.10), it follows that

$$\sum_{n=0}^{k} \sum c_n \partial_{\nu_1 \lambda_1'} \cdots \partial_{\nu_n \lambda_n'} \partial_{\nu_{n+1}}{}^{\lambda_{n+1}} \cdots \partial_{\nu_k}{}^{\lambda_k} \partial_{\eta \mu_1'} \times \epsilon_{\mu_2 \mu_2'} \cdots \epsilon_{\mu_n \mu_n'} \partial_{\lambda_{n+1} \cdots \lambda_b' \mu_{n+1} \cdots \mu_a'} D^{(0)} = 0. \tag{10.11}$$

Now, every term of the left-hand side of (10.11) except the term $n=k$ contains a factor of the form

$$\partial_{\nu_{n+1}}{}^{\lambda_{n+1}}\partial_{\lambda_{n+1}}\cdots{}_{\lambda_b'\mu_{n+1}}\cdots{}_{\mu_b'}D^{(0)}.$$

Thus, again since $\Box D^{(0)}=0$, all these terms are zero and (10.11) reduces to the formula

$$c_k\sum\partial_{\nu_1\lambda_1'}\cdots\partial_{\nu_k\lambda_k'}\partial_{\eta\mu_1'}\epsilon_{\mu_2\mu_2'}\cdots\epsilon_{\mu_k\mu_k'}\times\partial_{\lambda_{k+1}}\cdots{}_{\lambda_b'\mu_{k+1}}\cdots{}_{\mu_a'}D^{(0)}=0, \quad (10.12)$$

the symbol $\sum$ denoting summation over all permutations of the indices $\lambda_1'\cdots\lambda_b'$ and over all permutations of the indices $\mu_1'\cdots\mu_a'$. If we apply the differential operator $\partial_{\eta_2}{}^{\mu_2'}\cdots\partial_{\eta_k}{}^{\mu_k'}$ to the left of formula (10.12), we consequently find that

$$c_k\partial_{\nu_1}\cdots{}_{\nu_k\lambda_{k+1}}\cdots{}_{\lambda_b'\eta\eta_2}\cdots{}_{\eta_k\mu_{k+1}}\cdots{}_{\mu_a'}D^{(0)}=0,$$

i.e., that

$$c_k\partial_{j_1}\cdots\partial_{j_{a+b}}D^{(0)}\equiv 0.$$

Since

$$\partial_{j_1}\cdots\partial_{j_{a+b}}D^{(0)}\neq 0,$$

it follows that $c_k=0$ if $k>0$, which is what we desired to prove.

This established, we see at once from (10.4) that

$$[\psi_{\lambda_1}\cdots{}_{\lambda_a\mu_1}\cdots{}_{\mu_b},\psi_{\mu_1'}\cdots{}_{\mu_a'\lambda_1'}\cdots{}_{\lambda_b'}{}^*]=c\partial_{\lambda_1}\cdots{}_{\lambda_b'\mu_1}\cdots{}_{\mu_a'}D^{(0)}. \quad (10.13)$$

Thus, if we put

$$\chi_{\eta\lambda_2}\cdots{}_{\lambda_a\mu_1}\cdots{}_{\mu_b}=\partial_\eta{}^{\lambda_1}\psi_{\lambda_1}\cdots{}_{\lambda_a\mu_1}\cdots{}_{\mu_b},$$

it follows just as before that

$$[\chi_{\eta\lambda_2}\cdots{}_{\lambda_a\mu_1}\cdots{}_{\mu_b},\chi_{\nu\mu_2'}\cdots{}_{\mu_a'\lambda_1'}\cdots{}_{\lambda_b'}{}^*]=0,$$

and hence again as before that

$$\partial_\eta{}^{\lambda_1}\psi_{\lambda_1}\cdots{}_{\lambda_a\mu_1}\cdots{}_{\mu_b}=0. \quad (10.14)$$

We may prove in exactly the same way that

$$\partial_\nu{}^{\mu_1}\psi_{\lambda_1}\cdots{}_{\lambda_a\mu_1}\cdots{}_{\mu_b}=0. \quad (10.15)$$

Thus, every field component satisfying the second-order equation $\Box\psi=0$ must also satisfy the first-order equations (10.14) and (10.15). Our next aim is to use Eqs. (10.14) and (10.15) to show that if a and b are both nonzero, then there exists a system of fields $\chi_{\lambda_2}\cdots{}_{\lambda_a\mu_2}\cdots{}_{\mu_b}$ such that

$$\psi_{\lambda_1}\cdots{}_{\lambda_a\mu_1}\cdots{}_{\mu_b}=\partial_{\lambda_1\mu_1}\chi_{\lambda_2}\cdots{}_{\lambda_a\mu_2}\cdots{}_{\mu_b}. \quad (10.16)$$

This is most readily accomplished by passing to the Fourier transformed field $\psi_{\lambda_1}\cdots{}_{\lambda_a\mu_1}\cdots{}_{\mu_b}(p)$, which by virtue of (10.14) and (10.15) satisfies

$$p_\eta{}^{\lambda_1}\psi_{\lambda_1}\cdots{}_{\lambda_a\mu_1}\cdots{}_{\mu_b}(p)=p_\nu{}^{\mu_1}\psi_{\lambda_1}\cdots{}_{\lambda_a\mu_1}\cdots{}_{\mu_b}(p)\equiv 0. \quad (10.17)$$

On noting that the distribution $\psi(p)$ is carried by the light-cone $p^2=0$ in virtue of the equation $\Box\psi(x)=0$, we realize that formula (10.17) involves the matrix $P=(p_\mu{}^\lambda)=(p_{\mu\lambda'}\epsilon^{\lambda'\lambda})$, where $p_{\mu\lambda'}=p_j\sigma_{\mu\lambda'}{}^j$ for null vectors (p_0,p_1,p_2,p_3): Since $p_\mu{}^\lambda p_\lambda{}^\mu=\delta_\lambda{}^\lambda p^2=0$, it follows that P is a 2×2 matrix satisfying $P^2=0$, $P\neq 0$. Such a 2×2 matrix may be put in the Jordan canonical form

$$\begin{bmatrix}0 & 0\\ 1 & 0\end{bmatrix}.$$

On taking P in this canonical form, it is clear that any 2×2 matrix A satisfying $PA=0$ has all the elements in its first row equal to zero, while any 2×2 matrix A satisfying $AP=0$ has all the elements in its second column equal to zero. Thus, if $AP=PA=0$, A must be a multiple cP of P; it is self-evident that this constant c is uniquely determined. Equations (10.17) consequently give us

$$\psi_{\lambda_1}\cdots{}_{\lambda_a\mu_1}\cdots{}_{\mu_b}(p)=ip_{\lambda_1\mu_1}\chi_{\lambda_2}\cdots{}_{\lambda_a\mu_2}\cdots{}_{\mu_b}(p) \quad (10.18)$$

with a certain uniquely determined coefficient function χ. The unique determination of χ, together with the Lorentz invariance of the field ψ, shows immediately that χ does have the Lorentz-invariant law of transformation [corresponding to the representation $(a-1)\otimes b_1$, where $b_1=b-1$, of the proper Lorentz group indicated by its subscripts].

On iterating the argument of the preceding paragraph a number of times equal to $\min(a,b)$, and passing back from the Fourier transformed fields to the fields in ordinary space, we find that we have established the following conclusion: If $a\geq b$, there exists a field $\chi_{\lambda_1}\cdots{}_{\lambda_{a-b}}$ having the indicated law of transformation under the Lorentz group such that

$$\psi_{\lambda_1}\cdots{}_{\lambda_a\mu_1}\cdots{}_{\mu_b}=\partial_{\lambda_1\mu_1}\partial_{\lambda_2\mu_2}\cdots\partial_{\lambda_b\mu_b}\chi_{\lambda_{b+1}}\cdots{}_{\lambda_a}; \quad (10.19)$$

if $b>a$, an exactly corresponding formula may be written in terms of a field $\chi_{\mu_1}\cdots{}_{\mu_{b-a}}$; that is, the fields ψ are gradients of the fields χ.

Our next aim is to investigate the commutation relations of the fields χ. Let $\{\psi_{\lambda_1}\cdots{}_{\lambda_a\mu_1}\cdots{}_{\mu_b}\}$ be the class of field components occurring in formulas (10.13)–(10.19), and let $\{\psi_\alpha\}$ be any other class of field components. Then it follows [cf. (10.4)] that

$$[\psi_{\lambda_1}\cdots{}_{\lambda_a\mu_1}\cdots{}_{\mu_b}(x),\psi_\alpha(x')]=a_{\lambda_1}\cdots{}_{\lambda_a\mu_1}\cdots{}_{\mu_b,\alpha}(\partial)D^0(x-x'), \quad (10.20)$$

where $a(\partial)$ denotes a polynomial in the gradient ∂. Hence, passing to the Fourier-transformed fields, we find

$$[\psi_{\lambda_1}\cdots{}_{\lambda_a\mu_1}\cdots{}_{\mu_b}(p),\psi_\alpha(p')]=a_{\lambda_1}\cdots{}_{\lambda_a\mu_1}\cdots{}_{\mu_b,\alpha}(ip)\times\delta(-p,p')D^{(0)}(p) \quad (10.21)$$

for all $p,p'\epsilon h^{(0)}$ as in (8.26), cf. also (8.24), (8.18), and (8.19). This being the case, (10.17) evidently implies that

$$p_\mu{}^{\lambda_1}a_{\lambda_1}\cdots{}_{\lambda_a\mu_1}\cdots{}_{\mu_b,\alpha}(ip)=0=p^{\lambda\mu_1}a_{\lambda_1}\cdots{}_{\lambda_a\mu_1}\cdots{}_{\mu_b,\alpha}(ip); \quad (10.22)$$

so that, arguing precisely as in the paragraph following

formula (10.17), we find that a may be written in the form

$$a_{\lambda_1\cdots\lambda_a\mu_1\cdots\mu_b,\alpha}(ip) = ip_{\lambda_1\mu_1}b_{\lambda_2\cdots\lambda_a\mu_2\cdots\mu_b,\alpha}(ip), \quad (10.23)$$

the b being certain coefficients. Then, using (10.18) and the fact that $p_\lambda\mu\neq 0$, we find from (10.21) that the field $\chi_{\lambda_2\cdots\lambda_a\mu_2\cdots\mu_b}$ of (10.18) satisfies

$$[\chi_{\lambda_2\cdots\lambda_a\mu_2\cdots\mu_b}(p),\psi_\alpha(p')] = b_{\lambda_2\cdots\lambda_a\mu_2\cdots\mu_b,\alpha}(ip)\delta(-p,p')D^{(0)}(p). \quad (10.24)$$

By iterating this argument a suitable number of times, and passing back from the Fourier-transformed fields to the fields in ordinary space, we find that the field $\chi_{\lambda_1\cdots\lambda_{a-b}}$ of (10.19) satisfies

$$[\chi_{\lambda_1\cdots\lambda_{a-b}}(x),\psi_\alpha(x')] = c_{\lambda_1\cdots\lambda_{a-b},\alpha}(\partial)D^{(0)}(x-x'), \quad (10.25)$$

where $c(\partial)$ denotes a polynomial in the gradient ∂. This shows that the fields χ satisfy commutation relations of the same form as do the fields ψ.

If we replace each of the fields ψ satisfying the zero-mass equation $\Box\psi=0$ by the corresponding field χ, we consequently obtain a field theory which continues to satisfy the axioms of Sec. 3. By virtue of (10.19), this new theory is semiequivalent to the theory with which we begin. It follows that our original free-field theory is semiequivalent to one in which all the field operators which satisfy the equation $\Box\psi=0$ transform according to a representation of the Lorentz group having either the form $(k)\otimes(\dot{0})$ or the complex-conjugate form $(0)\otimes(\dot{j})$. Our fields will also satisfy $\partial_\mu{}^{\lambda_1}\psi_{\lambda_1\cdots\lambda_k}\equiv 0$ and $\partial^{\mu_1}\psi_{\mu_1\cdots\mu_j}\equiv 0$; in addition, they may without loss of generality be assumed linearly independent.

We may now carry over the development given in Sec. 8 for fields of positive mass to the components of mass zero.

It follows, exactly as in the corresponding passage in Sec. 8 [cf. formula (8.14) and the following material] that $[\psi_\alpha(x),\psi_\beta(x')]\equiv 0$ unless either both classes of indices α and β transform according to the same representation or to complex-conjugate representations of the proper Lorentz group. In the former case it would follow exactly as in (8.14) that

$$[\psi_{\lambda_1\cdots\lambda_n}(x),\psi'_{\lambda_1'\cdots\lambda_n'}(x')] \equiv c\epsilon_{\lambda_1\cdots\lambda_n;\,\lambda_1'\cdots\lambda_n'}D^{(0)}(x-x'), \quad (10.26)$$

where c is some complex constant. But then, if $n>0$, the equation $\partial_\mu{}^{\lambda_1}\psi_{\lambda_1\cdots\lambda_n}\equiv 0$ would immediately imply $c=0$. Thus we have

$$[\psi_{\lambda_1\cdots\lambda_n}(x),\psi'_{\lambda_1'\cdots\lambda_n'}(x')]\equiv 0 \quad \text{if } n>0. \quad (10.27)$$

If $n=0$, we can decompose the field ψ into its Hermitian and anti-Hermitian parts, and hence, passing to an equivalent theory, suppose without loss of generality that $\psi=\psi^*$. Of the expression

$$[\psi_{\lambda_1\cdots\lambda_n}(x),\bar{\psi}_{\mu_1\cdots\mu_n}{}^*(x')],$$

the Clebsch-Gordan formula tells us that the right-hand side of the general formula (10.4) can have only one nonzero term, this term involving exactly n derivatives; and the uniqueness of the required expression tells us at once that

$$[\psi_{\lambda_1\cdots\lambda_n}(x),\bar{\psi}_{\mu_1\cdots\mu_n}{}^*(x')] = c\partial_{\lambda_1\cdots\lambda_n\mu_1\cdots\mu_n}D^{(0)}(x-x'), \quad (10.28)$$

c being some complex constant. Then, arguing exactly as in Sec. 8, we show that

(a) (Pauli theorem) If the field component ψ_α transforms according to the representation $(n)\times(\dot{0})$ of the proper Lorentz group, then $\sigma(\alpha,\dot{\alpha})=(-1)^{n+1}$.

(b) Upon passing to an equivalent theory, we may suppose that $[\psi_\alpha,\psi_\beta]=0$ if α and β lie in distinct classes of indices, while if $\{\psi_{\lambda_1\cdots\lambda_n}\}=\{\psi_\alpha\}$ is a given class of fields with $m(\alpha)=0$, we have

$$[\psi_{\lambda_1\cdots\lambda_n}(x),\psi_{\lambda_1'\cdots\lambda_n'}(x')]=0 \quad (\text{if } n>0) \quad (10.29)$$

$$[\psi_{\lambda_1\cdots\lambda_n}(x),\psi^*{}_{\mu_1\cdots\mu_n}(x')] = i^{n+1}\partial_{\lambda_1\cdots\lambda_n,\mu_1\cdots\mu_n}D^{(0)}(x-x'), \quad (10.30)$$

$$\text{if } n=0 \text{ (10.30) and } \psi(x)=\psi^*(x) \text{ hold.} \quad (10.31)$$

Thus, the theorem stated in Sec. 4 is completely proved.

11. SPECIAL EXAMPLES OF FIELDS OF ZERO MASS

The simplest example of a field of zero mass with spin is the field ψ_λ transforming according to the representation $(1)\times(\dot{0})$ of the Lorentz group, and satisfying the equations

$$\begin{aligned} \partial_\mu{}^\lambda\psi_\lambda &= 0 \\ [\psi_\lambda,\psi_{\lambda'}]_+ &= 0 \\ [\psi_{\lambda 1}\psi_\mu{}^*]_+ &= -\partial_{\lambda\mu}D^{(0)}. \end{aligned} \quad (11.1)$$

This is the two-component neutrino theory; the first of Eqs. (11.1) tells us that the "particle" is always polarized parallel to its momentum, while the "anti-particle" is always polarized opposite to its momentum.

An attempt to set up a theory of a vector field A_j satisfying $\Box A_j\equiv 0$ leads, as in the preceding section, to the conclusion that A_j is the gradient of a scalar field: $A_j=\partial_j A$. Thus, as is well known, the electromagnetic vector potential cannot be quantized if all the axioms of Sec. 3 are required. Either positivity of the energy, positivity of the inner product of Hilbert space, or the status of the field components A_j as operators mapping Hilbert space into itself, rather than into a larger enveloping space, must be given up.

To obtain a quantization of the free electromagnetic field satisfying the axioms of Sec. 3, it is preferable to quantize the field-strength tensor f_{ij} instead of the vector potential. The representation of the Lorentz

group by an antisymmetric tensor f_{ij} decomposes, on restriction to the proper Lorentz group, into the direct sum of the two representations (2) and $(\bar{2})$. The corresponding quantities $\psi_{\lambda\lambda'}$ and $\chi_{\mu\mu'}$ are expressed in terms of f_{ij} by the formulas

$$\begin{aligned}\psi_{\lambda\lambda'}&=\epsilon^{\mu\mu'}\sigma_{\lambda\mu}{}^{i}\sigma_{\lambda'\mu'}{}^{j}f_{ij}\\ \chi_{\mu\mu'}&=\epsilon^{\lambda\lambda'}\sigma_{\lambda\mu}{}^{i}\sigma_{\lambda'\mu'}{}^{j}f_{ij}.\end{aligned}\tag{11.2}$$

The reality condition $f_{ij}{}^{*}=f_{ij}$ is expressible in terms of ψ and χ as

$$\begin{aligned}\psi_{\nu\nu'}{}^{*}&=\epsilon^{\eta\eta'}\bar{\sigma}_{\nu\eta}{}^{i}\bar{\sigma}_{\nu'\eta'}{}^{j}f_{ij}{}^{*}\\ &=\epsilon^{\eta\eta'}\sigma_{\eta\nu}{}^{i}\sigma_{\eta'\nu'}{}^{j}f_{ij}=\chi_{\nu\nu'}.\end{aligned}$$

Thus, the quantized electromagnetic field is determined in our notation by a field $\psi_{\lambda\lambda'}$ and its adjoint, transforming according to the representation (11.2) of the Lorentz group, and satisfying the equations

$$\begin{aligned}\partial_{\mu}{}^{\lambda}\psi_{\lambda\lambda'}&\equiv 0\\ [\psi_{\lambda\lambda_1},\psi_{\lambda'\lambda_1'}]_{-}&\equiv 0\\ [\psi_{\lambda\lambda'},\psi_{\mu\mu_1}{}^{*}]_{-}&\equiv -i\partial_{\lambda\lambda_1,\mu\mu_1}D^{(0)}.\end{aligned}\tag{11.3}$$

The first of Eqs. (11.3) tells us that the "photon" is always polarized parallel to its momentum, while the "antiphoton" is always polarized opposite to its momentum.

12. EQUATIONS SATISFIED BY THE FIELDS IN CANONICAL FORM

In the present section, we assume that we have a system of fields in the canonical form specified in the theorem of Sec. 4, and consider the following question: Does our system of field operators satisfy any Lorentz-invariant system

$$\sum_{k=0}^{K} a_{\beta,j_1\cdots j_k}{}^{\alpha}\,\partial^{j_1}\cdots\partial^{j_k}\psi_{\alpha}=0\tag{12.1}$$

of partial differential equations with constant coefficients? If Eq. (12.1) is valid, then since $\psi^{(m)}$ can be written in the form $\psi^{(m)}=P^{(m)}(\Box)\psi$ (cf. Sec. 5), so evidently is each equation

$$\sum_{k=0}^{K} a_{\beta,j_1\cdots j_k}{}^{\alpha}\,\partial^{j_1}\cdots\partial^{j_k}\psi_{\alpha}{}^{(m)}=0,\quad m\geq 0.\tag{12.2}$$

Thus, without loss of generality we may suppose that all the field operators occuring with nonzero coefficients in Eq. (80) satisfy the equation $\Box\psi_{\alpha}=m^2\psi_{\alpha}$, $m\geq 0$. For Lorentz invariance, the index β in Eq. (12.2) must belong to some representation of the Lorentz group $\mathfrak{L}_p$. At the possible cost of writing the single equation (12.2) as a set of several equations, we may evidently suppose without loss of generality that the index β belongs to an irreducible representation of $\mathfrak{L}_p$; say, for definiteness, to the representation $(m)\times(\bar{n})$. The fields ψ_{α} will consist of a set of fields $\psi_{\lambda_1\cdots\lambda_p}{}^{(R)}$, and a set of fields $\psi_{\mu_1\cdots\mu_p}{}^{*(R)}=(\psi_{\lambda_1\cdots\lambda_p}{}^{(R)})^{*}$. Here of course $p=p(R)$. Equation (12.1) may thus be written more explicitly as

$$\begin{aligned}\sum_{R=1}^{S}\sum_{k=0}^{K}&(\alpha_{\beta,j_1\cdots j_k}{}^{\lambda_1\cdots\lambda_p,R}\partial^{j_1}\cdots\partial^{j_k}\psi_{\lambda_1\cdots\lambda_p}{}^{(R)}\\ &+b_{\beta,j_1\cdots j_k}{}^{\mu_1\cdots\mu_p,R}\,\partial^{j_1}\cdots\partial^{j_k}\psi_{\mu_1\cdots\mu_p}{}^{*(R)})=0.\end{aligned}\tag{12.3}$$

First suppose that $m>0$. If we apply to this equation, and in particular to its index β, the process of reduction used previously to express arbitrary field operators ψ_{α} satisfying $\Box\psi_{\alpha}=m^2\psi_{\alpha}$, $m>0$, as Lorentz-invariant linear combinations with constant coefficients of gradients of fields belonging to representations $(m)\times(\bar{0})$ of the Lorentz group, we see immediately that the Eq. (12.3) may be written as a Lorentz-invariant linear combination with constant coefficients of gradients of equations having exactly the same form, but in which the index β belongs to a representation $(m)\times(\bar{0})$ of $\mathfrak{L}_p$. It is then evidently sufficient for the subsequent course of our analysis to suppose that the index β in Eq. (12.3) itself belongs to a representation $(m)\times(\bar{0})$ of $\mathfrak{L}_p$.

Now, since $\Box\psi=m^2\psi$, we may evidently suppose that the coefficients a and b in Eq. (12.3) are symmetric in $j_1\cdots j_k$ and also satisfy the condition

$$\begin{aligned}a_{\beta,j_1j_2\cdots j_k}{}^{\lambda_1\cdots\lambda_p,R}G^{j_1j_2}&\\ =b_{\beta,j_1j_2\cdots j_k}{}^{\mu_1\cdots\mu_p,R}G^{j_1j_2}&=0\end{aligned}\tag{12.4}$$

of "vanishing trace," $G^{j_1j_2}$ being the Lorentz metric tensor. Thus, the invariant coefficients $a_{\beta,j_1\cdots j_k}{}^{\lambda_1\cdots\lambda_p,R}$ and $b_{\beta,j_1\cdots j_k}{}^{\mu_1\cdots\mu_p,R}$ are scalars belonging to the representation $(k)\times(\bar{k})\times(m)\times(p)$ and $(k)\times(\bar{k})\times(m)\times(\bar{p})$, respectively. We find immediately by the Clebsch-Gordan formula that $a_{\beta,j_1j_2\cdots j_k}{}^{\lambda_1\cdots\lambda_p,R}=0$ unless $k=0$ and $m=p$, in which case a must be a multiple of $\epsilon_{\lambda_1'\cdots\lambda_p'\lambda_1\cdots\lambda_p}$; and that $b_{\beta,j_1j_2\cdots j_k}{}^{\mu_1\cdots\mu_p,R}=0$ unless $k=m=p$, in which case β is a set of k indices $\lambda_1\cdots\lambda_k$ and the differential operator $b_{\beta,j_1\cdots j_k}{}^{\mu_1\cdots\mu_p,R}\times\partial^{j_1}\cdots\partial^{j_k}$ is necessarily a constant multiple of

$$\partial_{\lambda_1}{}^{\mu_1}\partial_{\lambda_2}{}^{\mu_2}\cdots\partial_{\lambda_p}{}^{\mu_p}.\tag{12.5}$$

Thus, the differential equation (12.3) can be written in the form

$$\sum_{R=1}^{S} a^{r}\psi_{\lambda_1\cdots\lambda_p}{}^{(R)}+\sum_{R=1}^{S} b^{r}\partial_{\lambda_1}{}^{\mu_1}\cdots\partial_{\lambda_p}{}^{\mu_p}\psi_{\mu_1\cdots\mu_p}{}^{*(r)},\tag{12.6}$$

only fields $\psi^{(R)}$ with $p(R)=k$ occuring with a nonzero coefficient. If we now make use of the reality conditions, then this equation evidently states a linear dependence between the fields $\psi_{\lambda_1\cdots\lambda_p}{}^{(R)}$, and by assumption no nontrivial linear dependence exists. It follows that Eq. (12.6) is a linear combination of the Eqs. (8.11) and (8.12) giving the reality conditions for the fields $\psi^{(R)}$. Thus, every Lorentz-invariant partial differential equation with constant coefficients satisfied by the system of fields ψ_{α} with $\Box\psi_{\alpha}=m^2\psi_{\alpha}$, $m>0$ is obtainable

as a consequence from the equations $\Box\psi_\alpha=m^2\psi_\alpha$ and from the reality conditions by a combination of repeated differentiation and purely linear algebraic operations.

If $m=0$, we reason as follows. If we make use of the equations $\partial_\mu{}^\lambda\psi_\lambda\cdots=0$ and $\partial_\lambda{}^\mu\psi_\mu\cdots=0$ satisfied by the field components of mass zero to subtract terms from Eq. (12.3), we must arrive at an equation having the form

$$\sum_{R=1}^{S}\sum_{k=0}^{K}(a_\beta{}^{\lambda_1\cdots\lambda_{k+p},\mu_1\cdots\mu_k}\partial_{\lambda_1\cdots\lambda_k,\mu_1\cdots\mu_k}\psi_{\lambda_{k+1}\cdots\lambda_{k+p}}{}^{(R)}+b_\beta{}^{\lambda_1\cdots\lambda_k,\mu_1\cdots\mu_{k+p}}\partial_{\lambda_1\cdots\lambda_k,\mu_1\cdots\mu_k}\times\psi_{\mu_{k+1}\cdots\mu_{k+p}}{}^{(R)})=0. \quad (12.7)$$

Now the index β belongs to the representation $(m)\otimes(\bar{n})$ of $\mathfrak{L}_p$. The Clebsch-Gordan formula consequently tells us that every term in the Lorentz-invariant equation (12.7), except those for which $k+p=m$ and $k=n$ or $k+p=n$ and $k=m$, must vanish, and, indeed, that Eq. (12.7) must reduce either to the form

$$\sum_{R=0}^{S}\partial_{\lambda_1\cdots\lambda_k,\mu_1\cdots\mu_k}(c_R\psi_{\lambda_{k+1}\cdots\lambda_{k+p}}{}^{(R)})=0, \quad (12.8)$$

where $c_R=0$ unless the field $\psi^{(R)}$ transforms according to a representation $(p)\times(\bar{0})$ with given p, or to the form

$$\sum_{R=0}^{S}\partial_{\lambda_1\cdots\lambda_k,\mu\cdots\mu_k}(c_R\psi^{*(R)}{}_{\mu_{k+1}\cdots\mu_{k+p}})=0, \quad (12.9)$$

c_R being subject to this same restriction, or, if $p=0$, to the form

$$\sum_{R=0}^{S}\partial_{\lambda_1\cdots\lambda_k,\mu_1\cdots\mu_k}(c_R\psi^{(R)}+d_R\psi^{(R)*})=0, \quad (12.10)$$

where c_R and d_R equal zero unless the field $\psi^{(R)}$ transforms according to the representation $(0)\times(\bar{0})$. Now, unless all the coefficients c_R are zero, Eq. (12.8) evidently implies a linear dependence between the fields $\psi^{(R)}$; since we have assumed these fields are linearly independent, it follows that all the coefficients c_R in Eq. (12.8) are zero. In the same way, it follows that all the coefficients c_R in Eq. (12.9) are zero. In the case of Eq. (12.10), the same conclusion follows in the same way, once we make use of the reality condition $\psi=\psi^*$.

Summarizing, we may state the following theorem.

Theorem. Let $\{\psi_\alpha\}$ be a set of free fields defining a free-field theory satisfying the axioms of Sec. 3, and suppose that these fields are in the canonical form specified in the theorem of Sec. 4. Then every Lorentz-invariant partial differential equation with constant coefficients satisfied by our system of fields, is obtainable as a consequence from the equations

(i) $m^{-n}\partial_{\mu_1}{}^{\lambda_1}\cdots\partial_{\mu_n}{}^{\lambda_n}\psi_{\lambda_1\cdots\lambda_n}=\psi_{\mu_1\cdots\mu_n}{}^*$ (if $\Box\psi=m^2$ or if $n=0$),

$m^{-n}\partial_{\lambda_1}{}^{\mu_1}\cdots\partial_{\lambda_n}{}^{\mu_n}\psi_{\mu_1\cdots\mu_n}{}^*=\psi_{\lambda_1\cdots\lambda_n}$

(ii) $\partial_\mu{}^{\lambda_1}\psi_{\lambda_1\cdots\lambda_n}=0=\partial^{\mu_1}\psi_{\mu_1\cdots\mu_n}$ (if $\Box\psi=0$ and $n>0$)

(iii) $\psi=\psi^*$ (if $\Box\psi=0$ and $n=0$)

by a combination of repeated differentiation and purely linear algebraic operations.

ACKNOWLEDGMENTS

In the course of preparing the present article, the author has had invaluable assistance from K. O. Friedrichs and B. Zumino.

The work was supported by a grant from the ONR.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 2, NUMBER 3 MAY-JUNE, 1961

Null Electromagnetic Fields*

IVOR ROBINSON
Department of Physics, Syracuse University, Syracuse, New York
(Received September 9, 1960)

It is shown that a field of null rays is geodetic and shear-free if and only if the associated family of null bivectors includes a solution to Maxwell's equations for charge-free space.

A BIVECTOR $F^{kl}[=-F^{lk}]$, in a normal hyperbolic Riemann four-space, is said to be *null* if $F_{kl}F^{kl}=0=F_{kl}{}^*F^{kl}$, where ${}^*F_{kl}=\frac{1}{2}(-g)^{\frac{1}{2}}\epsilon_{klmn}F^{mn}$. For any real nonzero null bivector F_{kl}, the conditions[1]

$$F_{[kl}\sigma_{m]}=0, \quad F_{km}\sigma^m=0 \quad (1)$$

determine a real null direction: conversely, for any real nonzero null vector σ_k, these equations determine a real nonzero null bivector up to a change of amplitude and polarization,

$$F_{kl}-i^*F_{kl}\rightarrow e^w(F_{kl}-i^*F_{kl}),$$

where w is a disposable complex scalar. This note deals with the question: What conditions on a field of null rays are necessary and sufficient to ensure that the

* Supported by the U. S. Air Force Office of Scientific Research of the Air Research and Development Command.

[1] The operations of symmetrization and antisymmetrization are denoted by round and square brackets, respectively.

as a consequence from the equations $\Box\psi_\alpha=m^2\psi_\alpha$ and from the reality conditions by a combination of repeated differentiation and purely linear algebraic operations.

If $m=0$, we reason as follows. If we make use of the equations $\partial_\mu{}^\lambda\psi_\lambda\cdots=0$ and $\partial_\lambda{}^\mu\psi_\mu\cdots=0$ satisfied by the field components of mass zero to subtract terms from Eq. (12.3), we must arrive at an equation having the form

$$\sum_{R=1}^{S}\sum_{k=0}^{K}(a_\beta{}^{\lambda_1\cdots\lambda_{k+p},\mu_1\cdots\mu_k}\partial_{\lambda_1\cdots\lambda_k,\mu_1\cdots\mu_k}\psi_{\lambda_{k+1}\cdots\lambda_{k+p}}{}^{(R)}+b_\beta{}^{\lambda_1\cdots\lambda_k,\mu_1\cdots\mu_{k+p}}\partial_{\lambda_1\cdots\lambda_k,\mu_1\cdots\mu_k}\times\psi_{\mu_{k+1}\cdots\mu_{k+p}}{}^{(R)})=0. \quad (12.7)$$

Now the index β belongs to the representation $(m)\otimes(\bar{n})$ of $\mathfrak{L}_p$. The Clebsch-Gordan formula consequently tells us that every term in the Lorentz-invariant equation (12.7), except those for which $k+p=m$ and $k=n$ or $k+p=n$ and $k=m$, must vanish, and, indeed, that Eq. (12.7) must reduce either to the form

$$\sum_{R=0}^{S}\partial_{\lambda_1\cdots\lambda_k,\mu_1\cdots\mu_k}(c_R\psi_{\lambda_{k+1}\cdots\lambda_{k+p}}{}^{(R)})=0, \quad (12.8)$$

where $c_R=0$ unless the field $\psi^{(R)}$ transforms according to a representation $(p)\times(\bar{0})$ with given p, or to the form

$$\sum_{R=0}^{S}\partial_{\lambda_1\cdots\lambda_k,\mu\cdots\mu_k}(c_R\psi^{*(R)}{}_{\mu_{k+1}\cdots\mu_{k+p}})=0, \quad (12.9)$$

c_R being subject to this same restriction, or, if $p=0$, to the form

$$\sum_{R=0}^{S}\partial_{\lambda_1\cdots\lambda_k,\mu_1\cdots\mu_k}(c_R\psi^{(R)}+d_R\psi^{(R)*})=0, \quad (12.10)$$

where c_R and d_R equal zero unless the field $\psi^{(R)}$ transforms according to the representation $(0)\times(\bar{0})$. Now, unless all the coefficients c_R are zero, Eq. (12.8) evidently implies a linear dependence between the fields $\psi^{(R)}$; since we have assumed these fields are linearly independent, it follows that all the coefficients c_R in Eq. (12.8) are zero. In the same way, it follows that all the coefficients c_R in Eq. (12.9) are zero. In the case of Eq. (12.10), the same conclusion follows in the same way, once we make use of the reality condition $\psi=\psi^*$.

Summarizing, we may state the following theorem.

Theorem. Let $\{\psi_\alpha\}$ be a set of free fields defining a free-field theory satisfying the axioms of Sec. 3, and suppose that these fields are in the canonical form specified in the theorem of Sec. 4. Then every Lorentz-invariant partial differential equation with constant coefficients satisfied by our system of fields, is obtainable as a consequence from the equations

(i) $m^{-n}\partial_{\mu_1}{}^{\lambda_1}\cdots\partial_{\mu_n}{}^{\lambda_n}\psi_{\lambda_1\cdots\lambda_n}=\psi_{\mu_1\cdots\mu_n}{}^*$ (if $\Box\psi=m^2$ or if $n=0$),

$m^{-n}\partial_{\lambda_1}{}^{\mu_1}\cdots\partial_{\lambda_n}{}^{\mu_n}\psi_{\mu_1\cdots\mu_n}{}^*=\psi_{\lambda_1\cdots\lambda_n}$

(ii) $\partial_\mu{}^{\lambda_1}\psi_{\lambda_1\cdots\lambda_n}=0=\partial^{\mu_1}\psi_{\mu_1\cdots\mu_n}$ (if $\Box\psi=0$ and $n>0$)

(iii) $\psi=\psi^*$ (if $\Box\psi=0$ and $n=0$)

by a combination of repeated differentiation and purely linear algebraic operations.

ACKNOWLEDGMENTS

In the course of preparing the present article, the author has had invaluable assistance from K. O. Friedrichs and B. Zumino.

The work was supported by a grant from the ONR.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 2, NUMBER 3 MAY-JUNE, 1961

Null Electromagnetic Fields*

IVOR ROBINSON
Department of Physics, Syracuse University, Syracuse, New York
(Received September 9, 1960)

It is shown that a field of null rays is geodetic and shear-free if and only if the associated family of null bivectors includes a solution to Maxwell's equations for charge-free space.

A BIVECTOR $F^{kl}[=-F^{lk}]$, in a normal hyperbolic Riemann four-space, is said to be *null* if $F_{kl}F^{kl}=0=F_{kl}{}^*F^{kl}$, where ${}^*F_{kl}=\frac{1}{2}(-g)^{\frac{1}{2}}\epsilon_{klmn}F^{mn}$. For any real nonzero null bivector F_{kl}, the conditions[1]

$$F_{[kl}\sigma_{m]}=0, \quad F_{km}\sigma^m=0 \quad (1)$$

determine a real null direction: conversely, for any real nonzero null vector σ_k, these equations determine a real nonzero null bivector up to a change of amplitude and polarization,

$$F_{kl}-i^*F_{kl}\rightarrow e^w(F_{kl}-i^*F_{kl}),$$

where w is a disposable complex scalar. This note deals with the question: What conditions on a field of null rays are necessary and sufficient to ensure that the

* Supported by the U. S. Air Force Office of Scientific Research of the Air Research and Development Command.

[1] The operations of symmetrization and antisymmetrization are denoted by round and square brackets, respectively.

associated family of null bivector fields should include a nontrivial solution to Maxwell's equations for charge-free space?

If a vector field σ_k is real, nonvanishing, and null, the tangent space at any point contains a two-parameter family of real spacelike planes orthogonal to σ_k. For any such plane, one can calculate the rates of rotation, dilatation, and maximum shear under a displacement along σ_k. By this means, one obtains three scalar functions of the coordinates and parameters. If any one of these scalars is independent of the parameters, so are the other two. A necessary and sufficient condition for this is that σ_k should be geodetic. In that case, the scalars may be introduced more directly as follows. On denoting the covariant differential operator $\sigma^p\delta/\delta x^p$ by a dot, writing

$$\omega_k=\{\sigma_{[k;\,l]}-i^*(\sigma_{[k;\,l]})\}\sigma^l,$$

and using the expression for $g\epsilon_{abcd}\epsilon_{efgh}$ in terms of the metric tensor, one finds that if σ_k is null,

$$2\omega_{[k}\sigma_{l]}=\sigma_{[k}\dot{\sigma}_{l]}+i^*(\sigma_{[k}\dot{\sigma}_{l]}).$$

Thus, *a real null vector field σ_k satisfies the geodetic condition, $\sigma_{[k}\dot{\sigma}_{l]}=0$, if and only if there exists a complex scalar ζ such that*

$$\{\sigma_{[k;\,l]}-i^*(\sigma_{[k;\,l]})\}\sigma^l=(\tfrac{1}{2}\sigma^l{}_{;\,l}+\zeta)\sigma_k. \tag{2}$$

On writing

$$\gamma=[2\sigma^p{}_{;\,q}\sigma_{(p;q)}-(\sigma^r{}_{;\,r}+\zeta+\bar{\zeta})^2-(\zeta+\bar{\zeta})^2]^{\frac{1}{2}},$$

one finds that the ratios $\sigma_k:\zeta:\gamma$ are invariant under any transformation of the form $\sigma_k\rightarrow\lambda\sigma_k\neq 0$. As Dr. R. K. Sachs has in effect shown, the rates of rotation, dilatation, and maximum shear are proportional to $i(\zeta-\bar{\zeta})$, $\zeta+\bar{\zeta}$, and γ, respectively.

For any nonvanishing geodetic σ_k, if F_{kl} is a real bivector field subject to (1), so is its propagation derivative $\dot{F}_{kl}$; and there exists a scalar z such that $\dot{F}_{kl}-i^*\dot{F}_{kl}=z(F_{kl}-i^*F_{kl})$. Under a change of amplitude and polarization, $z\rightarrow z+\dot{w}$. We may therefore *normalize* the propagation of F_{kl} by demanding that

$$(F_{kl}-i^*F_{kl})_{;\,m}\sigma^m=\zeta(F_{kl}-i^*F_{kl}). \tag{3}$$

The normally propagated field is determined up to a change of amplitude and polarization for which $\dot{w}=0$.

Without assuming that σ_k is geodetic, writing $J^m=F^{mn}{}_{;\,n}$, $^*J^m=\,^*F^{mn}{}_{;\,n}$, one obtains

$$\dot{F}_{kl}+F_{kl}\sigma^p{}_{;\,p}+F^p{}_k\sigma_{l;\,p}+F_l{}^p\sigma_{k;\,p}=2J_{[k}\sigma_{l]},$$

from the covariant divergence of the first of Eqs. (1), and

$$\dot{F}_{kl}+F_k{}^p\sigma_{p;\,l}+F^p{}_l\sigma_{p;\,k}=2^*(\sigma_{[k}{}^*J_{l]}),$$

from the second of Eqs. (1) together with $3F_{[kl;\,m]}\sigma^m=2^*(\sigma_{[k}{}^*J_{l]})$. Thus, the equations

$$\tfrac{1}{2}F_{kl}\sigma^p{}_{;\,p}=F_k{}^p\sigma_{(p;\,l)}+\sigma_{(k;\,p)}F^p{}_l, \tag{4}$$

$$F_{kl}+\tfrac{1}{2}F_{kl}\sigma^p{}_{;\,p}+F_k{}^p\sigma_{[p;\,l]}-\sigma_{[k;\,p]}F^p{}_l=0, \tag{5}$$

are equivalent to $J_{[k}\sigma_{l]}=0=\,^*J_{[k}\sigma_{l]}$, which is a necessary and sufficient condition for the existence of a scalar a such that

$$J_k-i^*J_k=a\sigma_k. \tag{6}$$

On transvecting (4) or (5) with σ^k and $\epsilon^{klmn}\sigma_n$, one obtains $F^{mn}\dot{\sigma}_n=0=\,^*F^{mn}\dot{\sigma}_n$, which is necessary and sufficient for σ_n to be geodetic wherever F_{kl} is nonzero. It follows from the algebraic conditions on F_{kl} that there exist a scalar b and vectors ρ_k, τ_k, subject to

$$F_{kl}-i^*F_{kl}=2b\sigma_{[k}\tau_{l]},\quad g_{kl}=2\rho_{(k}\sigma_{l)}-2\tau_{(k}\bar{\tau}_{l)}. \tag{7}$$

Hence, one finds that if σ_k is geodetic, then (4) reduces to the nonshearing condition,

$$2\sigma^p{}_{;\,q}\sigma_{(p;q)}=(\sigma^r{}_{;\,r}+\zeta+\bar{\zeta})^2+(\zeta+\bar{\zeta})^2, \tag{8}$$

wherever F_{kl} is nonzero, and (5) is a necessary and sufficient condition for F_{kl} to propagate normally. Thus, the propagation vector σ_k is geodetic and shear-free in any null solution of Maxwell's empty-space equations $J_k=0=\,^*J_k$.

Suppose, conversely, that a real nonvanishing null vector field σ_k satisfies the geodetic and nonshearing conditions (2) and (8). Let F_{kl} be any real bivector field subject to (1) and (3). From (2), (6), and the conservation identities $J^k{}_{;\,k}=0=\,^*J^k{}_{;\,k}$, it follows that

$$\dot{J}_k=(\zeta+\bar{\zeta})J_k,\quad {}^*\dot{J}_k=(\zeta+\bar{\zeta})\,{}^*J_k. \tag{9}$$

By taking coordinates (x^λ, x^3, x^4) such that $\sigma^k=\delta_4{}^k$ and $\tau^3\neq 0$, in some region, we see from (9) that if Maxwell's equations are satisfied for one value of x^4, then they are satisfied for all x^4 in the region considered. The conditions on F_{kl} are invariant under any change of amplitude and polarization for which w is independent of x^4. From (6) and (7), it follows that the transformed field satisfies Maxwell's equations wherever

$$w_{,3}=-(a/b+w_{,\lambda}\tau^\lambda)/\tau^3;$$

and the last equation can be solved on a three-space of constant x^4. The requirement that σ_k should be geodetic and shear-free is therefore sufficient, as well as necessary, to ensure that σ_k is the propagation vector of a null solution to Maxwell's equations.

ACKNOWLEDGMENTS

The author is obliged to Dr. F. A. E. Pirani, Dr. J. Ehlers, and Dr. W. Kundt for their valuable comments.

Free Propagator Expansion in the Evaluation of the Lamb Shift. I*†

ARTHUR J. LAYZER‡
Columbia University, New York, New York

This is the first of two papers concerned with "order properties," with respect to the parameter αZ, of an expansion method for the evaluation of the bound electron self-energy ΔE and the application of these properties to the calculation of the new Lamb shift orders of $\alpha(\alpha Z)^6 \ln^2(\alpha Z)$ and $\alpha(\alpha Z)^6 \ln(\alpha Z)$. The expansion method is the free-propagator expansion (FPE); that is, the formal algebraic expansion of the bound electron propagator or Green's function in "powers" of the external (Coulomb) potential. The principal result of the general mathematical analysis is a theorem which asserts that the FPE is an order expansion for (only) those terms of ΔE that are nonanalytic in the parameter $w \equiv (\alpha Z)^2$ and is thus particularly suitable for the calculation of this class of terms. A practical result of the theorem is that the new logarithmic orders arise from only the first four terms of the FPE. The nonanalytic part of a fixed term I_n of the FPE can be attacked directly through a consideration of $\mathrm{Im}_+ I_n$, where $\mathrm{Im}_+ I_n$ denotes the imaginary part of I_n, regarded as a function of the complex variable w, on the upper side of a branch cut along the negative w axis. As an auxiliary result, boundedness properties in momentum space are derived for certain iterated operators related to the FPE of the bound nonrelativistic electron Green's function.

INTRODUCTION

IN this and a following paper we shall present the results of a mathematical investigation which had as its "practical" aim the calculation of the new Lamb shift orders of $\alpha(\alpha Z)^6 \ln^2(\alpha Z)$ and $\alpha(\alpha Z)^6 \ln(\alpha Z)$.[1] The mathematical results are, however, more general and, in addition to clarifying the mathematical analysis of the Lamb shift, should be of use in the analysis of the wider class of bound-state problems involving the bound electron Green's function as a component. Our plan is to give the more general results in the present paper and those pertaining particularly to the calculation of the new logarithmic orders in the following one.

Technically speaking, the new orders fall into the category of second-order radiative corrections to atomic energy levels to arbitrary orders in the Coulomb interaction parameter (αZ). Thus, they correspond to the bound interaction diagrams shown in Fig. 1 called respectively, "self-energy" and "vacuum polarization" diagrams.[2] The self-energy diagram represents the lowest order radiative correction to the expectation value of the "mass operator."[3]

We shall consider in detail only diagram (a). The contribution of the polarization diagram to the order of interest is easily evaluated and will be included in the final result.

The analytic expression for the shift ΔE corresponding to the self-energy diagram is of exactly the same form as the free electron self-energy except that the wave functions are bound-state (Coulomb) wave functions and the free electron propagator or Green's function $(\mathbf{p}-\mathbf{k}+m)^{-1}$ is replaced by the bound propagator $(\mathbf{p}-\mathbf{k}-\mathbf{V}+m)^{-1}$. Here m is the mass of the electron; $\mathbf{a} \equiv i\gamma_\mu a_\mu$; V_μ is the four potential, in this case a static Coulomb potential; p_μ and k_μ are the electron and photon four momentum, respectively, with the fourth component of p_μ set equal to iE, where E is the bound-state energy.

The method we shall use to evaluate ΔE will be the "free-propagator expansion"; that is, the formal algebraic expansion of the bound electron propagator in "powers" of $\mathbf{V}$. We shall be concerned with the mathematical study of some of the main properties of this expansion connected with an expansion of ΔE in orders of αZ.

The free-propagator expansion, although at first sight the simplest and most natural technique for the evaluation of ΔE, has the disadvantage that it is not, in spite of appearances, an expansion in orders of (αZ). In fact, though the lowest order of the Lamb shift is $\alpha(\alpha Z)^4 \ln(\alpha Z)$, corresponding to the nonrelativistic calculation of Bethe,[4] the individual terms of the expansion are of still lower order.[5] For this reason, the expansion was neglected as a method for the evaluation of the Lamb shift until the recent work of Fried and Yennie,[6] hereafter referred to as F-Y, in which this expansion was used to extract successfully the two lowest orders of $\alpha(\alpha Z)^4 \ln(\alpha Z)$ and $\alpha(\alpha Z)^4$. These authors employed a special "infrared" gauge for the photon propagator[7] in which the lowest-order contributions from each term of the expansion were canceled while at the same time infrared divergences associated with the renormalization of the first two terms of the expansion were also eliminated. In the present paper we shall employ the usual and simpler Feynman form

* Based on a dissertation submitted to Columbia University if partial fulfillment of the requirements for the degree of Doctor on Philosophy.

† This research was supported in part by the U. S. Atomic Energy Commission.

‡ Present address: Institute of Mathematical Sciences, New York University, New York, New York.

[1] A. J. Layzer, Phys. Rev. Letters 4, 580 (1960). The $\ln^2$ term was calculated independently by H. Fried and D. Yennie [*ibid.* 4, 583 (1960)], and by G. Erickson [dissertation, University of Minnesota, 1960 (unpublished)].

[2] Of course, both Figs. 1(a) and 1(b) contribute to the self-energy of the bound electron. The term "self-energy" for Fig. 1(a) is used because of its pictorial similarity to the diagram for the free electron self-energy.

[3] J. Schwinger, Proc. Natl. Acad. Sci. U. S. 37, 452 (1951).

[4] H. A. Bethe, Phys. Rev. 72, (1947).

[5] This "spurious" lowest order contribution may be eliminated by summing over all terms of the expansion (see Sec. 2).

[6] H. Fried and D. R. Yennie, Phys. Rev. 112, 1391 (1958).

[7] The nature of this type of gauge transformation was elucidated by B. Zumino, J. Math. Phys. 1, 1 (1960).

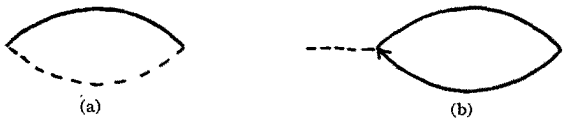

FIG. 1. "self-energy" and "polarization" diagrams.

of the photon propagator. The difficulty of the lowest order terms turns out to be not at all serious for our purposes.

The main result of the present mathematical investigation is a theorem[8] which asserts that the free-propagator expansion is an order expansion for those terms of ΔE that are nonanalytic in $w\equiv(\alpha Z)^2$ and is thus particularly suitable for the calculation of this class of terms.[9] A practical result of the precise form of the theorem is that the orders $\alpha w^3\ln^2 w$ and $\alpha w^3\ln w$ which we wish to calculate arise only for $n\leq 3$, while the previously computed orders of $\alpha w^2\ln w$[10] and $\alpha w^{\frac{5}{2}}$[11] arise only for $n\leq 2$. Here, n is the number of numerator V's, which serves as an index for the free-propagator expansion. The latter two assertions have been verified by explicit calculation[12] with results in agreement with the earlier calculations.[13] The proof of the theorem mentioned above depends on the properties of an auxiliary expansion for fixed n called the k expansion and is formal to the extent that the convergence of the k expansion is not rigorously established.

Although the theorem described above can be derived by real-variable methods, it is more natural to use a complex-variable approach, regarding $w\equiv(\alpha Z)^2$ as a complex variable. We show formally that the general term $I_n(w)$ of the free-propagator expansion can be continued analytically in a circular neighborhood of $w=0$ cut along the negative real axis, and that for the derivation of the theorem it is sufficient to consider $\mathrm{Im}_+\langle I_n\rangle$, where Im_+ denotes the imaginary part along the upper side of the cut. (The brackets denote a bound state expectation value.)

The treatment of mass and charge renormalization[14] becomes simpler when one considers $\mathrm{Im}_+\langle I_n\rangle$ rather than I_n itself. Since the mass renormalization is independent of w, $\mathrm{Im}_+\langle I_0\rangle$ does not have to be mass-renormalized. We show also that $\mathrm{Im}_+\langle I_0\rangle$ and $\mathrm{Im}_+\langle I_1\rangle$ can be charge renormalized without introducing infrared divergences into these terms.[15]

1. FREE-PROPAGATOR EXPANSION

With a convenient normalization,[16] the self-energy diagram part of the self-energy is given by[17]

$$\langle\Delta E\rangle=\langle\bar{v}\,|\,\Delta E\,|\,v\rangle \tag{1}$$

$$\Delta E=i\int_F\frac{d^4k}{k^2}\gamma_\mu\frac{1}{p-k-v+m}\gamma_\mu. \tag{2}$$

If the bound propagator is expanded in powers of V_0, we obtain formally

$$\Delta E=\sum_{n=0}^{\infty}I_n \tag{3}$$

$$I_0=i\int_F\frac{d^4k}{k^2}\gamma_\mu\frac{1}{p-k+m}\gamma_\mu$$

$$I_1=i\int_F\frac{d^4k}{k^2}\gamma_\mu\frac{1}{p-k+m}V\frac{1}{p-k+m}\gamma_\mu \tag{4}$$

$$I_2=i\int_F\frac{d^4k}{k^2}\gamma_\mu\frac{1}{p-k+m}V\frac{1}{p-k+m}V\frac{1}{p-k+m}\gamma_\mu,$$

and so forth.

The well-known expansion (3) will be called the free-propagator expansion. Of course, the expansion could be terminated at an arbitrary **n** by replacing the last free propagator by a bound propagator.

On "rationalizing" the Dirac algebra for the free

[8] A. J. Layzer, Bull. Am. Phys. Soc. 4, 280 (1959).

[9] Throughout this paper, functions analytic in w in a circular neighborhood including $w=0$ will be referred to simply as "analytic in w."

[10] Covariant calculations of the $\alpha w^2\ln w$ (and αw^2) terms were performed by N. M. Kroll and W. E. Lamb, Phys. Rev. **75**, 388 (1949); J. B. French and V. F. Weisskopf, *ibid.* **75**, 1240 (1949); R. P. Feynman, *ibid.* **74**, 1430 (1949); H. Fukuda, Y. Miyamoto, and F. Tomonaga, Progr. Theoret. Phys. (Kyoto) **4**, 47, 121 (1949).

[11] The $\alpha w^{5/2}$ term was calculated by M. Baranger, H. A. Bethe, and R. P. Feynman [Phys. Rev. **92**, 482 (1953)], and by R. Karplus, A. Klein, and J. Schwinger [*ibid.* **86**, 288 (1952)].

[12] A detailed calculation of the $\alpha w^{5/2}$ term using the free-propagator expansion can be found in the author's dissertation. This term was calculated also by Fried and Yennie using their special gauge (reference in footnote 1).

[13] The fact that the $\alpha w^{5/2}$ result is given by the three lowest terms of the free-propagator expansion could have been anticipated from a corresponding result of N. Kroll and F. Pollock in the similar hyperfine problem [Phys. Rev. **86**, 876 (1952), footnote 27]. It is not hard to see from the form of the magnetic potential that the order $\alpha(\alpha Z)E_F$ calculated by these authors for the hyperfine case corresponds mathematically to the Lamb shift order $\alpha w^{5/2}$. Here E_F is the Fermi energy.

[14] Somewhat inaccurately, we shall use the term "charge renormalization" for renormalizations other than mass renormalization. It is well known that renormalization in electrodynamics can be effected in a covariant way by introducing counter terms into the Lagrangian which can be ascribed to changes in the scale of mass, charge and the field quantities. In this way, the separate renormalizations of I_0 and I_1 can be correctly identified in the present noncovariant representation. From a computational point of view, the subtraction procedure for the renormalization of I_0 and I_1 is exactly analagous to the usual scattering prescriptions for the self-energy and vertex diagrams, the only difference being the occurrence of bound-state wave functions.

[15] The renormalization introduced here is a purely formal operation based on the usual Taylor's expansions *following* the analytic continuation.

[16] A factor of $-\alpha/(4\pi)^3$ is omitted.

[17] Here we use the more accurate notation $\langle\Delta E\rangle$ for the quantity denoted by ΔE in the introduction. We shall distinguish between the operators ΔE and I_n and their bound-state expectation values only when this is necessary to avoid confusion.

propagators, we obtain

$$I_n=i\int_F \frac{d^4k}{k^2}\gamma_\mu \frac{k-p+m}{k^2-2pk+\Delta} \overset{(1)}{V} \frac{k-p+m}{k^2-2pk+\Delta} \overset{(2)}{V} \cdots \frac{k-p+m}{k^2-2pk+\Delta} \overset{(n)}{V} \frac{k-p+m}{k^2-2pk+\Delta}\gamma_\mu, \quad (5)$$

where

$$\Delta \equiv p^2+m^2=\mathbf{p}^2+\epsilon^2 \quad (6)$$

$$\epsilon^2 \equiv m^2-E^2>0. \quad (7)$$

For I_0 and I_1, the k integrations are infinite at high k. These infinities are removed by mass renormalization of I_0 and I_1.[14]

The sequence of operators I_n is formally the same as in the similar expansion of the S matrix in the "free-interaction picture" corresponding to lowest order radiative corrections to (multiple) scattering in an external field, except that in the scattering case $\epsilon^2 \leq 0$.[18] One can, therefore, combine denominators, shift the origin of the k integration, and "rotate" the path of the k_0 integration to the imaginary axis according to the usual scattering prescriptions of Feynman.

For this purpose, we introduce a momentum representation, labeling the momentum coordinates from left to right by p_0, p_1, $p_2 \cdots p_n$ for I_n. Here p_0 is associated with the left wave function and p_n with the right wave function. Then according to the results and notation of F-Y, one obtains

$$\left[\prod_{i=0}^{n} k^2-2p_ik+\Delta_i\right]^{-1} = \int dY_{n-1}[k^2-2\bar{p}_nk+\bar{\Delta}_n]^{-(n+1)} \quad (8)$$

$$[k^2\prod_{i=0}^{n} k^2-2p_ik+\Delta_i]^{-1}=(n+1)\int dY_{n-1} \times \int_0^1 dx x^n[(k-x\bar{p}_n)^2+xD_n(x)]^{-(n+2)}, \quad (9)$$

where

$$\int dY_{n-1} \equiv n! \prod_{i=0}^{n-1}\int_0^1 dy_m y_m{}^m \quad (10)$$

and

$$D_n(x) \equiv \bar{\Delta}_n-x(\bar{p}_n)^2. \quad (11)$$

The "bar" operation for a given set of quantities $(a_1,a_2,\cdots,a_n)$ is defined by

$$\bar{a}_1=y_0a_0+(1-y_0)a_1$$
$$\bar{a}_n=y_{n-1}\bar{a}_{n-1}+(1-y_{n-1})a_n. \quad (12)$$

[18] This apparently minor difference of the sign of ϵ^2 is actually very important since it accounts for the fact that the bound-state matrix elements I_n are finite though the corresponding scattering matrix elements I_n are divergent in the infrared. [For the origin of the infrared divergence in the scattering case see, for example, J. M. Jauch and F. Rohrlich, *Theory of Photons and Electrons* (Addison-Wesley Publishing Company, Inc., Reading, Massachusetts, 1955), p. 394.] In effect, the binding energy of the bound state plays the role of a "photonmass" as far as the prevention of an infrared divergence is concerned.

Note that (8) is an example of the general relation

$$\int dY_{n-1}(\bar{a}_n)^{-(n+1)}=\prod_{i=0}^{n} a_i^{-1}. \quad (13)$$

From the definition (12), the bar operation is an average in the usual sense with respect to a y-dependent distribution; that is,

$$\bar{a}_n=\sum_{i=0}^{n-1} w_ia_n, \quad (14)$$

with

$$\Sigma w_i=1 \quad (15)$$
$$w_i>0 \quad i=0, 1, \cdots, (n-1).$$

One consequence of this fact is that the bar of a "constant" (n independent) quantity is the same constant. This enables us to write the expression (11) for $D_n(x)$ in three-dimensional notation as

$$D_n(x)=xm^2+\bar{\Delta}_n-x(\bar{p}_n)^2+\epsilon^2 \quad (16)$$

$$\bar{\Delta}=\bar{p}^2+\epsilon^2. \quad (17)$$

Another consequence is the inequality for three-dimensional operators p and $\bar{p}$,

$$\bar{p}^2-(\bar{p})^2\geq 0, \quad (18)$$

since the quantity on the left is a *dispersion*. Equation (18) implies, according to (16) and (7), that

$$D_n(x)>0. \quad (19)$$

To complete the standard scattering treatment of I_n, we perform the shift of the origin of the k integration,

$$k_\mu \Rightarrow k_\mu+x\bar{p}_\mu, \quad (20)$$

and then rotate the k_0 integration from a Feynman contour along the real axis to the imaginary axis. Note that I_n is invariant to the shift (20) for $n\geq 1$. For I_0, as is well known, an extra term must be added after the shift. In the remainder of this section and in the following two sections, it is understood that exceptions to the formulas and statements for general n may arise as a result of "ultraviolet" divergences in the k integration for $n=0$ and 1. The explicit treatment of I_0 and I_1 will be given in Sec. 4.

The expression for $\langle p_0|I_n|p_n\rangle$ is now

$$\langle p_0|I_n|p_n\rangle=(-)(n+1)\int d^4k\int_0^1 dx x^n\int dY_{n-1} \times d^3p_0d^3p_1\cdots d^3p_{n-1}\frac{\mathfrak{N}_n(k,p,\bar{p})}{[k^2+x\bar{p}_n(x)]^{n+2}} \quad (21)$$

$$\mathfrak{N}_n(k,p,\bar{p})=\gamma_\mu N(p_0)\prod_{i=1}^{n}\langle p_{i-1}|V|p_i\rangle N(p_i)\gamma_\mu$$
$$N(p_i)\equiv k+x\bar{p}_n-p_i+m. \quad (22)$$

In three-dimensional notation for the p's,

$$N(p_i)=\mathbf{k}-i\gamma\cdot(p_i-x\bar{p}_n)+2\Lambda_+m-x\gamma_4m - (m-E)\gamma_4(1-x), \quad (23)$$

where

$$\Lambda_\pm=\tfrac{1}{2}(1\pm\gamma_4). \quad (24)$$

According to (19) the denominator in the expression (21) for $\langle p_0|I_n|p_n\rangle$ is positive definite.

Let us note that if $f(x)$ is an operator function of the position vector $\mathbf{x}$, then we have the relation

$$F(q)=\langle p|f(x)|p'\rangle, \quad (25)$$

where

$$q\equiv p-p' \quad (26)$$

and $F(q)$ is the appropriately normalized Fourier transform of the "c number" $f(x)$:

$$F(q)=(2\pi)^{-3}\int e^{-iq\cdot x}f(x)d^3x. \quad (27)$$

We shall be interested in the particular case that V_μ has only a fourth component $V_4=iV$ and furthermore V is the Coulomb potential[19]:

$$V=-\alpha z/r;\quad V(q)=-(\alpha z/2\pi^2)/(1/q^2). \quad (28)$$

On returning now to formula (21) and using an obvious symbolic notation (the symbolic nature of the notation is indicated by quotation marks), we can write the expectation value of I_n (which we shall also denote by I_n) in the abbreviated form

$$I_n=-(n+1)\int d^4k\int dx x^n\int dY_{n-1} \quad \times\frac{\text{“}\langle\bar{v}|\gamma_\mu\mathbf{V}^n\prod_{i=0}^{n}N(p_i)\gamma_\mu|v\rangle\text{”}}{[k^2+xD_n(x)]^{n+2}}. \quad (29)$$

Instead of the parameter x, it often proves convenient in later work to employ the parameter $\sigma=1/x$. With this change of variables, the expression (29) for I_n reads

$$I_n=-(n+1)\int d^4k\int_1^\infty d\sigma\sigma^{n+2}\int dY_{n-1} \quad \times\frac{\text{“}\langle\bar{v}|\gamma_\mu V^n\prod_{i=0}^{n}N(p_i)\gamma_\mu|v\rangle\text{”}}{\{\sigma^2k^2+m^2+\sigma\bar{\Delta}-[(\bar{p})^2+\epsilon^2]\}^{n+2}}. \quad (30)$$

In the following, the existence of I_n in the form (29) or (30) will be *assumed*.

[19] It is worth noting that most of the analysis of this and the following sections is valid also when the Coulomb potential is generalized to a potential which, in momentum space, differs from the Coulomb potential by a factor which is a bounded function of $p/\alpha Zm$. This is the case, for example, for a screened Coulomb potential with a range of the order of the Bohr radius.

We wish now to estimate the order in $w\equiv(\alpha Z)^2$ of the various terms of I_n and to extract and evaluate certain lower orders in w.

This task will be simplified greatly by the introduction in the next section of the auxiliary "k expansion," which is a binomial expansion of the inverse denominator of I_n in powers of $(\bar{p})^2+\epsilon^2$. The existence of the individual terms of this expansion and their lowest orders will be established with a fair amount of rigor. It will be shown in particular that the "zeroth" term of the k expansion determines the lowest order of I_n and this is of lowest order w for each $n\geq 2$. This is consistent with the fact, noted by Fried and Yennie, that even after combination of I_0 and I_1 a "spurious" contribution of order w remains. Thus, as is well known, the free-propagator expansion, in spite of the appearance of the parameter (αZ) in the Coulomb potential, is not an expansion in order of (αZ).

To complete the present section, we shall derive some useful properties of I_n that are independent of the k expansion. In particular we shall employ a type of "scaling" that will exhibit "externally" or "nominally" the correct lowest order of w for I_n for $n\geq 2$ and we shall show how the nominal order of a particular term of I_n is related to the numerator type, that is the number of numerator k's, p's, x's, etc. This will also provide an opportunity for the introduction of a rather large amount of special notation that will be used in all of the following sections.

Definition and Labeling of "Elementary Terms"

Corresponding to the seven terms of the expression (23) for $N(p_i)$, one can express I_n as a linear combination of 7^{n+1} "elementary terms." We shall (partially) label an elementary term by the seven indexes q, r, s, t, u, v, w denoting, respectively, the number of numerator factors $\mathbf{k}$, $\mathbf{p}$, $x\bar{\mathbf{p}}$, Λ_+m, $x\gamma_4m$, $(m-E)\gamma_4$, $(m-E)\gamma_4x$. Since these indexes are connected by the obvious relation

$$q+r+s+t+u+v+w=n+1, \quad (31)$$

one of them can be omitted and we choose this to be the index $\mathbf{t}$ corresponding to Λ_+m. Furthermore, terms with nonzero v and w, which have "external" factors of $(m-E)$ are fairly trivial as far as the general analysis is concerned and the indexes v and w will therefore not be explicitly indicated. Accordingly, we will represent an elementary term of I_n by the notation

$$I_{n;qrsu\cdots}$$

where the dots represent the indexes v and w.

Division into $++$, $+-$, $-+$, and $--$ Components

The operator I_n or one of its elementary terms T_n is a 16-component matrix in "spinor space." It is often convenient to consider separately the four constituent 2×2 matrices which connect the large and small

component wave functions. This one can do by putting the projection operators $\Lambda_\pm$ defined in (24) to the left and right of I_n or T_n. Thus, we have

$$I_n=\Lambda_+I_n\Lambda_++\Lambda_+I_n\Lambda_-+\Lambda_-I_n\Lambda_++\Lambda_-I_n\Lambda_-. \quad (32)$$

In words, the first term of (32) is the "large-large" part of I_n, the second term is the "large-small" part of I_n and so on. To indicate this division for an elementary term $I_{n;qrsu\cdots}$, we add the subscript $++$, $+-$, $-+$, or $--$. Thus,

$$I_{n;qrsu\cdots;++}$$

denotes the large-large part of the elementary term $I_{n;qrsu\cdots}$.

Selection Rules on the Indexes q, r, s

There are two important restrictions on the values of the indexes q, r, s for nonvanishing elementary terms. These restrictions arise from symmetry properties of the operators I_n and will therefore be called "selection rules."

The first restriction follows from the symmetrical nature of the k integration.

Selection rule 1. For nonvanishing $I_{n;qrsu\cdots}$, $q=$even. In fact, in addition to selection rule 1, there is the stronger restriction that the number of "space" and "time" parts of the $\mathbf{k}$'s are separately even.

For the second selection rule, we assume that the potential V has only a fourth component:

$$\mathbf{V}=-\gamma_4 V. \quad (33)$$

Then, we have:

Selection rule 2. For nonvanishing $I_{n;qrs\cdots}$,

$$r+s=\text{even for } ++,\ --$$
$$r+s=\text{odd for } +-\text{ or } ++.$$

There are two ways of proving this restriction. The first method is independent of the wave functions and the momentum space integration but depends on the structure of the Dirac algebra and the restrictions imposed on the k integration. One uses the following properties of the projection operators Λ_+, Λ_-:

$$\Lambda_+\Lambda_-=\Lambda_-\Lambda_+=0$$
$$\Lambda_\pm\theta=\theta\Lambda_\mp;\quad \Lambda_\pm\epsilon=\epsilon\Lambda_\pm, \quad (34)$$

where θ and ϵ are "odd" and "even" Dirac matrices, respectively.[20] The result now follows immediately from the observation that $r+s$ is the number of odd matrices γ enclosed by the projection operators aside from "paired" odd matrices due to the photon k_μ and to the end factors of γ_μ.

In the second method of proof, the expectation value of I_n is considered. One inverts all momenta p_0, p_1, $\cdots p_n$ and takes note of the fact that the potential V and the denominator of I_n are invariant to spatial inversion, while the large and small component wave functions have definite but opposite spatial parity.

[20] L. L. Foldy and S. A. Wouthysen, Phys. Rev. **78**, 29 (1950).

"Scaling" Transformations

We turn now to the objective stated earlier of constructing an appropriate scaling transformation that will exhibit nominally the true order of w for I_n for the case that V is the Coulomb potential. For this purpose, we first scale coordinates and momenta according to

$$x \Rightarrow (\alpha zm)^{-1}x$$
$$p \Rightarrow (\alpha zm)p; \quad (35)$$

that is, we transform to "Z" atomic units

$$(\hbar=c=\alpha Zm=1).$$

The transformation (35) is to be thought of as a canonical transformation to new variables x and p. The wavekets $|v\rangle$ will undergo a corresponding transformation. If the new variables x and p defined by (35) are temporarily called x' and p' and the new ket is called $|v_s\rangle$, then it is easy to show, by requiring the invariance of the norm of $|v\rangle$, that we must have

$$\langle x'|v_s\rangle=(\alpha zm)^{-\frac{3}{2}}\langle x=(\alpha zm)^{-1}x'|v\rangle$$
$$\langle p'|v_s\rangle=(\alpha zm)^{\frac{3}{2}}\langle p=(\alpha zm)p'|v\rangle. \quad (36)$$

After performing this canonical scaling transformation there is an external factor of w^n due to V^n, and the denominator of I_n has the form

$$\sigma^2k^2+m^2+m^2(w\sigma\bar{\Delta}_s-w[(\bar{p})^2+\epsilon_s^2]), \quad (37)$$

where

$$\bar{\Delta}_s\equiv\bar{p}^2+\epsilon_s^2;\quad \epsilon_s^2\equiv(m^2-E^2)/m^2w.$$

In the following, we shall drop the subscripts s for the scaled quantities Δ and ϵ^2, since it will always be clear whether the scaled or unscaled quantities are meant.

The form of the denominator (37) suggests the two additional transformations

$$k \Rightarrow wkm$$
$$\sigma \Rightarrow w^{-1}\sigma, \quad (38)$$

which turn out indeed to yield the desired nominal order of w for I_n as a whole.

It is understood that the transformations (38) are to be applied only when the k integration is convergent. In the following we assume, in addition, that it is permissible to interchange the k and σ integrations.

On performing the transformations (35) and (38) on I_n in the form (30), we obtain, after inverting the order of the k and σ integrations,

$$I_n=-(n+1)mw\int_w^\infty d\sigma\sigma^{n+2}\int d^4k\int dY_{n+1}$$
$$\times\frac{\text{“}\langle\bar{v}_s|\gamma_\mu V_0{}^n\prod_{i=0}^{n}N_s(w^{\frac{1}{2}}p_i)\gamma_\mu|v_s\rangle\text{”}}{\{k^2\sigma^2+1+\sigma\bar{\Delta}-w[(\bar{p})^2+\epsilon^2]\}^{n+2}}; \quad (39)$$

$$V_0=-1/r; \tag{40}$$

$$N_s(w^{\frac{1}{2}}p_i)=w\mathbf{k}-i\gamma\cdot w^{\frac{1}{2}}\left(p-\frac{w}{\sigma}\bar{p}\right)+2\Lambda_+ -\frac{w}{\sigma}\gamma_4-\left(\frac{m-E}{m}\right)\gamma_4\left(1-\frac{w}{\sigma}\right). \tag{41}$$

Note that the dependence on m has been completely removed from the "interior" of I_n (if E is expressed in units of m) and appears only as a linear external factor.

The external factor of w was arrived at by simply counting n, $-(n+3)$, and $+4$ powers of w due to V^n, $d\sigma\sigma^{n+2}$, and d^4k, respectively.[21]

Approximations to the Large and Small Scaled Wave Functions[22]

Let the symbols ψ and χ denote, respectively, the large and small two-component spinors making up the exact wave function v. Thus,

$$v=\begin{pmatrix}\psi\\0\end{pmatrix}+\begin{pmatrix}0\\\chi\end{pmatrix}. \tag{42}$$

The subscript s for ψ and χ will denote, as usual, the corresponding scaled quantities. In the limit $w\Rightarrow 0$, ψ_s approaches the Schrödinger wave function u_s, which is independent of w.

The corresponding approximation for the small component is well known. One can write the Dirac equation in the form

$$\Lambda_-|v\rangle=\begin{pmatrix}0\\\chi\end{pmatrix}=\frac{1}{2m}[i\gamma_4\gamma\cdot p+V+(m-E)]|v\rangle. \tag{43}$$

By scaling both sides of this equation, we obtain formally in the limit $w\Rightarrow 0$

$$\begin{pmatrix}0\\\chi_s\end{pmatrix}\Rightarrow w^{\frac{1}{2}}\gamma_4 i\gamma\cdot p\begin{pmatrix}u_s\\0\end{pmatrix}=w^{\frac{1}{2}}\alpha\cdot p\begin{pmatrix}u_s\\0\end{pmatrix}. \tag{44}$$

This approximation for the small-component wave function is a scaled form of the usual "Pauli approximation." Here we shall use the phrase "generalized Schrödinger approximation" to denote the approximation to a matrix element in which $\psi_s\Rightarrow u_s$ and $\chi_s\Rightarrow w^{\frac{1}{2}}\alpha\cdot pu_s$ for the large and small component wave functions, respectively.

An exact expression for the small-component wave function is obtained by introducing the wave function ϕ according to

$$\begin{pmatrix}0\\\chi\end{pmatrix}=\frac{\alpha\cdot p}{2m}\begin{pmatrix}\phi\\0\end{pmatrix}. \tag{45}$$

In the limit $w\Rightarrow 0$; $\phi_s\Rightarrow u_s$. For S states, ϕ like ψ is spherically symmetric and has only one component for definite spin direction of u_s.

For S states, $\psi_s(p)$ and $\phi_s(p)$ have the general form

$$\begin{aligned}\psi_s(p)&=u_s(p)[f_1(p,w)+wpf_2(p,w)]\\ \phi_s(p)&=u_s(p)[f_2(p,w)],\end{aligned} \tag{46}$$

where the functions $f(p,w)$ are analytic functions of w and continuous functions of p. Furthermore, the functions $f(p,w)$ are finite at $p=0$ and behave at high p so that fu remains integrable over d^3p.

The representation (45) makes it possible to treat the large-small and small-small parts of I_n in the same formal manner as the large-large part. For this purpose, we use (34) to write

$$\begin{aligned}\Lambda_-\alpha\cdot p_n&=\alpha\cdot p_n\Lambda_+\\ \alpha\cdot p_0\Lambda_-&=\Lambda_+\alpha\cdot p_0,\end{aligned} \tag{47}$$

where p_0 and p_n are the momentum coordinates of the wave functions. After doing this, the appearance of the $+-$ and $--$ cases is like the $++$ case, except that effectively the index r is raised by 1 or 2 for $+-$ and $--$, respectively.

In the following calculations it is sometimes necessary to go beyond the generalized Schrödinger approximation. Since the wave functions ψ_s and ϕ_s are analytic functions of w, they can be expanded in powers of w. Only the first correction to the Schrödinger wave functions will be needed in the calculations. We introduce the notation u_s' and ϕ_s' for these "Dirac corrections" to u_s and ϕ_s, respectively. Thus,

$$\psi_s\simeq u_s+wu_s' \quad\text{and}\quad \phi_s\simeq u_s+w\phi_s'. \tag{48}$$

Form and Nominal Order of Elementary Term of I_n

We conclude this section by writing an expression for the nominal order of an elementary term labeled by q, D, S, and $u(v=w=0)$ in the generalized Schrödinger approximation. Here D is the degree of the numerator polynomial in the p's and $\bar{p}$:

$$\begin{aligned}D&=r+s &&++\\ &=r+s+1 &&+- \quad\text{and}\quad -+\\ &=r+s+2 &&--.\end{aligned} \tag{49}$$

The necessary modifications for Dirac correction terms and nonvanishing v and w will be fairly obvious.

[21] Note that if the photon had a finite mass, the transformations (35) and (38) would not have been useful and the free-propagator expanison would be an expansion in orders of αZ. Conversely, we can say that the failure of the free propagator expansion to be an expansion in orders of αZ is associated with the infrared divergences of the corresponding scattering matrix elements. (See also footnote 18.)

[22] For the basic properties and explicit forms of the relativistic Coulomb wave functions in position space see H. A. Bethe and E. E. Salpeter, *Quantum Mechanics of One and Two Electron Systems* (Academic Press Inc., New York, 1957), Chap. I. The momentum space wave functions may be obtained from references listed in that book or by taking the Fourier transforms with the aid of tables given by A. Erdelyi, *Tables of Integral Transforms*, Vol. 1., Bateman Manuscript Project (McGraw-Hill Book Company, Inc., New York, 1954).

From (39) and (41) one sees that the nominal order of our elementary term is w^ν with

$$\nu=D/2+q+s+u+1. \tag{50}$$

According to selection rule 2, D is an even number and therefore ν is always integral.

We shall show in the next section that according to the k expansion the nominal order ν is the true order up to an order of about $w^{(n+3)/2}$.

2. k-EXPANSION

The k expansion mentioned briefly in the last section is obtained by expanding I_n in powers of $(\bar{p})^2+\epsilon^2$ $\times[(\bar{p}^2+m^2)$ in four-dimensional notation$]$ according to the binomial expansion

$$\frac{1}{(a-b)^p}=\sum_{k=0}^{\infty}\binom{p+k-1}{k}\frac{b^k}{a^{m+k}} \tag{51}$$

with $p=n+2$, $\alpha=\sigma^2k^2+m^2+\sigma\bar{\Delta}$ and $b=(\bar{p})^2+\epsilon^2$.

Correspondingly,

$$I_n=\sum_{k=0}^{\infty}I_n{}^k, \tag{52}$$

where, after scaling, from (39)

$$I_n{}^k=(-)\binom{n+k-1}{k}(n+1)mw^{k+1}$$
$$\times\int_w^\infty d\sigma\sigma^{n+2}\int d^4k\int dY_{n-1}$$
$$\times\frac{\text{“}\langle\bar{v}_s|\gamma_\mu\mathbf{V}_0{}^n\prod_{i=0}^{n}N_s(w^{\frac{1}{2}}p_i)[(\bar{p})^2+\epsilon^2]^k\gamma_\mu|v_s\rangle\text{”}}{(k^2\sigma^2+1+\sigma\bar{\Delta})^{n+k+2}}. \tag{53}$$

From the inequalities (18) and (7) the expansion (51) is always convergent, but this is not necessarily true of the corresponding integrated expansion (52). We shall be able to prove the existence of each $I_n{}^k$, but the convergence of the k expansion of I_n will be *assumed*. The trouble here arises at the "point" $\sigma=1$ (before scaling) and some or all of the p's infinite. The assumption of convergence of the k expansion is clearly closely related to the previous assumption of the existence of I_n. For certain elementary terms these assumptions are in fact equivalent. This is the case, for example, when $D=0$ [see (49)]. For, using a momentum representation and dominating the Schrödinger wave functions by the $1-S$ wave functions, one sees from (28) that in this case the integrand is always of the same sign.

The k expansion has a number of useful properties which will become apparent in the following work. Here we note only that the k expansion simplifies the appearance of the parameter w. Except for the analytic, and therefore formally expandable, dependence of the scaled wave functions on w and the rather trivial and expandable dependence on w of the scaled ϵ^2 of the denominator $\bar{\Delta}$, w appears in the "interior" of $I_n{}^k$ only in the lower limit of the final s integration. One consequence of this is that whenever the lower limit can be replaced by 0 without introducing a divergence in the final s integration, the nominal order ν of a term of $I_n{}^k$ will be the true order. We shall show in this section that this is the case for ν less than $(n+3)/2$ (for $n\geqslant 3$). Thus, the "vertical" k expansion, unlike the "horizontal" free propagator expansion is initially an expansion in orders of w.

Formal Sum of Lowest Order Terms

If we assume the validity of the results mentioned in the last paragraph, then it follows that the "spurious" order w contributions for $n\geqslant 2$ arise only for the $k=0$ term $I_n{}^0$ and from the elementary term with $q=r=s=u=v=w=0$, that is, with $N(P_i)$ in (23) replaced by $2\Lambda_+m$.

It is now easy to show that these $k=0$ terms can be summed formally from $n=2$ to ∞ by making use of the identity

$$\frac{1}{A+B}=\frac{1}{A}+\frac{1}{A}(-B)\frac{1}{A}+\frac{1}{A}(-B)\frac{1}{A}(-B)\frac{1}{A}+\cdots, \tag{54}$$

yielding a result which in lowest order indeed cancels the known contribution of order w from I_0+I_1 as previously computed by F-Y. The formal sum results in the appearance of the bound nonrelativistic electron's Green's function acting on the Schrödinger wave function, which enables the calculation to be completed in an elementary way.

To carry through the algebraic summation one employs identity (13) to eliminate the y integrations and return to a "separable" or product representation. Since the denominator in (53), for $k=0$, is raised to the power $n+2$ rather than the power $n+1$ required by (13), one first performs a partial differentiation with respect to k^2. This also removes the numerical factor of $(n+1)$ in (53) and one finds now that identity (54) can be employed immediately with

$$A=\sigma^2k^2+1+\sigma\Delta;\quad B=2\sigma V_0.$$

It should be emphasized that for the calculation of the type of terms of interest (nonanalytic in w) the above formal sum will not be needed and will not be employed. For these terms the order w contributions (as well as those of order w^2) can be entirely "by-passed" as will be made clear in Sec. 3.

Boundedness Properties in Momentum Space of Certain Iterated Operators

We now turn to the question of the existence and lowest order of the $I_n{}^k$ and the corresponding elementary terms $I_{n;qrs\ldots}$. By using dominating arguments in

p space we shall show in the following subsection that the problem can be reduced essentially to the boundedness properties in a momentum representation of certain iterated operators characteristic of $k=0$ terms. The fact that these operators are iterated is essential, since it permits one to employ inductive methods of proof. For the sake of mathematical clarity, the definition and the derivation of the required properties of these operators will be given first, in the present subsection.

We introduce first the symmetric operators $\bar{K}_n$ defined by[23]

$$\bar{K}_n(p,p')=(-1)^n\Big\langle p\Big|\overset{(1)}{V}_0\frac{1}{1+p^2}\overset{(2)}{V}_0\frac{1}{1+p^2}\cdots\frac{1}{1+p^2}\overset{(n)}{V}_0\Big|p'\Big\rangle. \quad (55)$$

$\bar{K}_n$ is defined by iteration according to

$$\bar{K}_{n+1}=\bar{K}_n(-V_0)(1+p^2)^{-1} \quad (56)$$
$$\bar{K}_1=V_0(1+p^2)^{-1}V_0.$$

Since we shall dominate the wave functions by S-state wave functions, it will be sufficient to consider only the angular average of $\bar{K}_n(p,p')$ over the angles of p, p'. This two-dimensional function will be denoted by $K_n(p,p')$.

After the required integration over the angles of p, p' is performed, $K_n(p,p')$ is given inductively from (56) and (28) by the relations

$$K_n(p,p')=\frac{1}{(2\pi)^3}\frac{2\pi}{p'}\int_0^\infty K_{n-1}(p,x)\frac{x\,dx}{(1+x^2)}\ln\left|\frac{x+p'}{x-p'}\right|; \quad (57)$$

$$K_1(p,p')=\frac{1}{(2\pi^2)^2}\frac{\pi}{pp'}\int_0^\infty\frac{dx}{1+x^2}\ln\left|\frac{x+p}{x-p}\right|\ln\left|\frac{x+p'}{x-p'}\right|. \quad (58)$$

The properties of the $K_n(p,p')$ that we shall need in later work are listed in the following theorem. Here and in the following, $b(p,p')$ or $b(p)$ denote *bounded* functions of their arguments.

Theorem 1. $K_n(p,p')$ has the following properties:

1. $K_n(p,p')=K_n(p',p)$.
2. Let $R_n(p,p')=pK_n(p,p')p'$.

   Then, $R_n(p,p')=R_n(1/p,\,1/p')$.

3. $$K_1(p,p')=\frac{1}{(2\pi^2)^2}\frac{\pi^3}{pp'}\times(\tan^{-1}p\,\theta(p'-p)+\tan^{-1}p'\theta(p-p')-2/\pi\tan^{-1}p\tan^{-1}p']$$

   $\theta(x)=1,\frac{1}{2},0$ for $x>$, $=$, or <0, respectively.

4. $K_2(p,p')$ has the behavior

   (a) For $p<p'<1$,

   $$K_2(p,p')=b(p,p')+b_2(p)\ln p',$$

   where

   $$b_2(p)=-\frac{4\pi^4}{(2\pi^2)^3}(\tan^{-1}p)/p;$$

   (b) For $p\langle 1;\,p'\rangle 1$,

   $$K_2(p,p')=b_3(p,p')\frac{\ln p'}{p'^2}+\frac{b_4(p,p')}{p'^2}.$$

The behavior in other regions can be determined from properties 1 and 2.

5. For $n\geq 3$, $K_n(p,p')$ is a bounded function of p, p'. Moreover, for the case $p\langle 1;\,p'\rangle 1$, $K_n(p,p')$ has the behavior

$$K_n(p,p')=b(p,p')/p'^2.$$

The behavior in other regions can be found from properties 1 and 2.

6. $K_n(p,p')$ is a continuous function of p, p' for $n\geq 3$. For $n<3$, $K_n(p,p')$ is continuous except for $p=p'=0$.

Let us note that $K_1(p,p')$ diverges linearly if both p and p' approach zero together but does not diverge if one of them is held fixed. A similar statement holds for the logarithmic divergence of $K_2(p,p')$. Thus, $pK_n(p,p')$ and $pK_n(p,p')p'$ are bounded and continuous functions for $n\geq 1$.

The proof of theorem 1 is not difficult and will be only briefly described. Statement 1 of theorem 1 follows from the similar property of $\bar{K}_n$. Statement 2 follows by induction from (57) and (58) if one notes that $dx/(1+x^2)$ and $\ln(x+p)/(x-p)$ are separately invariant to the transformation $x\Rightarrow 1/x$, $p\Rightarrow 1/p$. After establishing 3 and 4, the relations 5 and 6 are easily proved by induction using the representation (57).

The evaluation of $K_1(p,p')$ given in statement 3 is easily obtained from (58) using contour integration. Since each of the two logarithmic factors is an odd function of x, the x integration can be extended to $-\infty$. Let us assume that the arguments of the four component logarithmic factors in the complex x plane run from $-\pi$ to π. Then, our procedure is to draw branch cuts downwards through the four points $x=\pm p$, $\pm p'$, close the contour from above, and pick up the residue at $x=+i$. The result is automatically real, but an unwanted

[23] The quantities $\bar{K}_n$ are essentially the iterated kernels of the integral equation satisfied by the nonrelativistic bound electron's Green's function $G\equiv(p^2/2+V+\beta)^{-1}$ with $2\beta=1$: $G=G_0+G_0VG$, where $G_0=(p^2/2+\beta)^{-1}$.

contribution arising from the product of the imaginary parts of the logarithms must be subtracted off.

In addition to the quantities $K_n(p,p')$ studied above, we will be interested in the more general quantities $K_n^{\nu}(p,p')$ defined by

$$K_n^{\nu}(p,p')=(-1)^{n+1}\left\langle p\left|p^{\nu_0}V_0\frac{p^{\nu_1}}{1+p^2}V_0\cdots\right.\right.$$
$$\left.\left.\frac{p^{\nu_n}}{1+p^2}V_0p^{\nu_{n+1}}\right|p'\right\rangle_{\text{av}} \quad (59)$$
$$\nu_i=0,1 \quad i=0 \text{ to } n.$$

ν is an abbreviation for the set of numbers $(\nu_0,\nu_1,\cdots\nu_n)$. The average indicated is an angular average over the angles of p and p'. The notation is really somewhat symbolic since it is understood that the intermediate integrations are to be carried out in p space and the numerator p's refer to the absolute values of the vector p's.

The quantities K_n will arise in the following subsection in connection with dominating arguments in p space for the general elementary term, which contains a numerator polynomial in the p's. To carry through these dominating arguments, it turns out to be sufficient to study the single iterated quantity $L_n(p,p')$ defined below, essentially the "worst" case of the $K_n(p,p')$:

$$L_n(p,p')=(-1)^{n+1}\left\langle p\left|\frac{1+p}{1+p^2}\overset{(1)}{V_0}\frac{1+p}{1+p^2}\cdots\right.\right.$$
$$\left.\left.\overset{(n+1)}{V_0}\frac{1+p}{1+p^2}\right|p'\right\rangle_{\text{av}}. \quad (60)$$

$L_n(p,p')$ is defined by iteration according to the relations

$$L_n=L_{n-1}(1+p^2)^{-1}(-V_0)(1+p)$$
$$L_1=(1+p)V_0[(1+p)/(1+p^2)]V_0(1+p) \quad (61)$$

or

$$L_n(p,p')=\frac{2\pi}{2\pi^2}\frac{(1+p')}{p'}\int_0^{\infty}L_{n-1}(p,x)\frac{xdx}{1+x^2}\ln\left|\frac{x+p'}{x-p'}\right| \quad (62)$$

$$L_1(p,p')=\frac{\pi}{(2\pi^2)^2}\frac{(1+p)(1+p')}{pp'}\int_0^{\infty}dx\frac{1+x}{1+x^2}$$
$$\times\ln\left|\frac{x+p}{x-p}\right|\ln\left|\frac{x+p'}{x-p'}\right|. \quad (63)$$

The properties of the $L_n(p,p')$ which we shall need are stated in the following theorem.

Theorem 2. $L_n(p,p')$ has the following properties:

1. $L_n(p,p')=L_n(p',p)$
2. $L_1(p,p')=b(p,p')+K_1(p,p')$
$$+c[\ln p'\theta(p'-p)\theta(1-p')$$
$$+\ln p\theta(p-p')\theta(1-p)].$$
3. $L_2(p,p')=b(p,p')+b(p,p')K_2(p,p')$.
4. For $n\geq 3$, $L_n(p,p')$ is a bounded function of p and p'.
5. For $n\geq 3$, $L_n(p,p')$ is a continuous function of p, p'. For $n<3$ $L_n(p,p')$ is continuous except at the point $p=p'=0$.

The proofs are straightforward and similar to those of theorem 1. It is convenient to prove inductively the statement that $L_n(p,p')=b(p,p')+K_n(p,p')b(p,p')$ for $n\geq 2$ and then use the result 5 of theorem 1.

Existence and Lowest Order of Elementary Terms of I_n^k

We are now ready to derive the result of main interest in the present section which is stated below.

Theorem 3. Let I_n^{ν} denote an elementary term of the k expansion with nominal order ν. Then I_n^{ν} exists for $n\geq 2$, and for $n>3$:

(a) $\nu<(n+3)/2$: $I_n^{\nu}=cw^{\nu}$+higher order.

(b) $\nu>(n+3)/2$: $I_n^{\nu}=cw^{(n+3)/2}$+higher order

(c) If n is odd and $\nu=(n+3)/2$: $I_n^{\nu}=cw^{(n+3)/2}\ln w$ +higher order.

For $n=3$:

(a) is still true but (b) and (c) are replaced by

(b) $I_s^{\nu>3}=cw^3\ln w$+higher order.

(c) $I_s^{\nu=3}=cw^3\ln^2 w$+higher order.

For $n=2$:

(a) $\nu=1$ $I_2^{\nu=1}=cw$+higher order.

(b) $\nu=2$ $I_2^{\nu=2}=cw^2\ln w$+higher order.

(c) $\nu>2$ $I_2^{\nu}=cw^2$+higher order.

Proof of theorem 3. With the aid of dominating arguments in p space and appropriate transformations of variables, we shall reduce the problem to the boundedness properties of the iterated operators K_n and L_n studied in the previous subsection. Essential to the dominating arguments is the fact that $\langle p|-V_0|p'\rangle>0$.

In order to eliminate trivial cases, we shall consider only elementary terms with $v=w=0$ and with the ϵ^2 of the numerator $[(\tilde{p})^2+\epsilon^2]^k$ and denominator $\bar{\Delta}$ replaced by ϵ_{sh}^2, twice the Schrödinger binding energy. Furthermore, we shall consider only the S-state case. Because of the improved behavior of the non-S state wave functions at high p, the dominating arguments given for the S-state case will apply even more strongly to the S-state case.

Let us first consider the large-large Schrödinger approximation. After carrying out the Dirac algebra and the angular part of the k integration, the relevant

elementary terms are of the form

$$I_n{}^\nu \equiv I_{n;qrsu}{}^k = mw^\nu \int_w^\infty d\sigma \sigma^{n-s-u+2} \int_0^\infty d\lambda \lambda^{\frac{1}{2}q+1} \int dY_{n-1}$$

$$\times \text{“} \left\langle u_s \left| \frac{(-V_0)^n Q_{n;rs}(p,\bar{p})[(\bar{p})^2+\epsilon^2]^k}{[\lambda\sigma^2+1+\sigma\bar{\Delta}]^{n+k+2}} \right| u_s \right\rangle \text{”} \quad (64)$$

$$\nu = k+q+(r+3s)/2+u+1, \quad (65)$$

where $\lambda \equiv k^2$ and Q is a scalar polynomial in the vectors p, $\bar{p}$.

We shall dominate I_n by dominating the momentum space wave functions by the $1-S$ wave function and replacing the vectors p and $\bar{p}$ in Q by their absolute values. Thus, we use

$$|Q_{n;rs}(p,p')| < \prod_{i=0}^{n} p_i{}^{\nu_i} \bar{p}^s, \quad (66)$$

where

$$\nu_i = 0, 1, \quad \Sigma \nu_i = r. \quad (67)$$

It is now convenient to divide the range of the σ integration into the intervals $(w,1)$ and $(1,\infty)$. Let us define $I_n(\sigma)$ and $I_n(\alpha,\beta)$ by the relation

$$I_n{}^\nu(\alpha,\beta) = \int_\alpha^\beta d\sigma I_n{}^\nu(\sigma). \quad (68)$$

Then, we divide $I_n{}^\nu$ into two terms according to

$$I_n{}^\nu = I_n{}^\nu(w,1) + I_n{}^\nu(1,\infty). \quad (69)$$

A. $I_n(w,1)$

The numerator dependence on $\bar{p}$ and ϵ^2 may be removed by using the inequalities (18) and (7). This changes the power of the denominator from $n+2+k$ to $n+2-s/2$ and changes the power of σ in the numerator from $n-s-u+2$ to $n-k-3s/2-u+2$. Next, we perform the transformation

$$\lambda \Rightarrow \lambda\sigma^{-2}. \quad (70)$$

To make use of the previously studied iterated operators, it is necessary to go over to a separable representation with the aid of the identity (13). For this purpose we must change the power of the denominator from $n+2-s/2$ to $n+1$. Since the limits of the λ integration run from 0 to ∞, this is easily accomplished by adjusting the power of λ in the numerator. Having gone over to a separable representation, we use the domination

$$\lambda+1+\sigma\Delta > \lambda+1+\sigma\langle p^2\rangle_{\text{av}} \quad (71)$$

and perform the transformations

$$p_i \Rightarrow [(\lambda+1)/\sigma]^{\frac{1}{2}} p_i, \quad i=1,2,\cdots,(n-1). \quad (72)$$

We obtain finally the inequality

$$|I_n{}^\nu(w,1)| < \text{const} mw^\nu \int_w^1 d\sigma \sigma^{(n+1)/2-\nu} \int_0^\infty d\lambda$$

$$\times \frac{\lambda^{(q+s)/2}}{[\lambda+1]^{(n+5-r)/2}} \int d^3p_0 d^3p_n \langle u_s | p_0 \rangle$$

$$\times \frac{1}{1+p^2} K_{n-1}{}^\nu(p,p') \frac{1}{1+p'^2} \langle p_n | u_s \rangle, \quad (73)$$

where

$$p \equiv [\sigma/(\lambda+1)]^{\frac{1}{2}} p_0; \quad p' \equiv [\sigma/(\lambda+1)]^{\frac{1}{2}} p_n. \quad (74)$$

From the definitions (59) and (60) we see that

$$K_{n-1}{}^\nu < L_{n-1}(p,p'). \quad (75)$$

The results of theorem 3 now follow easily from the known boundedness properties of the $L_n(p,p')$ given in theorem 2.

B. $I_n(1,\infty)$

$I_n(1,\infty)$ is formally proportional to ω^ν. We shall show that $I_n(1,\infty)$ exists and therefore the nominal order is in fact the true order.

This time, the ϵ^2 in the denominator $\bar{\Delta}$ of (64) is in general essential to the convergence of the σ integration at high σ and cannot be removed. Instead of the domination (71), we use therefore the domination

$$\lambda+1+\sigma\Delta > \lambda+\sigma\Delta. \quad (76)$$

After an elementary analysis similar to the one used in the preceding case, we obtain the inequality

$$|I_n{}^\nu(1,\infty)| < \text{const} mw^\nu \int_0^\infty d\lambda \frac{\lambda^{(q+s)/2}}{[\lambda+1]^{(n+5-r)/2}}$$

$$\times \int d^3p_0 d^3p_n \langle u_s | p_0 \rangle$$

$$\times \frac{1}{1+p^2} K_{n-1}{}^\nu(p,p') \frac{1}{1+p'^2} \langle p_n | u_s \rangle \quad (77)$$

with

$$p \equiv [1/(\lambda+1)]^{\frac{1}{2}} p_0; \quad p' \equiv [1/(\lambda+1)]^{\frac{1}{2}} p_n. \quad (78)$$

From the properties of the $L_{n-1}(p,p')$ it is easy to see that the right-hand side of (77) exists for $n \geqslant 2$.

This completes the discussion of the proof of theorem 3 for the large-large Schrödinger case. The analyses of the $++$ and $--$ Schrödinger cases and Dirac corrections are straightforward. For the Dirac corrections one uses the general form (46).

3. A THEOREM ON THE NONANALYTIC PART OF I_n

The main result of the present section is the following theorem which states that for the particular class of terms nonanalytic in w, the free-propagator expansion

is an expansion in orders of w. Thus, the free-propagator expansion is especially suitable for the calculation of this class of terms.

Theorem 4. (The "main result.") Assuming the validity of the k expansion, the nth term I_n of the free-propagator expansion has the following property. For $n>3$, the lowest order of terms in I_n nonanalytic in w is $w^{(n+3)/2}$ for n even and $w^{(n+3)/2}\ln w$ for n odd; for $n=3$, it is $w^3\ln^2 w$.

Note that this theorem coincides with theorem 3 for terms of the k expansion of nominal order

$$\nu \geqslant (n+3)/2.$$

According to the theorem, the previously calculated Lamb shift orders $w^2\ln w$ and $w^{\frac{5}{2}}$ should be calculated correctly by considering only $n=0$, 1, and 2. This result will be verified in the calculations of the following sections and is in agreement with the earlier results of F-Y and Kroll and Pollock mentioned in the introduction. The orders $w^3\ln^2 w$ and $w^3\ln w$ which we wish to calculate arise according to the theorem only for $n=0$, 1, 2, and 3.

Although the proof of the above theorem can be obtained by a refinement of the proof of theorem 3, the formulation and the result of the theorem become more natural, as will be seen, when w is regarded as a complex variable. We shall therefore first derive the theorem using a complex variable approach. Later on in this section we shall briefly show how the same results can be obtained when w is considered to be a real variable.

Preliminary Theorems

When one considers the Lamb shift orders $w^2\ln w$ and $w^{\frac{5}{2}}$ for which only a few terms of the free-propagator expansion are needed and seeks a simple criterion to distinguish them from orders such as w^2 for which the infinite expansion is necessary, the nonsingle-valued nature of the former terms as functions of the complex variable w suggests itself naturally. Thus, if one requires the functions $w^2\ln w$ and $w^{\frac{5}{2}}$ to be, like I_n, real on the positive real axis, then these functions are analytic in any circular neighborhood of $w=0$ cut along the negative real axis but have along the cut a finite imaginary discontinuity. (We shall show later that I_n can be analytically continued in this cut region.) The values of the discontinuity Δ of "upper" minus "lower" values are determined by the easily derived relations

$$\Delta\ln w=2i\pi, \quad \Delta w^{\frac{5}{2}}=2i|w|^{\frac{5}{2}}. \tag{79}$$

Obviously, $\Delta A(w)=0$ where $A(w)$ is any single-valued function and in particular an analytic function of w.

Let us now define the "original" corresponding to a given discontinuity to be any function which has the specified discontinuity. Then, we are interested in the extent to which the discontinuity determines the original. It turns out that under the conditions of interest the original is determined to within an analytic function of w. Moreover, the discontinuity, as in the examples (79), is always a purely imaginary quantity and equal to twice the imaginary part of the original on the upper side of the cut.

These results are contained in the elementary theorems 5–7 and given below. Theorem 5 is readily seen to follow from Morera's theorem[24]; theorem 6 is an immediate consequence of the so-called reflection principle[25]; theorem 7 is a corollary of theorems 5 and 6 since the only isolated singularities are poles and essential singularities and a function having these singularities can not be finite in a neighborhood of $w=0$.

Theorem 5. Let $f(w)$ and $g(w)$ be functions of the complex variable w. Let G be the annular region in the w plane formed by a circle of radius W_m[26] and a circle of radius ϵ where ϵ is arbitrarily small but nonzero. Let G' denote the same region cut along the negative real axis. Let $h(w)\equiv f(w)-g(w)$. We suppose that f and g satisfy the following conditions:

(1) Both functions are analytic in G';

(2) values of the functions along the cut called upper and lower values may be assigned such that the functions are continuous in closed upper and lower neighborhoods of the cut, respectively. The associated values of f and g are denoted by $f^+(w_c)$ and $g^+(w_c)$ for upper values of $\bar{f}(w_c)$, $\bar{g}(w_c)$ for lower values, where w_c is a value of w along the cut.

(3) $h^+(w_c)=h^-(w_c)\equiv h^*(w_c)$ for all values of w. Then $h(w)$ defined along the cut as $h^*(w_c)$ is analytic in G and therefore, since ϵ is arbitrarily small, $h(w)$ has at most an isolated singularity at $w=0$.

Theorem 6. Let $f(w)$ be a function satisfying the conditions (1) and (2) of theorem 5 and also real along the positive real w axis. Then the discontinuity $f^+(w_c)-f^-(w_c)$ along the cut is twice the imaginary part of $f^+(w_c)$.

Theorem 7. Let $f(w)$ be a function satisfying the conditions (1) and (2) of theorem 5. Let $\mathrm{Im}_+f(w)$ be defined as the imaginary part of $f(w)$ along the upper side of the cut. Then, if $f(w)$ is real along the positive real axis and finite in a neighborhood of $w=0$, a knowledge of $\mathrm{Im}_+f(w)$ is sufficient to determine $f(w)$ to within an analytic function of w.

In view of theorem 7, the "main result" admits of the following simpler formulation.

Theorem 8. Assuming the validity of the k expansion, the lowest order of Im_+I_n is $w^{(n+3)/2}$ for $n>3$ and $w^3\ln w$ for $n=3$.

---

[24] K. Knopf, *Theory of Functions* (Dover Publications, New York, 1945), p. 66.

[25] E. C. Titchmarsh, *The Theory of Functions* (Oxford University Press, New York, 1958), p. 155.

[26] In our problem of the analytic continuation of I_n, it seems necessary to take $W_m<1$ for the $1-S$ state since, for example, the formal expression for the Dirac energy $E=(1-w)^{\frac{1}{2}}$ has a branch point at $w=+1$. However, a rigorous derivation of the maximum value of W_m for the stated analyticity properties is not attempted.

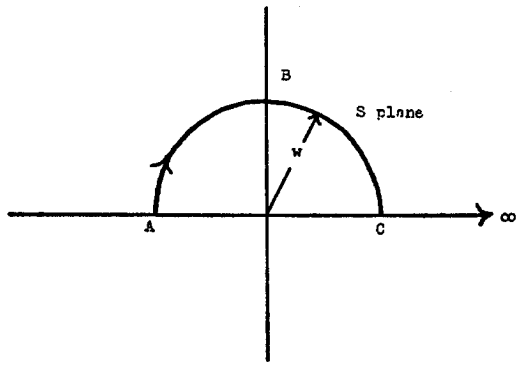

FIG. 2. Rotated path of integration for example 2.

Since I_n is in the form of a rather complicated multiple integral involving the parameter w, it is worthwhile, before proceeding to the derivation of theorem 8, to give a few illustrative examples of the determination of $\mathrm{Im}_+ f(w)$ when $f(w)$ involves a one-dimensional integral containing the parameter w.[27] The second example is similar to the situation which will arise for I_n in the derivation of theorem 8. The first example is introduced because of its simplicity and because it is similar to the treatment that will actually be employed in the detailed analysis of I_0 in Sec. 4.

Example 1.

$$f(w)\equiv\int_0^1 \frac{dx}{x+w}g(x) \quad 0\leq w<1. \tag{80}$$

$g(x)$ is a real continuous function of x. Note that $f(w)$ is real for real w and is analytic in the w plane cut along the negative real w axis. Furthermore, the analytic continuation in this region can be effected by simply letting the parameter w assume complex values.

Since

$$\mathrm{Im}_+ 1/(x+w)=-\pi\delta(x-w), \tag{71}$$

we have immediately

$$\mathrm{Im}_+ f(w)=-\pi g(w). \tag{82}$$

Example 2.

$$f(w)=g(w)\int_w^\infty ds s^\mu \exp(-s^{\frac{1}{2}}) \quad 0<|w|<1. \tag{83}$$

Here, μ is a real number and $g(w)$ is an analytic function of w, *real* along the positive real axis. Note that since it is analytic $g(w)$ is also real on the cut along the negative real axis.

To continue the integral, the lower limit of the path of integration will be "rotated," as shown in Fig. 2, from the point C on the positive real axis to the point A on the negative real axis along the semicircular arc CBA of radius w.

[27] In these examples and in the following work, Im_+ is to be regarded as an operation in which the argument of the function $f(w)$ acted upon is rotated in a circular arc of radius w from a point on the positive real axis to a point on the (upper) negative real axis. The imaginary part is extracted in the limit $\theta=\pi$, where θ is the angle of rotation.

The imaginary contribution arises only along the semicircular path ABC. Thus,

$$\mathrm{Im}_+ f(w)=g(-w)\,\mathrm{Im}_+\int_{ABC} s^\mu \exp(-s^{\frac{1}{2}})ds \tag{84}$$

or, letting

$$s\equiv we^{i\theta}, \tag{85}$$

$$\mathrm{Im}_+ f(w)=-g(-w)w^{\mu+1}\,\mathrm{Im}_+ ih(w)$$

$$=-g(-w)w^{\mu+1}Rh(w) \tag{86}$$

$$h(w)\equiv\int_0^\pi d\theta \exp\{i(\mu+1)\theta-w^{\frac{1}{2}}e^{i\theta/2}\}. \tag{87}$$

Analytic Continuation of I_n and Derivation of Theorem 8

We begin with the expression (39) for I_n. We observe that I_n is real for positive w. $v_s(w)$ is an analytic function of w in either a position or momentum representation. (More accurately, ψ_s and ϕ_s are analytic functions of w.) To analytically continue I_n we would like to transform the path of the σ integration as in example 2 of the preceding subsection to the path ABC of Fig. 2 while replacing w by $-w$ in the integrand. The expression as it stands is unsuitable for a formal analytic continuation since the denominator may become zero when σ is replaced by $-w$. To avoid these poles of the integrand we perform the transformation

$$k\Rightarrow\sigma^{-1}k \tag{88}$$

as in (70) and also the canonical scaling transformation

$$\begin{aligned} p&\Rightarrow\sigma^{-\frac{1}{2}}p \\ x&\Rightarrow\sigma^{+\frac{1}{2}}x \\ |v_s(w)\rangle&\Rightarrow|v(\sigma,w)\rangle. \end{aligned} \tag{89}$$

The new wave functions $v(\sigma,w)$ in position space reduce in the Schrödinger approximation to the wave-functions $u(\sigma)$ which for the $1-S$ state are given by

$$u(\sigma)=\mathrm{const}\sigma^{\frac{1}{2}}\exp(-\sigma^{\frac{1}{2}}r). \tag{90}$$

The denominator is now positive definite even when $\sigma\Rightarrow-w$ (and $w\Rightarrow-w$). The integral over the circular path ABC of Fig. 2 involves a "rotation" of the wave functions $v(\sigma,w)$ through complex values of σ but because of the negative real part of the exponent in $u(\sigma)$, formula (91), the integrals will remain convergent.

Because of "selection rules" 1 and 2 of Sec. 1 and the analytic nature of $v(\sigma,w)$ as a function of w, the integrand remains real after the replacement $w\Rightarrow-w$. Thus, I_n is real along the path C to ∞ as in example 2 of the previous subsection and as in that example, we can again restrict the path of integration to the circular path ABC. On introducing again the parameter θ by the relation

$$\sigma=we^{i\theta}, \tag{91}$$

we obtain finally, omitting an unimportant numerical factor,

$$\mathrm{Im}_+ I_n = \mathrm{Im}_+ imw^{n/2}\int_0^\pi d\theta\,\exp[i\theta(n/2-1)]\int d^4k\int dY_{n-1}$$

$$\times\text{“}\left(\bar v^*(we^{i\theta},-w)\left|\frac{\gamma_\mu V_0\prod_{i=0}^{n}N_s(ie^{-i\theta/2}p_i)\gamma_\mu}{[k^2+1+\langle p^2\rangle_{\mathrm{av}}+e^{-i\theta}(\bar p)^2+w(e^{i\theta}+1)\epsilon^2]^{n+2}}\right|v(we^{i\theta},-w)\right). \quad (92)$$

It is understood that the argument of ϵ^2 is now $-w$. Finally, in accordance with (72), we perform the transformations

$$p_0 \Rightarrow w^{\frac{1}{2}}p_0$$
$$p_n \Rightarrow w^{\frac{1}{2}}p_n. \quad (93)$$

Then

$$\langle p\,|\,v(\sigma,w)\rangle \Rightarrow w^{-\frac{1}{2}}\langle p\,|\,v(e^{i\theta},w)\rangle.$$

We obtain

$$\mathrm{Im}_+ I_n = \mathrm{Im}_+ imw^{(n+3)/2}\int_0^\pi d\theta\,\exp[i\theta(n/2-1)]\int d^3p_0 d^3p_n\langle\bar v^*(e^{i\theta},w)\,|\,p_0\rangle M(w^{\frac{1}{2}}p_0,w^{\frac{1}{2}}p_n)\langle p_n\,|\,v(e^{i\theta},w)\rangle, \quad (94)$$

where

$$M(p_0,p_n)\equiv\int d^4k\int dY_{n-1}\text{“}\left\langle p_0\left|\frac{\gamma_\mu V_0{}^n\prod_{i=0}^{n}N_s(ie^{-i\theta/2}p_i)\gamma_\mu}{[k^2+1+\langle p^2\rangle_{\mathrm{av}}+e^{-i\theta}(\bar p)^2+w(e^{i\theta}+1)\epsilon^2]^{n+2}}\right|p_n\right\rangle\text{”}. \quad (95)$$

The formal limit of this expression as $w\Rightarrow 0$ is

$$\mathrm{Im}_+ I_n \Rightarrow amw^{(n+3)/2}u_s{}^2(0) \quad (96)$$

$$a=\text{const}\,\mathrm{Im}_+ i\int_0^\pi d\theta\,\exp[i\theta(n+1)/2]M(0,0)\Big\rangle. \quad (97)$$

We have noted that the rotated position space wave functions $a(e^{i\theta})$ have the same value at the origin as the scaled Schrödinger wave functions u_s except for a factor of $e^{\frac{1}{2}i\theta}$.

If one admits the validity of the k expansion, in this case an expansion of $M(0,0)$ in powers of $e^{-i\theta}(\bar p)^2$, then for $n>3$, the existence of $M(0,0)$ is assured by the results of Sec. 2. Furthermore, for the individual terms of the k expansion, $M(0,0)$ is of the form

$$\text{const}e^{-(\nu-1)\theta},$$

where ν is the nominal order of the original elementary term of the k expansion as evaluated in (65).

The constant “a” in (97) is thus always finite. If n is odd and the k expansion is used, then we see that the coefficient “a” vanishes unless the condition

$$\nu=(n+3)/2$$

is satisfied.

For $n=3$, it is easy to see that the logarithmic singularity at small p_0, p_n, due to the behavior of $K_2(p,p')$ or $L_2(p,p')$ at small p, p', leads to the appearance of an “extra” factor of $\ln w$ in the lowest order of $\mathrm{Im}_+ I_n$.

This completes the derivation of theorem 8. We have also shown the following stronger results.

Theorem 9. The lowest order terms $w^{(n+3)/2}$, $w^{(n+3)/2}\times\ln w$, and (for $n=3$) $w^3\ln^2 w$ mentioned in the main results are proportional to $u_s{}^2(0)$, the square of the scaled Schrödinger wave-function at the origin in position space. Furthermore, if n is odd, then for the individual terms of the k expansion, the coefficients of $w^{(n+3)/2}\ln w$ or $w^3\ln^2 w$ vanish unless the normal order satisfies the condition

$$\nu=(n+3)/2.$$

The proof given of theorem 8 is formal and nonrigorous at a number of points, particularly in the treatment of the passage to the limit from Dirac to Schrödinger wave functions.

A “rigorization” of the derivation of the main result is probably easiest to achieve using the real-variable methods described briefly in the following subsection.

Real-Variable Methods. Differentiation Device

In distinction to the complex variable method just discussed in which a closed form could be preserved until the last step, real-variable methods have the disadvantage that they depend on performing the k expansion at an early stage.

As mentioned earlier, the “main result” can be derived by a refinement of the proof of theorem 3. We wish here to discuss a second real-variable method which is similar to the preceding complex-variable approach in that it also discards information about the analytic part of I_n and restricts the parameter σ to values of absolute magnitude w. This method will be called the “differentiation device.”

Let us consider the expression for a particular

elementary term. After performing the transformation (70) and carrying out the λ integration, this is of the form, corresponding to (64),

$$I_n^{\nu}=mw^{\nu}\int_w^{\infty}d\sigma\sigma^{n-1-\nu}\int dY_{n-1}$$
$$\times''\left\langle v_s(w)\left|\frac{(-V_0)^n Q_n{}^k(\sigma^{\frac{1}{2}}p,\sigma^{\frac{1}{2}}\bar{p})}{(1+\sigma\bar{\Delta})^{n+k-q/2}}\right|v_s(w)\right\rangle'', \quad (98)$$

where $Q_n{}^K(p,\bar{p})$ is a scalar polynomial in the vectors p_i, $\bar{p}$ for $i=0, 1, \cdots, n$. One can now expand the analytic dependence of $v_s(w)$ on w leaving an "internal" dependence on w only in the lower limit of the σ integration where it can be removed by differentiation.

For definiteness, let us consider the generalized Schrödinger approximation. Then, we form the differential equation

$$\frac{d}{dw}\left(\frac{I_n^{\nu}}{mw^{\nu}}\right)=-w^{n-1-\nu}\int dY_{n-1}$$
$$\times''\left\langle u_s\left|\frac{(-V_0)^n Q_n{}^k(w^{\frac{1}{2}}p,w^{\frac{1}{2}}\bar{p})}{(1+w\bar{\Delta})^{n+k-q/2}}\right|u_s\right\rangle'', \quad (99)$$

This is of the general form

$$(d/dw)[f(w)/w^{\nu}]=\text{const}w^{\mu}+\text{higher order} \quad (100)$$

which has as solution

$$\text{for } \mu\neq -1 \quad f(w)=c_1w^{\nu}+c_2w^{\mu+\nu+1}+0>w^{\mu+\nu+1} \quad (101)$$

$$\text{for } \mu=-1 \quad f(w)=c_3w^{\nu}\ln w+0>w^{\nu}\ln w, \quad (102)$$

where the symbol $0>$ means "higher order than."

In the case of our differential equation (99) and $n>3$ one verifies by familiar methods that $\mu=(n+1-\nu)/2$ and therefore, that, according to (101) and (102), the lowest order nonanalytic terms are proportional to $w^{(n+3)/2}$ and $w^{(n+3)/2}\ln w$ for n even or odd, respectively.

Additional Remarks

It is of interest to point out a few results of a more detailed nature than those obtained so far.

The derivation of theorem 8 shows clearly that the lowest order nonanalytic terms are given by the generalized Schrödinger approximation. In fact, a closer analysis shows that only the $++$ part enters in lowest order. This is due to the fact that $pL_n(p,p')$ and $pL_n(p,p')p'$ approach zero in the limit $p=p'=0$ for $n\geq 2$, where L_n is defined in Sec. 2. More generally and for the same reason, the coefficient of the lowest order $w^{(n+3)/2}$ or $w^{(n+3)/2}\ln w$ terms vanish whenever the powers ν_0, ν_n of p_0, p_n in the numerator are greater than zero.

Finally, we remark that the anomalous position of $n=2$ and 3 relative to the main result, namely, the appearance of the orders $w^2\ln w$ and $w^3\ln^2 w$ instead of $w^{\frac{5}{2}}$ and $w^3\ln w$ for $n=2$ and 3, respectively, occurs only for the case $k=r=s=0$ and $++$. This follows in a straightforward way from the boundedness properties of the quantities $K_1{}^{\nu}$ or $K_2{}^{\nu}$ given in the Appendix.

4. TREATMENT OF I_0 AND I_1

In the preceding sections we examined "order" properties of the general term I_n of the free-propagator expansion. We established in particular that the new orders $w^3\ln^2 w$ and $w^3\ln w$ which we wish to calculate arise only for $n\leq 3$. In this section and the following paper we shall study these lowest terms of the free-propagator expansion in more detail. In the actual calculations we shall use the usual k expansion and the "differentiation device" for I_n with $n>0$. On the other hand, for I_0 we shall use the "rotation trick" to obtain a closed expression for the terms of interest. That is, we shall form $\text{Im}_+\langle I_0\rangle$ and take the original to the desired order in w. This gives an independent and especially concise treatment of the non-analytic part of I_0.

We turn first to a consideration of the terms I_0 and I_1 for which "ultraviolet" divergences associated with mass and charge renormalization[14] have made inapplicable the general analysis of the first three sections.

It is a well-known consequence of gauge invariance that charge renormalization is unnecessary in this problem and that the charge renormalization subtractions cancel between I_0 and I_1.[28] If charge renormalization is performed, one introduces infrared divergences into I_0 and I_1 separately which cancel against each other. If charge renormalization is not performed then these infrared divergences are avoided but one obtains "spurious" lowest order terms of order $w\ln w$ which again cancel between I_0 and I_1 (see F-Y).

As we have noted in the introduction, the complications due to mass and charge renormalization are removed when one considers only the nonanalytic part of I_0.

Formation of $\text{Im}_+ I_0$

We shall form $\text{Im}_+ I_0$ directly without passing through the step of the mass renormalization of I_0. Let us note first that the nonanalytic part of I_0 is invariant to the customary shift in the origin of the k integration, though this is not true of I_0 itself due to the presence of a linearly divergent k integration.[29] The reason for this is that the "extra" term occasioned by the shift is proportional to $\langle\bar{v}|\mathbf{p}|v\rangle=\langle\bar{v}|\mathbf{V}|v\rangle-m$, which is an analytic function of w. Thus, we have from (5), with

[28] An explicit proof is given by F-Y. From the point of view of the general renormalization theory (compare footnote 14), the cancellation is a consequence of the identity of the renormalization constants usually labeled Z_1 and Z_2 (Ward's identity).

[29] We refer to the standard formal treatment of the linear divergence given by J. M. Jauch and F. Rohrlich (reference in footnote 18, Appendix).

$n=0$,

I_0(nonanalytic)

$$=-\int d^4k\int_0^1 dx\,\text{“}\left\langle\bar{v}\left|\frac{\gamma_\mu[m+(1-x)\mathbf{p}]\gamma_\mu}{[k^2+xD(x)]^2}\right|v\right\rangle\text{”},\quad(103)$$

where

$$D(x)=xm^2+(1-x)\Delta.\quad(104)$$

We have discarded in standard fashion the term linear in $\mathbf{k}$ on the grounds of symmetrical integration.

Let us now perform the usual scaling transformation (35) for the coordinates and momenta. Since Im_+I_0 is finite, we can interchange the order of the x and k integrations. Then we are interested in the evaluation of the quantity

$$F(w)\equiv\mathrm{Im}_+\int_0^\infty\frac{\lambda d\lambda}{[\lambda+a(w)]^2}\quad(105)$$

with $\lambda=k^2$ and $a(w)\equiv x[x+(1-x)w\Delta]$. On using (81) and partial differentiation, we see that

$$F(w)=\pi\theta\{-a(-w)\},\quad(106)$$

where in our case $\theta\{-a(-w)\}=\theta(w\Delta/(1+w\Delta)-x)$. In this way we obtain the following expression for Im_+I_0:

$$\mathrm{Im}_+I_0=\pi^3\left\langle\bar{v}_s(-w)\left|\int_0^{w\Delta/1+w\Delta}dx\{\gamma_\mu m-(1-x)\mathbf{p}_s\gamma_\mu\}\right|\right.$$
$$\left.\times v_s(-w)\right\rangle,\quad(107)$$

where

$$\mathbf{p}_s=imw^{\frac{1}{2}}\mathbf{p};\quad p_{s,0}=E(-w),\quad(108)$$

The argument $-w$ of Δ in the upper limit of the x integration has been omitted. We have used the fact that

$$0<w\Delta(-w)/[1+w\Delta(-w)]<1.\quad(109)$$

Performing the trivial x integration and carrying out the Dirac algebra, we obtain

$$\mathrm{Im}_+I_0=2m\pi^3\left\langle\bar{v}_s(-w)\left|\frac{w\Delta}{1+w\Delta}\left[1+\frac{1}{2}\frac{w\Delta}{1+w\Delta}\right]\right|v_s(-w)\right\rangle$$
$$-2\pi^3\left\langle\bar{v}_s(-w)\left|\frac{w\Delta}{1+w\Delta}\left[1-\frac{1}{2}\frac{w\Delta}{1+w\Delta}\right]\right.\right.$$
$$\left.\left.\times(\mathbf{p}_s+m)\right|v_s(-w)\right\rangle.\quad(110)$$

On separating off the lowest order term by the expansion $(1+w\Delta)^{-1}=1-w\Delta(1+w\Delta)^{-1}$ and using the Dirac equation for the second line of (110), we obtain

$$\mathrm{Im}_+I_0=2m\pi^3w\langle v_s(-w)|\gamma_4\Delta|v_s(-w)\rangle$$
$$-2m\pi^3w^2\left\langle v_s(-w)\left|\gamma_4\frac{\Delta^2}{1+w\Delta}\right|v_s(-w)\right\rangle$$
$$+m\pi^2w^2\left\langle v_s(-w)\left|\gamma_4\frac{\Delta^2}{(1+w\Delta)^2}\right|v_s(-w)\right\rangle$$
$$+2m\pi^3w^2\left\langle v_s(-w)\left|\frac{\Delta}{1+w\Delta}\right.\right.$$
$$\left.\left.\times\left[1-\frac{1}{2}\frac{w\Delta}{1+w\Delta}\right]V_0\right|v_s(-w)\right\rangle.\quad(111)$$

As in formula (110), it is understood that Δ should be replaced by $\Delta(-w)$.[30]

Let us form the "original" of the first term of (111) using the relation (for f analytic in w and real for real w)

$$f(w)\,\mathrm{Im}_+\ln w=\pi f(-w)\quad(112)$$

and also the identity $\Delta=2m(\mathbf{p}+m)-(\mathbf{p}+m)^2$ and the Dirac equation. We obtain

$$\mathrm{Orig}[\text{1st term of (111)}]=+4\pi^2mw\ln w\langle v_s|V_0|v_s\rangle$$
$$+2\pi^2mw^2\ln w\langle v_s|\gamma_4V_0^2|v_s\rangle.\quad(113)$$

We see that the first term of (113), and only this term of I_0, would be removed by "imaginary charge renormalization," that is, charge renormalization of Im_+I_0, and must therefore cancel against a similar term in I_1.

Treatment of I_1

As mentioned earlier, we shall employ the k expansion for I_1. It is easy to show that terms with $k\geqslant 1$ are at least of order $w^2\ln w$. It is clear that charge renormalization, if performed, would affect only $k=0$ terms. The term $I_{n;qrsu}{}^k=I_{1;2000}{}^0$ contains a logarithmically divergent k integration. One can use the "rotation trick" to extract directly the nonanalytic part of this term.

The lowest order $w\ln w$ contribution is contained in the finite term $I_{1;000}{}^0$. As in I_0, one can use the "rotation trick" to separate off the "renormalization term" proportional to $w\ln w\langle v_s|V_0|v_s\rangle$.

$w^2\ln w$ Calculation

On the basis of the previous analysis, the $w^2\ln w$ coefficient arises from I_0, I_1, and I_2, and it is a simple matter to verify this by an explicit calculation. This result was established earlier in F-Y. We remark only that after "imaginary charge renormalization" and the elimination of gauge-variant terms proportional

[30] For the special case of the $1-S$ state, a Coulomb potential and the usual gauge for V, the difference $\Delta(w)-(p^2+\epsilon_{sn}{}^2)$ vanishes and Δ is not a function of w. (ϵ_{sn} is twice the Schrödinger binding energy.)

to $\epsilon^2\langle V\rangle$ and $\langle V^2\rangle$, the entire $w^2 \ln w$ coefficient, proportional to $\langle q^2 V\rangle$, arises from the "vertex diagram," I_1. (Here, the brackets denote an expectation value with respect to Schrödinger wave functions.)

APPENDIX. BOUNDEDNESS PROPERTIES OF SOME K_n^r FOR $n=1, 2$

We give here boundedness properties of several of the $K_n^r(p,p')$ defined in Sec. 2 for $n=1$ and 2. The proofs are briefly outlined. The properties of these K_n supplement those of the $K_n(p,p')$ and $L_n(p,p')$ of Sec. 2 and permit stronger statements about I_n for the cases $n=2$ and 3 than theorem 3 of Sec. 2 and the "main result" of Sec. 3.

Theorem A. $K_1^{010}(p',p)$ has the following behavior:

(1) $K_1^{010}(p,p')=K_1^{010}(p',p)$.

(2) For $p<p'<1$: $K_1^{010}(p,p')=c\ln p'+b(p,p')$.

(3) For $p\langle 1; p'\rangle 1$: $K_1^{010}(p,p')$
$=[b(p,p')\ln p'/p'^2]+[b(p,p')/p'^2]$.

(4) For $p'>p>1$: $K_1^{010}(p,p')$
$=[b(p,p')\ln(p'/p)/p'^2]+[b(p,p')/p'^2]$.

Proof. The symmetry property (1) is clear from the definition of K_1^{010}. For the proof of (2)–(4), we write K_0^{010} in the form

$$K_1^{010}(p,p')\sim\frac{1}{pp'}\int_0^\infty dx\frac{x}{x^2+1}\ln\left|\frac{x+p}{x-p}\right|\ln\left|\frac{x+p'}{x-p'}\right|$$

and divide the range of x integration into the three intervals $(0,p)$, (p,p'), and (p',∞). For these separate regions, the analysis can be carried through in a simple manner by expandlng $K_1^{010}(p,p')$ in a doubly infinite series according to the expansion

$$\ln\left|\frac{1+x}{1-x}\right|=2\sum_0^\infty\frac{x^{\pm 2k+1}}{2k+1}\quad\begin{matrix}+: x^2<1\\ -: x^2>1\end{matrix}$$

for each of the two logarithmic factors.

Theorem B.

$$K_1^{020}(p,p')=\frac{2\pi^2}{(2\pi^2)^2}\left[\frac{\tan^{-1}p}{p}\frac{\tan^{-1}p'}{p'}+\frac{\pi}{2p}\left[1-\frac{\tan^{-1}p'}{p'}\right]\theta(p-p')+\frac{\pi}{2p'}\left(1-\frac{\tan^{-1}p}{p}\right)\theta(p'-p)\right].$$

Proof. This exact result is derived by contour integration in the same manner as the similar integral $K_1(p,p')$ of Sec. 2. Note that $K_1^{020}(p,p')$ is a bounded function.

Theorem C. $K_2^{0010}(p,p')$; $K_2^{0110}(p,p')$, and $K_2^{0200}(p,p')$ are bounded functions.

Proof. We can write these K_2 in the form.

$$K_2^{0110}(p,p')\sim\frac{1}{p'}\int_0^\infty dx K_1^{010}(p,x)\frac{x^2}{x^2+1}\ln\left|\frac{x+p'}{x-p'}\right| \tag{114}$$

$$K_2^{0010}(p,p')\sim\frac{1}{p'}\int_0^\infty dx K_1(p,x)\frac{x^2}{x^2+1}\ln\left|\frac{x+p'}{x-p'}\right| \tag{115}$$

$$K_2^{0200}(p,p')\sim\frac{1}{p'}\int_0^\infty dx K_1^{020}(p,x)\frac{x}{x^2+1}\ln\left|\frac{x+p'}{x-p'}\right|. \tag{116}$$

Since

$$\frac{1}{p'}\int dx\frac{x}{x^2+1}\ln\left|\frac{x+p'}{x-p'}\right|\sim\frac{\tan^{-1}p'}{p'}=b(p')$$

and since $K_1^{010}(p,x)x$, $K_1(p,x)x$, and $K_1^{020}(p,x)$ are bounded functions, as one can verify from theorems A, B, and theorem 1, the left-hand sides of equations (114)–(116) are bounded functions of (p,p').

Free Propagator Expansion in the Evaluation of the Lamb Shift. II*†

ARTHUR J. LAYZER‡
Columbia University, New York
(Received August 8, 1960)

The results of the preceding paper are extended and applied to the calculation of the new Lamb shift orders of $\alpha w^3 \ln^2 w$ and $\alpha w^3 \ln w$ [$w \equiv (\alpha Z)^2$]. The final result for the $2S-2P_{\frac{1}{2}}$ shift due to the $\ln^2$ and ln terms, which has been previously published, is $\Delta E(2S-2P_{\frac{1}{2}}) = -Lw[\frac{2}{3}\ln^2 w + \ln w(4\ln 2 + 1 + 7/48)]$ where L is Z^4 times the Lamb unit. In megacycles this is -0.25 for H and -9.5 for He^+. The corresponding new values for the total theoretical shift are 1057.70 ± 0.15, 1059.08 ± 0.16, and 14047.2 ± 3.0 for H, D, and He^+, respectively. These values incorporate more up to date estimates for the nuclear finite size effect in D and He^+ than those previously reported.

INTRODUCTION

THIS is a direct continuation of a preceding paper, which we shall refer to as I.[1] In I, order properties, with respect to the parameter $w \equiv (\alpha Z)^2$, were established for the general term of the free-propagator expansion of the bound electron self-energy. In this expansion, the bound electron propagator is expanded in "powers" of the external (Coulomb) potential. The main result of I (theorem 4) asserted that the free-propagator expansion is an expansion in orders of w for what may be called the nonanalytic part of the self-energy. In this paper, we shall be concerned primarily with the application of these results to the calculation of the new Lamb shift orders of $\alpha w^3 \ln^2 w$ and $\alpha w^3 \ln w$ (for hydrogenic atoms).[2] At the same time, however, the general analysis is extended through a discussion of gauge invariance (Sec. 2) and a qualitative consideration of the question of the convergence of the free-propagator expansion (concluding section). The results and notation of I will be used freely in this paper.

The outline of material presented is as follows. In Sec. 1, the sources of $\alpha w^{\frac{5}{2}}$, $\alpha w^3 \ln^2 w$, and $\alpha w^3 \ln w$ terms are studied in preparation for the later systematic calculation of the logarithmic terms.[3] It is shown that the new logarithmic terms arise from a finite and small number of terms of the auxiliary k expansion introduced in I. It turns out that the $\alpha w^{\frac{5}{2}}$ and $\alpha w^3 \ln^2 w$ contributions are associated with the divergence like $w^{-\frac{1}{2}}$ or $\ln w$ in the limit $w \Rightarrow 0$ of certain "standard" matrix elements of simple appearance.

In Sec. 2, we introduce a particular form of gauge-invariance test which proves very helpful in the calculation of the logarithmic terms. The test employs a special gauge transformation called a "w-gauge-transformation" in which E and V are shifted by an amount proportional to the parameter w. We show that the "main result" of I (theorem 4) is invariant to the w-gauge-transformation. In the application of this gauge-invariance test, the favored role of nonanalytic terms in the free-propagator expansion is again manifest. For the $\alpha w^3 \ln w$ coefficient in particular, it is shown that a large number of identities can be written down by inspection connecting the coefficients of matrix elements from I_n with $n \leq 3$ and guaranteeing the w-gauge-invariance of the $\alpha w^3 \ln w$ coefficient.

Section 3 is devoted to a presentation of the method and results of the detailed calculation of the new logarithmic orders. The organization of the calculation is based on the test of w-gauge-invariance outlined in Sec. 2. It is shown that this gauge-invariance test provides a strong check on the accuracy of the entire $\alpha w^3 \ln^2 w$ calculation and also of a portion of the $\alpha w^3 \ln w$ calculation. Special integration techniques are introduced which enable all auxiliary parameters of integration to be eliminated. The matrix elements from the various I_n ($n \leq 3$) then become freely convertible, with the aid of the Schrödinger equation, and may be combined arithmetically. This results in a decided improvement in the conciseness of the calculation. The final result is expressed in terms of a few simple and manifestly gauge-invariant matrix elements which can be easily evaluated for bound states of interest. The work of this section assumes that the bound state under consideration is an S state. The very slight modifications of the results necessary for non-S states are given in the next section (Sec. 4).

In Sec. 5, the new theoretical values of the Lamb shift are presented and compared with experimental values for several bound states of interest. In a concluding section, the question of the convergence of the free propagator expansion is briefly discussed in a qualitative way in the light of previous results.

* This work was supported in part by the U. S. Atomic Energy Commission.

† Based on a dissertation submitted to Columbia University in partial fulfillment of the requirements of doctor of philosophy.

‡ Present address: Institute of Mathematical Sciences, New York University, New York.

[1] A. J. Layzer, J. Math. Phys. (preceding article).

[2] A. J. Layzer, Phys. Rev. Letters 4, 580 1960. The $\ln^2$ term was calculated independently by H. Fried and D. Yennie, Phys. Rev. Letters 4, 583 1960, and by G. Erickson, dissertation, University of Minnesota, 1960 (unpublished).

[3] A detailed calculation of the $w^{\frac{5}{2}}$ term can be found in the author's dissertation.

1. SOURCES OF $w^{\frac{5}{2}}$, $w^3\ln^2 w$, AND $w^3\ln w$ TERMS[4]

In I, the calculation of the lowest-order $w^2\ln w$ coefficient was discussed briefly. We turn now to a consideration of the higher order nonanalytic terms of order $w^{\frac{5}{2}}$, $w^3\ln w$, and $w^3\ln^2 w$. The "main result" of I asserts that the first of these orders arises from terms I_n of the free-propagator expansion with $n\leq 2$, while for the latter two orders we must consider also I_3. We shall first make clear the nature and location of these higher-order terms without entering into the details of the actual calculations. The results of this section, and several of the mathematical devices introduced in their derivation, will greatly facilitate the later systematic calculations of the new logarithmic orders.

To determine the coefficients of the nonanalytic orders in w of interest, it is sufficient to analyze certain matrix elements in which the parameter x or $\sigma=1/x$ is absent. In the case of I_0, where the "rotation trick" has been used, these matrix elements occur directly in the expression (I.111) for $\mathrm{Im}_+ I_0$. For I_n with $n>0$, where the k expansion will be used, we refer to the matrix elements $M_n(w)$ of the following standard form of the differential equation (I.99):

$$\frac{d}{dw}\left(\frac{I_n(w)}{mw^\nu}\right)=-\pi^2 w^{-\nu+2}M_n(w), \tag{1}$$

where

$$M_n(w)=w^{D/2+n-3}\times\int dY\,\text{“}\left\langle u\left|\frac{V^n Q_n{}^\nu(p,\bar{p})}{(1+w\bar{\Delta})^{n+k-q/2}}\right|u\right\rangle\text{”}, \tag{2}$$

with

$$\nu=D/2+s+q+u+1 \tag{3}$$

$$\begin{aligned} D&=2k+r+s\quad ++\\ &=2k+r+s+1+-\\ &=2k+r+s+2--. \end{aligned} \tag{4}$$

Here, ν is the nominal order of the term $I_n{}^\nu$ labeled by k, q, r, s, u. $Q(p,\bar{p})$ is the corresponding numerator polynomial in the p's and $\bar{p}$; D is the degree of this numerator polynomial. For simplicity, we have considered only the generalized Schrödinger approximation and "nontrivial" terms, that is, terms with no external factors of $\epsilon_{\mathrm{sh}}{}^2$ or $E-E_{\mathrm{sh}}$.[5] For Dirac correction (to the large component wave function) terms, one replaces (say) the left wave function $\langle u|$ by the Dirac-correction "wave function" $\langle u'|$ and makes the replacement $D\Rightarrow D+2$ in (2) and (3). The modification for "trivial" terms is obvious.

The separation of the factor $w^{-\nu+2}$ in (1) arranges that $I_n(w)$ is of lowest order w^3 for $\nu\neq 3$ or $w^3\ln w$ for $\nu=3$, modulo cw^ν, provided that $M_n(w)$ approaches a constant in the limit $w\Rightarrow 0$. We shall show later that, except for possible lower-order terms proportional to w^{-1} or w^{-2}, the "spectrum" of $M_n(w)$ is of the form

$$M_n(w)=aw^{-\frac{1}{2}}+b\ln w+c+\text{higher order}. \tag{5}$$

Thus, from (1) for $w^{\frac{5}{2}}$ and $w^3\ln^2 w$ contributions, $M_n(w)$ must contain a $w^{-\frac{1}{2}}$ or $\ln w$ divergence, respectively, in the limit $w\Rightarrow 0$, and for $w^3\ln^2 w$ it is necessary in addition that $\nu=3$. For $w^3\ln w$ terms, the $\ln w$ divergence is necessary for $\nu\neq 3$, but if $\nu=3$, the constant term is sufficient.

Separable Representation for $M_n(w)$

The structure of the $M_n(w)$ of (1) can be simplified and brought closer to that of the matrix elements of $\mathrm{Im}_+ I_0$ by going over to a "separable" or "product" representation as was done in Sec. 2 of I with the aid of identity (I.13) and the single integration parameter λ, for $k=s=0$ terms.

To achieve the separable representation for arbitrary k and s, we shall borrow a mathematical device from Fried and Yennie[6] and introduce a term $\beta\cdot\bar{p}$ into the denominator, where β is an infinitesimal differentiation parameter. The required number of $\bar{p}$'s can be brought into the numerator by repeated applications of the operator ∇_β or $\nabla_\beta{}^2$. The parameter λ is used, similarly to its previous role, to arrange that the denominator is raised to the $n+1$ power before the start of the differentiation.

As an example, consider the matrix elements $M_n{}^{qk}=M_2{}^{qk}$ defined by

$$M_2{}^{qk}=w^{k-1}\int dY_1\,\text{“}\left\langle u\left|\frac{VV(\bar{p})^{2k}}{[1+w\bar{\Delta}]^{k+2-q/2}}\right|u\right\rangle\text{”} \tag{6}$$

$$q=0,\,2.$$

Then, by the method outlined above, we obtain

$$M_2{}^{qk}=C_{qk}\int_1^\infty d\lambda\frac{(\lambda-1)^{k+q/2}}{\lambda^{k+q}}M_2{}^k(\eta) \tag{7}$$

$$M_2{}^k(\eta)\equiv\lim_{\beta\to 0}(V_{\beta^2})^k\frac{1}{\eta}\left\langle u\left|\frac{1}{1+\eta\Delta+\beta\cdot\eta^{\frac{1}{2}}p}\right.\right.$$

$$\left.\left.\times V_0\frac{1}{1+\eta\Delta+\beta\cdot\eta^{\frac{1}{2}}p}V_0\frac{1}{1+\eta\Delta+\beta\cdot\eta^{\frac{1}{2}}p}\right|u\right\rangle, \tag{8}$$

where

$$\eta\equiv w/\lambda \tag{9}$$

and

$$C_{qk}=\left[\int_1^\infty d\lambda\frac{(\lambda-1)^{k+q/2}}{\lambda^{2k+3}}\right]^{-1}\frac{2}{(2k+2)!}.$$

[4] In accordance with the normalization of I, a factor of α is omitted.

[5] Here $\epsilon_{\mathrm{sh}}{}^2\equiv 2(m-E_{\mathrm{sh}})$ where $(m-E_{\mathrm{sh}})$ is the Schrödinger binding energy.

[6] H. M. Fried and D. R. Yennie, Phys. Rev. **112**, 1391 (1958).

By using this construction, every $M_n(w)$ can be written, as in the example, as a convergent integral over λ of the corresponding separable $M_n'(\eta)$ times a polynomial in $1/\lambda$. The nominal order in η of the $M_n(\eta)$ is the same as the nominal order in w of the $M_n(w)$ and the degrees D of the numerator polynomials are also the same.

We observe that an expansion of the $M_n(\eta)$ in orders of η is also an expansion of the $M_n(w)$ in orders of w. The "spectrum" of the $M_n(w)$ is thus fully determined by that of the matrix elements $M_n(\eta)$. The structure of the $M_n(\eta)$ is essentially the same as that of the matrix elements of $\mathrm{Im}_+ I_0$.

Standard Matrix Elements for $w^{\frac{1}{2}}$ and $w^3 \ln^2 w$ Terms

We have seen that a necessary condition for $w^{\frac{1}{2}}$ and $w^3 \ln^2 w$ terms is that the matrix elements $M_n(w)$ contain a $w^{-\frac{1}{2}}$ or $\ln w$ divergence, respectively, as $w \Rightarrow 0$. The same statement holds true with regard to the matrix elements of $w^{-3}\,\mathrm{Im}_+ I_0$ since we have

$$\mathrm{Im}_+ w^{-\frac{1}{2}} = -w^{-\frac{1}{2}} \tag{10}$$

$$\mathrm{Im}_+ \ln^2 w = 2\pi \ln w. \tag{11}$$

We seek, therefore, to recognize and locate matrix elements with such divergences.

As far as recognition is concerned, it turns out to be sufficient to consider a small number of "standard" matrix elements of separable form. These matrix elements are "nontrivial" and in addition have no external powers of w; that is, from (2) $(D/2)+n-3=0$ in the generalized Schrödinger approximation. Also, the denominator factors of $(1+wp^2)$ are reduced to the minimum power necessary for convergence of the matrix elements. All possible matrix elements meeting this description are given below. To save space, we use a dash "—" to indicate a factor $(1+wp^2)^{-1}$. For $n>0$, we really have in mind the variable η defined by (9) rather than w. The equivalences were obtained by using the Schrödinger equation:

$$\begin{aligned}
M_A &= \langle Vp^2 - V\rangle \doteq -\tfrac{1}{2}\langle Vp^4 -\rangle \doteq \tfrac{1}{4}\langle p^6 -\rangle \\
M_B &= \langle V - V^2\rangle \doteq -\tfrac{1}{2}\langle V - Vp^2\rangle \doteq \tfrac{1}{4}\langle p^2 - Vp^2\rangle \\
M_C &= \langle Vp_i - Vp_i\rangle \doteq -\tfrac{1}{2}\langle p^2 p_i - Vp_i\rangle \\
M_D &= \langle p_i p_i - V - p_i p_i\rangle \\
M_E &= \langle u'|\, p^2 - V\,|u\rangle \doteq -\tfrac{1}{2}\langle u'|\, p^4 -\,|u\rangle \\
M_F &= \langle p_i V^2 p_i\rangle \\
M_G &= \langle u'|\, V^2\,|u\rangle \doteq -\tfrac{1}{2}\langle u'|\, Vp^2\,|u\rangle \\
M_H &= \langle u'|\, p_i V p_i\,|u\rangle.
\end{aligned} \tag{12}$$

These forms arise for $n=0$, 1, 2, and 3. For $n>0$, the index n is determined by the number of numerator V's. Equivalences are with respect to the coefficients of $w^{-\frac{1}{2}}$ and $\ln w$. Terms with Dirac corrections to the small component wave functions or higher-order Dirac corrections to the wave functions are finite in the limit $w \Rightarrow 0$ and need not be considered.

In M_F and M_H, the p_i are to be interpreted as differential operators acting toward the wave functions. M_F, M_G, and M_H are finite in the limit $w \Rightarrow 0$ and therefore need not be further considered.

For the evaluation of the leading behavior of the matrix elements M_A to M_E, it is convenient to define the quantity $\psi(0)$ proportional to the Schrödinger position-space wave function at the origin:

$$\psi(0) \equiv \int d^3p \langle u_s|\, p\rangle. \tag{13}$$

The matrix element M_A diverges like $w^{-\frac{1}{2}}$ in the limit $w \Rightarrow 0$ with a coefficient proportional to $\psi^2(0)$. All the other matrix elements, M_B to M_E, diverge like $\ln w$ in the limit $w \Rightarrow 0$ with coefficients again proportional to $\psi^2(0)$.[7] It is not difficult to derive the following lowest-order estimates for these matrix elements:

$$M_A(w) = (1/2\pi^2) w^{-\frac{1}{2}} \psi^2(0) + \text{const} + \text{higher order} \tag{14}$$

$$\begin{aligned}
M_B = -M_C = \tfrac{1}{2} M_D = 2M_E \\
= (1/4\pi^2)\psi^2(0) \ln w + \text{higher order}.
\end{aligned} \tag{15}$$

The equivalence $M_C = -2M_E$ can be demonstrated by using the differential equation satisfied by u' (for S states):

$$\begin{aligned}
2 \cdot H_0 u_s' &= 2[(\Delta/2) + V_0] u_s' \\
&= (m-E) V_0 u_s - \tfrac{1}{2} p_i V_0 p_i u_s - \delta u_s,
\end{aligned} \tag{16}$$

where

$$\delta = \frac{1}{w}(\epsilon^2 - \epsilon_{\mathrm{sh}}^2) = \frac{1}{w}\left[\frac{m^2 - E^2}{m^2 w} - 2\frac{(m - E_{\mathrm{sh}})}{mw}\right]. \tag{17}$$

"Spectrum" of $M_n(w)$ for $n>0$

In spite of the infinite number of matrix elements introduced by the k expansion, the standard matrix elements M_A to M_E are the only ones that need be considered as far as recognizing $\ln w$ or $w^{-\frac{1}{2}}$ divergences are concerned, as we have previously stated. The reason is that the $\ln w$ part disappears as soon as the degree of the numerator polynomial is large enough that the external factor of w for the matrix element has a positive power. Furthermore, the parts of the matrix elements yielding $w^{-\frac{1}{2}}$ divergences can be isolated and are then seen to be derivable in a certain sense from the matrix element M_A; in particular, they have the same "spectrum."

These assertions follow from the statement and proof of theorems A and A' below. The most important consequence of these theorems, aside from the verification of the form (5), is that $w^3 \ln^2 w$ and also $w^2 \ln w$

[7] For the $2-V$ form of M_A and the $3-V$ form of M_B, the $\psi^2(0)$ factor can be separated off "automatically" by performing the transformation $p_i \Rightarrow w^{-\frac{1}{2}} p_i$ for $i=1$ to $n-1$ and then taking the limit $w \Rightarrow 0$.

terms arise from only a finite and small number of terms of the k expansion.

In the derivation of theorems A and A′ as well as the later systematic calculations, a good deal of use is made of the simple operation wD_w, where $D_w \equiv d/dw$. This operator applied to an arbitrary power of w, say w^α, leaves its form unchanged and in this sense is "spectrum preserving." However, it has the additional useful property that it singles out and removes $\ln w$ and "constant" terms since $wD_w \ln w = 1$ and wD_w constant $= 0$ and $(wD_w)^n \ln w = 0$ for $n > 2$. Furthermore, the operator wD_w applied to $\ln w$ and a "higher order" term $f(w)$ leaves $f(w)$ of "higher order" (than unity) as is easily shown with the aid of L'Hospital's rule provided suitable differentiability requirements are imposed on $f(w)$.

We are now ready to derive the theorems of main interest in the present section which are stated below.

Theorem A. Consider the matrix elements $M_n(w)$ of (1). If $D > D_n$, where $D_n = 0$, 2, and 4 for $n = 3$, 2, and 1, respectively, then $M_n(w)$ has no $\ln w$ part and the "spectrum" of $M_n(w)$ has the form

$$n=3: \quad M_n(w) = \text{const} + \text{higher order}$$
$$n=1,\ 2: \quad M_n(w) = aw^{-\frac{1}{2}} + \text{const} + \text{higher order}.$$

Theorem A′. Consider the Dirac correction matrix elements $M_n'(w)$. If $D > D_n$ where $D_n = 0$, 2 for $n = 2$ and 1, respectively, then $M_n'(w)$ has no $\ln w$ part and its "spectrum" has the form

$$M_n'(w) = \text{const} + \text{higher order}.$$

We consider first the proof of theorem A. For I_3 the proof is based on the boundness properties of the various $K_2^\nu(p,p')$ of I together with dominating arguments in p space. We assume here and in the remainder of the proof that a separable representation is employed. The λ integration is always well behaved and we restrict attention to the matrix elements $M_n(\eta)$ with $\eta = w/\lambda$.

Consider now $n = 2$ and 1. To save space, we use the abbreviation "**m**," a boldface m, to denote the factor $1/(1+\eta p^2)^m$. The proof proceeds according to one of the two following programs:

(1) Whenever this is possible, "synthesize" the $n = 3$ situation by domination of the numerator p's in p space and domination of $u(p)$ by the $1-S$ wave function $u_{1-S}(p)$ and the use of the Schrödinger equation. Then the proof for $n = 3$ applies. As an example, for $n = 2$, by using $|p_0 \cdot p_2| < \frac{1}{2}(p_0^2 + p_2^2)$, we obtain

$$\eta\langle u|\, p_i - Vp^2 2Vp_i - |u\rangle < \eta < u_{1-S}|\, -Vp^2 - Vp^2 - |u_{1-S}\rangle$$
$$= -2\eta\langle u_{1-S}|\, -Vp^2 - V - V\,|u_{1-S}\rangle$$
$$= +2\int d^3p_0 d^3p_3 u_{1-S}(p_0) u_{1-S}(p_3) K_2^{0200}(w^{\frac{1}{2}}p_0, w^{\frac{1}{2}}p_3)$$
$$\Rightarrow +2\psi^2(0) K_2^{0200}(0,0) = \text{finite}.$$

(2) For certain terms, it will not be possible to "synthesize" the $n = 3$ situation. As a matter of fact, these terms must be of one of the following types

$n = 2$:
$$M_A{}^k = \eta^k\langle u|\, \mathbf{r}V\mathbf{k}' + \mathbf{1}(p^2)^{k+1}V\mathbf{s}\,|u\rangle$$
$$M_B{}^k = \eta^k\langle u|\, \mathbf{r}V\mathbf{k}'(p^2)^k p_i V p_i \mathbf{s}\,|u\rangle,$$

$n = 1$:
$$M_C{}^k = \eta^k\langle u|\, \mathbf{k}' + \mathbf{1}(p^2)^{k+2}V\mathbf{s}\,|u\rangle$$
$$M_D{}^k = \eta^k\langle u|\, \mathbf{k}'(p^2)^{k+1} p_i V p_i \mathbf{s}\,|u\rangle,$$

where $k \geq 1$, $k' > k$.

By using the Schrödinger equation to the left, C and D, with $n = 1$, go over to types A and B, respectively, and therefore it is sufficient to consider A and B of the $n = 2$ situation. In A and B we expand the end factors of $1/(1+\eta p^2)$ using the identity

$$\frac{1}{(1+x)^n} = 1 - x\sum_{k=1}^{n} \frac{1}{(1+x)^k}$$

and then "synthesize" the $n = 3$ situation for the remainder terms by using the Schrödinger equation. For the remainder terms, the proof for $n = 3$ applies.

For the lowest-order term of this expansion, we will have

$$M_A{}^k \Rightarrow \eta^k\langle u|\, V\mathbf{k}' + \mathbf{1}(p^2)^{k+1}V\,|u\rangle$$
$$M_B{}^k \Rightarrow \eta^k\langle u|\, V\mathbf{k}'(p^2)^k p_i V p_i\,|u\rangle.$$

These forms for $M_A{}^k$ and $M_B{}^k$ can be expressed as linear combinations of $(\eta D_\eta)^m M'(\eta)$ with $m \geq 1$, where

$$M_A' = \langle u|\, Vp^2 - V\,|u\rangle$$
$$M_B' = \langle u|\, V - p_i V p_i\,|u\rangle.$$

These matrix elements are identical with the matrix elements M_A and M_C of (12) and we know, therefore, that M_A' diverges like $\eta^{-\frac{1}{2}}$ but has no $\ln\eta$ part, and M_B' diverges like $\ln\eta$. It follows from the properties of the operator ηD_η discussed earlier that the $M_B{}^k$ are finite while the $M_A{}^k$ diverge like $\eta^{-\frac{1}{2}}$ in the limit $\eta \Rightarrow 0$ but contain no $\ln\eta$ part (and, in fact, no "constant" part either).

This completes the outline of the proof of theorem A.

The proof of theorem A′ is straightforward. One uses the asymptotic behavior at high p of $u'(p)$, which is like $p\,u(p)$.

2. GAUGE INVARIANCE

We know that $\Delta E = \sum_n I_n$ remains invariant under the transformation

$$V \Rightarrow V + a; \quad E \Rightarrow E + a, \tag{18}$$

where a is an arbitrary constant. This invariance property has nothing to do with the wave functions. It is an attribute of the propagator $1/\mathbf{p} - \mathbf{k} - \mathbf{V} + m$

which is part of the operator whose expectation value is ΔE. The individual terms I_n of the free-propagator expansion are, however, not gauge invariant and for this reason a gauge-invariance test provides a useful check on the consistency and accuracy of the calculation.

The full gauge invariance of the propagator is expressed by the fact that it is a function of $(p_\mu - V_\mu)$. Thus, it is invariant to simultaneous and identical formal "displacements" of p_μ and V_μ by an arbitrary operator 0_μ: $p_\mu \Rightarrow p_\mu + 0_\mu$; $V_\mu \Rightarrow V_\mu + 0_\mu$. The usefulness of a gauge-invariance test depends on a suitable specialization of the operator 0_μ.

A natural limitation to impose on 0_μ is that it be a constant "four vector." In this way, the displacement will not interfere with the free-propagator expansion. In the present problem (Lamb shift) it is not convenient to carry the "space" part of this displacement since the actual potential (in the usual gauge) has no space part. For this reason we restrict the following considerations to displacements of the type (18).

An especially convenient and practical form of gauge-invariance test results if we further limit the gauge transformations to those of the type

$$\begin{aligned} V &\Rightarrow V + gwm \\ E &\Rightarrow E + gwm, \end{aligned} \tag{19}$$

where g is a parameter independent of w. We call this special gauge transformation a "w-gauge transformation." In terms of the scaled potential V_0 and the scaled *Schrödinger* binding energy, $\frac{1}{2}\epsilon_{sh}^2$, this transformation is independent of w:

$$\begin{aligned} V_0 &\Rightarrow V_0 + g \equiv V_0(g) \\ \tfrac{1}{2}\epsilon_{sh}^2 = (m - E_{sh})/mw &\Rightarrow \tfrac{1}{2}\epsilon_{sh}^2(g) = \tfrac{1}{2}\epsilon_{sh}^2 - g. \end{aligned} \tag{20}$$

This has the consequence that the formal scaling properties of a term I_n^ν of I_n are not affected by the w-gauge-transformation $I_n^\nu \Rightarrow I_n^\nu(g)$. One can then use the same formal process of reduction for $I_n^\nu(g)$ as for I_n^ν. In particular, one can employ the usual differentiation device in the form

$$\frac{d}{dw}\left(\frac{I_n^\nu(g)}{mw^\nu}\right) = w^{-\nu+2}M(w;g), \tag{21}$$

where $M(w;g)$ differs from $M(w)$ in that $V_0 \Rightarrow V_0(g)$ in the numerator and $\epsilon_{sh}^2 \Rightarrow \epsilon_{sh}^2(g)$ in the denominator and numerator. This has the desirable result that the gauge-invariance test can be applied at a late stage of the calculation.

We list below seven assertions which summarize the principal results about w-gauge invariance in connection with calculation of nonanalytic terms of interest with particular reference to the $w^3 \ln w$ calculation.

(1) The "main result" of I, Sec. 3, for $n>3$ is invariant to the w-gauge transformation provided that $|g|$ is suitably limited. A suitable limitation is

$$|g| < \tfrac{1}{2}\epsilon_{sh}^2 = 1/(2n^2)$$

for $n-S$ state.

(2) The w^5 and $w^3 \ln^2 w$ contributions to each I_n are w-gauge invariant.

(3) $I_n(w^3 \ln w; g) - I_n(w^3 \ln w) = w^3 \ln w(a_n^1 g + a_n^2 g^2 + a_n^3 g^3)$ where $a_n^i = 0$ for $n>3$. (More generally, $I_n(g) - I_n$ is a polynomial in g for orders nonanalytic in w but an infinite series for analytic orders.)

(4) We must have $\sum_{n=0}^{3} a_n^i = 0$ for $i = 1, 2$, and 3.

(5) The coefficients a_n^i are functionals of V. We can then classify the a_n^i according to the number of V's involved in the numerator of the corresponding "matrix element." This classification applies before the use of the Dirac or Schrödinger equation. The symbol $a_n^{i\nu}$ indicates that the matrix elements corresponding to this coefficient have ν V's. Then, we have

$$\begin{aligned} \nu &\leq n \\ i + \nu &\leq 3. \end{aligned} \tag{a}$$

Furthermore,

$$\sum_{n=0}^{3} a_n^{i\nu} = 0. \tag{b}$$

(6) The "mechanism" of the cancellation expressed by (b) is given by the following identities, all of which can be derived "by inspection." For g^3:

$$\begin{aligned} I_3(g^3) &= -I_0(g^3) \equiv A \\ I_1(g^3) &= -I_2(g^3) \equiv B. \end{aligned}$$

Furthermore, $B = 3A$

For g^2:	(a)	$1V$:	$I_1 = -\frac{1}{2}I_2 = I_3$
	(b)	$0V$'s:	$I_0 = -\frac{1}{2}I_1 = I_2$
For g:	(a)	$2V$'s:	$I_2 = -I_3$
	(b)	$1V$:	$I_1 = -I_2$
	(c)	$0V$'s:	$I_0 = -I_1$.

(7) Provided that the Dirac or Schrödinger equation has not been used in the reduction of matrix elements, assertion 5(b) and the identities of assertion 6 must hold separately for $++$, $+-$, and $--$ and for the generalized Schrödinger approximation and Dirac corrections.

A brief discussion of these assertions is given below. Details are easily filled in.

Assertion 1. The validity of this assertion is easy to see. The limitation on g is necessary in order that $\epsilon^2(g) > 0$ where $\epsilon^2 \equiv (m^2 - E^2)/m^2 w$, since for $\epsilon^2(g) = 0$ the integrals I_n will diverge, in general. One can show that the restriction on g given insures that $\epsilon^2(g) > 0$ for $|w| < 1$.

Consider now the effect of the replacement $V_0 \Rightarrow V_0(g)$. If one or more of the V_0's are replaced by a g, it is easy to show first of all that the existence of the integrals I_n is not destroyed, and secondly, that

the lowest order of the nonanalytic part is not lowered. In fact, it is always raised for $n>3$.

Assertion 2. The appearance of $w^{\frac{5}{2}}$ and $w^3\ln^2 w$ terms is associated with the divergence like $w^{-\frac{1}{2}}$ or $\ln w$, respectively, of certain matrix elements with no external factors of w or ϵ_{sh}^2. In every case, one finds that the replacement of a V by a g removes this divergence and, therefore, the $w^{-\frac{1}{2}}$ or $\ln w$ parts of these matrix elements are w-gauge invariant.

Assertion 3. In view of assertion 1, it is sufficient to consider I_n with $n<3$. Since the number of numerator V's is then finite and ≤ 3, the validity of the present assertion clearly depends on the effect on the coefficient of $w^3\ln w$ of the g dependent part of the denominator $[1+\sigma\bar{\Delta}(g)]^{n+k-q/2}$. To investigate this point, it is convenient to expand the inverse of this denominator in powers of $\sigma[\epsilon^2-\epsilon^2(g)]/(1+\sigma\bar{\Delta})$, an expansion which is permitted according to the inequality $|\epsilon^2-\epsilon^2(g)|<\epsilon^2$ which follows from the restriction on g mentioned in the statement of assertion 1. Due to the appearance of σ in the numerator of this expansion, successive terms of the expansion will have lowest order nonanalytic in w parts that are of successively higher order in w. From this fact, the truth of assertion 3 follows immediately.

It is clear from the method of proof that the result that $I_n(g)-I_n$ is a polynomial in g rather than an infinite series holds generally for nonanalytic in w terms but not for analytic orders such as w^2.

Assertion 4. The requirement of w-gauge invariance combined with assertion 3 says that

$$\sum_{i=1}^{3}\sum_{n=p}^{3} a_n{}^i g^i = 0.$$

Since, with the restriction on g given in assertion 1, this must hold for all g in a certain fixed (w independent) neighborhood of $g=0$, it follows that the coefficients of the g^i are separately zero.

Assertions 5 and 7. The proof is straightforward. The important point is that the gauge invariance is a property of the propagator alone. For assertion 5, we can make the formal replacement $V\Rightarrow\lambda V$ in the propagator. Then the coefficients $a_n{}^i$ in assertion 4 will be polynomials in λ.

Assertion 6. The trick here is to algebraically expand the g's out of the propagators to begin with and to use the restrictions of 5(a). Then the cancellations become obvious and appear as "structural identities."

3. $w^3\ln^2 w$ AND $w^3\ln w$ CALCULATIONS

The sources of $w^3\ln^2 w$ and $w^3\ln w$ terms were considered in Sec. 1. The relevance of gauge invariance to the $w^3\ln w$ calculation and a practical test of gauge invariance were discussed in the preceding section.

In this section, we shall present the "systematics" and results of the calculation of the $w^3\ln^2 w$ and $w^3\ln w$ coefficients. These terms will be calculated together since, as will be made clear, the relatively short $w^3\ln^2 w$ calculation is in a certain sense embedded in the much longer $w^3\ln w$ calculation. In particular, a gauge-invariance test applied to the $w^3\ln w$ terms automatically "covers" the gauge-invariant $w^3\ln^2 w$ contributions from each I_n.

We will begin the calculation of $w^3\ln^2 w$ and $w^3\ln w$ coefficients by considering the contributions due to I_0. In a way it would be more appropriate to start with I_3 since the separation of the $\psi^2(0)$ factor for $w^3\ln^2 w$ terms, without reference to the form of the wave functions, occurs "naturally" only for I_3. This is analogous to the case of the $w^{\frac{5}{2}}$ coefficient which occurs naturally in the above sense only for $n=2$.

There are two good reasons, though, for starting with I_0. In the first place, it will turn out that the entire $w^3\ln^2 w$ coefficient is due to I_0, a statement which is w-gauge invariant. The $w^3\ln w$ parts of I_1, I_2, and I_3 will be shown to cancel.

In the second place, the method that we have chosen to use for I_0, namely that of considering Im_+, is a concise alternative method to the k expansion used for $n>0$ and, therefore, the I_0 calculation can be presented independently and in detail. Incidentally, as a check the k expansion was tried for I_0 also and the final results obtained were the same as those presented here.

I_0 Calculation

An expression for $\mathrm{Im}_+ I_0$ is given in (I.111). The original of the first term of (I.111) is given in (I.113). The first term of (I.113) cancels against a similar term in I_1 and will be retained intact. The second term contributes to $w^3\ln w$ through large-large Dirac corrections and small-small Schrödinger. We obtain

$$\begin{aligned} I_0 &\text{ (first term of I.111)} \\ &= m\pi^2 w^3\ln w \text{ times } 4/w^2\ln w\langle v_s|V|v_s\rangle \\ &\quad +4\langle u'|V^2|u\rangle-\tfrac{1}{2}\langle u|p_iV^2p_i|u\rangle. \end{aligned} \tag{22}$$

Consider now the last three terms of (I.111) in the generalized Schrödinger approximation. After a straightforward algebraic reduction involving the use of the Schrödinger equation for the V_0 term, one obtains the following result. (Lower-order contributions of order w^2 have been dropped.)

$$\begin{aligned} &\mathrm{Im}_+ I_0 \text{ (Schrödinger; last 3 terms of I.111)} \\ &= m\pi^2 w^3 \text{ times } [5/4-\tfrac{1}{4}wD_w]\langle u|\Delta^3/1+w\Delta|u\rangle \\ &\quad +[1+\tfrac{1}{2}wD_w]\langle u|V(p_i/1+w\Delta)Vp_i|u\rangle \\ &\quad +\tfrac{1}{4}\epsilon^2\langle u|\Delta^2|u\rangle-6\delta\langle u|V|u\rangle, \end{aligned} \tag{23}$$

where the quantity δ is defined by the relation (17).

Let us now consider the Dirac correction part of the last three terms in (I.111). The Dirac corrections to small-small are of higher order than w^3 and, therefore, have an original of higher order than $w^3\ln w$. The Dirac corrections to large-large yield the following

expression, after some consolidation.

$$\mathrm{Im}_{+}I_0 \text{ (Dirac; last three terms of I.111)} = m\pi^3 w^3 \text{ times } [3-\tfrac{3}{2}wD_w]\langle u'|\Delta^2/1+w\Delta|u\rangle+4\langle u'|V^2|u\rangle. \quad (24)$$

To obtain the total $w^3 \ln w$ contribution, we must add the originals of (23) and (24) to (22). In order to combine gauge-variant terms with I_1, I_2, and I_3, it is convenient to use the following easily proved equivalence relative to $w^3 \ln^2 w$ and $w^3 \ln w$ but *not* $w^{\frac{1}{2}}$.

$$\text{Original} \quad \pi w^3 M(w) \doteq -w^3 \int_{w_0}^{w} dw(1/w)M(w) \quad (25)$$

where $M(w)$ is a matrix element with a "spectrum" of the form

$$M(w) = aw^{-\frac{1}{2}}+b \ln w+c+\text{higher order}. \quad (26)$$

By using the equivalence (25) we obtain the final result

$$I_0(w^3 \ln^2 w; w^3 \ln w) = I_0(\text{gauge-var})+I_0[\psi^2(0)] \quad (27)$$

$$I_0[\psi^2(0)] = m\pi^2 w^3 \ln w \psi^2(0)/4\pi^2[-1], \quad (28)$$

$$I_0(\text{gauge-var}) = m\pi^2 w^3 \text{ times } 4/w^2\langle v|V|v\rangle \ln w$$
$$+6\delta \ln w\langle u|V|u\rangle-(5/4)\int^{w}(dw/w)$$
$$\times\{\langle u|\Delta^3/1+w\Delta|u\rangle-\langle u|p_iV(p_i/1+w\Delta)V|u\rangle$$
$$-3\langle u'|\Delta^2/1+w\Delta|u\rangle\}-\tfrac{1}{2}\ln w\langle u|p_iV^2p_i|u\rangle$$
$$-\epsilon^2 \ln w\langle u|V^2|u\rangle. \quad (29)$$

Note that the $\psi^2(0)$ part arises from the wD_w operation acting on the $\ln w$ parts of certain matrix elements and is, therefore, w-gauge invariant.

The $w^3 \ln^2 w$ part comes from the Dirac and $\langle p_iVp_iV\rangle$ terms of the gauge-variant part. From (29) and the indefinite integral $\int \ln w/w = \frac{1}{2}\ln^2 w$, we find[8]

$$I_0(w^3 \ln^2 w) = mw^3 \ln^2 w[+\tfrac{1}{2}]. \quad (30)$$

I_3, I_2, and I_1 Calculations

The results of Sec. 1 concerning the sources of $w^3 \ln^2 w$ and $w^3 \ln w$ terms from $n=1$, 2, and 3 may be summarized in the following statements. The notation and terminology is that of Sec. 1.

$w^3 \ln^2 w$ terms arise from "nontrivial" terms which satisfy the condition

$$\nu = 3 \qquad (*)$$

and which, in addition, have corresponding matrix elements $M_n(w)$ with $\ln w$ parts. The latter condition restricts the degree D of the numerator polynomial in the p's to values satisfying $D \leq D_n$, where for the generalized Schrödinger approximation $D_n = 6-2n$. $M_n(w)$ with Dirac correction wave functions have $\ln w$ parts only for $n=1$ and $++$ with $D_n = D_1 = 2$.

For brevity, we shall call terms satisfying the condition (*) above ($\nu=3$) "star terms."

$w^3 \ln w$ terms arise, in addition to the above sources for $w^3 \ln^2 w$ from "star terms" with $D > D_n$ and from "nonstar" terms which have $M_n(w)$ with $\ln w$ parts and therefore with $D \leq D_n$.

Nonstar Terms

The simplest category of $w^3 \ln w$ terms to discuss is that of nonstar terms. Let us define the quantities A_n^{ν} by the relation

$$\sum M_n^{\nu}(w; \ln w) = A_n^{\nu} \ln w. \quad (31)$$

$M(w; \ln w)$ denotes the $\ln w$ part of $M(w)$. The sum is over all terms with a fixed value of ν and n. Note that A_n is proportional to $\psi^2(0)$. It follows from (31) and (1) that

$$I_n(\text{non}*; w^3 \ln w) = m\pi^2 w^3 \ln w \sum_{\nu} (A_n^{\nu}/\nu-3) \quad (32)$$
$$\nu = 1, 2, 4, \text{ and } 5.$$

ν cannot exceed 5 because of the restriction $D \leq D_n$ and the obvious condition $q+s+u \leq n+1$.

The evaluation of the coefficients A_n^{ν} is simple. The "standard" matrix elements yielding $\ln w$ parts are the matrix elements M_B to M_E listed in (12). The $\ln w$ coefficients are given in (15). The exact number of denominator factors does not affect the $\ln w$ coefficient. When $D = D_n$, one can, therefore, discard the denominators and perform the simple y integration for the numerator. The coefficient of $\ln w$ is then computed as a linear combination of the coefficients for the matrix elements M_B to M_E. For $D < D_n$, the procedure is similar but it is necessary to first expand out the lower-order contributions from the matrix element. It should be noted that this introduces a dependence on the power $n+k-q/2$ of the denominator factor $(1+w\bar{\Delta})$.

Elimination of Auxiliary Parameters; Definition of $*M$

In the systematic calculation of star terms, it proves very convenient to use a "separable" representation. In Sec. 1, the separable representation was achieved through the introduction of the single parameter λ. We will now show that as far as $w^3 \ln^2 w$ and $w^3 \ln w$ terms are concerned, even the parameter λ can be dispensed with.

A single example will suffice to illustrate the general method. Consider the matrix element for $n=3$ and $k=1$

$$M_3^{\nu}(w) = M(w) = w\int dY \text{ “}\left\langle u\left|\frac{V^3\bar{p}^2}{[1+w\bar{\Delta}]^4}\right|u\right\rangle\text{”} \quad (33)$$

[8] As stated in the introductory remarks to this section, we will show that this contribution from I_0 is the entire contribution of order $w^3 \ln^2 w$ to the Lamb shift. This value of the $w^3 \ln^2 w$ coefficient was announced at the Washington meeting of the APS, May, 1959.

Let the matrix element $*M$ be defined in the same way except for the power of the denominator $(1+w\bar{\Delta})$, which is raised from 4 to 6:

$$*M(w)=w\int dY\,\Big\langle u\Big|\frac{V^3\bar{p}^2}{(1+w\bar{\Delta})^6}\Big|u\Big\rangle''. \tag{34}$$

Then $*M(w)$ can be written in separable form without introducing the parameter λ. For we have

$$\frac{w\bar{p}^2}{[1+w\bar{\Delta}]^6}=\frac{1}{4.5}\lim_{\beta\to 0}\nabla_{\beta^2}\frac{1}{[1+w\bar{\Delta}+\beta\cdot w^{\frac{1}{2}}\bar{p}]^4}$$

and

$$\int\frac{dY}{[1+w\bar{\Delta}+\beta\cdot w^{\frac{1}{2}}\bar{p}]^4}=\prod_{i=v}^{3}\frac{1}{[1+w\bar{\Delta}_i+\beta\cdot w^{\frac{1}{2}}\mathbf{p}_i]}.$$

Now we can write $*M(w)$ in terms of $M(w)$ since

$$\frac{1}{[1+w\bar{\Delta}]^6}=(1+\tfrac{1}{5}wD_w)(1+\tfrac{1}{4}wD_w)\frac{1}{[1+w\bar{\Delta}]^4}$$

$$=\Big[1+\frac{9}{20}wD_w+\frac{1}{20}(wD_w)^2\Big]\frac{1}{[1+w\bar{\Delta}]^4}, \tag{35}$$

and we have also the commutation relations

$$\begin{aligned}[w,wD_w]&=-w\\ [w,(wD_w)^2]&=-2wD_ww+w.\end{aligned} \tag{36}$$

From (33), (34), (35), and (36) it follows that

$$*M(w)=[\tfrac{3}{5}+bwD_w+c(wD_w)^2]M(w). \tag{37}$$

Since $M(w)$ has no $\ln w$ part, we can restrict ourselves to the "constant" terms in (37):

$$*M(w)=\tfrac{3}{5}M(w); \tag{38}$$

or,

$$M(w)=5/3*M. \tag{39}$$

Thus, we have expressed the original matrix element $M(w)$ as a multiple of the matrix elements $*M(w)$ which can be written in separable form without the introduction of the auxiliary parameter λ.

In the following discussion $*M_n(w)$ will always denote $M_n(w)$ modified so that the denominator $(1+w\bar{\Delta})$ is raised to the right power to give a separable representation without the introduction of the parameter λ. M is expressed in terms of $*M$ by using the formula

$$\frac{1}{[1+w\bar{\Delta}]^{m+k}}=\prod_{j=m}^{m+k-1}\Big[1+\frac{wD_w}{j}\Big]\frac{1}{[1+w\bar{\Delta}]^m} \tag{40}$$

and the appropriate commutation relations between powers of w and wD_w, of which (36) is a particular example. This yields a relation of the type (37) giving $*M$ in terms of M which is then "inverted" to give M in terms of $*M$.

If $M(w)$ contains an $\ln w$ part, the situation is slightly more complicated than in the above example. In this case, the wD_w part of the relation between M and $*M$ must be retained. Thus, we have in general

$$\begin{aligned}*M&=[a+bwD_w+c(wD_w)^2+\cdots\text{const}(wD_w)^m]M\\ M&=a^{-1}[1-b/awD_w]*M.\end{aligned} \tag{41}$$

The equivalence is with respect to the $\ln w$ and "constant" terms only (not $w^{-\frac{1}{2}}$ terms, for example).

For wD_w part, we need pick up only the $\ln w$ part of $*M$ and, therefore, the wD_w part is always proportional to $\psi^2(0)$.

Star Terms with $D>D_n$

We consider first star terms with $D>D_n$. This case arises for $n=3$ and 2 only, not for $n=1$. These terms contribute to $w^3\ln w$ only and, as we shall show, they are all proportional to $\psi^2(0)$.

In the case of I_3, p_0 and p_3 in the numerator and in the denominator factor $(1+w\bar{\Delta})$ may be set equal to zero.[9] The corresponding separable $*M$ have the following form

$$X_{mn}(w)=w^2\,\Big\langle u\Big|\frac{(-V)^3Q_x(p_1,p_2)}{(1+wp_1^2)^m(1+wp_2)^n}\Big|u\Big\rangle''\Rightarrow$$

$$X_{mn}\psi^2(0) \tag{42}$$

$$Y_{mn}(w)=w\,\Big\langle u\Big|\frac{(-V)^3Q_y(p_1,p_2)}{(1+wp_1^2)^m(1+wp_2^2)^n}\Big|u\Big\rangle''\Rightarrow$$

$$Y_{mn}\psi^2(0) \tag{43}$$

$Z_{mn}=X_{mn}$ or Y_{mn}

$$\equiv\Big\langle 0\Big|\frac{(-V)^3Q(_zp_1,p_2)}{(1+p_1^2)^m(1+p_2^2)^n}\Big|0\Big\rangle'', \tag{44}$$

where for an arbitrary operator R

$$\langle 0|R|0\rangle=\lim_{p,p'\to 0}\langle p|R|p'\rangle. \tag{45}$$

The Q_x are quartic and the Q_y quadratic scalar polynomials in the vectors p_1, p_2. We list below all possible such polynomials with corresponding letter symbols introduced for convenience in later reference. We have taken into account "left-right" symmetry.

$$\begin{aligned}&D:Q_y=p_1p_2 \quad G:Q_x=(p_1p_2)^2 \quad I:Q_x=p_1^4\\ &F:Q_y=p_1^2 \quad H:Q_x=(p_1p_2)p_1^2 \quad J:Q_x=p_1^2p_2^2.\end{aligned} \tag{46}$$

By making use of the identity

$$p_1\cdot p_2=-\tfrac{1}{2}q^2+\tfrac{1}{2}(p_1^2+p_2^2)$$

where $q=p_1-p_2$ and the symmetry of the S-state wave function, we can use instead of the above integrals,

[9] For the justification of this statement, compare the concluding remarks to Sec. 3 of I.

the set

$$Y_{mn}=E_{mn},F_{mn} \quad X_{mn}=I_{mn},J_{mn},K_{mn} \tag{47}$$

where the integrals E_{mn}, K_{mn} are defined by

$$E:Q_y=q^2 \quad K:Q_x=p_1^2q^2. \tag{48}$$

The necessary relations are

$$\begin{aligned} D_{mn}&=-\tfrac{1}{2}E_{mn}+\tfrac{1}{2}(F_{mn}+F_{nm})\\ H_{mn}&=-\tfrac{1}{2}K_{mn}+\tfrac{1}{2}(I_{mn}+J_{mn})\\ G_{mn}&=-\tfrac{1}{4}(K_{mn}+K_{nm})+\tfrac{1}{2}J_{mn}+\tfrac{1}{4}(I_{mn}+I_{nm}). \end{aligned} \tag{49}$$

The quartic set X_{mn} can now be expressed in terms of the quadratic set $Y_{mn}=E_{mn}$, F_{mn} through the following relations, obvious from (44):

$$\begin{aligned} I_{mn}&=F_{mn}-F_{m-1,n};\\ J_{mn}&=F_{m,n-1}-F_{mn};\\ K_{mn}&=E_{m-1,n}-E_{mn}. \end{aligned} \tag{50}$$

The limit integrals E_{mn} and F_{mn} are given explicitly below.

$$E_{mn}=-\text{“}\left\langle 0\left|V\frac{1}{(1+p^2)^m}q^2V\frac{1}{(1+p^2)^n}V\right|0\right\rangle\text{”} \tag{51}$$

$$F_{mn}=-\left\langle 0\left|V\frac{p^2}{(1+p^2)^m}V\frac{1}{(1+p^2)^n}V\right|0\right\rangle. \tag{52}$$

By using the fact that $wD_wY_{mn}(w)\Rightarrow 0$ as $w\Rightarrow 0$ together with the defining relation (43) and formula (48), one can easily derive the following difference equation for $Y_{mn}=E_{mn}$, F_{mn}.

$$Y_{mn}=\frac{mY_{m+1,n}+nY_{m,n+1}}{m+n-1}. \tag{53}$$

A knowledge of Y_{m1} for all m is thus sufficient for the calculation successively of all Y_{m2}, all Y_{m3}, etc. and, therefore, all Y_{mn} through formula (53).

The evaluation of the integrals E_{mn}, F_{mn} is simple.[10] The following results are obtained.

$$E_{mn}=\frac{1}{2\pi^2}\frac{(2m-3)!!(2n-3)!!}{(m-1)!(n-1)!2^{m-1}2^{n-1}} \tag{54}$$

$(-1)!!\equiv 1$

$$F_{mn}=\frac{1}{2\pi^2}\left[\delta_{m1}\ln 2+\frac{(-1)^{m-1}(-1)^{n-1}}{(m-1)!(n-1)!}\times\lim_{y,z\to 1}D_y^{n-1}D_z^{m-1}\frac{I(y,z)}{y}\right] \tag{55}$$

$$I(y,z)\equiv\ln\left[\frac{(z^{\frac{1}{2}}+y^{\frac{1}{2}})}{2z^{\frac{1}{2}}}\right]. \tag{56}$$

[10] We used tables of sine, cosine, and Laplace transforms given by A. Erdelyi, editor, Bateman Manuscript Project, *Tables of Integral Transforms* (McGraw-Hill Book Company, Inc., New York, 1954), Vol. 1.

Consider now star terms with $D>D_n$ from I_2. These terms arise either in the generalized Schrödinger approximation for $M(w)$ with a quartic polynomial in p in the numerator or for Dirac correction terms with a quadratic numerator polynomial. A closer analysis shows that Dirac correction terms do not contribute; the corresponding $M(w)$ all approach zero in the limit $w\Rightarrow 0$.

The Schrödinger quartic terms have $w^3\ln w$ parts that also are proportional to $\psi^2(0)$ and, in fact, the integrals involved can again be expressed as integral linear combinations of the E_{mn} and F_{mn}. The necessary relations are given below. We use the notation

$$\begin{aligned} &M[0_1(p)V0_2(p)V0_3(p);k,m,n]\\ &\equiv\lim_{w\to 0}w^{D/2-1}\left\langle u\left|\frac{0_1(p)}{[1+wp^2]^k}V\frac{0_2(p)}{[1+wp^2]^m}\times V\frac{0_3(p)}{[1+wp^3]^n}\right|u\right\rangle. \end{aligned} \tag{57}$$

In the present case, $D=4$. There are a total of 21 possible scalar quartic numerator polynomials in the p's, in a momentum representation. Ten of these forms are eliminated by the first relation below. The remaining 11 are accounted for by relations 2–7 and left-right symmetry.

The proofs of relations 1–7 are simple. The basic formula is the expansion identity (with $x=wp^2$)

$$(1+x)^{-n}=1-x\sum_{k=1}^{n}(1+x)^{-k}. \tag{58}$$

Use is also made, according to need, of dominating arguments in p space, the boundedness properties of the $K_2{}^{\nu}(p,p')$ given in I, the Schrödinger equation (usually to "synthesize" the $n=3$ situation), the wD_w operation and the properties of the "standard" matrix elements given in Sec. 1. If p_1 enters quadratically (or quartically) in the numerator, the index m in (57) ≥ 2 and it is occasionally necessary to use this fact.

Integrals for I_2 (Quartic-Star)

1. $M(w)$ involving *both* p_0 and p_2 approach zero in the limit $w\Rightarrow 0$ and may thus be discarded.
2. $M(Vp_iVp_ip^2;1,m,n)=2D_{mn}\psi^2(0)$
3. $M(VVp^4;1,m,m)=2F_{nm}\psi^2(0)$
4. $M(Vp^2Vp^2;1,m,n=2F_{mn}\psi^2(0)$
5. $M(Vp^2p_iVp_i;1,m,n)=S_{mn}\psi^2(0)(m\geq 2)$
$$S_{mn}=\frac{1}{4\pi^2}1/(m-1)-2\sum_{k=1}^{n}H_{mk}$$
6. $M(Vp_ip_jVp_jp_i;1,m,n)=T_{mn}\psi^2(0)(m\geq 2)$
$$T_{mn}=D_{mn}+\tfrac{1}{2}S_{mn}$$
7. $M(Vp^4V;1,m,1)=\text{const}w^{-\frac{1}{2}}$
$$-4I_{m1}\psi^2(0)(m\geq 2).$$

(59)

Star Terms with $D \leq D_n$

These terms include all gauge-variant contributions to $w^3 \ln w$ and also all $w^3 \ln^2 w$ terms. The method of organization is adapted to the gauge-invariance test. We will write the various $M_n{}^\nu(w)$ for $n=1$, 2, and 3 (with $\nu=3$) in the following way:

$$M_n{}^\nu(w) = (M')_n{}^\nu(w) + a\psi^2(0) \tag{60}$$

where a is a numerical gauge-invariant constant and $(M')_n{}^\nu$ is a gauge-variant "standard" matrix element. The equivalence is relative to $\ln w$ and "constant" parts only.

The standard M' will be formed from the separable $^*M$ and will be of lowest order $w^{-\frac{1}{2}}$, $\ln w$ or "const" with the minimum power of denominator necessary for convergence. The matrix elements M_A to M_H of Sec. 1 are examples of standard M' with the dash "—" interpreted as $1/1+wp^2$. These matrix elements in fact exhaust all "nontrivial" standard matrix elements except for two from $n=1$ involving Dirac corrections to the small-component wave functions:

$$\begin{aligned} 1/w^2\langle v_s|\, V\Lambda_-|\, v_s\rangle_{\mathrm{Dir}} &\equiv M_I \\ 1/w^{\frac{3}{2}} i\langle v_s|\, V\gamma_4\boldsymbol{\gamma}\cdot\mathbf{p}|\, v_s\rangle_{\mathrm{Dir}} &\equiv M_J \end{aligned} \tag{61}$$

where "Dir" indicates the Dirac correction part. The "renormalization term" of $n=1$ is assumed to be separated out. In addition to these matrix elements, there are also a number of "trivial" matrix elements such as $\epsilon_{sh}{}^4\langle V\rangle$, $\epsilon_{sn}{}^2\langle V^2\rangle$, etc.

After testing gauge invariance by the method indicated in Sec. 2, the part of the I_n results containing the standard M' from $n=1$, 2, and 3 and the gauge-variant part of the I_0 results will be combined into a small number of manifestly gauge-invariant matrix elements. These will then be evaluated for the particular S state under consideration.

The purpose of the standardization of the M' is to permit this combination of matrix elements to be effected in a simple arithmetic way.

It is not obvious *a priori* that the desired separation (60) with constant "a" and standard M' is possible, but a consideration of all possible cases shows that it is. The derivations of the necessary relations are simple and use the same "ingredients" as the proofs of relations (59). The $\psi^2(0)$ integrals are again linear combinations of the E_{mn}, F_{mn} of (51), (52). As examples, we list the following equivalences for quartic star terms from I_1, using a notation that is an obvious extension of the notation of (57).

Integrals for I_1 (Quartic-Star)

1. $M(Vp^4; 1,n)$
$$= cw^{-\frac{1}{2}} + M(Vp^4; 0,1) + 4F_{n1}\psi^2(0).$$
2. $M(p_iVp_ip^2; m,n)$
$$= M(p_iVp_ip^2; 0,1) + 4\sum_{k=1}^{m} D_{kn}\psi^2(0) + \sum_{k=1}^{n-1} (1/k)(1/2\pi^2)\psi^2(0). \tag{62}$$
3. $M(p_ip_iVp_ip_i; m,n)$
$$= \tfrac{1}{2}[M(p_iVp_ip^2; m,n) + M(p_iVp_ip^2; n,m)],$$
(for S states only).
4. $M(p^2Vp^2; m,n)$
$$= M(p^2Vp^2; 1,0) + 4\sum_{k=1}^{n} F_{kn}\psi^2(0) + [\psi^2(0)/\pi^2]\sum_{k=1}^{m-1} (1/k).$$

Polynomials $Q(p,\bar{p})$ and Dirac Algebra for Star and Nonstar Terms

The form of the polynomials $Q_n(p,\bar{p})$ in formula (2) can be predicted in advance on the basis of rotational symmetry and the value of ν, n. These polynomials have rational (usually integral) coefficients which can be read off from the results of the Dirac algebra reduction. It is necessary to keep in mind also the integral binomial coefficients associated with the k expansion.

For S states, the Dirac algebra reduction may be simplified by using a trace technique which automatically eliminates "spin-orbit" terms. If the numerator is denoted by N, then for the $++$ part, for example, it is easy to show that rotational symmetry implies that $\Lambda_+N\Lambda_+$ can be effectively replaced by $\frac{1}{2}\,\mathrm{Tr}(\Lambda_+N\Lambda_+) = \frac{1}{2}\,\mathrm{Tr}(\Lambda_+N)$.

Results for I_3, I_2, and I_1

We list below the principal results for $n=1$, 2, and 3. Each I_n is divided into a "$\psi^2(0)$" part and a "gauge-variant" part, the latter containing the standard matrix elements M'. The notation $M_X{}^n$ with $X=A, B, --J$ refers to the standard matrix elements M_A, M_B — M_H of (12), in the form appropriate to to I_n (nV's in the numerator), and the matrix elements M_I and M_J of (61). The dash "—" in (12) is to be interpreted as $1/(1+wp^2)$. The "renormalization" term of I_1 proportional to $\langle v_s|\, V_0|\, v_s\rangle$, has been omitted.

For convenience in carrying out the later non-S-state calculation, we have retained the matrix element M_D which for S states is eliminated by relation 3 of (62). The *convention* is understood that for S states the

coefficient of M_D should be added to that of $M_C{}^1$ and only the latter matrix element used.

$\psi^2(0)$ results are in units of $m\pi^2w^3\ln w\psi^2(0)/4\pi^2$ while the gauge-variant part is in "units" of $m\pi^2w^3\int^w dw/w$.

I_3 (Results)

$$I_3(\psi^2(0))=-52/3\ln 2+41/6 \tag{63}$$

$$I_3(\text{gauge var})=-6M_B{}^3. \tag{64}$$

I_2 (Results)

$$I_2(\psi^2(0))=48\ln 2-166/9 \tag{65}$$

$$I_2(\text{gauge var})=-11/3M_A{}^2+16M_B{}^2-16/3M_C{}^2+3/2M_F-12M_G{}^2-9\epsilon^2\langle V^2\rangle. \tag{66}$$

I_1 (Results)

$$I_1[\psi^2(0)]=-20\ln 2+43/9 \tag{67}$$

$$\begin{aligned}I_1(\text{gauge var})=&16M_I-4M_J+158/15M_A{}^1\\&+166/15M_B{}^1+58/15M_C{}^1+22/15M_D{}^1\\&-20/3M_E{}^1+4/3M_H+40/3M_G{}^1\\&-7/6\epsilon^2\langle p_iVp_i\rangle-43/3\epsilon^2\langle Vp^2\rangle\\&-12\epsilon^2\langle u'|V|u\rangle-6\epsilon^4\langle V\rangle-6\delta\langle V\rangle.\end{aligned} \tag{68}$$

Cancellation of $w^3\ln^2 w$ Coefficients from $n=3$, 2, and 1

The $w^3\ln^2 w$ coefficient from $n=1$, 2, and 3 arises from the $\ln w$ part of the standard matrix elements M_B, M_C, and M_E in the gauge-variant part of the results above. By using the equivalences of (15) we can express all matrix elements in terms of M_B as far as the $w^3\ln^2 w$ coefficient is concerned. From the gauge-variant part of the above results, we obtain, in "units" of $m\pi^2w^3\int^w dw/w$

$$\begin{aligned}&I_3(w^3\ln^2 w)=-6M_B;\\&I_2(w^3\ln^2 w)=+(64/3)M_B;\\&I_1(w^3\ln^2 w)=(-46/3)M_B.\end{aligned} \tag{69}$$

Thus, the sum $\sum_{n=1}{}^3 I_n(w^3\ln^2 w)$ is zero. We will show in the next subsection that this result has a connection with and is consistent with the w-gauge invariance of the $w^3\ln w$ coefficient.

Verification of Gauge Invariance of $w^3\ln w$ Coefficient

In the present subsection, we will apply the gauge-invariance test outlined in assertions 1–6 of Sec. 2. The results for each I_n have been divided into a "gauge-variant" part and a $\psi^2(0)$ part. The latter can easily be shown to be w-gauge invariant.

Thus, to test gauge invariance we can apply the displacement $V\Rightarrow V+g$; $\epsilon_{\rm sh}{}^2\Rightarrow\epsilon_{\rm sh}{}^2-2g$ to the matrix elements of the gauge-variant parts of the results for I_n, $n=0$, 1, 2, 3.

For the check of the "one-g; no-V" relation of assertion 6: $I_0=-I_1$, it is convenient to reduce the $+-$ and $--$ Dirac correction forms of I_1 by means of the Dirac equation to forms involving at most the Dirac correction u' to the large-component wave function. We know this is possible because the results for I_0 do not contain the Dirac correction to the quantity φ, where $\chi=i\gamma_4(\gamma\cdot p/2m)\varphi$ and χ is the small-component wave function. The necessary identities for this purpose are given below. The proofs are simple. The most important element of the proofs is the use of the Dirac equation in the form

$$\Lambda_-|v\rangle=(2m)^{-1}[i\gamma_4\gamma\cdot p+V+(m-E)]|v\rangle. \tag{70}$$

We obtain, using the notation of (61)[11]:

$$M_I=\tfrac{1}{4}\langle u|p_iV^2p|u\rangle+\tfrac{1}{8}\epsilon_{sn}{}^2\langle p_iVp_i\rangle+\tfrac{1}{2}\langle u'|p_iVp_i|u\rangle \tag{71}$$

$$\begin{aligned}M_J=&\tfrac{1}{4}\langle u|p_iV^2p_i|u\rangle+\tfrac{1}{8}\epsilon_{sn}{}^2\langle p_iVp_i\rangle+\langle u'|p_iVp_i|u\rangle\\&-2\langle u'|V^2|u\rangle-\epsilon^2\langle u'|V|u\rangle\\&-(E_{\rm sh}-E/mw^2)\langle V\rangle.\end{aligned} \tag{72}$$

Furthermore, we obtain for the Dirac correction to the energy[11]

$$E_{\rm sh}-E=-\tfrac{1}{4}mw^2\langle p_iVp_i\rangle-\tfrac{1}{8}\epsilon_{\rm sh}{}^2\langle p^2\rangle mw^2. \tag{73}$$

The polynomial in g due to each I_n has the following form, where $X_n{}^{i\nu}$ for $X=A$, $B--S$, as in assertion 5 of Sec. 2 is the coefficient for a contribution to g^i with ν numerator V's (*before* the use of the Dirac or Schrödinger equation) due to I_n. To obtain this form, the Dirac and Schrödinger equations have been used for I_0 and in the reduction of the matrix elements M_I and M_J of I_1 but not elsewhere. The "renormalization terms" of I_0 and I_1, which cancel against each other, are omitted. We have used $\langle u|u\rangle=1$. A common factor of $m\pi^2w^3\ln w$ is omitted.

$$\begin{aligned}I_n(g)-I_n=&g^3A_n{}^{30}+g^2(B_n{}^{21}\langle V\rangle+C_n{}^{20}\langle p^2\rangle\\&+D_n{}^{20}\langle u'|u\rangle+E_n{}^{20}\epsilon^2)+g(F_n{}^{12}\langle V^2\rangle\\&+H_n{}^{11}\langle Vp^2\rangle+I_n{}^{11}\langle p_iVp_i\rangle+J_n{}^{11}\langle u'|V|u\rangle\\&+K_n{}^{11}\epsilon^2\langle V\rangle+M_n{}^{10}\epsilon^2\langle u'|u\rangle+N_n{}^{10}\langle u'|p^2|u\rangle\\&+O_n{}^{10}\langle p^4\rangle+P_n{}^{10}\langle p_iVp_i\rangle+Q_n{}^{10}\epsilon^2\langle p^2\rangle\\&+R_n{}^{10}\epsilon^4+S_n{}^{10}\delta).\end{aligned} \tag{74}$$

In a non-S-state calculation there would also be the forms

$$G_n{}^{11}i\langle\sigma_{ij}p_iVp_j\rangle \quad\text{and}\quad L_n{}^{10}i\langle\sigma_{ij}p_iVp_j\rangle$$

(cf. Sec. 4).

The letters X_n for $X=A$, $B--S$ are linear combinations of the coefficients of the matrix elements of the gauge-variant part of the results for I_n. For

[11] The corresponding expressions for non-S states are obtained by *adding* certain spin-orbit matrix elements to the right-hand side. These spin-orbit matrix elements may be obtained from the S-state expressions by replacing the combination $p_i\cdots p_i$ wherever it appears by $\sigma_{ij}p_i\cdots p_j$.

example,

$$H_2^{11}=2(-11/3+16)-16/3 \tag{75}$$

$$O_1^{10}=(158+166+58+22)/15. \tag{76}$$

Terms arising from the matrix elements M_I and M_J of I_1 contribute to the coefficients L, M, N, P, and Q using (71)–(73) and the Schrödinger equation.

It is now easy to verify that $\sum_{n=0}^{3} X_n=0$ for each $X=A$, $B--S$ and furthermore, the nature of the cancellation agrees with all the identities of assertion 6, Sec. 2. For example, we have $C_2=-31/2$ which (one can show) is made up of -10, -10, $+9/2$ for $++$, $+-$, and $--$, respectively. $C_1=+31=+20$, $+20$, -9 for $++$, $+-$, and $--$. $C_0=-31/2$. In general, the cancellations of assertion 6 apply to $++$, $+-$, and $--$ separately except for identities involving I_0.[12]

Since the $w^3\ln^2 w$ coefficient can be expressed in terms of the coefficients of certain gauge-variant matrix elements, the above verification of gauge invariance is a strong check on the accuracy of the $w^3\ln^2 w$ coefficient as previously calculated and, in particular, on the cancellation of this coefficient between $n=1$, 2, and 3. It is even true that the coefficient -6 in (64) and, therefore, the $w^3\ln^2 w$ coefficient from I_3 can be predicted on the basis of w-gauge invariance and the I_0 results. However, it is not possible to predict the cancellation of the $w^3\ln^2 w$ coefficient between $n=1$, 2, and 3 only on the basis of w-gauge invariance and the results of I_0 (or even I_0+I_1) as the example (75) shows.

Totals for "Non-$\psi^2(0)$" Part of ΔE

In this subsection we will add together, in matrix element form, the gauge-variant contributions to $w^3\ln w$ from $n=0$, 1, 2, and 3 and express the result in a concise, manifestly gauge-invariant form. This will be followed by an evaluation of the matrix elements for the $1-S$ and $2-S$ states. Finally, we will add in the $\psi^2(0)$ results and obtain a numerical value for the $w^3\ln w$ coefficient for the $1-S$ and $2-S$ cases.

In the following, we will use the name "non-$\psi^2(0)$" terms for terms previously designated as gauge-variant terms.

$\Delta E[\text{non-}\psi^2(0)]$ assumes a simpler form when the term of (72) involving $(E_{\rm sh}-E)$ is eliminated through the easily derived equivalence[11]

$$(E_{\rm sh}-E/mw^2)\langle V\rangle = -\tfrac{1}{2}\langle u'|p^2-V|u\rangle+\tfrac{1}{2}\langle u'|Vp^2|u\rangle -\tfrac{1}{4}\langle V-p_iVp_i\rangle-\tfrac{1}{8}\epsilon^2\langle Vp^2\rangle. \tag{77}$$

On adding together the gauge-variant parts of the results for I_n, $n=0$ to 3, and using the equivalence (77), one obtains after a lengthy but elementary reduction involving the use of the Schrödinger equation

$$\Delta E[\text{non-}\psi^2(0)] = m\pi^2w^3\int^w dw/w\{-8/3\langle u'|-q^2V|u\rangle -11/15\langle p_i-q^2V-p_i\rangle+16/15\langle p^2-q^2V\rangle +\tfrac{4}{3}\langle p^2-[p_i,V]p_i\rangle-4\langle[p_i,V]-[p_i,V]\rangle\}. \tag{78}$$

This form for $\Delta E[\text{non-}\psi^2(0)]$ is concise and obviously gauge invariant. Note that $\langle p_i-q^2V-p_i\rangle$ vanishes for S states.

By expanding the commutators in (78) and using the equivalences of (15) we can verify that the $w^3\ln^2 w$ coefficient is still $\frac{1}{2}m\psi^2(0)$, which is a check on the previous reduction.

Numerical Evaluation of "Non-$\psi^2(0)$" Coefficient of $w^3\ln w$ for 1-S and 2-S States

The evaluation for the $1-S$ and $2-S$ cases of the three nonvanishing Schrödinger matrix elements in (78) is simple.[10]

The Dirac correction wave function u' in position space for the $1-S$ and $2-S$ cases, may be derived from the explicit expression for the Dirac wave functions[13] or, more easily, from the differential Eq. (16) plus the normalization requirement[12] which in the present case ($V_0=-1/r$) is the same as

$$\langle u'|u\rangle=-1/(8n^2) \quad \text{for} \quad n-S \text{ state.} \tag{79}$$

The following expressions for u' are obtained:

For the $1-S$ state:

$$\langle u_s'|r\rangle=au_s-\tfrac{1}{2}\ln ru_s \tag{80}$$

where

$$u_s=Ce^{-r} \tag{81}$$

$$a=\tfrac{5}{8}-\tfrac{1}{2}\ln\gamma-\tfrac{1}{2}\ln 2 \tag{82}$$

$$\ln\gamma=0.577--. \tag{83}$$

For the $2-S$ state:

$$\langle u_s'|r\rangle=au_s-\tfrac{1}{2}\ln ru_s+bC_{2-S}re^{-r/2}+cC_{2-S}r^2e^{-r/2} \tag{84}$$

$$u_s=C_{2-S}(1-r/2)e^{-r/2} \tag{85}$$

$$a=\tfrac{7}{32}+\tfrac{1}{2}(1-\ln\gamma); \quad b=-\tfrac{5}{16}; \quad c=+\tfrac{1}{32}. \tag{86}$$

[12] Since the Dirac equation was used in the reduction of I_0 and of the matrix elements M_I and M_J for I_1, it is at first surprising that the cancellation between I_0 and I_1 of the forms with coefficients $M, N\cdots Q$ and the fulfillment of the relation $I_0=-\frac{1}{2}I_1$ for forms with coefficients C, D, and E takes place for each "letter" separately. However, a closer analysis shows that this was to be expected. In particular, the Dirac and Schrödinger forms must cancel separately in spite of the use of the Dirac equation (and the Schrödinger normalization $\langle u|u\rangle=1$). This is true because the differential equation for u' determines u' only to within λu where λ is an arbitrary number. λ is determined only by normalizing v (in addition to u). This imposes the additional restriction

$$\langle u'|u\rangle=-\tfrac{1}{8}\langle p^2\rangle,$$

but this normalization has not been used in writing (74).

[13] H. A. Bethe and E. E. Salpeter, *Quantum Mechanics of One and Electron Systems* (Academic Press, Inc., New York, 1957), p. 69.

Actually, a knowledge of the constant b and c in (84) is not required for the evaluation of the Dirac matrix element in (78). The explicit evaluation of the Dirac matrix element is

$$\langle u_s'|-q^2V|u\rangle = -(2/\pi)^{\frac{1}{2}}\psi(0)[a_0\{\ln(w^{\frac{1}{2}})+1-\ln\gamma\}+a_1], \quad (87)$$

where a_0 and a_1 are defined by the expansion

$$\langle u_s'|r\rangle = a_0\ln r + a_1 + \text{higher order}. \quad (88)$$

The following numerical values were obtained for the matrix elements of (78), dropping lower-order terms proportional to $w^{-\frac{1}{2}}$ and $\ln w$.[14]

$$\begin{array}{lrl} (1) & \langle u'|-q^2V\rangle = 2\ln 2-\frac{1}{2} & 1-S \\ & -7/64 & 2-S \\ (2) & \langle p^2-q^2V\rangle = 12 & 1-S \\ & =9/8 & 2-S \\ (3) & \langle[p_i,V]-[p_i,V]\rangle = -8\ln 2 & 1-S \\ & -\frac{5}{8} & 2-S \\ (4) & \langle p^2-[p_i,V]p_i\rangle = -8\ln 2-2 & 1-S \\ & -11/16 & 2-S. \end{array} \quad (89)$$

On substituting (89) into (78), we obtain, in units of $m\pi^2w^3\ln w$:

$$\Delta E(\text{non-}\psi^2(0);\, w^3\ln w) = 16\ln 2+172/15 \quad 1-S \quad (90)$$

$$\Delta E(\text{non-}\psi^2(0);\, w^3\ln w) = 3+3/40 \quad 2-S. \quad (91)$$

Totals for $\psi^2(0)$ Part

From (63), (65), (67), and (28) we obtain, in units of $mw^3\ln w\pi^2$,

$$\Delta E[\psi^2(0)] = \sum_{n=1}^{3} I_n[\psi^2(0)] = \frac{\psi^2(0)}{4\pi^2}\left[\frac{32}{3}\ln 2-\frac{47}{6}\right]. \quad (92)$$

Final Results

From (90)–(92) and (30), using also the normalization

$$\psi^2(0) = 8\pi^2/n^3,$$

we obtain[15]

$$\Delta E(1-S) = m\pi^2w^3[4\ln^2 w+\ln w[112/3\ln 2-21/5]] \quad (93)$$

$$\Delta E(2-S) = m\pi^2w^3[\tfrac{1}{2}\ln^2 w+\ln w(8/3\ln w+67/60)]. \quad (94)$$

4. NON-S-STATE CALCULATION OF LAMB SHIFT

The formal setup for non-S states is very similar to that of S states. In fact, only the wave functions are changed. The "scaling" goes through as before; the division into large-large, large-small, and small-small parts by means of the projection operators $\Lambda\pm$ is still possible and one can still use the generalized Schrödinger approximation. The k expansion, which is independent of the wave functions and the scaling, is unchanged.

There is one obvious formal difference, however, caused by the change of the wave functions from S state to non-S state. This is the appearance of additional "spin-orbit" terms, involving operators which are linear combinations of the Pauli spin matrices $\boldsymbol{\sigma}_i$ or their tensor equivalents $\sigma_{ij}\equiv(1/2i)[\gamma_i,\gamma_j]$ together with two component Schrödinger wave functions.[16] We call these terms "spin-orbit" terms since in every case that will occur in the present calculation, the corresponding operators can be expressed in terms of the scalar operator $\boldsymbol{\sigma}\cdot\mathbf{L}$. For example,

$$i\langle u|\sigma_{ij}p_iVp_j|u\rangle = -\left\langle u\left|\frac{1}{r}\frac{dV}{dr}\boldsymbol{\sigma}\cdot\mathbf{L}\right|u\right\rangle. \quad (95)$$

In spite of this additional feature, the non-S-state calculation turns out to be considerably simpler than the S state one. This is because of the improved behavior of the non-S-state Schrödinger wave functions at small r and high p, in particular the vanishing of $\psi^2(0)$. As a result, many terms that appear in the S-state case, namely all those proportional to $\psi^2(0)$, do not appear in the non-S-state case. In particular, $w^2\ln w$, $w^{\frac{1}{2}}$, and $w^3\ln^2 w$ terms are absent and the "nonstar" ($\nu\neq 3$) category of $w^3\ln w$ terms is eliminated.

---

[14] Dr. G. Erickson has kindly informed me that the following generic expressions exist for the values of the matrix elements (1)–(4) of formula (89) for the general n-S state:

$$\langle u'|-q^2V\rangle = 2/n^3[\ln(2/n)+1+\tfrac{1}{2}+\cdots+1/(n-1)-1/n+9/(4n^2)-\tfrac{3}{2}],$$

$$\langle p^2-q^2V\rangle = 8/n^3[1+1/(2n^2)],$$

$$\langle(p_i,V)-(p_i,V)\rangle = -8/n^3[\ln(2/n)+1+\tfrac{1}{2}+\cdots 1/(n-1)+1/(2n)+1/(6n^2)-\tfrac{2}{3}]$$

and, finally,

$$\langle p^2-(p_i,V)p_i\rangle = \langle(p_i,V)-(p_i,V)\rangle - 2/n^5.$$

[15] According to Erickson's generic formulas of the previous footnote for the non-$\psi^2(0)$ matrix elements of formula (89), the result for the $n-S$ state can be written in the form

$$\Delta E(n-S) = 4m\pi^2w^3/(n^3)\ln w\{\ln w+4[\ln(2/n)+1+\tfrac{1}{2}\cdots+1/n] +16/3\ln 2-601/180-75/(45n^2)\}.$$

Except for the last two terms, proportional to $1/n^3$ and $1/n^5$, this expression has been confirmed by an independent calculation of Erickson and Yennie, still in progress (private communication from Dr. Erickson). Their method employs a new formal treatment of the self-energy, in which, in distinction to the free-propagator expansion, the bound electron propagator is left in closed form until the last stages of the calculation (see also Dr. Erickson's dissertation, reference in footnote 1).

[16] In the S-state case, the trace operation eliminated such terms.

The principal results for the non-S-state case are summarized in the five assertions listed below.

1. The "main result" of I is still true.
2. There are no $w^2\ln w$, $w^{\frac{3}{2}}$, or $w^3\ln^2 w$ coefficients for non-S states.
3. The non-S-state contribution to $w^3\ln w$ can be obtained by adding "spin-orbit" terms to the "gauge-variant" part of the I_0, I_1, I_2, and I_3 results given in (64), (66), (68), and (29).
4. The spin-orbit terms, like the nonspin-orbit terms, are "star" terms ($\nu=3$). Furthermore, they arise only for $n=0$, 1, and 2 (not $n=3$), and for $n=2$, they occur only in the generalized Schrödinger approximation with a quadratic in p numerator ($D=2$).
5. The final results for the $w^3\ln w$ contribution for non-S states can be obtained by adding the spin-orbit term

$$\Delta E(s-0;\, w^3\ln w)=m\pi^2w^3\ln w(-\tfrac{1}{3}i\langle\sigma_{ij}p_iq^2Vp_j\rangle) \quad (96)$$

to the previous non-$\psi^2(0)$ results in the form (78).

The proofs of these assertions are straightforward and are omitted.

Evaluation of Non-$\psi^2(0)$ Matrix Elements for $2P_{\frac{1}{2}}$ and General Non-S State

Note that the Dirac matrix element of (78) and the Schrödinger matrix element $\langle p^2-q^2V\rangle$ vanish for non-S states. Also, we can remove the dash "—" in (78), and replace $\int^w dw/w$ by $\ln w$.

The evaluation of the spin-orbit matrix element (96) and the three nonvanishing matrix elements of (78) is simple. The following results are obtained.

1. $$P\equiv\langle p_iq^2Vp_i\rangle = -2\pi L(L+1)[1/r^2d/dr(rR_{nL})^2]_{r=0} \quad (97)$$
$$P=-\tfrac{1}{8} \quad \text{for} \quad 2P_{\frac{1}{2}}$$

2. $$Q\equiv i\langle\sigma_{ij}p_iq^2Vp_j\rangle = -2\frac{[J(J+1)-L(L+1)-\frac{3}{4}]}{L(L+1)}P \quad (98)$$
$$Q=-\tfrac{1}{4} \quad \text{for} \quad 2P_{\frac{1}{2}}$$

3. $$\langle[p_i,V][p_i,V]\rangle=-1/r^4 \quad (99)$$

4. $$\langle p^2[p_i,V]p_i\rangle=-2\langle V[p_i,V]p_i\rangle=-\langle 1/r^4\rangle. \quad (100)$$

Here,[17]

$$\langle 1/r^4\rangle=-\frac{1}{2}\frac{3n^2-L(L+1)}{n^5(L+\frac{3}{2})(L+1)(L+\frac{1}{2})L(L-\frac{1}{2})} = 1/24 \quad \text{for} \quad 2P_{\frac{1}{2}}. \quad (101)$$

[17] H. A. Bethe and E. E. Salpeter, reference in footnote 13, p. 17.

To derive (100), we used the identity

$$\langle r^{s-1}(\partial/\partial r)\rangle=-\tfrac{1}{2}(s+1)\langle r^{s-2}\rangle \quad (102)$$

for all s such that the right-hand side exists. This follows from $\langle[H,r^s]\rangle=0$. In (97) R_{nL} refers to the radial wave function of the n, L state normalized so that $4\pi\int|R_{NL}|^2r^2dr=1$.

Final Results for $2P_{\frac{1}{2}}$ State and $\Delta E(2S_{\frac{1}{2}}-2P_{\frac{1}{2}})$

On substituting (97)–(101) into (78) and (96), we obtain

$$\Delta E(2P_{\frac{1}{2}};\, w^3\ln w)=m\pi^2w^3\ln w\ 103/360. \quad (103)$$

Or, from (94),

$$\Delta E(2S-2P_{\frac{1}{2}};\, w^3\ln^2 w \quad \text{and} \quad w^3\ln w) = m\pi^2w^3[\tfrac{1}{2}\ln^2 w+\ln w(8/3\ln 2+1-61/360)]. \quad (104)$$

5. NEW THEORETICAL VALUES FOR THE LAMB SHIFT[18]

To obtain the absolute value of the $2S-2P_{\frac{1}{2}}$ shift due to the new orders calculated here, we insert the normalization factor of $-\alpha/4\pi^3$ omitted from our starting expression for ΔE and use the evaluation:

$$\psi^2(0)\equiv\left|\int d^3p\langle u|p\rangle\right|^2=8\pi^3|\langle x=0|u\rangle|^2 = 8\pi^2/(n^3) \quad \text{for} \quad S \text{ states}. \quad (105)$$

Then we obtain from (104) the result

$$\Delta E(2S-2P_{\frac{1}{2}})=-Lw[3/4\ln^2 w+\ln w(4\ln 2+1+27/40-103/240)], \quad (106)$$

where L is Z^4 times the Lamb constant:

$$L\equiv Z^4\alpha^3/(3\pi)Ry=Z^4(135.6\text{ Mc}).$$

To this we must add a small contribution to the $\ln w$ term due to vacuum polarization, easily derivable as the effect of the Dirac modification of the large component wave function on the expectation value of the Uehling potential[19]:

$$\Delta E(\text{vac. pol.})=-Lw\ln w(-1/10)\psi^2(0)/\pi^2. \quad (107)$$

Adding this to (106), we obtain

$$E(2S_{\frac{1}{2}}-2P_{\frac{1}{2}}, \text{ including vac. pol.}) = -Lw(3/4\ln^2 w+\ln w(4\ln 2+1+7/48)). \quad (108)$$

[18] The principal results of this section were reported previously (reference in footnote 2). However, the theoretical values for D and He^+ given in Table I of that article did not incorporate the latest values of the nuclear finite size correction. This oversight is corrected in Table I of this section.

[19] A. E. Uehling, Phys. Rev. **48**, 55 (1935). Our result follows immediately from expressions given in the papers of E. Wichmann and N. M. Kroll, Phys. Rev. **101**, 843 (1956) or N. M. Kroll and F. Pollock, Phys. Rev. **86**, 876 (1952).

TABLE I. Theoretical and experimental values in Mc of the Lamb shift in H, D, and He^+.

	H	D	He^+
Experimental	1057.77±0.10	1059.00±0.10	14040.2±4.5
Theoretical	1057.70±0.15	1059.08±0.16	14047.2±3.0

The corresponding numerical results in megacycles are -0.25 for H or D and -9.5 for He^+. These additions reduce markedly the disagreement between theory and experiment.[20] To obtain the new theoretical values for the Lamb shift, we shall make use of the previous theoretical values as tabulated by Petermann[21] with, however, more recent values for the finite size corrections in D and He^+. A newer value for the nuclear finite size correction for He^+ is given by Lipworth and Novick.[22] For D we use the value of the finite size correction given by Yennie, Lévy, and Ravenhall.[23] Adding in the new contributions calculated here, we obtain the theoretical values listed in Table I, which are compared with experimental values for H,[24] D,[24] and He^+.[25] Theoretical errors are copied from Petermann's article.

It is of interest to list also the contribution to the fine structure separation $2P_{\frac{1}{2}}-2P_{\frac{3}{2}}$ which is due, of course, to the spin-orbit matrix element (96). We obtain[26]

$$\Delta E(2P_{\frac{1}{2}}-2P_{\frac{3}{2}})=-Lw\ln w(-\tfrac{3}{16}). \qquad (109)$$

In megacycles, this is only $(-)0.01$ for $Z=1$ and is thus much smaller than the experimental error of 0.20 Mc for the separation in deuterium.[25]

CONCLUDING REMARKS

As outlined in the preceding work, we have used the free-propagator expansion and the k expansion to carry out detailed calculations of nonanalytic orders of the Lamb shift through the order $\alpha w^3 \ln w$. For the previously calculated orders we have obtained results in agreement with previous values, and for the new orders of $\alpha w^3 \ln^2 w$ and $\alpha w^3 \ln w$ we have obtained reasonable results that to the present experimental and theoretical accuracy remove the discrepancy between the theoretical and experimental values of the Lamb shift.

These detailed calculations were based on the more general mathematical analysis of I. This analysis, it should be stressed, was incomplete in two important respects: neither the convergence of the free-propagator expansion nor that of the k expansion was rigorously established. These interesting questions of convergence are more difficult than those considered here and must be regarded as open. However, the successful use of the two expansions in rederiving the coefficients of previously calculated nonanalytic orders justifies an optimistic attitude with regard to the validity of the methods used here for the calculation of the new orders of $\alpha w^3 \ln^2 w$ and $\alpha w^3 \ln w$.

In connection with the convergence of the free-propagator expansion there are a few comments of a qualitative nature that can be made which may help to clarify the situation. For the purposes of this discussion, the convergence of the k expansion will be assumed.

In the first place, it is not hard to see that for the general bound state the free-propagator expansion does *not* converge to (ΔE) in the strict sense. This is because each term I_n of the free propagator expansion is real while ΔE as a whole, for all states except the ground state, contains an imaginary part proportional to the transition rate for decay to lower states.[27] Closely connected with this difficulty is the fact that for each term of the free-propagator expansion one can immediately rotate the path of the k_0 integration to the imaginary axis while this is not possible for ΔE as a whole except for the $1-S$ state. One can get around this difficulty of the imaginary part in a formal way

[20] It should be mentioned that the uncomputed fourth- and sixth-order radiative corrections of order $\alpha^2 ZL$ and $\alpha^2 L$, respectively, should contain no logarithmic factors and are expected to be numerically negligible. Fourth-order radiative corrections to the Lamb shift were evaluated in lowest order αL by J. Weneser, R. Bersohn, and N. M. Kroll, Phys. Rev. **91**, 1257 (1953), except for the fourth-order vacuum polarization contribution which was evaluated by M. Baranger, F. J. Dyson, and E. E. Salpeter, Phys. Rev. **88**, 680 (1952). The method used by J. Weneser, R. Bersohn, and N. M. Kroll was subsequently justified by R. L. Mills and N. M. Kroll, Phys. Rev. **98**, 1489 (1955). The absence of logarithmic factors in (αZ) in 4th- and 6th-order radiative corrections at the present order was pointed out to the author by N. M. Kroll. For fourth-order corrections this is asserted without a detailed proof in the last section of the cited work of Mills and Kroll. The absence of $\alpha L \ln w$ terms in the lowest-order fourth-order radiative corrections suggests that logarithmic factors are not present in the lowest order sixth-order radiative corrections, of order $\alpha^2 L$.

[21] A. Petermann, "Atomic energy level shifts in hydrogen-like atoms," Fortschr. Physik **6**, 507 (1958).

[22] E. Lipworth and R. Novick, Phys. Rev. **108**, 1434 (1957).

[23] D. R. Yennie, M. M. Levy, and D. G. Ravenhall, Revs. Modern Phys. **29**, 144 (1957). The change of +0.12 Mc is due to the finite size of the proton. The charge form factor of the neutron is assumed to vanish.

[24] S. Triebwasser, E. Dayhoff, and W. Lamb, Phys. Rev. **89**, 98 (1953).

[25] E. Dayhoff, S. Triebwasser, and W. Lamb, Phys. Rev. **89**, 106 (1953).

[26] In the author's previously published note (reference in footnote 2) the left-hand side of this equation is erroneously written as $\Delta E(2P_{\frac{1}{2}}-2P_{\frac{3}{2}})$. We are grateful to Professor R. E. Cohen for pointing out the ambiguity in sign of this notation. [Our present notation is consistent with the fact that the $2P_{\frac{3}{2}}$ level lies above the $2P_{\frac{1}{2}}$ level.]

[27] The imaginary part of ΔE, which can be evaluated in closed form, has been treated relativistically and in detail by L. Kaagjarv and S. T. Ma, Nuovo cimento **8**, 432 (1958). This imaginary part is incorporated into the algebraic structure of the self-energy operator through the presence of small negative imaginary parts in the k^2 of the photon propagator and in the mass m of the electron in the expression for the bound electron propagator, in accordance with the usual Feynman prescription. As a consequence of this prescription, the energies of positive and negative energy intermediate states in the sum-over-states representation of the bound electron propagator have small negative and positive imaginary parts, respectively, in agreement with the prescription of Kaagjarv and Ma.

by regarding the free-propagator expansion as an expansion of the real part of ΔE alone, but in doing this we are covering up mathematical difficulties which may affect the convergence of the real part of ΔE as well.

To examine the question of the convergence of the free-propagator expansion, we must look at the remainder after n terms R_n of the expansion in finite form and see whether it approaches zero as $n \to \infty$. R_n has the same form as I_n except that the last free propagator is replaced by a bound propagator. It is not difficult to show by considering the pole at $k^2=0$ that one can recover the imaginary part of ΔE from R_n alone as must be the case. This shows, however, that, for states other than the ground state, the imaginary part of R_n certainly does not approach zero, and therefore, the free propagator-expansion does not converge to ΔE, as already noted. Since the same iterated structures are involved in the real part of R_n as in the imaginary part, these considerations also cast doubt on the convergence of the expansion for the real part of ΔE, especially for states other than the ground state.

In spite of the above remarks prejudicial to the convergence of the free-propagator expansion in the strict sense, the situation in practice is much brighter. It is a fact that the sum of the "spurious" lowest-order terms for the remainder after I_0+I_1 was obtained correctly as outlined in I (Sec. 2) by a formal algebraic sum over n from $n=2$ to infinity. Furthermore, in F–Y the coefficient of the order αw^2 was obtained correctly, in their formalism, by a formal sum technique. In this case the order is the same as that of the lowest order of the imaginary part of ΔE for the general bound state. It is thus plausible that by taking formal algebraic sums for a given order in w of interest one can "transcend" the free-propagator expansion and avoid convergence difficulties associated with the expansion.

Finally, we remark that the problem of the convergence of the free propagator expansion has much less force for the orders nonanalytic in w that have been calculated in the present paper since a given nonanalytic order is contained in only a finite number terms of the expansion. In this connection it is satisfying that the imaginary part of ΔE, as one can infer from expressions given by Kaagjarv and Ma,[27] is an analytic function of w for small w. Thus, the imaginary part of ΔE does not contribute to the discontinuity $\Delta(\Delta E)$ of ΔE on the cut along the negative w axis. We can then regard $2\sum_n \mathrm{Im}_+ I_n = \sum_n \Delta I_n$ as an expansion of $\Delta(\Delta E)$ and it is now reasonable to estimate the order of the remainder after n terms in this expansion by replacing the bound propagator of the exact expression for the remainder by a free propagator. In this way, according to theorem 4 of Sec. 3 of I, one obtains an estimate of order for the remainder that can be made as high as desired by picking n sufficiently large. Thus, even if one adopts a pessimistic view about the convergence of the free propagator expansion for ΔE, it is plausible that the expansion $2\sum_n \mathrm{Im}_+ I_n$ is valid for $\Delta(\Delta E)$, if only in the sense of an asymptotic expansion in orders of w.

ACKNOWLEDGMENTS

It is a pleasure to acknowledge my indebtedness to Professor N. M. Kroll for suggesting the desirability of the calculation of the new logarithmic orders as well as the probable existence of the $\ln^2$ term. I wish to thank him also for valuable advice and invaluable encouragement during the course of this work. I would like to thank Professor D. R. Yennie and Dr. H. M. Fried for providing me with a preprint of their work and for several beneficial discussions. A stimulating discussion of the results with Dr. J. Weneser is acknowledged. Finally, I wish to thank Dr. D. Zwanziger for many helpful comments.

Some Properties of Empty Space-Time*

E. NEWMAN
Department of Physics, University of Pittsburgh, Pittsburgh, Pennsylvania
(Received July 26, 1960)

In close analogy to the "vierbein" technique of representing world vectors as linear combinations of four fixed vectors, we have constructed from six linearly independent bivectors a (complete) set, or basis, of ten tensors of rank 4 that possess all the algebraic properties of the empty-space Riemann-Christoffel tensor (i.e., Weyl's tensor). Any actual Weyl tensor must be a linear combination of these ten; the expansion coefficients with respect to a given basis are uniquely determined. We have examined Petrov's classification scheme in terms of this formalism and have developed a further division into subclasses. Finally, we shall present in this paper, as an application of our technique, the derivation of plane-fronted waves.

I. INTRODUCTION

PETROV[1,2] was the first to classify the Riemann-Christoffel tensors of empty space (which are numerically identical with Weyl's tensor) according to their local algebraic properties. Pirani[2] used Petrov's classification to define, again locally, "pure radiation."

In this paper, we shall present a different approach to the problem of classifying empty-space geometries. Our approach is based on the construction of an algebraic basis for Weyl's tensors. We begin with a vierbein formalism, which differs from the usual construction only in that we replace the timelike and one spacelike unit vector by two null vectors. From the four world vectors, we construct six bivectors (i.e., antisymmetric tensors of rank 2), which are pairwise each others' duals. Finally, from the bivectors we can construct ten linearly independent tensors of rank 4, which have all the algebraic properties of Weyl's tensor, i.e., they are antisymmetric in each of two index pairs, symmetric with respect to the interchange of the two pairs, cyclically symmetric with respect to any triplet of indices, and all contractions vanish. These ten combinations, quadratic in the bivectors, form a basis; thus, any Weyl's tensor may be expressed as a (unique) linear combination of the chosen ten basis tensors. Petrov's classification may be formulated in terms of the expansion coefficients. Moreover, by studying the field of the coefficients in the neighborhood of a world point, we shall extend the algebraic classification scheme to include differential properties as well. Our results include, but go beyond, so-called propagation studies. Work with a "vierbein" formalism is, of course, not new. In our approach, the chosen vectors are, however, adapted to the preferred directions of the various "special" types of Weyl's tensors, including Pirani's radiation types. In the latter types, one of our null vectors plays the role of the propagation vector.

The concluding section of this paper deals specifically with plane[3] and plane-fronted[4] waves, which were originally studied by Bondi, Pirani, and Robinson.

II. CLASSIFICATION

Consider four independent vectors ξ_μ, χ_μ, ζ_μ, ρ_μ, ξ_μ, and χ_μ being null vectors satisfying $\xi_\mu\chi^\mu=1$, $\xi_\mu\xi^\mu=\chi_\mu\chi^\mu=0$, and ζ_μ and ρ_μ being spacelike, orthogonal to ξ_μ and χ_μ and satisfying $\zeta_\mu\rho^\mu=0$, $\zeta_\mu\zeta^\mu=\rho_\mu\rho^\mu=-1$. The signature of the metric tensor is $(+\cdots)$.

We form all six antisymmetric pairs of the vierbein vectors and introduce the notation

$$\begin{aligned} Q_{\alpha\beta}&=\xi_{[\alpha}\zeta_{\beta]}, \quad Q_{\alpha\beta}{}^*=-\xi_{[\alpha}\rho_{\beta]},\\ P_{\alpha\beta}&=\chi_{[\alpha}\zeta_{\beta]}, \quad P_{\alpha\beta}{}^*=\chi_{[\alpha}\rho_{\beta]},\\ L_{\alpha\beta}&=\zeta_{[\alpha}\rho_{\beta]}, \quad L_{\alpha\beta}{}^*=\xi_{[\alpha}\chi_{\beta]}. \end{aligned} \tag{2.1}$$

The square bracket indicates antisymmetrization and the * indicates the dual defined

$$B_{[\alpha\beta]}{}^*=\tfrac{1}{2}(-g)^{\frac{1}{2}}\epsilon_{\alpha\beta\gamma\delta}B^{[\gamma\delta]}.$$

The dual has the property

$$B_{[\alpha\beta]}{}^{**}=-B_{[\alpha\beta]}. \tag{2.2}$$

The Q, P, L and the duals are called bivectors. On taking all quadratic products of the bivectors, as for example $Q_{\alpha\beta}Q_{\gamma\delta}{}^*$, we construct four-index tensors with the algebraic properties of the Riemann tensor. An example of this would be $Q_{\alpha\beta}Q_{\gamma\delta}{}^*+Q_{\alpha\beta}{}^*Q_{\gamma\delta}$, it is antisymmetric in α,β and γ,δ and symmetric with respect to interchange of the pairs $\alpha\beta$, $\gamma\delta$. (The cyclic symmetry is automatically satisfied by the next condition.) If we now restrict the choice of these four-index tensors by demanding that they have a vanishing "Ricci" tensor (the "Ricci" tensor is defined by summing the four-index tensor over the first and last indices), this leaves just the following 10 tensors, which form a basis for algebraically possible Riemann tensors in empty space:

$$R^{\mathrm{I}}{}_{\alpha\beta\gamma\delta}=Q_{\alpha\beta}Q_{\gamma\delta}-Q_{\alpha\beta}{}^*Q_{\gamma\delta}{}^*, \tag{2.3a}$$

$$R^{\mathrm{II}}{}_{\alpha\beta\gamma\delta}=P_{\alpha\beta}P_{\gamma\delta}-P_{\alpha\beta}{}^*P_{\gamma\delta}{}^*, \tag{2.3b}$$

* This research was supported in part by a contract with the Air Research and Development Command, U. S. Air Force.

[1] A. Z. Petrov, Sci. Not. Kazan State Univ. **114**, 55 (1954).

[2] F. A. E. Pirani, Phys. Rev. **105**, 1089 (1957).

[3] H. Bondi, F. A. E. Pirani, and I. Robinson, Proc. Roy. Soc. (London) **A251**, 519 (1959).

[4] I. Robinson, report to the Royaumont Conference.

$$R^{\mathrm{III}}{}_{\alpha\beta\gamma\delta}=(Q_{\alpha\beta}L_{\gamma\delta}+L_{\alpha\beta}Q_{\gamma\delta})-(Q_{\alpha\beta}{}^{*}L_{\gamma\delta}{}^{*}+L_{\alpha\beta}{}^{*}Q_{\gamma\delta}{}^{*}), \quad (2.3c)$$

$$R^{\mathrm{IV}}{}_{\alpha\beta\gamma\delta}=(L_{\alpha\beta}P_{\gamma\delta}+P_{\alpha\beta}L_{\gamma\delta})-(P_{\alpha\beta}{}^{*}L_{\gamma\delta}{}^{*}+L_{\alpha\beta}{}^{*}P_{\gamma\delta}{}^{*}), \quad (2.3d)$$

$$R^{\mathrm{V}}{}_{\alpha\beta\gamma\delta}=(Q_{\alpha\beta}P_{\gamma\delta}+P_{\alpha\beta}Q_{\gamma\delta})-(Q_{\alpha\beta}{}^{*}P_{\gamma\delta}{}^{*}+P_{\alpha\beta}{}^{*}Q_{\gamma\delta}{}^{*})+2(L_{\alpha\beta}L_{\gamma\delta}-L_{\alpha\beta}{}^{*}L_{\gamma\delta}{}^{*}). \quad (2.3e)$$

The remaining five basic tensors are just the duals of the first five, the dual being defined as

$$R_{\alpha\beta\gamma\delta}{}^{*}=(-g)^{\frac{1}{2}}/2\epsilon_{\gamma\delta\mu\nu}R_{\alpha\beta}{}^{\mu\nu}. \quad (2.4)$$

From these tensors we can construct examples of the Petrov types as follows:

Petrov type I:

$$R_{\alpha\beta\gamma\delta}=\alpha_1 R^{\mathrm{V}}{}_{\alpha\beta\gamma\delta}+\alpha_2(R^{\mathrm{I}}{}_{\alpha\beta\gamma\delta}+R^{\mathrm{II}}{}_{\alpha\beta\gamma\delta})+\beta_1 R^{\mathrm{V}}{}_{\alpha\beta\gamma\delta}{}^{*}+\beta_2(R^{\mathrm{I}}{}_{\alpha\beta\gamma\delta}{}^{*}+R^{\mathrm{II}}{}_{\alpha\beta\gamma\delta}{}^{*}). \quad (2.5a)$$

Petrov type II:

$$R_{\alpha\beta\gamma\delta}=\sigma R^{\mathrm{I}}{}_{\alpha\beta\gamma\delta}+\alpha R^{\mathrm{V}}{}_{\alpha\beta\gamma\delta}+\beta R^{\mathrm{V}}{}_{\alpha\beta\gamma\delta}{}^{*}. \quad (2.5b)$$

Petrov type III:

$$R_{\alpha\beta\gamma\delta}=\sigma R^{\mathrm{III}}{}_{\alpha\beta\gamma\delta}. \quad (2.5c)$$

An important subclass is type II null defined by

$$R_{\alpha\beta\gamma\delta}=\sigma R^{\mathrm{I}}{}_{\alpha\beta\gamma\delta}. \quad (2.6)$$

The α's and β's are invariants of the Riemann tensor and σ is nonzero, but otherwise arbitrary.

Pirani calls the Petrov types II and III the radiation cases.

III. DIFFERENTIAL PROPERTIES

By differentiating the 10 possible Riemann tensors we can, after a lengthy calculation, obtain 10 relations, consisting of

$$R^{\mathrm{I}}{}_{\alpha\beta\gamma\delta;\mu}=2R^{\mathrm{I}}{}_{\alpha\beta\gamma\delta}a_\mu+2R^{\mathrm{I}}{}_{\alpha\beta\gamma\delta}{}^{*}f_\mu+R^{\mathrm{III}}{}_{\alpha\beta\gamma\delta}c_\mu-R^{\mathrm{III}}{}_{\alpha\beta\gamma\delta}{}^{*}b_\mu, \quad (3.1a)$$

$$R^{\mathrm{II}}{}_{\alpha\beta\gamma\delta;\mu}=-2R^{\mathrm{II}}{}_{\alpha\beta\gamma\delta}a_\mu-2R^{\mathrm{II}}{}_{\alpha\beta\gamma\delta}{}^{*}f_\mu+R^{\mathrm{IV}}{}_{\alpha\beta\gamma\delta}e_\mu+R^{\mathrm{IV}}{}_{\alpha\beta\gamma\delta}{}^{*}d_\mu, \quad (3.1b)$$

$$R^{\mathrm{III}}{}_{\alpha\beta\gamma\delta;\mu}=2R^{\mathrm{I}}{}_{\alpha\beta\gamma\delta}e_\mu+2R^{\mathrm{I}}{}_{\alpha\beta\gamma\delta}{}^{*}d_\mu+R^{\mathrm{III}}{}_{\alpha\beta\gamma\delta}a_\mu+R^{\mathrm{III}}{}_{\alpha\beta\gamma\delta}{}^{*}f_\mu+R^{\mathrm{V}}{}_{\alpha\beta\gamma\delta}c_\mu-R^{\mathrm{V}}{}_{\alpha\beta\gamma\delta}{}^{*}b_\mu, \quad (3.1c)$$

$$R^{\mathrm{IV}}{}_{\alpha\beta\gamma\delta;\mu}=2R^{\mathrm{II}}{}_{\alpha\beta\gamma\delta}c_\mu-2R^{\mathrm{II}}{}_{\alpha\beta\gamma\delta}{}^{*}b_\mu-R^{\mathrm{IV}}{}_{\alpha\beta\gamma\delta}a_\mu-R^{\mathrm{IV}}{}_{\alpha\beta\gamma\delta}{}^{*}f_\mu+R^{\mathrm{V}}{}_{\alpha\beta\gamma\delta}e_\mu+R^{\mathrm{V}}{}_{\alpha\beta\gamma\delta}{}^{*}d_\mu, \quad (3.1d)$$

$$R^{\mathrm{V}}{}_{\alpha\beta\gamma\delta;\mu}=3(R^{\mathrm{III}}{}_{\alpha\beta\gamma\delta}e_\mu+R^{\mathrm{III}}{}_{\alpha\beta\gamma\delta}{}^{*}d_\mu+R^{\mathrm{IV}}{}_{\alpha\beta\gamma\delta}c_\mu-R^{\mathrm{IV}}{}_{\alpha\beta\gamma\delta}{}^{*}b_\mu), \quad (3.1e)$$

and their duals.

The vector coefficients are defined as

$$a_\mu=\xi_{\nu;\mu}\chi^\nu, \quad d_\mu=\chi_{\nu;\mu}\zeta^\nu,$$
$$b_\mu=\xi_{\nu;\mu}\zeta^\nu, \quad e_\mu=\chi_{\nu;\mu}\rho^\nu,$$
$$c_\mu=\xi_{\nu;\mu}\rho^\nu, \quad f_\mu=\zeta_{\nu;\mu}\rho^\nu.$$

It is possible to give a subclassification[5] of the "Riemann" tensors by stating that the different vectors $a_\mu\cdots, f_\mu$ are zero or nonzero. That this is fruitful will be demonstrated in Sec. V.

IV. BIANCHI IDENTITIES

Up to this point, we have been considering a tensor with the symmetries of the Riemann tensor. A necessary condition, but by no means sufficient condition, that this tensor be the Riemann tensor in a Riemann space is that it satisfies the Bianchi identities,

$$R^{\alpha}{}_{\beta\gamma\delta;\mu}+R^{\alpha}{}_{\beta\mu\gamma;\delta}+R^{\alpha}{}_{\beta\delta\mu;\gamma}=0. \quad (4.1)$$

By summing over α and μ and remembering that $R_{\beta\gamma}=0$, we obtain

$$R^{\alpha}{}_{\beta\gamma\delta;\alpha}=0. \quad (4.2)$$

By applying this equation to Petrov type I degenerate, types II and III, Sachs[6] and others have shown that the ξ^μ field is tangent to a null geodesic congruence and satisfies the nonshearing condition,

$$\xi^{\alpha;\beta}(\xi_{\alpha;\beta}+\xi_{\beta;\alpha})=(\xi^{\mu}{}_{;\mu})^2, \quad (4.3)$$

which was first derived by Robinson[4] for type II null solutions (and for null solutions of Maxwell's empty-space equations).

V. PLANE AND PLANE-FRONTED WAVES

For the remainder of the paper we will restrict ourselves to Petrov type II null, and hence our Riemann tensor may be written as $R_{\alpha\beta\gamma\delta}=\sigma R^{\mathrm{I}}{}_{\alpha\beta\gamma\delta}$, Eq. (2.6), with the differential properties given by Eq. (3.1a). We now restrict ourselves further and require that $c_\mu=b_\mu=0$. The differential properties of the Riemann tensor now become

$$R_{\alpha\beta\gamma\delta;\mu}=(\log\sigma)_{,\mu}R_{\alpha\beta\gamma\delta}+2R_{\alpha\beta\gamma\delta}a_\mu+2R_{\alpha\beta\gamma\delta}{}^{*}f_\mu$$

or

$$R_{\alpha\beta\gamma\delta;\mu}=R_{\alpha\beta\gamma\delta}(\log\sigma_{,\mu}+2a_\mu)+2R_{\alpha\beta\gamma\delta}{}^{*}f_\mu. \quad (5.1)$$

The first theorem we wish to prove is that a space whose Riemann tensor has the foregoing properties, possesses a covariantly constant null vector field.

It is easily verified from (3.2) that

$$\xi_{\mu;\nu}=\xi_\mu a_\nu-\zeta_\mu b_\nu-\rho_\mu c_\nu \quad (5.2)$$

and

$$\chi_{\mu;\nu}=-\chi_\mu a_\nu-\zeta_\mu d_\nu-\rho_\mu e_\nu. \quad (5.3)$$

We define a new vector λ_μ by $\lambda_\mu=\phi\xi_\mu$, where ϕ is a scalar to be determined. The covariant derivative of λ_μ is

$$\lambda_{\mu;\nu}=\phi_{,\nu}\xi_\mu+\phi\xi_{\mu;\nu}=\phi_{,\nu}\xi_\mu+\phi\xi_\mu a_\nu=\xi_\mu(\phi_{,\nu}+\phi a_\nu) \quad (5.4)$$

from Eq. (5.2) and remembering $b_\mu=c_\mu=0$.

[5] This subclassification is similar to that proposed by J. Schell, who arrived at it by a study of the generators of the holonomy group.

[6] R. Sachs (private communication and preprints).

We now must show that a_ν can be expressed as the gradient of a scalar.

Consider the expression

$$\begin{aligned} a_{\alpha;\beta}-a_{\beta;\alpha}&=(\xi_{\nu;\alpha}\chi^\nu)_{;\beta}-(\xi_{\nu;\beta}\chi^\nu)_{;\alpha}\\ &=(\xi_{\nu;\alpha\beta}-\xi_{\nu;\beta\alpha})\chi^\nu+\xi_{\nu;\alpha}\chi^\nu{}_{;\beta}-\xi_{\nu;\beta}\chi^\nu{}_{;\alpha}\\ &=\xi_\mu R^\mu{}_{\nu\alpha\beta}\chi^\nu-a_\alpha a_\beta+a_\beta a_\alpha. \end{aligned} \tag{5.5}$$

The last two terms arise from (5.2) and (5.3). Direct calculation of $\xi_\mu R^\mu{}_{\nu\alpha\beta}$ from Eq. (2.3a) shows it to be zero. Hence,

$$a_{\alpha;\beta}-a_{\beta;\alpha}=0 \quad \text{and} \quad a_\nu=\psi_{,\nu}. \tag{5.6}$$

By choosing ϕ in (5.4) as $\phi=e^{-\psi}$, we obtain $\lambda_{\mu;\nu}=0$, proving the theorem.

We will now find the most general line element which satisfies Eq. (5.1). The method used is based on an elegant paper by Walker.[7]

From a theorem of Eisenhart,[8] one can put the metric of any four-dimensional space-time that possesses a covariantly constant vector field into the form

$$\begin{aligned} ds^2=g_{00}(dx^0)^2+2g_{0r}dx^0dx^r+g_{rs}dx^rdx^s\\ +2dx^0dx^1 \quad (r,s=2,3), \end{aligned} \tag{5.7}$$

the g's being independent of x^1. In addition, the Eisenhart theorem states $\lambda^\mu=\delta_1{}^\mu$. For the remainder of the paper r, s, t $\cdots$ will go from 2 to 3.

It is easy to see from (5.7) and $\lambda^\mu=\delta_1{}^\mu$ that $\lambda_\mu=\delta_\mu{}^0$, and hence $\xi_\mu=e^\psi\delta_\mu{}^0$. The Riemann tensor we are considering then takes the form

$$R_{\alpha\beta\gamma\delta}=\sigma e^{2\psi}(\delta^0{}_{[\alpha}\zeta_{\beta]}\delta^0{}_{[\gamma}\zeta_{\delta]}-\delta^0{}_{[\alpha}\rho_{\beta]}\delta^0{}_{[\gamma}\rho_{\delta]}), \tag{5.8}$$

with the consequence that the only nonvanishing components are those with two indices equal to zero. By direct calculation using the metric (5.7) we can calculate the 2323 component of the Riemann tensor with the result

$$R_{rstu}=\bar{R}_{rstu}=0. \tag{5.9}$$

$\bar{R}_{rstu}$ is the Riemann tensor in the two-dimensional space defined by the metric $d\bar{s}=g_{rs}dx^rdx^s$. Since the two-dimensional space is flat we can take the metric as $g_{rs}=-\delta_{rs}$. (The minus sign is due to our choice of the signature of the metric.)

From $R_{0rst}=0$, Walker shows that there exists a coordinate transformation which causes the g_{0r} to vanish.

With these two facts our metric (5.7) can then be written

$$ds^2=g_{00}(dx^0)^2-\delta_{rs}dx^rdx^s+2dx^0dx^1, \tag{5.10}$$

and

$$\begin{aligned} g^{00}=0, \quad g^{11}=-g_{00}, \quad g^{rs}=\delta^{rs}, \quad g^{01}=1,\\ g^{0s}=g^{1s}=0. \end{aligned} \tag{5.11}$$

The only nonvanishing components of the affine connection are

$$\left\{\begin{matrix} r \\ 00 \end{matrix}\right\}=-\tfrac{1}{2}g_{00,r}, \quad \left\{\begin{matrix} 1 \\ 00 \end{matrix}\right\}=\tfrac{1}{2}g_{00,0}, \quad \left\{\begin{matrix} 1 \\ 0s \end{matrix}\right\}=\tfrac{1}{2}g_{00,s}. \tag{5.12}$$

The components of the Ricci tensor then are

$$R_{rs}=R_{1s}=R_{0s}=R_{11}\equiv 0, \quad R_{00}=\tfrac{1}{2}g_{00,rr}.$$

The empty-space field equations are

$$g_{00,rr}=0. \tag{5.13}$$

Though it is not obvious, this solution (5.10) and (5.13) satisfies Eq. (5.1) identically and hence solves our problem completely. It is instructive to see how this is done.

Any metric tensor can be written in terms of the "vierbein" vectors as

$$g_{\mu\nu}=\xi_\mu\chi_\mu+\chi_\mu\xi_\nu-\zeta_\mu\zeta_\nu-\rho_\mu\rho_\nu. \tag{5.14}$$

If the metric is (5.10), the vierbein vectors may be chosen as

$$\begin{aligned} &\xi_\mu=e^\psi\delta^0{}_\mu, \quad \chi_\mu=e^{-\psi}[\delta^0{}_\mu(g_{00}/2)+\delta^1{}_\mu],\\ &\zeta_\mu=\cos\theta\delta^2{}_\mu+\sin\theta\delta^3{}_\mu,\\ &\rho_\mu=-\sin\theta\delta^2{}_\mu+\cos\theta\delta^3{}_\mu. \end{aligned} \tag{5.15}$$

$\delta^2{}_\mu$ and $\delta^3{}_\mu$ are unit vectors along the x^2 and x^3 axis and at this point θ is an arbitrary angle. The vectors a_μ and f_μ computed from their definition, Eq. (3.2), are

$$a_\mu=\psi_{,u}, \quad f_\mu=\theta_{,\mu}. \tag{5.16}$$

On using (5.15) in (5.8) we have the only independent nonvanishing components of the Riemann tensor

$$\begin{aligned} R_{0220}=-\sigma e^{2\psi}\cos2\theta, \quad R_{0330}=\sigma e^{2\psi}\cos2\theta,\\ R_{0230}=-\sigma e^{2\psi}\sin2\theta. \end{aligned} \tag{5.17}$$

On calculating the same components, this time from the metric tensor, we obtain

$$\begin{aligned} R_{0220}=g_{00,22}/2, \quad R_{0330}=g_{00,33}/2,\\ R_{0230}=g_{00,23}/2. \end{aligned} \tag{5.18}$$

Therefore,

$$g_{00,22}=-2\sigma e^{2\psi}\cos2\theta, \quad g_{00,23}=-2\sigma e^{2\psi}\sin2\theta \tag{5.19}$$

and

$$\sigma=-\frac{g_{00,22}e^{-2\psi}}{2\cos2\theta} \quad \text{or} \quad \sigma=-\frac{g_{00,23}e^{-2\psi}}{2\sin2\theta}. \tag{5.20}$$

From (5.20) we obtain

$$(\log\sigma)_{,\mu}=-2\psi_{,\mu}+\frac{g_{00,22\mu}}{g_{00,22}}+2(\tan2\theta)\theta_{,\mu}. \tag{5.21}$$

From (5.19) we also obtain

$$g_{00,23}/g_{00,22}=\tan2\theta. \tag{5.22}$$

[7] A. G. Walker, Proc. London Math. Soc. Ser. 2, **V52** (1951).
[8] L. Eisenhart, Ann. Math. **39**, 316 (1938).

On substituting (5.16), (5.18), and (5.21) into (5.1) and with the aid of (5.22), we have (5.1) identically satisfied. Though this completes our proof, a discussion of (5.22) sheds light on the nature of some particular solutions.

If $\tan\theta=\text{constant}=c$, our field equations are $g_{00,22}+g_{00,33}=0$ and $g_{00,23}=cg_{00,22}$. The solution to this is

$$g_{00}=A(x^0)(x_2^2-x_3^2+cx_2x_3).$$

There exists a simple coordinate transformation which eliminates the cross term. We can call this solution a plane-polarized wave, with θ as the angle of polarization.

If $\tan\theta=f(x^0)$, our general solution is

$$g_{00}=A(x^0)(x_2^2-x_3^2)+B(x^0)x_2x_3.$$

The angle of polarization changes with x^0 but is independent of the spatial coordinates.

The most general solution has the angle of polarization as a function of x_0, x_2, and x_3.

The first two solutions, whose properties are discussed in great detail in footnote 3, are referred to as plane waves, while the last given in footnote 4 is called a plane-fronted wave.

ACKNOWLEDGMENT

The author would like to express his appreciation to Dr. H. Bondi and Dr. F. Pirani for their hospitality at Kings College where this work was begun and to Dr. Synge at the Dublin Institute for Advanced Study where a large part of the work was completed.

In particular, he would like to thank Dr. Pirani for the many discussions which led to this problem and Dr. I. Robinson for reading the manuscript and making many valuable suggestions. Many of the results of this paper were arrived at independently by Dr. Robinson in unpublished works.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 2, NUMBER 3 MAY-JUNE, 1961

Some Applications of the Infinitesimal-Holonomy Group to the Petrov Classification of Einstein Spaces

J. N. GOLDBERG AND R. P. KERR
Aeronautical Research Laboratories, Wright-Patterson Air Force Base, Ohio

(Received September 15, 1960)

The classifications of Einstein spaces by Schell and Petrov are combined and certain nonlocal results are obtained. In particular, we show that an Einstein space cannot be type I with a rank four Riemann tensor in a four-dimensional region. On using the notion of a perfect or imperfect infinitesimal-holonomy group, we establish the conditions under which an Einstein space possesses a two-, four-, or six-parameter group. We find that two- and four-parameter groups are associated with special cases of type II null and type III, respectively.

1. INTRODUCTION

RECENTLY, Schell[1] has presented a classification of Riemannian spaces in terms of the bivector generators of the infinitesimal-holonomy group (ihg). He has shown that this approach is closely related to an earlier classification due to Petrov.[2] For empty Einstein spaces this relationship is also revealed by the work of Newman[3] who described the Petrov classes in terms of certain basis tensors of rank four with the symmetries of the curvature tensor.

In this paper we restrict our attention to empty Einstein spaces, $R_{\mu\nu}=0$. We study the three Petrov classes using additional information given by the differential properties of the curvature tensor, that is, the ihg. We shall write the Petrov classes in terms of their canonical bivectors. These bivectors always form part of a basis for the ihg and, indeed, as we shall see, always generate a subgroup of it. For the most general spaces this group by itself does not give much information. However, for type II and type III spaces the classification according to the ihg allows a further subdivision which depends on the dimension of this group. These types are of particular interest since Pirani[4] has defined pure gravitational radiation in terms of them.

Many of the ideas and results presented in this paper appear in Newman's work.[3] However, his language and notation are not suited to a discussion of the holonomy group. The motivation for this paper is derived from the classification by Schell and, therefore, the results are stated with respect to the ihg and give information about the structure of this group in empty Einstein space.

In the next section we shall discuss the holonomy group briefly. Following this, we examine the individual Petrov types.

[1] J. Schell (preprint).
[2] A. Z. Petrov, Sci. Not. Kazan State. Univ. **114**, 55 (1954).
[3] E. Newman (preprint).
[4] F. A. E. Pirani, Phys. Rev. **105**, 1089 (1957).

On substituting (5.16), (5.18), and (5.21) into (5.1) and with the aid of (5.22), we have (5.1) identically satisfied. Though this completes our proof, a discussion of (5.22) sheds light on the nature of some particular solutions.

If $\tan\theta=\text{constant}=c$, our field equations are $g_{00,22}+g_{00,33}=0$ and $g_{00,23}=cg_{00,22}$. The solution to this is

$$g_{00}=A(x^0)(x_2^2-x_3^2+cx_2x_3).$$

There exists a simple coordinate transformation which eliminates the cross term. We can call this solution a plane-polarized wave, with θ as the angle of polarization.

If $\tan\theta=f(x^0)$, our general solution is

$$g_{00}=A(x^0)(x_2^2-x_3^2)+B(x^0)x_2x_3.$$

The angle of polarization changes with x^0 but is independent of the spatial coordinates.

The most general solution has the angle of polarization as a function of x_0, x_2, and x_3.

The first two solutions, whose properties are discussed in great detail in footnote 3, are referred to as plane waves, while the last given in footnote 4 is called a plane-fronted wave.

ACKNOWLEDGMENT

The author would like to express his appreciation to Dr. H. Bondi and Dr. F. Pirani for their hospitality at Kings College where this work was begun and to Dr. Synge at the Dublin Institute for Advanced Study where a large part of the work was completed.

In particular, he would like to thank Dr. Pirani for the many discussions which led to this problem and Dr. I. Robinson for reading the manuscript and making many valuable suggestions. Many of the results of this paper were arrived at independently by Dr. Robinson in unpublished works.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 2, NUMBER 3 MAY-JUNE, 1961

Some Applications of the Infinitesimal-Holonomy Group to the Petrov Classification of Einstein Spaces

J. N. GOLDBERG AND R. P. KERR
Aeronautical Research Laboratories, Wright-Patterson Air Force Base, Ohio

(Received September 15, 1960)

The classifications of Einstein spaces by Schell and Petrov are combined and certain nonlocal results are obtained. In particular, we show that an Einstein space cannot be type I with a rank four Riemann tensor in a four-dimensional region. On using the notion of a perfect or imperfect infinitesimal-holonomy group, we establish the conditions under which an Einstein space possesses a two-, four-, or six-parameter group. We find that two- and four-parameter groups are associated with special cases of type II null and type III, respectively.

1. INTRODUCTION

RECENTLY, Schell[1] has presented a classification of Riemannian spaces in terms of the bivector generators of the infinitesimal-holonomy group (ihg). He has shown that this approach is closely related to an earlier classification due to Petrov.[2] For empty Einstein spaces this relationship is also revealed by the work of Newman[3] who described the Petrov classes in terms of certain basis tensors of rank four with the symmetries of the curvature tensor.

In this paper we restrict our attention to empty Einstein spaces, $R_{\mu\nu}=0$. We study the three Petrov classes using additional information given by the differential properties of the curvature tensor, that is, the ihg. We shall write the Petrov classes in terms of their canonical bivectors. These bivectors always form part of a basis for the ihg and, indeed, as we shall see, always generate a subgroup of it. For the most general spaces this group by itself does not give much information. However, for type II and type III spaces the classification according to the ihg allows a further subdivision which depends on the dimension of this group. These types are of particular interest since Pirani[4] has defined pure gravitational radiation in terms of them.

Many of the ideas and results presented in this paper appear in Newman's work.[3] However, his language and notation are not suited to a discussion of the holonomy group. The motivation for this paper is derived from the classification by Schell and, therefore, the results are stated with respect to the ihg and give information about the structure of this group in empty Einstein space.

In the next section we shall discuss the holonomy group briefly. Following this, we examine the individual Petrov types.

[1] J. Schell (preprint).
[2] A. Z. Petrov, Sci. Not. Kazan State. Univ. **114**, 55 (1954).
[3] E. Newman (preprint).
[4] F. A. E. Pirani, Phys. Rev. **105**, 1089 (1957).

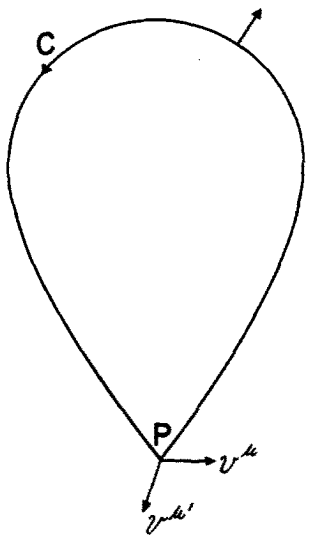

FIG. 1. The vector $v^{\mu'}$ obtained by parallel transfer of v^μ around the curve C will in general differ from v^μ by a rigid rotation.

2. HOLONOMY GROUP[5]

In a Riemannian space V_4 with signature $+2$, consider a closed path C starting and ending at the point P (Fig. 1). If an arbitrary vector v^μ is parallel transferred around C, then upon return to P it will, in general, differ from its original value. This operation of parallel transference around C induces a linear transformation, T_c, of the tangent space at P[6]:

$$v'^\mu = T_c{}^\mu{}_\nu v^\nu.$$

Since parallel transfer preserves all inner products, T_c will be an element of the rotation group at P defined by

$$T_c{}^\alpha{}_\mu g_{\alpha\beta} T_c{}^\beta{}_\nu = g_{\mu\nu}.$$

The linear transformations generated by the set of all such paths through P forms a subgroup of the rotation group at P, and is called the holonomy group. If we restrict ourselves to regions in which the connection is analytic, and if we further restrict ourselves to curves which are homotopic to zero, then we obtain the infinitesimal-holonomy group (ihg). Since the groups at P and P' are isomorphic, there is, in fact, only one ihg for the domain. Nijenhuis[7] has shown that the ihg is a Lie group, and that its generators, $L_{\mathrm{A}\mu}{}^\nu$, span the λ^ρ domain of the curvature tensor, $R_{\mu\nu\lambda}{}^\rho$, and its covariant derivatives. Latin letters, A..., run from 1 to r, where r is the dimension of the group space. If we lower ν with the metric tensor, we obtain the *generating bivectors* $L_{\mathrm{A}\mu\nu}$. The curvature tensor and its derivatives can be expressed *uniquely* in terms of these,

$$\nabla_\mu \cdots \nabla_\nu R_{\rho\sigma\lambda\tau} = h^{\mathrm{AK}}{}_{\mu\ldots\nu} L_{\mathrm{A}\rho\sigma} L_{\mathrm{K}\lambda\tau}, \tag{1}$$

where $h^{\mathrm{AK}}\cdots$ is symmetric in A and K. It may be seen from (1) that the covariant derivatives of the bivectors themselves are expressible as a linear combination of the basis bivectors,

$$\nabla_\mu L_{\mathrm{A}\rho\sigma} = \mathcal{C}^{\mathrm{K}}{}_{\mathrm{A}\mu} L_{\mathrm{K}\rho\sigma}, \tag{2}$$

with suitable coefficients, $\mathcal{C}^{\mathrm{K}}{}_{\mathrm{A}\mu}$.

If the rank of the matrix h^{AK}, which appears in the curvature tensor itself, is r then the ihg is called *perfect*. Conversely, if the rank is less than r then the ihg is called *imperfect*. Since the ihg is a property of a region rather than a point, it is still called perfect even if there are isolated points at which its rank is less than r.

The classification of Riemannian spaces devised by Schell depends on the number of parameters needed to describe the ihg and its degree of imperfectness. In Table I we present a comparison of his classification with that of Petrov.

3. CANONICAL FORMS FOR THE RIEMANN TENSOR

In his classification Schell has shown that the generating bivectors may be conveniently chosen to be simple bivectors constructed out of an orthogonal tetrad. Similarly, Petrov made use of such a tetrad to obtain his canonical forms for the Riemann tensor. In this section we shall show that the basis bivector generators of Schell are identical to the canonical bivectors of Petrov.

We introduce a null tetrad[8] $e_{a\mu}$,

$$\begin{aligned} e_{1\mu} &= x_\mu, \quad e_{2\mu} = y_\mu, \\ e_{3\mu} &= q_\mu, \quad e_{4\mu} = p_\mu, \end{aligned} \tag{3}$$

where p and q are null vectors, and x and y are spacelike. The orthogonality properties of the tetrad may be expressed as

$$e_{a\mu} e_b{}^\mu = \tilde{g}_{ab}, \tag{4}$$

where

$$\tilde{g}_{ab} = \tilde{g}^{ab} = \begin{bmatrix} 1 & \cdot & \cdot & \cdot \\ \cdot & 1 & \cdot & \cdot \\ \cdot & \cdot & \cdot & 1 \\ \cdot & \cdot & 1 & \cdot \end{bmatrix}. \tag{5}$$

Clearly, the $\tilde{g}_{ab}$ are the anholonomic components of the metric tensor, and they will be used to raise and lower tetrad indices. The metric tensor iteself is given by

$$g_{\mu\nu} = \tilde{g}^{ab} e_{a\mu} e_{b\nu}. \tag{6}$$

TABLE I. Comparison of the Petrov classification with that of Schell.

Number of group parameters	Bivector rank of curvature tensor	Petrov type
2	2	II null
4	4	III
6	2	II null
	4	III[a]
	6	I, II

[a] According to Schell's work, type I may also be rank 4. We show in Sec. 4 that such a solution of the field equations does not exist.

[5] J. A. Schouten, *Ricci-Calculus* (Springer-Verlag, Berlin, 1954), p. 361.

[6] Greek indices label the coordinates and run from 1 to 4.

[7] A. Nijenhuis, Koninkl. Ned. Akad. Wetenschap. Proc. Ser. A, 56, 233, 241 (1953).

[8] The Latin indices a, b, c from the early part of the alphabet run from 1 to 4 and label the different tetrad vectors.

From the tetrad (3) we construct the following six linearly independent simple bivectors:

$$L_{1\rho\sigma}=[px], \quad L_{2\rho\sigma}=[py],$$
$$L_{3\rho\sigma}=[pq], \quad L_{4\rho\sigma}=[xy], \tag{7}$$
$$L_{5\rho\sigma}=[qx], \quad L_{6\rho\sigma}=[qy].$$

In the preceding we have used the shorthand notations

$$[ab]=a_{[\rho}b_{\sigma]}=\tfrac{1}{2}(a_\rho b_\sigma-a_\sigma b_\rho).$$

One can easily show that the bivectors L_1 and L_2 generate the group of null rotations about p:

$$p_\mu'=p_\mu,$$
$$q_\mu'=q_\mu+ax_\mu+by_\mu-\tfrac{1}{2}(a^2+b^2)p_\mu,$$
$$x_\mu'=x_\mu-ap_\mu,$$
$$y_\mu'=y_\mu-bp_\mu.$$

Similarly, L_5 and L_6 generate the null rotations about q. L_3 and L_4 generate the Lorentz transformations in the (pq) plane and the rigid rotations in the (xy) plane, respectively. A Lorentz transformation in the (pq) plane only changes the scale of p and q:

$$p_\mu'=\alpha p_\mu, \quad q_\mu'=\alpha^{-1}q_\mu.$$

In addition to the subgroup generated by L_1 and L_2, we also have the four-parameter subgroup which preserves the direction of p, $\{L_1L_2L_3L_4\}$. From the canonical forms given by Petrov, the Riemann tensor for empty Einstein spaces may be written as follows:

Type I:

$$\begin{aligned}R_{\mu\nu\rho\sigma}=&(\alpha_1+\alpha_2)[L_{(1}L_{5)}+L_{(2}L_{6)}+L_3L_3-L_4L_4]\\&+(\alpha_1-\alpha_2)[L_1L_1-L_2L_2+L_5L_5-L_6L_6]\\&+(\beta_1+\beta_2)[L_{(2}L_{5)}-L_{(1}L_{6)}-2L_{(3}L_{4)}]\\&+(\beta_1-\beta_2)[L_{(1}L_{2)}-L_{(5}L_{6)}];\end{aligned} \tag{8}$$

Type II:

$$\begin{aligned}R_{\mu\nu\rho\sigma}=&2\alpha[L_{(1}L_{5)}+L_{(2}L_{6)}+L_3L_3-L_4L_4]\\&+2\beta[L_{(2}L_{5)}-L_{(1}L_{6)}-2L_{(3}L_{4)}]\\&+\sigma[L_1L_1-L_2L_2];\end{aligned} \tag{9}$$

Type III:

$$R_{\mu\nu\rho\sigma}=2\sigma[L_{(1}L_{4)}-L_{(2}L_{3)}]. \tag{10}$$

In the foregoing expressions the coordinate indices have been suppressed on the right-hand sides.

Since we are interested in properties of four-dimensional regions rather than local properties at a point, we shall restrict ourselves to neighborhoods in which the Petrov type is the same at every point. Any compact region may always be subdivided into a finite number of regions with the foregoing property, as well as a finite set of submanifolds of lower dimension on which the Petrov class is also constant. Consequently, this is not a serious restriction.

Further information about the Riemann tensor may be obtained by examining the properties of the ihg. In order to do so, we must construct the rotation coefficients and show their relationship to the coefficients $\mathfrak{C}^{K}{}_{A\mu}$ of (2). The rotation coefficients γ_{abc} are defined by the covariant derivatives of the tetrad vectors[9]:

$$\gamma_{abc}=e_a{}^\mu e_{b\mu;\rho}e_c{}^\rho. \tag{11}$$

They are skew symmetric in (a,b) because of the orthogonality relations (4). In terms of these, the matrix $\mathfrak{C}^{K}{}_{A\mu}$ is given by the following:

$$(\mathfrak{C}^{K}{}_{A\mu})=\begin{bmatrix}\gamma_{34\mu} & \gamma_{21\mu} & \gamma_{41\mu} & \gamma_{42\mu} & \cdot & \cdot\\ \gamma_{12\mu} & \gamma_{34\mu} & \gamma_{42\mu} & \gamma_{14\mu} & \cdot & \cdot\\ \gamma_{13\mu} & \gamma_{23\mu} & \cdot & \cdot & \gamma_{41\mu} & \gamma_{42\mu}\\ \gamma_{23\mu} & \gamma_{31\mu} & \cdot & \cdot & \gamma_{24\mu} & \gamma_{41\mu}\\ \cdot & \cdot & \gamma_{13\mu} & \gamma_{32\mu} & \gamma_{43\mu} & \gamma_{21\mu}\\ \cdot & \cdot & \gamma_{23\mu} & \gamma_{13\mu} & \gamma_{12\mu} & \gamma_{43\mu}\end{bmatrix}, \tag{12}$$

where

$$\gamma_{ab\mu}=\gamma_{acb}e^c{}_\mu=e_a{}^\rho e_{b\rho;\mu}.$$

From the Ricci identity,

$$\nabla_{[\nu}\nabla_{\lambda]}e_{a\mu}\equiv\tfrac{1}{2}R_{\nu\lambda\mu}{}^\rho e_{a\rho},$$

the anholonomic components of the curvature tensor may be expressed as

$$R_{abcd}=-2\gamma_{ab[c/d]}-2\gamma_{ea[c}\gamma^e{}_{|b|d]}+2\gamma_{abe}\gamma^e{}_{[cd]}. \tag{13}$$

where

$$V_{\ldots/a}=V_{\ldots;\mu}e_a{}^\mu.$$

4. INFINITESIMAL HOLONOMY GROUP FOR PETROV TYPES

Type I

According to the classification by the ihg, type I always possesses a six-parameter group. From the canonical form (8), it is evident that the curvature tensor is in general of rank six and, therefore, has a perfect ihg. When it is of rank four, the axes may always be chosen so that

$$\alpha_1+\alpha_2=\beta_1+\beta_2=0.$$

We shall now show that such a space cannot exist in a finite *four-dimensional* region. With α_1 and β_1 replaced by α and β, the curvature tensor becomes

$$\begin{aligned}R_{\mu\nu\rho\sigma}=&2\alpha[L_1L_1-L_2L_2+L_5L_5-L_6L_6]\\&+2\beta[L_{(1}L_{2)}-L_{(5}L_{6)}].\end{aligned} \tag{14}$$

If we substitute (14) into the Bianchi identities for an Einstein space,

$$R_{\mu\nu\lambda}{}^\rho{}_{;\rho}\equiv 2R_{\lambda[\mu;\nu]}=0, \tag{15}$$

and use (12), we obtain the following restrictions on the rotation coefficients:

$$\begin{aligned}&\gamma_{144}=\gamma_{244}=\gamma_{133}=\gamma_{233}=0,\\&\gamma_{241}=-\gamma_{142}, \quad \gamma_{231}=-\gamma_{132},\\&\gamma_{141}=\gamma_{242}, \quad \gamma_{131}=\gamma_{232}.\end{aligned} \tag{16}$$

[9] L. P. Eisenhart, *Riemannian Geometry* (Princeton University Press, Princeton, New Jersey, 1949), p. 96. (In a private discussion, Newman pointed out the efficacy of using the rotation coefficients to study the properties of the Riemann tensor.)

On using these equations in (13), we find that

$$R_{4141}-R_{4242}=0,$$
$$R_{4142}=0,$$

whereas, from the canonical form (14), we have

$$R_{4141}-R_{4242}=\tfrac{1}{2}\alpha,$$
$$R_{4142}=-\tfrac{1}{4}\beta.$$

Hence $\alpha=\beta=0$, and we have the following theorem:

Theorem 1. *There does not exist an empty Einstein space whose Riemann tensor is type I rank four in a finite four-dimensional region.* This shows that a type I curvature tensor always generates a *perfect* six-parameter ihg, except for the trivial case where the curvature tensor is zero and the space is flat.

Type I metrics represent the most general class of Einstein spaces and the Bianchi identities will not lead to algebraic resistrictions on the rotation coefficients, in general. Rather, they lead to equations defining the derivatives of the parameters, $\alpha_1, \alpha_2, \beta_1, \beta_2$. However, in the special degenerate case when $\alpha_1-\alpha_2=\beta_1=\beta_2=0$, there are a sufficient number of conditions imposed on the rotation coefficients to allow the metric to be diagonalized.[10]

Type II

The curvature tensor for type II nonnull is rank six and hence generates a six-parameter perfect ihg. From the Bianchi identities we obtain the following relations for the rotation coefficients:

$$\gamma_{144}=\gamma_{244}=0,$$
$$\gamma_{142}=-\gamma_{241}, \tag{17}$$
$$\gamma_{141}=\gamma_{242}.$$

The other equations are restrictions on the derivatives of α, β, and σ.

The first two of Eqs. (17) tell us that

$$p_{\mu;\nu}p^{\nu}=\gamma_{344}p_{\mu}.$$

Since a scale change in p and q,

$$p_{\mu}'=\exp\{\varphi\}p_{\mu},$$
$$q_{\mu}'=\exp\{-\varphi\}q_{\mu}, \tag{18}$$

leaves the orthogonality relations for the tetrad (4) invariant and only changes the Riemann tensor by

$$\sigma'=\exp\{-2\varphi\}\sigma,$$

we may choose φ so that

$$\gamma_{344}'=\varphi_{1\nu}p^{\nu}+\gamma_{344}=0;$$

p will then be geodesic,

$$p_{\mu;\nu}p^{\nu}=0. \tag{19}$$

It may be easily shown from (18) and (19) that p satisfies Robinson's equation,

$$(p_{(\mu;\nu)}p^{\mu;\nu})=\tfrac{1}{2}(p^{\mu}{}_{;\mu})^2.$$

[10] E. Newman and L. Tamborino (private communication).

Sachs has previously shown that this equation holds for type II null and type III metrics.[11]

For type II null $\alpha=\beta=0$, and the curvature tensor is given by

$$R_{\lambda\mu\nu\rho}=\sigma(L_{1\lambda\mu}L_{1\nu\rho}-L_{2\lambda\mu}L_{2\nu\rho}). \tag{20}$$

From Eqs. (2) and (12) we see that the first covariant derivative of the Riemann tensor will introduce new bivectors, unless

$$\gamma_{41a}=\gamma_{42a}=0. \tag{21}$$

Furthermore, it must bring in both L_3 and L_4, since otherwise we must have

$$\begin{vmatrix} \gamma_{41\mu} & \gamma_{42\mu} \\ -\gamma_{42\mu} & \gamma_{41\mu} \end{vmatrix}=0$$

for all μ. This implies (21). From (12) we also see that the covariant derivatives of L_3 and L_4 will bring in the remaining two bivectors, L_5 and L_6, unless (21) is satisfied. Consequently, the ihg associated with a type II null curvature tensor is either two-parameter perfect [when (21) is satisfied] or six-parameter imperfect.

In addition to (17) the Bianchi identities for (20) give the following conditions:

$$2\gamma_{124}=\gamma_{142}, \tag{22}$$
$$(\sigma^2p^{\mu})_{;\mu}=0. \tag{23}$$

Equation (23) is the conservation law found earlier by Ehlers and Sachs.[12] It is only true when we choose the noninvariant parameter σ so that p is geodesic (19).

When (21) holds, we have

$$p_{\mu;\nu}=p_{\mu}a_{\nu}. \tag{24}$$

From the Ricci identity, we have

$$p_{\mu;[\nu\lambda]}=p_{\mu}a_{[\nu,\lambda]}=\tfrac{1}{2}R_{\lambda\nu\mu}{}^{\rho}p_{\rho}=0.$$

It follows that a_{ν} is a gradient field $\phi_{,\nu}$, and so $\exp(-\phi)p_{\mu}$ is covariant constant. Hence we may choose p to be covariant constant itself,

$$p_{\mu;\nu}=0. \tag{25}$$

Theorem 2. *The necessary and sufficient condition for an empty Einstein space to have a two-parameter perfect ihg is that there exist a covariant constant null vector. In that case the curvature tensor is necessarily type II null.*

The necessity has been shown previously and the sufficiency proof is trivial since (25) implies (20) and (21).

The metric which gives rise to (25) is well known as the plane fronted wave solution[3,13,14]:

$$ds^2=dx^2+dy^2+2dpdq+Hdp^2,$$
$$H_{,q}=0, \quad H_{,xx}+H_{,yy}=0.$$

[11] R. K. Sachs (preprint).
[12] J. Ehlers and R. K. Sachs, Z. Physik **155**, 498 (1959).
[13] H. W. Brinkman, Math. Ann. **94**, 119 (1925).
[14] I. Robinson and A. Trautman, Phys. Rev. Letters **4**, 431 (1960).

Type III

From (10) it is evident that the curvature tensor for a type III metric is always rank four. As for type II null, (2) and (12) show that the ihg will be six parametric unless Eq. (21) is satisfied. In addition to the relations (17), the Bianchi identities yield

$$\begin{aligned}\gamma_{121}&=2\gamma_{324}+\gamma_{243},\\ \gamma_{122}&=2\gamma_{134}-\gamma_{143},\\ \gamma_{123}&=2\gamma_{231}-2\gamma_{132},\\ \gamma_{124}&=2\gamma_{142}.\end{aligned}\tag{26}$$

Furthermore, these identities also give the conservation law already found by Sachs[15] for type III:

$$(\sigma p^{\mu})_{;\mu}=0.\tag{27}$$

It is interesting to note that in this case (27) is independent of the choice of scale for p and σ, whereas for type II null the conservation law requires that p be geodesic. Of course, since the relations (17) are true here, we may always *choose* p to be geodesic for type III as well.

In the previous section we showed that if the ihg is two-parameter perfect, there always exists a covariant constant null vector. We shall now prove the following theorem:

Theorem 3. *In an empty Einstein space possessing a four-parameter ihg, there exists a null vector satisfying*

$$p_{\mu;\nu}=\kappa p_{\mu}p_{\nu},\tag{28a}$$

$$\kappa_{,[\lambda}p_{\nu]}\neq 0,\tag{28b}$$

and the space is type III. Conversely, whenever there exists a null vector satisfying (28) the ihg is four-parameter perfect.

When the scalar κ in (28) is zero throughout a finite four-dimensional region, then p is covariant constant and theorem 2 applies. To prove the first part of the theorem, we recall that the ihg is four-parameter only when the space is type III and Eq. (21) holds. As in the discussion of type II, (21) implies that

$$p_{\mu;\nu}=p_{\mu}a_{\nu}=\gamma_{34\nu}p_{\mu}.\tag{24'}$$

If we change the scale of p as in (18), the rotation coefficients transform and we have in particular,

$$\gamma_{34a}{}'=\gamma_{34a}+\phi_{/a}.\tag{29}$$

Obviously, we cannot hope to make $\gamma_{34a}{}'$ for all a, since then $a_{\nu}{}'$ would be zero and the space would possess a covariant constant null vector. Consequently, it would be type II null. We shall look at the integrability conditions for the existence of a scalar satisfying

$$\phi_{/s}=\gamma_{43s},\quad (s,t,u,\cdots=1,2,4).\tag{30}$$

On forming the second derivatives of ϕ, we obtain the following equations:

$$\phi_{/[st]}\equiv\phi_{/u}\gamma^{u}{}_{[st]}=\gamma_{43[s/t]},$$

since (21) implies

$$\gamma^{3}{}_{[st]}=\gamma_{4[st]}=0.$$

On substituting from (30), we obtain the integrability conditions,

$$\gamma_{43[s/t]}-\gamma_{43u}\gamma^{u}{}_{[st]}=0.$$

From (13) and (21), this reduces to

$$R_{34st}=0.$$

For a type III space with curvature tensor given by (10), this condition is satisfied identically. Consequently, we may always satisfy (30), and so may choose the scalar σ so that $a_{\nu}=\kappa p_{\nu}$. Hence, (28a) is satisfied.

The second part of the theorem is proved by observing that

$$\nabla_{[\lambda}\nabla_{\nu]}p_{\mu}=p_{\mu}p_{[\nu}\nabla_{\lambda]}\kappa=\tfrac{1}{2}R_{\lambda\nu\mu}{}^{\rho}p_{\rho}.\tag{31}$$

As Sachs[15] has shown, and as may be shown by using (7) as a basis for the curvature tensor, (31) implies that the space is type III when (28b) is satisfied and type II null otherwise. In the latter case we may always choose the scale of p such that it is covariant constant. Trivially, (28) implies (21).

We can establish a relationship between the scalars κ of (28) and the σ of the curvature tensor. On substituting (10) into (31), we obtain

$$\kappa_{,[\lambda}p_{\nu]}=\tfrac{1}{4}\sigma p_{[\nu}y_{\lambda]}.$$

Hence,

$$\begin{aligned}\sigma&=4\kappa_{,\lambda}y^{\lambda}\\ \kappa_{,\lambda}x^{\lambda}&=\kappa_{,\lambda}p^{\lambda}=0.\end{aligned}$$

Equation (28) tells us that p itself is a gradient, rather than just being surface orthogonal,

$$p_{\mu}=\chi_{,\nu}.\tag{32}$$

From this fact, one can show that the only allowable transformation which preserves the form of (28) is

$$p_{\mu}{}'=\exp\{\phi(\chi)\}p_{\mu},$$

giving

$$\begin{aligned}\kappa'&=\exp\{-\phi(\chi)\}(\kappa+d\phi/d\chi),\\ \sigma'&=\exp\{-\phi(\chi)\}\sigma.\end{aligned}$$

From these equations, we shall construct the most general type III metric with a perfect ihg in a subsequent paper.

5. CONCLUSION

In this paper we have used the Petrov canonical forms to study the infinitesimal-holonomy groups for empty Einstein spaces. By restricting ourselves to regions in which the Petrov class is the same at every point, we have been able to obtain the following results:

(a) A type I metric is always rank six and, therefore, must have a six-parameter perfect ihg. It should be noted that we do not claim that the Riemann tensor

[15] R. K. Sachs, Z. Physik **157**, 462 (1960).

cannot be type I rank four on a submanifold of lower dimension.

(b) The ihg of a space is always six parameter except when there exist a null vector satisfying

$$p_{\mu;\nu}=p_\mu a_\nu.$$

(c) When a_ν is a gradient field, p may be chosen to be covariant constant. The space is type II null and the ihg is two-parameter perfect.

(d) When a_ν is not a gradient field, it may be chosen to be proportional to p_ν. The space is then type III and the ihg is four-parameter perfect.

(e) Every space with an imperfect ihg is either type II null or type III. Therefore, from the discussion following (7) we have the theorem:

Theorem 4. *When the ihg is imperfect it is always six parameter, and the bivectors spanning the* $^{\nu}{}_{\mu}$*-domain of* $R\cdots_{\lambda\rho\mu}{}^{\nu}$ *generate a nontrivial subgroup of the ihg.*

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 2, NUMBER 3 MAY-JUNE, 1961

Einstein Spaces With Four-Parameter Holonomy Groups

R. P. KERR AND J. N. GOLDBERG
Aeronautical Research Laboratories, Wright-Patterson Air Force Base, Ohio

(Received September 16, 1960)

The metric tensor is constructed for Einstein spaces which are Petrov type III and whose holonomy group is four parametric. Together with the previously known plane fronted wave solutions, this completes the study of all metrics whose holonomy groups are less than six parameter. The Killing vector equations are studied and it is found that the space cannot admit more than two independent motions.

1. INTRODUCTION

IN a previous paper[1] we combined the Petrov[2] and holonomy group classification[3] in a study of Einstein spaces. In particular we found that the infinitesimal-holonomy group (ihg) of a region is six parametric, unless the space contains a null vector p_μ satisfying

$$p_{\mu;\nu}=p_\mu a_\nu. \tag{1}$$

When a_ν is gradient forming,

$$a_\nu=\varphi_{1\nu},$$

the scale of p may be changed, $p_\mu'=e^{-\phi}p_\mu$, so that the null vector is covariant constant. The space is then type II-null. As in I we introduce the null tetrad

$$e_{1\mu}=m_\mu, \quad e_{3\mu}=q_\mu,$$
$$e_{2\mu}=n_\mu, \quad e_{4\mu}=p_\mu,$$
$$e_{a\mu}e_b{}^\mu=\tilde{g}_{ab},$$

with

$$\tilde{g}_{ab}=\tilde{g}^{ab}=\begin{bmatrix}1&.&.&.\\.&1&.&.\\.&.&.&1\\.&.&1&.\end{bmatrix}.$$

By a canonical choice of the tetrad, the type II-null Riemann tensor may be written as

$$R_{\mu\nu\rho\sigma}=4\sigma\{p_{[\mu}m_{\nu]}p_{[\rho}m_{\sigma]}-p_{[\mu}n_{\nu]}p_{[\rho}n_{\sigma]}\}. \tag{2}$$

If a_ν is not a gradient, the right-hand side of Eq. (1) may still be simplified so that

$$p_{\mu;\nu}=\kappa p_\mu p_\nu.$$

The space is then type III, four-parameter perfect, and the canonical form for the Riemann tensor is

$$R_{\mu\nu\rho\sigma}=4\sigma\{p_{[\mu}m_{\nu]}m_{[\rho}n_{\sigma]}+m_{[\mu}n_{\nu]}p_{[\rho}m_{\sigma]}-p_{[\mu}n_{\nu]}p_{[\rho}q_{\sigma]}-p_{[\mu}q_{\nu]}p_{[\rho}n_{\sigma]}\}. \tag{3}$$

We summarize these results by

$$p_{\mu;\nu}=\kappa p_\mu p_\nu\begin{cases}\kappa=0, & \text{type II null, two-parameter perfect,}\\ \kappa,_{[\lambda}p_{\mu]}\neq 0, & \text{type III, four-parameter perfect.}\end{cases} \tag{4}$$

In this paper we shall derive the metric for the most general space admitting a null vector satisfying Eq. (4). To do so we shall use the surface forming properties of the basis tetrad. Type II null will be treated as a special case of the type III metric. The surface forming properties are most easily studied by use of the rotation coefficients

$$\gamma_{abc}=e_a{}^\mu e_{b\mu;\nu}e_c{}^\nu=-\gamma_{bac}.$$

2. CONSTRUCTION OF A COORDINATE SYSTEM

From Eq. (4) we see that p is a gradient field,

$$p_{[\mu,\nu]}=0, \quad p_\mu=\psi^4{}_{,\mu}. \tag{5}$$

Since p is orthogonal to p_μ, m_μ and n_μ, ψ^4 is a solution of the *outer problem*,

$$\psi^4{}_{,\mu}p^\mu=\psi^4{}_{,\mu}m^\mu=\psi^4{}_{,\mu}n^\mu=0. \tag{6}$$

Also, from Eq. (4) we have

$$\gamma_{4ab}=\gamma^3{}_{ab}=-\delta^3{}_a\delta^3{}_b\kappa, \tag{7.a}$$

which, in conjunction with Eqs. (I.22) and (I.26) gives

$$\gamma_{124}=0. \tag{7.b}$$

[1] J. N. Goldberg and R. P. Kerr, J. Math. Phys. 2, 327 (1961), preceding paper. This paper will be referred to as I and equations in it will be denoted by (I.2), for example.
[2] A. Z. Petrov Sci. Not Kazan State Univ. 114, 55 (1954).
[3] J. F. Schell, J. Math. Phys. 2, 202 (1961).

cannot be type I rank four on a submanifold of lower dimension.

(b) The ihg of a space is always six parameter except when there exist a null vector satisfying

$$p_{\mu;\nu}=p_\mu a_\nu.$$

(c) When a_ν is a gradient field, p may be chosen to be covariant constant. The space is type II null and the ihg is two-parameter perfect.

(d) When a_ν is not a gradient field, it may be chosen to be proportional to p_ν. The space is then type III and the ihg is four-parameter perfect.

(e) Every space with an imperfect ihg is either type II null or type III. Therefore, from the discussion following (7) we have the theorem:

Theorem 4. *When the ihg is imperfect it is always six parameter, and the bivectors spanning the* $^{\nu}{}_{\mu}$*-domain of* $R\cdots_{\lambda\rho\mu}{}^{\nu}$ *generate a nontrivial subgroup of the ihg.*

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 2, NUMBER 3 MAY-JUNE, 1961

Einstein Spaces With Four-Parameter Holonomy Groups

R. P. KERR AND J. N. GOLDBERG
Aeronautical Research Laboratories, Wright-Patterson Air Force Base, Ohio

(Received September 16, 1960)

The metric tensor is constructed for Einstein spaces which are Petrov type III and whose holonomy group is four parametric. Together with the previously known plane fronted wave solutions, this completes the study of all metrics whose holonomy groups are less than six parameter. The Killing vector equations are studied and it is found that the space cannot admit more than two independent motions.

1. INTRODUCTION

IN a previous paper[1] we combined the Petrov[2] and holonomy group classification[3] in a study of Einstein spaces. In particular we found that the infinitesimal-holonomy group (ihg) of a region is six parametric, unless the space contains a null vector p_μ satisfying

$$p_{\mu;\nu}=p_\mu a_\nu. \tag{1}$$

When a_ν is gradient forming,

$$a_\nu=\varphi_{1\nu},$$

the scale of p may be changed, $p_\mu'=e^{-\phi}p_\mu$, so that the null vector is covariant constant. The space is then type II-null. As in I we introduce the null tetrad

$$e_{1\mu}=m_\mu, \quad e_{3\mu}=q_\mu,$$
$$e_{2\mu}=n_\mu, \quad e_{4\mu}=p_\mu,$$
$$e_{a\mu}e_b{}^\mu=\tilde{g}_{ab},$$

with

$$\tilde{g}_{ab}=\tilde{g}^{ab}=\begin{bmatrix}1&.&.&.\\.&1&.&.\\.&.&.&1\\.&.&1&.\end{bmatrix}.$$

By a canonical choice of the tetrad, the type II-null Riemann tensor may be written as

$$R_{\mu\nu\rho\sigma}=4\sigma\{p_{[\mu}m_{\nu]}p_{[\rho}m_{\sigma]}-p_{[\mu}n_{\nu]}p_{[\rho}n_{\sigma]}\}. \tag{2}$$

If a_ν is not a gradient, the right-hand side of Eq. (1) may still be simplified so that

$$p_{\mu;\nu}=\kappa p_\mu p_\nu.$$

The space is then type III, four-parameter perfect, and the canonical form for the Riemann tensor is

$$R_{\mu\nu\rho\sigma}=4\sigma\{p_{[\mu}m_{\nu]}m_{[\rho}n_{\sigma]}+m_{[\mu}n_{\nu]}p_{[\rho}m_{\sigma]}-p_{[\mu}n_{\nu]}p_{[\rho}q_{\sigma]}-p_{[\mu}q_{\nu]}p_{[\rho}n_{\sigma]}\}. \tag{3}$$

We summarize these results by

$$p_{\mu;\nu}=\kappa p_\mu p_\nu\begin{cases}\kappa=0, & \text{type II null, two-parameter perfect,}\\ \kappa,_{[\lambda}p_{\mu]}\neq 0, & \text{type III, four-parameter perfect.}\end{cases} \tag{4}$$

In this paper we shall derive the metric for the most general space admitting a null vector satisfying Eq. (4). To do so we shall use the surface forming properties of the basis tetrad. Type II null will be treated as a special case of the type III metric. The surface forming properties are most easily studied by use of the rotation coefficients

$$\gamma_{abc}=e_a{}^\mu e_{b\mu;\nu}e_c{}^\nu=-\gamma_{bac}.$$

2. CONSTRUCTION OF A COORDINATE SYSTEM

From Eq. (4) we see that p is a gradient field,

$$p_{[\mu,\nu]}=0, \quad p_\mu=\psi^4{}_{,\mu}. \tag{5}$$

Since p is orthogonal to p_μ, m_μ and n_μ, ψ^4 is a solution of the *outer problem*,

$$\psi^4{}_{,\mu}p^\mu=\psi^4{}_{,\mu}m^\mu=\psi^4{}_{,\mu}n^\mu=0. \tag{6}$$

Also, from Eq. (4) we have

$$\gamma_{4ab}=\gamma^3{}_{ab}=-\delta^3{}_a\delta^3{}_b\kappa, \tag{7.a}$$

which, in conjunction with Eqs. (I.22) and (I.26) gives

$$\gamma_{124}=0. \tag{7.b}$$

[1] J. N. Goldberg and R. P. Kerr, J. Math. Phys. 2, 327 (1961), preceding paper. This paper will be referred to as I and equations in it will be denoted by (I.2), for example.
[2] A. Z. Petrov Sci. Not Kazan State Univ. 114, 55 (1954).
[3] J. F. Schell, J. Math. Phys. 2, 202 (1961).

We shall now examine the integrability conditions for the existence of a gradient field satisfying

$$p^\alpha \partial_\alpha F = n^\alpha \partial_\alpha F = 0. \tag{8}$$

The first derived system of these equations consists of Eq. (8) and

$$(p^\alpha \partial_\alpha n^\beta \partial_\beta - n^\alpha \partial_\alpha p^\beta \partial_\beta) F \overset{\text{def}}{=} (\mathfrak{L}_p n^\alpha)\partial_\alpha F = 0. \tag{9}$$

From Eq. (7) we see that the Lie derivative of n with respect to p is just

$$\mathfrak{L}_p n^\alpha = n^\alpha{}_{;\rho} p^\rho - p^\alpha{}_{;\rho} n^\rho = \gamma_{324} p^\alpha.$$

Consequently, the first derived system is Eq. (8) itself and so these equations are complete. From Schouten[4] we see that there are two functionally independent solutions, one of which must be ψ^4. We shall take ψ^1 as being the other, so that the most general solution is $\phi(\psi^1,\psi^4)$. Hence, we have

$$\psi^1{}_{,\mu} p^\mu = \psi^1{}_{,\mu} n^\mu = 0. \tag{10}$$

Similarly, we can show that there is another scalar function independent of ψ^4 and which is a solution of the outer problem

$$\psi^2{}_{,\mu} p^\mu = \psi^2{}_{,\mu} m^\mu = 0. \tag{11}$$

We shall use these three functions as coordinates. As our fourth coordinate we choose a solution of

$$\psi^3{}_{,\mu} p^\mu = 1. \tag{12}$$

This equation has four independent solutions.

It may be easily seen that these four functions are functionally independent at every point of the region. If $\psi^1{}_{,\mu} m^\mu$ were zero, then by Eqs. (10) and (6) ψ^1 and ψ^4 would not be functionally independent, contrary to our assumption. A similar argument holds for $\psi^2{}_{,\mu} n^\mu$. Hence, from the preceeding equations, we have

$$\begin{aligned}\det(e_a{}^\mu \psi^b{}_{,\mu}) &= \det(e_a{}^\mu)\ \det(\psi^b{}_{,\mu}) \\ &= (\psi^1{}_{,\mu} m^\mu)(\psi^2{}_{,\mu} n^\mu) \neq 0.\end{aligned}$$

The functional independence of the ψ^a follows immediately. In the coordinate system

$$(x,y,v,u) \overset{\text{def}}{=} x^\mu \overset{*}{=} \delta^\mu{}_a \psi^a,$$

the tetrad vectors have the components

$$\begin{aligned} m_\mu &= (m_1,0,0,m_4), \quad p_\mu = (0,0,0,1), \\ n_\mu &= (0,n_2,0,n_4), \quad q_\mu = (q_1,q_2,1,q_4), \end{aligned} \tag{13}$$

where m_1 and n_2 are not zero. Since

$$g_{\mu\nu} = m_\mu m_\nu + n_\mu n_\nu + 2q_{(\mu} p_{\nu)},$$

the line element takes the form

$$\begin{aligned} ds^2 = g_{11}dx^2 + g_{22}dy^2 + 2g_{41}dxdu \\ + 2g_{42}dydu + 2dudv + g_{44}du^2. \end{aligned} \tag{14}$$

From Eq. (4) we see that

$$\{\mu\nu,3\} = \Gamma_{\mu\nu}{}^4 = -\kappa \delta_\nu{}^4 \delta_\mu{}^4, \tag{15}$$

since $g^{4\mu} = \delta_3{}^\mu$. From this it follows that

$$\begin{aligned} &g_{\mu\nu,3} = 0, \quad (\mu\nu) \neq (44), \\ &g_{44,3} = 2\kappa. \end{aligned} \tag{16}$$

Also, in I we showed that

$$\kappa_{,\lambda} p^\lambda = \kappa_{,\lambda} m^\lambda = 0, \tag{17}$$

and hence κ itself is independent of v, since $p^\mu = \delta_3{}^\mu$. Consequently, g_{44} may be written as

$$\begin{aligned} &g_{44} = \omega + 2v\kappa, \\ &\omega_{,3} = \kappa_{,3} = 0. \end{aligned} \tag{18}$$

This shows that v only occurs in one element of the metric tensor and even there it does so linearly.

We shall now prove that there exists a coordinate transformation of the type

$$\begin{aligned} x' &= \mathfrak{F}^{1'}(x,y,u), \quad v' = v, \\ y' &= \mathfrak{F}^{2'}(x,y,u), \quad u' = u, \end{aligned}$$

which preserves the form of the metric tensor, but gives

$$g'_{11} = g'_{22} = 1.$$

If we use $r, s, t, u \cdots = 1, 2$, and if we denote the affine connection and the Riemann tensor in the (x,y) space by a bar, $\bar{\Gamma}_{st}{}^r$, $\bar{R}_{rstu}$, then we have

$$\begin{aligned} \Gamma_{st}{}^r &= \bar{g}^{ru}\{st,u\} + g^{r3}\{st,3\} = \bar{\Gamma}_{st}{}^r, \\ \Gamma_{3t}{}^r &= \bar{g}^{ru}\{3t,u\} + g^{r3}\{3t,3\} = 0. \end{aligned}$$

Consequently,

$$\bar{R}_{rst}{}^u = 2\partial_{[r}\bar{\Gamma}_{s]t}{}^u + 2\bar{\Gamma}_{v[r}{}^u \bar{\Gamma}_{s]t}{}^v = R_{rst}{}^u = 0,$$

from Eqs. (2), (3), and (14). Clearly, this is the necessary and sufficient condition for the existence of a coordinate transformation giving the line element the form

$$\begin{aligned} ds^2 = dx^2 + dy^2 + 2\alpha dxdu + 2\beta dydu \\ + 2dvdu + (\omega + 2\kappa v)du^2. \end{aligned} \tag{19}$$

It should be noted that while p is invariant under the transformation, and consequently (16) through (18) remain satisfied, the vectors m and n no longer have the same form as in Eq. (13).

Looking at the field equations now, we find that

$$R_{14} = \tfrac{1}{2}(\beta_{,1} - \alpha_{,2})_{,2} + \kappa_{,1} = 0, \tag{20a}$$

$$R_{24} = \tfrac{1}{2}(\alpha_{,2} - \beta_{,1})_{,1} + \kappa_{,2} = 0. \tag{20b}$$

We introduce a function ρ satisfying

$$\alpha_{,2} - \beta_{,1} = \rho_{,2}, \tag{21}$$

and then Eq. (20b) gives

$$\rho_{,21} = -2\kappa_{,2}.$$

[4] J. A. Schouten, *Ricci Calculus* (Springer-Verlag, Berlin, 1954), pp. 78–83.

Hence we may integrate this equation and express K as a function of p,

$$\kappa=-\tfrac{1}{2}\rho_{,1} \tag{22}$$

A function $f(x,u)$ should have been added to the right-hand side of (24). However, it can always be absorbed into ρ, without affecting Eq. (23), which does not define ρ completely. The solution of Eq. (21) is

$$\alpha=\rho+\phi_{,1},$$
$$\beta=\phi_{,2},$$

where ρ and ϕ are functions of x, y, and u, but not v. From Eq. (20a) we see that ρ is a surface harmonic,

$$\rho_{,11}+\rho_{,22}=0. \tag{23}$$

It may be easily seen that if we transform x^u by the following

$$v'=v+\phi(x,y,u);\quad x^{\mu\prime}=x^\mu,\quad (\mu\neq 3), \tag{24}$$

then the metric becomes

$$ds^2=dx^2+dy^2+2dudv+2\rho dxdu+(\omega-\rho_{,1}v)du^2. \tag{25}$$

The only remaining nonzero components of the affine connection are

$$\Gamma_{42}{}^1=\tfrac{1}{2}\rho_{,2},\qquad \Gamma_{44}{}^1=\rho_{,4}-\tfrac{1}{2}\rho\rho_{,1}-\tfrac{1}{2}g_{44,1},$$
$$\Gamma_{41}{}^2=-\tfrac{1}{2}\rho_{,2},\qquad \Gamma_{44}{}^2=-\tfrac{1}{2}g_{44,2},$$
$$\Gamma_{11}{}^3=\rho_{,1},\qquad \Gamma_{12}{}^3=\tfrac{1}{2}\rho_{,2},$$
$$\Gamma_{41}{}^3=\tfrac{1}{2}g_{44,1},\qquad \Gamma_{42}{}^3=-\tfrac{1}{2}\rho\rho_{,2}+\tfrac{1}{2}g_{44,2},$$
$$\Gamma_{43}{}^3=-\tfrac{1}{2}\rho_{,1},\qquad \Gamma_{44}{}^4=\tfrac{1}{2}\rho_{,1},$$
$$\Gamma_{44}{}^3=\tfrac{1}{2}g_{44,4}+\tfrac{1}{2}\rho g_{44,1}+\tfrac{1}{2}\rho_{,1}(\rho^2-g_{44})-\rho\rho_{,4}.$$

On using these values, one finds that the Einstein equations reduce to R_{14} and R_{44}, and that these give

$$\rho_{,11}+\rho_{,22}=0,$$
$$\omega_{,11}+\omega_{,22}=2\rho_{,41}-2\rho\rho_{,11}-(\rho_{,1})^2+(\rho_{,2})^2. \tag{26}$$

This is as far as we have been able to reduce the metric, in general.

The coordinate system used is not unique. In fact, there are three classes of transformations which preserve the form (25) of the metric (we shall use a dot to denote differentiation with respect to u):

(a) *Translations*:

$$x'=x+a(u),\quad v'=v-\dot{a}x-\dot{b}y,$$
$$y'=y+b(u),\quad u'=u;$$

(b) *Rotations*:

$$x'=x\cos\theta+y\sin\theta,\quad v'=v+\Phi(x,y,u),$$
$$y'=y\cos\theta-x\sin\theta,\quad u'=u,$$

where $\theta=\theta(u)$ and Φ is a function of ρ and θ. This Φ corresponds to that in Eq. (24);

(c) *Scale transformations*:

$$x'=x,\quad \dot{\Phi}v'=v+x^2\ddot{\Phi},$$
$$y'=y,\quad u'=\Phi(u).$$

We shall refer to these in the following discussion.

3. DISCUSSION

From the construction of the metric, Eq. (4) is satisfied, and hence we know the space is either type II null or type III. From Eqs. (22) and (4), it will be type II null only if

$$\rho_{,11}=-\rho_{,22}=\rho_{,12}=0, \tag{27}$$

that is, if ρ is linear in x and y:

$$\rho=a(u)x+b(u)y+c(u).$$

It is fairly easy to show that when ρ has this form, it may be transformed away completely by a scale transformation and a rotation. One then obtains the plane fronted wave solution of Brinkman[5] and Robinson[6]:

$$ds^2=dx^2+dy^2+2dudv+\omega du^2$$

where ω is a surface harmonic.

With the metric in this form $\kappa=0$ and p_μ is convariant constant. Therefore, it satisfies the Killing equation,

$$\xi_{\mu;\nu}+\xi_{\nu;\mu}=0. \tag{28}$$

In general, p is the only Killing vector although in special examples there may be many more.[7]

We shall now investigate under what conditions a type III perfect metric can possess a Killing vector. It will be shown that any Killing vector present in such a space must belong to one of two distinct classes and that the space can never possess more than two independent motions. When one of these is present the u dependence may be transformed away. The other is a rotation in the (xy) plane.

When $\mu,\nu\neq 4$, Eq. (28) may be integrated directly to give the following *contravariant* components:

$$\xi^1=a^1+by,$$
$$\xi^2=a^2-bx, \tag{29}$$
$$\xi^4=a^4,$$

where a^μ and b are functions of u. The remaining Killing equations are

$$\xi^3{}_{,1}=-\dot{\xi}^1-\rho_{,\mu}\xi^\mu-\rho\dot{a}^4 \tag{30a}$$
$$\xi^3{}_{,2}=-\dot{\xi}^2-b\rho \tag{30b}$$
$$\xi^3{}_{,3}=-\dot{a}^4 \tag{30c}$$
$$\xi^3{}_{,4}=-\rho\dot{\xi}^1-g_{44}\dot{a}^4-\tfrac{1}{2}g_{44,\mu}\xi^\mu. \tag{30d}$$

[5] H. W. Brinkman, Math. Ann. **94**, 119 (1925).
[6] I. Robinson and A. Trautman, Phys. Rev. Letters 4, 431 (1960).
[7] H. Bondi, F. A. E. Pirani, and I. Robinson, Proc. Roy. Soc. (London) **A251**, 519 (1959).

By considering the integrability relations for these equations, we obtain the following conditions:

$$\rho_{,2\mu}\xi^\mu = -2\dot{b} - \dot{a}^4\rho_{,2}, \tag{31a}$$

$$\rho_{,1\mu}\xi^\mu = -\dot{a}^4(1+2\rho_{,1}). \tag{32b}$$

The integrability conditions on these give

$$\dot{a}^4\rho_{,12} = 0.$$

When $\rho_{,12}$ is zero we may easily show from (31) that a^4 is zero. Consequently, we always have

$$\dot{a}^4 = 0, \tag{32}$$

that is, a^4 is a true constant.

First of all we observe that when a^4 and b are both zero, a^1 and a^2 must also be zero. Otherwise, from Eq. (31) ρ is linear in x and y and the metric can be reduced to the type II plane-fronted wave solution. Furthermore, under these conditions Eqs. (30) give

$$\xi^3 = \xi^3(u), \quad 2\dot{\xi}^3 = \rho_{,1}\xi^3.$$

The second of these again implies that the metric is type II, or $\xi^\mu = 0$. Consequently, there does not exist a Killing vector with both a^4 and b zero.

We shall now consider the case where the space possesses a motion $\xi_{(1)}{}^\mu$ with $a^4 \neq 0$. By Eq. (32) we may choose $a^4 = 1$. It may be shown that by a translation followed by a rotation we may always transform a^1, a^2, and b to zero. In the new coordinate system we have

$$\xi^1{}_{(1)} = \xi^2{}_{(1)} = 0, \quad \xi^4{}_{(1)} = 1.$$

From Eq. (31) ρ has the form

$$\rho = f(u) + g(x,y).$$

The function $f(u)$ may be transformed away by

$$v' = v + xf(u), \quad x^{\mu'} = x^\mu, \quad \mu \neq 3,$$

and then we have $\rho_{,4} = 0$. It follows from Eqs. (30a)–(30c) that ξ^3 is a function of u alone. Equation (30d) then gives

$$\omega = \tilde{\omega}(xy) + \Phi(u)\rho_{,1} - 2\dot{\Phi}(u), \tag{33}$$

where $\dot{\Phi}(u) = \xi^3$. Under the transformation

$$v' = v - \Phi(u), \tag{34}$$

ω reduces to $\tilde{\omega}$ alone, and ξ^3 to zero. Finally, we see that the metric is independent of u, and the Killing vector, as is to be expected, is simply

$$\xi^\mu{}_{(1)} \overset{*}{=} \delta^\mu{}_4, \quad g_{\mu\nu,4} = 0. \tag{35}$$

We shall now examine the case where the space possesses a motion $\xi^\mu{}_{(2)}$ with

$$\xi^\mu{}_{(2)}: \quad a^4 = 0, \quad b \neq 0.$$

As for $\xi^\mu{}_{(1)}$, we find that a *translation* will eliminate both a^1 and a^2. The Killing vector then has the form

$$\xi^1{}_{(2)} = by, \quad \xi^2{}_{(2)} = -bx, \quad \xi^4{}_{(2)} = 0.$$

The Killing equations may then be integrated in a straightforward manner. If we transform the resultant metric so that $\xi^\mu{}_{(2)}$ is in canonical form,

$$\xi^\mu{}_{(2)} \overset{*}{=} \delta^\mu{}_2, \tag{36}$$

then the metric may be written as follows:

$$ds^2 = dr^2 + \exp\{u\}r^2d\theta^2 + 2dudv + 2A(u)dud\theta + [In(r/r_0)(v - \tfrac{1}{8}r^2) + B(u)]du^2, \tag{37}$$

where r_0, A, and B are functions of u.

We have investigated all type III perfect metrics which possess a Killing vector and have seen that this must be either $\xi^\mu{}_{(1)}$ or $\xi^\mu{}_{(2)}$. We shall now examine the case where the space admits two or more motions. First of all, we observe that under these conditions the space must possess a Killing vector for which a^4 is zero. This follows from Eq. (32) and the fact that the motions from a linear space. Consequently, the metric can always be transformed to the form of Eq. (37). It is then a simple matter to show from Eq. (28) that the space does not possess a second Killing vector unless A and B have the form

$$A(u) = a\exp\{\tfrac{1}{2}u\}, \quad B(u) = b, \quad r_{0,u} = 0 \tag{37}$$

where a and b are independent of u. The second Killing vector is of the first type considered and may be expressed in this coordinate system as

$$\xi^\mu{}_{(1)} \overset{*}{=} \delta^\mu{}_4 - \tfrac{1}{2}\theta\delta^\mu{}_2. \tag{38}$$

Furthermore, the space does not admit a third Killing vector.

It is easy to see that the two motions cannot have the canonical form of Eqs. (35) and (36) simultaneously, nor, for that matter, can any two linearly independent combinations. This would require the existence of a scalar function, $\bar{\psi}^2$ satisfying

$$\xi^\mu{}_{(1)}\bar{\psi}^2{}_{,\mu} = 0, \quad \xi^\mu{}_{(2)}\bar{\psi}^2{}_{,\mu} = 1,$$

and the first derived system of equations for these are incompatible, since they include

$$(-\xi^\rho{}_{(1)}\xi^\mu{}_{(2),\rho} + \xi^\rho{}_{(2)}\xi^\mu{}_{(1),\rho})\bar{\psi}^2{}_{,\mu} \equiv -\tfrac{1}{2}\xi^\mu{}_{(2)}\bar{\psi}^2{}_{,\mu} = 0.$$

This shows that we cannot transform the metric to a form where it is a function of two coordinates alone.

Invariants. There are no scalar invariants of the second differential order for type III metrics, since the function σ of (3) is not completely determined. However, there are four invariantly determined vectors for the curvature tensor, namely,

$$m_\mu, \quad n_\mu, \quad \bar{p}_\mu = \sigma p_\mu, \quad \bar{q}_\mu = \sigma^{-1}q_\mu. \tag{39}$$

If we use a bar $\bar{e}_{a\mu}$ to denote this tetrad rather than $e_{a\mu}$, then the rotation coefficients and their tetrad derivatives form a complete set of scalar invariants for the metric:

$$\bar{\gamma}_{abc}, \quad \bar{\gamma}_{abc/d}, \quad \cdots.$$

We shall not compute the invariants for the general

metric but shall restrict ourselves to that of Eq. (37). For this, we have

$$r=(\bar{\gamma}_{211})^{-1} \tag{40a}$$

$$v=-r^2/8+r(\tfrac{1}{2}\bar{\gamma}_{311}-\bar{\gamma}_{322}) \tag{40b}$$

$$r_0(u)=r\exp\{3-r^{-1}(4\bar{\gamma}_{322}+2\bar{\gamma}_{311}) \tag{40c}$$

$$A(u)\exp\{-\tfrac{1}{2}u\}=2r(3\bar{\gamma}_{312}-2\bar{\gamma}_{321}) \tag{40d}$$

$$B(u)=-4r\bar{\gamma}_{311/3}+\mathfrak{F}(\bar{\gamma}_{abc}) \tag{40e}$$

where $\mathfrak{F}$ is a rational function of $\bar{\gamma}_{3rs}(r,s=1,2)$ and $\bar{\gamma}_{211}$. From these equations we see that there are always three functional independent scalar invariants unless there is a second Killing vector, and consequently Eq. (37) is satisfied. The coordinates r and v are always invariants, as are the functions $r_0(u)$ and $B(u)$. Since κ of Eq. (4) is just $-\ln(r/r_0)$, it is an invariant and so p_μ must be also. From Eq. (5) we see that u, i.e., ψ^4, is defined up to an additive constant

$$u'=u+c. \tag{41}$$

Consequently, $A(u)$ is also unique up to a multiplicative constant,

$$A'(u')=\exp\{-c/2\}A(u').$$

When there are two Killing vectors present then from Eqs. (37) and (40) we see that a, b, and r_0 are all scalar invariants of the metric and therefore cannot be transformed away.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 2, NUMBER 3 MAY–JUNE, 1961

Absolute Definition of Phase Shift in the Elastic Scattering of a Particle from Compound Systems

A. TEMKIN
Theoretical Division, Goddard Space Flight Center, National Aeronautics and Space Administration, Washington, D. C.

(Received November 4, 1960)

The projection of the target wave function on the total wave function of a scattered particle interacting with the target system is used to define an absolute phase shift including any multiples of π. With this definition of the absolute phase shift, one can prove rigorously in the limit of zero energy for s-wave electrons scattered from atomic hydrogen that the triplet phase shift must approach a nonzero multiple of π. One can further show that at least one π of this phase shift is not connected with the existence of a bound state of the H^- ion.

I. INTRODUCTION

IN the scattering of a particle from a local central potential, Levinson[1] has proved an interesting and important theorem; as it applies to attractive potentials, it says that the number of nodes in the zero-energy radial wave function is equal to the number of bound states which the potential will allow. (For the purposes of this discussion, we confine ourselves to s-wave scattering.) The zero-energy phase shift δ is related to the number of nodes by

$$\delta=n\pi, \quad n=0, 1, 2, \cdots. \tag{1.1}$$

(For an attractive potential, we use the usual convention of choosing the phase shift as positive.)

It should be noted that in this one-body problem the phase shift at any energy has an absolute significance. It is the absolute phase difference between the radial wave function and the spherical Bessel function measured from the origin to where the phase difference becomes constant. This phase difference is the absolute phase shift. It is sufficient for our purposes to define the absolute phase of a function (in radians) at a point as the ratio of the length from the previous node to the abscissa of the point divided by the length between the surrounding nodes, plus the number of nodes up to and including the previous node (but excluding the origin) all times π. (See Fig. 1.)

There has been some interest in extending Levinson's theorem to the case of a particle scattered from targets consisting of more than one particle. In these cases, however, there is as far as we know no universally accepted definition of an absolute phase shift. Clearly, such a definition is a necessary step in extending Levinson's theorem to the compound target case. An obvious approach to such a definition is to associate a one-particle radial wave function with the many-body scattering process; then one can use the above procedure to determine its absolute phase. We shall argue that

$$u_k(r_0)\equiv r_0\int\Phi_0^*(1,2,\cdots,N)\Psi_k(0;1,2,\cdots,N)dr_0^{-1} \tag{1.2}$$

[1] Norman Levinson, Kgl. Danske Videnskab. Selskab, Mat.-fys. Medd 25, No. 9 (1949). The reason these theorems are restricted to zero impacting energy is because for nonzero energy the scattered wave function is always sinusoidal at infinity and, therefore, contains an infinite number of nodes. In the zero-energy case, the scattered wave function degenerates into a straight line, and the number of nodes becomes finite. There are various conditions which the potential must have in order for Levinson's proof to be valid. The most important is that it cannot have more than a Coulomblike singularity at the origin.

metric but shall restrict ourselves to that of Eq. (37). For this, we have

$$r=(\bar{\gamma}_{211})^{-1} \tag{40a}$$

$$v=-r^2/8+r(\tfrac{1}{2}\bar{\gamma}_{311}-\bar{\gamma}_{322}) \tag{40b}$$

$$r_0(u)=r\exp\{3-r^{-1}(4\bar{\gamma}_{322}+2\bar{\gamma}_{311}) \tag{40c}$$

$$A(u)\exp\{-\tfrac{1}{2}u\}=2r(3\bar{\gamma}_{312}-2\bar{\gamma}_{321}) \tag{40d}$$

$$B(u)=-4r\bar{\gamma}_{311/3}+\mathfrak{F}(\bar{\gamma}_{abc}) \tag{40e}$$

where $\mathfrak{F}$ is a rational function of $\bar{\gamma}_{3rs}(r,s=1,2)$ and $\bar{\gamma}_{211}$. From these equations we see that there are always three functional independent scalar invariants unless there is a second Killing vector, and consequently Eq. (37) is satisfied. The coordinates r and v are always invariants, as are the functions $r_0(u)$ and $B(u)$. Since κ of Eq. (4) is just $-\ln(r/r_0)$, it is an invariant and so p_μ must be also. From Eq. (5) we see that u, i.e., ψ^4, is defined up to an additive constant

$$u'=u+c. \tag{41}$$

Consequently, $A(u)$ is also unique up to a multiplicative constant,

$$A'(u')=\exp\{-c/2\}A(u').$$

When there are two Killing vectors present then from Eqs. (37) and (40) we see that a, b, and r_0 are all scalar invariants of the metric and therefore cannot be transformed away.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 2, NUMBER 3 MAY–JUNE, 1961

Absolute Definition of Phase Shift in the Elastic Scattering of a Particle from Compound Systems

A. TEMKIN
Theoretical Division, Goddard Space Flight Center, National Aeronautics and Space Administration, Washington, D. C.

(Received November 4, 1960)

The projection of the target wave function on the total wave function of a scattered particle interacting with the target system is used to define an absolute phase shift including any multiples of π. With this definition of the absolute phase shift, one can prove rigorously in the limit of zero energy for s-wave electrons scattered from atomic hydrogen that the triplet phase shift must approach a nonzero multiple of π. One can further show that at least one π of this phase shift is not connected with the existence of a bound state of the H^- ion.

I. INTRODUCTION

IN the scattering of a particle from a local central potential, Levinson[1] has proved an interesting and important theorem; as it applies to attractive potentials, it says that the number of nodes in the zero-energy radial wave function is equal to the number of bound states which the potential will allow. (For the purposes of this discussion, we confine ourselves to s-wave scattering.) The zero-energy phase shift δ is related to the number of nodes by

$$\delta=n\pi, \quad n=0, 1, 2, \cdots. \tag{1.1}$$

(For an attractive potential, we use the usual convention of choosing the phase shift as positive.)

It should be noted that in this one-body problem the phase shift at any energy has an absolute significance. It is the absolute phase difference between the radial wave function and the spherical Bessel function measured from the origin to where the phase difference becomes constant. This phase difference is the absolute phase shift. It is sufficient for our purposes to define the absolute phase of a function (in radians) at a point as the ratio of the length from the previous node to the abscissa of the point divided by the length between the surrounding nodes, plus the number of nodes up to and including the previous node (but excluding the origin) all times π. (See Fig. 1.)

There has been some interest in extending Levinson's theorem to the case of a particle scattered from targets consisting of more than one particle. In these cases, however, there is as far as we know no universally accepted definition of an absolute phase shift. Clearly, such a definition is a necessary step in extending Levinson's theorem to the compound target case. An obvious approach to such a definition is to associate a one-particle radial wave function with the many-body scattering process; then one can use the above procedure to determine its absolute phase. We shall argue that

$$u_k(r_0)\equiv r_0\int\Phi_0^*(1,2,\cdots,N)\Psi_k(0;1,2,\cdots,N)dr_0^{-1} \tag{1.2}$$

[1] Norman Levinson, Kgl. Danske Videnskab. Selskab, Mat.-fys. Medd 25, No. 9 (1949). The reason these theorems are restricted to zero impacting energy is because for nonzero energy the scattered wave function is always sinusoidal at infinity and, therefore, contains an infinite number of nodes. In the zero-energy case, the scattered wave function degenerates into a straight line, and the number of nodes becomes finite. There are various conditions which the potential must have in order for Levinson's proof to be valid. The most important is that it cannot have more than a Coulomblike singularity at the origin.

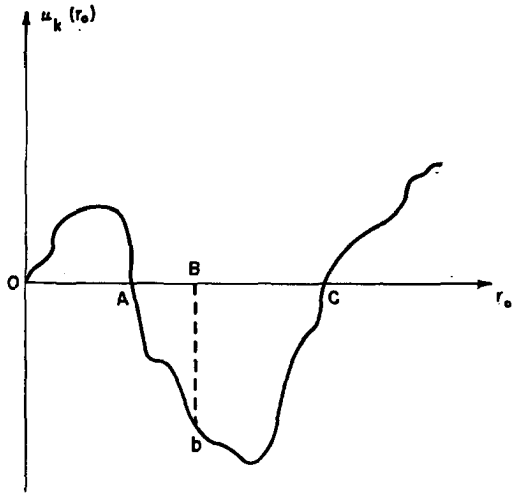

FIG. 1. Absolute phase of a function $u_k(r_0)$. In the diagram, phase of u_k at b is $(1+\langle AB\rangle/\langle AC\rangle)\pi$. Note that bumps in the function $u_k(r_0)$ do not contribute integral multiples of π to phase unless they cause u_k to cross the axis.

provides a consistent definition of such an equivalent one-particle radial wave function.[2] (u_k should always be understood as r_0 times the radial wave function. dr_0^{-1} signifies integration over all coordinates but r_0.) Indeed, the above expression, which is the projection on the target wave function Φ_0 of the exact wave function of the target system plus incoming particle Ψ_k, has tacitly been assumed by some people. However, alternate means of defining the phase shift have been used. For instance, in the scattering of neutrons from nuclei, it is customary to compute the phase shift from formulas relating it to the logarithmic derivative at a point where the interaction is assumed to be negligible. In so doing, one has renounced altogether defining a phase shift except modulo π. There is some justification in this approach on the ground that when the incoming particle is interacting strongly with the target system, the wave function is highly nonseparable; then the idea of a phase shift loses all meaning. However, no matter how complicated the wave function, it is clear that (1.2) does tell how much on the average the incoming particle is being attracted or repelled by the target. What we are really asking is whether this definition of the phase shift, which traces the buildup of the phase shift as a function of r_0, will not alter (correctly) calculated results.

However, that the definition (1.2) cannot change results modulo π is abundantly clear. For in order that a phase shift be defined in a given process, Ψ_k must have the asymptotic form

$$\lim_{r_0\to\infty}\Psi_k=\sin(kr_0+\delta')\Phi_0+\sum_{j\neq 0}h_j(r_0)\Phi_j(1,\cdots,N),$$

where the Φ_j for $j\neq 0$ refer to other states (including perhaps dissociated states) of the target system (for which the h_j do not vanish). All these states have in common that they are orthogonal to the original state of the target. Thus, substituting the above into (1.2), we see that

$$\delta=\delta'+l\pi.$$

In Sec. II we shall discuss the Hartree-Fock approximation for scattering, in particular Swan's[3] conjecture of the extension of Levinson's theorem, which is intimately connected with it. A different method of determining phase shifts is also analyzed in that section. Section III contains a rigorous yet trival demonstration that the triplet, zero-energy phase shift in the scattering of electrons from atomic hydrogen is a nonzero multiple of π. One can show that this behavior need have nothing to do with the existence of triplet bound state of the composite system which is the H^- ion.

II. ALTERNATE METHODS OF DETERMINING THE MANY-PARTICLE PHASE SHIFT AND SWAN'S CONJECTURE

A particularly important case of scattering from a compound target is the scattering of electrons from atoms. One of the most established methods of treating this problem is the exchange or Hartree-Fock approximation.[4] According to this approximation, one makes an Ansatz for the total wave function of an antisymmetrized product of an undetermined function $\tilde{u}_k(r_0)$ of the scattered particle times the ground state wave function of the atom Φ_0:

$$r_0\Psi_k=\mathcal{A}\{\tilde{u}_k(r_0)\Phi_0(1,2,\cdots,N)\}. \tag{2.1}$$

$\mathcal{A}$ is the antisymmetrizer; Φ_0 is considered already antisymmetrized. The variational principle (H is the Hamiltonian of the total system and E the total energy),

$$\delta\int\Psi_k^*(H-E)\Psi_k d\tau=0, \tag{2.2}$$

is used to derive a one-dimensional integro-differential equation for $\tilde{u}_k$, from which the phase shift is determined precisely as in the one-body problem. Here then, one has a prescription for calculating an absolutely defined phase shift, which, however, has the disadvantage of being tied to an approximate Ansatz for the wave function. Within this approximation, Swan[4] has conjunctured an extension of Levinson's theorem which says, concerning the zero-energy phase shift δ,

$$\delta=(n+m)\pi, \tag{2.3}$$

[2] We assume that Ψ_k is real. We have always found that one can so choose Ψ_k providing one is dealing with a given partial wave (such that Ψ_k is a state of good total angular momentum) below the threshold for any inelastic process.

[3] P. Swan, Proc. Roy. Soc. (London) **A228**, 10 (1955). This important paper is unfortunately not clear. In particular, the conjecture is put in the form of a theorem, but it is not clear to what extent his subsequent arguments are intended to be a proof of the theorem.

[4] P. M. Morse and W. P. Allis, Phys. Rev. **44**, 269 (1933). See also E. Feenberg, Phys. Rev. **40**, 40 (1932); **42**, 17 (1932).

where n is, as above, the number of composite bound states that the potential will allow, and m is the number of states from which the particle is excluded by the Pauli principle. An example of the meaning of m is the following. Consider the s-wave scattering of a particle from an atom whose (Hartree-Fock) wave function has filled $1s$ and $2s$ shells. Both these shells are excluded so that the zero-energy wave function would be expected to have at least two nodes according to Swan's theorem. If, in addition, the negative ion would have a bound state in which the additional particle were in a $3s$ state, then an additional node would be induced in the zero-energy scattered wave function.

We wish to investigate the connection of the function $\tilde{u}_k$ with u_k of (1.2), assuming we replace the exact Ψ_k by the exchange approximate (2.1).

From the considerations of Sec. I we know that $\tilde{u}_k$ and u_k will give rise to the same phase shift modulo π. What we want to know is how these functions compare for smaller values of r_0. Because of the antisymmetry of (2.1), one can add to $\tilde{u}_k$ any amount of any orbitals which have the same angular and spin dependence (assumed to be s orbitals in the case of s-wave scattering) without changing Ψ_k. For instance, in the example discussed above, one can add any amount of $1s$ or $2s$ to $\tilde{u}_k$ without affecting Ψ_k, and in general a $\tilde{u}_k = \tilde{U}_k$ can always be found which is orthogonal to all the orbitals of the Hartree-Fock wave function. *The virtue of (1.2) as it applies to the Hartree-Fock approximation is that it automatically generates such a* $\tilde{U}_k$. It is to such a $\tilde{U}_k$ that Swan's theorem is meant to apply.[5] From the theory of linear second-order equations (in particular the radial hydrogen equations), one would expect the scattered "orbital" in the example to have at least one more node that u_{2s}, which would mean that it would have at least two nodes. [See example connected with Eq. (3.5a).] The existence of bound states would then, in the usual way, induce additional nodes, and this is the heuristic basis of Swan's conjecture. (The argument is not rigorous because $\tilde{u}_k$ does not satisfy an ordinary differential equation, but rather an integro-differential equation.) The only part of Swan's assertion which can be rigorously salvaged from the theorem of Sec. III is that the zero-energy s-wave phase shift of an electron scattered from an atom with a filled $1s$ shell must approach at least π.

There arises here as in the one-body problem an ambiguity in the over-all sign of the phase shift. In the one-body problem, this ambiguity is settled, where the radial function $u_k(r)$ is known, by choosing that sign for $u_k(r)$ such that $u_k'(0)>0$. In our case, this means choosing that sign for Ψ_k such that the same condition for $u_k(r_0)$ applies.[6] In both cases, the phase shift is the difference between the phase of $u_k(r_0)$ and that of $+\sin kr_0$. This assures that an attractive potential produces a positive phase shift.

It may be thought at this point that we are getting very much out of this procedure for determining the absolute phase shift. However, one should bear in mind that we are using considerably more information contained in the total wave function than its asymptotic form. In particular, by using the projection of the total wave function on the ground state, we are including some aspects of its behavior right down to the origin.

In some applications, the phase shift has been determined by the condition that it approach zero as the energy becomes infinite, and that it vary continuously as the energy varies. We should like to point out that the first condition is not necessarily true and that the second can be misleading. An obvious counter example of the first condition is a system with hard-core potentials. Here, the region in which the wave function differs from its asymptotic form does not vanish as the energy is increased indefinitely, and the phase shift does not approach zero. Concerning the second condition, it must be emphasized that although the phase shift is a continuous function of the energy, its slope need not be continuous and will in fact be discontinuous whenever the threshold for some competing process is reached. Since in any numerical calculation one can only find phase shifts for some finite set of k's, it is possible that a discontinuity in slope of δ as a function of k may appear as a discontinuity in δ itself. In some variational calculations where the variational expression may have several relative minima as a function of k, such a behavior might cause the wrong branch to be followed.

III. PROOF THAT THE ZERO-ENERGY TRIPLET PHASE SHIFT IN THE SCATTERING OF ELECTRONS FROM HYDROGEN IS A NONZERO MULTIPLE OF π

Let $\Psi_-(\mathbf{r}_0,\mathbf{r}_1)$ be the exact wave function of the zero-energy triplet $e-$H system. Then,

$$\Psi_-(\mathbf{r}_0,\mathbf{r}_1) = -\Psi_-(\mathbf{r}_1,\mathbf{r}_0). \tag{3.1}$$

Now expand

$$\Psi_-(\mathbf{r}_0,\mathbf{r}_1) = \sum_n \psi_n(\mathbf{r}_0)\varphi_n(\mathbf{r}_1), \tag{3.2}$$

where φ_n are the states of the hydrogen atom. According to our definition, the absolute phase shift is determined from

$$\psi_0(\mathbf{r}_0) = \int \varphi_0^*(\mathbf{r}_1)\Psi_-(\mathbf{r}_0,\mathbf{r}_1)d\mathbf{r}_1. \tag{3.3}$$

[5] The Hartree-Fock orbitals of nonclosed shell atoms are not in general orthogonal. Nevertheless one can always orthonormalize them without changing the Slater determinant. We shall always think of the determinantal Hartree-Fock wave function whose orbitals are orthonormalized. The fact that $\tilde{u}_k$ could be made orthogonal to the Hartree-Fock orbitals was realized by Feenberg (footnote 4).

[6] This prescription only applies where the wave function is defined over all space. It therefore is not suitable without modification for such things as hard-core potentials. Assuming the potential to be allowable, the prescription implies that a repulsive potential can only give rise to a phase shift of 0 as $k \to 0$; i.e., it cannot be any other negative multiple of π. In the one-body case, this is implicit in Levinson's theorem (footnote 1).

Now use (3.1) and (3.2) on the rhs of (3.3), and multiply both sides by $\varphi_0^*(r_0)$ and integrate over $\mathbf{r}_0$ to get

$$\int \varphi_0^*(\mathbf{r}_0)\psi_0(\mathbf{r}_0)d\mathbf{r}_0 = -\int \varphi_0^*(\mathbf{r}_1)\psi_0(\mathbf{r}_1)d\mathbf{r}_1. \quad (3.4)$$

From this, we conclude finally

$$\int_0^\infty u_0(r_0)u_{1s}(r_0)dr_0 = 0, \quad (3.5)$$

where

$$\psi_0(\mathbf{r}_0) = [u_0(r_0)/r_0]Y_{00}(\Omega_0),$$

and

$$\varphi_0(\mathbf{r}_0) = [u_{1s}(r_0)/r_0]Y_{00}(\Omega_0) \equiv \varphi_{10}(\mathbf{r}_0).$$

This proof is implicit in the work of Mittleman,[7] in which he constructs an equivalent (nonlocal) potential for $\psi_0(\mathbf{r}_0)$ for which the orthogonality property is preserved at every stage of approximation. However, by virtue of our definition of absolute phase shift, we can readily go one step further and say that this implies the triplet phase is a nonzero multiple of π. For the function $u_{1s}=2r_0e^{-r_0}$ has no nodes; therefore, the function $u_0(r_0)$ must have at least one node. From our method of computing phase shifts, it follows that this is equivalent to the phase shift being $n\pi$ where $n>0$ is the number of nodes in $u_0(r_0)$ (neglecting the measure zero probability that the slope of u_0 is horizontal at $r_0=\infty$, which would add another $\pi/2$ onto the phase shift[1]). To show that at least one of these nodes is not connected with the existence of a triplet bound state of the H^- ion, one need only consider the case where the repulsion between the electrons was increased and the attraction of the electrons to the nucleus was decreased. It is clear that a point would be reached for it to be physically inconceivable for a bound state to exist. Yet everything in the above proof would go through, and there would still have to be at least one node in $u_0(r_0)$.

It is worth noting in the proof of orthogonality that the eigenfunction character of $\Psi_-(\mathbf{r}_0,\mathbf{r}_1)$ with respect to the total Hamiltonian is not used. This has as a consequence the fact that if one uses any antisymmetric Ansatz for Ψ_- and computes an equivalent one-particle orbital and phase shift as defined above, then the nodal behavior of this function and the orthogonality still apply. As a trivial example, consider the Ansatz

$$\Psi_-(r_0,r_1) = (4\pi)^{-\frac{1}{2}}\left\{\frac{\sin kr_0}{r_0}\varphi_{10}(r_1) - \frac{\sin kr_1}{r_1}\varphi_{10}(r_0)\right\}. \quad (3.5a)$$

If we compute $u_k(r_0)$ according to (1.2), confining ourselves to the $k=0$ case, we find

$$\lim_{k\to 0} u_k(r_0)/k = r_0(1-8e^{-r_0}).$$

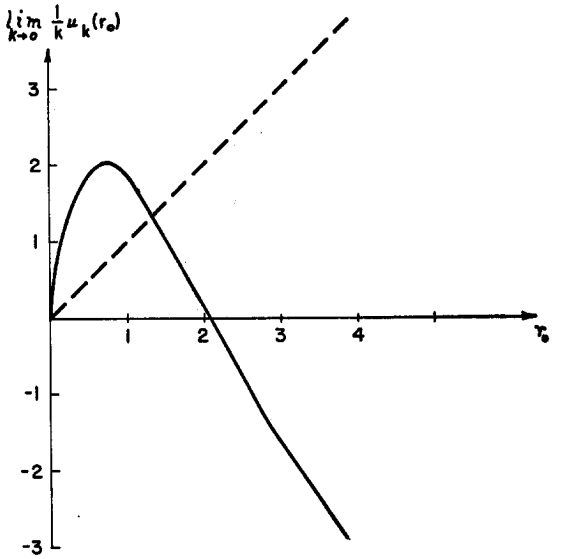

FIG. 2. Solid curve is plot of $k^{-1}u_k(r_0)=k^{-1}r_0\int\varphi_{10}\Psi_-dr_1$, where Ψ_- is given in Eq. (3.5a) for $k=0$. The phase shift in this case is π. Dashed line is scattered function before orthogonalization to ground state function. For that curve phase shift is zero.

A simple integration shows that this function is orthogonal to $u_{1s}(r_0)$, and it is also clear that the function has one node at $r_0=\ln 8$. [Note that $8r_0e^{-r_0}$ is a multiple of $u_{1s}(r_0)$.] If we had concentrated our attention on $k^{-1}\sin kr_0 \to r_0$ as $k\to 0$, then we would have said the phase shift approached 0 (see Fig. 2). The function $\sin kr_0$, for arbitrary k, is the analog of the function $\tilde{u}_k(r_0)$ in the Hartree-Fock approximation before orthogonalization. The function $u_k(r_0)$ is then the orthogonalized form of $\tilde{u}_k$ and, as discussed in Sec. II, it can be substituted in (3.5a) without changing Ψ_-.

In this context, it should be mentioned that in the work of Mittleman and Watson,[8] which concerns the construction of equivalent one-body potentials in the case that an incoming particle is different from the particles in the target, the functions for which the potentials are constructed are equivalent to those defined in (1.2).

We shall conclude with an example of a nontrivial method in which the phase shift is determined by a procedure different from any of those discussed above. It has been shown[9] that the s-wave scattering of electrons by hydrogen can be described in zeroth order (which approximation nevertheless can be expected to give results correct to within 25%) by the equation

$$\left[-\frac{\partial^2}{\partial r_<^2} - \frac{\partial^2}{\partial r_>^2} - \frac{2}{r_<} + 1 - k^2\right]\Psi_0^{(0)}(r_0,r_1) = 0. \quad (3.6)$$

$r_<$ is the lesser and $r_>$ is the greater of r_0 and r_1. [$\Psi_0^{(0)}$ is an approximation of (r_0r_1 times the) s-wave function

[7] M. Mittleman, University of California Radiation Lab. Rept. UCRL-5711 (1959) (unpublished).

[8] M. Mittleman and K. Watson, Phys. Rev. **113**, 198 (1958).

[9] A. Temkin, Phys. Rev. Letters **4**, 511 (1960).

of the $e-H$ system.] The equation can be considered completely in the region $r_0>r_1$:

$$\left[-\frac{\partial^2}{\partial r_0^2}-\frac{\partial^2}{\partial r_1^2}-\frac{2}{r_1}+1-k^2\right]\Phi_0^{(0)}=0, \quad r_0>r_1, \quad (3.6a)$$

providing one adds the additional boundary condition corresponding to the triplet $\Psi_0^{(0)}$ that $\Phi_0^{(0)}$ be zero along the line $r_0=r_1$. $\Phi_0^{(0)}$ can be expanded in the form

$$\Phi_0^{(0)}(r_0,r_1)=\sin(kr_0+\delta)u_{10}(r_1) + \sum_{n=2} C_n e^{-\kappa_n r_0}u_{n0}(r_1), \quad (3.7)$$

where

$$\kappa_n=(1-n^{-2}-k^2)^{\frac{1}{2}}.$$

The $u_{n0}(r_1)$ are r_1 times the radial s state wave functions of the hydrogen atom. The sum includes the states of the continuum (for which $n\rightarrow i/p$, where p is the momentum of a continuous state). Each term is separately a solution of the above equation, thus obeying one of the boundary conditions of being 0 along $r_1=0$. The C_n and δ are determined by the remaining boundary condition $\Phi_0^{(0)}(r_0=r_1)=0$. In actual practice, we use a finite number of terms and determine the C_n and δ by minimizing the expression $\int_0^\infty|\Phi_0^{(0)}(r_0=r_1)|^2dr_0$. The minimization leads in the case that we only include the first two terms of (3.7) to the expressions

$$\tan2\delta=\frac{-\int_0^\infty e^{-2\kappa_2 r}u_{20}^2dr\int_0^\infty \sin2kru_{10}^2dr+2\mathfrak{N}_s\mathfrak{N}_c}{\int_0^\infty e^{-2\kappa_2 r}u_{20}^2dr\int_0^\infty \cos2kru_{10}^2dr+\mathfrak{N}_s^2-\mathfrak{N}_c^2},$$

and

$$C_2=\frac{-\int_0^\infty \sin(kr+\delta)e^{-\kappa_2 r}u_{10}u_{20}dr}{\int_0^\infty e^{-2\kappa_2 r}u_{20}^2dr}.$$

Here

$$\mathfrak{N}_{(s;c)}=\int_0^\infty \left\{\begin{matrix}\sin kr\\ \cos kr\end{matrix}\right\}e^{-(\kappa_2+\frac{1}{2})r}u_{10}u_{20}dr.$$

The integrations are trivial, and one can select the correct quadrant of 2δ by testing to see which 2δ actually minimizes $\int_0^\infty|\Phi_0^{(0)}|^2dr$. This uniquely determines C_2.

At this point δ is still undetermined modulo π, and we would expect to use (1.2) to make δ unique. However,

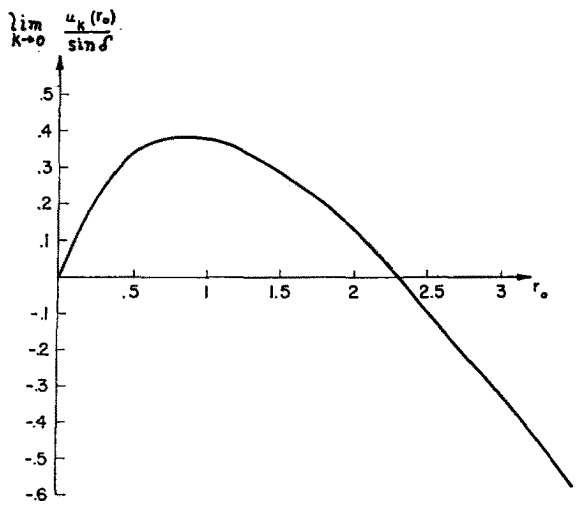

FIG. 3. Zero energy form of u_k given in (3.8), where $\Phi_0^{(0)}$ is given by (3.9) and (3.7). Only two terms in (3.7) are used. The function $\lim k\rightarrow 0\ (\sin\delta)^{-1}u_k(r_0)$ is plotted. It is orthogonal to $u_{1s}(r_0)=2r_0e^{-r_0}$.

to use (1.2), which reduces here to

$$u_k(r_0)=\int_0^\infty u_{10}(r_1)\Psi_0^{(0)}(r_0,r_1)dr_1, \quad (3.8)$$

one must know how $\Psi_0^{(0)}$ is related to $\Phi_0^{(0)}$. This is not completely obvious since if $\Phi_0^{(0)}(r_0,r_1)$ is a solution of (3.6) in the region $r_0>r_1$, then the two solutions $\pm\Phi_0^{(0)}(r_1,r_0)$ are solution of (3.6) in the region $r_1>r_0$. Physically, it is obvious that if $\Psi_0^{(0)}$ is to correspond to the space antisymmetric (triplet) solution

$$\Psi_0^{(0)}(r_0,r_1)=-\Psi_0^{(0)}(r_1,r_0),$$

then $\Psi_0^{(0)}$ must be the combination

$$\begin{aligned}\Psi_0^{(0)}&=\Phi_0^{(0)}(r_0,r_1) \quad r_0>r_1\\ &=-\Phi_0^{(0)}(r_1,r_0) \quad r_0<r_1.\end{aligned} \quad (3.9)$$

With this, the function $\Psi_0^{(0)}$ is completely defined and the integration (3.8) can be carried out from which δ can be found uniquely. On using the two term approximation of $\Phi_0^{(0)}$, we get the curve of Fig. 3, from which it is clear that $\delta\rightarrow\pi$.

ACKNOWLEDGMENTS

We are grateful to Dr. I. C. Percival for pointing out to us, after we had shown him a draft of this material, that L. Rosenberg and L. Spruch were working along similar lines. A subsequent discussion with Dr. Spruch revealed that these authors have in fact adopted the same definition and proven the same theorem as has been done here.

Operator Formalism in Statistical Mechanics

ELLIOTT LIEB
IBM Research Laboratory, Yorktown Heights, New York
(Received December 28, 1960)

We consider the configuration partition function, Z_N, of a classical imperfect gas of particles having rigid cores. The rigid cores result in a geometric simplification which makes it possible to find a finite recursion relation between Z_N and Z_{N-1}. From this, it is possible to express Z_N in terms of the vacuum expectation value of a finite Boson operator raised to the Nth power, N being the number of particles.

AT the present time there exist many points of view, formalisms, and exact relations, pertinent to the evaluation of the partition function of a classical gas. None of these have proved sufficiently powerful to solve the problem in any except weak coupling or low density situations; yet, they are nonetheless valuable in that, taken together, they may ultimately provide the needed insight into the structure of the true solution. In the present note, we consider a gas of pairwise interacting particles which have an infinite repulsive core interaction together with an arbitrary short range finite part. Since a real gas always has some repulsive core, the situation discussed here is quite general. We shall present a formalism appropriate to this problem which, while it does not offer an immediate solution, is sufficiently different from other approaches that it may be of value.

Mainly two ideas are involved. The first is the geometric simplification resulting from the fact that it is always impossible to find a configuration with two or more particles overlapping (by the word overlapping we of course mean overlapping of the hard cores). This idea has been used before,[1] but we believe not quite in the same form as in the present work. In any event, the geometric simplification is essential for the second part which consists in the reduction of the partition function to the evaluation of the expectation value of an operator raised to the Nth power, N being the number of particles. That is to say, we reduce the evaluation of a $(3N)$-dimensional integral to something similar to the trace of an iterated linear operator. Such a procedure is conveniently employed in many problems of statistical mechanics, for the problem then becomes one of finding the largest eigenvalue of the matrix in question. In our case, unfortunately, we cannot use this well-known device because the matrix is not Hermitian, and indeed has no eigenvalues. All the same, it is our hope that this formal reduction may be perspicuous.

We shall begin by discussing the case of rigid spheres of diameter a, no additional finite potential being present. In this case, the configuration partition function Z_N depends only upon the density and is independent of the temperature. The configuration partition function is defined by

$$Z_N=\frac{1}{N!}\int_V \prod_1^N d^3\mathbf{x}_i \; I_N(\mathbf{x}_1, \cdots, \mathbf{x}_N), \tag{1}$$

where

$$\begin{aligned} I_N&=1, \quad \text{all } |\mathbf{x}_i-\mathbf{x}_j|>a \\ &\quad 0, \quad \text{any } |\mathbf{x}_i-\mathbf{x}_j|<a. \end{aligned} \tag{2}$$

The function I_N may be expressed in terms of the Mayer f functions by

$$I_N=\prod_{\langle i,j\rangle}(1+f_{ij}), \tag{3}$$

where

$$\begin{aligned} f_{ij}&=-1, \quad |\mathbf{x}_i-\mathbf{x}_j|<a \\ &=\ 0, \quad |\mathbf{x}_i-\mathbf{x}_j|>a. \end{aligned} \tag{4}$$

Therefore,

$$I_N(\mathbf{x}_1, \cdots, \mathbf{x}_N)=\{\prod_{j=2}^N(1+f_{1j})\}I_{N-1}(\mathbf{x}_2, \cdots, \mathbf{x}_N). \tag{5}$$

In Eq. (5), we have explicitly extracted the part of I_N depending upon the variable $\mathbf{x}_1$.

On expanding the first factor in Eq. (5), we have[2]

$$\begin{aligned} I_N(\mathbf{x}_1, \cdots, \mathbf{x}_N)=I_{N-1}(\mathbf{x}_2, \cdots, \mathbf{x}_N)\{1+\sum_{j=2}^N(f_{1j} \\ +\frac{1}{2!}\sum_{j_1,j_2=2}^N f_{1j_1}f_{2j_2}+\cdots+\frac{1}{11!}\sum_{j_1\cdots j_{11}}(f\cdots f) \\ +\frac{1}{12!}\sum_{j_1\cdots j_{12}}(f\cdots f)+\cdots\}. \end{aligned} \tag{6}$$

Now if we inspect the series in Eq. (6), we see that all terms involving 12 or more f's may be deleted. This is a simple consequence of the fact that in three dimensions it is impossible to arrange more than 11 rigid spheres of diameter $=a$ in such a way that all their centers are *within* a distance $a/2$ of a given point. In other words, a term such as

$$\sum_{j_1\cdots j_{12}} f_{1j_1}\cdots f_{1j_{12}}$$

is nonvanishing only when $I_{N-1}(\mathbf{x}_2, \cdots, \mathbf{x}_N)$ vanishes.

[1] T. L. Hill, *Statistical Mechanics* (McGraw-Hill Book Company, Inc., New York, 1956), p. 256.

[2] The sums appearing in Eqs. (6) and (8) are to be understood as not including terms in which two or more j's are equal.

Consequently, we may write

$$I_N(\mathbf{x}_1, \cdots, \mathbf{x}_N) = \{\sum_{s=0}^{11} G_s(\mathbf{x}_1 | \mathbf{x}_2 \cdots \mathbf{x}_N)\} I_{N-1}(\mathbf{x}_2 \cdots \mathbf{x}_N), \quad (7)$$

where

$$G_s(\mathbf{x}_1 | \mathbf{x}_2 \cdots \mathbf{x}_N) = \frac{1}{s!} \sum_{j_1=2}^{N} \cdots \sum_{j_s=2}^{N} [f(\mathbf{x}_1 - \mathbf{x}_{j_1}) \cdots f(\mathbf{x}_1 - \mathbf{x}_{j_s})]. \quad (8)$$

The point is that only 12 different G's appear in Eq. (7) instead of N, as would be the case if no rigid core were present. We might remark in passing, that in two dimensions the terms beyond G_5 would be deleted, and in one dimension only G_0, G_1, and G_2 would appear. For rigid oriented squares in two dimensions, only $G_0 \cdots G_4$ would be present, one less term than for circles.

It is easy to see how a finite range potential would modify Eq. (7). One must ask how many rigid spheres of diameter$=a$ can come *within* a distance $a/2+b$ of a given point, where $b=$range of the potential. Whatever the value of b, only a finite number of G's are required. In this case, of course, the $f(ij)$, and hence the G_s, will depend upon the temperature.

We next expand I_N in a Fourier series

$$I_N(\mathbf{x}_1, \cdots, \mathbf{x}_N) = \sum_{\mathbf{p}_1 \cdots \mathbf{p}_N} \exp(i \sum_{1}^{N} \mathbf{p}_i \cdot \mathbf{x}_i) \phi_N(\mathbf{p}_1, \cdots, \mathbf{p}_N), \quad (9)$$

where

$$\phi_N(\mathbf{p}_1, \cdots, \mathbf{p}_N) = V^{-N} \int_V \prod_{1}^{N} d^3\mathbf{x}_i \; I_N(\mathbf{x}_1, \cdots, \mathbf{x}_N) \times \exp(-i \sum_{1}^{N} \mathbf{p}_i \cdot \mathbf{x}_i). \quad (10)$$

A Fourier series is obviously convenient since the terms in Eq. (7) are in the form of a convolution.

The Fourier transform [Eq. (10)] of the first term in (7) (i.e., in the $s=0$ term) is given by

$$F[I_{N-1}(\mathbf{x}_2, \cdots, \mathbf{x}_N)] = \delta(\mathbf{p}_1) \phi_{N-1}(\mathbf{p}_2, \cdots, \mathbf{p}_N), \quad (11)$$

where $F[\;]$ means Fourier transform, and $\delta(\mathbf{p}_1)$ is the three-dimensional Kroenecker delta function. For the second term, we have

$$F[\sum_{j=2}^{N} f(\mathbf{x}_1 - \mathbf{x}_j) I_{N-1}(\mathbf{x}_2, \cdots, \mathbf{x}_N)] = \frac{1}{V} \sum_{j=2}^{N} \sum_{\mathbf{q}} F(\mathbf{q}) \phi_{N-1}(\mathbf{p}_2, \cdots, \mathbf{p}_j + \mathbf{q}, \cdots, \mathbf{p}_N) \times \delta(\mathbf{p}_1 - \mathbf{q}), \quad (12)$$

where $V^{-1}F(\mathbf{q})$ is the Fourier transform of $f(\mathbf{x})$, i.e.,

$$F(\mathbf{q}) = \int_V \exp(-i\mathbf{q} \cdot \mathbf{x}) f(\mathbf{x}) d^3\mathbf{x}. \quad (13)$$

For the rigid-sphere gas in three dimensions,

$$F(\mathbf{q}) = -\int_{r<a} \exp(-i\mathbf{q} \cdot \mathbf{x}) d^3\mathbf{x} = (4\pi a/q^2)\{\cos qa - (1/qa) \sin qa\}. \quad (14)$$

The third term in (7) yields

$$F[\sum_{\langle i,l \rangle} f(\mathbf{x}_1 - \mathbf{x}_j) f(\mathbf{x}_1 - \mathbf{x}_l) I_{N-1}(\mathbf{x}_2 \cdots \mathbf{x}_N)] = \sum_{\langle i,l \rangle} \sum_{\mathbf{q}_1} \sum_{\mathbf{q}_2} F(\mathbf{q}_1) F(\mathbf{q}_2) \delta(\mathbf{p} - \mathbf{q}_1 - \mathbf{q}_2) \times \phi_{N-1}(\mathbf{p}_2, \cdots, \mathbf{p}_j + \mathbf{q}_1, \cdots, \mathbf{p}_l + \mathbf{q}_2, \cdots, \mathbf{p}_N). \quad (15)$$

It is easy to see from Eq. (15) the form of the remaining terms in (7).

At this point we introduce the notion of second quantization. Since I_N, and hence ϕ_N, are symmetric function of N variables, we seek an algebraic formulation of Eq. (7) which automatically keeps track of the $N-1$ similar terms in Eq. (12), the $(N-1; 2)$ terms of Eq. (15), etc. Such an algebraic formalism has already been developed in connection with Boson fields and is called second quantization, but the formalism itself has nothing to do with the esoteric nature of field theory and is merely a convenient "bookkeeping" arrangement.

To each value of the "momentum" $\mathbf{k}$, we associate the operator $a_\mathbf{k}$ and its Hermitian conjugate $a_\mathbf{k}^+$ with the well-known Boson commutation relations:

$$[a_\mathbf{k}, a_{\mathbf{k}'}] = 0, \; [a_\mathbf{k}^+, a_{\mathbf{k}'}^+] = 0, \; [a_\mathbf{k}, a_{\mathbf{k}'}^+] = \delta(\mathbf{k} - \mathbf{k}'). \quad (16)$$

From any function $\phi_N(\mathbf{k}_1, \cdots, \mathbf{k}_N)$, we generate the corresponding "state" ψ_N in the following way:

$$\psi_N \equiv \sum_{\mathbf{k}_1} \cdots \sum_{\mathbf{k}_N} \phi_N(\mathbf{k}_1, \cdots, \mathbf{k}_N) a_{\mathbf{k}_1}^+ \cdots a_{\mathbf{k}_N}^+ |0\rangle, \quad (17)$$

where $|0\rangle$ is the "vacuum" and has the property

$$a_\mathbf{k} |0\rangle = 0, \quad \text{all } \mathbf{k}, \quad (18)$$

It then follows that the state generated by the right-hand side of Eq. (11) is given by $O_0 \psi_{N-1}$, where

$$O_0 = a_0^+ \quad (19)$$

and ψ_{N-1} is the state generated by ϕ_{N-1}. The state generated by the right-hand side of Eq. (12) is $O_1 \psi_{N-1}$, where

$$O_1 = V^{-1} \sum_{\mathbf{q},\mathbf{p}} F(\mathbf{q}) a_\mathbf{q}^+ a_\mathbf{p}^+ a_{\mathbf{p}+\mathbf{q}}. \quad (20)$$

In general, the state generated by the term $G_s I_{N-1}$ in

Eq. (7) is given by $O_s\psi_{N-1}$, where

$$O_s=\frac{1}{s!}\frac{1}{V^s}\sum_{\mathbf{p}_1\cdots\mathbf{p}_{s+1}}\sum_{\mathbf{q}_1\cdots\mathbf{q}_s}F(\mathbf{q}_1)\cdots F(\mathbf{q}_s)a^+(\mathbf{p}_{s+1})\cdots a^+(\mathbf{p}_1)a(\mathbf{p}_1+\mathbf{q}_1)\cdots a(\mathbf{p}_s+\mathbf{q}_s)\times\delta(\mathbf{p}_{s+1}-\sum_{i=1}^{s}\mathbf{q}_i). \quad (21)$$

We, therefore, have

$$\psi_N=O\psi_{N-1}, \quad (22)$$

where

$$O=\sum_{s=0}O_s, \quad (23)$$

the summation on s being the same as in Eq. (7). It is to be noticed that each O_s is essentially a momentum conserving creation operator, i.e., acting on a state of N particles it creates a state of $N+1$ particles. With this formalism, we have achieved our goal, for the relation Eq. (22) does not contain N explicitly. The state ψ_N is simply given by a definite finite N-independent operator acting on the state ψ_{N-1}.

Now, since $I_1(\mathbf{x}_1)=1$,

$$\psi_1=\sum_{\mathbf{k}_1}\phi_1(\mathbf{k}_1)a_{\mathbf{k}1}{}^+|0\rangle=\sum_{\mathbf{k}_1}\delta(\mathbf{k}_1)a_{\mathbf{k}1}{}^+|0\rangle=a_0{}^+|0\rangle=O|0\rangle. \quad (24)$$

Hence, from Eq. (22),

$$\psi_N=O^N|0\rangle. \quad (25)$$

It is next necessary to convert our knowledge of ψ_N into Z_N. To accomplish this we notice that, by definition,

$$Z_N=(1/N!)V^N\phi_N(0,0,\cdots,0) \quad (26)$$

which, aside from normalization, is the projection of the state ψ_N onto the state $(a_0{}^+)^N|0\rangle$.

We, therefore, have the result

$$Z_N=[V^N/(N!)^2]\langle 0|a_0{}^N O^N|0\rangle. \quad (27)$$

In the case of rigid particles *without* a finite potential, Z_N may be expressed in still another way. Since in this case $I_N=(I_N)^2$, we have

$$Z_N=\frac{1}{N!}\int_V I_N{}^2(\mathbf{x}_1,\cdots,\mathbf{x}_N)\prod_1^N d^3\mathbf{x}_i=\frac{V^N}{N!}\sum_{\mathbf{p}_1\cdots\mathbf{p}_N}|\phi_N(\mathbf{p}_1,\cdots,\mathbf{p}_N)|^2 \quad (28)$$

which, aside from a constant, is the norm of the state ψ_N. Hence,

$$Z_N=[V^N/(N!)^2]\langle 0|(O^+)^N O^N|0\rangle. \quad (29)$$

The advantage of (29) over (27), if any, is that Z_N is expressed in (29) as the expectation value of a Hermitian operator. If a finite potential is present, Z_N, O, $F(\mathbf{q})$, and I_N are functions of $\beta=(kT)^{-1}$. By using an obvious notation, $I_N(2\beta)=[I_N(\beta)]^2$, and consequently,

$$Z_N(2\beta)=[V^N/(N!)^2]\langle 0|[O^+(\beta)]^N[O(\beta)]^N|0\rangle. \quad (30)$$

Unfortunately, O is not a Hermitian operator and, in fact, possesses no right eigenfunctions[3] (consider the operator $O_0=a_0{}^+$). If O possessed a complete set of orthogonal eigenfunctions, to evaluate Z_N would merely require finding the largest eigenvalue of O. The representations, Eqs. (27) and (30), may nevertheless be perspicuous.

It is also possible to generate the grand partition function Q using the operator O. By definition

$$Q(\beta,z)=\sum_{N=0}^{\infty}Z_N(\beta)z^N, \quad (31)$$

where z is the activity. Since O is a creation operator and O^+ is a destruction operator, it follows that

$$\langle 0|(O^+)^N O^M|0\rangle=\delta_{N,M}V^{-N}(N!)^2Z_N(2\beta). \quad (32)$$

Consequently,

$$\langle 0|\exp(\lambda O^+)\exp(\lambda O)|0\rangle=\sum_{n=0}^{\infty}(\lambda^n/n!)^2V^{-n}(n!)^2Z_n(2\beta). \quad (33)$$

If we now set

$$\lambda^2=zV, \quad (34)$$

we have the result

$$Q(2\beta,z)=\langle 0|\exp[\lambda O^+(\beta)]\exp[\lambda O(\beta)]|0\rangle. \quad (35)$$

Equation (35) was derived from (30). Equation (27) could have been used as well.

[3] The operator O, however, possesses a complete set of *left* eigenfunctions.

Theories of Transport in Fluids

H. S. GREEN
University of Adelaide, Adelaide, South Australia
(Received December 22, 1960)

The object of this paper is to establish the equivalence of the Kubo-type transport coefficients with those obtained by the method of Chapman and Enskog. Kubo-type coefficients are derived by a simple method, based on classical mechanics, and these are found to be in general agreement with those found by Mori, who has, however, discarded some important relaxation terms. The neglect of these terms by Kubo and other authors has the effect of leaving their coefficients of diffusion and electrical conductivity divergent. It is shown quite generally that the computation of the corrected Kubo-type coefficients for dilute gases leads to the same results, and even the same calculations, as the method of Chapman and Enskog. The equivalence of the two methods for dense systems is also briefly discussed.

1. INTRODUCTION

THERE has been a renewed interest in the theory of transport processes in the last few years, stimulated to some extent by Kubo's quantum-mechanical formula for the electrical conductivity.[1] It was recognized that the same method might be used to produce formulas for other transport coefficients, and this has been done by Mori,[2] and by Montroll,[3] Kirkwood,[4] and McLennan.[5] Similar results had been obtained previously by Green,[6] by a different method.

The expressions, representing the transport coefficients, obtained by different methods often differ slightly from one another; and, as they are not easy to evaluate, it is not a simple matter to decide on their relative merits. The correct formulas should presumably be in agreement with those obtained by the Chapman-Enskog method for dilute gases, whose generalization to dense fluids was developed to some extent by Born and Green.[7] In gas theory, a partial equivalence between the Kubo-type and the Chapman-Enskog formulas has been demonstrated by Mori,[8] for Maxwellian molecules, and by Montroll,[9] for the coefficient of self-diffusion. The purpose of this paper is to present a general proof of the equivalence of the two methods, for both dilute gases and dense fluids, which will also serve to establish the correct Kubo-type coefficients. Classical mechanics will be used for the sake of simplicity, but a similar quantum-mechanical formalism is easy to construct.

2. NOTATION AND FUNDAMENTALS

A system of particles, constituting a fluid, is constrained within a region of volume V by a potential barrier, which contributes to the external forces. The positions and velocities of the particles, which may be of different types, are $\mathbf{x}_i$ and $\boldsymbol{\xi}_i$ $(i=1\cdots N)$. The mass of the ith particle, assumed to be of the ath type, is $m_i=m_a$; its potential energy due to external forces is $\psi_a(\mathbf{x}_i)$ and its interaction energy with the jth particle, assumed to be of the bth type, is $\phi_{ab}(\mathbf{r}_{ij})$, where $\mathbf{r}_{ij}=\mathbf{x}_i-\mathbf{x}_j$. For brevity, $\psi_a(\mathbf{x}_i)$, $\phi_{ab}(\mathbf{r}_{ij})$ and similar expressions will often be written as ψ_a, ϕ_{ab}, $\cdots$ in the following. Thus, the total potential energy of the system is $\Phi_N=\Sigma_i(\psi_a+\frac{1}{2}\Sigma_j\phi_{ab})$.

The phase-space distribution function for the isolated system of particles F_N satisfies Liouville's equation:

$$\frac{dF_N}{dt_N}=\frac{\partial F_N}{\partial t}+\Sigma_i\left(\boldsymbol{\xi}_i\cdot\frac{\partial F_N}{\partial \mathbf{x}_i}-\frac{1}{m_i}\frac{\partial \Phi_N}{\partial \mathbf{x}_i}\cdot\frac{\partial F_N}{\partial \boldsymbol{\xi}_i}\right)=0. \quad (1)$$

The normalization and other properties of this function are detailed in an earlier paper by the author.[10] It can be used to construct the single-particle velocity distribution function f_a which, for a dilute gas, satisfies the Maxwell-Boltzmann equation

$$\frac{\partial f_a}{\partial t}+\boldsymbol{\xi}_i\cdot\frac{\partial f_a}{\partial \mathbf{x}_i}-\frac{1}{m_a}\frac{\partial \psi_a}{\partial \mathbf{x}_i}\cdot\frac{\partial f_a}{\partial \boldsymbol{\xi}_i}=\Sigma_b[f_a,f_b],$$
$$[f_a,f_b]=\int\int(f_a'f_b'-f_af_b)g\,d\mathbf{b}\,d\boldsymbol{\xi}_j, \quad (2)$$

where $g=|\boldsymbol{\xi}_i-\boldsymbol{\xi}_j|$ and $d\mathbf{b}$ is an element of cross section for the encounter $(\boldsymbol{\xi}_i',\boldsymbol{\xi}_j')\rightarrow(\boldsymbol{\xi}_i,\boldsymbol{\xi}_j)$. It has been shown by the author[11] that (2) is a rigorous consequence of Liouville's equation at low densities.

The number density of particles of the ath type (obtained by integrating f_a over $\boldsymbol{\xi}_i$), is n_a; the corresponding mass density is $\rho_a=m_an_a$, and the total mass density is $\rho=\Sigma_a\rho_a$. Two different types of conditional means will be required: one, symbolized by $\langle\cdots\rangle_a$, is conditional on the presence of a particle of

[1] R. Kubo, J. Phys. Soc. Japan **12**, 570 (1957).
[2] H. Mori, Phys. Rev. **112**, 1829 (1958).
[3] E. Montroll, in *Termodinamica dei Processi Irreversibili* (Nicola Zanichelli, Bologna, 1960), pp. 217–261.
[4] J. G. Kirkwood, reference 3, pp. 205–216.
[5] J. A. McLennan, Phys. Fluids **3**, 493 (1960).
[6] M. S. Green, J. Chem. Phys. **22**, 398 (1954).
[7] M. Born and H. S. Green, Proc. Roy. Soc. (London) **A190**, 27 (1947).
[8] H. Mori, Phys. Rev. **111**, 694 (1958).
[9] E. Montroll, reference 3, pp. 255–260.
[10] H. S. Green, Proc. Phys. Soc. (London) **B69**, 269 (1956).
[11] H. S. Green, *Molecular Theory of Fluids* (North-Holland Publishing Company, Amsterdam, 1952), Chap. 8; also, M. Born and H. S. Green, Proc. Roy. Soc. (London) **A188**, 2 (1946).

the ath type at $\mathbf{x}_i$; the other, symbolized by $\langle\cdots\rangle_i$, is conditional on the presence of a particle, irrespective of type, at $\mathbf{x}_i$. These are related by $\rho\langle\cdots\rangle_i=\Sigma_a\rho_a\langle\cdots\rangle_a$. The mean velocity of a particle of the ath type at $\mathbf{x}_i$ is $\mathbf{u}_a=\langle\boldsymbol{\xi}_i\rangle_a$, but the local mass velocity is $\mathbf{u}=\langle\boldsymbol{\xi}_i\rangle_i$. If $\mathbf{v}_i=\boldsymbol{\xi}_i-\mathbf{u}$, the local diffusion velocity of particles of the ath type is $\mathbf{w}_a=\langle\mathbf{v}_i\rangle_a$, but $\langle\mathbf{v}_i\rangle_i=0$. The local temperature T, also a function of $\mathbf{x}_i$, is given by $3kT=\langle m_i\mathbf{v}_i^2\rangle$ and $\beta=1/(kT)$.

If

$$E_i=\tfrac{1}{2}\mathbf{v}_i^2+\tfrac{1}{2}\Sigma_j\phi_{ab}/m_i, \tag{3}$$

the internal energy per unit mass is $U=\langle E_i\rangle_i$. As shown elsewhere,[7] if $\nabla=\partial/\partial\mathbf{x}_i$ and $\mathbf{P}_i$ is the tensor

$$\mathbf{P}_i=\mathbf{v}_i\mathbf{v}_i-\tfrac{1}{2}\Sigma_j\mathbf{r}_{ij}\nabla\phi_{ab}/m_i, \tag{4}$$

the local pressure tensor is given by $\mathbf{p}=\rho\langle\mathbf{P}_i\rangle_i$. Also, if μ_a is the chemical potential, per unit mass, of the ath constituent, and $\mathbf{Q}_i$ is the vector

$$\mathbf{Q}_i=(E_i-\mu_a)\mathbf{v}_i-\tfrac{1}{2}\Sigma_j\mathbf{r}_{ij}\mathbf{v}_i\cdot\nabla\phi_{ab}/m_i, \tag{5}$$

the local thermal flux is $\mathbf{q}=\rho\langle\mathbf{Q}_i\rangle$.

To summarize the required macroscopic equations, if

$$D=(\partial/\partial t)+\mathbf{u}\cdot(\partial/\partial\mathbf{x}_i), \tag{6}$$

the hydrodynamical equations read

$$\begin{aligned} D\rho+\rho\nabla\cdot\mathbf{u}=\rho D(\rho_a/\rho)+\nabla\cdot(\rho_a\mathbf{w}_a)=0,\\ \rho D\mathbf{u}+\nabla\cdot\mathbf{p}+\Sigma_a\rho_a\nabla\psi_a=0. \end{aligned} \tag{7}$$

The coefficients of volume and shearing viscosity (ζ and η) are defined in

$$\begin{aligned} \mathbf{p}=p\boldsymbol{\delta}-\zeta\nabla\cdot\mathbf{u}\boldsymbol{\delta}-2\eta\nabla_s\mathbf{u},\\ \nabla_s\mathbf{u}=\tfrac{1}{2}(\nabla\mathbf{u}+\mathbf{u}\nabla)-\nabla\cdot\mathbf{u}\boldsymbol{\delta}/3, \end{aligned} \tag{8}$$

where p is the hydrostatic pressure and $\boldsymbol{\delta}$ is the unit tensor. The equation of energy transport reads

$$\rho DU+(\mathbf{p}\cdot\nabla)\cdot\mathbf{u}+\nabla\cdot(\mathbf{q}+\Sigma_a\rho_a\mu_a\mathbf{w}_a)+\Sigma_a\rho_a\mathbf{w}_a\cdot\nabla\psi_a=0,$$

and the coefficients of diffusion and thermal conduction are defined in

$$\begin{aligned} \rho_a\mathbf{w}_a=-\Sigma_b D_{ab}\nabla\lambda_b-D_a\nabla T/T,\\ \mathbf{q}=-\Sigma_b D_b\nabla\lambda_b-K\nabla T/T, \end{aligned} \tag{9}$$

where

$$\lambda_b=\mu_b+\psi_b. \tag{10}$$

The entropy per unit mass, S, satisfies

$$\begin{aligned} TS=U+(p-\Sigma_a\rho_a\mu_a)/\rho\\ S\nabla T=(\nabla p-\Sigma_a\rho_a\nabla\mu_a)/\rho. \end{aligned} \tag{11}$$

3. DERIVATION OF TRANSPORT COEFFICIENTS

The Chapman-Enskog method[12] of solving Boltzmann's equation (2) proceeds, in effect, by expanding the distribution functions f_a, f_b, $\cdots$ as Taylor's series in the gradients $\nabla\mathbf{u}$, ∇T, and $\nabla\lambda_a$ (and the higher-space derivatives $\nabla\nabla\mathbf{u}$, $\nabla\nabla T$, etc., if required). Thus, f_a is expressed in the form $f_a=f_a^0+f_a^{(1)}+f_a^{(2)}+\cdots$, where $f_a^{(1)}$ is linear in the gradients, $f_a^{(2)}$ is quadratic in the gradients and linear in the second derivatives, etc., and

$$f_a^0=n_a(\tfrac{1}{2}\beta m_a/\pi)^{\frac{3}{2}}\exp(-\tfrac{1}{2}\beta m_a\mathbf{v}_i^2) \tag{12}$$

is Maxwellian in form but depends on the space and time variables n_a, β, and $\mathbf{u}$ (in $\mathbf{v}_i=\boldsymbol{\xi}_i-\mathbf{u}$) instead of the constants of equilibrium. The generalization of this method requires a similar expansion of the function $F_N=F_N^0+F_N^{(1)}+F_N^{(2)}+\cdots$. For the determination of the transport coefficients, only the first two terms in the expansion are needed, so that one can consistently discord terms quadratic in the gradients or involving higher derivatives. Thus (1) reduces to

$$dF_N^{(1)}/dt_N=-dF_N^0/dt_N. \tag{13}$$

The correct choice of F_N^0 is clearly of great importance; it is essential that the mean values $\langle\boldsymbol{\xi}_i\rangle_i^0$, $\langle m_i\mathbf{v}_i^2\rangle_i^0$, $\langle E_i\rangle_i^0$, computed with F_N^0 instead of F_N, should yield the exact values $\mathbf{u}$, $3kT$, and U, etc. These requirements are met by taking

$$\begin{aligned} F_N^0=\exp[-B+\Sigma_i\beta m_i(\mu_a-E_i)],\\ B=\int\beta p\,d\mathbf{x}_i, \end{aligned} \tag{14}$$

where B, β, μ_a and $\mathbf{u}$ (in $\mathbf{v}_i$) are functions of time, and β, μ_a, and $\mathbf{u}$ depend on $\mathbf{x}_i$ as well. Then one has

$$\begin{aligned} dF_N^0/dt_N=F_N^0\Big[-\int\rho D(\beta p/\rho)d\mathbf{x}_i\Big]\\ +F_N^0\Sigma_i[m_i(\mu_a-E_i)(D+\mathbf{v}_i\cdot\nabla)\beta\\ +m_i\beta(D+\mathbf{v}_i\cdot\nabla)\mu_a+m_i\beta(D\mathbf{u}+\mathbf{v}_i\cdot\nabla\mathbf{u})\cdot\mathbf{v}_i\\ -\tfrac{1}{2}\Sigma_j(r_{ij}\cdot\nabla\mathbf{u})\cdot\nabla\phi_{ab}+\tfrac{1}{2}\mathbf{v}_i\cdot\Sigma_j(\mathbf{r}_{ij}\cdot\nabla\beta)\nabla\phi_{ab}\\ +m_i\beta\mathbf{v}_i\cdot\nabla\psi_a]. \end{aligned} \tag{15}$$

To eliminate the time derivatives, one uses the macroscopic equations, which yield

$$\begin{aligned} D\mathbf{u}=-(\nabla p+\Sigma_a\rho_a\nabla\psi_a)/\rho\\ =-S\nabla T-\Sigma_a\rho_a\nabla\lambda_a/\rho;\\ DS=D(\rho_a/\rho)=0, \end{aligned} \tag{16}$$

correct to terms linear in the gradients. The last two results show that, if θ is any thermodynamic variable,

$$D\theta=(\partial\theta/\partial\rho)D\rho=-\rho(\partial\theta/\partial\rho)\nabla\cdot\mathbf{u}, \tag{17}$$

where $\partial\theta/\partial\rho$ is to be computed keeping S and ρ_a/ρ constant. With these substitutions, (15) reduces to

[12] See S. Chapman and T. G. Cowling, *Mathematical Theory of Nonuniform Gases* (Cambridge University Press, New York, 1939).

$dF_N^0/dt_N=F_N^0R_N$, where

$$R_N=(\Sigma_i m_i\beta P_i+\Delta)\nabla\cdot\mathbf{u}+\Sigma_i m_i\beta\mathbf{P}_i:\nabla_s\mathbf{u}+\Sigma_i m_i\beta\mathbf{v}_i\cdot(\nabla\lambda_a-\Sigma_b\rho_b\nabla\lambda_b/\rho)+\Sigma_i m_i\beta(\mathbf{Q}_i-TS\mathbf{v}_i)\cdot\nabla T/T, \quad (18)$$

where

$$P_i=(\mathbf{v}_i^2-\tfrac{1}{2}\Sigma_j\mathbf{r}_{ij}\cdot\nabla\phi_{ab})/3-\rho(\partial U_a/\partial\rho)+(E_i-U_a)\rho kT(\partial\beta/\partial\rho),$$

$$U_a=\mu_a+TS-p/\rho, \quad (19)$$

$$\Delta=\int\rho^2\partial(\beta p/\rho)/\partial\rho d\mathbf{x}_i-\Sigma_i m_i\rho\partial(\beta p/\rho)/\partial\rho.$$

To obtain the Kubo-type formulas, one integrates (13) in the form

$$F_N^{(1)}=F_N^0\int_{-\infty}^{t}R_N(\mathbf{x}',\boldsymbol{\xi}')dt', \quad (20)$$

where $\mathbf{x}_i'$ and $\boldsymbol{\xi}_i'$ are the positions and velocities of a set of particles at time t', which by their natural motions reach the positions $\mathbf{x}_i$ with velocities $\boldsymbol{\xi}_i$ at time t. (At this stage the variability of β, μ_a, and $\mathbf{u}$ with time and place may be forgotten.) By using the approximation $F_N=F_N^0+F_N'$ to compute the means $\mathbf{p}=\rho\langle\mathbf{P}_i\rangle_i$, $\mathbf{w}_a=\langle\boldsymbol{\xi}_i\rangle_a$, and $\mathbf{q}=\rho\langle\mathbf{Q}_i\rangle_i$, one recovers the macroscopic equations (8) and (9), with the following values for the coefficients:

$$\zeta=\beta\rho\left\langle P_i\int_{-\infty}^{t}\Sigma_j m_j P_j'dt'\right\rangle_i^0$$

$$\eta=(\beta\rho/5)\left\langle\mathbf{P}_i:\int_{-\infty}^{t}\Sigma_j m_j\mathbf{P}_j'dt'\right\rangle_i^0$$

$$D_{ab}=(\beta\rho_a/3)\left\langle\mathbf{v}_i\cdot\int_{-\infty}^{t}(\Sigma_j^{(b)}m_j\mathbf{v}_j'-\rho_b\Sigma_j m_j\mathbf{v}_j'/\rho)dt'\right\rangle_a^0 \quad (21)$$

$$D_a=(\beta\rho_a/3)\left\langle\mathbf{v}_i\cdot\int_{-\infty}^{t}\Sigma_j m_j(\mathbf{Q}_j'-TS\mathbf{v}_j')dt'\right\rangle_a^0$$

$$K=(\beta\rho/3)\left\langle\mathbf{Q}_i\cdot\int_{-\infty}^{t}\Sigma_j(\mathbf{Q}_j'-TS\mathbf{v}_j')dt'\right\rangle_i^0,$$

where $\langle\cdots\rangle_i^0$ and $\langle\cdots\rangle_a^0$ are means formed with the equilibrium distribution function, the summation $\Sigma_j^{(b)}$ is restricted to particles of the bth type, and the primed quantities have $\mathbf{x}_i'$ and $\boldsymbol{\xi}_i'$ as arguments instead of $\mathbf{x}_i$ and $\boldsymbol{\xi}_i$.

Apart from the coefficients ζ and D_{ab}, these results agree with Mori's.[2] The discrepancy affecting the coefficients of diffusion is perhaps the most serious, as it also affects Kubo's electrical conductivity formula; the latter is (as we shall see) actually divergent. Kubo's method is to focus attention on the effects of the external force on the system, and ignores relaxation effects. Physically, an external force tends to produce a compensating pressure gradient, which is an independent cause of diffusion, and modifies the direct relationship between the external force and the drift velocities. The fact that $\nabla\lambda_a-\Sigma_b\rho_b\nabla\lambda_b/\rho$, rather than $\nabla\lambda_a$, must appear in (18), is because only differences in force per unit mass can cause diffusion.

It is convenient to notice here the relation

$$D_{aa}=-D_{ab}=-D_{ba}=D_{bb}=(\beta n^2/n_a n_b)\left(\frac{\rho_a\rho_b}{\rho}\right)^2D_{12} \quad (22)$$

between the coefficients of (9), when only two different types of particles are present, and the coefficient D_{12} discussed in Chapman and Cowling's book.[12] The coefficient of self-diffusion is obtained from D_{12} by identifying the two types of particles.

4. EVALUATION FOR DILUTE GASES

In this section it will be shown that the evaluation of the formulas (21) for dilute gases always leads to exactly the same results as the Chapman-Enskog method.[12] For this purpose the coefficient of volume viscosity can be disregarded, as this tends to zero with the density; the same applies to contributions to the viscosity and thermal conduction which depend explicitly on the interparticle potential ϕ_{ab} in (21). One therefore may set $\mathbf{P}_i=\mathbf{v}_i\mathbf{v}_i$ and

$$\mathbf{Q}_i-TS\mathbf{v}_i=\tfrac{1}{2}(\mathbf{v}_i^2-C_akT/m_a),\quad C_a=5+2\log(n_a/T^{\frac{3}{2}})-2\Sigma_b(m_an_b/\rho)\times\log(n_b/T^{\frac{3}{2}}). \quad (23)$$

The task is then to evaluate expressions of the general type

$$I(\sigma,\tau)=n_a\left\langle\sigma_a(\mathbf{v}_i)\int_{-\infty}^{t}\Sigma_j\tau_j(\mathbf{v}_j')dt'\right\rangle_a^0. \quad (24)$$

As these are independent of the time t, it is permitted to set $t=0$; also, because the microscopic motion is reversible, it is permitted to change the sign of t', $\mathbf{v}_i$, and $\mathbf{v}_j'$ simultaneously. Assuming that σ and τ_j are either both even, or both odd, functions of the velocities, one then has

$$I(\sigma,\tau)=n_a\left\langle\sigma_a(\mathbf{v}_i)\int_{0}^{\infty}\Sigma_j\tau_j(\mathbf{v}_j')dt'\right\rangle_a^0. \quad (25)$$

To evaluate this expression, one has to separate, from the grand ensemble of statistical equilibrium, the subensemble for which the velocity $\mathbf{v}_i$ of the ith particle has a definite value, $\mathbf{v}_0$ say, at time $t=0$. If $f_b(t,\mathbf{v})$ denotes the special velocity distribution function for

this subensemble, one has

$$I(\sigma,\tau)=V\int f_a^0(\mathbf{v}_0)\sigma_a(\mathbf{v}_0)\times\int_0^\infty\Sigma_b\int f_b(t,\mathbf{v})\tau_b(\mathbf{v})d\mathbf{v}dtd\mathbf{v}_0, \quad (26)$$

where V is the volume. By hypothesis, at the initial time $t=0$, f_b must reduce to

$$f_b(0,\mathbf{v})=f_b^0(\mathbf{v})+\delta_{ab}\delta(\mathbf{v}-\mathbf{v}_0)/V. \quad (27)$$

At the time $t=\infty$, f_b will have reverted to Maxwellian form, but owing to contributions from the distinguished particle, the mass velocity, temperature and the density of particles of the ath type will be slightly displaced from those of the distribution f_b^0. Calling the displaced values $\mathbf{u}+\boldsymbol{\delta}\mathbf{u}$, $T+\delta T$, and $n_a+\delta n_a$, conservation of mass, momentum, and energy require that

$$\begin{aligned} V\delta n_a&=1,\\ \rho V\boldsymbol{\delta}\mathbf{u}&=m_a\mathbf{v}_0,\\ 3nVk\delta T&=m_av_0^2-3kT. \end{aligned} \quad (28)$$

Then, as one sees from (12),

$$f_b(\infty,\mathbf{v})=f_b^0[1+\delta_{ab}\delta n_b/n_b+\beta m_b\mathbf{v}\cdot\boldsymbol{\delta}\mathbf{u}+\tfrac{1}{2}(\beta m_bv^2-3)\delta T/T]. \quad (29)$$

Hence, if

$$f_b(t,\mathbf{v})=f_b^0(\mathbf{v})+g_b(t,\mathbf{v})/V, \quad (30)$$

the boundary conditions on g_b are

$$\begin{aligned} g_b(0,\mathbf{v})&=\delta_{ab}\delta(\mathbf{v}-\mathbf{v}_0)\\ g_b(\infty,\mathbf{v})&=f_b^0[\delta_{ab}/n_b+\beta m_am_b\mathbf{v}\cdot\mathbf{v}_0/\rho\\ &\quad+\tfrac{1}{2}(\beta m_a\mathbf{v}_0^2-3)(\beta m_bv^2-3)/(3n)]. \end{aligned} \quad (31)$$

On substituting (30) into (26), it is clear that if $I(\sigma,\tau)$ is to be volume independent, one must have

$$\Sigma_b\int f_b^0(\mathbf{v})\tau_b(\mathbf{v})d\mathbf{v}=0; \quad (32)$$

also, if the integration with respect to t is to converge, one must have

$$\Sigma_b\int g_b(\infty,\mathbf{v})\tau_b(\mathbf{v})d\mathbf{v}=0. \quad (33)$$

Both conditions are actually satisfied by all the substitutions for $\tau_b(\mathbf{v})$ which have to be made in connection with formulas (21), but by the same criterion Kubo's electrical conductivity formula is divergent.

To determine the function g_b, one substitutes (30) into Boltzmann's equation (2) and neglects the terms of order $1/V^2$, thus obtaining

$$\partial g_b/\partial t=\Sigma_c([g_b,f_c^0]+[f_b^0,g_c]). \quad (34)$$

For computational purposes, only the integral

$$G_b(\mathbf{v})=\int_0^\infty g_b(t,\mathbf{v})dt \quad (35)$$

is required, and this satisfies

$$\Sigma_c([G_b,f_c^0]+[f_b^0,G_c])=g_b(\infty,\mathbf{v})-g_b(0,\mathbf{v}), \quad (36)$$

with $g_b(\infty,\mathbf{v})$ and $g_b(0,\mathbf{v})$ given by (31). It is satisfactory to notice that the right-hand side of this equation satisfies the integrability conditions

$$\begin{aligned} &\int[g_b(\infty)-g_b(0)]d\mathbf{v}=0\\ &\Sigma_bm_b\int[g_b(\infty)-g_b(0)]\mathbf{v}d\mathbf{v}=0\\ &\Sigma_bm_b\int[g_b(\infty)-g_b(0)]v^2d\mathbf{v}=0 \end{aligned} \quad (37)$$

found by Hilbert.[13] A solution, therefore, exists.

Of course G_b, like g_b, is implicitly a function of $\mathbf{v}_0$ as well as $\mathbf{v}$, and if one defines

$$G_b^\sigma(\mathbf{v})=\int f_a^0(\mathbf{v}_0)\sigma_a(\mathbf{v}_0)G_b(\mathbf{v})d\mathbf{v}_0, \quad (38)$$

it follows from (36) that

$$\begin{aligned} &\Sigma_c([G_b^\sigma,f_c^0]+[f_b^0,G_c^\sigma])=g_b^\sigma,\\ &g_b^\sigma(\mathbf{v})=\int f_a^0(\mathbf{v}_0)\sigma_a(\mathbf{v}_0)[g_b(\infty,\mathbf{v})-g_b(0,\mathbf{v})]d\mathbf{v}_0; \end{aligned} \quad (39)$$

also

$$I(\sigma,\tau)=\int\Sigma_bG_b^\sigma(\mathbf{v})\tau_b(\mathbf{v})d\mathbf{v}. \quad (40)$$

Thus, it is strictly necessary to solve only Eq. (39) to compute $I(\sigma,\tau)$. It is now a matter of routine to verify that, for the substitutions for σ_a required to determine the coefficients of viscosity, thermal conduction, and diffusion, (39) reduces to the equations solved by Chapman and Enskog for the same purpose.

[1] Set $\sigma_a=m_a\mathbf{v}_0\mathbf{v}_0$; then

$$g_b^\sigma=-f_b^0[\delta_{ab}(m_b\mathbf{v}\mathbf{v}-kT\boldsymbol{\delta})-n_akT\boldsymbol{\delta}(\beta m_bv^2-3)/(3n)]$$

and

$$\Sigma_ag_b^\sigma=-f_b^0m_b(\mathbf{v}\mathbf{v}-v^2\boldsymbol{\delta}/3).$$

Thus, by summing Eq. (39) with respect to the implicit suffix a, one obtains the equation used by Chapman and Enskog to determine the viscosity of a gas mixture.

[13] D. Hilbert, Math. Ann. 72, 562 (1912).

[2] Set $\sigma_a=\frac{1}{2}m_a v_0^2 v_0$; then

$$g_b{}^\sigma=-\tfrac{1}{2}f_b{}^0(\delta_{ab}m_b v^2-5kT\rho_a/\rho)\mathbf{v},$$

and

$$\Sigma_a g_b{}^\sigma=-\tfrac{1}{2}f_b{}^0(m_b v^2-5kT)\mathbf{v}$$

which, in conjunction with (39), yields the equation used by Chapman and Enskog to determine the thermal conductivity of a gas mixture.

[3] Finally, set $\sigma_a=m_a\mathbf{v}_0$; then

$$g_b{}^\sigma=-f_b{}^0 m_b(\delta_{ab}-\rho_a/\rho)\mathbf{v}.$$

Thus, for a binary mixture, (39) reduces to

$$\Sigma_c([G_a{}^\sigma,f_c{}^0]+[f_a{}^0,G_c{}^\sigma])=-(\rho_a\rho_b/\rho)(f_a{}^0/n_a)\mathbf{v}$$
$$\Sigma_c([G_b{}^\sigma,f_c{}^0]+[f_b{}^0,G_c{}^\sigma])=(\rho_a\rho_b/\rho)(f_b{}^0/n_b)\mathbf{v},$$

which are equivalent to the equations used by Chapman and Enskog to determing the coefficients of diffusion. Bearing in mind the relation (22) between Chapman's D_{12} and the coefficients D_{ab} of (21), the formula

$$D_{ab}=(\beta/3)\int\Sigma_c G_c{}^\sigma(\mathbf{v})m_c(\delta_{bc}-\rho_b/\rho)\cdot\mathbf{v}d\mathbf{v}$$

derived from (21) is in exact agreement with that deduced by the Chapman-Enskog method.

5. EXTENSION TO DENSE SYSTEMS

It is, of course, to be hoped that ways will be found to evaluate the Kubo-type transport coefficients, other than the reduction to the Chapman-Enskog method outlined for dilute gases in the last section. However, the most straightforward approach to the theory of dense fluids seems to be along similar lines. To begin with, the necessary generalization of the Chapman-Enskog method will be summarized.

In integrating the equation $dF_N{}^{(1)}/dt_N=F_N{}^0R_N$, which was derived in Sec. (3), the simplest procedure is to regard $F_N{}^{(1)}$ as a linear form in the gradients $\nabla\mathbf{u}$, ∇T and $\nabla\lambda_a$, with coefficients which are functions of the relative coordinates and velocities. When this is done, the fact that $\mathbf{u}$ and the thermodynamic functions depended originally on time and position can be ignored, and the equation to be solved becomes simply

$$\Sigma_i\left(\mathbf{v}_i\cdot\frac{\partial F_N{}^{(1)}}{\partial\mathbf{x}_i}-\frac{1}{m_i}\frac{\partial\Phi_N}{\partial\mathbf{x}_i}\cdot\frac{\partial F_N{}^{(1)}}{\partial\mathbf{v}_i}\right)=F_N{}^0R_N$$
$$=F_N{}^0\Sigma_i m_i(\mathbf{a}\cdot\mathbf{v}_i+bE_i+c+\mathbf{A}:\mathbf{P}_i+\mathbf{B}\cdot\mathbf{Q}_i),\quad(41)$$

with constant coefficients $\mathbf{a}$, b, c, $\mathbf{A}$, and $\mathbf{B}$ which can be inferred from (18). Equation (41) can be integrated explicitly, the result being

$$F_N{}^{(1)}=F_N{}^0[\Sigma_i m_i(\mathbf{a}\cdot\mathbf{x}_i+bE_i\tau+c\tau+\mathbf{A}_i(\mathbf{x}_i\mathbf{v}_i)$$
$$+\mathbf{B}\cdot\mathbf{x}_iE_i)]_{-\infty}{}^t,\quad(42)$$

where τ is the timelike solution of the equation $d_N F_N{}^{(1)}/dt_N=1$. As $[\tau]_{-\infty}{}^t$ is divergent, this cannot appear in the expression for any physical quantity. From the above expression for $F_N{}^{(1)}$ it is possible, in principle, to compute all the transport coefficients, though the difficulties are, of course, very great in practice. Several methods of approximation are available, based mostly on the substitution of a slightly retarded time $t-\tau$, in place of $-\infty$ in (42). These will not be discussed here.

It may be mentioned, however, that the result of (42) can be used to deduce an exact solution of the linearized Boltzmann's equation. For instance, an exact solution of

$$[G,f^0]=-mf^0(\mathbf{vv}-v^2\boldsymbol{\delta}/3)$$

obtained in this way is

$$G=C[\mathbf{xv}-\mathbf{x}\cdot\mathbf{v}\boldsymbol{\delta}/3],$$

where C is a constant, equal to $15(4\pi a^3\beta n)^{-1}$ for rigid spheres, if a is the diameter. The coordinate $\mathbf{x}$ is to be interpreted as the mean displacement of a particle with velocity $\mathbf{v}$ from the mass center, but the point of origin does not affect the value of $[G,f^0]$.

Finally, it may be remarked that the coefficients (21) are exactly what one would compute using the formal solution (42) of Liouville's equation. For instance, the expression for D_{ab} is simply

$$D_{ab}=(\beta m_a/3V)\Sigma_N\int F_N{}^0\mathbf{v}_i$$
$$\times[\Sigma_j{}^{(b)}m_j\mathbf{x}_j-\rho_b\Sigma_j m_j\mathbf{x}_j/\rho]_{-\infty}{}^t d\Omega_N,$$

where $d\Omega_N=\Pi_i d\mathbf{x}_i d\mathbf{v}_i/\Pi_a N_a!$

Nonequilibrium Processes in Isotopically Disordered Crystals*

ALEXEI A. MARADUDIN†
Westinghouse Research Laboratories, Pittsburgh, Pennsylvania, and
Institute for Fluid Dynamics and Applied Mathematics, University of Maryland, College Park, Maryland

AND

GEORGE H. WEISS AND DONALD W. JEPSEN‡§
Institute for Fluid Dynamics and Applied Mathematics, University of Maryland, College Park, Maryland
(Received November 4, 1960)

The equations of motion of the atoms in an isotopically disordered crystal, which contains a fraction p of atoms of mass M_1, and a fraction $1-p$ of atoms of mass M_2, are expanded in terms of the normal coordinates of a monatomic lattice whose atoms all have mass $M=pM_1+(1-p)M_2$. The equations of motion of these normal coordinates are derived and are then solved by Laplace transform methods. The perturbed normal coordinates are found to decay exponentially into the future and into the past until an inverse power dependence on time becomes dominant. Calculations of the mean lifetime and frequency shift of each normal coordinate are carried out for the one-dimensional case. A theory of the optical absorption spectrum of an isotopically disordered ionic crystal is obtained, and the distribution function for the energies of the normal modes and the mean energy in a normal mode are found. The generalization of the methods of this paper to three-dimensional lattices is discussed.

I. INTRODUCTION

THE problem we consider in this paper is the following one. We have a crystal lattice at every lattice point of which there is either an atom of mass M_1 or an atom of mass M_2, so that the total number of atoms with mass M_1 is pN, while the total number of atoms of mass M_2 is $(1-p)N$, where N is the total number of atoms in the crystal. We now expand the displacements of the atoms from their equilibrium positions in terms of normal coordinates appropriate to a monatomic lattice, each of whose atoms has mass $pM_1+(1-p)M_2$. In the limit of long waves, the vibrational properties of this mean-mass lattice closely approximate those of the isotopically disordered lattice.[1] We seek to find the evolution in time of these normal coordinates in the presence of the perturbation resulting from the random array of isotopic impurities. We are concerned in this paper with the classical formulation of this problem, and defer a quantum mechanical discussion to a subsequent paper.

In recent years, the vibrational properties of randomly disordered lattices have been studied in some detail theoretically, and a good deal of qualitative and some quantitative information is now available about the effect of randomly distributed impurities on the frequency spectrum[2–5] and thermodynamic functions[6–9] of crystal lattices. Although it is well known that the presence of even small numbers of impurities and defects can profoundly alter the electrical, optical, chemical, and tensile properties of crystals,[10–13] there have been virtually no experiments carried out to test directly any of the predictions of the above-mentioned theories regarding the effects of a random distribution of impurities on the purely vibrational properties of crystals.[14]

The recent emergence of neutron spectrometry as a powerful tool in the study of the dynamical properties of crystal lattices furnishes us with a probe by means of which we can study the effects of defects and disorder on the individual normal modes of a crystal. It is known that in a purely harmonic crystal, a plane wave disturbance of a definite frequency can propagate freely from one end of the crystal to the other unattenuated. In the presence of a random array of isotopic scatterers,

* Portions of this research were supported by the U. S. Air Force Office of Scientific Research.

† Present address: Westinghouse Research Laboratories, Beulah Road, Churchill Borough, Pittsburgh 35, Pennsylvania.

‡ General Motors Fellow in Applied Mathematics and Mathematical Physics.

§ Present address: IBM Research Center, Yorktown Heights, New York.

[1] A. A. Maradudin and G. H. Weiss, J. Chem. Phys. **29**, 631 (1958).

[2] C. Domb, A. A. Maradudin, E. W. Montroll, and G. H. Weiss, Phys. Rev. **115**, 18, 24 (1959).

[3] I. M. Lifshitz and G. I. Stepanova, J. Exptl. Theoret. Phys. (U.S.S.R.) **30**, 938 (1956) [English translation: Soviet Phys.—JETP **3**, 656 (1956)].

[4] P. Dean, Proc. Roy. Soc. (London) **A254**, 507 (1960).

[5] J. Hori, Progr. Theoret. Phys. (Kyoto) **18**, 367 (1957); **23**, 475 (1960).

[6] G. H. Weiss and A. A. Maradudin, J. Phys. Chem. Solids **7**, 327 (1958).

[7] J. Mahanty, A. A. Maradudin, and G. H. Weiss, Progr. Theoret. Phys. (Kyoto) **20**, 369 (1958).

[8] E. W. Montroll, A. A. Maradudin, and G. H. Weiss, *Proceedings of the 1957 Stevens Institute Many Body Conference* (Interscience Publishers, Inc., New York, to be published).

[9] A. A. Maradudin, P. Mazur, E. W. Montroll, and G. H. Weiss, Revs. Modern Phys. **30**, 175 (1958).

[10] See for example, W. Shockley, *Electrons and Holes in Semiconductors* (D. Van Nostrand Company, Inc., Princeton, New Jersey, 1950), Chap. 1.

[11] R. F. Wallis and A. A. Maradudin, Progr. Theoret. Phys. (Kyoto) **24**, 1055 (1960).

[12] See, for example, *Catalysis*, edited by P. H. Emmett (Reinhold Publishing Corp., New York, 1955), Chap. 3.

[13] A. H. Cottrell, *Dislocations and Plastic Flow in Crystals* (Clarendon Press, Oxford, England, 1953), Chap. IV.

[14] In this regard, however, see A. T. Stewart and B. N. Brockhouse, Revs. Modern Phys. **30**, 250 (1958).

however, the various normal modes of the unperturbed lattice are not independent any longer and, in fact, interact with each other. A plane wave disturbance of given wavelength and corresponding frequency in such a lattice undergoes scattering by the impurities and suffers a complex frequency shift in the range of times which are dealt with here. The real part of this shift gives the change in the normal mode frequency of the normal mode associated with the given wavelength. The imaginary part of the shift is directly related to the mean lifetime of the normal coordinate. Both the frequency shift and mean lifetime of a normal coordinate can be measured experimentally by neutron spectrometry as has been shown beautifully by Larsson and co-workers[15] in a study of phonons in an anharmonic crystal. For this reason, it seems worthwhile to examine carefully the problem of the mean lifetime and frequency shift of a normal coordinate in an isotopically disordered lattice, since, hopefully, experimental investigations of these quantities will not be long in coming.

In addition to their effects on the equilibrium properties of lattices, randomly distributed impurities can affect nonequilibrium or transport properties of crystals. There have been several theoretical and experimental studies of the effects of a random array of isotopic impurities on lattice thermal conductivity.[16,17]

Klemens[16] formulated the problem of determining the thermal resistance caused by a random array of isotope impurities as a quantum mechanical scattering problem and calculated the scattering of phonons by the isotope impurities.

More recently, the mean free path of phonons in isotopically disordered crystals has been studied by Mattis.[18] After stating the problem classically, he adapted quantum mechanical time-dependent perturbation theory to obtain the mean lifetime of a phonon to lowest order in the perturbation parameter. However, Mattis makes an approximation in his treatment which is unnecessary, and which, in addition, would have to be removed before corrections to his result could be obtained.

Recent work of Kubo,[19] Mori,[20] and Green[21] has resulted in expressions for transport coefficients in the form of integrals over time-relaxed correlation functions of coordinates and/or momenta. In particular, their expressions for the dielectric susceptibility and thermal conductivity for systems which obey classical statistics are

$$\chi_{\mu\nu}(\omega)=\lim_{\epsilon\to 0+}\beta\int_0^\infty e^{-i\omega t-\epsilon t}\langle\dot{M}_\nu(0)M_\mu(t)\rangle dt, \quad (1.1)$$

$$\kappa_{\mu\nu}=k\beta^2\int_0^\infty\langle J_\mu(0)J_\nu(t)\rangle dt, \quad (1.2)$$

where $\beta=1/kT$. In Eq. (1.1) $\chi_{\mu\nu}(\omega)$ is the dielectric susceptibility tensor, whose imaginary part is directly related to the absorption coefficient of the lattice, and $\mathbf{M}(t)$ expresses the natural motion of the dipole moment of the lattice. In Eq. (1.2), $\kappa_{\mu\nu}$ is the thermal conductivity tensor, while $\mathbf{J}(t)$ represents the natural motion of the heat current density in the lattice. The brackets $\langle\ \rangle$ indicate an average over a canonical ensemble at time $t=0$. Since both $\mathbf{M}(t)$ and $\mathbf{J}(t)$ can be expressed simply in terms of the normal coordinates of the lattice, and since in the present case the value of the normal coordinates at time t can be related to their values at time $t=0$, the evaluation of the correlation functions is greatly simplified. This fact imparts additional interest to the calculation of the time evolution of the normal coordinates of an isotopically disordered lattice. In this paper, we calculate only the optical absorption of an isotopically disordered crystal and defer the discussion of thermal conductivity to a subsequent paper, since the latter result has some independent interest apart from the techniques employed in its calculation.

The effect of a random distribution of isotope defects on the distribution function of the energies of the normal modes has been calculated to lowest order by George,[22] who used the techniques developed by Prigogine and co-workers[23] in their study of irreversible processes in many-body systems.

Although in the initial statement of our problem we postulated that our crystal contains exactly Np atoms of mass M_1 and exactly $N(1-p)$ atoms of mass M_2, it will prove to be convenient in what follows to relax slightly this restriction in the following way.[6] We assume instead that the atom at each lattice site has a mass M_1 with probability p, or a mass M_2 with probability $1-p$. Then the lattice will contain k atoms of mass M_1 with probability

$$\binom{N}{k}p^k(1-p)^{N-k}.$$

If we now consider an ensemble of binomial lattices, then the distribution of the number of lattices with k M_1 atoms is found to be approximately Gaussian with mean Np and standard deviation equal to $[N\cdot p(1-p)]^{\frac{1}{2}}$. As N goes to infinity, the law of large numbers assures us that the probability of a large deviation from the mean number goes to zero. Hence, our following remarks

[15] K. E. Larsson, U. Dahlborg, and S. Holmryd, Arkiv Fysik (to be published).
[16] P. G. Klemens, Proc. Phys. Soc. (London) **A68**, 1113 (1955).
[17] R. Berman, E. L. Foster, and J. M. Ziman, Proc. Roy. Soc. (London) **A237**, 344 (1956).
[18] D. C. Mattis, Phys. Rev. **106**, 721 (1957).
[19] R. Kubo, J. Phys. Soc. Japan **12**, 570 (1957).
[20] H. Mori, Phys. Rev. **112**, 1829 (1958).
[21] M. S. Green, J. Chem. Phys. **22**, 398 (1954).
[22] C. George, Bull. Acad. roy. Belg. **45**, 239 (1959).
[23] R. Brout and I. Prigogine, Physica **22**, 621 (1956), and subsequent papers.

will apply to a good approximation to the problem of a lattice containing exactly Np atoms of mass M_1 randomly placed.

The equations of motion of the normal coordinates in an isotopically disordered lattice are obtained in Sec. II, and are solved formally in Sec. III. The resulting expressions for the normal coordinates are discussed in Sec. IV. In Sec. V, the results of the preceding sections are applied to the calculation of the optical absorption of such lattices. In Sec. VI, the techniques of this paper are employed in a brief discussion of the problem studied by George, viz., the determination of the distribution function of the energies of the normal modes of an isotopically disordered lattice.

The calculations presented here were stimulated by the recent work by Prigogine and co-workers[23] which has gone so far in explaining the origins of irreversibility. The mathematical techniques employed in the present paper, however, are more closely related to those employed by Van Hove and his collaborators[24] in their treatments of "persistent" and "dissipative" effects in many particle systems.

II. EQUATIONS OF MOTION OF THE DISORDERED LATTICE

For simplicity, in the body of this paper we study a lattice which contains only one atom per unit cell, and furthermore, restrict our calculations to the case of a one-dimensional chain. The generalization to three-dimensional lattices containing more than one atom per unit cell is indicated in Appendix A, where it is shown that the results obtained in the one-dimensional case can be carried over to the more general case by a suitable interpretation of the summation index.

We thus consider a one-dimensional chain of N atoms in which each atom interacts with its two nearest neighbors only. The mass m_l at the lth lattice point is either M_1 or M_2.

The equations of motion for our lattice are

$$m_l\ddot{u}_l=\gamma(u_{l+1}-2u_l+u_{l-1}), \tag{2.1}$$

where u_l is the displacement of the lth atom from its equilibrium position, and γ is the force constant for nearest neighbor interactions. Introducing the mean mass of the constituent atoms, which is defined by

$$M=pM_1+(1-p)M_2, \tag{2.2}$$

we rewrite Eq. (2.1) as

$$\ddot{u}_l=\frac{\gamma}{M}(u_{l+1}-2u_l+u_{l-1})+\left(\frac{\gamma}{m_l}-\frac{\gamma}{M}\right)\times(u_{l+1}-2u_l+u_{l-1}). \tag{2.3}$$

We will regard the second term on the right-hand side of Eq. (2.3) as a perturbation on the equations of motion of the monatomic mean-mass lattice given by the first term.

This breakup of the equations of motion into an unperturbed and a perturbed part, with the monatomic mean-mass lattice playing the part of the unperturbed lattice, corresponds to a perturbation of the potential energy of the lattice. An alternative derivation of the equations of motion in which the monatomic mean-mass lattice still plays the role of the unperturbed lattice, but in which the perturbation caused by the isotopic impurities is now associated with the kinetic energy of the lattice is presented in Appendix B.

If we assume that the displacements u_l satisfy the cyclic boundary condition

$$u_l=u_{l+N}, \tag{2.4}$$

we can introduce the normal coordinates Q_k for the unperturbed monatomic mean mass lattice by the relation

$$u_l=\frac{1}{(NM)^{\frac{1}{2}}}\sum_{k=-N/2+1}^{N/2}Q_k\exp\left(\frac{2\pi ikl}{N}\right). \tag{2.5}$$

The normal coordinate Q_k satisfies the reality condition

$$Q_{-k}=Q_k^*. \tag{2.6}$$

If we substitute this expression for u_l into Eq. (2.3), multiply the resulting equation by $\exp\{-2\pi ik'l/N\}$, sum over l, and use the orthogonality properties of the exponentials, we obtain the result that

$$\ddot{Q}_k+\omega_k^2Q_k=-\sum_{k'}\Phi_{kk'}Q_{k'}. \tag{2.7}$$

In Eq. (2.7), ω_k is the kth normal mode frequency of the unperturbed lattice and is given by

$$\omega_k=\omega_L|\sin(\pi k/N)|, \tag{2.8}$$

where

$$\omega_L=(4\gamma/M)^{\frac{1}{2}}. \tag{2.9}$$

The "matrix elements" $\Phi_{kk'}$ are given by

$$\Phi_{kk'}=\frac{\omega_{k'}^2}{N}\sum_l\left(\frac{M}{m_l}-1\right)\exp\left(\frac{2\pi i(k'-k)l}{N}\right). \tag{2.10}$$

Equation (2.7) is the basic equation of this paper.

III. SOLUTION OF THE EQUATIONS OF MOTION

We denote the Laplace transform of $Q_k(t)$ by $q_k(s)$,

$$q_k(s)=\int_0^\infty e^{-st}Q_k(t)dt. \tag{3.1}$$

On taking the Laplace transform of Eq. (2.7), we obtain

$$q_k(s)=\frac{c_k(s)}{d_k}-\frac{1}{d_k}\sum_{k_1}\Phi_{kk_1}q_{k_1}(s), \tag{3.2}$$

[24] L. Van Hove, Physica 21, 901 (1955), and subsequent papers.

where

$$sQ_k(0)+\dot{Q}(0)=c_k(s)$$
$$s^2+\omega_k^2=d_k. \tag{3.3}$$

This set of equations can be solved by iteration,

$$q_k(s)=D_{kk}(s)c_k(s)+\sum_{n(\neq k)} D_{kn}(s)c_n(s), \tag{3.4}$$

where

$$D_{kn}(s)=\frac{\delta_{kn}}{d_k}-\frac{1}{d_kd_n}\Big\{\Phi_{kn}-\sum_{k_1}\frac{\Phi_{kk_1}\Phi_{k_1n}}{d_{k_1}} + \sum_{k_1k_2}\frac{\Phi_{kk_1}\Phi_{k_1k_2}\Phi_{k_2n}}{d_{k_1}d_{k_2}}-\cdots\Big\}. \tag{3.5}$$

We assume that this expansion converges.

We shall now classify and regroup the terms in $D_{kn}(s)$ to obtain results which express explicitly the effects of the perturbation on the unperturbed normal coordinates. We consider first the structure of $D_{kk}(s)$, which for simplicity we relabel $D_k(s)$. It is given explicitly by

$$D_k(s)=\frac{1}{d_k}-\frac{1}{d_k^2}\Big\{\Phi_{kk}-\sum_{k_1}\frac{\Phi_{kk_1}\Phi_{k_1k}}{d_{k_1}} + \sum_{k_1k_2}\frac{\Phi_{kk_1}\Phi_{k_1k_2}\Phi_{k_2k}}{d_{k_1}d_{k_2}}-\cdots\Big\}. \tag{3.6}$$

We now introduce a function $G_k(s)$ which is defined by

$$G_k(s)=\Phi_{kk}-\sum_{k_1}'\frac{\Phi_{kk_1}\Phi_{k_1k}}{d_{k_1}}+\sum_{k_1k_2}'\frac{\Phi_{kk_1}\Phi_{k_1k_2}\Phi_{k_2k}}{d_{k_1}d_{k_2}}-\cdots, \tag{3.7}$$

where the prime on the summations means that no summation index k_j equals k. The function defined by the expansion in curly brackets in Eq. (3.6) can be re-expressed as the sum of all terms in which no summation index equals k, plus the sum of all terms in which one index equals k, $\cdots$, and so on. In terms of the function $G_k(s)$, we can thus express $D_k(s)$ as

$$D_k(s)=\frac{1}{d_k}-\frac{1}{d_k^2}\Big\{G_k(s)-\frac{1}{d_k}G_k^2(s)+\frac{1}{d_k^2}G_k^2(s)-\cdots\Big\} =\frac{1}{d_k}-\frac{1}{d_k^2}\frac{G_k(s)}{1+(1/d_k)G_k(s)}=\frac{1}{d_k+G_k(s)}. \tag{3.8}$$

The function $G_k(s)$ is analogous to the simple diagonal part of an operator introduced by Van Hove.[24]

It should be mentioned at this point that it is also possible to regroup and resum the terms in the expansion of $G_k(s)$ in such a way that the occurrence of repeated intermediate indices $k_1, k_2, \cdots, k_n$, in $(n+1)$st order, is eliminated. This elimination is accompanied by a "renormalization" of the denominators $(s^2+\omega_{k_j}^2)^{-1}$ which appear in the expansion of $G_k(s)$, in the sense that they are replaced by expressions of the form $s^2+\omega_{k_j}^2+G_{k_j}'(s)$ where the prime on $G_{k_j}'(s)$ indicates that there are certain restrictions on the intermediate states in the expansion of the function. In the absence of such a renormalization, we see from Eq. (3.7) that each term in the expansion of $G_k(s)$, regarded as a function of the complex variable s, is analytic everywhere in the complex s plane except for a cut along the imaginary axis in the interval $(-i\omega_L, i\omega_L)$, where ω_L is the maximum frequency of the unperturbed lattice. The effect of the renormalization is to alter the length of the cut along the imaginary axis to bring it into closer conformity with the true range of the perturbed frequency spectrum. However, since the results of this renormalization procedure are not used in the present paper, we omit a discussion of this point here and refer the interested reader to footnote 25.

A similar expansion can be obtained for the off-diagonal elements $D_{kn}(s)$. However, in light of the remarks following Eq. (3.16), we could omit an explicit discussion of these elements, at least insofar as they appear in the calculation of the normal coordinates, since they will vanish on taking a configuration average of each $Q_k(t)$. Nevertheless, it seems worthwhile to outline the derivation of the result for $D_{kn}(s)$ since this result is directly applicable to the discussions in Sec. VI and Appendix D.

Explicitly, we have that

$$D_{kn}(s)=-\frac{1}{d_k}\Big\{\Phi_{kn}-\sum_{k_1}\frac{\Phi_{kk_1}\Phi_{k_1n}}{d_{k_1}} + \sum_{k_1k_2}\frac{\Phi_{kk_1}\Phi_{k_1k_2}\Phi_{k_2n}}{d_{k_1}d_{k_2}}-\cdots\Big\}\frac{1}{d_n}. \tag{3.9}$$

We now sum all terms in which no summation index equals k, those terms in which one summation index equals k, $\cdots$, and so on. We thus obtain the result that

$$D_{kn}(s)=-\frac{1}{d_k}D_{kn}^{(k)}(s)\Big[1-\frac{G_k^{(k)}(s)}{d_k} +\Big(\frac{G_k^{(k)}(s)}{d_k}\Big)^2-\cdots\Big]\frac{1}{d_n} =-\frac{D_{kn}^{(k)}(s)}{d_k+G_k^{(k)}(s)}\frac{1}{d_n}, \tag{3.10}$$

where $G_k^{(k)}(s)\equiv G_k(s)$ is defined by Eq. (3.7), and $D_{kn}^{(k)}(s)$ is given by

$$D_{kn}^{(k)}(s)=\Phi_{kn}-\sum_{k_1\neq k}\frac{\Phi_{kk_1}\Phi_{k_1n}}{d_{k_1}} =\sum_{\substack{k_1\neq k\\ k_2\neq k}}\frac{\Phi_{kk_1}\Phi_{k_1k_2}\Phi_{k_2n}}{d_{k_1}d_{k_2}}-\cdots. \tag{3.11}$$

[25] A. A. Maradudin, G. H. Weiss, and D. W. Jepsen, Westinghouse Research Lab. Sci. Paper 6-41402-6-P7 (September 7, 1960).

The superscript k means that no summation index equals k.

We continue this process by summing all terms in the expansion (3.11) in which no summation index equals n, those terms in which one summation index equals n, $\cdots$, and so on. The result of these operations is that

$$D_{kn}^{(k)}(s)=D_{kn}^{(kn)}(s)\left[1-\frac{G_n^{(kn)}(s)}{d_n}+\left(\frac{G_n^{(kn)}(s)}{d_n}\right)^2-\cdots\right]=\frac{d_n D_{kn}^{(kn)}(s)}{d_n+G_n^{(kn)}(s)},\tag{3.12}$$

where $G_n^{(kn)}(s)$ is given by

$$G_n^{(kn)}(s)=\Phi_{nn}-\sum_{k_1\neq k,n}\frac{\Phi_{nk_1}\Phi_{k_1n}}{d_{k_1}}+\sum_{\substack{k_1\neq k,n\\ k_2\neq k,n}}\frac{\Phi_{nk_1}\Phi_{k_1k_2}\Phi_{k_2n}}{d_{k_1}d_{k_2}}-\cdots,\tag{3.13}$$

and $D_{kn}^{(kn)}(s)$ is given by

$$D_{kn}^{(kn)}(s)=\Phi_{kn}-\sum_{k_1\neq k,n}\frac{\Phi_{kk_1}\Phi_{k_1n}}{d_{k_1}}+\sum_{\substack{k_1\neq k,n\\ k_2\neq k,n}}\frac{\Phi_{kk_1}\Phi_{k_1k_2}\Phi_{k_2n}}{d_{k_1}d_{k_2}}-\cdots.\tag{3.14}$$

Combining Eqs. (3.12) and (3.10), we find that

$$D_{kn}(s)=-\frac{1}{s^2+\omega_k^2+G_k^{(k)}(s)}\times D_{kn}^{(kn)}(s)\frac{1}{s^2+\omega_n^2+G_n^{(kn)}(s)}.\tag{3.15}$$

We could continue to regroup and sum terms in the expansion for $D_{kn}^{(kn)}(s)$ by picking out all terms in which no summation index equals k_1, all terms in which one index equals k_1, $\cdots$, and so on. The net result would be to replace d_{k_1} by $d_{k_1}+G_{k_1}^{(knk_1)}$ and to impose the additional restriction that no summation index can now equal k_1. Clearly, this process can be continued indefinitely until every function d_p is replaced by $d_p+G_p^{(\cdots)}$ and in any term no summation index can equal k, n, or any other summation index which appears in that term. Then every function $G_p^{(\cdots)}$ appearing in the expansion for $D_{kn}(s)$ can itself be "renormalized" in the way described in the discussion following Eq. (3.8).[26] However, in this case at least, what is possible in principle becomes rather difficult to realize in practice, and we restrict ourselves here to writing out the explicit expansion for $D_{kn}(s)$ to the two lowest orders in the perturbation:

$$D_{kn}(s)=\frac{1}{s^2+\omega_k^2+G_k^{(k)}(s)}\Phi_{kn}\frac{1}{s^2+\omega_n^2+G_n^{(kn)}(s)}+\frac{1}{s^2+\omega_k^2+G_k^{(k)}(s)}\sum_{k_1\neq k,n}\frac{\Phi_{kk_1}\Phi_{k_1n}}{s^2+\omega_{k_1}^2+G_{k_1}^{(knk_1)}(s)}\times\frac{1}{s^2+\omega_n^2+G_n^{(kn)}(s)}-\cdots.\tag{3.16}$$

In obtaining expressions for the normal coordinates and for quantities which depend on them, we will not be interested so much in results appropriate to any particular configuration of the isotope impurities at the lattice sites of our crystal, but rather in ensemble averages of these quantities over all possible configurations of the impurities in the crystal.[6] Thus for example, we will have expressions of the form

$$\left\langle\frac{s}{s^2+\omega_k^2+G_k(s)}\right\rangle_{\text{configuration}}$$

to evaluate. It is, however, more convenient to evaluate an expression which contains averages only of the form $\langle G_k(s)\rangle$ rather than the more complicated average indicated above. It turns out to be the case for the present model that

$$\left\langle\frac{1}{s^2+\omega_k^2+G_k(s)}\right\rangle_{\text{config}}=\frac{1}{s^2+\omega_k^2+\langle G_k(s)\rangle},\tag{3.17}$$

where this result is exact to any order of the perturbation. Equation (3.17) is most readily obtained by taking the configuration average of $D_k(s)$, Eq. (3.6), term-by-term with the aid of the results of Appendix C, and then regrouping the terms in the expansion in the manner described already. This point is discussed more fully by Langer in a forthcoming paper.[27]

This result together with the result established in Appendix C that the configuration average of a product of matrix elements $\langle\Phi_{kk_1}\Phi_{k_1k_2}\cdots\Phi_{k_jn}\rangle$ vanishes unless $k=n$, means that only the diagonal part of $Q_k(t)$ is nonvanishing after averaging over all configurations. This conclusion, of course, does not apply to the configuration average of expressions such as $Q_k(t)Q_k^*(t)$, to which the nondiagonal terms of $Q_k(t)$ contribute.

IV. TIME DEPENDENCE OF $Q_k(t)$

At this point we must refer to two systems in a discussion of asymptotic properties of $Q_k(t)$. The first consists of a finite lattice of N atoms, and the second consists of a strictly infinite number of atoms. The physical properties of either system depend, through

[26] P. W. Anderson, Phys. Rev. **109**, 1492 (1958).

[27] J. S. Langer (private communication).

$Q_k(t)$, on the function $G_k(s)$. In a finite system, the solution for $Q_k(t)$ can be written

$$Q_k(t)=\sum_n A_{kn}e^{i\lambda_{kn}t},$$

where the λ_{kn} are real, i.e., the roots of $s^2+\omega_k^2+G_k(s)=0$ all lie on the imaginary axis, and in fact there are $2N$ poles since there are N degrees of freedom. Furthermore, no matter what the collective behavior of the $Q_k(t)$ might appear to be over a fixed length of time, all of the $Q_k(t)$ cannot ultimately tend to zero since there is no physical mechanism available to dissipate energy. The situation is otherwise in a system with an infinite number of particles. From a mathematical point of view this can be inferred from the behavior of $G_k(iy)$. For a finite system, to second order in the perturbation $G_k(iy)$ is given by

$$G_k(iy)\sim\Phi_{kk}-\sum_{k_1\neq k}^{N}\frac{\Phi_{kk_1}\Phi_{k_1k}}{\omega_{k_1}^2-y^2}, \tag{4.1}$$

and it is seen that $G_k(iy)$ has $2N$ poles at $y=\pm\omega_1, \pm\omega_2, \cdots, \pm\omega_N$. These poles all lie in a fixed interval, say $(-\omega_L, \omega_L)$. If the upper limit on the sum is taken to be infinity, $G_k(iy)$ has an infinite number of poles which are still confined to $(-\omega_L, \omega_L)$. Since these poles are densely distributed throughout this interval, a branch cut must be introduced in the s plane along the s interval $(-i\omega_L, i\omega_L)$. This branch cut introduces the property

$$\lim_{\eta\to 0+}G_k(iy\pm\eta)=-K_k(y)\pm iJ_k(y), \tag{4.2}$$

where $K_k(y)$ and $J_k(y)$ are real functions of the real variable y, which in turn gives rise, as we shall see, to dissipative properties. In considering a system with an infinite number of atoms, we are in effect making the Poincare cycle infinite. Thus, we cannot expect that the expressions obtained below are valid for all time in describing finite systems, no matter how large these systems may be.

According to the result of the preceding section, the solution for the configuration average of the normal coordinate $Q_k(t)$ to all orders in the perturbation can be expressed as

$$\langle Q_k(t)\rangle=\frac{Q_k(0)}{2\pi i}\int_{c-i\infty}^{c+i\infty}\frac{se^{st}}{s^2+\omega_k^2+\langle G_k(s)\rangle}ds + \frac{\dot{Q}_k(0)}{2\pi i}\int_{c-i\infty}^{c+i\infty}\frac{e^{st}}{s^2+\omega_k^2+\langle G_k(s)\rangle}ds. \tag{4.3}$$

Since the first integral is just the time derivative of the second, we need only consider the latter, which we denote by

$$I_k(t)=\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty}\frac{e^{st}}{s^2+\omega_k^2+\langle G_k(s)\rangle}ds. \tag{4.4}$$

The function $\langle G_k(s)\rangle$ satisfies the condition expressed by Eq. (4.2). This result follows from the definition of the function $G_k(s)$, Eq. (3.7), in the limit that the sums over the indices k_j are replaced by integrals. The leading terms in the expansions of $K_k(y)$ and $J_k(y)$ are given explicitly by[28]

$$K_k(y)=-\Phi_{kk}+\sum_{k_1}{}'\frac{\Phi_{kk_1}\Phi_{k_1k}}{(\omega_{k_1}^2-y^2)_P}, \tag{4.5a}$$

$$J_k(y)=\pi\,\mathrm{sgn}y\sum_{k_1}{}'\Phi_{kk_1}\Phi_{k_1k}\delta(\omega_{k_1}^2-y^2), \tag{4.5b}$$

so that with the aid of the results of Appendix C we find that

$$\langle K_k(y)\rangle=-\omega_k^2\mu_1+\omega_k^2(\mu_2-\mu_1^2)\frac{1}{N}\sum_{k_1}\frac{\omega_{k_1}^2}{(\omega_{k_1}^2-y^2)_P}, \tag{4.6a}$$

$$\langle J_k(y)\rangle=\pi\omega_k^2(\mu_2-\mu_1^2)\frac{\mathrm{sgn}y}{N}\sum_{k_1}\omega_{k_1}^2\delta(\omega_{k_1}^2-y^2), \tag{4.6b}$$

where it is understood that the sums are to be evaluated only in the limit as $N\to\infty$, and it is seen that

$$K_k(-y)=K_k(y), \quad J_k(y)=-J_k(-y). \tag{4.7}$$

These results are general, and hold in all orders of the perturbation.

In Eq. (4.4) we make a change of variable, $s=iy+c$, and pass to the limit as $c\to 0+$. With the aid of Eq. (4.2) we obtain

$$I_k(t)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{e^{iyt}}{-y^2+\omega_k^2-\langle K_k(y)\rangle+i\langle J_k(y)\rangle}dy \tag{4.8a}$$

$$\equiv\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{e^{iyt}}{U_k(y)+iV_k(y)}dy. \tag{4.8b}$$

The function $V_k(y)$ vanishes for all $|y|>y_0$, where y_0 is the magnitude of the maximum frequency of the unperturbed lattice. Similarly, the function $U_k(y)$ has simple zeroes for values of y which we denote by $y=\pm y_k$, where $|y_k|\leqslant y_0$. In the absence of the perturbation resulting from the random distribution of isotopic impurities, $V_k(y)$ vanishes identically, and the roots $\pm y_k$ of the equation $U_k(y)=0$ are the unperturbed normal mode frequencies of the monatomic mean mass lattice. In the presence of the perturbation, the denominator of the integrand in Eq. (4.8b) never vanishes. Following van Hove[24] and Hugenholtz,[29] however, we assume that for large t the dominant contribution to the integral

[28] The expression $(1/x)_P$ is defined by

$$\left(\frac{1}{x}\right)_P=\lim_{\epsilon\to 0+}\frac{x}{x^2+\epsilon^2}.$$

In taking the double limit $\epsilon\to 0+$ and $N\to\infty$, we always ensure that ϵ remains larger than the spacing between consecutive unperturbed frequencies.

[29] N. M. Hugenholtz and L. Van Hove, Physica 24, 363 (1958); N. M. Hugenholtz, *The Many Body Problem* (John Wiley & Sons, Inc., New York, 1959), p. 33.

comes from the portion of the y axis about the points where $U_k(y)$ vanishes. If we expand $U_k(y)+iV_k(y)$ about the points $\pm y_k$ and retain the leading nonvanishing contributions to the real and imaginary parts, $I_k(t)$ becomes approximately

$$I_k(t)\sim\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{iyt}\left\{\frac{1}{(y-y_k)U_k'(y_k)+iV_k(y_k)}+\frac{1}{(y+y_k)U_k'(-y_k)+iV_k(-y_k)}\right\}dy$$
$$=-\frac{N_k}{2\pi}\int_{-\infty}^{\infty}e^{iyt}\left\{\frac{1}{y-y_k-i\Gamma_k}-\frac{1}{y+y_k-i\Gamma_k}\right\}dy, \quad (4.9)$$

where

$$N_k^{-1}=|U_k'(y_k)| \quad \Gamma_k=N_kV_k(y_k), \quad (4.10)$$

where $y_k>0$, and we have used the fact that since $U_k(y)$ is a function of y^2 only, $U_k'(-y)=-U_k'(y)$. On closing the integration contour with a semicircle of infinite radius in the upper half-plane, we obtain finally

$$I_k\sim 2N_ke^{-\Gamma_kt}\sin y_kt. \quad (4.11)$$

With this result, we can write finally

$$\langle Q_k(t)\rangle\sim 2N_k\{Q_k(0)e^{-\Gamma_kt}y_k\cos y_kt+\dot{Q}_k(0)e^{-\Gamma_kt}\sin y_kt\}. \quad (4.12)$$

Had we initially been concerned with negative times, we would have found the same result as that just given except that the factor $e^{-\Gamma_kt}$ would now be replaced by $e^{-\Gamma_k|t|}$. $Q_k(t)$ thus depends on both t and $|t|$, and its dependence on $|t|$ demonstrates the irreversible character of the motion.

We now solve for y_k in the first approximation. The equation $U_k(y_k)=0$ becomes explicitly

$$-y_k^2+\omega_k^2=-\omega_k^2\mu_1+\omega_k^2(\mu_2-\mu_1^2)\times\frac{1}{N}\sum_{k_1}\frac{\omega_{k_1}^2}{(\omega_{k_1}^2-y_k^2)_P}. \quad (4.13)$$

We solve iteratively, and to second order obtain

$$y_k=\omega_c+\omega_k\left\{(\tfrac{1}{2}\mu_1-\tfrac{1}{2}\mu_2+\tfrac{3}{8}\mu_1^2)-\tfrac{1}{2}(\mu_2-\mu_1^2)\frac{\omega_k^2}{N}\sum_{k_1}\frac{1}{(\omega_{k_1}^2-\omega_k^2)_P}\right\}. \quad (4.14)$$

A completely analogous result is obtained in higher dimensions. The shift in the unperturbed normal-mode frequencies is given by the second term in Eq. (4.14).

For our one-dimensional model, the mean shift in the kth normal-mode frequency is given to second order by

$$\Delta\omega_k=(\tfrac{1}{2}\mu_1-\tfrac{1}{2}\mu_2+\tfrac{3}{8}\mu_1^2)\omega_k-\tfrac{1}{2}(\mu_2-\mu_1^2)\frac{\omega_k^3}{\omega_L^2}\frac{1}{N}\sum_{k_1=1}^{N}\frac{1}{[\sin^2(\pi k_1/N)-\sin^2(\pi k/N)]_P}$$
$$=(\tfrac{1}{2}\mu_1-\tfrac{1}{2}\mu_2+\tfrac{3}{8}\mu_1^2)\omega_L\sin\frac{\pi k}{N}+\tfrac{1}{2}(\mu_2-\mu_1^2)\omega_L\begin{cases}\dfrac{1}{2N}\left(\dfrac{\sin(\pi k/N)}{\cos^2(\pi k/N)}-2\sin(\pi k/N)\right) & k\neq N/2,\\ \tfrac{1}{3}N & k=N/2\end{cases} \quad (4.15)$$

where the evaluation of the sum is discussed in Appendix B. This result for $\Delta\omega_k$ can be used to evaluate the change in the zero-point energy caused by disorder

$$\Delta E_0=\tfrac{1}{2}\hbar\sum_k\Delta\omega_k$$
$$=N\frac{\hbar\omega_L}{\pi}\left\{(\tfrac{1}{2}\mu_1+\tfrac{3}{8}\mu_1^2-\tfrac{1}{2}\mu_2)+\frac{\pi}{8}p(1-p)\frac{M^2(M_2-M_1)^2}{M_1^2M_2^2}\right\}. \quad (4.16)$$

This expression differs from the result one obtains by keeping the perturbation in the kinetic-energy terms. The alternative calculation is carried out in Appendix B. That the two results for ΔE_0 should be slightly different is not surprising because of the difference in the fundamental nature of the perturbation.

At least in lowest order, the expression given by Eq. (4.12) is not the long time asymptotic solution for $\langle Q_k(t)\rangle$. To obtain the asymptotic solution, we need some properties of the function $J_k(y)$. We establish these properties here only in the first approximation.

The value of the function $V_k(y)=\langle J_k(y)\rangle$ to lowest order is

$$V_k(y)=\pi\omega_k{}^2p(1-p)M^2\left(\frac{M_2-M_1}{M_1M_2}\right)^2\frac{1}{N}\sum_{k_1=1}^{N}\omega_L{}^2\sin^2\frac{\pi k_1}{N}\,\delta\left(\omega_L{}^2\sin^2\frac{\pi k_1}{N}-y^2\right)$$

$$=\pi\omega_k{}^2p(1-p)M^2\left(\frac{M_2-M_1}{M_1M_2}\right)^2\frac{\omega_L{}^2}{\pi}\int_0^{\pi}\sin^2\phi\,\delta(\omega_L{}^2\sin^2\phi-y^2)d\phi$$

$$=\omega_k{}^2p(1-p)M^2\left(\frac{M_2-M_1}{M_1M_2}\right)^2\frac{y}{(\omega_L{}^2-y^2)^{\frac{1}{2}}}\quad 0\leqslant|y|<\omega_L$$

$$=0\quad\text{otherwise.}\tag{4.17}$$

This result, together with the fact that

$$N_k{}^{-1}=|U_k'(y_k)|=|-2y_k-\langle K_k'(y_k)\rangle|\approx 2\omega_k\tag{4.18}$$

to lowest order in the perturbation, leads finally to the following expression for the damping constant Γ_k:

$$\Gamma_k=\tfrac{1}{2}p(1-p)M^2\left(\frac{M_2-M_1}{M_1M_2}\right)^2\frac{\omega_k{}^2}{(\omega_L{}^2-\omega_k{}^2)^{\frac{1}{2}}}\quad 0\leqslant\omega_k<\omega_L$$

$$=0\quad\text{otherwise.}\tag{4.19}$$

The constant factor $p(1-p)M^2(M_2-M_1/M_1M_2)^2$ will sometimes be denoted by ϵ^2 for brevity.

We must now examine the effect of the singularities of $\langle J_k(y)\rangle$ on the long-time behavior of $Q_k(t)$, since by a theorem due to Lighthill,[30] the two are directly related. We expand the integrand in Eq. (4.8b) about $y=\pm\omega_L$ and obtain

$$I_k{}^{(1)}(t)\sim\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{iyt}\left\{\left[U_k(\omega_L)+i\epsilon^2\omega_k{}^2\left(\frac{\omega_L}{2}\right)^{\frac{1}{2}}\frac{H(\omega_L-y)}{(\omega_L-y)^{\frac{1}{2}}}\right]^{-1}+\left[U_k(-\omega_L)-i\epsilon^2\omega_k{}^2\left(\frac{\omega_L}{2}\right)^{\frac{1}{2}}\frac{H(\omega_L+y)}{(\omega_L+y)^{\frac{1}{2}}}\right]^{-1}\right\}dy,\tag{4.20}$$

where $H(x)$ is the Heaviside unit function. With some manipulation we obtain

$$I_k{}^{(1)}(t)\sim-\frac{1}{\pi}\,\mathrm{Re}\left\{\frac{e^{i\omega_Lt}}{U_k(\omega_L)}\int_0^{\infty}e^{itx}dx+\frac{e^{i\omega_Lt}}{i\epsilon^2\omega_k{}^2(\omega_L/2)^{\frac{1}{2}}}\int_0^{\infty}x^{\frac{1}{2}}e^{-itx}dx\right\}=-\frac{1}{\pi U_k(\omega_L)}\frac{\sin\omega_Lt}{t}+O\left(\frac{1}{t^{\frac{3}{2}}}\right).\tag{4.21}$$

The correct long-time behavior of $\langle Q_k(t)\rangle$ to lowest order in the perturbation is thus given by

$$\langle Q_k{}^{(1)}(t)\rangle\sim-\frac{\omega_LQ_k(0)}{\pi U_k(\omega_L)}\frac{\cos\omega_Lt}{t}-\frac{\dot{Q}_k(0)}{\pi U_k(\omega_L)}\frac{\sin\omega_Lt}{t},\tag{4.22}$$

since for sufficiently large t the term of $O(t^{-1})$ will dominate the contribution from the terms of $O(e^{-\Gamma_k|t|})$. The result (4.22) in fact gives us a basis for discussing the time interval over which our solution for $\langle Q_k(t)\rangle$, Eq. (4.12), can be expected to hold. We write our result for $\langle Q_k(t)\rangle$ schematically as

$$Q_k(t)\sim A\exp(-\epsilon^2t)+B(\epsilon^2/\epsilon^2t)\quad t>0.\tag{4.23}$$

For times such that $\epsilon^2t\sim 1$, the second term is negligible compared with the first term since the ϵ^2 factor, which we have assumed to be small in this analysis, is uncompensated by any power of t to keep it appreciable as t increases. For times such that $\epsilon^2t\gg 1$, the second term overwhelms the first, regardless of the magnitude of ϵ^2. The approximation given by Eq. (4.12) must fail when the two terms become the same order of magnitude. A more accurate asymptotic expansion may assign a somewhat greater range of validity to Eq. (4.12); however, the existence of these correction terms should not be ignored since they help establish, independently of any other considerations, the time scale for which our one-to-one correspondence between the perturbed and unperturbed normal modes remains valid before they decay into more complicated states.

To conclude this section, we give a formal expression for the damping constant Γ_{kj} for a disordered three-dimensional cubic Bravais lattice to lowest order in the perturbation. Combining Eqs. (4.5), (4.10), (4.18),

[30] M. J. Lighthill, *An Introduction to Fourier Analysis and Generalized Functions* (Cambridge University Press, New York, 1958), p. 52.

and (C12), we have

$$\Gamma_{kj}=\frac{\pi}{2\omega(\mathbf{k};j)}\sum_{\mathbf{k}'j'}\left\langle\Phi\begin{pmatrix}\mathbf{k} & \mathbf{k}'\\ j & j'\end{pmatrix}\Phi\begin{pmatrix}\mathbf{k}' & \mathbf{k}\\ j' & j\end{pmatrix}\right\rangle\delta\left[\omega^2\begin{pmatrix}\mathbf{k}'\\ j'\end{pmatrix}-\omega^2\begin{pmatrix}\mathbf{k}\\ j\end{pmatrix}\right]$$

$$=\frac{\pi}{2\omega(\mathbf{k};j)}\omega^2\begin{pmatrix}\mathbf{k}\\ j\end{pmatrix}(\mu_2-\mu_1^2)\frac{1}{3N}\sum_{\mathbf{k}'j'}\omega^2\begin{pmatrix}\mathbf{k}'\\ j'\end{pmatrix}\delta\left[\omega^2\begin{pmatrix}\mathbf{k}'\\ j'\end{pmatrix}-\omega^2\begin{pmatrix}\mathbf{k}\\ j\end{pmatrix}\right]$$

$$=\frac{\pi}{2}\omega^3\begin{pmatrix}\mathbf{k}\\ j\end{pmatrix}(\mu_2-\mu_1^2)\frac{1}{3N}\sum_{\mathbf{k}'j'}\delta\left[\omega^2\begin{pmatrix}\mathbf{k}'\\ j'\end{pmatrix}-\omega^2\begin{pmatrix}\mathbf{k}\\ j\end{pmatrix}\right]. \tag{4.24}$$

We now make use of the result that

$$G(\omega^2)=\frac{1}{3N}\sum_{\mathbf{k}j}\delta\left[\omega^2\begin{pmatrix}\mathbf{k}\\ j\end{pmatrix}-\omega^2\right], \tag{4.25}$$

where $G(\omega^2)$ is the distribution function (normalized to unity) for the squares of the normal-mode frequencies, to rewrite Eq. (4.24) as

$$\Gamma_{kj}=\frac{\pi}{2}\omega^3\begin{pmatrix}\mathbf{k}\\ j\end{pmatrix}p(1-p)M^2\left(\frac{M_2-M_1}{M_1M_2}\right)^2G\left[\omega^2\begin{pmatrix}\mathbf{k}\\ j\end{pmatrix}\right]$$

$$=\frac{\pi}{4}p(1-p)M^2\left(\frac{M_2-M_1}{M_1M_2}\right)^2\omega^2\begin{pmatrix}\mathbf{k}\\ j\end{pmatrix}g\left[\omega\begin{pmatrix}\mathbf{k}\\ j\end{pmatrix}\right], \tag{4.26}$$

where $g(\omega)=2\omega G(\omega^2)$ is the distribution function for the normal-mode frequencies.

In the long wavelength limit, $g(\omega)=C\omega^2$, and Eq. (4.26) becomes

$$\Gamma_{kj}\sim\frac{\pi}{4}p(1-p)M^2\left(\frac{M_2-M_1}{M_1M_2}\right)^2C\omega^4\begin{pmatrix}\mathbf{k}\\ j\end{pmatrix}, \tag{4.27}$$

which is the result appropriate for the case of so-called Rayleigh scattering.

For three dimensional lattices, for which in lowest order $g(\omega)\propto(\omega_L-\omega)^{\frac{1}{2}}$ in the limit as $\omega\to\omega_L^-$, the analysis leading to Eq. (4.22) predicts a long time behavior for $\langle Q_k(t)\rangle$ which is proportional to $t^{-\frac{3}{2}}$.

An alternative method of obtaining the results of this section consists in studying the equation $s^2+\omega_k^2=-G_k(s)$ in the complex s plane. It can be shown that as it stands this equation has no solutions except for values of s on the imaginary axis. However, if we introduce the functions $G_k^+(s)$ and $G_k^-(s)$, which are the values of $G_k(s)$ for $\mathrm{Re}s>0$ and $\mathrm{Re}s<0$, respectively, then the analytic continuation of the function $s^2+\omega_k^2+G_k^+(s)$ through the branch cut along the imaginary axis onto the Riemann surface which is contiguous with the surface in the right-hand half-plane upon which $G_k^+(s)$ is defined will have zeroes in the left-hand half-plane. It should be stressed that the analytic continuation of $G_k^+(s)$ through the cut into the left-hand half-plane is not the same as $G_k^-(s)$ since they are defined on different sheets of the Riemann surface. The contribution (4.22) arises from swinging the contour around the branch points at $\pm i\omega_L$. A discussion of this method for the evaluation of integrals of the type (4.3) has been given by Migdal and Galitskii.[31]

V. OPTICAL ABSORPTION IN AN ISOTOPICALLY DISORDERED CRYSTAL

In this section we apply the results of the preceding sections to the calculation of the infrared lattice absorption of isotopically disordered crystals. In this calculation, we start with the exact expression for the dielectric susceptibility which has been derived by Kubo.[19] As in earlier sections of this paper, we restrict ourselves to the classical case.

The dielectric susceptibility tensor $\chi_{\mu\nu}(\omega)$ is expressed by Kubo in the classical limit in the following form

$$\chi_{\mu\nu}(\omega)=\lim_{\epsilon\to 0+}\beta\int_0^\infty e^{-i\omega t-\epsilon t}\langle\mathfrak{M}_\nu(0)\mathfrak{M}_\mu(t)\rangle dt, \tag{5.1}$$

Where $\mathfrak{M}$ is the dipole moment of the lattice, ω is the frequency of the incident electromagnetic radiation, and $\langle\ \rangle$ represents an average over a canonical distribution. The absorption coefficient is directly related to the imaginary part of χ.

Since our primary aim is to display the techniques for calculating the susceptibility, and since the three-dimensional case gives us nothing fundamentally different, we restrict ourselves here, as in previous sections, to the one-dimensional case.

For a chain of alternately charged ions, the dipole moment is given by

$$\mathfrak{M}=\sum_{l=1}^{N}(-1)^l eu_l, \tag{5.2}$$

where e is the magnitude of the charge on each ion, and N is now an even number. If we substitute for u_l its normal coordinate expansion, Eq. (2.5), we obtain

$$\mathfrak{M}=e(N/M)^{\frac{1}{2}}Q_{N/2}. \tag{5.3}$$

It is important to note that $Q_{N/2}$ is a real quantity, $Q_{N/2}=Q_{N/2}^*$.

The dielectric susceptibility in the present case

[31] V. M. Galitskii and A. B. Migdal, J. Exptl. Theoret. Phys. (U.S.S.R.) 34, 139 (1958) [English translation: Soviet Phys.—JETP 7, 96 (1958)].

becomes

$$\chi_{xx}(\omega)=\beta\frac{Ne^2}{M}\lim_{\epsilon\to 0+}\int_0^\infty e^{-i\omega t-\epsilon t}\langle\dot{Q}_{N/2}(0)Q_{N/2}(t)\rangle dt. \quad (5.4)$$

From the results of Sec. III, we know that we can express $Q_{N/2}(t)$ as

$$Q_{N/2}(t)=\dot{I}_{N/2}(t)Q_{N/2}(0)+I_{N/2}(t)\dot{Q}_{N/2}(0)+\sum_{n(\neq N/2)}\{\dot{I}_{N/2,n}(t)Q_n(0)+I_{N/2,n}(t)\dot{Q}_n(0)\}, \quad (5.5)$$

so that

$$\chi_{xx}(\omega)=\beta\langle|\dot{Q}_{N/2}(0)|^2\rangle\frac{Ne^2}{M}\lim_{\epsilon\to 0+}\int_0^\infty e^{-i\omega t-\epsilon t}I_{N/2}(t)dt. \quad (5.6)$$

We recognize the integral on the right-hand side of Eq. (5.6) as the Laplace transform of the function $I_{N/2}(t)$, with the customary parameter s replaced by $i\omega+\epsilon$. This is particularly convenient for our purposes, since it is not $I_{N/2}(t)$ that we know explicitly, but its Laplace transform, which is given by Eq. (3.9) as

$$D_{N/2}(s)=[s^2+\omega_{N/2}{}^2+G_{N/2}(s)]^{-1}. \quad (5.7)$$

If in Eq. (5.7) we replace s by $i\omega+\epsilon$ and pass to the limit as $\epsilon\to 0+$, we find for $\chi_{xx}(\omega)$ the result that

$$\chi_{xx}(\omega)=\beta\langle|\dot{Q}_{N/2}(0)|^2\rangle\times\frac{Ne^2}{M}\frac{1}{\omega_{N/2}{}^2-\omega^2-K_{N/2}(\omega)+iJ_{N/2}(\omega)}. \quad (5.8)$$

The imaginary part of $\chi_{xx}(\omega)$ is given by

$$\chi_{xx}{}^{(2)}(\omega)=-\beta\langle|\dot{Q}_{N/2}(0)|^2\rangle\times\frac{Ne^2}{M}\frac{J_{N/2}(\omega)}{[\omega_{N/2}{}^2-\omega^2-K_{N/2}(\omega)]^2+J_{N/2}{}^2(\omega)}. \quad (5.9)$$

If we evaluate the thermal average by using the Hamiltonian appropriate to the monatomic mean mass lattice, then to lowest order in the perturbation the imaginary part of the susceptibility normalized to unit volume becomes[32]

$$\chi_{xx}{}^{(2)}(\omega)=-\frac{e^2}{Ma_0{}^3}\frac{J_{N/2}(\omega)}{[\omega_{N/2}{}^2-\omega^2-K_{N/2}(\omega)]^2+J_{N/2}{}^2(\omega)}, \quad (5.10)$$

where a_0 is the nearest-neighbor separation between ions.

The complex dielectric constant is given by

$$\epsilon_{\mu\nu}=\delta_{\mu\nu}+4\pi\chi_{\mu\nu}(\omega), \quad (5.11)$$

which in the present case reduces to

$$\epsilon_{xx}=1+4\pi\chi_{xx}(\omega)=n^2, \quad (5.12)$$

[32] We associate a volume $a_0{}^3$ with each ion.

where

$$n=\lambda+i\mu \quad (5.13)$$

is the complex refractive index. λ is known as the ordinary index of refraction, while μ is the extinction coefficient. If χ_{xx} is so small that its square can be neglected compared with χ_{xx} itself, we find that

$$\mu=2\pi\chi_{xx}{}^{(2)}(\omega). \quad (5.14)$$

The linear absorption coefficient α_{xx} is now given by

$$\alpha_{xx}(\omega)=-2(\omega/c)\mu=-4\pi(\omega/c)\chi_{xx}{}^{(2)}(\omega) \quad (5.15)$$

$$=4\pi\frac{\omega}{c}\frac{e^2}{Ma_0{}^3}\times\frac{\gamma_{N/2}(\omega)}{[(\omega_{N/2}+\Delta\omega_{N/2})^2-\omega^2]^2+\gamma_{N/2}{}^2(\omega)}, \quad (5.16)$$

where

$$\Delta\omega_{N/2}(\omega)=[\omega_{N/2}{}^2-K_{N/2}(\omega)]^{\frac{1}{2}}-\omega_{N/2}, \quad (5.17a)$$

$$\gamma_{N/2}(\omega)=J_{N/2}(\omega), \quad (5.17b)$$

and c is the speed of light.

Strictly speaking, the result given by Eq. (5.16) is appropriate to a particular configuration of isotopic impurities at the lattice points. What we require, however, is the average of $\alpha_{xx}(\omega)$ over all configurations of the impurities:

$$\langle\alpha_{xx}(\omega)\rangle=4\pi\frac{\omega}{c}\frac{e^2}{Ma_0{}^3}\times\left\langle\frac{\gamma_{N/2}(\omega)}{[(\omega_{N/2}+\Delta\omega_{N/2}(\omega))^2-\omega^2]^2+\gamma_{N/2}{}^2(\omega)}\right\rangle. \quad (5.18)$$

In accord with the arguments put forth in Sec. III, we assert that in the limit as $N\to\infty$ we can replace the configuration average of the function inside the braces by the function of the configuration averages of the quantities $\Delta\omega_{N/2}$ and $\gamma_{N/2}$:

$$\langle\alpha_{xx}(\omega)\rangle=4\pi\frac{\omega}{c}\frac{e^2}{Ma_0{}^3}\times\frac{\langle\gamma_{N/2}(\omega)\rangle}{[(\omega_{N/2}+\langle\Delta\omega_{N/2}(\omega)\rangle)^2-\omega^2]^2+\langle\gamma_{N/2}(\omega)\rangle^2}. \quad (5.19)$$

An almost identical result holds for each dispersion oscillator in a three-dimensional crystal in the special case of plane-polarized incident radiation.

In the application of Eq. (5.19) to the calculation of the optical absorption in a mixed alkali-halide crystal, it must be kept in mind that in the most usual cases only the anion (cation) sublattice is disordered, while the cation (anion) sublattice remains ordered. This means that the matrix elements $\Phi_{kk'}$ given by Eqs.

(2.10) become

$$\Phi_{kk'}=\frac{\omega_{k'}^2}{N}\sum_{l\ \text{odd}}\left(\frac{M}{M_1}-1\right)\exp\left[\frac{2\pi i}{N}(k'-k)l\right]$$
$$+\frac{\omega_{k'}^2}{N}\sum_{l\ \text{even}}\left(\frac{M}{m_l}-1\right)\exp\left[\frac{2\pi i}{N}(k'-k)l\right], \quad (5.20)$$

in the case that the sublattice defined by the even-numbered ions becomes disordered, while the sublattice defined by the odd-numbered ions remains unperturbed. The mean mass M is defined in the present case by

$$M=\tfrac{1}{2}M_1+\tfrac{1}{2}[pM_2+(1-p)M_3], \quad (5.21)$$

where M_1 is the anion (cation) mass, while M_2 and M_3 are the masses of the two cation (anion) species. p and $(1-p)$ are the fractions of cations (anions) with masses M_2 and M_3 on the disordered sublattice.

The configuration average of the diagonal matrix elements is found to be

$$\langle\Phi_{kk}\rangle=\omega_k^2\left\{\frac{M}{2}\left[\frac{1}{M_1}+\frac{p}{M_2}+\frac{1-p}{M_3}\right]-1\right\}, \quad (5.22)$$

while the configuration average of the product of matrix elements $\langle\Phi_{(N/2)k_1}\Phi_{k_1(N/2)}\rangle$, which is needed in the evaluation of $\langle K_{N/2}(\omega)\rangle$ and $\langle J_{N/2}(\omega)\rangle$, is given by

$$\langle\Phi_{(N/2)k_1}\Phi_{k_1(N/2)}\rangle=\frac{\omega_{N/2}^2\omega_{k_1}^2}{2N}p(1-p)M^2$$
$$\times\left(\frac{M_3-M_2}{M_2M_3}\right)^2, \quad k_1\neq N/2. \quad (5.23)$$

With these results, we obtain finally that to second order

$$\langle K_{N/2}(\omega)\rangle=-\omega_{N/2}^2\left\{\frac{M}{2}\left[\frac{1}{M_1}+\frac{p}{M_2}+\frac{1-p}{M_3}\right]-1\right\}$$
$$+\tfrac{1}{2}\omega_{N/2}^2p(1-p)M^2\left(\frac{M_3-M_2}{M_2M_3}\right)^2$$
$$\times\frac{1}{N}\sum_{k_1}\frac{\omega_{k_1}^2}{(\omega_{k_1}^2-\omega^2)_P}, \quad (5.24)$$

$$\langle J_{N/2}(\omega)\rangle=\frac{\pi}{2}\omega_{N/2}^2p(1-p)M^2\left(\frac{M_3-M_2}{M_2M_3}\right)^2$$
$$\times\frac{1}{N}\sum_{k_1}\omega_{k_1}^2\delta(\omega_{k_1}^2-\omega^2).$$

It is, of course, clear that the problem of determining the time evolution of $\langle Q_{N/2}(t)\rangle$ could have been formulated in an alternative way in which the unperturbed lattice is taken to be a diatomic lattice with two atoms in each unit cell. One of the two unperturbed masses would be that of the cation or anion which is not affected in the preparation of the mixed crystal, while the second would be the mean of the masses of the two kinds of ions which could occupy the second position in the unit cell. The generalization of the present formalism to cover this choice for the unperturbed lattice is presented in Appendix A.

In conclusion, we note that the damping constant we have obtained in our expression for the absorption coefficient is frequency dependent. We can expect this frequency dependence to explain the departure of the absorption spectrum from a pure Lorentz shape and perhaps in some cases, even to give subsidiary maxima in the spectrum. We have obtained this frequency dependence by working with Laplace transforms without making any approximations other than using the thermal average appropriate to the unperturbed lattice. In particular, we did not have to find the actual time behavior of the system. It seems to be almost essential to work with Laplace transforms in calculating optical absorption with Kubo's formalism if it is desired to obtain the damping constant $\gamma_{N/2}$ as an explicit function of the frequency ω in a simple manner. This is because of frequency ω does not appear in the equations of motion for the normal coordinates and is introduced only in the Laplace integral (5.1). Thus, if the expression (4.12) for $Q_{N/2}(t)$ were substituted into Eq. (5.1) and the resulting integral evaluated, the expression for $\chi_{\mu\nu}$ would still be of a Lorentz form, but damping constant $\gamma_{N/2}$ would no longer be a function of ω. This is because we have had to make certain approximations, discussed in the preceding section, to obtain Eq. (4.12). To obtain the result expressed by Eq. (5.8), we would have to know the *exact* time dependence of $\langle\dot{Q}_{N/2}(0)Q_{N/2}(t)\rangle$ to this order. These observations remain valid for calculations of anharmonic optical absorption in ionic crystals, where frequency dependent damping constants are needed to explain the subsidiary maxima observed experimentally in infrared absorption studies.

VI. DISTRIBUTION FUNCTION FOR THE ENERGIES OF NORMAL MODES

In conclusion, we apply the methods employed in this paper to the problem studied by George.[22] He has examined the time evolution of the phase-space distribution function ρ which satisfies the Liouville equation

$$i(\partial\rho/\partial t)=(L_0+\delta L)\rho=L\rho, \quad (6.1)$$

where L_0 is the unperturbed Liouville operator, and δL is the change in L resulting from the random array of isotopic impurities. In order that L take its simplest form, it is expressed in terms of action-angle variables, $\{J_k\}$, $\{\alpha_k\}$. The distribution function is then expanded in terms of the eigenfunctions of the unperturbed

Liouville operator,

$$\rho=\sum_{\{n_k\}}\rho_{\{n_k\}}(\{J_k\},t)\exp[i\textstyle\sum_k n_k(\alpha_k-\omega_k t)], \quad (6.2)$$

and the "Fourier coefficients" $\rho_{\{n_k\}}(\{J_k\},t)$ are shown to satisfy an equation of the form

$$\frac{\partial}{\partial t}\rho_n(t)=\sum_{n'}\langle n|\delta L|n'\rangle \times\exp[it\textstyle\sum_k(n_k-n_k')\omega_k]\rho_{n'}(t), \quad (6.3)$$

where for convenience we have put $\rho_n=\rho_{\{n_k\}}$. The matrix elements in the present formulation are differential operators with respect to the action variables.[33] If we replace $\sum_k n_k\omega_k$ by Ω_n, and $\langle n|\delta L|n'\rangle$ by $-i\lambda V_{nn'}$, Eq. (6.3) reduces to

$$\frac{\partial}{\partial t}\rho_n(t)=-i\lambda\sum_{n'}V_{nn'}\exp[i(\Omega_n-\Omega_{n'})t]\rho_{n'}(t). \quad (6.4)$$

George assumes that at time $t=0$,

$$\rho_0(0)=O(1), \quad (6.5a)$$

and that the Fourier coefficients ρ_n which couple directly to ρ_0 satisfy the initial conditions

$$\rho_n(0)=O(\lambda) \quad n\neq 0. \quad (6.5b)$$

All other Fourier coefficients are at least of $O(\lambda^2)$ at $t=0$.

To solve Eq. (6.4), we put

$$\rho_n(t)=a_n(t)e^{i\Omega_n t}, \quad (6.6)$$

where

$$a_0(0)=O(1) \quad (6.7a)$$

and

$$a_n(0)=O(\lambda) \quad n\neq 0, \quad (6.7b)$$

for those coefficients which couple directly to a_0. Equation (6.4) becomes

$$\dot{a}_n(t)+i\Omega_n a_n(t)=-i\lambda\textstyle\sum_{n'}V_{nn'}a_{n'}(t). \quad (6.8)$$

We take the Laplace transform of this equation and find for $\alpha_n(s)=\mathfrak{L}\{a_n(t)\}$:

$$\alpha_n(s)=\frac{a_n(0)}{s+i\Omega_n}-\frac{i\lambda}{s+i\Omega_n}\sum_{n'}V_{nn'}\alpha_{n'}(s). \quad (6.9)$$

The solution to this equation can be cast in the form

$$\alpha_n(s)=D_n(s)a_n(0)+\sum_{n'(\neq n)}D_{nn'}(s)a_{n'}(0). \quad (6.10)$$

In view of the initial conditions (6.5) or (6.7), we see that, if we wish to obtain $a_0(t)$ to lowest order in λ only, we need to consider the diagonal transitions which are described by $D_0(s)$. With the aid of the general results of Sec. III, we can write down the expression for $D_0(s)$ immediately:

$$D_0(s)=[s+i\Omega_0-G_0(s)]^{-1}, \quad (6.11)$$

where

$$G_0(s)=(i\lambda)^2{\sum_{n_1}}'\frac{V_{0n_1}V_{n_10}}{s+i\Omega_{n_1}} -(i\lambda)^3{\sum_{n_1n_2}}'\frac{V_{0n_1}V_{n_1n_2}V_{n_20}}{(s+i\Omega_{n_1})(s+i\Omega_{n_2})}+\cdots \quad (6.12)$$

is the sum of the contributions with no intermediate state equal to the initial state. It is again clear that $G_0(s)$ satisfies the relation $G_0(iy+x)=-G_0^*(iy-x)$, so that

$$\lim_{\eta\to 0+}G_0(iy\pm\eta)=\mp K_0(y)+iJ_0(y), \quad (6.13)$$

where now $K_0(y)$ and $J_0(y)$ are differential operators with respect to the action variables. We thus obtain the result that

$$\{a_0(t)\}_{\text{diagonal}}=\left\{\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty}\frac{e^{st}}{s+i\Omega_0-G_0(s)}ds\right\}a_0(0). \quad (6.14)$$

Through the use of arguments similar to those employed in Sec. IV, we find that the long-time asymptotic form of $\rho_0(t)$ is given by (recalling that $\Omega_0\equiv 0$)

$$\rho_0(t)=a_0(t)=N_0\exp[iy_0t-N_0K_0(y_0)t]\rho_0(0), \quad (6.15)$$

where, to lowest order in λ,

$$J_0(y)=\lambda^2\sum_{n_1}\frac{V_{0n_1}V_{n_10}}{(y+\Omega_{n_1})_P}$$

$$K_0(y)=\lambda^2\sum_{n_1}V_{0n_1}V_{n_10}\delta(y+\Omega_{n_1}) \quad (6.16)$$

$$N_0^{-1}=[1-J_0'(y_0)],$$

and y_0 is the solution to the equation

$$y_0=J_0(y_0). \quad (6.17)$$

As long as we are interested only in the approach to equilibrium of our system, we may neglect the "energy shift" term in the exponential of Eq. (6.15) and write the equation for $\rho_0(t)$ as

$$(\partial/\partial t)\rho_0(t)=-N_0K_0(y_0)\rho_0(t)=\Omega(\{J_k\})\rho_0(t). \quad (6.18)$$

An explicit expression for the operator Ω is given by George:

$$\Omega=-\frac{\pi}{4}\sum_{kk'}\omega_k{}^2\Phi_{kk'}\Phi_{k'k}\delta(\omega_k-\omega_{k'}) \times\left(\frac{\partial}{\partial J_k}-\frac{\partial}{\partial J_{k'}}\right)J_kJ_{k'}\left(\frac{\partial}{\partial J_k}-\frac{\partial}{\partial J_{k'}}\right). \quad (6.19)$$

[33] Explicitly, for energy conserving processes we have that

$$\langle n|\delta L|n'\rangle=-\frac{i}{2}\sum_{kk'}(\omega_k\omega_{k'}J_kJ_{k'})^{\frac{1}{2}}\Phi_{kk'}\left(\frac{n_k'}{2J_k}+\frac{n_{k'}'}{2J_{k'}}+\frac{\partial}{\partial J_k}-\frac{\partial}{\partial J_{k'}}\right) \times\delta_{n_k,n_{k'}'+1}\delta_{n_{k'},n_{k'}'-1}\prod_{P\neq k,k'}\delta_{n_p,n_p'}.$$

From its definition,

$$\rho_0(\{J_k\};t)=\frac{1}{(2\pi)^N}\int_0^{2\pi}\cdots\int\rho(\{J_k\},\{\alpha_k\};t)\prod_k d\alpha_k,$$

we see that ρ_0 is the distribution function for the action variables of our system. Since the action variable J_k is proportional to the energy in the kth normal mode,

$$E_k=J_k\omega_k, \tag{6.20}$$

we can equally well regard ρ_0 as the distribution function of the normal-mode energies, $\rho(\{E_k\};t)$. If we assume a factored distribution for $\rho(\{E_k\};t)$,

$$\rho(\{E_k\};t)=\prod_n f_n(E_n,t), \tag{6.21}$$

substitute this form into both sides of Eq. (6.18) and integrate over all E_k save E_n, we obtain the equation satisfied by $f_n(E_n,t)$:

$$\frac{\partial}{\partial t}f_n=\frac{\pi}{2}c_n\omega_n^2\left[f_n+(E_n+kT)\frac{\partial f_n}{\partial E_n}+kTE_n\frac{\partial^2 f_n}{\partial E_n^2}\right], \tag{6.22}$$

where

$$c_n=\sum_k\Phi_{nk}\Phi_{kn}\delta(\omega_n-\omega_k). \tag{6.23}$$

In terms of the dimensionless variables

$$x=E_n/kT \quad \tau=c_nt=t/\tau_n, \tag{6.24}$$

it is readily shown that the solution to Eq. (6.22) which satisfies the initial condition

$$f(x,0)=\delta(x-x_0) \tag{6.25}$$

is given by[34]

$$f_n(x,\tau)=e^{-x}\sum_{m=0}^{\infty}\frac{e^{-m\tau}}{(m!)^2}L_m(x_0)L_m(x), \tag{6.26}$$

where $L_m(x)$ is the mth Laguerre polynomial.

The mean energy in the nth normal mode $\bar{x}$ is given by

$$\bar{x}=\int_0^{\infty}xf_n(x,\tau)dx=1-(1-x_0)e^{-\tau}, \tag{6.27}$$

or in dimensional variables,

$$\langle E_n(t)\rangle=kT-[kT-E_n(0)]e^{-t/\tau_n}. \tag{6.28}$$

This result shows that for any initial value of the mean energy in the nth normal mode $E_n(0)$ the interaction between normal modes induced by the isotope impurities will cause the energy to approach the equipartition value kT in the limit of long times.

The results of this section, together with the one-dimensional kinetic theory expression for thermal conductivity,

$$\kappa=\sum_p C_v(p)v_pl_p=\sum_p C_v(p)v_p^2l_p, \tag{6.29}$$

where $C_v(p)$ is the contribution to the specific heat from the pth normal mode and v_p and l_p are the group velocity and mean free path of the pth normal mode (see Appendix B), lead to a sum for κ which diverges at the long-wavelength limit. This divergence is associated with the fact that in the long-wavelength limit the disordered lattice is equivalent to a monatomic mean-mass lattice,[1] and such a lattice in the harmonic approximation cannot give rise to a finite thermal resistance. We remark in conclusion that the chief merit of the methods developed in this paper as applied to the calculations in this section is the ease with which it yields Eq. (6.18), as well as the equations for the other Fourier coefficients.

[34] I. Prigogine and R. Balescu, Physica **23**, 555 (1957).

ACKNOWLEDGMENTS

We would like to thank Dr. P. G. Klemens, Dr. J. S. Langer, Dr. R. Rubin, Dr. S. Teitler, Dr. R. F. Wallis, and Dr. R. Zwanzig, for helpful discussions concerning various points discussed in this paper.

APPENDIX A

We derive here the equation of motion of a normal coordinate in an isotopically disordered general three-dimensional lattice containing N unit cells with n atoms per unit cell.

The position vectors of the nuclei in such a lattice can be represented as[35]

$$\mathbf{x}(l;\kappa)=\mathbf{x}(l)+\mathbf{x}(\kappa). \tag{A1}$$

Here $\mathbf{x}(l)$ is the position vector of the lth unit cell and is given by

$$\mathbf{x}(l)=l_1\mathbf{a}_1+l_2\mathbf{a}_2+l_3\mathbf{a}_3, \tag{A2}$$

where $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$ are the three primitive translation vectors of the lattice and l_1, l_2, l_3 are integers, positive, negative, or zero. The vector $\mathbf{x}(\kappa)$ gives the position of the κth atom in a unit cell ($\kappa=1, 2, \cdots, n$).

If each atom undergoes a small vector displacement $\mathbf{u}(l;\kappa)$ from its equilibrium position, the equations of motion of the lattice can be written as

$$M_\kappa{}^l\ddot{u}_\alpha\binom{l}{\kappa}=-\sum_{l'\kappa'\beta}\Phi_{\alpha\beta}\binom{ll'}{\kappa\kappa'}u_\beta\binom{l'}{\kappa'}, \tag{A3}$$

where $M_\kappa{}^l$ is the mass at the κth position in the lth unit cell, $\alpha(=x,y,z)$ labels the Cartesian axes, and $\Phi_{\alpha\beta}(ll';\kappa\kappa')$ is a general force constant of the lattice which depends on l and l' only through the difference $l-l'$. We introduce the mean mass of the atom in the κth position in a unit cell by

$$M_\kappa=\frac{1}{N}\sum_l(M_\kappa{}^l), \tag{A4}$$

[35] M. Born and K. Huang, *Dynamical Theory of Crystal Lattices* (Clarendon Press, Oxford, England, 1954), Chap. V.

and rewrite Eq. (A3) as

$$\ddot{u}_\alpha\binom{l}{\kappa}=-\frac{1}{M_\kappa}\sum_{l'\kappa'\beta}\Phi_{\alpha\beta}\binom{ll'}{\kappa\kappa'}u_\beta\binom{l'}{\kappa'}+\left(\frac{1}{M_\kappa}-\frac{1}{M_\kappa{}^l}\right)\sum_{l'\kappa'\beta}\Phi_{\alpha\beta}\binom{ll'}{\kappa\kappa'}u_\beta\binom{l'}{\kappa'}. \quad \text{(A5)}$$

The normal coordinate transformation for the unperturbed mean mass lattice, described by the masses M_κ, is given by

$$u_\alpha\binom{l}{\kappa}=\frac{1}{(NM_\kappa)^{\frac{1}{2}}}\sum_{\mathbf{k},j}e_\alpha\left(\kappa\middle|\begin{matrix}\mathbf{k}\\ j\end{matrix}\right)Q\binom{\mathbf{k}}{j}\times\exp[2\pi i\mathbf{k}\cdot\mathbf{x}(l)]. \quad \text{(A6)}$$

In this equation, $\mathbf{k}$ is a propagation vector whose N allowed values are determined, e.g., by the cyclic boundary condition, and are uniformly distributed throughout a unit cell of the reciprocal lattice. The index j labels the n branches of the frequency spectrum, and $e_\alpha(\kappa|\mathbf{k};j)$ is the α component of an eigenvector of the dynamical matrix whose elements are given by

$$D_{\alpha\beta}\binom{\mathbf{k}}{\kappa\kappa'}=\frac{1}{(M_\kappa M_{\kappa'})^{\frac{1}{2}}}\sum_{\bar{l}}\Phi_{\alpha\beta}\binom{\bar{l}}{\kappa\kappa'}\times\exp[-2\pi i\mathbf{k}\cdot\mathbf{x}(\bar{l})], \quad \text{(A7)}$$

where $\bar{l}=l-l'$. The $\{e_\alpha(\kappa|\mathbf{k};j)\}$ satisfy the following equations:

$$\sum_{\kappa'\beta}D_{\alpha\beta}\binom{\mathbf{k}}{\kappa\kappa'}e_\beta\left(\kappa'\middle|\begin{matrix}\mathbf{k}\\ j\end{matrix}\right)=\omega^2\binom{\mathbf{k}}{j}e_\alpha\left(\kappa\middle|\begin{matrix}\mathbf{k}\\ j\end{matrix}\right) \quad \text{(A8a)}$$

$$\sum_{\kappa\alpha}e_\alpha{}^*\left(\kappa\middle|\begin{matrix}\mathbf{k}\\ j'\end{matrix}\right)e_\alpha\left(\kappa\middle|\begin{matrix}\mathbf{k}\\ j\end{matrix}\right)=\delta_{jj'} \quad \text{(A8b)}$$

$$\sum_{j}e_\beta{}^*\left(\kappa'\middle|\begin{matrix}\mathbf{k}\\ j\end{matrix}\right)e_\alpha\left(\kappa\middle|\begin{matrix}\mathbf{k}\\ j\end{matrix}\right)=\delta_{\kappa\kappa'}\delta_{\alpha\beta}. \quad \text{(A8c)}$$

The eigenvalue $\omega^2(\mathbf{k};j)$ appearing in Eq. (A8a) is the square of the jth normal mode frequency associated with the propagation vector $\mathbf{k}$. The normal coordinate $Q(\mathbf{k};j)$ satisfies the reality condition

$$Q(-\mathbf{k};j)=Q^*(\mathbf{k};j).$$

If $u_\alpha(l;\kappa)$ given by Eq. (A6) is substituted into Eq. (A5), and use is made of the orthogonality of the exponentials and the $\{e_\alpha(\kappa|\mathbf{k};j)\}$, we obtain

$$\ddot{Q}\binom{\mathbf{k}}{j}+\omega^2\binom{\mathbf{k}}{j}Q\binom{\mathbf{k}}{j}=-\sum_{\mathbf{k}'j'}\Phi\binom{\mathbf{k}\mathbf{k}'}{jj'}Q\binom{\mathbf{k}'}{j'}, \quad \text{(A9)}$$

where the matrix elements $\Phi(\mathbf{k}\mathbf{k}';jj')$ are given by

$$\Phi\binom{\mathbf{k}\mathbf{k}'}{jj'}=\frac{\omega^2\binom{\mathbf{k}'}{j'}}{N}\sum_{l\kappa\alpha}\left(\frac{M_\kappa}{M_\kappa{}^l}-1\right)e_\alpha{}^*\left(\kappa\middle|\begin{matrix}\mathbf{k}\\ j\end{matrix}\right)e_\alpha\left(\kappa\middle|\begin{matrix}\mathbf{k}'\\ j'\end{matrix}\right)\times\exp[2\pi i(\mathbf{k}'-\mathbf{k})\cdot\mathbf{x}(l)]. \quad \text{(A10)}$$

Since the indices $\mathbf{k}$ and j always appear paired in the three-dimensional formulation of the present problem, we see that Eq. (2.7) and the results obtained from it can be translated directly into the three-dimensional language by replacing the single summation index k everywhere by the pair $(\mathbf{k};j)$.

The results of this Appendix would be required, for example, in a discussion of the lattice vibration infrared optical absorption of Ge-Si alloys, which possess no first-order moment by a second-order moment. The unperturbed lattice in this case could be a monatomic mean mass lattice each of whose unit cells contains two atoms.

APPENDIX B

In Sec. II we obtained the equations of motion of an isotopically disordered lattice in a form in which the perturbation is transferred to the potential energy terms. This leads to a simple set of equations of motion for the normal coordinates. However, it is perhaps more natural to recast the equations of motion in a form in which the perturbation is retained in the kinetic energy terms. Although the equations of motion become somewhat more complicated in this formalism, for completeness we present their derivation and a discussion of their consequences in this appendix.

We begin by rewriting Eq. (2.1) as

$$M\ddot{u}_l-\gamma(u_{l+1}-2u_l+u_{l+1})=(M-m_l)\ddot{u}_l, \quad \text{(B1)}$$

where M is defined by Eq. (2.2). We now carry out the normal coordinate transformation defined by Eq. (2.5) to obtain

$$\ddot{Q}_k+\omega_k{}^2Q_k=\sum_{k'}\Phi_{kk'}\ddot{Q}_{k'}, \quad \text{(B2)}$$

where

$$\Phi_{kk'}=\frac{1}{N}\sum_l\left(\frac{M-m_l}{M}\right)\exp\left[\frac{2\pi i(k'-k)l}{N}\right]. \quad \text{(B3)}$$

The configuration average of the diagonal matrix elements vanishes in this formulation,

$$\langle\Phi_{kk}\rangle=\frac{1}{N}\sum_l\frac{M-m_l}{M}\equiv 0. \quad \text{(B4)}$$

We proceed to solve these equations with the aid of Laplace transforms as before. The Laplace transform of Eq. (B2) is

$$s^2q_k(s)-sQ_k(0)-\dot{Q}_k(0)+\omega_k{}^2q_k(s)=\sum_{k'}\Phi_{kk'}[s^2q_{k'}(s)-sQ_{k'}(0)-\dot{Q}_{k'}(0)],$$

or

$$q_k(s)=\frac{sQ_k(0)+\dot{Q}_k(0)}{s^2+\omega_k{}^2}-\sum_{k'}\Phi_{kk'}\frac{sQ_{k'}(0)+\dot{Q}_{k'}(0)}{s^2+\omega_{k'}{}^2}+\frac{s^2}{s^2+\omega_k{}^2}\sum_{k'}\Phi_{kk'}q_{k'}(s). \quad \text{(B5)}$$

We solve this set of equations by iteration. As in the text, only the diagonal contribution to $q_k(s)$ is non-vanishing on taking the configuration average. We find that

$$\{q_k(s)\}_d=\Big\{\frac{1}{d_k}-\Big(\frac{s^2}{d_k}-\frac{s^4}{d_k^2}\Big)\sum_{k_1}\frac{\Phi_{kk_1}\Phi_{k_1k}}{d_{k_1}} -s^2\Big(\frac{s^2}{d_k}-\frac{s^4}{d_k^2}\Big)\sum_{k_1k_2}\frac{\Phi_{kk_1}\Phi_{k_1k_2}\Phi_{k_2k}}{d_{k_1}d_{k_2}} -\cdots\Big\}c_k(s) =\mathfrak{D}_k(s)c_k(s), \tag{B6}$$

where $c_k(s)$ has been defined by Eqs. (3.3). We again regroup the terms in the expansion for $\mathfrak{D}_k(s)$ by first summing all terms in which no intermediate index equals k, in which one intermediate index equals k, $\cdots$, etc. To do this, it is convenient to rewrite $\mathfrak{D}_k(s)$ as

$$\mathfrak{D}_k(s)=\frac{1}{d_k}-\Big(\frac{1}{d_k}-\frac{s^2}{d_k^2}\Big)\Big[s^2\sum_{k_1}\frac{\Phi_{kk_1}\Phi_{k_1k}}{d_{k_1}} +s^4\sum_{k_1k_2}\frac{\Phi_{kk_1}\Phi_{k_1k_2}\Phi_{k_2k}}{d_{k_1}d_{k_2}}+\cdots\Big]. \tag{B7}$$

We now introduce a function $\mathcal{G}_k(s)$ which is defined by

$$\mathcal{G}_k(s)=s^2\sum_{k_1}{}'\frac{\Phi_{kk_1}\Phi_{k_1k}}{d_{k_1}}+s^4\sum_{k_1k_2}{}'\frac{\Phi_{kk_1}\Phi_{k_1k_2}\Phi_{k_2k}}{d_{k_1}d_{k_2}}+\cdots, \tag{B8}$$

where the prime on the summations means that no intermediate index equals k. In terms of $\mathcal{G}_k(s)$, the function $\mathfrak{D}_k(s)$ becomes

$$\mathfrak{D}_k(s)=\frac{1-\mathcal{G}_k(s)}{d_k-s^2\mathcal{G}_k(s)}. \tag{B9}$$

Thus, the configuration average of the normal coordinate $Q_k(t)$ is given by

$$\langle Q_k(t)\rangle=\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty}\frac{e^{st}}{s^2}\Big[1-\frac{\omega_k^2}{s^2+\omega_k^2-s^2\langle\mathcal{G}_k(s)\rangle}\Big] \times[sQ_k(0)+\dot{Q}_k(0)]ds, \tag{B10}$$

which yields the result that

$$\langle Q_k(t)\rangle=Q_k(0)+t\dot{Q}_k(0) -\frac{\omega_k^2}{2\pi i}\int_{c-i\infty}^{c+i\infty}e^{st}\frac{sQ_k(0)+\dot{Q}_k(0)}{s^2[s^2+\omega_k^2-s^2\langle\mathcal{G}_k(s)\rangle]}ds. \tag{B11}$$

From its defining equation, (B8), we see that

$$\mathcal{G}_k(s^*)=\mathcal{G}_k{}^*(s),\quad \mathcal{G}_k(-s)=\mathcal{G}_k(s), \tag{B12}$$

so that

$$\lim_{\eta\to 0+}\mathcal{G}_k(iy\pm\eta)=\mathcal{J}_k(y)\pm i\mathcal{K}_k(y), \tag{B13}$$

where $\mathcal{J}_k(y)$ and $\mathcal{K}_k(y)$ are real functions of y, and which, to the lowest order in the perturbation, are given by

$$\mathcal{J}_k(y)=\sum_{k_1}{}'\Phi_{kk_1}\Phi_{k_1k}\Big\{1-\frac{\omega_{k_1}^2}{(\omega_{k_1}^2-y^2)_P}\Big\}, \tag{B14a}$$

$$\mathcal{K}_k(y)=\pi\,\mathrm{sgn}y\sum_{k_1}{}'\omega_{k_1}^2\Phi_{kk_1}\Phi_{k_1k}\delta(\omega_{k_1}^2-y^2). \tag{B14b}$$

With the aid of the convolution theorem for Laplace transforms, we can write

$$\langle Q_k(t)\rangle=Q_k(0)+t\dot{Q}_k(0)-\omega_k^2Q_k(0)\int_0^t(t-t_1)F'(t_1)dt_1 -\omega_k^2\dot{Q}_k(0)\int_0^t(t-t_1)F(t_1)dt_1, \tag{B15}$$

where

$$F(t)=\mathcal{L}^{-1}\{s^2+\omega_k^2-s^2\langle\mathcal{G}_k(s)\rangle\}^{-1}. \tag{B16}$$

The function $F(t)$ can be obtained in the large t limit by writing it as

$$F(t)=\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{iyt}\frac{dy}{-y^2+\omega_k^2+y^2[\langle\mathcal{J}_k(y)\rangle+i\langle\mathcal{K}_k(y)\rangle]}. \tag{B17}$$

By expanding the denominator of the integrand of the integral in Eq. (B17) about the values of y for which the real part vanishes, we have

$$F(t)\sim\frac{1}{4\pi}\sum_{\nu_k}\int_{-\infty}^{\infty}e^{iyt}\frac{N_k}{y-y_k+i\Gamma_k}dy, \tag{B18}$$

where

$$N_k{}^{-1}=-y_k[1-\langle\mathcal{J}_k(y_k)\rangle]+\tfrac{1}{2}y_k^2\langle\mathcal{J}_k{}'(y_k)\rangle<0 \quad \text{for } y_k>0, \tag{B19a}$$

$$\Gamma_k=\tfrac{1}{2}N_ky_k^2\langle\mathcal{K}_k(y_k)\rangle. \tag{B19b}$$

To lowest order in the perturbation, y_k is given by the solution to

$$y_k^2=\omega_k^2+y_k^2\sum_{k_1}{}'\langle\Phi_{kk_1}\Phi_{k_1k}\rangle\Big\{1-\frac{\omega_{k_1}^2}{(\omega_{k_1}^2-y_k^2)_P}\Big\},$$

i.e.,

$$y_k^2=\omega_k^2+\omega_k^2\sum_{k_1}{}'\langle\Phi_{kk_1}\Phi_{k_1k}\rangle\Big\{1-\frac{\omega_{k_1}^2}{(\omega_{k_1}^2-\omega_k^2)_P}\Big\}, \tag{B20}$$

or

$$y_k=\pm\omega_k\Big\{1+\tfrac{1}{2}\sum_{k_1}{}'\langle\Phi_{kk_1}\Phi_{k_1k}\rangle\Big[1-\frac{\omega_{k_1}^2}{(\omega_{k_1}^2-\omega_k^2)_P}\Big]\Big\}. \tag{B21}$$

The evaluation of $F(t)$ proceeds as follows:

$$F(t)\sim-\frac{|N_k|}{4\pi}\int_{-\infty}^{\infty}e^{iyt} \times\Big\{\frac{1}{y-|y_k|-i|\Gamma_k|}-\frac{1}{y+|y_k|-i|\Gamma_k|}\Big\}dy =|N_k|e^{-|\Gamma_k|t}\sin|y_k|t. \tag{B22}$$

In what follows, we omit the modulus symbols and understand that it is the magnitudes of the various quantities which enter into our expressions. After an integration by parts we have

$$\langle Q_k(t)\rangle = Q_k(0)\left\{1-\omega_k^2\int_0^t F(t_1)dt_1\right\}+\dot{Q}_k(0)\left\{t-\omega_k^2\int_0^t F(t_1)(t-t_1)dt_1\right\}$$

$$=Q_k(0)\left\{1+\frac{N_k\omega_k^2}{\Gamma_k^2+y_k^2}[y_k(e^{-\Gamma_k t}\cos y_k t-1)+\Gamma_k e^{-\Gamma_k t}\sin y_k t]\right\}+\dot{Q}_k(0)\left\{t+t\frac{N_k\omega_k^2}{y_k^2+\Gamma_k^2}[y_k(e^{-\Gamma_k t}\cos y_k t-1)\right.$$

$$+\Gamma_k e^{-\Gamma_k t}\sin y_k t]+\frac{2N_k\omega_k^2\Gamma_k}{y_k^2+\Gamma_k^2}[y_k(1-e^{-\Gamma_k t}\cos y_k t)-\Gamma_k e^{-\Gamma_k t}\sin y_k t]$$

$$\left.-\frac{N_k\omega_k^2}{y_k^2+\Gamma_k^2}[y_k t e^{-\Gamma_k t}\cos y_k t-(1-\Gamma_k t)e^{-\Gamma_k t}\sin y_k t]\right\},$$

$$\sim Q_k(0)e^{-\Gamma_k t}\cos y_k t+\dot{Q}_k(0)e^{-\Gamma_k t}[\sin(y_k t)/y_k] \quad \Gamma_k t\ll 1, \tag{B23}$$

a result which is much more complicated than that obtained in the text, Eq. (4.12).

For a one-dimensional lattice, we can obtain y_k to lowest order in the perturbation. Since it is readily shown that

$$\langle\Phi_{kk_1}\Phi_{k_1k}\rangle=\frac{1}{N}p(1-p)\frac{(M_1-M_2)^2}{M^2}=\frac{1}{N}p(1-p)\mu^2, \tag{B24}$$

we have from Eqs. (2.8) and (B21) that

$$y_k=\omega_k\left\{1-\frac{1}{2N}p(1-p)\mu^2\right.$$

$$\left.\times\sum_{k_1=1}^{N}{}'\frac{\sin^2(\pi k/N)}{[\sin^2(\pi k_1/N)-\sin^2(\pi k/N)]_P}\right\}. \tag{B25}$$

It turns out that, at least in the one-dimensional case, it is not accurate enough to replace the sum in Eq. (B25) by an integral in the limit as $N\to\infty$, so that we must evaluate the sum exactly. This is a reasonably straightforward, though tedious calculation, which yields the result that[11]

$$\sum_{k_1=1}^{N}\left[\left(\sin^2\frac{\pi k_1}{N}-\sin^2\frac{\pi k}{N}\right)_P\right]^{-1}$$

$$=2\sum_{k_1=1}^{N}{}''\left(\cos\frac{2\pi k}{N}-\cos\frac{2\pi k_1}{N}\right)^{-1}$$

$$=\frac{2\cos(2\pi k/N)}{\sin^2(2\pi k/N)} \qquad k\neq N/2$$

$$=-\tfrac{1}{3}N^2 \qquad k=N/2, \tag{B26}$$

where the double prime on the second sum in Eq. (B26) excludes the values $k_1=k$ and $k_1=N-k$. Thus, we find that

$$y_k=\omega_k+\Delta\omega_k$$

$$=\omega_L\sin\frac{\pi k}{N}-\frac{\omega_L}{4N}p(1-p)\mu^2$$

$$\times\sin\frac{\pi k}{N}\frac{1-2\sin^2(\pi k/N)}{\cos^2(\pi k/N)} \qquad k\neq N/2$$

$$=\omega_L+(\omega_L/6)p(1-p)N\mu^2 \qquad k=N/2 \tag{B27}$$

to second order in the perturbation. The correctness of this result can be verified by computing the zero-point energy of the disordered lattice. This is given by

$$E_0=E_0^{(0)}+\Delta E_0=\tfrac{1}{2}\hbar\sum_k(\omega_k+\Delta\omega_k)$$

$$=\tfrac{1}{2}\hbar\omega_L\sum_{k=1}^{N}\sin\frac{\pi k}{N}+\frac{\hbar\omega_L}{8N}p(1-p)\mu^2\sum_{k=1}^{N}{}'\sin\frac{\pi k}{N}$$

$$\times\frac{1-2\cos^2(\pi k/N)}{\cos^2(\pi k/N)}+\tfrac{1}{12}\hbar\omega_L p(1-p)N\mu^2, \tag{B28}$$

where the prime on the second sum excludes the value $k=N/2$.

The first sum, which gives the unperturbed zero-point energy, is readily found to be $2N/\pi$ in the limit of large N, so that $E_0^{(0)}=N\hbar\omega_L/\pi$.

The change in zero-point energy ΔE_0 can be expressed compactly as

$$\Delta E_0=E_0^{(0)}\frac{\pi}{8}p(1-p)\mu^2\left[\frac{2}{3}+\frac{1}{N^2}\sum_{k=1}^{N}{}'\frac{\sin(\pi k/N)}{\cos^2(\pi k/N)}\right], \tag{B29}$$

where the prime on the sum excludes the value $k=N/2$, and where we have neglected contributions of $O(1/N)$ relative to the terms retained. The sum is evaluated by

contour integration with the result that

$$\frac{1}{N^2}\sum_{k=1}^{N}{}'\frac{\sin(\pi k/N)}{\cos^2(\pi k/N)}=\frac{1}{3}. \tag{B30}$$

Our final result, then, is that

$$\Delta E_0=E_0^{(0)}(\pi/8)p(1-p)\mu^2, \tag{B31}$$

in agreement with the results of earlier calculations.[36] Thus, in spite of the unphysical divergence in the dispersion relation (B27) in the short-wave limit, which had been discussed previously by Rubin and Fleishman[37] for a chain with fixed ends, we recover the correct expression for the change in zero-point energy of a disordered linear chain. This result confirms the correctness of our basic theory.

To conclude, and for completeness, we calculate the mean lifetime and mean free path of a normal mode for the present model. The former is given by

$$\left\langle\frac{1}{\tau_k}\right\rangle=|\Gamma_k|=\frac{\langle y_k^2\mathcal{K}_k(y_k)\rangle}{2y_k[1-\langle\mathcal{J}_k(y_k)\rangle]-y_k^2\langle\mathcal{J}_k'(y_k)\rangle}$$

$$y_k>0$$

$$\approx\tfrac{1}{2}y_k\langle\mathcal{K}_k(y_k)\rangle\approx\tfrac{1}{2}\omega_k\langle\mathcal{K}_k(\omega_k)\rangle, \tag{B32}$$

to lowest order in the perturbation. With the aid of Eqs. (2.8) and (B24), we have that

$$\left\langle\frac{1}{\tau_k}\right\rangle=\frac{\pi}{2}\omega_k\frac{1}{N}p(1-p)\mu^2\sum_{k_1}{}'\omega_{k_1}^2\delta(\omega_{k_1}^2-\omega_k^2)$$

$$\rightarrow\frac{\omega_k}{2}p(1-p)\mu^2\int_0^{\pi}\omega_L^2$$

$$\times\sin^2\phi\,\delta(\omega_L^2\sin^2\phi-\omega_L^2\sin^2\phi_k)d\phi$$

$$=\tfrac{1}{2}p(1-p)\mu^2[\omega_k^2/(\omega_L^2-\omega_k^2)^{\frac{1}{2}}]\quad 0\leqslant\omega_k<\omega_L$$

$$=0\quad\text{otherwise}$$

$$=(\pi/4)p(1-p)\mu^2\omega_k^2 g(\omega_k), \tag{B33}$$

where $g(\omega)$ is the frequency distribution function for the unperturbed one-dimensional chain. This result, Eq. (B33), is in agreement with the results of George.[38]

The mean free path of a normal mode is defined as the group velocity of the normal mode multiplied by the mean lifetime of the mode:

$$l_k=\frac{a_0}{2}\omega_L\left|\cos\frac{\pi k}{N}\right|[(\pi/4)p(1-p)\mu^2\omega_k^2 g(\omega_k)]^{-1}$$

$$=\frac{a_0}{\omega_k^2}\frac{\omega_L^2-\omega_k^2}{p(1-p)\mu^2}\quad 0\leqslant\omega_k<\omega_L$$

$$=0\quad\text{otherwise},$$

where a_0 is the nearest neighbor separation. A similar result has also been obtained by Travernier.[39]

APPENDIX C. EVALUATION OF THE AVERAGE MATRIX ELEMENTS

In this Appendix we evaluate average matrix element expressions of the form

$$\langle\Phi_{kk_1}\rangle,\quad\langle\Phi_{kk_1}\Phi_{k_1k_2}\rangle,\quad\langle\Phi_{kk_1}\Phi_{k_1k_2}\Phi_{k_2k_3}\rangle,$$

These are required for the computation of $\langle G_k(s)\rangle$. In performing the averaging, we assume that each lattice site is occupied by a particle of mass M_1 with probability p and by a particle of mass M_2 with probability $q(=1-p)$. This same set of probabilities is associated with each lattice site, and all sites are independent.

Each perturbation matrix element in the expressions we wish to average contains a sum over lattice sites. As a result, each expression contains a multiple sum over lattice sites. We separate the terms in this multiple sum in which two or more summation indices refer to the same lattice sites from the terms in which they refer to different sites. In the second case, the separate sites are occupied with independent probabilities, while in the first case only one probability is involved. As long as we know that different summation indices always refer to different lattice sites, we may take independent averages at the sites associated with different indices.

Consequently, we split the multiple sum into a sum with all indices unequal to each other, plus sums with two indices equal in the various possible combinations, plus sums with three equal, plus sums with two pairs set equal separately, and so forth, down to a single sum with all indices equal. In each sum, indices which are not explicitly set equal are restricted to be different from each other. We may now average each sum over the quantities associated with each index separately.

Taking these averages leads to a great deal of simplification because, on the average, each lattice point is like any other and most of the position dependence disappears. By adding and subtracting terms, we can express the lattice sums in terms of various combinations of Kronecker δ's.

The general procedure is made quite clear by examples. We restrict ourselves here to the linear chain, although the method is applicable to lattices

[36] I. Prigogine, R. Bingen, and J. Jeener, Physica **20**, 383, 516 (1954).

[37] R. J. Rubin and B. A. Fleishman, Phys. Rev. **99**, 656 (1955); private communication.

[38] Strictly speaking, the result obtained for $\langle 1/\tau_k\rangle$ by George is twice that given by Eq. (B33). This is due to the fact that he considers the decay of the energy in the kth normal mode, and this is related to $|Q_k|^2$ and $|\dot{Q}|^2$ (see Sec. VI).

[39] J. Tavernier, Compt. rend. **245**, 1705 (1957).

of higher dimensions. We begin by obtaining the configuration average of the matrix element Φ_{kk_1} which is given by

$$\Phi_{kk_1}=\frac{\omega_{k_1}^2}{N}\sum_{l=1}^{N}\left(\frac{M}{m_l}-1\right)\exp\left[-\frac{2\pi i(k-k_1)l}{N}\right]. \quad \text{(C1)}$$

Since the operations of summing over and averaging over the configuration of the masses at the lattice points commute, we obtain

$$\langle\Phi_{kk_1}\rangle=\frac{\omega_{k_1}^2}{N}\left[p\left(\frac{M}{M_1}-1\right)+(1-p)\left(\frac{M}{M_2}-1\right)\right]\times\sum_{l=1}^{N}\exp\left[-\frac{2\pi i(k-k_1)l}{N}\right]$$
$$=\omega_{k_1}^2\mu_1\Delta(k-k_1), \quad \text{(C2)}$$

where we have put

$$\mu_n=\left\langle\left(\frac{M}{m_l}-1\right)^n\right\rangle=p\left(\frac{M}{M_1}-1\right)^n+(1-p)\left(\frac{M}{M_2}-1\right)^n, \quad \text{(C3)}$$

and

$$\Delta(k-k_1)=\frac{1}{N}\sum_{l=1}^{N}\exp\left[-\frac{2\pi i(k-k_1)}{N}l\right]=\begin{cases}1 & \text{if } k=k_1 \pmod N\\ 0 & \text{otherwise.}\end{cases} \quad \text{(C4)}$$

For the average of the product of two matrix elements; we have

$$\langle\Phi_{kk_1}\Phi_{k_1k_2}\rangle=\frac{\omega_{k_1}^2\omega_{k_2}^2}{N^2}\sum_{j,l}\left\langle\left(\frac{M}{m_j}-1\right)\left(\frac{M}{m_l}-1\right)\right\rangle\times\exp\left\{-\frac{2\pi i}{N}[(k-k_1)j+(k_1-k_2)l]\right\}. \quad \text{(C5)}$$

Here we split the sum into two sums with $j=l$ and $j\neq l$, respectively:

$$\langle\Phi_{kk_1}\Phi_{k_1k_2}\rangle=\frac{\omega_{k_1}^2\omega_{k_2}^2}{N^2}\left\{\sum_{j=1}^{N}\left\langle\left(\frac{M}{m_j}-1\right)^2\right\rangle\exp\left[-\frac{2\pi i}{N}(k-k_2)j\right]+\sum_{\substack{jl\\(j\neq l)}}\left(\frac{M}{m_j}-1\right)\left(\frac{M}{m_l}-1\right)\times\exp\left(-\frac{2\pi i}{N}[(k-k_1)j+(k_1-k_2)l]\right)\right\}. \quad \text{(C6)}$$

Each term in each sum can be averaged separately, and in the second sum we find that the average at point j is not affected by the value of l. This would not be true if l were allowed to equal j. Averaging gives

$$\langle\Phi_{kk_1}\Phi_{k_1k_2}\rangle=\frac{\omega_{k_1}^2\omega_{k_2}^2}{N^2}\left\{\mu_2\sum_{j=1}^{N}\exp\left[-\frac{2\pi i}{N}(k-k_2)j\right]+\mu_1^2\sum_{\substack{jl\\(j\neq l)}}\exp\left(-\frac{2\pi i}{N}[(k-k_1)j+(k_1-k_2)l]\right)\right\}. \quad \text{(C7)}$$

We introduce the $j=l$ terms into the second sum and subtract them off again from the full expression to obtain finally

$$\langle\Phi_{kk_1}\Phi_{k_1k_2}\rangle=\omega_{k_1}^2\omega_{k_2}^2\left\{\frac{1}{N}(\mu_2-\mu_1^2)+\mu_1^2\Delta(k-k_1)\right\}\times\Delta(k-k_2). \quad \text{(C8)}$$

The averages of the products of more Φ's become complicated, though the method is the same. For the product of three Φ's, we obtain

$$\langle\Phi_{kk_1}\Phi_{k_1k_2}\Phi_{k_2k_3}\rangle=\omega_{k_1}^2\omega_{k_2}^2\omega_{k_3}^2\left\{\frac{1}{N^2}\mu_3+\mu_2\mu_1\left[\frac{1}{N}\Delta(k_1-k_2)-\frac{3}{N^2}+\frac{1}{N}\Delta(k_2-k_3)+\frac{1}{N}\Delta(k-k_1)\right]+\mu_1^3\left[\Delta(k-k_1)\Delta(k-k_2)-\frac{1}{N}\Delta(k-k_1)-\frac{1}{N}\Delta(k_2-k_3)+\frac{2}{N^2}\right]\right\}\Delta(k-k_3). \quad \text{(C9)}$$

Most of the terms in these expressions do not make a contribution to $\langle G_k(s)\rangle$. We note that every average will be zero unless the initial and final wave vectors on the two ends of the expression are equal to each other. On the average, there are no off-diagonal contributions. This is a perfectly general result, valid in all orders of the perturbation. It follows from the translational invariance of the mean-mass lattice after the ensemble average over all configurations has been taken. The result

$$\langle\Phi_{kk_1}\Phi_{k_1k_2}\rangle=(1/N)\omega_{k_1}^2\omega_{k_2}^2(\mu_2-\mu_1^2)\Delta(k-k_2)$$

is our analog of Van Hove's "diagonal singularity condition."[40]

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

It must be mentioned that had we defined the mean mass M by the relation

$$\frac{1}{M}=\frac{p}{M_1}+\frac{1-p}{M_2}$$

rather than by Eq. (2.2), the results of this Appendix would have simplified greatly in that all terms involving μ_1 would vanish identically. The motivation for the choice of the definition (2.2) for M is discussed in the Introduction.

We conclude by calculating $\langle\Phi_{kk_1}\Phi_{k_1k}\rangle$ explicitly for a three-dimensional cubic Bravais lattice. From Eq. (A10), we have that

$$\left\langle\Phi\binom{\mathbf{k}\mathbf{k}'}{jj'}\Phi\binom{\mathbf{k}'\mathbf{k}}{j'j}\right\rangle=\frac{\omega^2(\mathbf{k};j)\omega^2(\mathbf{k}';j')}{N^2}\sum_{\substack{l\alpha\\ l'\beta}}\left\langle\left(\frac{M}{M_l}-1\right)\left(\frac{M}{M_{l'}}-1\right)\right\rangle$$

$$\times e_\alpha\binom{\mathbf{k}}{j}e_\alpha\binom{\mathbf{k}'}{j'}e_\beta\binom{\mathbf{k}}{j}e_\beta\binom{\mathbf{k}'}{j}\exp\{2\pi i(\mathbf{k}'-\mathbf{k})\cdot[\mathbf{x}(l)-\mathbf{x}(l')]\} \quad \text{(C10)}$$

$$=\frac{\omega^2(\mathbf{k};j)\omega^2(\mathbf{k}';j')}{N^2}\{N\mu_2+N^2\mu_1^2\Delta(\mathbf{k}'-\mathbf{k})\Delta(\mathbf{k}-\mathbf{k}')-N\mu_1^2\}$$

$$\times\sum_\alpha e_\alpha\binom{\mathbf{k}}{j}e_\alpha\binom{\mathbf{k}'}{j'}\sum_\beta e_\beta\binom{\mathbf{k}'}{j}e_\beta\binom{\mathbf{k}}{j}$$

$$=\frac{\omega^2(\mathbf{k};j)\omega^2(\mathbf{k}';j')}{N}(\mu_2-\mu_1^2)\left[\mathbf{e}\binom{\mathbf{k}}{j}\cdot\mathbf{e}\binom{\mathbf{k}'}{j'}\right]^2+\omega^4(\mathbf{k};j)\mu_1^2\Delta(\mathbf{k}'-\mathbf{k})\delta_{jj'}. \quad \text{(C11)}$$

Now, $[\mathbf{e}(\mathbf{k};j)\cdot\mathbf{e}(\mathbf{k}';j')]^2$ is a quadratic form in the components of $\mathbf{e}(\mathbf{k};j)$ which has cubic symmetry. A quadratic form which has cubic symmetry must be isotropic, so that we can replace $[\mathbf{e}(\mathbf{k};j)\cdot\mathbf{e}(\mathbf{k}';j')]^2$ in a sum over $\mathbf{k}'$ and j' by $\frac{1}{3}$. Thus we have finally

$$\left\langle\Phi\binom{\mathbf{k}\mathbf{k}'}{jj'}\Phi\binom{\mathbf{k}'\mathbf{k}}{j'j}\right\rangle=\frac{\omega^2(\mathbf{k};j)\omega^2(\mathbf{k}';j')}{3N}(\mu_2-\mu_1^2)$$

$$+\omega^4(\mathbf{k};j)\mu_1^2\Delta(\mathbf{k}'-\mathbf{k})\delta_{jj'}. \quad \text{(C12)}$$

APPENDIX D. USE OF THE RESOLVENT METHOD ON THE DIATOMIC LINEAR CHAIN

To clarify the effect of our assumption of a random distribution of isotopic defects, in this Appendix we apply the formalism of the preceding sections to the case of an ordered diatomic lattice. For simplicity we restrict ourselves to the one-dimensional case. In this case, the unperturbed normal-mode frequencies are given by

$$\omega_k^2=\omega_L^2\sin^2(\pi k/N)\quad(k=1,2,\cdots,N). \quad \text{(D1)}$$

The parameter ω_L is the maximum normal-mode frequency of the monatomic mean-mass lattice,

$$\omega_L^2=4\gamma/M, \quad \text{(D2)}$$

where γ is the force constant for nearest neighbor interactions.

We assume that all the even-numbered lattice points are occupied by masses M_1, while the odd-numbered lattice points are occupied by masses M_2. The expression for the matrix elements $\Phi_{kk'}$ reduces to

$$\Phi_{kk'}=\frac{\omega_{k'}}{N}\left\{\left(1-\frac{M}{M_1}\right)\sum_{l\,\text{even}}\exp\left[\frac{2\pi i(k'-k)l}{N}\right]\right.$$

$$\left.+\left(1-\frac{M}{M_2}\right)\sum_{l\,\text{odd}}\exp\left[\frac{2\pi i(k'-k)l}{N}\right]\right\}. \quad \text{(D3)}$$

We put $N=2K$ and sum the resulting expressions to obtain

$$\Phi_{kk'}=\frac{\omega_{k'}}{2}\left\{\frac{M_1-M_2}{2M_1}\Delta(k'-k)\right.$$

$$\left.+\frac{M_2-M_1}{2M_2}\exp\left[\frac{\pi i(k'-k)}{K}\right]\Delta(k'-k)\right\}, \quad \text{(D4)}$$

where $\Delta(k'-k)$ has been defined in Appendix C, Eq. (C4). In the present context, the N appearing in Eq. (C4) is to be replaced by $2K$.

The diagonal elements

$$\Phi_{kk}=-\frac{\omega_k}{4}\frac{(M_1-M_2)^2}{M_1M_2} \quad \text{(D5)}$$

can be absorbed into the left-hand side of Eq. (2.7) by defining a new frequency Ω_k^2 by

$$\Omega_k^2=\omega_k^2\frac{(M_1+M_2)^2}{4M_1M_2}. \quad \text{(D6)}$$

The off-diagonal elements can be expressed as

$$\Phi_{kk'}\equiv V_{kk'}=\eta\Omega_{k'}{}^2(\delta_{k,k'+K}+\delta_{k,k'-K}),\quad \eta=\frac{M_1-M_2}{M_1+M_2}. \tag{D7}$$

Since both k and k' are restricted to the interval $(1,2K)$, this expression gives all the nonvanishing off-diagonal elements. With these simple matrix elements, the expansion for the function $G_k(s)$ is found to be

$$G_k(s)=-\sum_{k_1}{}'\frac{V_{kk_1}V_{k_1k}}{s^2+\Omega_{k_1}{}^2}+\sum_{k_1k_2}{}'\frac{V_{kk_1}V_{k_1k_2}V_{k_2k}}{(s^2+\Omega_{k_1}{}^2)(s^2+\Omega_{k_2}{}^2)}-\cdots \tag{D8}$$

and can be seen to terminate after the first term. For example, in the second term, if the index k lies in the interval $0<k\leqslant K$, it is linked by a nonzero matrix element only to the index $k_1=k+K$. The index k_1 is also linked to only one index for which $0<k\leqslant 2K$, namely, $k_2=k$. However, the second summation index k_2 cannot be equal to k because this is explicitly excluded in the definition of $G_k(s)$. Hence, this second summation is vacuous and this term in the expression for $G_k(s)$ is zero. A similar argument holds for a k in the interval $K\leqslant k\leqslant 2K$, and the higher terms in the series are zero in the same way.

In the following, we shall assume that $0<k\leqslant K$ and write an index in the upper half of the $2K$ range as $k+K$. The two normal modes with indices k and $k+K$ are naturally associated with each other. The functions $G_k(s)$ and $G_{k+K}(s)$ are given by

$$G_k(s)=-\frac{V_{k,k+K}V_{k+K,k}}{s^2+\Omega_{k+K}{}^2}=-\eta^2\frac{\Omega_{k+K}{}^2\Omega_k{}^2}{s^2+\Omega_{k+K}{}^2} \tag{D9a}$$

$$G_{k+K}(s)=\frac{-V_{k+K,k}V_{k,k+K}}{s^2+\Omega_k{}^2}=-\eta^2\frac{\Omega_k{}^2\Omega_{k+K}{}^2}{s^2+\Omega_k{}^2}. \tag{D9b}$$

We note from Eq. (D6) that

$$\Omega_L{}^2=\Omega_k{}^2+\Omega_{k+K}{}^2, \tag{D10}$$

where

$$\Omega_L{}^2=[(M_1+M_2)^2/4M_1M_2]\omega_L{}^2 \tag{D11}$$

is the maximum frequency of the alternating diatomic lattice which we are studying. The equations which give the poles of $D_k(s)$ and $D_{k+K}(s)$ both turn out to be

$$(s^2+\Omega_{k+K}{}^2)(s^2+\Omega_K{}^2)=\eta^2\Omega_k{}^2\Omega_{k+K}{}^2; \tag{D12}$$

this gives

$$s^2=(i\omega_\pm)^2=\frac{\Omega_L{}^2}{2}\left\{-1\pm\left[1-(1-\eta^2)\sin^2\frac{\pi k}{N}\right]^{\frac{1}{2}}\right\}, \tag{D13}$$

which is the usual expression for the dispersion relation of the alternating diatomic linear chain.

The expansion for $D_{kn}(s)$ also breaks off after the first term so that the off diagonal contributions are also very simple. These expressions have the same poles as $D_k(s)$ and $D_{k+K}(s)$. The final expressions obtained for $q_k(s)$ and $q_{k+K}(s)$ are

$$q_k(s)=\frac{(s^2+\Omega_{k+K}{}^2)[sQ_k(0)+\dot{Q}_k(0)]-\eta\Omega_{k+K}{}^2[sQ_{k+K}(0)+\dot{Q}_{k+K}(0)]}{(s^2+\Omega_{k+K}{}^2)(s^2+\Omega_k{}^2)-\eta^2\Omega_k{}^2\Omega_{k+K}{}^2},$$

$$q_{k+K}(s)=\frac{(s^2+\Omega_k{}^2)[sQ_{k+K}(0)+\dot{Q}_{k+K}(0)]-\eta\Omega_k{}^2[sQ_k(0)+\dot{Q}_k(0)]}{(s^2+\Omega_{k+K}{}^2)(s^2+\Omega_k{}^2)-\eta^2\Omega_k{}^2\Omega_{k+K}{}^2}. \tag{D14}$$

The inverse Laplace transform of $q_k(s)$ is easily found to be

$$Q_k(t)=\frac{\omega_+{}^2-\Omega_{k+K}{}^2}{\omega_+{}^2-\omega_-{}^2}\left(Q_k(0)\cos\omega_+t+\frac{\dot{Q}_k(0)}{\omega_+}\sin\omega_+t\right)+\frac{\omega_-{}^2-\Omega_{k+K}{}^2}{\omega_-{}^2-\omega_+{}^2}\left[Q_k(0)\cos\omega_-t+\frac{\dot{Q}_k(0)}{\omega_-}\sin\omega_-t\right]$$
$$+\frac{\eta\Omega_{k+K}{}^2}{\omega_+{}^2-\omega_-{}^2}\left[Q_{k+K}(0)\cos\omega_+t+\frac{\dot{Q}_{k+K}(0)}{\omega_+}\sin\omega_+t\right]+\frac{\eta\Omega_{k+K}{}^2}{\omega_-{}^2-\omega_+{}^2}\left(Q_{k+K}(0)\cos\omega_-t+\frac{\dot{Q}_{k+K}(0)}{\omega_-}\sin\omega_-t\right), \tag{D15}$$

where $s=\pm i\omega_+$ and $s=\pm i\omega_-$ are the four roots of Eq. (D12). The corresponding expression for $Q_{k+K}(t)$ is obtained by replacing k by $k+K$ and $k+K$ by k.

We see that, as a result of the interaction $\Phi_{kk'}$, the coordinates Q_k and Q_{k+K} no longer describe dynamically independent modes of motion, although they are still not coupled to any Q corresponding to wave vectors different from k and $k+K$. The normal coordinates for the alternating diatomic linear chain associated with Q_k and Q_{k+K} are the two linear combinations of Q_k and Q_{k+K} which oscillate independently at the frequencies ω_+ and ω_-. The first is in the optical branch and the second is in the acoustical branch. We can obtain them by performing the transformation

$$\xi_+=C_+\{\eta\Omega_k{}^2Q_k+(\Omega_{k+K}{}^2-\omega_-{}^2)Q_{k-K}\}$$
$$\xi_-=C_-\{\eta\Omega_k{}^2Q_k+(\Omega_{k+K}{}^2-\omega_+{}^2)Q_{k+K}\}, \tag{D16}$$

where C_+ and C_- are constants chosen to make the coefficients in the kinetic energy unity.

It is interesting that we have been able to obtain the normal-mode frequencies by looking at only the diagonal part, and also that we have been able to follow the development of the motion in time without transforming to normal coordinates and thus have avoided a certain amount of complication. This procedure should also work for periodic lattices with several kinds of atoms arrayed periodically but in a more complicated way.

We may extrapolate qualitatively from these results, at least in lowest order, in the following way. It is not difficult to show that $\langle \Phi_{kk_1}\Phi_{k_1k}\rangle$ is essentially the Fourier transform of a correlation function for the relative separations of pairs of impurity atoms in the lattice.[41] As long as the order in the lattice is sufficiently longranged that the average product $\langle \Phi_{kk_1}\Phi_{k_1k}\rangle$ is nonvanishing for only a finite number of values of k_1 for a given value of k, the expansion for $G_k(s)$ will terminate after a few terms. This is a result of the difficulty of forming a product $\Phi_{kk_1}\Phi_{k_1k_2}\cdots\Phi_{k_jk}$ in which no intermediate index can equal k with only a few nonvanishing matrix elements. The equation $s^2+\omega_k^2=-G_k(s)$ for the frequencies of the perturbed lattice becomes an algebraic equation of (perhaps) high degree, whose various solutions will correspond to the various branches of the frequency spectrum. In such cases, dissipative effects are not expected. In the opposite case, where the range of order is so short that there is essentially an infinity of nonvanishing products $\langle \Phi_{kk_1}\Phi_{k_1k}\rangle$ for a given k, which is the case we have considered, we have seen that dissipative effects arise. A quantative discussion of the relation between the extent of long-range order in the lattice and the occurrence of dissipative effects has been given by Zwanzig.[41]

[41] R. Zwanzig (private communication).

Nonequilibrium Processes in Isotopically Disordered Crystals. Dependence on Degree of Order

ROBERT ZWANZIG
National Bureau of Standards, Washington, D. C.
(Received December 9, 1960)

A previous paper [A. A. Maradudin, G. H. Weiss, and D. H. Jepsen, J. Math. Phys. 2, 349 (1961)] has shown that approximate normal modes of isotopically disordered crystals decay irreversibly when the arrangement of isotopes is completely random. These results are generalized to crystals with an arbitrary degree of order. In particular, it is shown irreversible behavior occurs whenever the spatial correlation between isotopic species extends over a fixed finite range, as the size of the crystal tends to infinity.

A RECENT treatment, by Maradudin, Weiss, and Jepsen,[1] of nonequilibrium processes in isotopically disordered crystals suggests the following question. How can one describe the transition from reversible behavior in an ordered lattice to irreversible behavior in a disordered lattice? How much disorder does one need for irreversibility? Maradudin, Weiss, and Jepsen have treated the two extreme cases of perfect order and complete randomness in their article; we shall fill in the middle here.

Our answer is that irreversible behavior occurs whenever the correlation between isotopic species has a fixed and finite range as the size of the crystal tends to infinity. If the correlation extends over a region that grows in size as the size of the crystal grows, reversible behavior occurs instead.

Our answer is correct to the same order of perturbation theory as the results of Maradudin, Weiss, and Jepsen, having been found by essentially the same methods.

In order to avoid duplication of material, we shall copy some pertinent equations from I, with just enough discussion to define the symbols.

We study a one-dimensional crystal, with one atom per unit cell, and with periodic boundary conditions. (There does not appear to be anything significantly different about three-dimensional crystals as far as the present discussion is concerned.)

The displacement of the atom at the jth lattice site from its equilibrium position is u_j. Its mass is m_j. The index j runs from 1 to N. Each atom interacts only with its two nearest neighbors; the force constant is γ. Therefore, the equations of motion are

$$m_j\ddot{u}_j=\gamma(u_{j+1}-2u_j+u_{j-1}). \tag{1}$$

In I, the properties of the lattice with an arbitrary mass distribution are related by perturbation theory to those of an average lattice, in which each mass m_j is replaced by the average[2] mass M,

$$M=\langle m\rangle=\frac{1}{N}\sum_{j=1}^{N} m_j. \tag{2}$$

The frequency spectrum of the average lattice is given by

$$\omega_k=\omega_L|\sin(\pi k/N)|, \tag{3}$$

where ω_L is the maximum frequency

$$\omega_L=(4\gamma/M)^{\frac{1}{2}}. \tag{4}$$

The normal coordinates of the average lattice are Q_k. They are related to displacements by

$$u_j=\frac{1}{(NM)^{\frac{1}{2}}}\sum_k Q_k\exp\left(\frac{2\pi ikj}{N}\right). \tag{5}$$

In the arbitrary lattice, Q_k is no longer a normal coordinate. However, it is still by definition a Fourier component of the displacements. In a perturbation theory, where the arbitrary lattice is assumed to be close to the average lattice, the Fourier components Q_k form a logical basis for studying dynamical behavior. They are "approximate" normal coordinates.

In terms of the Q_k, the equations of motion of the arbitrary lattice are

$$\ddot{Q}_k+\omega_k^2Q_k=-\textstyle\sum_{k'}\Phi_{kk'}Q_{k'}. \tag{6}$$

The "perturbation" $\Phi_{kk'}$ is

$$\Phi_{kk'}=\frac{\omega_{k'}^2}{N}\sum_i\left(\frac{M}{m_j}-1\right)\exp\left[\frac{2\pi i}{N}(k'-k)j\right]; \tag{7}$$

it is proportional to the $(k'-k)$th Fourier component of the distribution of $(1/m_j)-(1/M)$.

Because the equations of motion are linear, $Q_k(t)$ is a linear combination of the initial values $Q_{k'}(0)$ and $\dot{Q}_{k'}(0)$,

$$Q_k(t)=\textstyle\sum_{k'}\{A_{kk'}(t)\dot{Q}_{k'}(0)+B_{kk'}(t)Q_k(0)\}. \tag{8}$$

We follow Maradudin, Weiss, and Jepsen, and consider the evolution of a particular component $Q_k(t)$, but we assume that only the same component is excited initially. Then we are interested only in the diagonal

[1] A. A. Maradudin, G. H. Weiss, and D. W. Jepsen, J. Math. Phys. 2, 349 (1961), preceding paper (hereafter referred to as I).

[2] Maradudin, Weiss, and Jepsen average over an ensemble of crystals having a completely random mass distribution. We prefer to consider a single crystal, to interpret the symbol $\langle\ \rangle$ as an arithmetic mean over the lattice sites, and at the end to invoke the law of large numbers. The difference is mainly one of point of view, and will not affect results.

elements of A and B. These are related by

$$B_{kk}(t)=\dot{A}_{kk}(t). \tag{9}$$

It is shown in I that

$$A_{kk}(t)\equiv I_k(t)=\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} ds\frac{e^{sr}}{s^2+\omega_k^2+G_k(s)}, \tag{10}$$

where c is any positive number. It is also shown there that the first two terms in the perturbation expansion of $G_k(s)$ are

$$G_k(s)=\Phi_{kk}-\sum_{k'\neq k}\frac{\Phi_{kk'}\Phi_{k'k}}{s^2+\omega_{k'}^2}+\cdots. \tag{11}$$

This is the starting point for our discussion of the effects of long-range order.

The first-order term in G_k is simply[2]

$$\Phi_{kk}=\omega_k^2\langle(M/m)-1\rangle. \tag{12}$$

The second-order term, which we shall call $\Delta G_k(s)$, is

$$\begin{aligned}\Delta G_k(s)&=-\sum_{k'\neq k}\frac{\Phi_{kk'}\Phi_{k'k}}{s^2+\omega_{k'}^2}\\&=-\sum_{k'\neq k}\frac{\omega_k^2\omega_{k'}^2}{N^2}\frac{1}{s^2+\omega_{k'}^2}\sum_j\sum_{j'}\left(\frac{M}{m_j}-1\right)\\&\quad\times\left(\frac{M}{m_{j'}}-1\right)\exp\left[\frac{2\pi i}{N}(k'-k)(j-j')\right].\end{aligned} \tag{13}$$

Because of the periodic boundary conditions, the sums over j and j' can be replaced by sums over j and $\nu=j'-j$. Then we obtain

$$\begin{aligned}\Delta G_k(s)=-\sum_{k'\neq k}\frac{\omega_k^2\omega_{k'}^2}{s^2+\omega_{k'}^2}\frac{1}{N}\sum_\nu R(\nu)\\\times\exp\left[-\frac{2\pi i}{N}(k'-k)\nu\right],\end{aligned} \tag{14}$$

where $R(\nu)$ is an autocorrelation function of the mass distribution,

$$\begin{aligned}R(\nu)&=\frac{1}{N}\sum_j\left(\frac{M}{m_j}-1\right)\left(\frac{M}{m_{j+\nu}}-1\right)\\&=\left\langle\left(\frac{M}{m_j}-1\right)\left(\frac{M}{m_{j+\nu}}-1\right)\right\rangle.\end{aligned} \tag{15}$$

It will be convenient to separate $R(\nu)$ into two parts,

$$R(\nu)=R_0+R_1(\nu), \tag{16}$$

where

$$R_0=\left(\left\langle\frac{M}{m}\right\rangle-1\right)^2, \tag{17}$$

and

$$R_1(\nu)=\left\langle\left(\frac{M}{m_j}-\left\langle\frac{M}{m}\right\rangle\right)\left(\frac{M}{m_{j+\nu}}-\left\langle\frac{M}{m}\right\rangle\right)\right\rangle \tag{18}$$

The part R_0 is independent of ν and does not contribute to $\Delta G_k(s)$ because

$$\frac{1}{N}\sum_\nu R_0\exp\left[-\frac{2\pi i}{N}(k'-k)\nu\right]=R_0\delta_{kk'}, \tag{19}$$

and because we need only $k\neq k'$. So we shall drop R_0 entirely.

It will be convenient also to factor out the term $R_1(0)$, and to define a function $g(\nu)$ by

$$R_1(\nu)=R_1(0)g(\nu). \tag{20}$$

The term we have factored out has the value

$$R_1(0)=\left\langle\left(\frac{M}{m}-\left\langle\frac{M}{m}\right\rangle\right)^2\right\rangle. \tag{21}$$

In the limit of an infinite, completely random lattice, $R_1(0)$ is the same as the ϵ^2 in I [following Eq. (4.19)]. The function $g(\nu)$ is so normalized that $g(0)=1$.

$R(\nu)$, or $g(\nu)$, is all one needs to know about the mass distribution as long as second-order perturbation theory is sufficient. Let us see what it looks like in several instances.

If there is no correlation at all between different lattice sites, then $g(\nu)$ vanishes[3] except for $\nu=0$,

$$g(\nu)=\delta_{\nu 0}. \tag{22}$$

Suppose that the lattice is perfectly ordered, as for example $ABABABAB\cdots$; then

$$(M/m_j)-\langle M/m_j\rangle$$

merely changes in sign as one moves from one lattice site to the next, and evidently

$$g(\nu)=(-1)^\nu. \tag{23}$$

If the ordering process involves nearest-neighbor interactions of the kind that occur in the one-dimensional Ising lattice, then in the limit of an infinite crystal,

$$g(\nu)=(-\tanh J/kT)^{|\nu|}. \tag{24}$$

The interaction energies of neighboring pairs AA, AB, and BB are $+J$, $-J$, and $+J$; kT is Boltzmann's constant times temperature. At zero temperature (and positive J), Eq. (24) reduces to Eq. (23). At infinite temperature, the completely disordered lattice is obtained.

[3] In a finite lattice this is not strictly true. According to Eqs. (18) and (20), $\Sigma g(\nu)=0$, while we have asserted by Eq. (22) that $\Sigma g(\nu)=1$. The reason for the discrepancy is that in a finite lattice $g(\nu)$ will be of order $(1/N)$ for $\nu\neq 0$. There are N of these terms, and their net effect is to cancel out $g(0)=1$. Later it will become clear that this uncritical passage to the limit of an infinite crystal does not cause any trouble.

In the notation we have just established, the function $G_k(s)$ is

$$G_k(s)=\langle(M/m)-1\rangle\omega_k{}^2+\Delta G_k(s); \tag{25}$$

$$\Delta G_k(s)=-R_1(0)\frac{1}{N}\sum_{k'\neq k}\frac{\omega_k{}^2\omega_{k'}{}^2}{s^2+\omega_{k'}{}^2}\sum_\nu g(\nu) \times\exp\left[-\frac{2\pi i}{N}(k'-k)\nu\right]. \tag{26}$$

On using this, we have to calculate the inverse Laplace transform in Eq. (10) for various $g(\nu)$.

The method of Laplace inversion used by Maradudin, Weiss, and Jepsen depends critically on the singularities of $G_k(s)$ in the complex s plane. In a finite crystal, $G_k(s)$ has simple poles at $s=\pm i\omega_k$, $2N$ in number, and they all lie between $-i\omega_L$ and $+i\omega_L$. In the limit of an infinite crystal, these poles become densely distributed in the interval, and $G_k(s)$ approaches a function having a branch cut from $-i\omega_L$ to $+i\omega_L$, *provided that the coefficients*

$$f(k'-k)=\sum_\nu g(\nu)\exp\left[-\frac{2\pi i}{N}(k'-k)\nu\right] \tag{27}$$

do not vanish for too many values of k'. If the number of values of k' for which they vanish approaches N as N becomes infinite, then the branch-cut method will not work.

For example, if the correlation function $g(\nu)$ is significant over a range of ν that is proportional to N, then $f(k'-k)$ will be of order N for a few values of k', and it will be otherwise negligible. The factor $1/N$ at the front of $\Delta G_k(s)$ will have to be combined with $f(k'-k)$ as N tends to infinity. It cannot then be combined with the sum over k', but the sum over k' will in this case contain only a few values of k' anyhow. In other words, $G_k(s)$ will have only a few simple poles, on the imaginary s axis, in the limit of an infinite crystal. The inverse Laplace transform $I_k(t)$ will show a rather simple periodic behavior, without irreversibility.

If, on the other hand, the correlation function $g(\nu)$ has a fixed and finite range as N tends to infinity, then $f(k'-k)$ will be of the order of unity for roughly N values of k'. The factor $1/N$ at the front of $\Delta G_k(s)$ can be combined with the sum over k', and the sum over k' approaches an integral as N tends to infinity. This is what gives rise to the branch cut and to irreversible behavior.

In the latter case, where $g(\nu)$ has a fixed, finite range, application of the formalism in I is straightforward. There it is shown that for times t that are not too long, $I_k(t)$ is approximately

$$I_k(t)\sim N_k e^{-\Gamma_k t}\sin y_k t, \tag{28}$$

where N_k, Γ_k, and y_k are real positive constants. Since our main interest is in the irreversible character of $I_k(t)$, we shall focus attention on Γ_k and ignore the other constants.

On applying the prescription given in I for calculating Γ_k, we find

$$\Gamma_k=\tfrac{1}{2}R_1(0)\frac{\omega_k{}^2}{(\omega_L{}^2-\omega_k{}^2)^{\frac{1}{2}}}\sum_\nu g(\nu)\cos^2\frac{2\pi k}{N}\nu. \tag{29}$$

When the lattice is completely disordered, so that Eq. (22) applies, we recover Eq. (4.19) of I.

When $g(\nu)$ is the correlation function corresponding to the Ising interaction, Eq. (24), then the sum over ν is elementary. We get

$$\sum_\nu g(\nu)\cos^2\nu\theta=\frac{1}{1-x}+x\frac{\cos2\theta-x}{1-2x\cos2\theta+x^2}, \tag{30}$$

using the abbreviations

$$x=-\tanh J/kT \tag{31}$$

and

$$\theta=2\pi(k/N). \tag{32}$$

As the temperature tends to infinity, x tends to zero, and we again recover the result of Maradudin, Weiss, and Jepsen for the completely random lattice.

More interesting is the behavior of Γ_k at low temperatures. We shall assume that J is positive, so that the preferred configuration is $ABABAB\cdots$. Then x approaches -1 at low temperatures. According to Eqs. (29) and (30), Γ_k is proportional to $x+1$ when $x+1$ is small,

$$\Gamma_k\sim1-\tanh(J/kT)\sim\exp(-2J/kT), \tag{33}$$

so the rate of decay of Q_k vanishes as the temperature drops. There is, however, one expection: if $1+\cos2\theta=0$, then the rate of decay diverges at low temperatures. This happens for values of k such that

$$\lim_{N\to\infty}\frac{k}{N}=\pm\frac{1}{4}. \tag{34}$$

The corresponding Fourier components are actually normal modes of the perfectly ordered $ABABAB$ lattice, but they are very special ones in that appropriate linear combinations of Q_k and Q_{-k} can be assigned to either the optical or the acoustical branch, with different frequencies. Because of this, they are abnormally sensitive to flaws in the $ABABAB$ order. This is why they decay so rapidly at low temperatures.

All other Fourier components damp out very slowly at low temperatures.

Statistical Dynamics of Simple Cubic Lattices. Model for the Study of Brownian Motion. II

ROBERT J. RUBIN
National Bureau of Standards, Washington 25, D. C.
(Received December 13, 1960)

New results concerning the statistical dynamics of a heavy particle in an n-dimensional (nD) cubic lattice are presented. It is demonstrated that this model exhibits many properties which are familiar in the phenomenological theory of Brownian motion. In a well-defined sense, the random thermal motions of a heavy particle in a $1D$ lattice and a $3D$ lattice are accurately described by Kramers' equation for a free particle and a harmonically bound particle, respectively. A related, but not independent, result is that the velocity $v(t)$ and position $u(t)$ of a heavy particle in a $1D$ lattice and a $3D$ lattice constitute two-dimensional stationary Gaussian Markoff processes. It is definitely established that in the case of a $2D$ lattice the stationary Gaussian process $\{v(t),u(t)\}$ is non-Markoffian. In the course of the analysis, several interesting connections between solutions of the discrete lattice equations of motion and solutions of the corresponding continuum equation of motion (the nD wave equation) are uncovered.

I. INTRODUCTION

THE purpose of this paper is to present additional results concerning the statistical dynamics of a heavy particle in an n-dimensional (nD) cubic lattice.[1] The principal objective of the work reported here and in I is the investigation of a mathematically tractable, nontrivial, many-body system which can serve as a model for the study of Brownian motion. One such system satisfying these requirements is a modified nD simple cubic lattice with nearest-neighbor central and noncentral forces. The modification is made by increasing the mass of one of the lattice particles to a very large value. The classical equations of motion for the system can be solved, and the velocity and position of the heavy particle can be expressed as linear combinations of the initial positions and velocities for the entire system. Only one statistical, or nonmechanical, element is introduced: the initial positions and velocities are assumed to be canonically distributed as in thermal equilibrium. On starting with this assumption in the present paper, the rth-order probability distribution function $W_r(v_0,u_0,t_0;\cdots;v_{r-1},u_{r-1},t_{r-1})$ is obtained which expresses the joint probability of finding the velocity v and position u of the heavy particle in the ranges $(v_0, v_0+dv_0; u_0, u_0+du_0)$, $\cdots$, $(v_{r-1}, v_{r-1}+dv_{r-1}; u_{r-1}, u_{r-1}+du_{r-1})$ at the successive times t_0, $\cdots$, t_{r-1}. In addition, the time dependence of the parameters in W_r is examined in detail for the $1D$, $2D$, and $3D$ lattices. Conditional probability distribution functions (pdf's) are also obtained. For example, $P(v_1,u_1,t_1|v_0,u_0,t_0)$ expresses the conditional probability of finding the velocity and position of the heavy particle in the ranges $(v_1, v_1+dv_1; u_1, u_1+du_1)$ at the time t_1 when the velocity and position were in the ranges $(v_0, v_0+dv_0; u_0, u_0+du_0)$ at the time t_0. The conditional pdf $P(v_1,u_1,t_1|v_0,u_0,t_0)$ is of considerable interest because the same type of distribution function is sought in the phenomenological approach to the problem of Brownian motion starting with the Kramers' equation.[2-4] Only one pdf was obtained in I, namely, $P(v_1,t_1|v_0,t_0)$. The time-dependent parameters of this distribution were determined for the $1D$ and $2D$ lattices only.

Instead of repeating the type of analysis used in I to derive the pdf's in the present paper, we make use of known results concerning the properties of multivariate normal pdf's.[5] A basic feature of the work presented here and in I is the linearity of the equations of motion. Thus, the position and velocity of the heavy particle at time t can be expressed as a linear combination of the initial conditions of the entire lattice. Assuming that the initial positions and velocities of the lattice particles are canonically or normally distributed as in thermal equilibrium, it is a routine matter to obtain the pdf of a set of linear combinations of initial conditions such as $W_r(v_0,u_0,t_0;\cdots;v_{r-1},u_{r-1},t_{r-1})$.

Section II contains a description of the lattice model and a summary of the results obtained in paper I which will be needed. In Sec. III, the various pdf's are obtained. In Sec. IV and V, the time-dependent behavior of the pdf's is considered in detail for the $1D$, $2D$, and $3D$ lattices. It is shown that in a well-defined sense the random motion of a heavy particle in a $1D$ lattice and a $3D$ lattice is accurately described by Kramers' equation for a free particle and a harmonically bound particle, respectively. In Sec. VI, the lattice model of Brownian motion is considered from the point of view of the theory of stochastic processes. It is shown that *in the case of a heavy particle in the 1D and 3D lattices, the velocity $v(t)$ and position $u(t)$ together constitute a two-dimensional stationary Gaussian Markoff process.* These results are not independent of those in Sec. V. On the other hand, *in the corresponding case of the 2D lattice, $v(t)$ and $u(t)$ constitute a two-dimensional*

[1] R. J. Rubin, J. Math. Phys. 1, 309 (1960), hereafter referred to as I.

[2] H. A. Kramers, Physica 7, 284 (1940).
[3] S. Chandrasekhar, Revs. Modern Phys. 15, 1 (1943).
[4] M. C. Wang and G. E. Uhlenbeck, Revs. Modern Phys. 17, 323 (1945).
[5] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis* (John Wiley & Sons, Inc., New York, 1958), Chap. 2.

Gaussian process which is not Markoffian. These different types of statistical behavior are related to the different effective (or collective) dynamic properties of the lattice medium with which the heavy particle interacts.

II. SUMMARY OF SOME PROPERTIES OF THE MODEL

The pertinent results in I which will be needed will now be summarized. The system considered is a modification of a uniform nD cubic crystal lattice with nearest-neighbor central and noncentral forces. The modification is made by increasing the mass of one particle located at lattice point $\mathbf{0}=\{0,\cdots,0\}$ to a relatively large value. The equations of motion can be written in matrix notation as[1,6]

$$\mathbf{M}\mathbf{x}_{tt}(t)=-\mathbf{V}\mathbf{x}(t), \tag{1}$$

where $\mathbf{M}$ and $\mathbf{V}$ are the kinetic and potential energy matrices,[7] respectively, the vector $\mathbf{x}(t)$ specifies the configuration of the lattice[8] at time t, and each subscript t denotes differentiation with respect to the time. Periodic boundary conditions are imposed and the number of particles in each lattice direction is $2N+1$. It is shown in I that the position $x[\mathbf{0},t]$ and velocity $x_t[\mathbf{0},t]$ of the heavy particle $\mathbf{0}$ at time t can be expressed as linear combinations of the initial positions and velocities $\mathbf{x}(0)$ and $\mathbf{x}_t(0)$,

$$x[\mathbf{0},t]=M^{-1}\{\mathbf{X}(t)^T\mathbf{M}\mathbf{x}_t(0)+\mathbf{X}_t(t)^T\mathbf{M}\mathbf{x}(0)\} \tag{2}$$

and

$$x_t[\mathbf{0},t]=M^{-1}\{\mathbf{X}_t(t)^T\mathbf{M}\mathbf{x}_t(0)-\mathbf{X}(t)^T\mathbf{V}\mathbf{x}(0)\}, \tag{3}$$

where M is the mass of the heavy particle, and the superscript T denotes the transpose of a matrix. The time-dependent coefficients $\mathbf{X}(t)$ have the very important property that they are a solution of the equations of motion (1) for the special initial condition

$$x[\mathbf{R},0]=0, \quad \text{all } \mathbf{R} \tag{4}$$

and

$$x_t[\mathbf{R},0]=\begin{cases}1, & \mathbf{R}=\mathbf{0}\\ 0, & \mathbf{R}\neq\mathbf{0},\end{cases}$$

i.e.,

$$\mathbf{M}\mathbf{X}_{tt}(t)=-\mathbf{V}\mathbf{X}(t). \tag{4a}$$

This property plays an essential role in obtaining simple statistical formulas.

The initial conditions $\mathbf{x}(0)$ and $\mathbf{x}_t(0)$ are assumed to be canonically distributed as in thermal equilibrium:

$$\mathcal{W}[\mathbf{x}_t(0),\mathbf{x}(0)]=\mathfrak{N}\exp\left\{-\tfrac{1}{2}\begin{pmatrix}\mathbf{x}_t(0)\\ \mathbf{x}(0)\end{pmatrix}^T \times\begin{pmatrix}(kT)^{-1}\mathbf{M} & \mathbf{0}\\ \mathbf{0} & (kT)^{-1}\mathbf{V}\end{pmatrix}\begin{pmatrix}\mathbf{x}_t(0)\\ \mathbf{x}(0)\end{pmatrix}\right\}, \tag{5}$$

where

$$\begin{pmatrix}\mathbf{x}_t(0)\\ \mathbf{x}(0)\end{pmatrix}$$

is a partitioned vector whose first $(2N+1)^n$ components are $\mathbf{x}_t(0)$ and whose second $(2N+1)^n$ components are $\mathbf{x}(0)$.

It will be seen in Sec. III that, as a consequence of the thermal equilibrium assumption, Eq. (5), the joint pdf W_r is a multivariate normal distribution. Although the procedure by which W_r is obtained involves the use of a cumbersome notation, the final results are remarkably simple. The first moments (means) and the second moments (covariances), which completely define W_r, are expressible in terms of the $\mathbf{0}$ component of $\mathbf{X}(t)$ only. The results obtained in Sec. III constitute a generalization of the results of Mazur and Montroll[9] in two respects: (1) the mass of one of the lattice particles is different from the others, and (2) the second moments of the position variables as well as the velocity variables are obtained. Mazur and Montroll investigate questions of ergodicity, Poincaré cycles, and irreversibility for lattices containing a large, but finite, number of identical particles. Their arguments can be applied unchanged to the present generalized problem. However, the primary question which we consider in this paper is the effect of the large mass of lattice particle $\mathbf{0}$ on its statistical dynamical behavior. To eliminate the consideration of Poincaré cycles, which we consider irrelevant to our problem, we treat the limit in which the number of lattice particles $(2N+1)^n$ is infinite. In this limit, when the masses are all equal, there are two characteristic times, the vibrational period associated with the mass m and the lattice force constant κ, and a longer characteristic time associated with the oscillatory decay[9] of the velocity autocorrelation function. [The velocity autocorrelation function decays as t^{-c}, where the value of c, which is of order unity, depends on the dimensionality of the lattice.[9]] In our model, the effect of making the mass of particle $\mathbf{0}$ very large is to introduce a third characteristic time, which is associated with the rate of momentum transfer from the heavy particle to the rest of the lattice.

In this paper, we examine in detail the statistical dynamical behavior of the heavy particle in the $1D$, $2D$, and $3D$ lattices. It will be helpful in understanding

[6] R. J. Rubin, J. Math. Phys. 2, 266(E) (1961).

[7] The potential energy matrix $\mathbf{V}$ has been modified in one respect, namely, each lattice particle is fastened to its equilibrium position by an additional weak spring (force constant $=\kappa'$). This modification is introduced only as a formal convenience to make $\mathbf{V}$ positive definite. Ultimately the results are examined in the limit $\kappa'\to 0$.

[8] The $\mathbf{R}$th component of $\mathbf{x}(t)$ is $x[\mathbf{R},t]$ where $\mathbf{R}$ is a vector whose n components define the equilibrium position of particle $\mathbf{R}$. $x[\mathbf{R},t]$ is the displacement of particle $\mathbf{R}$ from its equilibrium position in one of the lattice directions at time t. It is a property of the lattice that motions in different lattice directions are independent.

[9] P. Mazur and E. W. Montroll, J. Math. Phys. 1, 70 (1960).

some of the dimension-dependent properties of our model to keep in mind the following connection between the lattice equations of motion (1) and the nD wave equation. Equation (1), with all masses equal, is the finite difference analog of the nD wave equation which is obtained by replacing the nD Laplacian by the appropriate second differences.

III. CALCULATION OF PROBABILITY DISTRIBUTION FUNCTIONS

The canonical distribution function, Eq. (5), which characterizes thermal equilibrium in the lattice, is an example of a multivariate normal pdf. We will summarize three theorems concerning these distributions which will be useful in our analysis.[5]

Theorem 1. If the pdf of the components of an m-dimensional vector $\mathbf{Y}$ is

$$(2\pi)^{-m/2}(|\Sigma^{-1}|)^{\frac{1}{2}}\exp\{-\tfrac{1}{2}(\mathbf{Y}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\mathbf{Y}-\boldsymbol{\mu})\}, \quad (6)$$

where $\boldsymbol{\Sigma}^{-1}$ is an $m\times m$ positive definite matrix and $|\Sigma^{-1}|$ denotes the determinant of $\boldsymbol{\Sigma}^{-1}$, then the mean value of $\mathbf{Y}$ is $\boldsymbol{\mu}$ and the covariance matrix is $\boldsymbol{\Sigma}$ (the inverse of $\boldsymbol{\Sigma}^{-1}$). The i, jth element of $\boldsymbol{\Sigma}$ is the covariance or expected value of the product of the ith and jth components of $\mathbf{Y}-\boldsymbol{\mu}$,

$$\Sigma_{ij}=\langle(Y_i-\mu_i)(Y_j-\mu_j)\rangle. \quad (7)$$

A multivariate normal distribution such as (6) is completely specified by $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$; thus the pdf (6) can be denoted as $\mathfrak{n}\{\mathbf{Y}|\boldsymbol{\mu},\boldsymbol{\Sigma}\}$.

Theorem 2. If $\mathbf{Y}$ is distributed according to the pdf $\mathfrak{n}\{\mathbf{Y}|\boldsymbol{\mu},\boldsymbol{\Sigma}\}$, then a set of p linearly independent combinations of the m components of $\mathbf{Y}$, $\mathbf{Z}=\mathbf{D}\mathbf{Y}$, is distributed according to the pdf $\mathfrak{n}\{\mathbf{Z}|\mathbf{D}\boldsymbol{\mu}, \mathbf{D}\boldsymbol{\Sigma}\mathbf{D}^T\}$, where $\mathbf{D}$ is a $p\times m$ matrix.

Theorem 3. Let the components of $\mathbf{Y}$ be divided into two groups composing the subvectors $\mathbf{Y}^{(1)}$ and $\mathbf{Y}^{(2)}$. Suppose the mean $\boldsymbol{\mu}$ is similarly divided into $\boldsymbol{\mu}^{(1)}$ and $\boldsymbol{\mu}^{(2)}$ and suppose the covariance matrix $\boldsymbol{\Sigma}$ of $\mathbf{Y}$ is divided into $\boldsymbol{\Sigma}_{11}$, $\boldsymbol{\Sigma}_{12}$, $\boldsymbol{\Sigma}_{22}$, the covariance matrices of $\mathbf{Y}^{(1)}$, of $\mathbf{Y}^{(1)}$ and $\mathbf{Y}^{(2)}$, and of $\mathbf{Y}^{(2)}$, respectively. Then if the pdf of $\mathbf{Y}$ is $\mathfrak{n}\{\mathbf{Y}|\boldsymbol{\mu},\boldsymbol{\Sigma}\}$, the conditional pdf of $\mathbf{Y}^{(1)}$ given $\mathbf{Y}^{(2)}$ is

$$\mathfrak{n}\{\mathbf{Y}^{(1)}|\boldsymbol{\mu}^{(1)}+\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{Y}^{(2)}-\boldsymbol{\mu}^{(2)}), \boldsymbol{\Sigma}_{11}-\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\}.$$

A. Derivation of the Joint pdf, $W_r(v_0,u_0,t_0;\cdots;v_{r-1},u_{r-1},t_{r-1})$

Let us now consider the problem of obtaining the joint probability that the velocity and position of the heavy particle are v_0 and u_0 at t_0, v_1 and u_1 at t_1, $\cdots$, and v_{r-1} and u_{r-1} at t_{r-1}. The set of linear relations between v_0, $\cdots$, v_{r-1}, u_0, $\cdots$, u_{r-1} and the initial conditions can be represented by the partitioned matrix expression

$$\begin{pmatrix} v_0 \\ \vdots \\ v_{r-1} \\ u_0 \\ \vdots \\ u_{r-1} \end{pmatrix} = \begin{pmatrix} M^{-1}[\mathbf{M}\dot{\mathbf{X}}_t(t_0)]^T & -M^{-1}[\mathbf{V}\mathbf{X}(t_0)]^T \\ \vdots & \vdots \\ M^{-1}[\mathbf{M}\dot{\mathbf{X}}_t(t_{r-1})]^T & -M^{-1}[\mathbf{V}\mathbf{X}(t_{r-1})]^T \\ M^{-1}[\mathbf{M}\mathbf{X}(t_0)]^T & M^{-1}[\mathbf{M}\mathbf{X}_t(t_0)]^T \\ \vdots & \vdots \\ M^{-1}[\mathbf{M}\mathbf{X}(t_{r-1})]^T & M^{-1}[\mathbf{M}\mathbf{X}_t(t_{r-1})]^T \end{pmatrix} \times \begin{pmatrix} \mathbf{x}_t(0) \\ \mathbf{x}(0) \end{pmatrix}, \quad (8)$$

or

$$\begin{pmatrix} \mathbf{v} \\ \mathbf{u} \end{pmatrix} = \mathbf{D}\begin{pmatrix} \mathbf{x}_t(0) \\ \mathbf{x}(0) \end{pmatrix},$$

where the components of the partitioned vector

$$\begin{pmatrix} \mathbf{v} \\ \mathbf{u} \end{pmatrix}$$

are v_0, $\cdots$, v_{r-1}, u_0, $\cdots u_{r-1}$, and $\mathbf{D}$ is the partitioned matrix on the right-hand side of Eq. (8). The pdf of the initial conditions Eq. (5) is

$$\mathfrak{n}\left\{\begin{pmatrix} \mathbf{x}_t(0) \\ \mathbf{x}(0) \end{pmatrix}\middle|\begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \begin{pmatrix} kT\mathbf{M}^{-1} & \mathbf{0} \\ \mathbf{0} & kT\mathbf{V}^{-1} \end{pmatrix}\right\}.$$

From theorem 2, the pdf of

$$\begin{pmatrix} \mathbf{v} \\ \mathbf{u} \end{pmatrix}$$

is

$$W(v_0,u_0,t_0;\cdots;v_{r-1},u_{r-1},t_{r-1}) = \mathfrak{n}\left\{\begin{pmatrix} \mathbf{v} \\ \mathbf{u} \end{pmatrix}\middle|\begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \mathbf{D}\begin{pmatrix} kT\mathbf{M}^{-1} & \mathbf{0} \\ \mathbf{0} & kT\mathbf{V}^{-1} \end{pmatrix}\mathbf{D}^T\right\}. \quad (9)$$

The covariance matrix

$$\mathbf{D}\begin{pmatrix} kT\mathbf{M}^{-1} & \mathbf{0} \\ \mathbf{0} & kT\mathbf{V}^{-1} \end{pmatrix}\mathbf{D}^T$$

reduces to a symmetric partitioned matrix

$$\mathbf{D}\begin{pmatrix} kT\mathbf{M}^{-1} & \mathbf{0} \\ \mathbf{0} & kT\mathbf{V}^{-1} \end{pmatrix}\mathbf{D}^T = \begin{pmatrix} \boldsymbol{\Sigma}^{\mathbf{v},\mathbf{v}} & \boldsymbol{\Sigma}^{\mathbf{v},\mathbf{u}} \\ \boldsymbol{\Sigma}^{\mathbf{u},\mathbf{v}} & \boldsymbol{\Sigma}^{\mathbf{u},\mathbf{u}} \end{pmatrix}, \quad (10)$$

where the elements of the $r\times r$ submatrices are

$$\Sigma^{vv}{}_{ij}=\langle v_iv_j\rangle = kTM^{-2}\{\mathbf{X}_t(t_i)^T\mathbf{M}\mathbf{X}_t(t_j)+\mathbf{X}(t_i)^T\mathbf{V}\mathbf{X}(t_j)\}, \quad (11)$$

$$\Sigma^{vu}{}_{ij}=\langle v_iu_j\rangle = kTM^{-2}\{\mathbf{X}_t(t_i)^T\mathbf{M}\mathbf{X}(t_j)-\mathbf{X}(t_i)^T\mathbf{M}\mathbf{X}_t(t_j)\}, \quad (12)$$

$$\Sigma^{uv}{}_{ij}=\langle u_iv_j\rangle = kTM^{-2}\{\mathbf{X}(t_i)^T\mathbf{M}\mathbf{X}_t(t_j)-\mathbf{X}_t(t_i)^T\mathbf{M}\mathbf{X}(t_j)\}, \quad (13)$$

and

$$\Sigma^{uu}{}_{ij}=\langle u_i u_j\rangle = kTM^{-2}\{\mathbf{X}(t_i)^T\mathbf{M}\mathbf{X}(t_j) + [\mathbf{M}\mathbf{X}_t(t_i)]^T\mathbf{V}^{-1}[\mathbf{M}\mathbf{X}_t(t_j)]\}. \quad (14)$$

Expressions (11)–(14), which are sums over the entire set of lattice points, can be considerably simplified.

First, we will simplify:

$$-(kT)^{-1}M\,\Sigma^{vv}{}_{ij}=M^{-1}\{\mathbf{X}_t(t_i)^T\mathbf{M}[-\mathbf{X}_t(t_j)] - \mathbf{X}(t_i)^T\mathbf{V}[\mathbf{X}(t_j)]\}. \quad (15)$$

The right-hand side of Eq. (15) has the form of Eq. (3), the expression for the velocity of particle **0** as a linear combination of initial conditions. In this case, the initial positions are $\mathbf{x}(0)=\mathbf{X}(t_j)$ and the initial velocities are the reversed velocities $\mathbf{x}_t(0)=-\mathbf{X}_t(t_j)$. Recalling the fact that $\mathbf{X}(t_j)$ and $\mathbf{X}_t(t_j)$ express the state of the lattice at time t_j corresponding to the special initial state (4), we see that with the velocities reversed at time t_j all lattice particles retrace their paths[10] and

$$\Sigma^{vv}{}_{ij}=kTM^{-1}X_t[\mathbf{0}, t_j-t_i]. \quad (11a)$$

Correspondingly simple expressions can be obtained for $\Sigma^{vu}{}_{ij}$ and $\Sigma^{uv}{}_{ij}$ by using Eq. (2). The results are

$$\Sigma^{vu}{}_{ij}=kTM^{-1}X[\mathbf{0}, t_j-t_i] \quad (12a)$$

and

$$\Sigma^{uv}{}_{ij}=-kTM^{-1}X[\mathbf{0}, t_j-t_i]. \quad (13a)$$

Finally, we will simplify the expression for $\Sigma^{uu}{}_{ij}$. First note that the integral of $\mathbf{M}\mathbf{X}_{tt}(t)=-\mathbf{V}\mathbf{X}(t)$ with respect to the time is

$$\mathbf{M}\mathbf{X}_t(t)=M\boldsymbol{\Delta}-\mathbf{V}\int_0^t\mathbf{X}(\sigma)d\sigma, \quad (16)$$

where $\boldsymbol{\Delta}$ is a vector whose components are all zero except for the component **0** which is unity. On substituting Eq. (16) in (14), we find

$$\Sigma^{uu}{}_{ij}=kTM^{-2}\Big\{\mathbf{X}(t_i)^T\mathbf{M}\mathbf{X}(t_j)+M^2V^{-1}{}_{\mathbf{00}} - M\int_0^{t_i}X[\mathbf{0},\sigma]d\sigma - M\int_0^{t_j}X[\mathbf{0},\sigma]d\sigma + \Big[\int_0^{t_i}\mathbf{X}(\sigma)d\sigma\Big]^T\mathbf{V}\Big[\int_0^{t_j}\mathbf{X}(\sigma)d\sigma\Big]\Big\}, \quad (17)$$

where $V^{-1}{}_{\mathbf{00}}$ is the **0,0** element of $\mathbf{V}^{-1}$. Now consider Eq. (2) for $x[\mathbf{0},t_j]$ when $\mathbf{x}_t(0)=-\mathbf{X}_t(t_i)$ and $\mathbf{x}(0)=\mathbf{X}(t_i)$. It can be written as

$$X[\mathbf{0}, t_i-t_j]=M^{-1}\{-\mathbf{X}(t_j)^T\mathbf{M}\mathbf{X}_t(t_i)+\mathbf{X}_t(t_j)^T\mathbf{M}\mathbf{X}(t_i)\} =M^{-1}\{-\mathbf{X}(t_j)^T\mathbf{M}\mathbf{X}_t(t_i)+\mathbf{X}(t_i)^T\mathbf{M}\mathbf{X}_t(t_j)\}. \quad (18)$$

[10] In I, the only covariance considered was $\Sigma^{vv}{}_{ii}=kTM^{-2}\times\{\mathbf{X}_t(t_i)^T\mathbf{M}\mathbf{X}_t(t_i)+\mathbf{X}(t_i)^T\mathbf{V}\mathbf{X}(t_i)\}$. The term in braces was identified with twice the lattice energy associated with the special initial condition (4), i.e., $2(M/2)$.

By using Eq. (16) for $\mathbf{M}\mathbf{X}_t(t_i)$, Eq. (18) can be rewritten as

$$X[\mathbf{0}, t_i-t_j]=M^{-1}\Big\{-MX[\mathbf{0},t_j]+\mathbf{X}(t_i)^T\mathbf{M}\mathbf{X}_t(t_j) + \mathbf{X}(t_j)^T\mathbf{V}\int_0^{t_i}\mathbf{X}(\sigma)d\sigma\Big\}. \quad (19)$$

By integrating both sides of Eq. (19) with respect to t_j and rearranging, one obtains

$$\Big[\int_0^{t_i}\mathbf{X}(\sigma)d\sigma\Big]^T\mathbf{V}\Big[\int_0^{t_j}\mathbf{X}(\sigma)d\sigma\Big]+\mathbf{X}(t_i)^T\mathbf{M}\mathbf{X}(t_j) =M\int_0^{t_j}X[\mathbf{0}, t_i-\sigma]d\sigma+M\int_0^{t_j}X[\mathbf{0},\sigma]d\sigma =M\int_{t_i-t_j}^{t_i}X[\mathbf{0},\sigma]d\sigma+M\int_0^{t_j}X[\mathbf{0},\sigma]d\sigma. \quad (20)$$

On substituting the right-hand side of Eq. (20) in Eq. (17), the final expression for $\Sigma^{uu}{}_{ij}$ is

$$\Sigma^{uu}{}_{ij}=kTV^{-1}{}_{\mathbf{00}}-kTM^{-1}\int_0^{t_i-t_j}X[\mathbf{0},\sigma]d\sigma. \quad (14a)$$

We have shown that, aside from the factor kTM^{-1}, all the covariances of the joint pdf (9) can be expressed in terms of a single function $X[\mathbf{0}, t_i-t_j]$, its integral and derivative, and the quantity $V^{-1}{}_{\mathbf{00}}$. The functions $X_t[\mathbf{0},t]$ and $\int_0^t X[\mathbf{0},\sigma]d\sigma$ are even functions of t, and $X[\mathbf{0},t]$ is an odd function of t. Note that the quantity $V^{-1}{}_{\mathbf{00}}$ approaches infinity in the physically interesting limit in which the added spring constant $\kappa'\to 0$.[7] Hence, the covariance approaches infinity. This limiting behavior is to be expected for the model considered here in which periodic boundary conditions are imposed. However, it will be shown that the conditional pdf's of interest have conditional covariance matrices which remain finite in the limit $\kappa'\to 0$.

B. Derivation of Conditional pdf's

We consider some specific examples of conditional pdf's.

1. $P(v_1,t|v_0,0)$

The conditional distribution $P(v_1,t_1|v_0,0)$ was determined in I. We will rederive the result from the joint pdf $W(v_0,0;v_1,t)$ for the values of the velocity v_0 at $t_0=0$ and v_1 at t. According to Eqs. (9) and (11a),

$$W(v_0,0;v_1,t)=\mathfrak{n}\Big\{\begin{pmatrix}v_0\\ v_1\end{pmatrix}\Big|\begin{pmatrix}0\\ 0\end{pmatrix}, \begin{pmatrix}kTM^{-1} & kTM^{-1}X_t[\mathbf{0},t]\\ kTM^{-1}X_t[\mathbf{0},t] & kTM^{-1}\end{pmatrix}\Big\}, \quad (21)$$

where $X_t[\mathbf{0},0]=1$. On using theorem 3 for the pdf (21), with $\mathbf{Y}^{(1)}=v_1$, $\mathbf{Y}^{(2)}=v_0$, $\mathbf{\Sigma}_{11}=\mathbf{\Sigma}_{22}=kTM^{-1}$, and $\mathbf{\Sigma}_{12}=kTM^{-1}X_t[\mathbf{0},t]$, we obtain the conditional pdf

$$P(v_1,t|v_0,0) = \mathfrak{n}\{v_1|v_0X_t[\mathbf{0},t],\ kTM^{-1}(1-X_t^2[\mathbf{0},t])\}, \quad (22)$$

which is identical with the result in I.

2. $P(u_1,t|u_0,0)$

The joint pdf $W(u_0,0;u_1,t)$ is, according to Eqs. (9) and (14a),

$$W(u_0,0;u_1,t)=\mathfrak{n}\left\{\begin{pmatrix}u_0\\u_1\end{pmatrix}\middle|\begin{pmatrix}0\\0\end{pmatrix},\ \begin{bmatrix} kTV^{-1}{}_{00} & kTV^{-1}{}_{00}-kTM^{-1}\int_0^t X[\mathbf{0},\sigma]d\sigma \\ kTV^{-1}{}_{00}-kTM^{-1}\int_0^t X[\mathbf{0},\sigma]d\sigma & kTV^{-1}{}_{00}\end{bmatrix}\right\}. \quad (23)$$

Again using theorem 3, we obtain the conditional pdf

$$P(u_1,t|u_0,0)=\mathfrak{n}\left\{u_1\middle|\left[1-\nu\int_0^t X[\mathbf{0},\sigma]d\sigma\right]u_0,\ 2kTM^{-1}\int_0^t X[\mathbf{0},\sigma]d\sigma\left[1-\tfrac{1}{2}\nu\int_0^t X[\mathbf{0},\sigma]d\sigma\right]\right\}, \quad (24)$$

where $\nu=(MV^{-1}{}_{00})^{-1}$. The variance $\langle(u_1-\langle u_1\rangle)^2\rangle$ approaches a finite limit for fixed t as $\kappa'\to 0$ whereas the elements of the 2×2 covariance matrix in (23) are infinite in the same limit.

3. $P(v_1,u_1,t|v_0,u_0,0)$

The joint pdf $W(v_0,u_0,0;v_1,u_1,t)$ is

$$\mathfrak{n}\left\{\begin{bmatrix}v_0\\v_1\\u_0\\u_1\end{bmatrix}\middle\|\begin{bmatrix}0\\0\\0\\0\end{bmatrix},\ \begin{bmatrix} kTM^{-1} & kTM^{-1}X_t[\mathbf{0},t] & 0 & +kTM^{-1}X[\mathbf{0},t] \\ kTM^{-1}X_t[\mathbf{0},t] & kTM^{-1} & -kTM^{-1}X[\mathbf{0},t] & 0 \\ 0 & -kTM^{-1}X[\mathbf{0},t] & kTV^{-1}{}_{00} & kT\left(V^{-1}{}_{00}-M^{-1}\int_0^t X[\mathbf{0},\sigma]d\sigma\right) \\ +kTM^{-1}X[\mathbf{0},t] & 0 & kT\left(V^{-1}{}_{00}-M^{-1}\int_0^t X[\mathbf{0},\sigma]d\sigma\right) & kTV^{-1}{}_{00}\end{bmatrix}\right\}.$$

Using theorem 3, with

$$\mathbf{Y}^{(2)}=\begin{pmatrix}v_0\\u_0\end{pmatrix},\quad \mathbf{Y}^{(1)}=\begin{pmatrix}v_1\\u_1\end{pmatrix},\quad \mathbf{\Sigma}_{11}=\begin{pmatrix}kTM^{-1} & 0\\ 0 & kTV^{-1}{}_{00}\end{pmatrix},$$

$$\mathbf{\Sigma}_{22}=\mathbf{\Sigma}_{11}\quad\text{and}\quad \mathbf{\Sigma}_{12}=\begin{bmatrix}kTM^{-1}X_t[\mathbf{0},t] & -kTM^{-1}X[\mathbf{0},t]\\ kTM^{-1}X[\mathbf{0},t] & kT\left[V^{-1}{}_{00}-M^{-1}\int_0^t X[\mathbf{0},\sigma]d\sigma\right]\end{bmatrix},$$

we obtain the joint conditional pdf

$$P(v_1,u_1,t|v_0,u_0,0)=\mathfrak{n}\left\{\begin{pmatrix}v_1\\u_1\end{pmatrix}\middle|\begin{bmatrix}v_0X_t[\mathbf{0},t]-u_0\nu X[\mathbf{0},t]\\ v_0X[\mathbf{0},t]+u_0\left[1-\nu\int_0^t X[\mathbf{0},\sigma]d\sigma\right]\end{bmatrix},\ \mathbf{\Sigma}\right\}, \quad (25)$$

where

$$\mathbf{\Sigma}=\begin{bmatrix} kTM^{-1}\{1-X_t^2[\mathbf{0},t]-\nu X^2[\mathbf{0},t]\} & kTM^{-1}X[\mathbf{0},t]\left\{1-X_t[\mathbf{0},t]-\nu\int_0^t X[\mathbf{0},\sigma]d\sigma\right\} \\ kTM^{-1}X[\mathbf{0},t]\left\{1-X_t[\mathbf{0},t]-\nu\int_0^t X[\mathbf{0},\sigma]d\sigma\right\} & kTM^{-1}\left\{2\int_0^t X[\mathbf{0},\sigma]d\sigma-X^2[\mathbf{0},t]-\nu\left[\int_0^t X[\mathbf{0},\sigma]d\sigma\right]^2\right\}\end{bmatrix}. \quad (26)$$

IV. VALUE OF ν AND TIME-DEPENDENT BEHAVIOR OF $X_t[0,t]$, $X[0,t]$, AND $\int_0^t X[0,\sigma]d\sigma$.

In the remainder of this paper, we will confine our discussion to the limit in which the number of particles in the nD system is infinite and in which the added force constant[7] κ' is zero. It was shown in Sec. III that the means and covariances of the various pdf's were expressible in terms of the parameter ν and the function $X[\mathbf{0},t]$ and its integral and derivative. Note that $X_t[\mathbf{0},t]$ and $X[\mathbf{0},t]$ are the velocity and position of the heavy particle at time t corresponding to the special initial conditions (4). This special initial value problem will be referred to as Problem 1. The quantity

$$\int_0^t X[\mathbf{0},\sigma]d\sigma$$

has a similar physical interpretation which is suggested by the following form of Eq. (16),

$$\mathbf{M}\mathbf{x}_{tt}(t)=M\boldsymbol{\Delta}-\mathbf{V}\mathbf{x}(t). \tag{16a}$$

Equation (16a) is the set of equations of motion for an nD lattice in which a constant force of magnitude M is suddenly applied to the heavy particle; and $\int_0^t \mathbf{X}(\sigma)d\sigma$ is the solution of Eq. (16a) corresponding to the special initial conditions

$$\mathbf{x}(0)=\mathbf{0} \quad \text{and} \quad \mathbf{x}_t(0)=\mathbf{0}. \tag{16b}$$

Thus it is seen that $\int_0^t X[\mathbf{0},\sigma]d\sigma$ is the position of the heavy particle at time t when a force M is suddenly applied to $\mathbf{0}$ in a lattice which is initially at rest with all particles at their equilibrium position. This special initial value problem will be referred to as Problem 2. In this Section we consider the foregoing quantities in detail.

A. Value of ν in nD Lattice

First consider the value of $\nu=(MV^{-1}{}_{\mathbf{00}})^{-1}$. Montroll and Potts[11] obtained an expression for $V^{-1}{}_{\mathbf{00}}$ in a study of the effect of defects on lattice vibrations. Their expression for $V^{-1}{}_{\mathbf{00}}$ in the limit of an infinite system is, when translated into our notation,

$$V^{-1}{}_{\mathbf{00}}=\left(\frac{1}{2\pi}\right)^n\int_0^{2\pi}\cdots\int_0^{2\pi}d\theta_1\cdots d\theta_n \times\{\kappa'+2\sum_{j=1}^{n}\kappa_j(1-\cos\theta_j)\}^{-1}, \tag{27}$$

where κ_j is the force constant associated with displacements relative to the jth lattice direction. In the case of the $1D$ and $2D$ lattices, in the limit $\kappa'\to 0$, the integral in Eq. (27) diverges as $(\kappa')^{-\frac{1}{2}}$ and $-\ln\kappa'$, respectively. However, in the nD lattice ($n\geqslant 3$), the integral is finite.[12] For the $3D$ lattice and $\kappa_j=\kappa, j=1, 2, 3$ the value of $V^{-1}{}_{\mathbf{00}}$ is[13]

$$V^{-1}{}_{\mathbf{00}}=\tfrac{1}{4}\zeta_0\kappa^{-1}, \tag{28}$$

where

$$\zeta_0=8\pi^{-2}[18+12\sqrt{2}-10\sqrt{3}-7\sqrt{6}] \times\mathcal{K}^2[(2-\sqrt{3})(\sqrt{3}-\sqrt{2})]\cong 1.019, \tag{28a}$$

and $\mathcal{K}[a]$ is the complete elliptic integral of the first kind. Thus, for the $1D$, $2D$, and $3D$ lattices, with $\kappa_1=\kappa_2=\kappa_3=\kappa$, the values of

$$\nu_n=\lim_{\kappa'\to 0}\lim_{N\to\infty}\nu$$

are, respectively, $\nu_1=0$, $\nu_2=0$, and $\nu_3=4\kappa(M\zeta_0)^{-1}$.

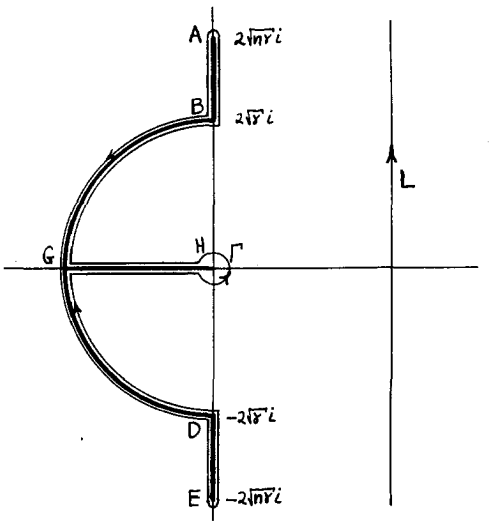

FIG. 1. The cut in the p plane and the paths of integration. The integral referred to as L' is the line integral which starts near G, encircles the origin H, and returns to G.

B. Time-Dependent Behavior of $X_t[0,t]$, $X[0,t]$, and $\int_0^t X[0,\sigma]d\sigma$ in the $1D$, $2D$, and $3D$ Lattices

The integral representation obtained in I for $X[\mathbf{0},t]$ is

$$X[\mathbf{0},t]=\frac{Q+1}{2\pi i}\int_L\frac{p^{-1}e^{pt}dp}{Qp+(p\zeta[\mathbf{0},p])^{-1}}, \tag{29}$$

where $Q+1=M/m$ is the ratio of the masses of the heavy and light particles, and where L, the path of integration, is a line parallel to the imaginary p axis and to the right of all singularities of the integrand (see Fig. 1). The function $\zeta[\mathbf{0},p]$ is, in the limit $N\to\infty$

[11] E. W. Montroll and R. B. Potts, Phys. Rev. **102**, 72 (1956).

[12] The value of the integral in Eq. (27) as a function of n for $\kappa'=0$ and for all $\kappa_j=\kappa$ has been given by E. W. Montroll, J. Soc. Indust. Appl. Math. **4**, 241 (1956), in connection with a study of random walks on nD lattices. In this reference the divergence of the integral in the $1D$ and $2D$ lattices is associated with the eventual certain return of a random walker to his starting point in an infinitely long walk on these lattices.

[13] G. N. Watson, Quar. J. Math., Oxford Series (2) **10**, 266 (1939).

and $\kappa' \to 0$, with $\kappa_j = \kappa$

$$\zeta[\mathbf{0},p] = \left(\frac{1}{2\pi}\right)^n \int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} d\theta_1 \cdots d\theta_n \times \{p^2 + 2m^{-1}\kappa \sum_{j=1}^{n} (1-\cos\theta_j)\}^{-1}, \quad (30)$$

an integral similar to the one which appears in Eq. (27). The fact that Q is large compared to unity plays a dominant role in the following analysis. On the other hand, it should be noted that if the mass of particle **0** is the same as the mass of the other lattice particles (i.e., $Q=0$), then it follows from Eq. (29) that $\zeta[\mathbf{0},p]$ is the Laplace transform of $X[\mathbf{0},t]$. The integrations in Eq. (30) can be carried out explicitly for the $1D$ and $2D$ lattices to give, respectively,

$$\zeta[\mathbf{0},p] = p^{-1}[p^2+4\gamma]^{-\frac{1}{2}} \quad (31)$$

and

$$\zeta[\mathbf{0},p] = [p^2+4\gamma]^{-1}\, {}_2F_1[\tfrac{1}{2}, \tfrac{1}{2}; 1; 16\gamma^2(p^2+4\gamma)^{-2}], \quad (32)$$

where ${}_2F_1[\ ]$ is a hypergeometric function and $\gamma = m^{-1}\kappa$.

In I, the time-dependent behavior of $X_t[\mathbf{0},t]$ was determined for the $1D$ and $2D$ lattices. The method of analysis developed there will now be used for the function $\int_0^t X[\mathbf{0},\sigma]d\sigma$ in the $1D$, $2D$, and $3D$ lattices, and the results for the functions $X[\mathbf{0},t]$ and $X_t[\mathbf{0},t]$ will merely be stated. Before considering the detailed behavior for each lattice, we note the following five facts concerning $\zeta[\mathbf{0},p]$ and the integral representation

$$\frac{Q+1}{2\pi i} \int_L \frac{p^{-s}e^{pt}dp}{Qp+(p\zeta[\mathbf{0},p])^{-1}}, \quad s=0,\ 1,\ \text{and}\ 2, \quad (33)$$

for an nD lattice:

(1) The only possible singularities of $\zeta[\mathbf{0},p]$ as a function of p are located on the imaginary axis at $p=0$, $\pm 2\gamma^{\frac{1}{2}}i$, $\cdots$, $\pm 2(n\gamma)^{\frac{1}{2}}i$. Consequently, a cut can be constructed in the p plane as shown in Fig. 1, and the path of integration can be closed around the cut.

(2) The closed path of integration can be shrunk so that in the limit, the contour integral Eq. (33) is the sum of three parts: (i) the difference of the line integrals on either side of the portion of the cut $ABGDE$, (ii) the contributions of the residues of the integrand at poles which are zeros of the denominator

$$D(p) = Qp + (p\zeta[\mathbf{0},p])^{-1}, \quad (34)$$

and (iii) the line integral L' around the origin following the cut GH and the small circle Γ. Thus the radius of the circle Γ in Fig. 1 is taken to be so small that any zeros of $D(p)$ lie outside Γ.

(3) The difference $\delta_n(t)$ of the line integrals on either side of the portion of the cut $ABGDE$ is

$$\delta_n^{(s)}(t) = \frac{Q+1}{2\pi i} \int_{E\to A} p^{-s}e^{pt} \left\{ \left[Qp + \frac{1}{p\zeta_R[\mathbf{0},p]} \right]^{-1} - \left[Qp + \frac{1}{p\zeta_L[\mathbf{0},p]} \right]^{-1} \right\} dp, \quad (35)$$

where $\int_{E\to A}$ denotes the line integral along $ABGDE$ from E to A, and $\zeta_R[\mathbf{0},p]$ and $\zeta_L[\mathbf{0},p]$ denote, respectively, the values of $\zeta[\mathbf{0},p]$ on the right- and left-hand sides of the cut. On simplifying Eq. (35) for $\delta_n^{(s)}(t)$, one obtains

$$\delta_n^{(s)}(t) = \frac{Q+1}{2\pi i} \int_{E\to A} p^{-s}e^{pt} \left\{ \frac{(p\zeta_L[\mathbf{0},p])^{-1} - (p\zeta_R[\mathbf{0},p])^{-1}}{[Qp+(p\zeta_R[\mathbf{0},p])^{-1}][Qp+(p\zeta_L[\mathbf{0},p])^{-1}]} \right\} dp. \quad (36)$$

It can be verified from Eq. (36) and the expressions for $\zeta[\mathbf{0},p]$ in Eqs. (31), (32), and (30) that the magnitude of $\delta_n^{(s)}(t)$ is bounded by $C_n(4\gamma)^{-\frac{1}{2}}Q^{-1}$ for all times t, where C_n depends on n but is of order unity. It was shown in detail in I that $C_1 = \sqrt{2}$. Although the foregoing estimate of $|\delta_n^{(s)}(t)|$ is sufficient for our purposes, a sharper estimate is obtainable. The method which can be used is illustrated for the $1D$ lattice in Appendix A where it is shown that

$$|\delta_1^{(s)}(t)| < (4\gamma)^{-s/2} 2^{\frac{1}{2}} Q^{-1} \min\{1, \pi^{-\frac{1}{2}}(2\gamma^{\frac{1}{2}}t)^{-\frac{1}{2}}\}, \quad (37)$$

where $\min\{a,b\}$ denotes the smaller of a and b.

(4) Since $Q \gg 1$, the only zeros of $D(p)$, Eq. (34), are located near $p=0$. Therefore, the expansion of $\zeta[\mathbf{0},p]$ for $|p| \ll 1$ can be used to locate the zeros of $D(p)$ for each lattice.

(5) It follows from (3) and (4) that appreciable contributions to the line integral (33) can arise only from zeros of $D(p)$ in the neighborhood of $p=0$ and from the line integral L' near $p=0$. There is a further simplification, however, in the case of the $1D$ and $3D$ lattices which we now consider. It was noted in I that in the case of the $1D$ lattice, the denominator

$$D(p) = Qp + [p^2+4\gamma]^{\frac{1}{2}}$$

is an analytic function of p in the neighborhood of $p=0$ and at $p=0$, and so

$$D(p) = Qp + 2\gamma^{\frac{1}{2}}(1 + p^2/8\gamma + \cdots).$$

Consequently, the line integral L' is equal to the sum of the residues of the integrand at $p=0$ and at any zeros of $D(p)$ on the negative real axis near $p=0$. The final result for the integral in Eq. (33) in the $1D$ case is

$$\frac{Q+1}{2\pi i} \int_L \frac{p^{-s}e^{pt}dp}{Qp+[p^2+4\gamma]^{\frac{1}{2}}} \doteq \frac{Q}{2\pi i} \oint_C \frac{p^{-s}e^{pt}dp}{Qp+2\gamma^{\frac{1}{2}}}, \quad (38)$$

where C is a closed contour surrounding $p=0$ and all zeros of $Qp+[p^2+4\gamma]^{\frac{1}{2}}$. We have introduced the special approximate equality sign $\doteq$ to indicate that the magnitude of the difference between the right- and left-hand sides of the equality is bounded by a quantity at most of order $C_n(4\gamma)^{-s/2}Q^{-1}$.

A similar simplification is found for the $3D$ lattice. The expansion of $\zeta[\mathbf{0},p]$, Eq. (30), for $|p|\ll 1$ in the case of the $3D$ lattice has been given by Kac and Berlin[14] in their paper on the spherical approximation in the Ising Model. The result is[15]

$$\zeta[\mathbf{0},p]=(4\gamma)^{-1}\{\zeta_0-(\gamma^{\frac{1}{2}}\pi)^{-1}p+\cdots\}, \tag{39}$$

where ζ_0 is defined in Eqs. (28) and (28a). The principal feature of the expansion in Eq. (39) is that $p=0$ is not a branch point. As in the $1D$ lattice, the only significant contributions to the value of the integral (33) come from the residues of the integrand at the poles near $p=0$, i.e.,

$$\frac{Q+1}{2\pi i}\int_L \frac{p^{-s}e^{pt}dp}{Qp+(p\zeta[\mathbf{0},p])^{-1}} \doteq \frac{Q}{2\pi i}\oint_C \frac{p^{-s}e^{pt}dp}{Qp+4\gamma[\zeta_0 p-(\gamma^{\frac{1}{2}}\pi)^{-1}p^2]^{-1}}. \tag{40}$$

Although we will not give the details here, it can be shown that in $(2r+1)D$ lattices there is no branch point singularity at $p=0$, but that in $(2r)D$ lattices there is a logarithmic branch point at $p=0$. This difference is consistent with known fundamental differences in the nature of the solutions of the $(2r+1)D$ and $2rD$ wave equations.[16]

(a) 1D Lattice

We will obtain approximate expressions [in the sense of Eq. (38)] for $\int_0^t X[\mathbf{0},\sigma]d\sigma$, $X[\mathbf{0},t]$, and $X_t[\mathbf{0},t]$. The expression for $\int_0^t X[\mathbf{0},\sigma]d\sigma$ is

$$\int_0^t X[\mathbf{0},\sigma]d\sigma \doteq \frac{Q}{2\pi i}\oint_C \frac{p^{-2}e^{pt}dp}{Qp+2\gamma^{\frac{1}{2}}} \tag{41}$$

or

$$\int_0^t X[\mathbf{0},\sigma]d\sigma \doteq \frac{Q}{4\gamma}\{2\gamma^{\frac{1}{2}}t-Q[1-\exp(-2\gamma^{\frac{1}{2}}Q^{-1}t)]\}. \tag{42}$$

The values of $X[\mathbf{0},t]$ and $X_t[\mathbf{0},t]$ are

$$X[\mathbf{0},t]\doteq Q(2\gamma^{\frac{1}{2}})^{-1}\{1-\exp(-2\gamma^{\frac{1}{2}}Q^{-1}t)\} \tag{43}$$

and

$$X_t[\mathbf{0},t]\doteq \exp(-2\gamma^{\frac{1}{2}}Q^{-1}t). \tag{44}$$

The last expression was obtained in I. It is seen that in Problem 1 [Eqs. (4) and (4a)] the velocity and position of the heavy particle behave as though the particle were moving in a viscous medium with a resistance proportional to its velocity. In Problem 2 [Eqs. (16a) and (16b)], the position of the heavy particle increases linearly with the time for large t. Thus the heavy particle reaches a terminal velocity, as it would if it were moving through a viscous medium. In addition, it should be noted that the equations of motion for the lattice are the discrete analog of the continuum equation of motion for an elastic string. The result for the analog of Problem 2 for the wave equation is the same as that found above.

(b) 3D Lattice

Next consider the functions $\int_0^t X[\mathbf{0},\sigma]d\sigma$, $X[\mathbf{0},t]$, and $X_t[\mathbf{0},t]$ for the $3D$ lattice. On introducing the new variable $p=2\gamma^{\frac{1}{2}}\rho$ in the integral representation for $\int_0^t X[\mathbf{0},\sigma]d\sigma$, we have

$$\int_0^t X[\mathbf{0},\sigma]d\sigma \doteq \frac{Q(4\gamma)^{-1}}{2\pi i}\oint_C \frac{\exp(2\gamma^{\frac{1}{2}}\rho t)}{Q\rho^2+(\zeta_0-2\pi^{-1}\rho)^{-1}}\frac{d\rho}{\rho}. \tag{45}$$

The two zeros, $\rho_\pm$, of $D(\rho)=Q\rho+(\zeta_0\rho-2\pi^{-1}\rho^2)^{-1}$ can be found by substituting $\omega e^{\pm(\frac{1}{2}\pi+\delta)i}$ for ρ in

$$Q\rho^2\doteq -(\zeta_0-2\pi^{-1}\rho)^{-1}, \tag{46}$$

and solving the following pair of simultaneous equations for the real and imaginary parts of (46):

$$Q\omega^2\cos 2\delta \doteq \frac{\zeta_0+2\pi^{-1}\omega\sin\delta}{\zeta_0^2+4\pi^{-2}\omega^2+4\pi^{-1}\omega\zeta_0\sin\delta}, \tag{47}$$

and

$$Q\omega^2\sin 2\delta \doteq \frac{2\pi^{-1}\omega\cos\delta}{\zeta_0^2+4\pi^{-2}\omega^2+4\pi^{-1}\omega\zeta_0\sin\delta}. \tag{48}$$

Both ω and δ are small compared to one, so it can be seen that the solution of (47) and (48) is

$$\omega\doteq(Q\zeta_0)^{-\frac{1}{2}} \tag{49}$$

and

$$\delta\doteq\pi^{-1}\zeta_0^{-\frac{3}{2}}Q^{-\frac{1}{2}}. \tag{50}$$

On using the theory of residues,[17] the expression for $\int_0^t X[\mathbf{0},\sigma]d\sigma$ is

$$\int_0^t X[\mathbf{0},\sigma]d\sigma \doteq \frac{Q}{4\gamma}\left\{\zeta_0+\frac{\exp(2\gamma^{\frac{1}{2}}\rho_+ t)}{\rho_+^2 D'(\rho_+)}+\frac{\exp(2\gamma^{\frac{1}{2}}\rho_- t)}{\rho_-^2 D'(\rho_-)}\right\}, \tag{51}$$

where

$$D'(\rho_\pm)=Q+2\pi^{-1}\rho_\pm^{-1}(\zeta_0-2\pi^{-1}\rho_\pm)^{-2}-\rho_\pm^{-2}(\zeta_0-2\pi^{-1}\rho_\pm)^{-1},$$

and $\rho_\pm=\omega e^{\pm(\frac{1}{2}\pi+\delta)i}$. Equation (51) can be simplified with the aid of the relation $Q\rho_\pm^2=-(\zeta_0-2\pi^{-1}\rho_\pm)^{-1}$.

[14] M. Kac and T. H. Berlin, Phys. Rev. **86**, 621 (1952).

[15] Some of the higher order terms in the expansion which are actually not needed here are given by A. A. Maradudin, E. W. Montroll, G. H. Weiss, R. Herman, and H. W. Milnes, Acad. roy. Belg. Cl. sci. Mém. **XIV**, 7 (1960).

[16] R. Courant and D. Hilbert, *Methoden der Mathematischen Physik* (Interscience Publishers, Inc., New York, 1943), Vol. 2, Chap. 6.

[17] L. A. Pipes, *Applied Mathematics for Engineers and Physicists* (McGraw-Hill Book Company, Inc., New York, 1946), p. 464.

TABLE I. Solution of equations (57) and (58) for several different values of Q.

Q	ϵ	δ	$\ln(8/\epsilon^2)$	$Q\epsilon^2\pi^{-1}$
127	$10^{-1.5}$	0.1863	0.1112	0.0404
811	$10^{-2.0}$	0.1483	0.0886	0.0258
5590	$10^{-2.5}$	0.1232	0.0736	0.0178
40800	$10^{-3.0}$	0.1053	0.0629	0.0130

After some reduction and after neglecting terms of order Q^{-1} with respect to unity, the final expression for $\int_0^t X[\mathbf{0},\sigma]d\sigma$ is

$$\int_0^t X[\mathbf{0},\sigma]d\sigma \doteq \frac{Q\zeta_0}{4\gamma}\{1-e^{-\frac{1}{2}\beta t} \times[\cos\omega t+\beta(2\omega)^{-1}\sin\omega t]\}, \quad (52)$$

where $\beta=4\zeta_0^{-2}\pi^{-1}\gamma^{\frac{1}{2}}Q^{-1}$ and $\omega=2\gamma^{\frac{1}{2}}(Q\zeta_0)^{-\frac{1}{2}}$. In the same way, the following values can be obtained for $X[\mathbf{0},t]$ and $X_t[\mathbf{0},t]$,

$$X[\mathbf{0},t]\doteq(Q\zeta_0/4\gamma)^{\frac{1}{2}}\sin\omega t e^{-\frac{1}{2}\beta t} \quad (53)$$

and

$$X_t[\mathbf{0},t]\doteq e^{-\frac{1}{2}\beta t}[\cos\omega t-\beta(2\omega)^{-1}\sin\omega t]. \quad (54)$$

From the point of view of Problems **1** and 2, it is seen that the heavy particle in a $3D$ lattice appears to behave like a damped harmonic oscillator. There is also a connection with the analog of Problem 2 for a continuous elastic medium, i.e., for the $3D$ wave equation.

(c) 2D Lattice

Finally we consider the functions $\int_0^t X[\mathbf{0},\sigma]d\sigma$, $X[\mathbf{0},t]$, and $X_t[\mathbf{0},t]$ for the case of the $2D$ lattice in which there is a significant contribution from the line integral L' around the cut GH. On introducing the variable $2\gamma^{\frac{1}{2}}\rho$ for p, the exact expression for $\int_0^t X[\mathbf{0},\sigma]d\sigma$, according to Eqs. (33) and (32), is

$$\int_0^t X[O,\sigma]d\sigma=\frac{(Q+1)(4\gamma)^{-1}}{2\pi i} \times\int_L \frac{\exp(2\gamma^{\frac{1}{2}}t\rho)}{Q\rho^2+(\rho^2+1)/{}_2F_1[\frac{1}{2},\frac{1}{2};\,1;\,(\rho^2+1)^{-2}]}\frac{d\rho}{\rho}. \quad (55)$$

The problem of determining the zeros of $Q\rho^2+(\rho^2+1)/{}_2F_1[\frac{1}{2},\frac{1}{2};\,1;\,(\rho^2+1)^{-2}]$ has been considered in I. In the analysis, the hypergeometric function is replaced by the equivalent expression[18]

$$\begin{aligned}&{}_2F_1[\tfrac{1}{2},\tfrac{1}{2};\,1;\,(\rho^2+1)^{-2}]\\ &=\pi^{-1}\ln\frac{(\rho^2+1)^2}{\rho^2(\rho^2+2)}\,{}_2F_1\left[\tfrac{1}{2},\tfrac{1}{2};\,1;\,\frac{\rho^2(\rho^2+2)}{(\rho^2+1)^2}\right]\\ &\quad+2\pi^{-1}\sum_{n=0}^{\infty}\{\psi(n+1)-\psi(n+\tfrac{1}{2})\} \times\frac{(\frac{1}{2})_n(\frac{1}{2})_n}{(1)_n(1)_n}\left[\frac{\rho^2(\rho^2+2)}{(\rho^2+1)^2}\right]^n\\ &=\pi^{-1}\ln(8/\rho^2)+\cdots.\end{aligned} \quad (56)$$

As shown in I, there are two zeros which have the form $\rho_\pm=\epsilon e^{\pm(\frac{1}{4}\pi+\delta)i}$, where ϵ and δ are solutions of the pair of transcendental equations

$$Q\epsilon^2\cos2\delta\doteq\pi\ln(8/\epsilon^2)/\{[\ln(8/\epsilon^2)]^2+(\pi+2\delta)^2\} \quad (57)$$

and

$$Q\epsilon^2\sin2\delta\doteq\pi(\pi+2\delta)/\{[\ln(8/\epsilon^2)]^2+(\pi+2\delta)^2\}. \quad (58)$$

Pairs of values of ϵ and δ for different values of Q are given in Table I, which is reproduced from I. On using the theory of residues, the expression for $\int_0^t X[\mathbf{0},\sigma]d\sigma$ is

$$\int_0^t X[\mathbf{0},\sigma]d\sigma\doteq-\frac{\exp(-2\gamma^{\frac{1}{2}}\epsilon t\sin\delta)}{4\gamma\epsilon^2}\left\{\frac{\cos(2\gamma^{\frac{1}{2}}\epsilon t\cos\delta-2\delta)-Q\epsilon^2\pi^{-1}\cos(2\gamma^{\frac{1}{2}}\epsilon t\cos\delta-3\delta)}{1-2Q\epsilon^2\pi^{-1}\cos\delta+Q^2\epsilon^4\pi^{-2}}\right\} +\frac{Q(4\gamma)^{-1}}{2\pi i}\int_{L'}\frac{\exp(2\gamma^{\frac{1}{2}}t\rho)}{Q\rho^2+\pi[\log(8/\rho^2)]^{-1}}\frac{d\rho}{\rho}. \quad (59)$$

The corresponding expressions for $X[\mathbf{0},t]$ and $X_t[\mathbf{0},t]$ are

$$X[\mathbf{0},t]\doteq\frac{\exp(-2\gamma^{\frac{1}{2}}\epsilon t\sin\delta)}{2\gamma^{\frac{1}{2}}\epsilon}\left\{\frac{\sin(2\gamma^{\frac{1}{2}}\epsilon t\cos\delta-\delta)-Q\epsilon^2\pi^{-1}\sin(2\gamma^{\frac{1}{2}}\epsilon t\cos\delta-2\delta)}{1-2Q\epsilon^2\pi^{-1}\cos\delta+Q^2\epsilon^4\pi^{-2}}\right\} +\frac{Q\gamma^{-\frac{1}{2}}}{\pi}\int_0^1\frac{\exp(-2\gamma^{\frac{1}{2}}ty)dy}{[1+Qy^2\pi^{-1}\ln(8/y^2)]^2+4Q^2y^4} \quad (60)$$

and

$$X_t[\mathbf{0},t]\doteq\exp(-2\gamma^{\frac{1}{2}}\epsilon t\sin\delta)\left\{\frac{\cos(2\gamma^{\frac{1}{2}}\epsilon t\cos\delta)-Q\epsilon^2\pi^{-1}\cos(2\gamma^{\frac{1}{2}}\epsilon t\cos\delta-\delta)}{1-2Q\epsilon^2\pi^{-1}\cos\delta+Q^2\epsilon^4\pi^{-2}}\right\} -\frac{2Q}{\pi}\int_0^1\frac{y\exp(-2\gamma^{\frac{1}{2}}ty)dy}{[1+Qy^2\pi^{-1}\ln(8/y^2)]^2+4Q^2y^4}, \quad (61)$$

[18] A. Erdelyi, W. Magnus, F. Oberhettinger, and F. Tricomi, *Higher Transcendental Functions* (McGraw-Hill Book Company, Inc., New York, 1953), Vol. 1, p. 110.

where the line integrals L' around the cut GH in Eqs. (60) and (61) have been replaced by the differences of the line integrals on either side of GH. Equation (61) was obtained in Appendix C of I. The author has been unable to express the integrals in (59)–(61) in terms of known functions. For this reason, we obtain only estimates of their magnitude. First, note that the integral in Eq. (59) dominates the right-hand side for sufficiently large times; it is shown in Appendix B that the asymptotic formula for $\int_0^t X[\mathbf{0},\sigma]d\sigma$ is

$$\int_0^t X[\mathbf{0},\sigma]d\sigma \sim Q(2\pi\gamma)^{-1}\ln(2\gamma^{\frac{1}{2}}t). \tag{62}$$

The integrands in Eqs. (60) and (61) are positive. Consequently, the integrals are monotone decreasing functions of the time. The initial values of these integrals are

$$I_1(0)\doteq\frac{Q\gamma^{-\frac{1}{2}}}{\pi}\int_0^1\{[1-Qy^2\pi^{-1}\ln(8/y^2)]^2+4Q^2y^4\}^{-1}dy$$
$$\doteq(2\gamma^{\frac{1}{2}}\epsilon)^{-1}\left\{\frac{\sin\delta-Q\epsilon^2\pi^{-1}\sin 2\delta}{1-2Q\epsilon^2\pi^{-1}\cos\delta+Q^2\epsilon^4\pi^{-2}}\right\} \tag{63}$$

and

$$I_2(0)\doteq\frac{2Q}{\pi}\int_0^1\{[1-Qy^2\pi^{-1}\ln(8/y^2)]^2+4Q^2y^4\}y\,dy$$
$$\doteq-1+\frac{1-Q\epsilon^2\pi^{-1}\cos\delta}{1-2Q\epsilon^2\pi^{-1}\cos\delta+Q^2\epsilon^4\pi^{-2}}$$
$$\doteq\frac{Q\epsilon^2\pi^{-1}\cos\delta-Q^2\epsilon^4\pi^{-2}}{1-2Q\epsilon^2\pi^{-1}\cos\delta+Q^2\epsilon^4\pi^{-2}}. \tag{64}$$

On referring to Table I, it is seen that for $Q>10^4$, $Q\epsilon^2\pi^{-1}$ is small (<0.02). By retaining the leading terms of $I_1(0)$ and $I_2(0)$ in an expansion in powers of $Q\epsilon^2\pi^{-1}$, we have

$$I_1(0)\cong(2\gamma^{\frac{1}{2}}\epsilon)^{-1}\sin\delta \tag{65}$$

and

$$I_2(0)\cong Q\epsilon^2\pi^{-1}. \tag{66}$$

$I_1(t)$ in Eq. (60) and $I_2(t)$ in Eq. (61) are Laplace transforms. Their asymptotic values, for t sufficiently large, follow from Watson's lemma,[19]

$$I_1(t)\sim Q(2\gamma\pi t)^{-1} \tag{67}$$

and

$$I_2(t)\sim Q(2\gamma\pi t^2)^{-1}. \tag{68}$$

The contribution of the integral $I_2(t)$ to the value of $X_t[\mathbf{0},t]$, which is initially $-Q\epsilon^2\pi^{-1}$, decreases ultimately as in Eq. (68). Thus, there is an initial period, noted in I, in which the function $X_t[\mathbf{0},t]$ is approximately

$$X_t[\mathbf{0},t]\cong\exp(-2\gamma^{\frac{1}{2}}\epsilon t\sin\delta)\cos(2\gamma^{\frac{1}{2}}\epsilon t\cos\delta), \tag{69}$$

where δ is of the order of 0.1 or smaller. The estimate of the contribution of $I_1(t)$ to $X[\mathbf{0},t]$ is less precise,

$$X[\mathbf{0},t]\cong(2\gamma^{\frac{1}{2}}\epsilon)^{-1}\exp(-2\gamma^{\frac{1}{2}}\epsilon t\sin\delta)\sin(2\gamma^{\frac{1}{2}}\epsilon t\cos\delta-\delta)$$
$$+(2\gamma^{\frac{1}{2}}\epsilon)^{-1}\sin\delta,$$

where the initial contributions of the residues and of $I_1(0)$ compensate each other.

From the point of view of Problem 1, the heavy particle in a $2D$ lattice behaves as though it were bound to its equilibrium position, but it is seen that the displacement of the heavy particle in problem 2 increases asymptotically as $Q(2\pi\gamma)^{-1}\ln(2\gamma^{\frac{1}{2}}t)$. Clearly the same asymptotic type of result is to be expected for the continuum analog of Problem 2, i.e., the two-dimensional wave equation for an elastic membrane.

V. PROBABILITY DISTRIBUTION FUNCTIONS IN THE 1D, 3D, AND 2D LATTICES

In this section we will assemble the results obtained in Sec. III and IV.

A. 1D Lattice

On substituting the $1D$ values of ν_1, $\int_0^t X[\mathbf{0},\sigma]d\sigma$, $X[\mathbf{0},t]$, and $X_t[\mathbf{0},t]$, Eqs. (42)–(44), in the general expression for $P(v_1,u_1,t\,|\,v_0,u_0,0)$, Eq. (25), one obtains

$$P^{(1)}(v_1,u_1,t\,|\,v_0,u_0,0)=n\left\{\begin{pmatrix}v_1\\u_1\end{pmatrix}\middle|\begin{pmatrix}\langle v_1\rangle\\\langle u_1\rangle\end{pmatrix},\boldsymbol{\Sigma}\right\}, \tag{70}$$

with

$$\begin{pmatrix}\langle v_1\rangle\\\langle u_1\rangle\end{pmatrix}\doteq\begin{pmatrix}v_0e^{-bt}\\v_0b^{-1}[1-e^{-bt}]+u_0\end{pmatrix}, \tag{71}$$

and

$$\boldsymbol{\Sigma}\doteq\begin{pmatrix}kTM^{-1}[1-e^{-2bt}] & kTM^{-1}b^{-1}[1-e^{-bt}]^2\\kTM^{-1}b^{-1}[1-e^{-bt}]^2 & kTM^{-1}b^{-2}[2bt-3+4e^{-bt}-e^{-2bt}]\end{pmatrix}, \tag{72}$$

where $b=2\gamma^{\frac{1}{2}}Q^{-1}$. It is noteworthy that the conditional mean value $\langle v_1\rangle$ is independent of u_0, and is simply an exponential function of the time. Such behavior is characteristic of a free particle undergoing Brownian motion. In fact, the result in Eqs. (70)–(72) is identical with the conditional pdf obtained from the Kramers' equation[2,20] for a free Brownian particle, i.e., the fundamental solution of the equation

$$\partial P/\partial t=-v(\partial P/\partial u)+(\partial/\partial v)(bvP)+q(\partial^2P/\partial v^2), \tag{73}$$

[19] See A. Erdelyi, *Asymptotic Expansions* (Dover Publications, Inc., New York, 1956), p. 34.

[20] See Eqs. (279), (281), and (284)–(286) of reference 3.

corresponding to the initial condition $P(v,u,0)=\delta(v-v_0)\times\delta(u-u_0)$. Thus, *to within the accuracy of the equalities in Eqs.* (71) *and* (72), *the random thermal motion of a heavy particle in a* 1*D lattice at a temperature T is described by Kramers' equation for a free Brownian particle, Eq.* (73), in which the phenomenological constants b and q are, respectively, $2\gamma^{\frac{1}{2}}Q^{-1}$ and $kTM^{-1}b$. It is also possible to determine a diffusion coefficient for the heavy particle from the variance of the pdf $P(u_1,t\,|\,u_0,0)$, given in Eq. (24). For the 1*D* lattice, the variance is

$$\langle(u_1-\langle u_1\rangle)^2\rangle \doteq 2kTM^{-1}Q(4\gamma)^{-1}\{2\gamma^{\frac{1}{2}}t-Q[1-e^{-bt}]\}. \quad (74)$$

For times which are large compared to the velocity relaxation time b^{-1}, the phenomenological analog of Eq. (74) is

$$\langle(u_1-\langle u_1\rangle)^2\rangle = 2Dt;$$

and so the diffusion coefficient in Eq. (74),

$$D = kTM^{-1}Q(4\gamma)^{-1}2\gamma^{\frac{1}{2}} = kTm^{-1}(2\gamma^{\frac{1}{2}})^{-1},$$

is independent of the mass of the heavy particle and depends only on the properties of the medium with which the heavy particle interacts, as should be expected.

B. 3*D* Lattice

The value of ν_3 obtained in Sec. IV.A is

$$\nu_3 = 4\kappa(M\zeta_0)^{-1} = 4\gamma(Q\zeta_0)^{-1}.$$

On substituting this value of ν_3 and the results for $\int_0^t X[\mathbf{0},\sigma]d\sigma$, $X[\mathbf{0},t]$, and $X_t(\mathbf{0},t)$, Eqs. (52)–(54), in the general expression for $P(v_1,u_1,t\,|\,v_0,u_0,0)$, Eq. (25), one obtains

$$P^{(3)}(v_1,u_1,t\,|\,v_0,u_0,0) = \mathfrak{n}\left\{\begin{pmatrix} v_1 \\ u_1\end{pmatrix}\middle|\begin{pmatrix}\langle v_1\rangle \\ \langle u_1\rangle\end{pmatrix}, \boldsymbol{\Sigma}\right\}, \quad (75)$$

with

$$\begin{pmatrix}\langle v_1\rangle \\ \langle u_1\rangle\end{pmatrix} \doteq \begin{pmatrix} v_0e^{-\frac{1}{2}\beta t}[\cos\omega t-\beta(2\omega)^{-1}\sin\omega t]-u_0\omega e^{-\frac{1}{2}\beta t}\sin\omega t \\ v_0\omega^{-1}e^{-\frac{1}{2}\beta t}\sin\omega t+u_0e^{-\frac{1}{2}\beta t}[\cos\omega t+\beta(2\omega)^{-1}\sin\omega t]\end{pmatrix} \quad (76)$$

and

$$\boldsymbol{\Sigma} \doteq \begin{pmatrix} kTM^{-1}\{1-e^{-\beta t}[1-\beta\omega^{-1}\sin\omega t\cos\omega t]\} & kTM^{-1}\beta\omega^{-2}e^{-\beta t}\sin^2\omega t \\ kTM^{-1}\beta\omega^{-2}e^{-\beta t}\sin^2\omega t & kTM^{-1}Q\zeta_0(4\gamma)^{-1}\{1-e^{-\beta t}[1+\beta\omega^{-1}\sin\omega t\cos\omega t]\}\end{pmatrix}. \quad (77)$$

The remarkable feature of Eqs. (75)–(77) is that $P^{(3)}(v_1,u_1,t\,|\,v_0,u_0,0)$ is identical in form with the fundamental solution of Kramers' equation for an underdamped harmonic oscillator,[21]

$$(\partial P/\partial t) = -v(\partial P/\partial u)+(\partial/\partial v)[(\beta v+\omega^2 u)P] + q(\partial^2 P/\partial v^2). \quad (78)$$

Thus, *to within the accuracy of the equalities in* (76) *and* (77), *the random thermal motion of a heavy particle in a* 3*D lattice at a temperature T is described by Kramers' equation for a harmonically bound particle, Eq.* (78), in which the frequency ω and the phenomenological constants β and q are, respectively, $\omega=2\gamma^{\frac{1}{2}}(Q\zeta_0)^{-\frac{1}{2}}$, $\beta=4\gamma^{\frac{1}{2}}\zeta_0^{-2}\pi^{-1}Q^{-1}$, and $q=kTM^{-1}\beta$. [The effective force constant associated with the frequency ω is $4\kappa\zeta_0^{-1}$.] As $t\to\infty$, the mean and covariance matrices, (76) and (77), become

$$\begin{pmatrix}\langle v_1\rangle \\ \langle u_1\rangle\end{pmatrix} \to \begin{pmatrix}0\\0\end{pmatrix}$$

and

$$\boldsymbol{\Sigma} \to \begin{pmatrix} kTM^{-1} & 0 \\ 0 & kTM^{-1}Q\zeta_0(4\gamma)^{-1}\end{pmatrix}.$$

Thus,

$$P^{(3)}(v_1,u_1,\infty\,|\,v_0,u_0,0) = \mathfrak{n}\{v_1\,|\,0,\,kTM^{-1}\}\mathfrak{n}\{u_1\,|\,0,\,kT\zeta_0(4\kappa)^{-1}\}.$$

The position distribution $\mathfrak{n}\{u_1\,|\,0,\,kT\zeta_0(4\kappa)^{-1}\}$ has also been obtained by Montroll[22] in a study of the localizability of lattice particles.

C. 2*D* Lattice

The heavy particle in a 2*D* lattice exhibits properties intermediate between those of a free and a harmonically bound particle. For example, since $\nu_2=0$, the conditional mean velocity obtained from $P^{(2)}(v_1,u_1,t\,|\,v_0,u_0,0)$, Eq. (25), is independent of the initial position of the heavy particle,

$$\langle v_1\rangle = v_0X_t[\mathbf{0},t].$$

In addition, the conditional covariance, $\langle(u_1-\langle u_1\rangle)^2\rangle$, grows logarithmically with time,

$$\langle(u_1-\langle u_1\rangle)^2\rangle \sim kTM^{-1}Q(\pi\gamma)^{-1}\ln(2\gamma^{\frac{1}{2}}t) \sim kT(\pi\kappa)^{-1}\ln(2\gamma^{\frac{1}{2}}t).$$

This free particle behavior, however, is in apparent contradiction with the fact that $\langle v_1\rangle$, to a good approximation, is

$$\langle v_1\rangle \cong v_0\exp(-2\gamma^{\frac{1}{2}}\epsilon t\sin\delta)\cos(2\gamma^{\frac{1}{2}}\epsilon t\cos\delta).$$

[21] See Eqs. (55) of reference 4.

[22] E. W. Montroll, *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability* (University of California Press, Berkeley, California, 1956), Vol. 3. p. 209. A factor 2^{n-1} is omitted in Eq. (45) of this reference. When this correction is made, the agreement referred to in the text is obtained.

VI. *nD* LATTICE MODEL AND THEORY OF STOCHASTIC PROCESSES[23]

In this section we discuss the relationship between the *nD* lattice model of Brownian motion and the theory of stochastic processes. It has been assumed in our model that the initial positions and velocities of the lattice particles are normally distributed, Eq. (5). An important consequence of this assumption is that the position $u(t)$ and velocity $v(t)$ of particle **0** are stationary Gaussian random (sg) processes,[24] independent of the mass of particle **0** [see Eqs. (10), (11a)–(14a)]. The two functions together $\{v(t),u(t)\}$ constitute a two dimensional sg process. A question arises as to whether this two dimensional sg process is Markoffian.[25] An answer can be obtained by considering the correlation matrix $\mathbf{R}(t_j-t_i)$, $t_j>t_i$, which is formed from the covariances $\Sigma^{vv}{}_{ij}=\langle v(t_i)v(t_j)\rangle$, $\Sigma^{vu}{}_{ij}=\langle v(t_i)u(t_j)\rangle$, $\Sigma^{uv}{}_{ij}=\langle u(t_i)v(t_j)\rangle$, and $\Sigma^{uu}{}_{ij}=\langle u(t_i)u(t_j)\rangle$, Eqs. (11a)–(14a); it is

$$\mathbf{R}(t_j-t_i)=\begin{pmatrix}\langle v(t_i)v(t_j)\rangle/\langle v^2\rangle & \langle v(t_i)u(t_j)\rangle/\langle v^2\rangle^{\frac{1}{2}}\langle u^2\rangle^{\frac{1}{2}}\\ \langle u(t_i)v(t_j)\rangle/\langle v^2\rangle^{\frac{1}{2}}\langle u^2\rangle^{\frac{1}{2}} & \langle u(t_i)u(t_j)\rangle/\langle u^2\rangle\end{pmatrix}$$
$$=\begin{bmatrix}X_t[\mathbf{0},t_j-t_i] & \nu_n^{\frac{1}{2}}X[\mathbf{0},t_j-t_i]\\ -\nu_n^{\frac{1}{2}}X[\mathbf{0},t_j-t_i] & 1-\nu_n\int_0^{t_j-t_i}X[\mathbf{0},\sigma]d\sigma\end{bmatrix}, \quad (79)$$

where $\langle v^2\rangle=\langle v^2(t_i)\rangle=kTM^{-1}$, $\langle u^2\rangle=\langle u^2(t_i)\rangle=kTV^{-1}\omega_0$, and $\nu_n=(MV^{-1}\omega_0)^{-1}$. The elements of the correlation matrix Eq. (79) have been normalized so that $\mathbf{R}(0)=\mathbf{I}$, the unit matrix. A necessary and sufficient condition that the sg process $\{v(t),u(t)\}$ be Markoffian is that $\mathbf{R}(t)$ satisfy the condition[26]

$$\mathbf{R}(t_3-t_1)=\mathbf{R}(t_2-t_1)\mathbf{R}(t_3-t_2), \quad (80)$$

where $t_3\geqslant t_2\geqslant t_1$.

In the case of the $1D$ and $2D$ lattices (where $\nu_1=0$ and $\nu_2=0$), the correlation matrix $\mathbf{R}(t_j-t_i)$ has the simple diagonal form

$$\mathbf{R}(t_j-t_i)=\begin{pmatrix}X_t[\mathbf{0},t_j-t_i] & 0\\ 0 & 1\end{pmatrix}; \quad (81)$$

and the Markoffian condition, Eq. (80), is

$$\mathbf{R}(t_3-t_1)=\begin{pmatrix}X_t[\mathbf{0},t_3-t_2]X_t[\mathbf{0},t_2-t_1] & 0\\ 0 & 1\end{pmatrix}. \quad (82)$$

Thus in the case of the $1D$ and $2D$ lattices, the necessary and sufficient condition that the sg process $\{v(t),u(t)\}$ be Markoffian is that

$$X_t[\mathbf{0},t_3-t_1]=X_t[\mathbf{0},t_3-t_2]X_t[\mathbf{0},t_2-t_1]. \quad (83)$$

Since the only nontrivial solution of the functional equation (83) is $X_t[\mathbf{0},t]=e^{-\alpha t}$, and since the integral representation[27] for $X_t[\mathbf{0},t]$,

$$X_t[\mathbf{0},t]=\frac{Q+1}{2\pi i}\int_L\frac{p^{-1}e^{pt}dp}{Qp+(p\zeta[\mathbf{0},p])^{-1}}, \quad (84)$$

is not of the form $(1/2\pi i)\int_L e^{pt}(p+\alpha)^{-1}dp$, it is clear that the condition (83) is not satisfied exactly. However, in the $1D$ lattice we have the result for $Q\gg1$, Eq. (44), that $X_t[\mathbf{0},t]\doteq\exp(-2\gamma^{\frac{1}{2}}Q^{-1}t)$ and conclude that, *to the extent that the equality*

$$X_t[\mathbf{0},t_3-t_1]\doteq X_t[\mathbf{0},t_3-t_2]X_t(\mathbf{0},t_2-t_1]$$

can be regarded as exact, the sg process $\{v(t),u(t)\}$ is Markoffian for the $1D$ lattice. In the $2D$ lattice, when $Q\gg1$, we have from Eqs. (61) and (64)

$$X_t[\mathbf{0},t]\doteq\exp(-2\gamma^{\frac{1}{2}}\epsilon t\sin\delta)\times\left\{\frac{\cos(2\gamma^{\frac{1}{2}}\epsilon t\cos\delta)-Q\epsilon^2\pi^{-1}\cos(2\gamma^{\frac{1}{2}}\epsilon t\cos\delta-\delta)}{1-2Q\epsilon^2\pi^{-1}\cos\delta+Q^2\epsilon^4\pi^{-2}}\right\}-I_2(t)$$

where

$$0<I_2(t)<\frac{Q\epsilon^2\pi^{-1}\cos\delta-Q^2\epsilon^4\pi^{-2}}{1-2Q\epsilon^2\pi^{-1}\cos\delta+Q^2\epsilon^4\pi^{-2}}.$$

The upper bound on $I_2(t)$ is of the order of 0.015 or less for $Q>10^4$ (see Table I). Thus, it is clear that *the Markoffian condition* (83) *is not satisfied for a heavy particle in a 2D lattice.*

In the case of the $3D$ lattice, we have from Eqs. (79) and (52)–(54)

$$\mathbf{R}(t)\doteq\begin{pmatrix}e^{-\frac{1}{2}\beta t}[\cos\omega t-\beta(2\omega)^{-1}\sin\omega t] & e^{-\frac{1}{2}\beta t}\sin\omega t\\ -e^{-\frac{1}{2}\beta t}\sin\omega t & e^{-\frac{1}{2}\beta t}[\cos\omega t+\beta(2\omega)^{-1}\sin\omega t]\end{pmatrix}.$$

[23] The reader is referred to reference 4 and J. L. Doob, *Stochastic Processes* (John Wiley & Sons, Inc., New York, 1953), for successively more detailed discussions of stochastic processes.

[24] A random process $y(t)$ is stationary if the joint p.d.f. $W_r(y_0,t_0;\ y_1,t_1;\ \cdots;\ y_{r-1},t_{r-1})$ is independent of t_0 and depends only on the time differences t_i-t_j. A random process $y(t)$ is Gaussian if $W_r(y_0,t_0;\ \cdots;\ y_{r-1},t_{r-1})=\mathfrak{n}\{\mathbf{y}|\langle\mathbf{y}\rangle,\ \mathbf{\Sigma}\}$.

[25] A random process $y(t)$ is Markoffian if, for $t_n>t_{n-1}>\cdots>t_1$, $P(\mathbf{y}_n,t_n|\mathbf{y}_{n-1},t_{n-1};\ \cdots;\ \mathbf{y}_1,t_1)=P(\mathbf{y}_n,t_n|\mathbf{y}_{n-1},t_{n-1})$.

[26] J. L. Doob, Annals of Mathematical Statistics **15**, 229 (1944). See also Note II in the Appendix of reference 4 for a proof which follows naturally from Sec. III of this paper.

[27] The exact value of $X_t[\mathbf{0},t]$ is known for the $1D$ lattice when $Q=0$ ($M=m$) and $Q=1$ ($M=2m$). The values are, respectively, $J_0(2\gamma^{\frac{1}{2}}t)$ and $(\gamma^{\frac{1}{2}}t)^{-1}J_1(2\gamma^{\frac{1}{2}}t)$. The former result is an old one due first to Hamilton and later rediscovered several times, and the latter result was obtained in I. (See 1 for references.)

It can be verified that this matrix satisfies the Markoffian condition, Eq. (80), *in the same sense as in the* 1*D lattice, namely,*

$$\mathbf{R}(t_3-t_1)\doteq\mathbf{R}(t_2-t_1)\mathbf{R}(t_3-t_2).$$

Note that the results just obtained concerning the Markoffian character of $\{v(t),u(t)\}$ for the 1*D* and 3*D* lattices are not independent of the results in the last section involving Kramers' equation. For it has been pointed out by Wang and Uhlenbeck[4] that the random process $\{v(t),u(t)\}$ governed by Kramers' equation for the Brownian motion of a particle in a force field $K(u)$,

$$\partial P/\partial t=-v(\partial P/\partial u)+(\partial/\partial v)[(\beta v+K(u))P]+q(\partial^2 P/\partial v^2),$$

is, in general, Markoffian but not Gaussian. In the case of the linear Kramers' equations (73) for the 1*D* lattice and (78) for the 3*D* lattice, the random processes $\{v(t), u(t)\}$ are Markoffian *and* Gaussian. On the other hand, we have seen that in the case of the 2*D* lattice, the random process $\{v(t),u(t)\}$ is Gaussian but not Markoffian.

ACKNOWLEDGMENT

I wish to thank Dr. Marvin Zelen for bringing reference 5 to my attention.

APPENDIX A. IMPROVED ESTIMATE OF THE REMAINDER $\delta_1^{(0)}(t)$

Consider the integral

$$X_t[\mathbf{0},t]=\frac{Q+1}{2\pi i}\int_L\frac{\exp(2\gamma^{\frac{1}{2}}t\rho)d\rho}{Q\rho+[\rho^2+1]^{\frac{1}{2}}},$$

and construct the two cuts as shown in Fig. 2.

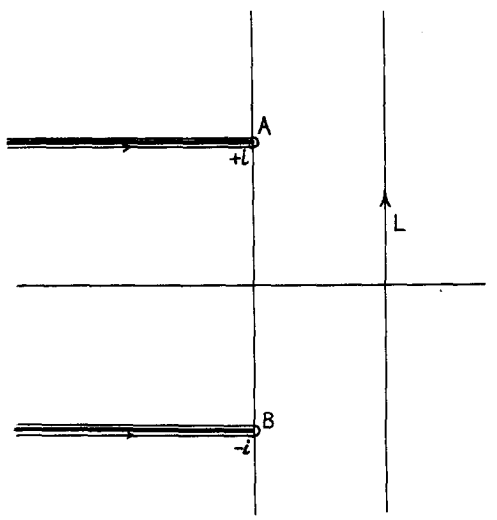

FIG. 2. The horizontal cuts start at A and B and extend to minus infinity.

The path of integration L can be shrunk so that $X_t[\mathbf{0},t]$ is the sum of the residue from the pole at $\rho\doteq-Q^{-1}$ and the contribution $\delta_1^{(0)}(t)$ from the line integrals around the two cuts,

$$\delta_1^{(0)}(t)=\frac{Q+1}{2\pi i}\Big\{-\int_\infty^0\frac{\exp[(-x+i)2\gamma^{\frac{1}{2}}t]dx}{Q(-x+i)+[2xe^{-\frac{1}{2}\pi i}-x^2]^{\frac{1}{2}}}-\int_0^\infty\frac{\exp[(-x+i)2\gamma^{\frac{1}{2}}t]dx}{Q(-x+i)+[2xe^{\frac{1}{2}\pi i}-x^2]^{\frac{1}{2}}}$$

$$-\int_0^\infty\frac{\exp[-(x+i)2\gamma^{\frac{1}{2}}t]dx}{-Q(x+i)+[2xe^{\frac{1}{2}\pi i}-x^2]^{\frac{1}{2}}}-\int_\infty^0\frac{\exp[-(x+i)2\gamma^{\frac{1}{2}}t]dx}{-Q(x+i)+[2xe^{\frac{1}{2}\pi i}-x^2]^{\frac{1}{2}}}\Big\}$$

$$=\frac{Q+1}{2\pi i}\Big\{\exp(2\gamma^{\frac{1}{2}}ti)\int_0^\infty dx\exp(-2\gamma^{\frac{1}{2}}tx)$$

$$\times\Big[\frac{(2xe^{\frac{1}{2}\pi i}-x^2)^{\frac{1}{2}}-(2xe^{-\frac{1}{2}\pi i}-x^2)^{\frac{1}{2}}}{Q^2(-x+i)^2+Q(-x+i)[(2xe^{-\frac{1}{2}\pi i}-x^2)^{\frac{1}{2}}+(2xe^{\frac{1}{2}\pi i}-x^2)^{\frac{1}{2}}]+(4x^2e^{\pi i}+x^4)^{\frac{1}{2}}}\Big]$$

$$+\exp(-2\gamma^{\frac{1}{2}}ti)\int_0^\infty dx\exp(-2\gamma^{\frac{1}{2}}tx)\Big[\frac{(2xe^{\frac{1}{2}\pi i}-x^2)^{\frac{1}{2}}-(2xe^{-\frac{1}{2}\pi i}-x^2)^{\frac{1}{2}}}{Q^2(x+i)^2-Q(x+i)[(2xe^{\frac{1}{2}\pi i}-x^2)^{\frac{1}{2}}+(2xe^{-\frac{1}{2}\pi i}-x^2)^{\frac{1}{2}}]+(4x^2e^{-\pi i}+x^4)^{\frac{1}{2}}}\Big]\Big\}. \quad \text{(A1)}$$

Each of the integrals in Eq. (A1) is a Laplace transform. Asymptotic series for these integrals can be obtained using Watson's lemma.[19] The leading terms in the asymptotic series are

$$\delta_1^{(0)}(t)\sim\frac{Q+1}{2\pi i}\Big\{\exp(2\gamma^{\frac{1}{2}}ti)[e^{3\pi i/4}-e^{-\pi i/4}]\frac{\sqrt{2}}{-Q^2}\frac{\Gamma(\frac{3}{2})}{(2\gamma^{\frac{1}{2}}t)^{\frac{3}{2}}}$$

$$+\exp(-2\gamma^{\frac{1}{2}}ti)[e^{\pi i/4}-e^{-3\pi i/4}]\frac{\sqrt{2}}{-Q^2}\frac{\Gamma(\frac{3}{2})}{(2\gamma^{\frac{1}{2}}t)^{\frac{3}{2}}}\Big\}$$

$$\sim-(2/\pi)^{\frac{1}{2}}Q^{-1}(2\gamma^{\frac{1}{2}}t)^{-\frac{3}{2}}\cos(2\gamma^{\frac{1}{2}}t+\pi/4).$$

The corresponding results for $\delta_1^{(1)}(t)$ and $\delta_1^{(2)}(t)$ are

$$\delta_1^{(1)}(t)\sim-(2\gamma^{\frac{1}{2}}Q)^{-1}(2/\pi)^{\frac{1}{2}}(2\gamma^{\frac{1}{2}}t)^{-\frac{3}{2}}\sin(2\gamma^{\frac{1}{2}}t+\pi/4)$$

and

$$\delta_1^{(2)}(t)\sim(4\gamma Q)^{-1}(2/\pi)^{\frac{1}{2}}(2\gamma^{\frac{1}{2}}t)^{-\frac{3}{2}}\cos(2\gamma^{\frac{1}{2}}t+\pi/4).$$

It is clear that even when $t\sim Q(2\gamma^{\frac{1}{2}})^{-1}$ and the exponential contributions from the residue are $\sim e^{-1}$, the contributions of $\delta_1^{(s)}(t)$ are considerably smaller.

APPENDIX B. ASYMPTOTIC BEHAVIOR OF $\int_0^t X[\mathbf{0},\sigma]d\sigma$ FOR LARGE t IN THE 2D LATTICE

In the integral representation for $\int_0^t X[\mathbf{0},\sigma]d\sigma$, Eq. (55), the dominant asymptotic behavior arises from the singularity ρ^{-1} at the logarithmic branch point $\rho=0$. For this reason, we add and subtract

$$-\frac{Q(4\gamma)^{-1}}{2\pi i}\int_L\pi^{-1}\rho^{-1}\exp(2\gamma^{\frac{1}{2}}t\rho)\ln(\rho^2/8)d\rho$$

in Eq. (55),

$$\int_0^t X[\mathbf{0},\sigma]d\sigma=-\frac{Q(2\pi\gamma)^{-1}}{2\pi i}\int_L\exp(2\gamma^{\frac{1}{2}}t\rho)[\ln\rho-\ln 2^{\frac{1}{2}}]d\rho/\rho$$

$$+\frac{Q(4\gamma)^{-1}}{2\pi i}\int_L\exp(2\gamma^{\frac{1}{2}}t\rho)$$

$$\times\Big\{\frac{{}_2F_1[\frac{1}{2},\frac{1}{2};1;(\rho^2+1)^{-2}]}{Q\rho^2\,{}_2F_1[\frac{1}{2},\frac{1}{2};1;(\rho^2+1)^{-2}]+\rho^2+1}$$

$$+\pi^{-1}\ln(\rho^2/8)\Big\}\frac{d\rho}{\rho}. \quad \text{(B1)}$$

The second integral in Eq. (B1) can be treated as before, except that the cut along the negative real axis from H to G must be extended to minus infinity. Deforming the path L so that it follows the cut, it can be seen that there is no longer a ρ^{-1} singularity at $\rho=0$. Further, it can be shown that the contributions from the line integrals around the cut and the residues of the integrand are bounded functions of the time. Now consider the first integral in Eq. (B1). Its value is[28]

$$-\frac{Q(2\pi\gamma)^{-1}}{2\pi i}\int_L\exp(2\gamma^{\frac{1}{2}}t\rho)[\ln\rho-\ln 2^{\frac{1}{2}}]d\rho/\rho$$

$$=Q(2\pi\gamma)^{-1}\{\ln(2\gamma^{\frac{1}{2}}t)+\ln(\gamma_e 2^{\frac{1}{2}})\},$$

where γ_e is the Euler-Mascheroni constant. Thus the asymptotic formula for $\int_0^t X[\mathbf{0},\sigma]d\sigma$ is

$$\int_0^t X[\mathbf{0},\sigma]d\sigma\sim Q(2\pi\gamma)^{-1}\ln(2\gamma^{\frac{1}{2}}t).$$

[28] A. Erdelyi, W. Magnus, F. Oberhettinger, and F. Tricomi, *Tables of Integral Transforms* (McGraw-Hill Book Company, Inc., New York, 1954), Vol. 1, p. 250.

Equipartition of Energy for Nonlinear Systems

JOSEPH FORD*
Physics Department, University of Miami, Coral Gables, Florida
(Received January 11, 1961)

A system of harmonic oscillators weakly coupled by nonlinear forces will not achieve equipartition of energy as long as the uncoupled frequencies ω_k are linearly independent on the integers, i.e., as long as there is no collection of integers $\{n_k\}$ for which $\Sigma n_k\omega_k=0$ other than all $n_k=0$. This result is shown to follow from the general form of the Kryloff and Bogoliuboff series solution to the equations of motion. Physically, the linear independence of the uncoupled frequencies means that none of the interacting oscillators drives another at its resonant frequency, and this lack of internal resonance precludes appreciable energy sharing in the limit as the coupling tends to zero. It is shown that the lack of equipartition of energy observed by Ulam, Fermi, and Pasta for certain nonlinear systems may be explained in terms of the preceding remarks. Moreover, a Kryloff and Bogoliuboff series solution to the appropriate equations of motion is shown to yield qualitative agreement with the Ulam, Fermi, and Pasta computer solution. Finally, a particular system of linear differential equations is solved which illustrates a mechanism whereby oscillator systems may achieve equipartition of energy.

I. INTRODUCTION

IT is generally believed that when a large number of independent linear harmonic oscillators are suitably coupled by nonlinear forces, an approach to equilibrium will ensue.[1] For a finite number of oscillators, deviations from equilibrium are expected[2] and the system will periodically return to its initial conditions because of Poincaré recurrences. However, when the number of oscillators is large, the Poincaré period is presumed large[3] and deviation from equilibrium rare.[2] Ulam, Fermi, and Pasta[4] tried to illustrate the expected approach to equilibrium by observing the equipartition of energy among normal modes for a system of one-dimensional oscillators obeying equations of the type

$$\ddot{y}_i=(y_{i+1}-2y_i+y_{i-1})+\alpha[(y_{i+1}-y_i)^2-(y_i-y_{i-1})^2], \quad (1)$$

where i runs over the positive integers from unity to as high as 64. Here α is to be chosen sufficiently small that the nonlinear terms can be treated as a small perturbation. Using MANIAC I at Los Alamos, they numerically solved Eq. (1) and then calculated the energy in each normal mode as a function of time. These calculations revealed no tendency toward equipartition of energy among the normal modes. Various initial conditions were used and cubic and broken linear as well as quadratic couplings were considered without materially changing the outcome. Energy was shared by only a few modes in a periodic fashion reminiscent of the energy sharing between two vertical pendulums hanging from the same horizontal string. A plot of the normal mode energies as a function of time for a typical case appears in Fig. 2 .

The results of the computer calculations of UFP thus appear to be in sharp contrast, if not in actual contradiction, to the widely held notions concerning the approach to equilibrium. Strictly speaking, there is no contradiction here because the theories on this subject usually attempt such generality that they make no effort to categorize the specific types of nonlinear forces needed to cause an approach to equilibrium. Thus, one could argue that the simple algebraic nonlinearity of Eq. (1) is not of the appropriate type. However, since the nonlinear force that does give rise to equilibrium might be expected to possess a power series expansion valid for small displacements, the nonlinear coupling of Eq. (1) should be adequate to illustrate the approach to equilibrium. Consequently, aside from any element of contradiction, a number of physicists have been puzzled by the failure of Eq. (1) to lead to equilibrium behavior.

This paper provides an explanation for the non-equilibrium behavior of Eq. (1) using a perturbation technique due to Kryloff and Bogoliuboff.[5,6] Section II states the basic reason for the failure of the UFP solution and then illustrates the mathematical techniques to be used by applying them to an interesting linear problem which can be solved exactly. Section III then solves the nonlinear equations obtaining qualitative agreement with the computer solution. The final section summarizes the main conclusion of the paper by illustrating how they follow from the general form of the Kryloff and Bogoliuboff series solution.

II. HARMONIC CHAIN

The reason why the UFP oscillator system fails to approach equilibrium can be explained by such a simple physical argument that it is worth exposing immediately unincrusted by mathematical detail. This is particularly cogent in view of the fact that the Kryloff-Bogoliuboff technique is subject to question. Equation (1) can be

* Present address: Physics Department, Johns Hopkins University, Baltimore 18, Maryland.

[1] For a discussion of recent work on this problem and for references to past work, see I. Prigogene, *Proceedings of the International Symposium on Transport Processes in Statistical Mechanics* (Interscience Publishers, Inc., New York, 1958).

[2] P. Mazur and E. Montroll, J. Math. Phys. 1, 70 (1960).

[3] See, for example, the discussion by J. A. McLennan, Jr., Phys. Fluids 2, 92 (1959), and his accompanying bibliography.

[4] S. Ulam, E. Fermi, and J. Pasta, "Studies of nonlinear problems I," Los Alamos Sci. Lab. Rept. LA-1940 (1955), hereafter referred to as UFP.

[5] N. Kryloff and N. Bogoliuboff, *Introduction to Nonlinear Mechanics* (Princeton University Press, Princeton, New Jersey, 1947).

[6] N. Minorsky, *Introduction to Nonlinear Mechanics* (Edwards Brothers, Inc., Ann Arbor, Michigan, 1947).

transformed to normal mode coordinates x_k via the transformation

$$y_l = (2/N)^{\frac{1}{2}} \sum_{k=1}^{(N-1)} x_k \sin(kl\pi/N) \tag{2}$$

to read

$$\ddot{x}_k = -\omega_k^2 x_k + \alpha \sum_{r,s=1}^{(N-1)} A_{krs} x_r x_s, \tag{3}$$

where $k=1, \cdots, (N-1)$, where the A_{krs} are determinable constants, and where

$$\omega_k = 2 \sin(k\pi/2N), \tag{4}$$

provided $y_0 \equiv y_N \equiv 0$.

For purposes of discussion, we may suppose that the normal coordinates x_k describe the small amplitude motion of a system of pendulums coupled by nonlinear springs. We would like to determine whether or not the pendulums described by Eq. (3) can exchange energy. In order to do this, first consider just two pendulums coupled by a linear spring. Here appreciable energy exchange will occur only when the lengths of the pendulums are very nearly equal, i.e., when $(\omega_1 - \omega_2) \approx 0$. Stated another way, appreciable energy sharing occurs only when one pendulum can drive the other pendulum at its resonant frequency. Kryloff and Bogoliuboff refer to this type of resonance as internal resonance. When the coupling between the two pendulums is nonlinear, one must expect a Fourier analysis of the motion to reveal the presence of sum and difference frequencies $(n_1\omega_1 \pm n_2\omega_2)$. Consequently, in the presence of nonlinear forces, we must expect internal resonance and energy sharing when $(n_1\omega_1 + n_2\omega_2) \approx 0$, where n_1 and n_2 are integers which may be positive or negative. Thus, the only effect of nonlinearity in regard to energy sharing is to generalize the condition for internal resonance from $(\omega_1 - \omega_2) \approx 0$ to $(n_1\omega_1 + n_2\omega_2) \approx 0$. The question of how closely $(n_1\omega_1 + n_2\omega_2)$ must approximate zero depends on the strength of the coupling α between pendulums. As $\alpha \to 0$, the frequency sums must also tend to zero. In what follows, we generally assume that α is very small and, therefore, we write the resonance condition as an equality.

On returning to Eq. (3), we now expect appreciable energy sharing only if the $\{\omega_k\}$ are such as to cause internal resonance, i.e., only if

$$\sum_{k=1}^{(N-1)} n_k \omega_k = 0 \tag{5}$$

for some nonzero collection of integers $\{n_k\}$. Numbers ω_k satisfying Eq. (5) for nontrivial $\{n_k\}$ are said to be linearly dependent on the integers. The question of energy sharing for Eq. (3) then turns on whether or not the frequencies ω_k of Eq. (4) are linearly dependent on the integers. Hemmer[7] has shown that the frequencies ω_k of Eq. (4) are linearly dependent except for N equal to a prime or a power of 2. The only cases considered by UFP were for N equal a power of 2. Thus, there are no Eq. (5) dependency relations for the UFP chains and hence no appreciable energy sharing. The austerity of this conclusion must be altered slightly since the actual values of α used by UFP did not approximate $\alpha \approx 0$; therefore, the equal sign of Eq. (5) must be replaced by the "approximately equal to" sign and some energy sharing occurs. From Eq. (4) we have that $\omega_1 \approx \omega_2/2 \approx \omega_3/3 \approx \omega_4/4 \ldots$, where the approximation becomes poorer as one reads to the right. Thus, as the value of α is increased from zero, one would expect appreciable energy sharing between higher and higher modes. UFP happened to choose a value of α such that appreciable energy sharing occured among the first few modes.

These heuristic conclusions are supported in detail by the Kryloff-Bogoliuboff calculations of the following two sections. From the standpoint of mathematical rigor, however, it is worthwhile to note that these conclusions also find support in the work of Balescu[8] who shows that the nonlinear Eqs. (3) do not possess any analytic constants of the motion other than the total energy provided, among other things, that the ω_k are linearly dependent. Consequently, when the ω_k are linearly independent, as is true for the UFP chain, we must anticipate the possibility of finding analytic constants of the motion other than the total energy; and in particular, one must anticipate the possibility that each normal mode energy E_k generates an analytic constant of the motion ϕ_k of Eq. (3) given by

$$\phi_k = E_k(t) + \sum_{n=1}^{\infty} \alpha^n \phi_k^{(n)}(t), \tag{6}$$

such that, for $\alpha \ll 1$, $E_k(t) \approx \phi_k$. In what follows, we confine our attention to the Kryloff and Bogoliuboff approach. We hope to discuss the full application of Balescu's work to these oscillator systems in another place.

Since we have now made it amply clear that there is, in general, no energy sharing without internal resonance, we might be expected in some later section to solve a nonlinear system of equations possessing internal resonance and show that equipartition of energy obtains. Unfortunately, except for the case of linear forces, the Kryloff and Bogoliuboff method becomes unmanageable in the presence of internal resonance; consequently, this task must be left for subsequent work on a computer. In view of these analytical difficulties, it is worthwhile to ask whether or not a system of pendulums weakly coupled by linear springs can provide any insight into the question of equipartition of energy. We now investigate such a linear system. In order to provide maximum internal resonance, we require that all the ω_k be equal. If we further assume

[7] P. Chr. Hemmer, Dynamic and stochastic types of motion in the linear chain, thesis, Trondheim, 1959.

[8] R. Balescu, Bull. acad. roy. Belg. Cl. Sci. **42**, 622 (1956).

nearest neighbor coupling, the differential equations governing the system may be written

$$\ddot{x}_k = -\omega^2 x_k + \alpha(x_{k+1} + x_{k-1}), \tag{7}$$

where again $k=1, \cdots, (N-1)$ and where $x_0 \equiv x_N \equiv 0$. Here the length of each pendulum has been adjusted so that the coefficient of each y_k is $(-\omega^2)$. Equation (7), being linear, cannot exhibit ergodic behavior; nonetheless, for selected initial conditions, we might observe equipartition of energy among the x_k degrees of freedom because of internal resonance.

We now compute the solution to Eq. (7) by using the Kryloff and Bogoliuboff technique. That the method leads to a correct solution for this case can be verified by exact integration of Eq. (7). We begin by observing that for $\alpha=0$, Eq. (7) has the solution

$$x_k = A_k \cos(\omega t + \varphi_k), \tag{8}$$

where A_k and φ_k are constants. If we now seek a power series solution to Eq. (7) of the form

$$x_k = A_k \cos(\omega_k t + \varphi_k) + \sum_{l=1}^{\infty} \alpha^l x_{kl}, \tag{9}$$

we encounter secular terms of the form

$$x_k \sim t \sin(\omega t + \varphi_k).$$

Since we are interested in using a truncated form of series (9), such terms are inadmissible. x_k is known to be a bounded, multiple-periodic function, while these secular terms are unbounded and aperiodic. Kryloff and Bogoliuboff observe that these secular terms arise because of a need to shift the frequency and vary the amplitude of the zeroth order terms in solution (9). In order to anticipate this amplitude and frequency modulation from the onset, Kryloff and Bogoliuboff suggest that one seek a power series solution of the form

$$x_k = A_k \cos\tau_k + \sum_{l=1}^{\infty} \alpha^l x_{kl}(A_1, \cdots, A_N, \tau_1, \cdots, \tau_N), \tag{10a}$$

where

$$\dot{A}_k = \sum_{l=1}^{\infty} \alpha^l A_{kl}(A_1, \cdots \tau_N) \tag{10b}$$

and

$$\dot{\tau}_k = \omega + \sum_{l=1}^{\infty} \alpha^l \omega_{kl}(A_1, \cdots, \tau_N). \tag{10c}$$

Here the x_{kl} are required to be periodic functions of each τ_j of period 2π. The series (10a) is to be put into Eq. (7). Equating the coefficient of each power of α to zero leads to a system of differential equations which formally can be solved sequentially for the x_{kl} as functions of the A_l and τ_l. The A_{kl} and ω_{kl} of Eqs. (10b) and (10c) are determined as functions of the A_l and τ_l at each step such that they eliminate secular terms. The resulting coupled differential Eqs. (10b) and (10c) can then be solved, at least for the linear case, yielding A_k and τ_k as functions of time. Carrying out this calculation for Eq. (7) requires some straightforward, but laborious, algebraic and trigonometric manipulations which we omit; to first order in α, the solution is found to be

$$x_k = \left[\frac{2}{N}\right] \sum_{r,s=1}^{(N-1)} \sin\left(\frac{kr\pi}{N}\right) \sin\left(\frac{rs\pi}{N}\right) \times [x_s(0) \cos\Omega_r t + \omega^{-1}\dot{x}_s(0) \sin\Omega_r t], \tag{11a}$$

where

$$\Omega_r = \omega - \omega^{-1}\alpha \cos(r\pi/N), \tag{11b}$$

and where $x_s(0)$ and $\dot{x}_s(0)$ are the initial position and speed of the sth pendulum.

Having found the solution to Eq. (7), we may now determine if there is some initial state which leads to equipartition. At $t=0$, let all $\dot{x}_k(0)$ be zero and let all $x_k(0)=0$ except for $x_1(0)=1$; thus, all the energy is initially possessed by the end pendulum. Equations (11a) and (11b) may then be written

$$x_k = (2/N) \sum_{r=1}^{(N-1)} \sin(r\pi/N) \sin(kr\pi/N) \times \cos(\omega - \beta_r)t, \tag{12a}$$

where

$$\beta_r = \omega^{-1}\alpha \cos(r\pi/N). \tag{12b}$$

It is obvious that Eq. (12a) satisfies the initial conditions if one verifies that

$$\sum_{l=1}^{(N-1)} \sin(lr\pi/N) \sin(ls\pi/N) = (N/2)\delta_{rs}, \tag{13}$$

where δ_{rs} is the Kronecker δ. For convenience, we assume that $(N-1)$ is even. Equation (12a) then becomes for k odd

$$x_k = 4N^{-1} \cos\omega t \sum_{r=1}^{(N-1)/2} \sin(r\pi/N) \sin(kr\pi/N) \cos\beta_r t, \tag{14a}$$

and for k even

$$x_k = 4N^{-1} \sin\omega t \sum_{r=1}^{(N-1)/2} \sin(r\pi/N) \sin(r\pi k/N) \sin\beta_r t. \tag{14b}$$

Each pendulum oscillates with a frequency equal to its uncoupled frequency and with an amplitude which consists of a sum of amplitude modulations, each modulating at a distinct frequency. The total energy E_T of the pendulum system divided by $\omega^2/2$ equals unity. Let E_i be the energy of the ith pendulum given by

$$E_k = 2^{-1}[\dot{x}_k^2 + \omega^2 x_k^2] \tag{15}$$

to zeroth order in α. Putting Eqs. (14a) and (14b) into Eq. (15) yields

$$(2E_k/\omega^2)=N^{-1}+\{\text{Multiple-periodic time varying terms}\} \quad (16a)$$

for $i=2, 3, \cdots, N-2$. For either end particle, we find

$$(2E_{\text{end}}/\omega^2)=N^{-1}(\tfrac{3}{2})+\{\text{Multiple-periodic time varying terms}\}. \quad (16b)$$

By averaging Eqs. (16a) and (16b) over a near period, we see that equipartition certainly occurs on the time average. By itself, however, this is a rather weak result since it does not preclude, for instance, each pendulum in turn, having all the energy for $1/N$th of the time. However, when the number $(N-1)$ of oscillators is large, we evidently have the much stronger result that each oscillator almost always possesses $1/N$th of the total energy. Moreover, this result is valid for a single system starting from a definite state and there is no need to average the motion over some ensemble of initial states as has been required in the past.[2] These points can be understood heuristically in the following way. The multiple-periodic terms in Equation (16a) consist of several sums S which can be brought to the general form

$$S=\sum_{l=1}^{\sim N} B_l \cos[t\cos(l\pi/N)], \quad (17)$$

where each B_l is a real number. Since most of the numbers $\cos(l\pi/N)$ are linearly independent on the integers, the sum S may be thought of as a sum of random variables similar to that appearing in the problem of a random walk. Consequently, as $N\to\infty$, we expect the probability distribution density for values of S to be a sharply peaked Gaussian with values of S at or near zero being overwhelmingly probable.

A detailed proof of these points, modeled on the arguments presented in Appendix III of the paper by Mazur and Montroll,[2] will be given elsewhere along with a more detailed consideration of this particular linear system. Our main purpose here is to try to present some insight into how internal resonance can bring about equipartition of energy, and the multiple-periodic amplitude modulation of Eqs. (14a) and (14b) certainly illustrates a possible mechanism for providing equipartition.

III. ANHARMONIC CHAIN

We now apply the Kryloff and Bogoliuboff method to obtain a solution to Eq. (3) which may be compared with the UFP results. Equation (3) contains constants A_{krs} not explicitly defined. Consequently, we now write out the normal mode differential equations in complete detail.

$$\ddot{x}_k=-\omega_k{}^2x_k+\frac{2^{\frac{3}{2}}\alpha}{N^{\frac{1}{2}}}\Bigg\{\sum_{r=N-(k-1)}^{N-1} x_rx_{2N-r-k}\sin\left[\frac{\pi(r+k)}{N}\right]\times\left\{-1+\cos\left[\frac{\pi r}{N}\right]\right\}+\sum_{r=1}^{N-k-1} x_rx_{r+k}\left\{\sin\left[\frac{\pi(r+k)}{N}\right]-\sin\left[\frac{r\pi}{N}\right]-\sin\left[\frac{k\pi}{N}\right]\right\}+\sum_{r=1}^{k-1} x_rx_{k-r}\sin\left[\frac{\pi(k-r)}{N}\right]\left\{-1+\cos\left[\frac{r\pi}{N}\right]\right\}\Bigg\}, \quad (18)$$

where $k=1, 2, \cdots, N-1$, and where

$$\omega_k=2\sin(k\pi/2N). \quad (19)$$

Equations (18) need be solved only for values of $N=2^k$, where k is a positive integer, for which the ω_k are linearly independent on the integers. We thus seek a Kryloff and Bogoliuboff solution of the form

$$x_k=A_k\sin\tau_k+\sum_{l=1}^{\infty}\alpha^l x_{kl}(\tau_1,\cdots,\tau_{N-1}), \quad (20a)$$

where

$$\tau_k=\Omega_k t+\phi_k \quad (20b)$$

and

$$\Omega_k=\omega_k+\sum_{l=1}^{\infty}\alpha^l\Omega_{kl}. \quad (20c)$$

The x_{kl} are required to be periodic functions of period 2π in each variable τ_j. The Ω_{kl} are constants independent of time determined so as to eliminate secular terms. The A_k and ϕ_k are constants determined by the initial conditions. On comparing Eqs. (20) with Eqs. (10), one sees that Eq. (20) retain the capacity for frequency modulation while discarding the mechanism for amplitude modulation. The lack of amplitude modulation, of course, means that one does not expect any appreciable energy sharing between normal modes. For large N and small k, however, we see from Eq. (19) that

$$\omega_1\approx\omega_2/2\approx\omega_3/3\approx\text{etc.};$$

hence, for these values of N and k, the ω_k are "almost" linearly dependent and, strictly speaking, we should at least allow for the possibility of some internal resonance, i.e., amplitude modulation. To allow for amplitude modulation, however, the Kryloff and Bogoliuboff method requires that one solve coupled nonlinear equations at least as difficult as Eqs. (18) themselves. We are, therefore, reduced to finding solution (20) while bearing in mind its limitations.

On putting Eqs. (20) into Eqs. (18), we find after some laborious algebraic and trigonometric manipula-

tions that to first order in α

$$x_{2l+1}=A_{2l+1}\sin\tau_{2l+1}+\frac{\alpha}{(2N)^{\frac{1}{2}}}\Bigg\{\sum_{r=N-l}^{N-1}A_rA_{(2N-r-2l-1)}\times\Bigg\{\cot\left[\frac{(2l+1)\pi}{4N}\right]\cos[\tau_r-\tau_{(2N-r-2l-1)}]+\tan\left[\frac{(2l+1)}{4N}\right]\cos[\tau_r+\tau_{(2N-r-2l-1)}]\Bigg\}+\sum_{r=1}^{N-2l-2}A_rA_{r+2l+1}\Bigg\{-\cot\left[\frac{(2l+2)\pi}{4N}\right]\times\cos[\tau_r-\tau_{r+2l+1}]-\tan\left[\frac{(2l+1)}{4N}\right]\times\cos[\tau_r+\tau_{r+2l+1}]\Bigg\}+\sum_{r=1}^{l}A_rA_{2l+1-r}\times\Bigg\{-\tan\left[\frac{(2l+1)\pi}{4N}\right]\cos[\tau_r-\tau_{2l+1-r}]-\cot\left[\frac{(2l+1)\pi}{4N}\right]\cos[\tau_r+\tau_{2l+1-r}]\Bigg\}\Bigg\}, \quad (21a)$$

and

$$x_{2l}=A_{2l}\sin\tau_{2l}+\frac{\alpha}{(2N)^{\frac{1}{2}}}\Bigg\{A_{N-l}^2\Bigg\{\tfrac{1}{2}\cot\left[\frac{l\pi}{2N}\right]+\tfrac{1}{2}\tan\left[\frac{l\pi}{2N}\right]\cos[2\tau_{N-l}]\Bigg\}+\sum_{r=N-l+1}^{N-1}A_rA_{(2N-r-2l)}\times\Bigg\{\cot\left[\frac{l\pi}{2N}\right]\cos[\tau_r-\tau_{(2N-r-2l)}]+\tan\left[\frac{l\pi}{2N}\right]\cos[\tau_r+\tau_{(2N-r-2l)}]\Bigg\}+\sum_{r=1}^{N-2l-1}A_rA_{r+2l}\Bigg\{-\cot\left[\frac{l\pi}{2N}\right]\cos[\tau_r-\tau_{r+2l}]-\tan\left[\frac{l\pi}{2N}\right]\cos[\tau_r+\tau_{2l+r}]\Bigg\}+\sum_{r=1}^{l}A_rA_{2l-r}\times\Bigg\{-\tan\left[\frac{l\pi}{2N}\right]\cos[\tau_r-\tau_{2l-r}]-\cot\left[\frac{l\pi}{2N}\right]\times\cos[\tau_r+\tau_{2l-r}]\Bigg\}+A_l^2\Bigg\{-\tfrac{1}{2}\tan\left[\frac{l\pi}{2N}\right]-\tfrac{1}{2}\cot\left[\frac{l\pi}{2N}\right]\cos[2\tau_l]\Bigg\}\Bigg\}. \quad (21b)$$

Here $\tau_k=\omega_k t+\varphi_k$ since $\Omega_k=\omega_k$ to first order in α.

Let us now specialize solutions (21a) and (21b) to the case $N=32$, $\alpha=\frac{1}{4}$, and all $x_k(0)$ and $\dot{x}_k(0)$ equal to zero except for $x_1(0)$. These, of course, are the conditions for one of the solutions obtained by UFP. Since solutions (21a) and (21b) are expected to be valid only when relatively little energy finds its way into the higher modes, we neglect all x_k and A_k for $k\geqq 6$. Under these conditions, Eqs. (21a) and (21b) take the form

$$x_{2l+1}=A_{2l+1}\cos[\omega_{2l+1}t]+\tfrac{1}{32}\cot\left[\frac{(2l+1)\pi}{128}\right]\times\Bigg\{-\sum_{r=1}^{5}A_rA_{r+2l+1}\cos[(\omega_r-\omega_{r+2l+1})t]+\sum_{r=1}^{l}A_rA_{2l+1-r}\cos[(\omega_r+\omega_{2l+1-r})t]\Bigg\}, \quad (22a)$$

$$x_{2l}=A_{2l}\cos[\omega_{2l}t]+\tfrac{1}{32}\cot\left[\frac{l\pi}{64}\right]\Bigg\{-\sum_{r=1}^{4}A_rA_{r+2l}\times\cos[(\omega_r-\omega_{r+2l})t]+\sum_{r=1}^{l-1}A_rA_{2l-r}\times\cos[(\omega_r+\omega_{2l-r})t]+\frac{A_l^2}{2}\cos[2\omega_l t]\Bigg\}, \quad (22b)$$

where $l\leqq 2$. Here we have neglected the tan terms of Eqs. (21a) and (21b) since they are small relative to the cot terms, and we have deferred until later the problem of determining the A_k in terms of $x_l(0)$. We see that Eqs. (22a) and (22b) already satisfy $\dot{x}_l(0)=0$.

Before evaluating the A_k, we show that solutions (22a) and (22b) exhibit a relatively short-term periodic behavior independent of the value of the A_k. To illustrate this behavior, we calculate the energy in the first normal mode as given following UFP by

$$E_1=\dot{x}_1^2/2+(\omega_1^2x_1^2)/2. \quad (23)$$

(This expression neglects the energy in the coupling which is proportional to α and higher powers of α; this neglect is no doubt justified if the resulting normal mode energies always sum to the correct total energy.) The calculation for the energy in the higher modes reveals the same short-term periodicity, and, therefore, we perform the calculation only for E_1. Putting Eq. (22a) into Eq. (23) yields, to first order in α,

$$(2E_1/\omega_1^2)=A_1^2-\tfrac{5}{2}\{A_1^2A_2\cos[(2\omega_1-\omega_2)t]+A_1A_2A_3\cos[(\omega_1+\omega_2-\omega_3)t]+A_1A_2A_4\cos[(\omega_1+\omega_3-\omega_4)t]+A_1A_4A_5\cos[(\omega_1+\omega_4-\omega_5)t]+A_1A_5A_6\cos[(\omega_1+\omega_5-\omega_6)t], \quad (24)$$

where we set $(\omega_2-\omega_1)=\omega_1$, $(\omega_4-\omega_5)=\omega_1$, etc., when such terms appear in an amplitude factor. Now from Eq.

(19), we have

$$(2\omega_1-\omega_2)=4\sin(\pi/64)-2\sin(\pi/32). \quad (25)$$

Expanding each sin term in Eq. (25) to third order in its argument. we have

$$(2\omega_1-\omega_2)=2(\pi/64)^3. \quad (26)$$

Making the same calculation for each frequency factor shows that each, to third order, is an integral multiple of $2(\pi/64)^3$. Equation (24) thus may be written

$$(2E_1/\omega_1^2)=A_1^2-\tfrac{5}{2}\{A_1^2A_2\cos\Omega t+A_1A_2A_3\cos3\Omega t+A_1A_3A_4\cos6\Omega t+A_1A_4A_5\cos10\Omega t+A_1A_5A_6\cos15\Omega t\}, \quad (27)$$

where $\Omega=2(\pi/64)^3$. Moreover, it is easy to show that this periodicity holds also for the E_1 terms second order in α where frequency factors like $[2\omega_2-\omega_1-\omega_3]=2\Omega$ arise.

We now must evaluate the A_k of Eqs. (22a) and (22b) for the case in which all $x_k(0)$ equal zero except $x_1(0)$. Setting $t=0$ in Eqs. (22a) and (22b), one finds

$$x_{2l+1}(0)=A_{2l+1}+\tfrac{1}{32}\cot[(2l+1)\pi/128]\times\left\{-\sum_{r=1}^{5}A_rA_{r+2l+1}+\sum_{r=1}^{l}A_rA_{2l+1-r}\right\}, \quad (28a)$$

and

$$x_{2l}(0)=A_{2l}+\tfrac{1}{32}\cot[l\pi/64]\times\left\{-\sum_{r=1}^{5}A_rA_{r+2l}+\sum_{r=1}^{l-1}A_rA_{2l-r}+A_l^2/2\right\}, \quad (28b)$$

where $l\leqq2$. Now $A_k\approx\alpha^{-1}$ is as large a value of A_k as it is prudent to use in Eqs. (28a) and (28b) without closely investigating higher order terms in the α-series expansion. The coefficient of α^n in a Kryloff and Bogoliuboff series in general contains powers of the A_k^m with $m>n$. Thus, setting $A_k\approx\alpha^{-1}$ introduces a diverging factor in the coefficient of α^n which the smallness of α^n alone does not overcome. In view of this fact, we chose to solve Eqs. (28a) and (28b) by setting $A_1=1$ and then computing the remaining A_k subject to the conditions $x_k(0)=0$, $k=2,\ \cdots,\ 5$. On carrying out this procedure, one calculates from Eqs. (28a) and (28b) that $A_2=-0.249$, $A_3=0.088$, $A_4=-0.035$, and $A_5=0.013$, while $x_1(0)=1.35$. On putting these values of the A_k into Eqs. (22a) and (22b) and calculating the energy in the first three normal modes via the equations,

$$E_k=\dot{x}_k^2/2+(\omega_k^2x_k^2)/2 \quad (29)$$

provides the data for the curves plotted in Fig. 1. The UFP computer curves to be compared with Fig. 1 are presented in Fig. 2. All data are the same except for the magnitude of $x_1(0)$ which equals 1.35 for the series solution and 250 for the UFP solution. Quantitatively, then, the two solutions are not comparable; qualitatively, however, they are similar. Both solutions appear to be periodic with periods the same order of magnitude, and for both solutions the amount of energy sharing decreases monotonically with increasing mode number. The greater amount of energy sharing present in the UFP solution indicates the presence of internal resonance for these first few modes. The Kryloff and Bogoliuboff solution indicates its own need to include internal resonance, i.e., amplitude modulation, for this comparison by its evident divergence for $x_1(0)=250$.

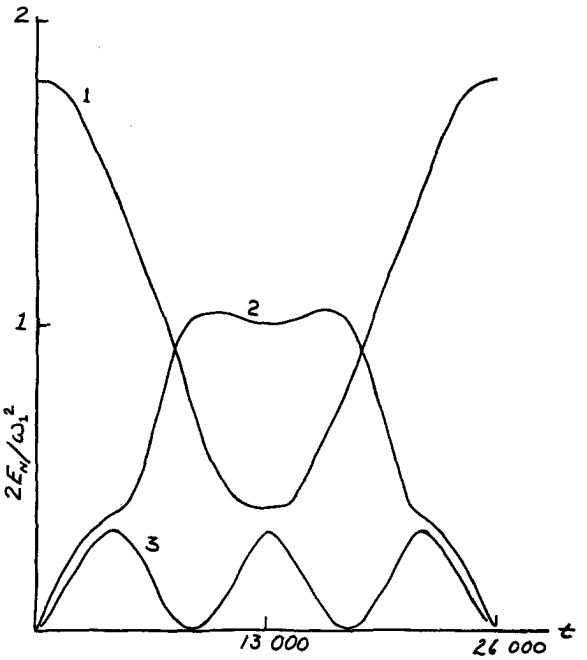

FIG. 1. A plot of the energy in the first three normal modes for $N=32$, $\alpha=\frac{1}{4}$, and $x_1(0)=1.35$ as predicted by the Kryloff and Bogoliuboff solution. Initially, mode 1 possesses the total energy. The energy in modes higher than mode 3 is negligible on this scale.

IV. CONCLUSIONS

The central purpose of this paper has been to expose a reason for the failure of the UFP computer calculation to reveal equipartition of energy. This purpose has certainly been achieved. We now add one additional note which also summarizes the reasons for the failure. We have remarked that equipartition for very weakly coupled oscillator systems can never occur as long as the uncoupled frequencies are linearly independent, i.e., as long as

$$\sum_k n_k\omega_k\neq0$$

for any nonzero collection of integers $\{n_k\}$. Moreover, we have heuristically connected this notion with the concept of resonance remarking that $\sum_k n_k\omega_k\neq0$ means that the interacting oscillators cannot drive each other at resonance. It is interesting to note that all

these ideas are revealed in the form of the Kryloff and Bogoliuboff solution. The solution to Eqs. (18) of the last section may be written in the form

$$x_{2l+1}=A_{2l+1}\sin\tau_{2l+1}+\frac{\alpha}{(2N)^{\frac{1}{2}}}\sum_{r=1}^{N-2l-2}A_rA_{r+2l+1}\times\frac{\cos(\tau_r-\tau_{r+2l+1})}{[\omega_{2l+1}^2-(\omega_r-\omega_{r+2l+1})^2]}\left\{\sin\left[\frac{\pi(r+k)}{N}\right]-\sin\left[\frac{r\pi}{N}\right]-\sin\left[\frac{k\pi}{N}\right]+\cdots,\right. \quad (30)$$

where the resonance denominator

$$[\omega_{2l+1}^2-(\omega_r-\omega_{r+2l+1})^2]$$

has been explicitly written out. Equations (21a) and (21b) of the previous section hid these resonance denominators in the cot and tan factors. Solution (30) clearly indicates that as $\alpha\to 0$, there will be no energy sharing unless some resonance denominator is zero, i.e., unless the uncoupled frequencies are linearly dependent. Moreover, the form of the resonance denominators almost demands an interpretation in terms of internal resonance. Seen in this light, the lack of equipartition of energy for the "linearly independent" UFP oscillator system is a property only of the form of the Kryloff and Bogoliuboff solution and hence might be expected to be independent of the convergence properties of these solutions.

That the UFP oscillator systems cannot exhibit equipartition is clear both from our heuristic arguments and the basic work of Balescu.[8] It is not expected that this result will be affected one way or another by the convergence questions regarding the Kryloff-Bogoliuboff series. The validity of Sec. III, however, does rest on such questions. Unfortunately, no really satisfactory answer has been given to the question of convergence. The author chooses to believe that the similarity of the series solution to the accurate computer solution is an indication of its validity, but this is not offered as a compelling argument. The validity of the solution found in Sec. III could most easily be verified using a computer, and it is hoped that this may be accomplished in the near future.

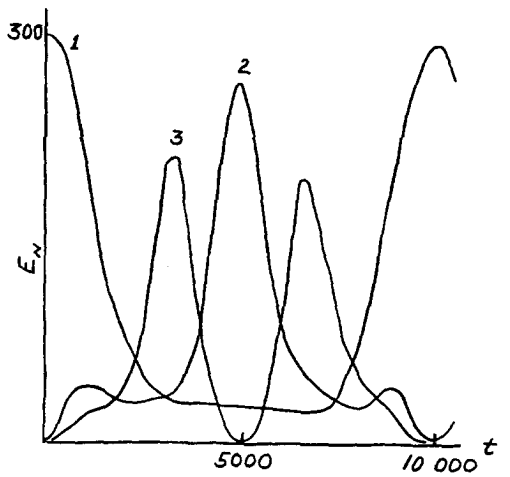

FIG. 2. A plot of the energy in the first three normal modes for $N=32$, $\alpha=\frac{1}{4}$, and $x_1(0)\approx 250$ as computed by UFP. Initially, mode 1 possesses the total energy. The energy in modes higher than mode 5 is negligible on this scale.

As to the future, by using as a guide the linear problem of Sec. II, the form of the Kryloff and Bogoliuboff series, and the criteria given in the work of Balescu, it should be possible to find a system of differential equations which can be solved on a computer to demonstrate equipartition of energy. Indeed such calculations might shed light on the whole question of the approach to equilibrium. Perhaps even more interesting, however, is the possibility of generalizing or modifying the Kryloff and Bogoliuboff technique to cover the case of internal resonance. Even a formal solution to the problem, independent of the covergence question, would be useful in view of the fact that a computer might be used to establish the range of usefulness of such a solution.

ACKNOWLEDGMENTS

It is a pleasure to thank Professor T. H. Berlin for suggesting this problem. The author also gratefully acknowledges the financial support given for this work by the National Science Foundation.

Higher-Order Approximations in Multiple Scattering.* I. Two-Dimensional Scalar Case

NORMAN ZITRON†
Institute of Mathematical Sciences, New York University, New York, New York

AND

SAMUEL N. KARP
Institute of Mathematical Sciences, New York University, New York, New York

(Received June 23, 1960)

A formula is derived which expresses the perturbed scattering amplitudes of a combination of two arbitrary cylinders as a function of the unperturbed scattering amplitudes of the individual cylinders. The formula is valid when the spacing of the scatterers is large compared to their dimensions. It involves derivatives of the scattering amplitudes with respect to the angles of incidence and of observation. Interaction terms of degrees $d^{-\frac{1}{2}}$, d^{-1}, and $d^{-\frac{3}{2}}$ are taken into account, where d is the spacing. Verification is obtained in a special case. The result is employed to calculate the total scattering cross section.

1. INTRODUCTION

THE present paper deals with the diffraction of plane electromagnetic or acoustic waves by a pair of parallel cylinders of arbitrary shape. The diffraction by each cylinder, in isolation, is assumed known, and the diffraction by the configuration is calculated explicitly, in terms of these data. An approximation is involved which will be described below.

The question of multiple scattering has already been treated, but in less detail, by a number of writers. A brief sketch of the history of the problem follows, with emphasis on those treatments of the problem whose accuracy increases with the spacing. We recall the work of Reiche and Schaefer,[1] who were the first to have given a wave theoretical discussion of the finite grating of circular cylinders. These authors neglected the interaction between the cylinders; their work was therefore valid, in principle, in the limit of large spacing. A very general expression for the diffraction by an arbitrary assemblage of circular cylinders was given by Twersky,[2] who took all orders of interaction into account. The method depended heavily upon the separability of the circular geometry, and the most general form of the result was too complicated to be discussed. However, Twersky found that, if he proceeded to the limit of large spacing, he could simplify his result immensely and achieve a perspicuous discussion of the correction to single scattering. This work gave a correct account of terms of degree $d^{-\frac{1}{2}}$ and d^{-1} in the spacing.

This success led to a general investigation of the large spacing approximation by Karp,[3] who showed that for cylinders of arbitrary shape, the leading terms of the *interaction correction* could themselves be expressed explicitly in terms of *noninteraction* or *single scattering* results, the latter being regarded as given. In fact, the interaction term could be regarded as being composed of the response of each cylinder to a plane wave arriving from the direction of the other cylinders. The techniques of Karp[3] were exploited by Karp and Radlow[4] and by Karp[5] in the analysis of a grating of cylinders. Similar methods were used by Karp and Russek[6] in expressing the approximate solution to the problem of diffraction by a wide slit in terms of the well-known solution for the half-plane problem.

The purpose of the present paper is to extend the work of Karp[3] so as to take into account higher order terms. Just as in reference 3, the cylinders are arbitrary, and the scattering by each cylinder in isolation is assumed as given. But, it is found[7] that even the higher order correction terms can be calculated generally, simply, and explicitly in terms of the single scattering data used in reference 3 for calculation of the leading terms. This is the principal result of the present paper.

The general calculation was carried out so as to include all effects of order of magnitude $d^{-\frac{1}{2}}$, d^{-1}, and $d^{-\frac{3}{2}}$, where d is the spacing of the cylinders. For purpose of comparison, Twersky's calculation for a pair of circular cylinders was continued so as to include terms of order $d^{-\frac{3}{2}}$; this special calculation by the repeated application of additional theorems was then shown to agree with the general result of the present work, when the latter result is specialized to the case of circular cylinders. The result is used to calculate the total cross section for a pair of circular cylinders as a function of the spacing and the known unperturbed (or noninteraction) scattering amplitude functions.

* The research reported in this article was sponsored by the Air Force Cambridge Research Center, Air Research and Development Command, under contract.

† Present Address: Dept. of Engineering Science, The Florida State University, Tallahassee, Florida.

[1] F. Reiche and C. Schaefer, Physik **35**, 817 (1911).

[2] V. Twersky, J. Appl. Phys. **23**, 407 (1952).

[3] S. Karp, "Diffraction by a combination of cylinders," Proc. McGill Symposium, McGill University, Montreal, Canada, June, 1953. AF CRC-TR-59-118(II). AST1A #AD211500.

[4] S. Karp and J. Radlow, IRE Trans. on Antennas and Propagation. AP-4, No. 4 (1956).

[5] S. N. Karp, N. Y. U. Institute of Mathematical Sciences, Division of EM Research, Research Rept. No. EM-85, October, 1955. See also Phys. Rev. **86**, 586(A) (1952).

[6] S. Karp and A. Russek, J. Appl. Phys. **27**, 886 (1956).

[7] N. R. Zitron and S. N. Karp, Bull. Am. Phys. Soc. **3**, 184 (1958).

2. STATEMENT OF THE PROBLEM AND OF THE METHOD OF ANALYSIS

We would like to know the scattering pattern of a combination of two infinite parallel cylinders in terms of the scattering patterns which these cylinders would have if they were isolated from each other. In other words, we want to obtain a functional relationship between the unperturbed and the perturbed scattering patterns of the cylinders. Such a relation is desirable because it simplifies the calculation of the pattern for the combination. If we can calculate the unperturbed patterns, we need only insert them in this relation to obtain the perturbed patterns. The relation is useful, moreover, even if the shapes of the cylinders are so complicated that we cannot separate variables or if calculation by separation of variables is too tedious.[8] Also, in such cases, the unperturbed patterns might be measured experimentally for all angles of observation and these unperturbed patterns might then be substituted into the relation obtained here to yield the perturbed patterns.

The situation is the following: A plane wave of unit amplitude is incident upon the two parallel cylinders A and B (Fig. 1).

To avoid ambiguity in the definition of the spacing d, we define a coordinate system for each cylinder. Let A' and B' be circular cylinders circumscribed about A and B, respectively. Let a and b be the respective radii of A' and B'. We shall let Z_a be the axis of A' and Z_b be the axis of B'. The problem is two dimensional, and we shall operate in a plane perpendicular to the Z axes. The respective intersections of this plane with the Z_a and Z_b axes will then be the origins of the coordinate systems of A and B. We can now define the spacing d as the distance between the axes of A' and B', that is, the distance between the two origins (Fig. 1).

We make the following assumptions:

(1) The individual complex scattering pattern is known when each cylinder stands alone in space.

(2) $d \gg \lambda$ where λ is the wavelength of the incident wave.[9]

(3) $d \gg a$ and $d \gg b$.

Our object is to find a functional relation between the scattering pattern of the combination and those of the isolated component cylinders. This cannot be accomplished by simple superposition of the unperturbed patterns of the components since the individual pattern of each component cylinder is modified by the field scattered by the other cylinder. We must, therefore, consider the interaction.

[8] See reference 4 for the use of an even less accurate relation for theoretical purposes.

[9] Numerical calculation in reference 5, which are based on a less accurate procedure, showed that d could be as small as one wavelength without impairing the accuracy materially. The present result would, therefore, allow an even smaller spacing, provided the cylinders are kept sufficiently small compared to the spacing.

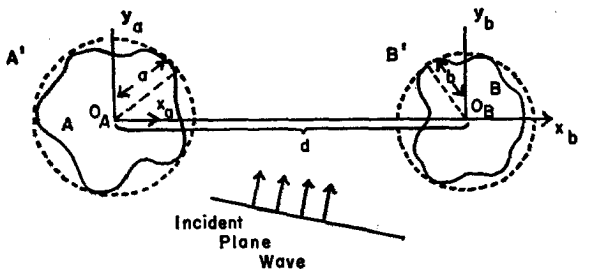

FIG. 1. Plane wave incident upon two parallel cylinders.

We shall assume that the response U_s of each cylinder to a plane wave is of the form

$$U_s = \frac{e^{ikr}}{r^{\frac{1}{2}}} \sum_{n=0}^{\infty} \frac{f_n(\theta,\theta_0)}{r^n} \quad \text{for large } r, \tag{1}$$

where r is the distance from the axis of the circumscribed circular cylinder, θ is the angle of observation, and θ_0 is the angle of incidence. Both θ and θ_0 are measured from the x axis of each cylinder. The x axes are collinear with d.

As will be explained below, at a sufficiently large distance from the cylinder, the scattered field U_s resembles a plane wave. This approximate plane wave elicits a response from the other cylinder, perturbing its scattered field. This response also has the form (1) and in turn perturbs the field scattered by the first cylinder. We can carry out successive calculations for this process until the desired order of accuracy is obtained. When the perturbed patterns have been calculated, they can be superposed.

We shall deal, in this paper, with interaction terms of degrees $d^{-\frac{1}{2}}$, d^{-1}, and $d^{-\frac{3}{2}}$. The procedure involves a new kind of expansion of the waves scattered by each cylinder about the origin located in the other cylinder. In order to ensure that all terms up to order $d^{-\frac{1}{2}}$ are contained in the result, we must include them in the first expansion. We then find that the field scattered by a given cylinder can be represented, in the neighborhood of the other cylinder, as a plane wave, plus additional terms which are recognized as derivatives of a plane wave with respect to its angle of incidence. The simple way of expressing the higher-order excitations is what enables us to calculate the higher-order responses conveniently.

3. EXPANSION OF THE SCATTERED WAVES IN TERMS OF PLANE WAVES

A. Expansion of the Response of Cylinder A in a Neighborhood of Cylinder B

Let us consider what happens when a plane wave is incident upon A. The wave function for a plane wave (Fig. 2) is

$$U_i = \exp[ik(x_a \cos\theta_0 + y_a \sin\theta_0)]. \tag{2}$$

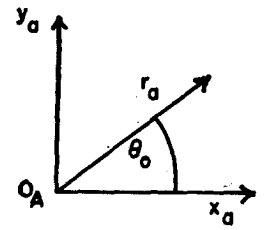

FIG. 2. Coordinate system for cylinder A.

The wave scattered by A in response to the plane wave is represented by an asymptotic solution of the reduced wave equation.

$$\text{D.E.} \qquad (\Delta+k^2)U=0. \tag{3}$$

The boundary conditions are such that the diffraction problem for each cylinder, and for the combination, is well posed. They may be, for example,

$$\text{B.C.} \qquad \begin{cases} \text{(a)}\ U=0 \\ \text{or (b)}\ \partial U/\partial n=0 \\ \text{or (c)}\ CU+D(\partial U/\partial n)=0, \end{cases} \tag{4}$$

where $\partial U/\partial n$ is the normal derivative of U, and C and D are constants. Alternatively, one or both of the cylinders may be filled with dielectric materials.

The radiation condition for an outgoing wave, i.e.,

$$\lim_{r\to\infty} r^{\frac{1}{2}}\left[\frac{\partial U}{\partial r}-ikU\right]=0, \tag{5}$$

is imposed on all scattered fields which occur.

The solution of (3) has the form

$$U=U_i+U_a, \tag{6}$$

where U_i is the incident field and U_a is the field scattered by cylinder A. We assume that U_a may be represented in the asymptotic form

$$U_a=\frac{e^{ikr_a}}{(r_a)^{\frac{1}{2}}}\left[f_0^{a0}(\theta_a,\theta_0)+\frac{f_1^{a0}}{r_a}(\theta_a,\theta_0)+\mathcal{O}(r_a^{-2})\right]. \tag{7}$$

The letter "a" signifies that the variable in question refers to cylinder A, and the superscript "$a0$" signifies that the pattern is unperturbed. If, as we have assumed, B is sufficiently small in relation to its distance from A, then the wave scattered by A is practically a plane wave in the neighborhood of B, and we may imagine such a wave incident on B. We can demonstrate an explicit representation of this approximately plane wave by expanding U_a in a neighborhood of B. We shall express the expansion in powers of $d^{-\frac{1}{2}}$ in a rectangular coordinate system with origin at the center of B' and carry out the calculation up to order $d^{-\frac{3}{2}}$. The calculation proceeds as follows. Let P be a point in the neighborhood of B (Fig. 3). Let r_a = the distance from the axis of A' to P. We see from Fig. 3 that $r_a=[(d+x_b)^2+y_b^2]^{\frac{1}{2}}$, and we shall henceforth omit the subscripts "a",

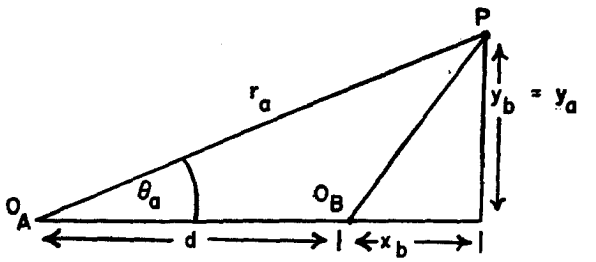

FIG. 3. Coordinate system for expansion in a neighborhood of cylinder B.

"b" as a matter of convenience. Then

$$e^{ikr}=e^{ik(d+x)}\left[1+\frac{iky^2}{2d}+\frac{ikxy^2-\frac{1}{4}k^2y^4}{d^2}+\mathcal{O}(d^{-3})\right], \tag{8}$$

$$r^{-\frac{1}{2}}=(d^2+2xd+x^2+y^2)^{-\frac{1}{4}} =(1/d^{\frac{1}{2}})-(x/2d^{\frac{3}{2}})+\mathcal{O}(d^{-\frac{5}{2}}), \tag{9}$$

$$r^{-\frac{3}{2}}=d^{-\frac{3}{2}}[1-\tfrac{3}{2}(x/d)+\mathcal{O}(d^{-2})] =(1/d^{\frac{3}{2}})+\mathcal{O}(d^{-\frac{5}{2}}), \tag{10}$$

$$\theta_a=\arctan(y/d+x) =(y/d)-(xy/d^2)+\mathcal{O}(d^{-3}), \tag{10a}$$

$$f_n^{a0}(\theta,\theta_0)=f_n^{a0}(0,\theta_0)+[D_\theta f_n^{a0}(0,\theta_0)]\theta +[D_\theta^2 f_n^{a0}(0,\theta_0)](\theta^2/2)+\cdots, \tag{10b}$$

where $D_\theta\equiv\partial/\partial\theta$. By using (10a), we can rewrite (10b) in the form

$$f_n^{a0}(\theta,\theta_0)=f_n^{a0}(0,\theta)+[D_\theta f_n^{a0}(0,\theta_0)](y/d) -[D_\theta f_n^{a0}(0,\theta_0)](xy/d^2) +[D_\theta^2 f_n^{a0}(0,\theta_0)](y^2/2d^2)+\mathcal{O}(d^{-3}). \tag{11}$$

On substituting Eqs. (8)–(11) into (7), we obtain

$$U_a^0=e^{ik(d+x)}\left[1+\frac{iky^2}{2d}+\frac{ikxy^2-\frac{1}{4}k^2y^4}{d^2}+\mathcal{O}(d^{-3})\right] \times\left\{\begin{aligned}&\left[\frac{1}{d^{\frac{1}{2}}}-\frac{x}{2d^{\frac{3}{2}}}+\mathcal{O}(d^{-\frac{5}{2}})\right]\\ &\quad[f_0^{a0}(0,\theta_0)+D_\theta f_0^{a0}(0,\theta_0)(y/d)+\mathcal{O}(d^{-2})]\\ &\qquad+\frac{1}{d^{\frac{3}{2}}}[f_1^{a0}(0,\theta_0)+\mathcal{O}(d^{-1})]\end{aligned}\right.$$

$$U_a^0=e^{ik(d+x)}\left[\frac{f_0^{a0}(0,\theta_0)}{d^{\frac{1}{2}}} +\frac{\frac{1}{2}(-x+iky^2)f_0^{a0}(0,\theta_0)+yD_\theta f_0^{a0}(0,\theta_0)+f_1^{a0}(0,\theta_0)}{d^{\frac{3}{2}}} +\mathcal{O}(d^{-\frac{5}{2}})\right]. \tag{12}$$

We have thus obtained the field scattered by A as it appears in a neighborhood of B. We must now carry out the corresponding expansion for the field initially scattered by B. Note that the leading term of (12) is a plane wave, i.e., a constant multiple of e^{ikx}, as we explained earlier.

B. Expansion in a Neighborhood of A for the Wave Scattered by B

The same type of procedure as that used above gives us the expansion in a neighborhood of A of the wave initially scattered by B. The expression here corresponding to (12) is

$$U_b{}^0=e^{ik(d-x)}\left[\frac{f_0{}^{b0}(\pi,\theta_0)}{d^{\frac{1}{2}}}+\frac{\frac{1}{2}(x+iky^2)f_0{}^{b0}(\pi,\theta_0)-yD_\theta f_0{}^{b0}(\pi,\theta_0)+f_1{}^{b0}(\pi,\theta_0)}{d^{\frac{3}{2}}}+\mathcal{O}(d^{-\frac{5}{2}})\right]. \quad (13)$$

The unperturbed fields scattered by either cylinder, (12) or (13), are the excitations of the other cylinder. We assumed, to begin with, that we knew the unperturbed response of each cylinder to a plane-wave excitation for all angles of incidence. If (12) and (13) were plane waves, we could calculate the responses for the second scattering.[10] (12) and (13), however, are not plane waves, but this impediment does not prevent the calculation of the effect of further scattering. The reason is that (12) and (13) may be represented in terms of plane waves by appropriate substitutions. This representation, which will enable us to calculate successive scattering, will now be given.

C. Expression of the Scattered Waves in Terms of Plane Waves

We can reduce the further scattering of singly-scattered waves to the scattering of plane waves by expressing (12) and (13) in terms of plane waves and derivatives of plane waves. On noting that a plane wave is represented by

$$v(\theta_0)=\exp[ik(x\cos\theta_0+y\sin\theta_0)], \quad (14)$$

we observe that

$$ikye^{ikx}=v_{\theta_0}(0) \quad (15)$$

$$-ikye^{-ikx}=v_{\theta_0}(\pi) \quad (16)$$

$$ik(-x+iky^2)e^{ikx}=v_{\theta_0\theta_0}(0) \quad (17)$$

$$ik(x+iky^2)e^{-ikx}=v_{\theta_0\theta_0}(\pi). \quad (18)$$

Substitution of (15) and (17) in (12), and of (16) and (18) in (13) yields the following representation of the scattered fields in terms of plane waves:

$$U_a{}^0=e^{ikd}\left[\frac{v(0)f_0{}^{a0}(0,\theta_0)}{d^{\frac{1}{2}}}+\frac{(1/2ik)[D_{\theta_0}{}^2v(0)]f_0{}^{a0}(0,\theta_0)+(1/ik)[D_{\theta_0}v(0)]D_\theta f_0{}^{a0}(0,\theta_0)+v(0)f_1{}^{a0}(0,\theta_0)}{d^{\frac{3}{2}}}+\mathcal{O}(d^{-\frac{5}{2}})\right] \quad (19)$$

$$U_b{}^0=e^{ikd}\left[\frac{v(\pi)f_0{}^{b0}(\pi,\theta_0)}{d^{\frac{1}{2}}}+\frac{(1/2ik)[D_{\theta_0}{}^2v(\pi)]f_0{}^{b0}(\pi,\theta_0)+(1/ik)[D_{\theta_0}v(\pi)]D_\theta f_0{}^{b0}(\pi,\theta_0)+v(\pi)f_1{}^{b0}(\pi,\theta_0)}{d^{\frac{3}{2}}}+\mathcal{O}(d^{-\frac{5}{2}})\right]. \quad (20)$$

D. Elimination of f_1

We note that the numerators of (19) and (20) are sums of terms, consisting of plane waves and their derivatives, namely, the v's, and coefficients which are independent of θ, namely, the f's. We see that these formulas are expressed in terms of both f_0 and f_1. (For the meaning of f_0 and f_1 see reference 7.) An advantage would result from the elimination of f_1, since we could then express the result in terms of the scattering amplitude of the far field without having to know the scattering amplitudes of further asymptotic terms.

We eliminate f_1 by expressing it in terms of f_0. This can be done by means of a recursion formula. The recursion is obtained by substitution into (3) of the assumed asymptotic[11] representation (7) of any radiating solution of the reduced-wave equation. When we equate the corresponding inverse powers of r, we find that

$$f_n=(1/2ikn)[(n-\tfrac{1}{2})^2f_{n-1}+D_\theta{}^2f_{n-1}]. \quad (21)$$

This recursion is useful, also, for calculations of higher degree than we are considering here. Since we want to express f_1 in terms of f_0, we need use it only for the value $n=1$.

$$f_1=(1/2ik)[\tfrac{1}{4}f_0+D_\theta{}^2f_0]. \quad (22)$$

This is a different type of recursion from that obtained for large k by Keller, Lewis, and Seckler[12] although it is similar in form.

If we now substitute (22) into (19) and (20), we obtain

[10] This would be the method used in reference 3.

[11] The representation is asymptotic for large r, θ, and k being held fixed.

[12] J. B. Keller, R. M. Lewis, and B. D. Seckler, Communs. Pure and Appl. Math. 9, 207 (1956).

$$U_a^0=e^{ikd}\Bigg[\frac{v(0)f_0^{a0}(0,\theta_0)}{d^{\frac{1}{2}}}+\frac{(1/2ik)[D_{\theta_0}{}^2v(0)]f_0^{a0}(0,\theta_0)+(1/ik)[D_{\theta_0}v(0)]D_\theta f_0^{a0}(0,\theta_0)+(1/2ik)v(0)[\frac{1}{4}f_0^{a0}(0,\theta_0)+D_\theta{}^2f_0^{a0}(0,\theta_0)]}{d^{\frac{3}{2}}}+\mathcal{O}(d^{-\frac{5}{2}})\Bigg], \tag{23}$$

$$U_b^0=e^{ikd}\Bigg[\frac{v(\pi)f_0^{b0}(\pi,\theta_0)}{d^{\frac{1}{2}}}+\frac{(1/2ik)[D_{\theta_0}{}^2v(\pi)]f_0^{b0}(\pi,\theta_0)+(1/ik)[D_{\theta_0}v(\pi)]D_\theta f_0^{b0}(\pi,\theta_0)+(1/2ik)v(\pi)[\frac{1}{4}f_0^{b0}(\pi,\theta_0)+D_\theta{}^2f_0^{b0}(\pi,\theta_0)]}{d^{\frac{3}{2}}}+\mathcal{O}(d^{-\frac{5}{2}})\Bigg]. \tag{24}$$

Expressions (23) and (24) represent the responses of the cylinders to the original incident plane wave and these responses are given near the other cylinder in terms of the plane waves. Expressions (23) and (24) are also excitations for the second scattering. Successive application of these formulas will yield the desired degree of interaction.

4. CALCULATION OF THE INTERACTION

We have expressed in (23) and (24), the fields singly scattered by A and by B, as they appear in a neighborhood of the second scatterer and have, moreover, expressed them in terms of plane waves and derivatives of plane waves. We may now imagine this combination of plane waves and their derivatives to be incident upon the second scatterer. The linearity of these expressions enables us to say that the responses of A and B to incident derivatives of plane waves are equal to the derivatives of the responses of A and B to the incident plane waves. Since we already know, by assumption, the unperturbed responses of A and B to an incident plane wave, we have reduced the second scattering to the previous case, namely, the first scattering.

If we carry out this process to the extent of three successive scatterings, we can obtain interaction terms of degrees $d^{-\frac{1}{2}}$, d^{-1}, and $d^{-\frac{3}{2}}$. The terms of degree $d^{-\frac{1}{2}}$ result from double scattering, those of degree d^{-1} result from triple scattering, and those of degree $d^{-\frac{3}{2}}$ result partly from double and partly from quadruple scattering.

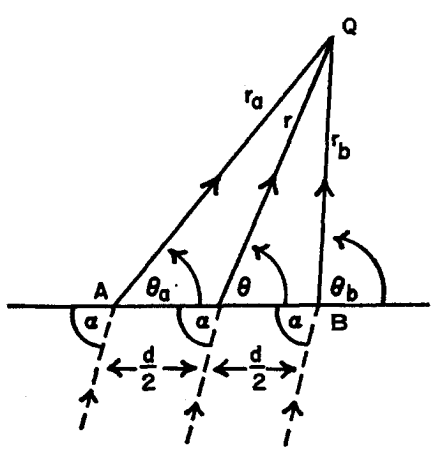

FIG. 4. Normalized coordinate systems.

For the sake of simplicity, the following results will be expressed in a normalized coordinate system (Fig. 4) with a common origin midway between the origins of the coordinate systems located in the scatterers. As a matter of convenience, we shall omit the subscript zero from the f's.

Let α be the angle of incidence of the original plane wave. The perturbed patterns will then be

$$\begin{aligned}f^a(\theta,\alpha)=&e^{-ik\frac{1}{2}d\cos\alpha}f^{a0}(\theta,\alpha)\\&+\frac{e^{ik\frac{1}{2}d\cos\alpha}e^{ikd}f^{b0}(\pi,\alpha)f^{a0}(\theta,\pi)}{d^{\frac{1}{2}}}\\&+\frac{e^{-ik\frac{1}{2}d\cos\alpha}e^{ik2d}f^{a0}(0,\alpha)f^{b0}(\pi,0)f^{a0}(\theta,\pi)}{d}\\&+\frac{e^{ik\frac{1}{2}d\cos\alpha}e^{ik3d}}{d^{\frac{3}{2}}}f^{b0}(\pi,\alpha)f^{a0}(0,\pi)f^{b0}(\pi,0)f^{a0}(\theta,\pi)\\&+\frac{e^{ik\frac{1}{2}d\cos\alpha}e^{ikd}}{d^{\frac{1}{2}}2ik}(f^{b0}(\pi,\alpha)D_{\theta_0}{}^2f^{a0}(\theta,\pi)\\&+2D_\theta f^{b0}(\pi,\alpha)D_{\theta_0}f^{a0}(\theta,\pi)+[\tfrac{1}{4}f^{b0}(\pi,\alpha)\\&+D_\theta{}^2f^{b0}(\pi\alpha)]f^{a0}(\theta,\pi))+\mathcal{O}(d^{-\frac{3}{2}}),\end{aligned} \tag{25}$$

and

$$\begin{aligned}f^b(\theta,\alpha)=&e^{ik\frac{1}{2}d\cos\alpha}f^{b0}(\theta,\alpha)+\frac{e^{-ik\frac{1}{2}d\cos\alpha}e^{ikd}}{d^{\frac{1}{2}}}f^{a0}(0,\alpha)f^{b0}(\theta,0)\\&+\frac{e^{ik\frac{1}{2}d\cos\alpha}e^{ik2d}}{d}f^{b0}(\pi,\alpha)f^{a0}(0,\pi)f^{b0}(\theta,0)\\&+\frac{e^{-ik\frac{1}{2}d\cos\alpha}e^{ik3d}}{d^{\frac{3}{2}}}f^{a0}(0,\alpha)f^{b0}(\pi,0)f^{a0}(0,\pi)f^{b0}(\theta,0)\\&+\frac{e^{-ik\frac{1}{2}d\cos\alpha}e^{ikd}}{d^{\frac{3}{2}}2ik}(f^{a0}(0,\alpha)D_{\theta_0}{}^2f^{b0}(\theta,0)\\&+2D_\theta f^{a0}(0,\alpha)D_{\theta_0}f^{b0}(\theta,0)+[\tfrac{1}{4}f^{a0}(0,\alpha)\\&+D_\theta{}^2f^{a0}(0,\alpha)]f^{b0}(\theta,0))+\mathcal{O}(d^{-\frac{3}{2}}).\end{aligned} \tag{26}$$

We note that the successive powers of $d^{-\frac{1}{2}}$ in (25) and (26) represent the various degrees of interaction. The scattered far fields resulting from the interaction will have the form

$$U_a=\frac{\exp[ik(r+\frac{1}{2}d\cos\theta)]}{r^{\frac{1}{2}}}f^a(\theta,\alpha), \tag{27}$$

and

$$U_b=\frac{\exp[ik(r-\frac{1}{2}d\cos\theta)]}{r^{\frac{1}{2}}}f^b(\theta,\alpha), \tag{28}$$

where $f^a(\theta,\alpha)$ and $f^b(\theta,\alpha)$ are the perturbed scattering amplitudes given by (25) and (26). The sum of the scattered fields has the following far-field representation which is the sum of (27) and (28).

$$U_s=U_a+U_b=(e^{ikr}/r^{\frac{1}{2}})(2/\pi k)^{\frac{1}{2}}e^{-i(\pi/4)}F(\theta,\alpha), \tag{29}$$

where $F(\theta,\alpha)$ is the scattering amplitude of the combination of cylinders. We see by comparison with (27) and (28) that

$$F(\theta,\alpha)=(\pi k/2)^{\frac{1}{2}}e^{i(\pi/4)}[\exp(ik\tfrac{1}{2}d\cos\theta)f^a(\theta,\alpha)+\exp(-ik\tfrac{1}{2}d\cos\theta)f^b(\theta,\alpha)]. \tag{30}$$

A more explicit formula for the far field is obtained by combining (25), (26), (29), and (30). The far field can then be written as follows:

$$\begin{aligned}u\cong\frac{e^{ikr}}{r^{\frac{1}{2}}}\Big\{&\exp[ik\tfrac{1}{2}d(\cos\theta-\cos\alpha)]f^{a0}(\theta,\alpha)+\exp[-ik\tfrac{1}{2}d(\cos\theta-\cos\alpha)]f^{b0}(\theta,\alpha)\\
&+\frac{e^{ikd}}{d^{\frac{1}{2}}}[\exp[ik\tfrac{1}{2}d(\cos\theta+\cos\alpha)]f^{b0}(\pi,\alpha)f^{a0}(\theta,\pi)+\exp[-ik\tfrac{1}{2}d(\cos\theta+\cos\alpha)]f^{a0}(0,\alpha)f^{b0}(\theta,0)]\\
&+\frac{e^{ik2d}}{d}[\exp[ik\tfrac{1}{2}d(\cos\theta-\cos\alpha)]f^{a0}(0,\alpha)f^{b0}(\pi,0)f^{a0}(\theta,\pi)+\exp[-ik\tfrac{1}{2}d(\cos\theta-\cos\alpha)]f^{b0}(\pi,\alpha)f^{a0}(0,\pi)f^{b0}(\theta,0)]\\
&+\frac{e^{ik3d}}{d^{\frac{3}{2}}}[\exp[ik\tfrac{1}{2}d(\cos\theta+\cos\alpha)]f^{b0}(\pi,\alpha)f^{a0}(0,\pi)f^{b0}(\pi,0)f^{a0}(\theta,\pi)\\
&+\exp[-ik\tfrac{1}{2}d(\cos\theta+\cos\alpha)]f^{a0}(0,\alpha)f^{b0}(\pi,0)f^{a0}(0,\pi)f^{b0}(\theta,0)]\\
&+\frac{e^{ikd}}{2ikd^{\frac{3}{2}}}\Big[\exp[ik\tfrac{1}{2}d(\cos\theta+\cos\alpha)]\big(f^{b0}(\pi,\alpha)D_{\theta_0}{}^2f^{a0}(\theta,\pi)+2D_\theta f^{b0}(\pi,\alpha)D_{\theta_0}f^{a0}(\theta,\pi)\\
&\qquad+\{\tfrac{1}{4}f^{b0}(\pi,\alpha)+D_\theta{}^2f^{b0}(\pi,\alpha)\}f^{a0}(\theta,\pi)\big)\\
&\quad+\exp[-ik\tfrac{1}{2}d(\cos\theta+\cos\alpha)]\big(f^{a0}(0,\alpha)D_{\theta_0}{}^2f^{b0}(\theta,0)+2D_\theta f^{a0}(0,\alpha)D_{\theta_0}f^{b0}(\theta,0)\\
&\qquad+\{\tfrac{1}{4}f^{a0}(0,\alpha)+D_\theta{}^2f^{a0}(0,\alpha)\}f^{b0}(\theta,0)\big)\Big]+\mathcal{O}(d^{-\frac{5}{2}})\Big\}.\end{aligned} \tag{31}$$

We have presented in (31) a relation between the scattered far field of the combination and those of the component cylinders. We wish to point out that a similar relation holds between the corresponding fields at *all* points of space. But this relation will not be detailed here.

5. SPECIAL CASE. SCATTERING BY TWO CONDUCTING CIRCULAR CYLINDERS

The abstract relations obtained above can be verified in the special case of scattering by two parallel, arbitrary circular cylinders A and B with corresponding sets of parameters a and b. For a single cylinder, say A, we have

$$U_a{}^0=\sum_{n=-\infty}^{\infty}i^nC_n{}^aH_n{}^{(1)}(kr_a)e^{in(\theta_a-\alpha)}, \tag{32}$$

where $C_n{}^a$ is the appropriate scattering coefficient for any of the usual boundary conditions; e.g., if the field vanishes on the cylinder's surface $r=a$, then $C_n{}^a=-J_n(ka)/H_n{}^{(1)}(ka)$; if the normal derivative vanishes, we replace the functions $J_n(ka)$, $H_n{}^{(1)}(ka)$ by their derivatives with respect to their arguments, etc. The use of the asymptotic form of $H_n{}^{(1)}(kr_a)$ in (32) yields the far field

$$U_a{}^0=\frac{e^{ikr_a}}{(r_a)^{\frac{1}{2}}}\Big\{e^{-i\pi/4}\Big(\frac{2}{\pi k}\Big)^{\frac{1}{2}}\sum_n C_n{}^ae^{in(\theta_a-\alpha)}\Big\}=\frac{e^{ikr_a}}{(r_a)^{\frac{1}{2}}}f^{a0}(\theta_a,\alpha), \tag{33}$$

where the function in braces is the unperturbed pattern. Similarly, for cylinder B (whose radius and boundary conditions differ in general from those of A), we replace a by b in (32) and (33).

The corresponding perturbed pattern may be obtained by specializing the general series of "Neumann type" in Eq. (3) of reference 2; see reference 13 for a derivation of the series and for a discussion of its physical significance. If we now expand this result in inverse powers of d and retain all terms of degree $d^{-\frac{1}{2}}$, we find (with reference to an origin midway between those of the cylinders)

$$f_0^a(\theta,\alpha)=\exp[-ik\tfrac{1}{2}d\cos\alpha]e^{-i\pi/4}\left(\frac{2}{\pi k}\right)^{\frac{1}{2}}\sum_n C_n{}^a e^{in(\theta-\alpha)}$$

$$+\frac{\exp[ik\tfrac{1}{2}d\cos\alpha]}{d^{\frac{1}{2}}}e^{-i\pi/2}\left(\frac{2}{\pi k}\right)\sum_n(-1)^{n-1}C_n{}^b e^{-in\alpha}\sum_{n'}(-1)C_{n'}{}^a e^{in'(\theta-\pi)}$$

$$+\frac{\exp[-ik\tfrac{1}{2}d\cos\alpha]e^{ik2d}}{d}e^{-i3\pi/4}\left(\frac{2}{\pi k}\right)^{\frac{3}{2}}\sum_n C_n{}^a e^{-in\alpha}\sum_{n'}(-1)^{n'-1}C_{n'}{}^b\sum_{n''}e^{in''(\theta-\pi)}$$

$$+\frac{\exp[ik\tfrac{1}{2}d\cos\alpha]e^{ik3d}}{d^{\frac{3}{2}}}e^{-i\pi}\left(\frac{2}{\pi k}\right)^{2}\sum_n(-1)^{n-1}C_n{}^b e^{-in\alpha}\sum_{n'}(-1)^{n'-1}C_{n'}{}^a\sum_{n''}(-1)^{n''-1}C_{n''}{}^b$$

$$\times\sum_{n'''}(-1)C_{n'''}{}^a e^{in'''(\theta-\pi)}$$

$$+\frac{\exp[ik\tfrac{1}{2}d\cos\alpha]e^{ikd}}{d^{\frac{3}{2}}}\frac{e^{-i\pi/2}}{2ik}\left(\frac{2}{\pi k}\right)\sum_n(-1)^n C_n{}^b e^{in\alpha}\sum_{n'}(-1)[(n-n')^2-\tfrac{1}{4}]C_{n'}{}^a e^{in'(\theta-\pi)}+\mathcal{O}(d^{-\frac{5}{2}}),\quad(34)$$

$$f_0^b(\theta,\alpha)=\exp[ik\tfrac{1}{2}d\cos\alpha]e^{-i\pi/2}\left(\frac{2}{\pi k}\right)^{\frac{1}{2}}\sum_n C_n{}^b e^{in(\theta-\alpha)}$$

$$+\frac{\exp[ik\tfrac{1}{2}d\cos\alpha]e^{ikd}}{d^{\frac{1}{2}}}e^{-i\pi/2}\left(\frac{2}{\pi k}\right)\sum_n(-1)C_n{}^a e^{-in\alpha}\sum_{n'}(-1)C_{n'}{}^b e^{in'\theta}$$

$$+\frac{\exp[ik\tfrac{1}{2}d\cos\alpha]e^{ik2d}}{d}e^{-i3\pi/4}\left(\frac{2}{\pi k}\right)^{\frac{3}{2}}\sum_n(-1)^n C_n{}^b e^{-in\alpha}\sum_{n'}(-1)^{n'-1}C_{n'}{}^a\sum_{n''}(-1)C_{n''}{}^b e^{in''\theta}$$

$$+\frac{\exp[-ik\tfrac{1}{2}d\cos\alpha]e^{ik3d}}{d^{\frac{3}{2}}}e^{-i\pi}\left(\frac{2}{\pi k}\right)^{2}\sum_n C_n{}^a e^{-in\alpha}\sum_{n'}(-1)^{n'-1}C_{n'}{}^b\sum_{n''}(-1)^{n''-1}C_{n''}{}^a$$

$$\times\sum_{n'''}(-1)C_{n'''}{}^b e^{in'''\theta}$$

$$+\frac{\exp[-k\tfrac{1}{2}d\cos\alpha]e^{ikd}}{d^{\frac{3}{2}}}\frac{e^{-i\pi/2}}{2ik}\left(\frac{2}{\pi k}\right)\sum_n C_n{}^a e^{-in\alpha}\sum_{n'}(-1)[(n'-n)^2-\tfrac{1}{4}]C_{n'}{}^b e^{in\theta}+\mathcal{O}(d^{-\frac{5}{2}}),\quad(35)$$

which we shall first compare with the expansion of the closed-form approximation given in Eq. (6) of reference 2. That approximation was first obtained by keeping only the largest term of each order of scattering and hence its expansion is not quite as accurate as (34) and (35). To carry out the comparison, we specialize (34) and (35) to the case of two identical cylinders ($a=b$). Then, comparison shows agreement so far as the non-interaction terms and the terms of degrees $d^{-\frac{1}{2}}$ and d^{-1} are concerned. The first of our terms of degree $d^{-\frac{3}{2}}$ agrees with the term of order $d^{-\frac{3}{2}}$ in the expansion of Twersky's result. But our second term of degree $d^{-\frac{3}{2}}$ is new. This is to be expected for the following reason. The closed form referred to above is obtained by summing the leading terms only, of the successive orders of scattering, that is, the successive bounces. Our term of degree $d^{-\frac{3}{2}}$, on the other hand, contains higher-order contributions from the second bounce, in addition to the leading term of the fourth bounce.

The procedure used above for obtaining (34) and (35) is long and tedious. These results need not be obtained by that procedure. The use of the abstract formulas (25) and (26) simplifies the calculation considerably and should do the same in any other case where the scattering problems for the component cylinders are separable. All we need do is substitute the specific unperturbed patterns $f^{a0}(\theta,\alpha)$ and $f^{b0}(\theta,\alpha)$ into (25) and (26), respectively. The results follow immediately and agree with (34) and (35), respectively.

[13] V. Twersky, J. Acoust. Soc. Am. 24, 42 (1952).

6. EXPLANATION OF THE RESULT

Expression (31) for the scattered far field of the combination of two cylinders appears, at first glance, to be complicated. However, it is not as abstruse as it seems. A closer examination of these expressions reveals the significance of these various terms and factors. We note first that the terms are grouped in increasing orders of accuracy. We note also that they are grouped in pairs. The first members of each pair represent fields ultimately scattered by cylinder A, while the second members of the pairs represent fields ultimately scattered by cylinder B. The first pair corresponds to single scattering while the other pairs correspond to multiple scattering. The factors

$$\exp(\pm ik\tfrac{1}{2}d\cos\theta)$$

represent the phase differences for the scatterers relative to the point of observation, while the factors

$$\exp(\pm ik\tfrac{1}{2}d\cos\alpha)$$

take into account the phase of the incident wave at the center of a scatterer when the incident wave has zero phase at the origin. The f's containing a θ dependence are scattering patterns, whereas the f's containing specific values for θ are excitation factors accumulated in the multiple scattering. We note also that the factors of the form e^{iknd} where $n=0, 1, 2, 3$ refer to the increase ikd in the phase of a wave in going from one scatterer to another and that n signifies the number of bounces.

As an illustration of the above remarks, let us consider some of the terms in more detail. The term

$$(e^{ikr}/r^{\frac{1}{2}})\exp[ik\tfrac{1}{2}d(\cos\theta-\cos\alpha)]f^{a0}(\theta,\alpha)$$

would occur in the case of no interaction, i.e., in the limit of infinite spacing. The factor $\exp[\frac{1}{2}ikd(\cos\theta)]$ takes account of the fact that the origin is not at the center of A. The positive sign preceding $\cos\theta$ shows that the scattered wave came from cylinder A. The negative sign in the exponent of the factor $\exp(-ik\frac{1}{2}d\times\cos\alpha)$ shows that cylinder A received the initial excitation. It gives the phase of the incident wave at A. The term

$$(e^{ikr}/r^{\frac{1}{2}})\exp[ik\tfrac{1}{2}d(\cos\theta+\cos\alpha)]f_{(\pi,\alpha)}{}^{b0}f_{(\theta,\pi)}{}^{a0}$$

differs from the first term in the following respects. The positive sign in the factor $\exp(+\frac{1}{2}ikd\cos\alpha)$ shows that cylinder B received the original excitation as does the factor $f_{(\pi,\alpha)}{}^{b0}$. The phase factor e^{ikd} represents the increase of the phase of the wave in going from B to A.

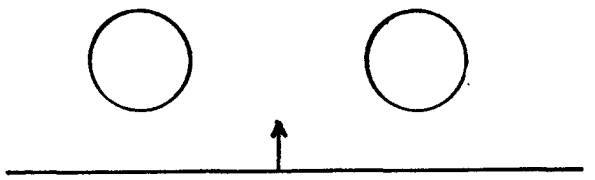

FIG. 5. Plane wave normally incident upon two parallel circular cylinders.

We shall now explain a typical term of the last square bracket. The first term in the last square bracket of (31) represents the scattering by A of a term of order $d^{-\frac{1}{2}}$ initially scattered by B. The phase factor e^{ikd} takes account of the travel of this wave from B to A.

The positive sign preceding $\cos\alpha$ shows that cylinder B was excited initially. The differentiation of the scattering amplitudes shows the effect of a higher-order excitation of A by the field initially scattered by B, since the higher-order response of B, (which acts as an excitation for A), is representable, near A, as a derivative of a plane wave with respect to angle of incidence.

The explanation of further terms in the final result proceeds on the same lines as the explanations given above. We omit these explanations for the sake of brevity.

7. TOTAL SCATTERING CROSS SECTION OF A COMBINATION OF TWO IDENTICAL CIRCULAR CYLINDERS

We consider a plane wave normally incident on a pair of identical circular cylinders. The circumstances are illustrated in Fig. 5.

The computation of the total scattering cross section σ of the two identical circular cylinders is facilitated by the use of (30) in conjunction with the following well-known theorem[14,15]:

$$\sigma=-(4/k)\,\mathrm{Re}F(\pi/2,\pi/2), \tag{36}$$

where $F(\theta,\alpha)$ is the scattering amplitude (30) of the combination. The phase factors are simplified for the values $\theta=\pi/2$, $\alpha=\pi/2$. The fact that the cylinders are identical enables us to write $f^0=f^{a0}=f^{b0}$. Further simplifications result from the geometrical symmetry of the problem, namely,

$$f^0(0,\pi)=f^0(\pi,0),\quad f^0(\pi/2,0)=f^0(0,\pi/2)$$
$$=f^0(\pi,\pi/2)=f^0(\pi/2,\pi).$$

The total scattering cross section in terms of the

[14] V. Twersky, J. Appl. Phys. 25, 859 (1954).
[15] C. H. Papas, J. Appl. Phys. 21, 318 (1950).

spacing is then

$$\sigma=-(4/k)\ \mathrm{Re}\left\{(2\pi k)^{\frac{1}{2}}e^{i\pi/4}\left\{f^0\left(\frac{\pi}{2},\frac{\pi}{2}\right)+\frac{e^{ikd}}{d^{\frac{1}{2}}}\left[f^0\left(0,\frac{\pi}{2}\right)\right]^2+\frac{e^{ik2d}}{d}\left[f^0\left(0,\frac{\pi}{2}\right)\right]^2 f^0(\pi,0)\right.\right.$$
$$+\frac{e^{ik3d}}{d^{\frac{3}{2}}}\left[f^0\left(0,\frac{\pi}{2}\right)\right]^2[f^0(\pi,0)]^2$$
$$+\frac{e^{ikd}}{d^{\frac{1}{2}}}\frac{1}{2ik}\left(f^0\left(0,\frac{\pi}{2}\right)D_{\theta_0}{}^2f^0\left(0,\frac{\pi}{2}\right)+2D_\theta f^0\left(0,\frac{\pi}{2}\right)D_{\theta_0}f_\phi{}^0\left(0,\frac{\pi}{2}\right)\right.$$
$$\left.\left.\left.+\left[\tfrac{1}{4}f^0\left(0,\frac{\pi}{2}\right)+D_\theta{}^2f^0\left(0,\frac{\pi}{2}\right)\right]f^0\left(0,\frac{\pi}{2}\right)\right)+\mathcal{O}(d^{-\frac{3}{2}})\right\}\right\} \tag{37}$$

8. ADDITIONAL REMARKS

It is clear that interactions of degree greater than $d^{-\frac{3}{2}}$ can be computed by the inclusion of more terms in the expansions used to obtain these results.

It is also clear that the method applies to cases of more than two scatterers, but the computations would be more tedious than in the case of two scatterers. The computations might be simplified by using a "consistency" method employed in reference 3 rather than tracing the successive scattering in detail. This method involves a steady-state point of view. The response of each cylinder is expanded in a neighborhood of each of the other cylinders. Each cylinder will then be excited by the incident plane wave and by an approximately plane wave from each of the other cylinders. These considerations introduce certain undetermined coefficients which can be determined by imposing the requirement that the fields scattered by the various cylinders be consistent with one another. Evaluation of the coefficients will provide the solution.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 2, NUMBER 3 MAY-JUNE, 1961

Higher-Order Approximations in Multiple Scattering.* II. Three-Dimensional Scalar Case

NORMAN ZITRON†‡
Gordon McKay Laboratory, Harvard University, Cambridge, Massachusetts

AND

SAMUEL N. KARP
Institute of Mathematical Sciences, New York University, New York, New York

(Received June 23, 1960)

The method of Part I is extended to cover the three-dimensional scalar problem for two bodies of arbitrary shape. All interaction terms of order d^{-1} and d^{-2} are given.

1. STATEMENT OF THE PROBLEM

THE method employed previously in the case of multiple scattering of plane waves by two widely spaced cylinders of arbitrary shape can be applied, also, to the corresponding three-dimensional scalar problem for two bodies of arbitrary shape. The assumptions of spacing large compared to the wavelength and the dimensions of the bodies and that the individual responses of the scatterers are known apply also to this case.

The situation is the following. A plane wave of unit amplitude

$$u=\exp[ik(x\sin\theta_0\cos\phi_0+y\sin\theta_0\sin\phi_0+z\cos\theta_0)], \tag{1}$$

where θ_0 and ϕ_0 are the angles of incidence (see Fig. 1), is incident upon the combination of two bodies. The response of each body in isolation to the incident plane wave is of the form

$$U_s=\frac{e^{ikr}}{r}\sum_{n=0}^{\infty}\frac{f_n(\theta,\theta_0,\phi,\phi_0)}{r^n}. \tag{2}$$

Here we are referring to spherical coordinates, r, θ, ϕ,

* The research reported in this article was sponsored by the Air Force Cambridge Research Center, Air Research and Development Command, under contract.

† Present Address: Dept. of Engineering Science, The Florida State University, Tallahassee, Florida.

‡ Norman Zitron's work in connection with Part II was performed at Harvard University and was supported by the Office of Naval Research under contract with the Gordon McKay Laboratory.

spacing is then

$$\sigma=-(4/k)\ \mathrm{Re}\left\{(2\pi k)^{\frac{1}{2}}e^{i\pi/4}\left\{f^0\left(\frac{\pi}{2},\frac{\pi}{2}\right)+\frac{e^{ikd}}{d^{\frac{1}{2}}}\left[f^0\left(0,\frac{\pi}{2}\right)\right]^2+\frac{e^{ik2d}}{d}\left[f^0\left(0,\frac{\pi}{2}\right)\right]^2 f^0(\pi,0)\right.\right.$$
$$+\frac{e^{ik3d}}{d^{\frac{3}{2}}}\left[f^0\left(0,\frac{\pi}{2}\right)\right]^2[f^0(\pi,0)]^2$$
$$+\frac{e^{ikd}}{d^{\frac{1}{2}}}\frac{1}{2ik}\left(f^0\left(0,\frac{\pi}{2}\right)D_{\theta_0}{}^2f^0\left(0,\frac{\pi}{2}\right)+2D_\theta f^0\left(0,\frac{\pi}{2}\right)D_{\theta_0}f_\phi{}^0\left(0,\frac{\pi}{2}\right)\right.$$
$$\left.\left.\left.+\left[\tfrac{1}{4}f^0\left(0,\frac{\pi}{2}\right)+D_\theta{}^2f^0\left(0,\frac{\pi}{2}\right)\right]f^0\left(0,\frac{\pi}{2}\right)\right)+\mathcal{O}(d^{-\frac{3}{2}})\right\}\right\} \tag{37}$$

8. ADDITIONAL REMARKS

It is clear that interactions of degree greater than $d^{-\frac{3}{2}}$ can be computed by the inclusion of more terms in the expansions used to obtain these results.

It is also clear that the method applies to cases of more than two scatterers, but the computations would be more tedious than in the case of two scatterers. The computations might be simplified by using a "consistency" method employed in reference 3 rather than tracing the successive scattering in detail. This method involves a steady-state point of view. The response of each cylinder is expanded in a neighborhood of each of the other cylinders. Each cylinder will then be excited by the incident plane wave and by an approximately plane wave from each of the other cylinders. These considerations introduce certain undetermined coefficients which can be determined by imposing the requirement that the fields scattered by the various cylinders be consistent with one another. Evaluation of the coefficients will provide the solution.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 2, NUMBER 3 MAY-JUNE, 1961

Higher-Order Approximations in Multiple Scattering.* II. Three-Dimensional Scalar Case

NORMAN ZITRON†‡
Gordon McKay Laboratory, Harvard University, Cambridge, Massachusetts

AND

SAMUEL N. KARP
Institute of Mathematical Sciences, New York University, New York, New York

(Received June 23, 1960)

The method of Part I is extended to cover the three-dimensional scalar problem for two bodies of arbitrary shape. All interaction terms of order d^{-1} and d^{-2} are given.

1. STATEMENT OF THE PROBLEM

THE method employed previously in the case of multiple scattering of plane waves by two widely spaced cylinders of arbitrary shape can be applied, also, to the corresponding three-dimensional scalar problem for two bodies of arbitrary shape. The assumptions of spacing large compared to the wavelength and the dimensions of the bodies and that the individual responses of the scatterers are known apply also to this case.

The situation is the following. A plane wave of unit amplitude

$$u=\exp[ik(x\sin\theta_0\cos\phi_0+y\sin\theta_0\sin\phi_0+z\cos\theta_0)], \tag{1}$$

where θ_0 and ϕ_0 are the angles of incidence (see Fig. 1), is incident upon the combination of two bodies. The response of each body in isolation to the incident plane wave is of the form

$$U_s=\frac{e^{ikr}}{r}\sum_{n=0}^{\infty}\frac{f_n(\theta,\theta_0,\phi,\phi_0)}{r^n}. \tag{2}$$

Here we are referring to spherical coordinates, r, θ, ϕ,

* The research reported in this article was sponsored by the Air Force Cambridge Research Center, Air Research and Development Command, under contract.

† Present Address: Dept. of Engineering Science, The Florida State University, Tallahassee, Florida.

‡ Norman Zitron's work in connection with Part II was performed at Harvard University and was supported by the Office of Naval Research under contract with the Gordon McKay Laboratory.

such that

$$x=r\sin\theta\cos\phi,$$
$$y=r\sin\theta\sin\phi, \tag{2a}$$
$$z=r\cos\theta.$$

The letters x, y, z are Cartesian coordinates of a point P with respect to a coordinate system located appropriately. The quantities θ_0 and ϕ_0 which specify the direction of travel of the incident wave can be thought of as the colatitude and longitude, respectively, of a point on the unit sphere. That point is determined as the intersection with the unit sphere of a line passing through the origin in the direction of travel of the incident wave.

The object is to obtain the perturbed complex scattering amplitude of the combination as a function of the individual unperturbed scattering amplitudes. This can be done, as in the previous case, by first expanding the field scattered by each body in a neighborhood of the other body and then expressing these scattered fields in terms of plane waves and their derivatives. The *a priori* assumption that the unperturbed responses to an incident plane wave are known permits the calculation of the perturbed responses.

We must establish coordinate systems for the problem. We first circumscribe spheres A' and B' about the respective bodies A and B. This establishes the origins of the coordinate systems of A and B. The origins of bodies A and B are established by means of spheres A' and B' circumscribed about them (Fig. 2). The spacing d is defined as the line joining the two centers. We then choose the x_a axis, i.e., the x axis for the coordinate system of A, to be collinear with d and such that x_a increases in the direction of B. The y_a and z_a axes can then be chosen to form a right-handed system and the coordinate system of B is parallel to that of A. (Choice of the z axis along d, instead of the x axis, would be more symmetrical, but has been avoided for reasons which will be explained later.)

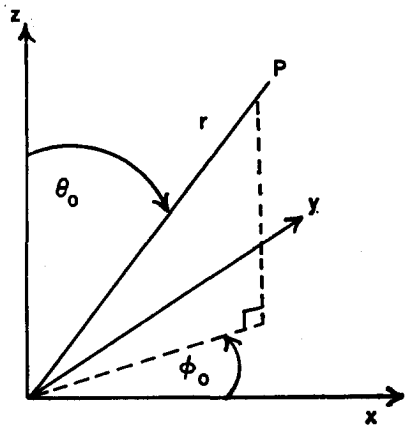

FIG. 1. Spherical coordinate system.

2. EXPANSION OF SCATTERED WAVES IN TERMS OF PLANE WAVES

We proceed to expand the response of each body in a neighborhood of the other body. This expansion yields the following field scattered by A in a neighborhood of B where the Cartesian coordinates are in the B system:

$$U_{s_a}{}^0=\exp[ik(d+x_b)]\Big[\frac{1}{d}f_0{}^{a0}\Big(\frac{\pi}{2},\theta_0,0,\phi_0\Big)$$
$$+\frac{1}{d^2}\Big\{\frac{ik(y^2+z^2)}{2}f_0{}^{a0}\Big(\frac{\pi}{2},\theta_0,0,\phi_0\Big)$$
$$-x_bf_0{}^{a0}\Big(\frac{\pi}{2},\theta_0,0,\phi_0\Big)-zD_\theta f_0{}^{a0}\Big(\frac{\pi}{2},\theta_0,0,\phi_0\Big)$$
$$+yD_\phi f_0{}^{a0}\Big(\frac{\pi}{2},\theta_0,0,\phi_0\Big)+f_1{}^{a0}\Big(\frac{\pi}{2},\theta_0,0,\phi_0\Big)\Big\}$$
$$+\mathcal{O}(d^{-3})\Big]. \tag{3}$$

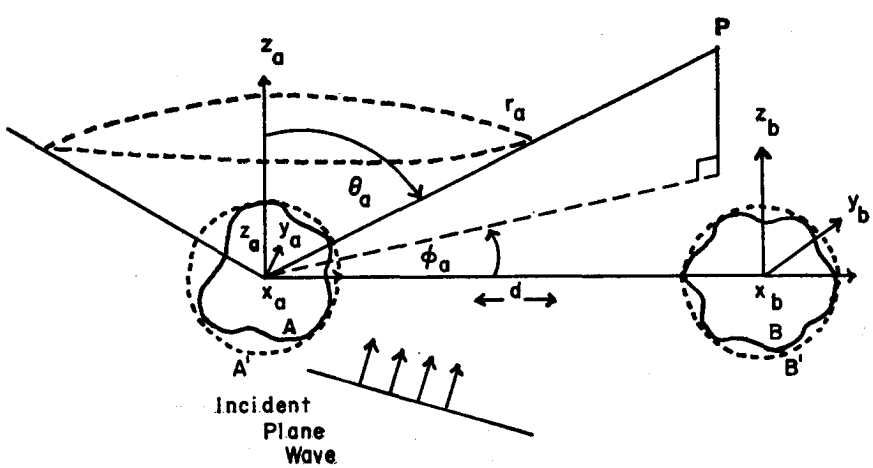

FIG. 2. Plane wave incident upon two bodies.

In calculating Eq. (3), the phase of the incident wave at the "center" of A has been ignored. The appropriate phase factor will be restored in the final result. A similar expansion of the field scattered by B in a neighborhood of A yields

$$U_{s_b}{}^0=\exp[ik(d-x_a)]\Big[\frac{1}{d}f_0{}^{b0}\Big(\frac{\pi}{2},\theta_0,\pi,\phi_0\Big)+\frac{1}{d^2}\Big\{\frac{ik(y^2+z^2)}{2}f_0{}^{b0}\Big(\frac{\pi}{2},\theta_0,\pi,\phi_0\Big)+x_a f_0{}^{b0}\Big(\frac{\pi}{2},\theta_0,\pi,\phi_0\Big)-zD_\theta f_0{}^{b0}\Big(\frac{\pi}{2},\theta_0,\pi,\phi_0\Big)-yD_\phi f_0{}^{b0}\Big(\frac{\pi}{2},\theta_0,\pi,\phi_0\Big)+f_1{}^{b0}\Big(\frac{\pi}{2},\theta_0,\pi,\phi_0\Big)\Big\}+\mathcal{O}(d^{-3})\Big], \quad (4)$$

where $D_\theta=\partial/\partial\theta$ and $D_\phi\equiv\partial/\partial\phi$. We now express these scattered fields in terms of plane waves by means of the following relations:

$$v(\theta_0,\phi_0)=\exp[ik(x\sin\theta_0\cos\phi_0+y\sin\theta_0\sin\phi_0+z\cos\theta_0)],$$

$$v(\pi/2,0)=e^{ikx},\quad v(\pi/2,\pi)=e^{-ikx},$$

$$v_{\theta_0}(\pi/2,0)=-ikze^{ikx},\quad v_{\theta_0}(\pi/2,\pi)=-ikze^{-ikx},$$

$$v_{\phi_0}(\pi/2,0)=ikye^{ikx},\quad v_{\phi_0}(\pi/2,\pi)=-ikye^{-ikx},$$

$$\frac{1}{2ik}\Big[v_{\theta_0\theta_0}\Big(\frac{\pi}{2},0\Big)+v_{\phi_0\phi_0}\Big(\frac{\pi}{2},0\Big)\Big]=\Big[\frac{ik(y^2+z^2)}{2}-x\Big]e^{ikx},$$

$$\frac{1}{2ik}\Big[v_{\theta_0\theta_0}\Big(\frac{\pi}{2},\pi\Big)+v_{\phi_0\phi_0}\Big(\frac{\pi}{2},\pi\Big)\Big]=\Big[\frac{ik(y^2+z^2)}{2}+x\Big]e^{-ikx}.$$

Substitution of these relations into (3) and (4) yields

$$U_{s_a}{}^0=e^{ikd}\Big[\frac{1}{d}v\Big(\frac{\pi}{2},0\Big)f_0{}^{a0}\Big(\frac{\pi}{2},\theta_0,0,\phi_0\Big)+\frac{1}{d^2}\Big(\frac{1}{2ik}\Big[v_{\theta_0\theta_0}\Big(\frac{\pi}{2},0\Big)+v_{\phi_0\phi_0}\Big(\frac{\pi}{2},0\Big)\Big]\times f_0{}^{a0}\Big(\frac{\pi}{2},\theta_0,0,\phi_0\Big)+v_{\theta_0}\Big(\frac{\pi}{2},0\Big)D_\theta f_0{}^{a0}\Big(\frac{\pi}{2},\theta_0,0,\phi_0\Big)+v_{\phi_0}\Big(\frac{\pi}{2},0\Big)D_\phi f_0{}^{a0}\Big(\frac{\pi}{2},\theta_0,0,\phi_0\Big)+v\Big(\frac{\pi}{2},0\Big)f_1{}^{a0}\Big(\frac{\pi}{2},\theta_0,0,\phi_0\Big)\Big)+\mathcal{O}(d^{-3})\Big], \quad (5)$$

and

$$U_{s_b}{}^0=e^{ikd}\Big[\frac{1}{d}v\Big(\frac{\pi}{2},\pi\Big)f_0{}^{b0}\Big(\frac{\pi}{2},\theta_0,\pi,\phi_0\Big)+\frac{1}{d^2}\Big(\frac{1}{2ik}\Big[v_{\theta_0\theta_0}\Big(\frac{\pi}{2},\pi\Big)+v_{\phi_0\phi_0}\Big(\frac{\pi}{2},\pi\Big)\Big]\times f_0{}^{b0}\Big(\frac{\pi}{2},\theta_0,\pi,\phi_0\Big)+v_{\theta_0}\Big(\frac{\pi}{2},\pi\Big)D_\theta f_0{}^{b0}\Big(\frac{\pi}{2},\theta_0,\pi,\phi_0\Big)+v_{\phi_0}\Big(\frac{\pi}{2},\pi\Big)D_\phi f_0{}^{b0}\Big(\frac{\pi}{2},\theta_0,\pi,\phi_0\Big)+v\Big(\frac{\pi}{2},\pi\Big)f_1{}^{b0}\Big(\frac{\pi}{2},\theta_0,\pi,\phi_0\Big)\Big)+\mathcal{O}(d^{-3})\Big]. \quad (6)$$

These fields can then be expressed in terms of f_0 alone by means of a recursion formula (7) obtained by Sommerfeld[1]:

$$2ik(n+1)f_{n+1}=\Big\{n(n+1)+\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\Big(\sin\theta\frac{\partial}{\partial\theta}\Big)+\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\Big\}f_n. \quad (7)$$

In our case, $n=0$ and the recursion reduces to

$$2ikf_1=\{\cot\theta D_\theta+D_\theta{}^2+\csc^2\theta D_\phi{}^2\}f_0. \quad (8)$$

It is here that the proper choice of a coordinate system is seen to be convenient. We note that the recursion involves the Beltrami Operator which is singular at $\theta=0$. This singularity occurs in the direction joining A to B, if the z axis is chosen collinear with d. Another difficulty in using such a coordinate system is the occurrence of terms in the expansions (3) and (4) which are not

[1] A. Sommerfeld, *Partial Differential Equations in Physics* (Academic Press, Inc., New York, 1949), p. 192.

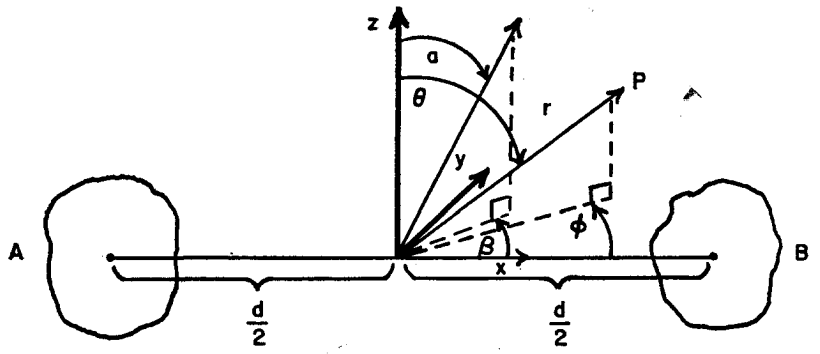

FIG. 3. Normalized coordinate system.

easily replaceable by derivatives of plane waves. Our choice of coordinate system avoids these difficulties.

For $\theta=\pi/2$ and $\phi=0$, Eq. (8) becomes

$$f_1(\pi/2,\theta_0,0,\phi_0)=(1/2ik)\{D_\theta^2+D_\phi^2\}f_0(\pi/2,\theta_0,0,\phi_0). \quad (9)$$

For $\theta=\pi/2$ and $\phi=\pi$, Eq. (8) becomes

$$f_1(\pi/2,\theta_0,\pi,\phi_0)=(1/2ik)\{D_\theta^2+D_\phi^2\}f_0(\pi/2,\theta_0,\pi,\phi_0). \quad (10)$$

Substitution of Eqs. (9) and (10) into (5) and (6) results in expressions (11) and (12) for $U_{s_a}{}^0$ and $U_{s_b}{}^0$, which do not contain $f_1{}^{a0}$ and $f_1{}^{b0}$:

$$\begin{aligned}U_{s_a}{}^0=e^{ikd}\Big[&\frac{1}{d}v\left(\frac{\pi}{2},0\right)f_0{}^{a0}\left(\frac{\pi}{2},\theta_0,0,\phi_0\right)\\&+\frac{1}{d^2}\Big\{\frac{1}{2ik}\Big[\{D_{\theta_0}{}^2+D_{\phi_0}{}^2\}v\left(\frac{\pi}{2},0\right)\Big]f_0{}^{a0}\left(\frac{\pi}{2},\theta_0,0,\phi_0\right)\\&+D_{\theta_0}v\left(\frac{\pi}{2},0\right)D_\theta f_0{}^{a0}\left(\frac{\pi}{2},\theta_0,0,\phi_0\right)\\&+D_{\phi_0}v\left(\frac{\pi}{2},0\right)D_\phi f_0{}^{a0}\left(\frac{\pi}{2},\theta_0,0,\phi_0\right)\\&+\frac{1}{2ik}v\left(\frac{\pi}{2},0\right)\{D_\theta{}^2+D_\phi{}^2\}f_0{}^{a0}\left(\frac{\pi}{2},\theta_0,0,\phi_0\right)\Big\}\\&+\mathcal{O}(d^{-3})\Big], \quad (11)\end{aligned}$$

$$\begin{aligned}U_{s_b}{}^0=e^{ikd}\Big[&\frac{1}{d}v\left(\frac{\pi}{2},\pi\right)f_0{}^{b0}\left(\frac{\pi}{2},\theta_0,\pi,\phi_0\right)\\&+\frac{1}{d^2}\Big\{\frac{1}{2ik}\Big[\{D_{\theta_0}{}^2+D_{\phi_0}{}^2\}v\left(\frac{\pi}{2},\pi\right)\Big]f_0{}^{b0}\left(\frac{\pi}{2},\theta_0,\pi,\phi_0\right)\\&+D_{\theta_0}v\left(\frac{\pi}{2},\pi\right)D_\theta f_0{}^{b0}\left(\frac{\pi}{2},\theta_0,\pi,\phi_0\right)\\&+D_{\phi_0}v\left(\frac{\pi}{2},\pi\right)D_\phi f_0{}^{b0}\left(\frac{\pi}{2},\theta_0,\pi,\phi_0\right)\\&+\frac{1}{2ik}v\left(\frac{\pi}{2},\pi\right)\{D_\theta{}^2+D_\phi{}^2\}f_0{}^{b0}\left(\frac{\pi}{2},\theta_0,\pi,\phi_0\right)\Big\}\\&+\mathcal{O}(d^{-3})\Big]. \quad (12)\end{aligned}$$

3. CALCULATION OF THE PERTURBED FAR-FIELD AMPLITUDE

The method employed previously in the two-dimensional case will now yield, up to interaction terms of degree d^{-2}, the field scattered by the two bodies.

The final result must take into account the difference in phase of the original plane wave with respect to the two scatterers and also the difference in phase of the scattered fields of the two bodies at the point of observation. This can be done by a normalization of the coordinate systems and subsequent consideration of the phases in the normalized coordinate system. By a normalized coordinate system we mean a central coordinate system whose origin lies midway between the origins for A and B and whose x, y, and z axes are parallel to those emanating from the origins of A and B. If we let α and β denote the angles of incidence corresponding to colatitude and longitude, respectively, (Fig. 3), we obtain the following expression for the field scattered by the two bodies. (The subscript "0" is omitted from the f's for convenience. Here the incident wave is taken to have zero phase at the origin of the normalized system of coordinates.)

$$
\begin{aligned}
U_s \cong \frac{e^{ikr}}{r}\Bigg[& \exp[ik\tfrac{1}{2}d(\sin\theta\cos\phi-\sin\alpha\cos\beta)]f^{a0}(\theta,\alpha,\phi,\beta)+\exp[-ik\tfrac{1}{2}d(\sin\theta\cos\phi-\sin\alpha\cos\beta)]f^{b0}(\theta,\alpha,\phi,\beta) \\
& \times\exp[ik\tfrac{1}{2}d(\sin\theta\cos\phi+\sin\alpha\cos\beta)]\frac{e^{ikd}}{d}f^{a0}\left(\theta,\frac{\pi}{2},\phi,\pi\right)f^{b0}\left(\frac{\pi}{2},\alpha,\pi,\beta\right) \\
& +\exp[-ik\tfrac{1}{2}d(\sin\theta\cos\phi+\sin\alpha\cos\beta)]\frac{e^{ikd}}{d}f^{b0}\left(\theta,\frac{\pi}{2},\phi,0\right)f^{a0}\left(\frac{\pi}{2},\alpha,0,\beta\right) \\
& +\exp[ik\tfrac{1}{2}d(\sin\theta\cos\phi-\sin\alpha\cos\beta)]\frac{e^{ik2d}}{d^2}f^{a0}\left(\theta,\frac{\pi}{2},\phi,\pi\right)f^{b0}\left(\frac{\pi}{2},\frac{\pi}{2},\pi,0\right)f^{a0}\left(\frac{\pi}{2},\alpha,0,\beta\right) \\
& +\exp[-ik\tfrac{1}{2}d(\sin\theta\cos\phi-\sin\alpha\cos\beta)]\frac{e^{ik2d}}{d^2}f^{b0}\left(\theta,\frac{\pi}{2},\phi,0\right)f^{a0}\left(\frac{\pi}{2},\frac{\pi}{2},0,\pi\right)f^{b0}\left(\frac{\pi}{2},\alpha,\pi,\beta\right) \\
& +\exp[ik\tfrac{1}{2}d(\sin\theta\cos\phi+\sin\alpha\cos\beta)]\frac{e^{ikd}}{d^2}\Big\{\frac{1}{2ik}[D_{\theta_0}{}^2+D_{\phi_0}{}^2]f^{a0}\left(\theta,\frac{\pi}{2},\phi,\pi\right)f^{b0}\left(\frac{\pi}{2},\alpha,\pi,\beta\right) \\
& +D_{\theta_0}f^{a0}\left(\theta,\frac{\pi}{2},\phi,\pi\right)D_\theta f^{b0}\left(\frac{\pi}{2},\alpha,\pi,\beta\right)+D_{\phi_0}f^{a0}\left(\theta,\frac{\pi}{2},\phi,\pi\right)D_\phi f^{b0}\left(\frac{\pi}{2},\alpha,\pi,\beta\right) \\
& +\frac{1}{2ik}f^{a0}\left(\theta,\frac{\pi}{2},\phi,\pi\right)[D_\theta{}^2+D_\phi{}^2]f^{b0}\left(\frac{\pi}{2},\alpha,\pi,\beta\right)\Big\} \\
& +\exp[-ik\tfrac{1}{2}d(\sin\theta\cos\phi+\sin\alpha\cos\beta)]\frac{e^{ikd}}{d^2}\Big\{\frac{1}{2ik}[D_{\theta_0}{}^2+D_{\phi_0}{}^2]f^{b0}\left(\theta,\frac{\pi}{2},\phi,0\right)f^{a0}\left(\frac{\pi}{2},\alpha,0,\beta\right) \\
& +D_{\theta_0}f^{b0}\left(\theta,\frac{\pi}{2},\phi,0\right)D_\theta f^{a0}\left(\frac{\pi}{2},\alpha,0,\beta\right)+D_{\phi_0}f_0{}^{b0}\left(\theta,\frac{\pi}{2},\phi,0\right)D_\phi f^{a0}\left(\frac{\pi}{2},\alpha,0,\beta\right) \\
& +\frac{1}{2ik}f^{b0}\left(\theta,\frac{\pi}{2},\phi,0\right)[D_\theta{}^2+D_\phi{}^2]f^{a0}\left(\frac{\pi}{2},\alpha,0,\beta\right)\Big\}+\mathcal{O}(d^{-3})\Bigg]
\end{aligned}
\tag{13}
$$

4. EXPLANATION OF THE RESULT

The expression (13) for the scattered field of the combination of two bodies appears, at first glance, to be complicated. However, it is not as abstruse as it seems. A closer examination of this expression along the lines followed in the two-dimensional case reveals the significance of the various terms and factors. We observe first that there are eight terms up to the order of accuracy d^{-2} calculated here, and we shall refer to them by number in the order of their appearance above. The first two terms represent single scattering while the others represent multiple scattering.

The functions f represent the complex scattering amplitude of the respective bodies with respect to an origin in the center of the relevant body, when the incident wave has zero phase at the center. The odd-numbered terms represent the fields ultimately scattered by body A, whereas the even-numbered terms represent fields ultimately scattered by B.

The factors $\exp[\pm ik\frac{1}{2}d\sin\theta\cos\phi]$ represent the phase differences for the scatterers as a result of the use of the normalized coordinate system, while the factors $\exp[\pm ik\frac{1}{2}d\sin\alpha\cos\beta]$ take into account the phase of the incident wave at the center of the respective scatterers. The f's containing a θ and ϕ dependence are scattering patterns, while the f's with specific values for θ and ϕ are excitation factors accumulated in the multiple scattering. We note, finally, that the factors of the form e^{iknd} where $n=0$, 1, 2 refer to the increase ikd in the phase of a wave in going from one scatterer to another, and that n signifies the number of bounces. If one uses the above comments as a guide, all the terms can be explained in a way similar to the explanation, given earlier, of the final expression for the scattered field in the two-dimensional case.

Brownian Motion of a Quantum Oscillator

JULIAN SCHWINGER*
Harvard University, Cambridge, Massachusetts
(Received November 28, 1960)

An action principle technique for the direct computation of expectation values is described and illustrated in detail by a special physical example, the effect on an oscillator of another physical system. This simple problem has the advantage of combining immediate physical applicability (e.g., resistive damping or maser amplification of a single electromagnetic cavity mode) with a significant idealization of the complex problems encountered in many-particle and relativistic field theory. Successive sections contain discussions of the oscillator subjected to external forces, the oscillator loosely coupled to the external system, an improved treatment of this problem and, finally, there is a brief account of a general formulation.

INTRODUCTION

THE title of this paper refers to an elementary physical example that we shall use to illustrate, at some length, a solution of the following methodological problem. The quantum action principle[1] is a differential characterization of transformation functions, $\langle a't_1|b't_2\rangle$, and thus is ideally suited to the practical computation of transition probabilities (which includes the determination of stationary states). Many physical questions do not pertain to individual transition probabilities, however, but rather to expectation values of a physical property for a specified initial state,

$$\langle X(t_1)\rangle_{b't_2}=\sum_{a'a''}\langle b't_2|a't_1\rangle\langle a't_1|X(t_1)|a''t_1\rangle\langle a''t_1|b't_2\rangle,$$

or, more generally, a mixture of states. Can one devise an action principle technique that is adapted to the direct computation of such expectation values, without requiring knowledge of the individual transformation functions?

The action principle asserts that ($\hbar=1$),

$$\delta\langle a't_1|b't_2\rangle=i\left\langle a't_1\left|\delta\left[\int_{t_2}^{t_1}dtL\right]\right|b't_2\right\rangle,$$

and

$$\delta\langle b't_2|a't_1\rangle=-i\left\langle b't_2\left|\delta\left[\int_{t_2}^{t_1}dtL\right]\right|a't_1\right\rangle,$$

in which we shall take $t_1>t_2$. These mutually complex-conjugate forms correspond to the two viewpoints whereby states at different times can be compared, either by progressing forward from the earlier time, or backward from the later time. The relation between the pair of transformation functions is such that

$$\delta[\sum_{a'}\langle b't_2|a't_1\rangle\langle a't_1|b''t_2\rangle]=0,$$

which expresses the fixed numerical value of

$$\langle b't_2|b''t_2\rangle=\delta(b',b'').$$

But now, imagine that the positive and negative senses of time development are governed by different dynamics. Then the transformation function for the closed circuit will be described by the action principle

$$\delta\langle t_2|t_2\rangle=\delta[\langle t_2|t_1\rangle\times\langle t_1|t_2\rangle]$$
$$=i\left\langle t_2\left|\delta\left[\int_{t_2}^{t_1}dtL_+-\int_{t_2}^{t_1}dtL_-\right]\right|t_2\right\rangle,$$

in which abbreviated notation the multiplication sign symbolizes the composition of transformation functions by summation over a complete set of states. If, in particular, the Lagrangian operators $L_\pm$ contain the dynamical term $\lambda_\pm(t)X(t)$, we have

$$\delta_\lambda\langle t_2|t_2\rangle=i\left\langle t_2\left|\int_{t_2}^{t_1}dt(\delta\lambda_+-\delta\lambda_-)X(t)\right|t_2\right\rangle,$$

and, therefore,

$$-i\frac{\delta}{\delta\lambda_+(t_1)}\langle t_2|t_2\rangle=i\frac{\delta}{\delta\lambda_-(t_1)}\langle t_2|t_2\rangle$$
$$=\langle t_2|X(t_1)|t_2\rangle,$$

where $\lambda_\pm$ can now be identified. Accordingly, if a system is suitably perturbed[2] in a manner that depends upon the time sense, a knowledge of the transformation function referring to a closed time path determines the expectation value of any desired physical quantity for a specified initial state or state mixture.

OSCILLATOR

To illustrate this remark we first consider an oscillator subjected to an arbitrary external force, as described by the Lagrangian operator

$$L=iy^\dagger(dy/dt)-\omega y^\dagger y-y^\dagger K(t)-yK^*(t),$$

* Supported by the Air Force Office of Scientific Research (ARDC).

[1] Some references are: Julian Schwinger, Phys. Rev. **82**, 914 (1951); **91**, 713 (1953); Phil. Mag. 44, 1171 (1953). The first two papers also appear in *Selected Papers on Quantum Electrodynamics* (Dover Publications, New York, 1958). A recent discussion is contained in Julian Schwinger, Proc. Natl. Acad. Sci. U. S. 46, 883 (1960).

[2] Despite this dynamical language, a change in the Hamiltonian operator of a system can be kinematical in character, arising from the consideration of another transformation along with the dynamical one generated by the Hamiltonian. See the last paper quoted in footnote 1, and Julian Schwinger, Proc. Natl. Acad. Sci. U. S. 46, 1401 (1960).

in which the complementary pair of non-Hermitian operators y, $iy^\dagger$, are constructed from Hermitian operators q, p by

$$y=2^{-\frac{1}{2}}(q+ip)$$
$$iy^\dagger=2^{-\frac{1}{2}}(p+iq).$$

The equations of motion implied by the action principle are

$$i(dy/dt)-\omega y=K$$
$$-i(dy^\dagger/dt)-\omega y^\dagger=K^*,$$

and solutions are given by

$$y(t)=e^{-i\omega(t-t_2)}y(t_2)-i\int_{t_2}^{t}dt'e^{-i\omega(t-t')}K(t'),$$

together with the adjoint equation. Since we now distinguish between the forces encountered in the positive time sense, $K_+(t)$, $K_+{}^*(t)$, and in the reverse time direction, $K_-(t)$, $K_-{}^*(t)$, the integral must be taken along the appropriate path. Thus, when t is reached first in the time evolution from t_2, we have

$$y_+(t)=e^{-i\omega(t-t_2)}y_+(t_2)-i\int_{t_2}^{t}dt'e^{-i\omega(t-t')}K_+(t'),$$

while on the subsequent return to time t,

$$y_-(t)=e^{-i\omega(t-t_2)}y_+(t_2)-i\int_{t_2}^{t_1}dt'e^{-i\omega(t-t')}K_+(t')+i\int_{t}^{t_1}dt'e^{-i\omega(t-t')}K_-(t').$$

Note that

$$y_-(t_1)-y_+(t_1)=0,$$
$$y_-(t_2)-y_+(t_2)=i\int_{t_2}^{t_1}dte^{i\omega(t-t_2)}(K_--K_+)(t).$$

We shall begin by constructing the transformation function referring to the lowest energy state of the unperturbed oscillator, $\langle 0t_2|0t_2\rangle^{K\pm}$. This state can be characterized by

$$\langle 0t_2|y^\dagger y(t_2)|0t_2\rangle=0$$

or, equivalently, by the eigenvector equations

$$y(t_2)|0t_2\rangle=0,\quad \langle 0t_2|y^\dagger(t_2)=0.$$

Since the transformation function simply equals unity if $K_+=K_-$ and $K_+{}^*=K_-{}^*$, we must examine the effect of independent changes in K_+ and K_-, and of $K_+{}^*$ and $K_-{}^*$, as described by the action principle

$$\delta_K\langle 0t_2|0t_2\rangle^{K\pm}=-i\left\langle 0t_2\left|\left[\int_{t_2}^{t_1}dt(\delta K_+{}^*y_+-\delta K_-{}^*y_-)+\int_{t_2}^{t_1}dt(y_+{}^\dagger\delta K_+-y_-{}^\dagger\delta K_-)\right]\right|0t_2\right\rangle^{K\pm}$$

The choice of initial state implies effective boundary conditions that supplement the equations of motion,

$$y_+(t_2)\to 0,\quad y_-{}^\dagger(t_2)\to 0.$$

Hence, in effect we have

$$y_+(t)=-i\int_{t_2}^{t_1}dt'e^{-i\omega(t-t')}\eta_+(t-t')K_+(t')$$

and

$$y_-(t)=-i\int_{t_2}^{t_1}dt'e^{-i\omega(t-t')}K_+(t')+i\int_{t_2}^{t_1}dt'e^{-i\omega(t-t')}\eta_-(t-t')K_-(t'),$$

together with the similar adjoint equations obtained by interchanging the $\pm$ labels. For convenience, step functions have been introduced:

$$\eta_+(t-t')=\begin{cases}1, & t-t'>0\\ 0, & t-t'<0\end{cases},$$

$$\eta_-(t-t')=\begin{cases}1, & t-t'<0\\ 0, & t-t'>0\end{cases},$$

$$\eta_+(t-t')+\eta_-(t-t')=1,\quad \eta_+(0)=\eta_-(0)=\tfrac{1}{2}.$$

We shall also have occasion to use the odd function

$$\epsilon(t-t')=\eta_+(t-t')-\eta_-(t-t').$$

The solution of the resulting integrable differential expression for $\log\langle 0t_2|0t_2\rangle^{K\pm}$ is given by

$$\langle 0t_2|0t_2\rangle^{K\pm}=\exp\left[-i\int_{t_2}^{t_1}dtdt'K^*(t)G_0(t-t')K(t')\right],$$

in a matrix notation, with

$$K(t)=\begin{pmatrix}K_+(t)\\ K_-(t)\end{pmatrix}$$

and

$$iG_0(t-t')=e^{-i\omega(t-t')}\begin{pmatrix}\eta_+(t-t') & 0\\ -1 & \eta_-(t-t')\end{pmatrix}.$$

The requirement that the transformation function reduce to unity on identifying K_+ with K_-, $K_+{}^*$ with $K_-{}^*$, is satisfied by the null sum of all elements of G_0, as assured by the property $\eta_++\eta_-=1$.

An operator interpretation of G_0 is given by the second variation

$$-\delta_{K^*}\delta_K\langle 0t_2|0t_2\rangle^{K\pm}|_{K=K^*=0}=i\int dtdt'\delta K^*(t)G_0(t-t')\delta K(t').$$

Generally, on performing two distinct variations in the structure of L that refer to parameters upon which

the dynamical variables at a given time are not explicitly dependent, we have

$$-\delta_1\delta_2\langle t_2|t_2\rangle=\Big\langle t_2\Big|\int_{t_2}^{t_1}dtdt'\{(\delta_1L_+(t)\delta_2L_+(t'))_+ -\delta_1L_-(t)\delta_2L_+(t')-\delta_2L_-(t)\delta_1L_+(t') +(\delta_1L_-(t)\delta_2L_-(t'))_-\}\Big|t_2\Big\rangle,$$

in which the multiplication order follows the sense of time development. Accordingly,

$$iG_0(t-t')\begin{pmatrix}\langle(y(t)y^\dagger(t'))_+\rangle_0 & -\langle y^\dagger(t')y(t)\rangle_0\\ -\langle y(t)y^\dagger(t')\rangle_0 & \langle(y(t)y^\dagger(t'))_-\rangle_0\end{pmatrix},$$

where the expectation values and operators refer to the lowest state and the dynamical variables of the unperturbed oscillator. The property of G_0 that the sum of rows and columns vanishes is here a consequence of the algebraic property

$$(y(t)y^\dagger(t'))_+ + (y(t)y^\dagger(t'))_- = \{y(t),y^\dagger(t')\}.$$

The choice of oscillator ground state is no essential restriction since we can now derive the analogous results for any initial oscillator state. Consider, for this purpose, the impulse forces

$$K_+(t)=iy''\delta(t-t_2),$$
$$K_-^*(t)=-iy^{\dagger\prime}\delta(t-t_2),$$

the effects of which are described by

$$y_+(t_2+0)-y_+(t_2)=y'',$$
$$y_-^\dagger(t_2+0)-y_-^\dagger(t_2)=y^{\dagger\prime}.$$

Thus, under the influence of these forces, the states $|0t_2\rangle$ and $\langle 0t_2|$ become, at the time t_2+0, the states $|y''t_2\rangle$ and $\langle y^{\dagger\prime}t_2|$, which are right and left eigenvectors, respectively, of the operators $y(t_2)$ and $y^\dagger(t_2)$. On taking into account arbitrary additional forces, the transformation function for the closed time path can be expressed as

$$\langle y^{\dagger\prime}t_1|y''t_2\rangle^{K\pm} =\exp\Big[y^{\dagger\prime}y''-y^{\dagger\prime}\Big(\int_{t_2}^{t_1}dtG_0(t_2-t)K(t)\Big) +\Big(\int_{t_2}^{t_1}dtK^*(t)G_0(t-t_2)\Big)_+y'' -i\int_{t_2}^{t_1}dtdt'K^*(t)G_0(t-t')K(t')\Big],$$

in which

$$\Big(\int_{t_2}^{t_1}dtG_0(t_2-t)K(t)\Big)=-i\int_{t_2}^{t_1}dte^{i\omega(t-t_2)}(K_--K_+)(t)$$

and

$$\Big(\int dtK^*(t)G_0(t-t_2)\Big)_+ =-i\int_{t_2}^{t_1}dte^{-i\omega(t-t_2)}(K_+^*-K_-^*)(t).$$

The eigenvectors of the non-Hermitian canonical variables are complete and have an intrinsic physical interpretation in terms of q and p measurements of optimum compatibility.[3] For our immediate purposes, however, we are more interested in the unperturbed oscillator energy states. The connection between the two descriptions can be obtained by considering the unperturbed oscillator transformation function

$$\langle y^{\dagger\prime}t_1|y''t_2\rangle=\langle y^{\dagger\prime}|\exp[-i(t_1-t_2)\omega y^\dagger y]|y''\rangle.$$

Now

$$i(\partial/\partial t_1)\langle y^{\dagger\prime}t_1|y''t_2\rangle=\langle y^{\dagger\prime}t_1|\omega y^\dagger(t_1)y(t_1)|y''t_2\rangle =\omega y^{\dagger\prime}e^{-i\omega(t_1-t_2)}y''\langle y^{\dagger\prime}t_1|y''t_2\rangle,$$

since

$$y(t_1)=e^{-i\omega(t_1-t_2)}y(t_2),$$

which yields

$$\langle y^{\dagger\prime}t_1|y''t_2\rangle=\exp[y^{\dagger\prime}e^{-i\omega(t_1-t_2)}y''] =\sum_{n=0}^{\infty}\frac{(y^{\dagger\prime})^n}{(n!)^{\frac12}}e^{-in\omega(t_1-t_2)}\frac{(y'')^n}{(n!)^{\frac12}}.$$

We infer the nonnegative integer spectrum of $y^\dagger y$, and the corresponding wave functions

$$\langle y^{\dagger\prime}|n\rangle=(y^{\dagger\prime})^n/(n!)^{\frac12},\quad \langle n|y'\rangle=(y')^n/(n!)^{\frac12}.$$

Accordingly, a non-Hermitian canonical variable transformation function can serve as a generator for the transformation function referring to unperturbed oscillator energy states,

$$\langle y^{\dagger\prime}t_2|y''t_2\rangle^{K\pm}=\sum_{n,n'=0}^{\infty}\frac{(y^{\dagger\prime})^n}{(n!)^{\frac12}}\langle nt_2|n't_2\rangle^{K\pm}\frac{(y'')^{n'}}{(n'!)^{\frac12}}.$$

If we are specifically interested in $\langle nt_2|nt_2\rangle^{K\pm}$, which supplies all expectation values referring to the initial state n, we must extract the coefficient of $(y^{\dagger\prime}y'')^n/n!$ from an exponential of the form

$$\exp[y^{\dagger\prime}y''+y^{\dagger\prime}\alpha+\beta y''+\gamma] =\sum_{kl}\frac{(y^{\dagger\prime})^k}{k!}\frac{(y'')^l}{l!}\alpha^k\beta^l\exp[y^{\dagger\prime}y''+\gamma].$$

All the terms that contribute to the required coefficient

[3] A discussion of non-Hermitian representations is given in *Lectures on Quantum Mechanics* (Les Houches, 1955), unpublished.

are contained in

$$\sum_{k=0}^{\infty}\frac{(y^{\dagger\prime}y^{\prime\prime})^k}{(k!)^2}(\alpha\beta)^k\exp[y^{\dagger\prime}y^{\prime\prime}+\gamma]$$
$$=\frac{1}{2\pi i}\oint\frac{d\lambda}{\lambda}e^{\lambda}\exp[y^{\dagger\prime}y^{\prime\prime}(1+\lambda^{-1}\alpha\beta)+\gamma],$$

where the latter version is obtained from

$$\frac{1}{2\pi i}\oint\frac{d\lambda}{\lambda^{k+1}}e^{\lambda}=\frac{1}{k!},$$

and

$$\langle nt_2|nt_2\rangle^{K\pm}=\exp\left[-i\int dtdt'K^*(t)G_0(t-t')K(t')\right]$$
$$\times L_n\left[\left(\int dtK^*(t)G_0(t-t_2)\right)_+\right.$$
$$\left.\times\left(\int dtG_0(t_2-t)K(t)\right)_-\right],$$

in which the nth Laguerre polynomial has been introduced on observing that

$$\frac{1}{2\pi i}\oint\frac{d\lambda}{\lambda}e^{\lambda}(1-\lambda^{-1}x)^n=\frac{1}{n!}e^x\left(\frac{d}{dx}\right)^n x^ne^{-x}=L_n(x).$$

One obtains a much neater form, however, from which these results can be recovered, on considering an initial mixture of oscillator energy states for which the nth state is assigned the probability

$$(1-e^{-\beta\omega})e^{-n\beta\omega},$$

and

$$\beta^{-1}=\vartheta$$

can be interpreted as a temperature. Then, since

$$(1-e^{-\beta\omega})\sum_{n=0}^{\infty}e^{-n\beta\omega}L_n(x)=(1-e^{-\beta\omega})\frac{1}{2\pi i}\oint\frac{d\lambda}{\lambda}$$
$$\times e^{\lambda}[1-e^{-\beta\omega}+\lambda^{-1}e^{-\beta\omega}x]^{-1}=\exp\left[-\frac{x}{e^{\beta\omega}-1}\right],$$

we obtain

$$\langle t_2|t_2\rangle_{\vartheta}^{K\pm}=\exp\left[-i\int_{t_2}^{t_1}dtdt'K^*(t)G_{\vartheta}(t-t')K(t')\right],$$

with

$$iG_{\vartheta}(t-t')=iG_0(t-t')+(e^{\beta\omega}-1)^{-1}G_0(t-t_2)_+{}_-G_0(t_2-t'),$$

and in which

$$iG_0(t-t_2)_+=e^{-i\omega(t-t_2)}\begin{pmatrix}1\\0\end{pmatrix},$$

$$i_-G_0(t_2-t)=e^{i\omega(t-t_2)}\begin{pmatrix}0&1\end{pmatrix}.$$

Thus,

$$iG_{\vartheta}(t-t')=e^{-i\omega(t-t')}\begin{pmatrix}\eta_+(t-t')+\langle n\rangle_{\vartheta}, & -\langle n\rangle_{\vartheta}\\ -1-\langle n\rangle_{\vartheta}, & \eta_-(t-t')+\langle n\rangle_{\vartheta}\end{pmatrix},$$

where we have written

$$\langle n\rangle_{\vartheta}=(e^{\beta\omega}-1)^{-1},$$

and since the elements of G_{ϑ} are also given by unperturbed oscillator thermal expectation values

$$iG_{\vartheta}(t-t')=\begin{pmatrix}\langle(y(t)y^{\dagger}(t'))_+\rangle_{\vartheta} & -\langle y^{\dagger}(t')y(t)\rangle_{\vartheta}\\ -\langle y(t)y^{\dagger}(t')\rangle_{\vartheta} & \langle(y(t)y^{\dagger}(t'))_-\rangle_{\vartheta}\end{pmatrix},$$

the designation $\langle n\rangle_{\vartheta}$ is consistent with its identification as $\langle y^{\dagger}y\rangle_{\vartheta}$.

The thermal forms can also be derived directly by solving the equations of motion, in the manner used to find $\langle 0t_2|0t_2\rangle^{K\pm}$. On replacing the single diagonal element

$$\langle 0t_2|0t_2\rangle^{K\pm}=\langle 0t_2|U|0t_2\rangle$$

by the statistical average

$$(1-e^{-\beta\omega})\sum_0^{\infty}e^{-n\beta\omega}\langle nt_2|nt_2\rangle^{K\pm}$$
$$=(1-e^{-\beta\omega})\,\mathrm{tr}[\exp(-\beta\omega y^{\dagger}y)U],$$

we find the following relation,

$$y_-(t_2)=e^{\beta\omega}y_+(t_2),$$

instead of the effective initial condition $y_+(t_2)=0$. This is obtained by combining

$$\exp(-\beta\omega y^{\dagger}y)y\exp(\beta\omega y^{\dagger}y)=\exp(\beta\omega)y$$

with the property of the trace

$$\mathrm{tr}[\exp(-\beta\omega y^{\dagger}y)yU]=\mathrm{tr}[\exp(\beta\omega)y\exp(-\beta\omega y^{\dagger}y)U]$$
$$=\mathrm{tr}[\exp(-\beta\omega y^{\dagger}y)U\exp(\beta\omega)y].$$

We also have

$$y_-(t_2)-y_+(t_2)=-i\int_{t_2}^{t_1}dte^{i\omega(t-t_2)}(K_+-K_-)(t),$$

and therefore, effectively,

$$y_+(t_2)=-i\frac{1}{e^{\beta\omega}-1}\int_{t_2}^{t_1}dte^{i\omega(t-t_2)}(K_+-K_-)(t).$$

Hence, to the previously determined $y_{\pm}(t)$ is to be added the term

$$-i\langle n\rangle_{\vartheta}\int_{t_2}^{t_1}dt'e^{-i\omega(t-t')}(K_+-K_-)(t'),$$

and correspondingly

$$\langle t_2|t_2\rangle_{\vartheta}^{K\pm}=\langle t_2|t_2\rangle_0^{K\pm}\exp\left[-\langle n\rangle_{\vartheta}\int_{t_2}^{t_1}dtdt'\right.$$
$$\left.\times(K_+{}^*-K_-{}^*)(t)e^{-i\omega(t-t')}(K_+-K_-)(t')\right],$$

which reproduces the earlier result.

As an elementary application let us evaluate the expectation value of the oscillator energy at time t_1 for a system that was in thermal equilibrium at time t_2 and is subsequently disturbed by an arbitrarily time-varying force. This can be computed as

$$\langle t_2|\omega y^\dagger y(t_1)|t_2\rangle_\vartheta{}^K$$
$$=\omega\frac{\delta}{\delta K_-(t_1)}\frac{\delta}{\delta K_+{}^*(t_1)}\langle t_2|t_2\rangle_\vartheta{}^{K^\pm}|_{K_+=K_-,\ K_+{}^*=K_-{}^*}.$$

The derivative $\delta/\delta K_+{}^*(t_1)$ supplies the factor

$$-i\left(\int_{t_2}^{t_1}dtG_\vartheta(t_1-t')K(t')\right)_+,$$

the subsequent variation with respect to $K_-(t_1)$ gives

$$-iG_\vartheta(0)_{+-}+\left(\int dtK^*(t)G_\vartheta(t-t_1)\right)_-$$
$$\times\left(\int dt'G_\vartheta(t_1-t')K(t')\right)_+,$$

and the required energy expectation value equals

$$\omega\langle n\rangle_\vartheta+\omega\left|\int_{t_2}^{t_1}dte^{i\omega t}K(t)\right|^2.$$

More generally, the expectation values of all functions of $y(t_1)$ and $y^\dagger(t_1)$ are known from that of

$$\exp\{-i[\lambda y^\dagger(t_1)+\mu y(t_1)]\},$$

and this quantity is obtained on supplementing K_+ and $K_+{}^*$ by the impulsive forces (note that in this use of the formalism a literal complex-conjugate relationship is not required)

$$K_+(t)=\lambda\delta(t-t_1),$$
$$K_+{}^*(t)=\mu\delta(t-t_1).$$

Then

$$\langle t_2|\exp\{-i[\lambda y^\dagger(t_1)+\mu y(t_1)]\}|t_2\rangle_\vartheta{}^K$$
$$=\exp\left[-\lambda\mu(\langle n\rangle_\vartheta+\tfrac{1}{2})+\lambda\int_{t_2}^{t_1}dte^{i\omega(t_1-t)}K^*(t)\right.$$
$$\left.-\mu\int_{t_2}^{t_1}dte^{-i\omega(t_1-t)}K(t)\right],$$

which involves the special step-function value

$$\eta_+(0)=\tfrac{1}{2}.$$

Alternatively, if we choose

$$K_+(t)=\lambda\delta(t-t_1),$$
$$K_+{}^*(t)=\mu\delta(t-t_1+0),$$

there appears

$$\langle t_2|\exp[-i\lambda y^\dagger(t_1)]\exp[-i\mu y(t_1)]|t_2\rangle_\vartheta{}^K$$
$$=\exp\left[-\lambda\mu\langle n\rangle_\vartheta+\lambda\int_{t_2}^{t_1}dte^{i\omega(t_1-t)}K^*(t)\right.$$
$$\left.-\mu\int_{t_2}^{t_1}dte^{-i\omega(t_1-t)}K(t)\right].$$

It may be worth remarking, in connection with these results, that the attention to expectation values does not deprive us of the ability to compute individual probabilities. Indeed, if probabilities for specific oscillator energy states are of interest, we have only to exhibit, as functions of y and $y^\dagger$, the projection operators for these states, the expectation values of which are the required probabilities. Now

$$P_n=|n\rangle\langle n|$$

is represented by the matrix

$$\langle y^{\dagger\prime}|P_n|y''\rangle=(y^{\dagger\prime}y'')^n/n!$$
$$=[(y^{\dagger\prime}y'')^n/n!]\exp(-y^{\dagger\prime}y'')\langle y^{\dagger\prime}|y''\rangle,$$

and, therefore,

$$P_n=\frac{1}{n!}(y^\dagger)^n\left[\sum_{k=0}^{\infty}\frac{(-1)^k}{k!}(y^\dagger)^k y^k\right]y^n$$
$$=\frac{1}{n!}(y^\dagger)^n\exp(-y^\dagger;y)y^n,$$

in which we have introduced a notation to indicate this ordered multiplication of operators. A convenient generating function for these projection operators is

$$\sum_{n=0}^{\infty}\alpha^nP_n=\exp[-(1-\alpha)y^\dagger;y],$$

and we observe that

$$\sum_0^\infty\alpha^nP_n=\exp\left[(1-\alpha)\frac{\partial}{\partial\lambda}\frac{\partial}{\partial\mu}\right]$$
$$\times\exp(-i\lambda y^\dagger)\exp(-i\mu y)|_{\lambda=\mu=0}.$$

Accordingly,

$$\sum_0^\infty\alpha^np(n,\vartheta,K)=\exp\left[(1-\alpha)\frac{\partial}{\partial\lambda}\frac{\partial}{\partial\mu}\right]\exp\left[-\lambda\mu\langle n\rangle_\vartheta\right.$$
$$+\lambda e^{i\omega t_1}\int dte^{-i\omega t}K^*(t)$$
$$\left.-\mu e^{-i\omega t_1}\int dte^{i\omega t}K(t)\right]\Big|_{\lambda=\mu=0}$$

gives the probability of finding the oscillator in the nth energy state after an arbitrary time-varying force

has acted, if it was initially in a thermal mixture of states.

To evaluate

$$X=\exp\left[(1-\alpha)\frac{\partial}{\partial\lambda}\frac{\partial}{\partial\mu}\right]\exp[-\lambda\mu\langle n\rangle+\lambda\gamma^*-\mu\gamma]|_{\lambda=\mu=0},$$

we first remark that

$$\frac{\partial}{\partial\gamma^*}X=\exp\left[(1-\alpha)\frac{\partial}{\partial\lambda}\frac{\partial}{\partial\mu}\right]\lambda\exp[\ ]|_{\lambda=\mu=0}$$

$$=(1-\alpha)\exp\left[(1-\alpha)\frac{\partial}{\partial\lambda}\frac{\partial}{\partial\mu}\right]\frac{\partial}{\partial\mu}\exp[\ ]|$$

$$=(1-\alpha)\exp\left[(1-\alpha)\frac{\partial}{\partial\lambda}\frac{\partial}{\partial\mu}\right](-\lambda\langle n\rangle-\gamma)\exp[\ ]|,$$

from which follows

$$\frac{\partial}{\partial\gamma^*}X=-\frac{\gamma(1-\alpha)}{1+\langle n\rangle(1-\alpha)}X$$

or

$$X=X_0\exp\left[-|\gamma|^2\frac{1-\alpha}{1+\langle n\rangle(1-\alpha)}\right].$$

Here

$$X_0=\exp\left[(1-\alpha)\frac{\partial}{\partial\lambda}\frac{\partial}{\partial\mu}\right]\exp[-\lambda\mu\langle n\rangle]|_{\lambda=\mu=0}$$

$$=[1+\langle n\rangle(1-\alpha)]^{-1},$$

as one shows with a similar procedure, or by direct series expansion. Therefore,

$$\sum_0^\infty\alpha^n p(n,\vartheta,K)=\frac{1-e^{-\beta\omega}}{1-\alpha e^{-\beta\omega}}\exp\left[-|\gamma|^2\frac{1-e^{-\beta\omega}}{1-\alpha e^{-\beta\omega}}(1-\alpha)\right],$$

where

$$|\gamma|^2=\left|\int dte^{i\omega t}K(t)\right|^2,$$

and on referring to the previously used Laguerre polynomial sum formula, we obtain

$$p(n,\vartheta,K)=(1-e^{-\beta\omega})e^{-n\beta\omega}\exp[-|\gamma|^2(1-e^{-\beta\omega})]\times L_n[-4|\gamma|^2\sinh^2(\beta\omega/2)].$$

In addition to describing the physical situation of initial thermal equilibrium, this result provides a generating function for the individual transition probabilities between oscillator energy states,

$$\sum_{n'=0}^\infty p(n,n',K)e^{-(n'-n)\beta\omega}$$

$$=\exp[-|\gamma|^2(1-e^{-\beta\omega})]L_n[-(1-e^{-\beta\omega})(e^{\beta\omega}-1)|\gamma|^2].$$

This form, and the implied transition probabilities, have already been derived in another connection,[4] and we shall only state the general result here:

$$p(n,n',K)=\frac{n_<!}{n_>!}(|\gamma|^2)^{n_>-n_<}[L_{n_<}^{(n_>-n_<)}(|\gamma|^2)]^2\times\exp(-|\gamma|^2),$$

in which $n_>$ and $n_<$ represent the larger and smaller of the two integers n and n'.

Another kind of probability is also easily identified, that referring to the continuous spectrum of the Hermitian operator

$$q=2^{-\frac{1}{2}}(y+y^\dagger)$$

[or $p=-2^{-\frac{1}{2}}i(y-y^\dagger)$]. For this purpose, we place $\lambda=\mu=-2^{-\frac{1}{2}}p'$ and obtain

$$\langle t_2|e^{ip'q(t_1)}|t_2\rangle_\vartheta{}^K=\exp[-\tfrac{1}{2}p'^2(\langle n\rangle_\vartheta+\tfrac{1}{2})+ip'\langle q(t_1)\rangle^K],$$

with

$$\langle q(t_1)\rangle^K=2^{-\frac{1}{2}}i\left[e^{i\omega t_1}\int_{t_2}^{t_1}dte^{-i\omega t}K^*(t)-e^{-i\omega t_1}\int_{t_2}^{t_1}dte^{i\omega t}K(t)\right].$$

If we multiply this result by $\exp(-ip'q')$ and integrate with respect to $p'/2\pi$ from $-\infty$ to ∞, we obtain the expectation value of $\delta[q(t_1)-q']$ which is the probability of realizing a value of $q(t_1)$ in a unit interval about q':

$$p(q't_1,\beta,K)=(\pi^{-1}\tanh\tfrac{1}{2}\beta\omega)^{\frac{1}{2}}\times\exp[-(\tanh\tfrac{1}{2}\beta\omega)(q'-\langle q(t_1)\rangle^K)^2].$$

Still another derivation of the formula giving thermal expectation values merits attention. Now we let the return path terminate at a different time $t_2'=t_2-T$, and on regarding the resulting transformation function as a matrix, compute the trace, or rather the trace ratio

$$\mathrm{tr}\langle t_2'|t_2\rangle^{K\pm}/\mathrm{tr}\langle t_2'|t_2\rangle,$$

which reduces to unity in the absence of external forces. The action principle again describes the dependence upon $K_\pm{}^*(t)$, $K_\pm(t)$ through the operators $y_\pm(t)$, $y_\pm{}^\dagger(t)$ which are related to the forces by the solutions of the equations of motion, and, in particular,

$$y_-(t_2')=e^{-i\omega(t_2'-t_2)}y_+(t_2)-i\int_{t_2}^{t_1}dte^{i\omega(t-t_2')}K_+(t)+i\int_{t_2'}^{t_1}dte^{i\omega(t-t_2')}K_-(t).$$

Next we recognize that the structure of the trace implies the effective boundary condition

$$y_-(t_2')=y_+(t_2).$$

[4] Julian Schwinger, Phys. Rev. **91**, 728 (1953).

Let us consider

$$\mathrm{tr}\langle t_2'|y_-(t_2')|t_2\rangle=\sum_{a'}\langle a't_2'|y_-(t_2')|a't_2\rangle,$$

where we require of the a representation only that it have no explicit time dependence. Then

$$\langle a't_2'|y_-(t_2')=\sum_{a''}\langle a'|y|a''\rangle\langle a''t_2'|$$

and

$$\mathrm{tr}\langle t_2'|y_-(t_2')|t_2\rangle=\sum_{a'a''}\langle a''t_2'|a't_2\rangle\langle a'|y|a''\rangle$$
$$=\mathrm{tr}\langle t_2'|y_+(t_2)|t_2\rangle,$$

which is the stated result.

The effective initial condition now appears as

$$y_+(t_2)=-\frac{1}{e^{-i\omega T}-1}i\left[\int_{t_2}^{t_1}dte^{i\omega(t-t_2)}K_+(t)-\int_{t_2'}^{t_1}dte^{i\omega(t-t_2)}K_-(t)\right],$$

and the action principle supplies the following evaluation of the trace ratio:

$$\exp\left[-i\int dtdt'K^*(t)G_0(t-t')K(t')\right]\times\exp\left[-(e^{-i\omega T}-1)^{-1}\left|\int dte^{i\omega t}(K_+-K_-)(t)\right|^2\right],$$

where the time variable in K_+ and K_- ranges from t_2 to t_1 and from t_2' to t_1, respectively. To solve the given physical problem we require that $K_-(t)$ vanish in the interval between t_2' and t_2 so that all time integrations are extended between t_2 and t_1. Then, since

$$\langle t_2'|=\langle t_2|e^{-i\omega(t_2'-t_2)n},\quad n=y^\dagger y(t_2),$$

what has been evaluated equals

$$\mathrm{tr}\langle t_2|e^{i\omega Tn}|t_2\rangle^{K\pm}/\mathrm{tr}\langle t_2|e^{i\omega Tn}|t_2\rangle,$$

and by adding the remark that this ratio continues to exist on making the complex substitution

$$-iT\rightarrow\beta>0,$$

the desired formula emerges as

$$\mathrm{tr}\langle t_2|e^{-\beta\omega n}|t_2\rangle^{K\pm}/\mathrm{tr}\langle t_2|e^{-\beta\omega n}|t_2\rangle=\exp\left[-i\int dtdt'K^*(t)G_\vartheta(t-t')K(t')\right].$$

EXTERNAL SYSTEM

This concludes our preliminary survey of the oscillator and we turn to the specific physical problem of interest: An oscillator subjected to prescribed external forces and loosely coupled to an essentially macroscopic external system. All oscillator interactions are linear in the oscillator variables, as described by the Lagrangian operator

$$L=iy^\dagger(dy/dt)-\omega_0y^\dagger y-y^\dagger K(t)-yK^*(t)-2^{\frac{1}{2}}qQ+L_{\mathrm{ext}},$$

in which L_{ext} characterizes the external system and $Q(t)$ is a Hermitian operator of that system.

We begin our treatment with a discussion of the transformation function $\langle t_2|t_2\rangle_{\vartheta_0\vartheta}{}^{K\pm}$ that refers initially to a thermal mixture at temperature ϑ for the external system, and to an independent thermal mixture at temperature ϑ_0 for the oscillator. The latter temperature can be interpreted literally, or as a convenient parametric device for obtaining expectation values referring to oscillator energy states. To study the effect of the coupling between the oscillator and the external system we supply the coupling term with a variable parameter λ, and compute

$$\frac{\partial}{\partial\lambda}\langle t_1|t_2\rangle^{K\pm}=-i\left\langle t_2\left|\int_{t_2}^{t_1}dt[2^{\frac{1}{2}}q_+(t)Q_+(t)-2^{\frac{1}{2}}q_-(t)Q_-(t)]\right|t_2\right\rangle^{K\pm}$$

where the distinction between the forward and return paths arises only from the application of different external forces $K_\pm(t)$ on the two segments of the closed time contour. The characterization of the external system as essentially macroscopic now enters through the assumption that this large system is only slightly affected by the coupling to the oscillator. In a corresponding first approximation, we would replace the operators $Q_\pm(t)$ by the effective numerical quantity $\langle Q(t)\rangle_\vartheta$. The phenomena that appear in this order of accuracy are comparatively trivial, however, and we shall suppose that

$$\langle Q(t)\rangle_\vartheta=0,$$

which forces us to proceed to the next approximation.

A second differentiation with respect to λ gives

$$-\frac{1}{2}\frac{\partial^2}{\partial\lambda^2}\langle t_2|t_2\rangle^{K\pm}=\left\langle t_2\left|\int_{t_2}^{t_1}dtdt'[(qQ(t)qQ(t'))_+-2q_-Q_-(t)q_+Q_+(t')+(qQ(t)qQ(t'))_-]\right|t_2\right\rangle^{K\pm}.$$

The introduction of an approximation based upon the slight disturbance of the macroscopic system converts this into

$$-\frac{1}{2}\frac{\partial^2}{\partial\lambda^2}\langle t_2|t_2\rangle^{K\pm}=\left\langle t_2\left|\int_{t_2}^{t_1}dtdt'[(y(t')y^\dagger(t))_+A_{++}(t-t')-y_-(t')y_+{}^\dagger(t)A_{+-}(t-t')-y_-{}^\dagger(t)y_+(t')A_{-+}(t-t')+(y(t')y^\dagger(t))_-A_{--}(t-t')]\right|t_2\right\rangle^{K\pm}$$

where

$$A(t-t')=\begin{pmatrix}\langle(Q(t)Q(t'))_+\rangle_\vartheta & \langle Q(t')Q(t)\rangle_\vartheta\\ \langle Q(t)Q(t')\rangle_\vartheta & \langle(Q(t)Q(t'))_-\rangle_\vartheta\end{pmatrix},$$

and we have also discarded all terms containing $y(t)y(t')$ and $y^\dagger(t)y^\dagger(t')$. The latter approximation refers to the assumed weakness of the coupling of the oscillator to the external system, for, during the many periods that are needed for the effect of the coupling to accumulate, quantities with the time dependence $e^{\pm i\omega_0(t+t')}$ will become suppressed in comparison with those varying as $e^{\pm i\omega_0(t-t')}$. At this point we ask what effective term in an action operator that refers to the closed time path of the oscillator would reproduce this approximate value of $(\partial/\partial\lambda)^2\langle t_2|t_2\rangle$ at $\lambda=0$. The complete action that satisfies this requirement, with λ^2 set equal to unity, is given by

$$W=\int_{t_2}^{t_1}dt\left[iy^\dagger\frac{dy}{dt}-\omega_0 y^\dagger y-y^\dagger K-yK^*\Big|_+-\Big|_-\right]$$
$$+i\int_{t_2}^{t_1}dtdt'[(y^\dagger(t)y(t'))_+A_{++}(t-t')$$
$$-y_-{}^\dagger(t)y_+(t')A_{-+}(t-t')-y_-(t')y_+{}^\dagger(t)A_{+-}(t-t')$$
$$+(y^\dagger(t)y(t'))_-A_{--}(t-t')].$$

The application of the principle of stationary action to this action operator yields equations of motion that are nonlocal in time, namely,

$$i\frac{dy_+}{dt}-\omega_0 y_+ +i\int_{t_2}^{t_1}dt'[A_{++}(t-t')y_+(t')$$
$$-A_{+-}(t-t')y_-(t')]=K_+(t)$$
$$i\frac{dy_-}{dt}-\omega_0 y_- -i\int_{t_2}^{t_1}dt'[A_{--}(t-t')y_-(t')$$
$$-A_{-+}(t-t')y_+(t')]=K_-(t),$$

together with

$$-i\frac{dy_+{}^\dagger}{dt}-\omega_0 y_+{}^\dagger+i\int_{t_2}^{t_1}dt'[y_+{}^\dagger(t')A_{++}(t'-t)$$
$$-y_-{}^\dagger(t')A_{-+}(t'-t)]=K_+{}^*(t)$$
$$-i\frac{dy_-{}^\dagger}{dt}-\omega_0 y_-{}^\dagger-i\int_{t_2}^{t_1}dt'[y_-{}^\dagger(t')A_{--}(t'-t)$$
$$-y_+{}^\dagger(t')A_{+-}(t'-t)]=K_-{}^*(t).$$

The latter set is also obtained by combining the formal adjoint operation with the interchange of the + and − labels attached to the operators and $K(t)$. Another significant form is conveyed by the pair of equations

$$\left(i\frac{d}{dt}-\omega_0\right)(y_--y_+)-i\int_{t_2}^{t_1}dt'(A_{--}-A_{+-})(t-t')$$
$$\times(y_--y_+)(t')=K_--K_+$$

and

$$\left(i\frac{d}{dt}-\omega_0\right)(y_++y_-)+i\int_{t_2}^{t_1}dt'(A_{++}-A_{+-})(t-t')$$
$$\times(y_++y_-)(t')-i\int_{t_2}^{t_1}dt'(A_{+-}+A_{-+})(t-t')$$
$$\times(y_--y_+)(t')=K_++K_-,$$

where

$$(A_{--}-A_{+-})(t-t')=-(A_{++}-A_{-+})(t-t')$$
$$=\langle[Q(t),Q(t')]\rangle_\vartheta\eta_-(t-t'),$$
$$(A_{++}-A_{+-})(t-t')=-(A_{--}-A_{-+})(t-t')$$
$$=\langle[Q(t),Q(t')]\rangle_\vartheta\eta_+(t-t'),$$

and

$$(A_{+-}+A_{-+})(t-t')=\langle\{Q(t),Q(t')\}\rangle_\vartheta.$$

The nonlocal character of these equations is not very marked if, for example, the correlation between $Q(t)$ and $Q(t')$ in the macroscopic system disappears when $|t-t'|$ is still small compared with the period of the oscillator. Then, since the behavior of $y(t)$ over a short time interval is given approximately by $e^{-i\omega t}$, the matrix $A(t-t')$ is effectively replaced by

$$\int_{-\infty}^{\infty}d(t-t')e^{i\omega(t-t')}A(t-t')=A(\omega),$$

and the equations of motion read

$$[i(d/dt)-\omega_-](y_--y_+)=K_--K_+,$$
$$[i(d/dt)-\omega_+](y_++y_-)-ia(y_--y_+)=K_++K_-.$$

Here we have defined

$$\omega_-=\omega_0+i(A_{--}-A_{+-})(\omega)=\omega+\tfrac{1}{2}i\gamma,$$
$$\omega_+=\omega_0-i(A_{++}-A_{+-})(\omega)=\omega-\tfrac{1}{2}i\gamma,$$

and

$$a(\omega)=(A_{+-}+A_{-+})(\omega).$$

It should be noted that $A_{+-}(\omega)$ and $A_{-+}(\omega)$ are real positive quantities since

$$A_{-+}(\omega)=\lim_{T\to\infty}\frac{1}{T}\left\langle\left(\int_{-\frac{1}{2}T}^{\frac{1}{2}T}dte^{-i\omega t}Q(t)\right)^\dagger\right.$$
$$\left.\times\left(\int_{-\frac{1}{2}T}^{\frac{1}{2}T}dte^{-i\omega t}Q(t)\right)\right\rangle$$

and

$$A_{+-}(\omega)=A_{-+}(-\omega).$$

One consequence is

$$a(\omega)=a(-\omega)\geq 0.$$

It also follows from

$$\omega_--\omega_+=i(A_{--}+A_{++}-2A_{+-})(\omega)$$
$$=i(A_{-+}-A_{+-})(\omega)$$

that

$$\gamma(\omega)=A_{-+}(\omega)-A_{+-}(\omega)$$
$$=-\gamma(-\omega)$$

is real. Furthermore

$$\omega=\omega_0-\tfrac{1}{2}i(A_{++}-A_{--})(\omega),$$

where

$$(A_{++}-A_{--})(t-t')=\langle[Q(t),Q(t')]\rangle_\vartheta\epsilon(t-t')$$
$$=(A_{-+}-A_{+-})(t-t')\epsilon(t-t'),$$

so that

$$-i(A_{++}-A_{--})(\omega)=\frac{1}{\pi}P\int_{-\infty}^{\infty}\frac{d\omega'}{\omega-\omega'}\gamma(\omega'),$$

and ω emerges as the real quantity

$$\omega=\omega_0-\frac{1}{\pi}P\int_0^{\infty}\frac{\omega' d\omega'}{\omega'^2-\omega^2}\gamma(\omega').$$

We have not yet made direct reference to the nature of the expectation value for the macroscopic system, which is now taken as the thermal average:

$$\langle X\rangle_\vartheta = C\,\mathrm{tr}e^{-\beta H}X$$
$$C^{-1}=\mathrm{tr}e^{-\beta H},$$

where H is the energy operator of the external system. The implication for the structure of the expectation values is contained in

$$\langle Q(t)Q(t')\rangle_\vartheta = C\,\mathrm{tr}e^{-\beta H}Q(t)Q(t')$$
$$=\langle Q(t')Q(t+i\beta)\rangle_\vartheta,$$

which employs the formal property

$$e^{-\beta H}Q(t)e^{\beta H}=Q(t+i\beta).$$

On introducing the time Fourier transforms, however, this becomes the explicit relation

$$A_{-+}(\omega)=e^{\beta\omega}A_{+-}(\omega),$$

and we conclude that

$$e^{-\frac{1}{2}\beta\omega}A_{-+}(\omega)=e^{\frac{1}{2}\beta\omega}A_{+-}(\omega)$$
$$=a(\omega)/2\cosh\tfrac{1}{2}\beta\omega,$$

which is a positive even function of ω. As a consequence, we have

$$\gamma(\omega)=a(\omega)\tanh\tfrac{1}{2}\beta\omega,$$
$$\geq 0,\quad \beta\omega>0,$$

which can also be written as

$$a(\omega)=2\gamma(\omega)[(e^{\beta\omega}-1)^{-1}+\tfrac{1}{2}].$$

The net result of this part of the discussion is to remove all explicit reference to the external system as a dynamical entity. We are given effective equations of motion for y_+ and y_- that contain the prescribed external forces and three parameters, the angular frequency $\omega(\simeq\omega_0)$, γ, and a, the latter pair being related by the temperature of the macroscopic system. The accompanying boundary conditions are

$$(y_--y_+)(t_1)=0$$

and, for the choice of an initial thermal mixture,

$$y_-(t_2)=e^{\beta_0\omega}y_+(t_2).$$

We now find that

$$(y_--y_+)(t)=i\int_{t_2}^{t_1}dt'e^{i\omega-(t'-t)}\eta_-(t-t')(K_--K_+)(t'),$$

which supplies the initial condition for the second equation of motion,

$$(y_++y_-)(t_2)=\coth(\tfrac{1}{2}\beta_0\omega)i\int_{t_2}^{t_1}dte^{i\omega-(t-t_2)}(K_--K_+)(t),$$

and the required solution is given by

$$i(y_++y_-)(t)$$
$$=\int_{t_2}^{t_1}dt'e^{-i\omega+(t-t')}\eta_+(t-t')(K_++K_-)(t')$$
$$-\coth(\tfrac{1}{2}\beta\omega)\int_{t_2}^{t_1}dt'[e^{-i\omega+(t-t')}\eta_+(t-t')$$
$$+e^{i\omega-(t'-t)}\eta_-(t-t')](K_--K_+)(t')$$
$$+(\coth\tfrac{1}{2}\beta\omega-\coth\tfrac{1}{2}\beta_0\omega)\int_{t_2}^{t_1}dt'e^{-i\omega+(t-t_2)}$$
$$\times e^{i\omega-(t'-t_2)}(K_--K_+)(t').$$

The corresponding solutions for $y_\pm^\dagger(t)$ are obtained by interchanging the $\pm$ labels in the formal adjoint equation.

The differential dependence of the transformation function $\langle t_2|\,t_2\rangle_{\vartheta_0\vartheta}{}^{K\pm}$ upon the external forces is described by these results, and the explicit formula obtained on integration is

$$\langle t_2|\,t_2\rangle_{\vartheta_0\vartheta}{}^{K\pm}$$
$$=\exp\left[-i\int dtdt'K^*(t)G_{\vartheta_0\vartheta}(t-t_2,\,t'-t_2)K(t')\right],$$

where $[n_0=\langle n\rangle_{\vartheta_0},\ n=\langle n\rangle_\vartheta]$

$$iG_{\vartheta_0\vartheta}(t-t_2,\,t'-t_2)$$
$$=e^{-i\omega+(t-t')}\eta_+(t-t')\begin{pmatrix}n+1, & -n\\ -n-1, & n\end{pmatrix}$$
$$+e^{-i\omega-(t-t')}\eta_-(t-t')\begin{pmatrix}n, & -n\\ -n-1, & n+1\end{pmatrix}$$
$$+e^{-i\omega+(t-t_2)}e^{i\omega-(t'-t_2)}(n_0-n)\begin{pmatrix}1 & -1\\ -1 & 1\end{pmatrix}.$$

Another way of presenting this result is

$$iG_{\vartheta_0\vartheta}(t-t_2,\,t'-t_2)$$
$$=e^{-i\omega(t-t')}e^{-\frac{1}{2}\gamma|t-t'|}\begin{pmatrix}\eta_+(t-t')+n, & -n\\ -n-1, & \eta_-(t-t')+n\end{pmatrix}$$
$$+e^{-i\omega(t-t')}e^{-\gamma[\frac{1}{2}(t+t')-t_2]}(n_0-n)\begin{pmatrix}1 & -1\\ -1 & 1\end{pmatrix},$$

although the simplest description of G is supplied by

the differential equation

$$\left[\left(i\frac{d}{dt}-\omega_+\right)G-\delta(t-t')\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\right]\left(-i\frac{d^T}{dt'}-\omega_-\right)$$
$$=-i\delta(t-t')\gamma\begin{pmatrix}n, & -n\\ -n-1, & n+1\end{pmatrix},$$

(where d^T indicates differentiation to the left) in conjunction with the initial value

$$iG_{\vartheta_0\vartheta}(0,0)=\begin{pmatrix}n_0+\frac{1}{2}, & -n_0\\ -n_0-1, & n_0+\frac{1}{2}\end{pmatrix}$$

and the boundary conditions

$$[i(d/dt)-\omega_+]G=0, \quad t>t'$$
$$[i(d/dt)-\omega_-]G=0, \quad t<t'.$$

A more symmetrical version of this differential equation is given by

$$\left(i\frac{d}{dt}-\omega_+\right)\left(-i\frac{d}{dt'}-\omega_-\right)G$$
$$-\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\left(i\frac{d}{dt}-\omega\right)\delta(t-t')$$
$$=-i\delta(t-t')\gamma\begin{pmatrix}n+\frac{1}{2}, & -n\\ -n-1, & n+\frac{1}{2}\end{pmatrix}.$$

We note the vanishing sum of all G elements, and that the role of complex conjugation in exchanging the two segments of the closed time path is expressed by

$$-\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}G(t',t)^{T*}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}=G(t,t'),$$

which is to say that

$$-G(t',t)_{+-}{}^*=G(t,t')_{+-}, \quad -G(t',t)_{-+}{}^*=G(t,t')_{-+}$$
$$-G(t',t)_{--}{}^*=G(t,t')_{++}.$$

It will be observed that only when

$$\langle n\rangle_{\vartheta_0}=\langle n\rangle_{\vartheta}$$

is $G_{\vartheta_0\vartheta}(t-t_2, t'-t_2)$ independent of t_2 and a function of $t-t'$. This clearly refers to the initial physical situation of thermal equilibrium between the oscillator and the external system at the common temperature $\vartheta_0=\vartheta>0$, which equilibrium persists in the absence of external forces. If the initial circumstances do not constitute thermal equilibrium, that will be established in the course of time at the macroscopic temperature $\vartheta>0$. Thus, all reference to the initial oscillator temperature disappears from $G_{\vartheta_0\vartheta}(t-t_2, t'-t_2)$ when, for fixed $t-t'$,

$$\gamma[\tfrac{1}{2}(t+t')-t_2]\gg 1.$$

The thermal relaxation of the oscillator energy is derived from

$$\langle t_2|y^\dagger y(t_1)|t_2\rangle_{\vartheta_0\vartheta}=\frac{\delta}{\delta K_-(t_1)}\frac{\delta}{\delta K_+{}^*(t_1)}\langle t_2|t_2\rangle_{\vartheta_0\vartheta}{}^{K\pm}\Big|_{K_\pm=0}$$
$$=-iG_{\vartheta_0\vartheta}(t_1-t_2, t_1-t_2)_{+-},$$

and is expressed by

$$\langle n(t_1)\rangle=\langle n\rangle_\vartheta+(\langle n\rangle_{\vartheta_0}-\langle n\rangle_\vartheta)e^{-\gamma(t_1-t_2)}.$$

The previously employed technique of impulsive forces applied at the time t_1 gives the more general result

$$\langle t_2|\exp[-i(\lambda y^\dagger(t_1)+\mu y(t_1))]|t_2\rangle_{\vartheta_0\vartheta}{}^K$$
$$=\exp\left[-\lambda\mu(\langle n(t_1)\rangle+\tfrac{1}{2})+\lambda\int_{t_2}^{t_1}dte^{i\omega-(t_1-t)}K^*(t)\right.$$
$$\left.-\mu\int_{t_2}^{t_1}dte^{-i\omega+(t_1-t)}K(t)\right],$$

from which a variety of probability distributions and expectation values can be obtained.

The latter calculation illustrates a general characteristic of the matrix $G(t,t')$, which is implied by the lack of dependence on the time t_1. Indeed, such a terminal time need not appear explicitly in the structure of the transformation function $\langle t_2|t_2\rangle^{K\pm}$ and all time integrations can range from t_2 to $+\infty$. Then t_1 is implicit as the time beyond which K_+ and K_- are identified, and the structure of G must be such as to remove any reference to a time greater than t_1. In the present situation, the use of an impulsive force at t_1 produces, for example, the term

$$\int_{t_2}^{\infty}dtG(t_1-t_2, t-t_2)K(t),$$

in which K_+ and K_- are set equal. Hence it is necessary that

$$G(t,t')\begin{pmatrix}1\\ 1\end{pmatrix}=0, \quad t<t'$$

and similarly that

$$\overbrace{1 \quad 1}\ G(t,t')=0, \quad t>t',$$

which says that adding the columns of $G(t,t')$ gives retarded functions of $t-t'$, while the sum of rows supplies a vector that is an advanced function of $t-t'$. In each instance, the two components must have a zero sum. These statements are immediately verified for the explicitly calculated $G_{\vartheta_0\vartheta}(t-t_2, t'-t_2)$ and follow more generally from the operator construction

$$iG=\begin{pmatrix}\langle(y(t)y^\dagger(t'))_+\rangle, & -\langle y^\dagger(t')y(t)\rangle\\ -\langle y(t)y^\dagger(t')\rangle, & \langle(y(t)y^\dagger(t'))_-\rangle\end{pmatrix},$$

for, as we have already noted in connection with Q products,

$$(y(t)y^\dagger(t'))_+-y^\dagger(t')y(t)=y(t)y^\dagger(t')-(y(t)y^\dagger(t'))_-$$
$$=\eta_+(t-t')[y(t),y^\dagger(t')]$$

and

$$(y(t)y^\dagger(t'))_+ - y(t)y^\dagger(t') = y^\dagger(t')y(t) - (y(t)y^\dagger(t'))_- \\ = -\eta_-(t-t')[y(t), y^\dagger(t')].$$

Our results show, incidentally, that

$$\langle [y(t), y^\dagger(t')]\rangle_{\vartheta_0\vartheta} = e^{-i\omega_+(t-t')}\eta_+(t-t') + e^{-i\omega_-(t-t')}\eta_-(t-t') \\ = e^{-i\omega(t-t')}e^{-\frac{1}{2}\gamma|t-t'|}.$$

Another general property can be illustrated by our calculation, the positiveness of $-iG(t,t')_{+-}$,

$$-\int dtdt' K(t) iG(t-t_2, t'-t_2)_{+-} K^*(t') \\ = \left\langle t_2 \middle| \left(\int dt K(t) y(t)\right)^\dagger \left(\int dt K(t) y(t)\right) \middle| t_2 \right\rangle > 0.$$

We have found that

$$-iG_{\vartheta_0\vartheta}(t-t_2, t'-t_2)_{+-} \\ = \exp\{-i\omega(t-t') - \gamma[\tfrac{1}{2}(t+t') - t_2]\}\langle n\rangle_\vartheta \\ + e^{-i\omega(t-t')}[e^{-\frac{1}{2}\gamma|t-t'|} - e^{-\gamma[\frac{1}{2}(t+t')-t_2]}]\langle n\rangle_{\vartheta_0},$$

and it is clearly necessary that each term obey separately the positiveness requirement. The first term is trivial,

$$\int dtdt' K(t) \exp\{-i\omega(t-t') - \gamma[\tfrac{1}{2}(t+t') - t_2]\} K^*(t') \\ = \left| \int dt e^{-i\omega_+(t-t_2)} K(t) \right|^2 > 0,$$

and the required property of the second term follows from the formula

$$e^{-\frac{1}{2}\gamma|t-t'|} - e^{-\gamma[\frac{1}{2}(t+t')-t_2]} \\ = \frac{2\gamma}{\pi} \int_0^\infty d\omega' \frac{\sin\omega'(t-t_2)\,\sin\omega'(t'-t_2)}{\omega'^2 + (\frac{1}{2}\gamma)^2}.$$

All the information that has been obtained about the oscillator is displayed on considering the forces

$$K_\pm(t) = \lambda_\pm(t) + K(t), \quad K_\pm^*(t) = \mu_\pm(t) + K^*(t),$$

and making explicit the effects of $\lambda_\pm(t)$, $\mu_\pm(t)$ by equivalent time-ordered operators:

$$\left\langle t_2 \middle| \left(\exp\left[i\int_{t_2}^\infty dt(\lambda_- y^\dagger + \mu_- y)\right]\right)_- \right. \\ \left. \times \left(\exp\left[-i\int_{t_2}^\infty dt(\lambda_+ y^\dagger + \mu_+ y)\right]\right)_+ \middle| t_2 \right\rangle_{\vartheta_0\vartheta}^K \\ = \exp\left[-i\int dtdt' \mu(t) G(t-t_2, t'-t_2)\lambda(t') \right. \\ + \int dtdt' K^*(t) e^{-i\omega_-(t-t')}\eta_-(t-t')(\lambda_+ - \lambda_-)(t') \\ \left. - \int dtdt' (\mu_+ - \mu_-)(t) e^{-i\omega_+(t-t')}\eta_+(t-t') K(t') \right].$$

This is a formula for the direct computation of expectation values of general functions of $y(t)$ and $y^\dagger(t)$. A less explicit but simpler result can also be given by means of expectation values for functions of the operators

$$[i(d/dt) - \omega_+]y(t) - K(t) = K_f(t), \\ [-i(d/dt) - \omega_-]y^\dagger(t) - K^*(t) = K_f^\dagger(t).$$

Let us recognize at once that

$$\langle K_f(t)\rangle = 0, \quad \langle K_f^\dagger(t)\rangle = 0,$$

and therefore that the fluctuations of $y(t)$, $y^\dagger(t)$ can be ascribed to the effect of the forces K_f, $K_f^\dagger$, which appear as the quantum analogs of the random forces in the classical Langevin approach to the theory of the Brownian motion. The change in viewpoint is accomplished by introducing

$$\lambda_\pm(t) = [i(d/dt) - \omega_-]u_\pm(t) \\ \mu_\pm(t) = [-i(d/dt) - \omega_+]v_\pm(t),$$

where we assume, just for simplicity, that the functions $u(t)$, $v(t)$ vanish at the time boundaries. Then, partial time integrations will replace the operators y, $y^\dagger$ with K_f, $K_f^\dagger$.

To carry this out, however, we need the following lemma on time-ordered products:

$$\left(\exp\left[\int dt(A(t) + (d/dt)B(t))\right]\right)_+ \\ = \left(\exp\left[\int dt A(t)\right]\right)_+ \exp\left(\int dt[A + \tfrac{1}{2}(dB/dt), B]\right),$$

which involves the unessential assumption that $B(t)$ vanishes at the time terminals, and the hypothesis that $[A(t), B(t)]$ and $[dB(t)/dt, B(t)]$ are commutative with all the other operators. The proof is obtained by replacing $B(t)$ with $\lambda B(t)$ and differentiating with respect to λ,

$$\frac{\partial}{\partial\lambda}\left(\exp\left[\int_{t_2}^{t_1} dt\left(A + \lambda\frac{d}{dt}B\right)\right]\right)_+ \\ = \int_{t_2}^{t_1} dt\left(\exp\left[\int_t^{t_1}\right]\right)_+ \frac{d}{dt}B(t)\left(\exp\left[\int_{t_2}^{t}\right]\right)_+.$$

Then, a partial integration yields

$$\int_{t_2}^{t_1} dt\left(\exp\left[\int_t^{t_1}\right]\right)_+ \left[A(t) + \lambda\frac{dB(t)}{dt}, B(t)\right] \\ \times \left(\exp\left[\int_{t_2}^{t}\right]\right)_+ \\ = \left(\exp\left[\int_{t_2}^{t_1}\right]\right)_+ \int_{t_2}^{t_1} dt\left[A + \lambda\frac{dB}{dt}, B\right],$$

according to the hypothesis, and the stated result follows on integrating this differential equation.

The structure of the lemma is given by the rearrangement

$$-i(\lambda y^\dagger+\mu y)=-i[u(K^*+K_f^\dagger)+v(K+K_f)] +(d/dt)(uy^\dagger-vy),$$

and we immediately find a commutator that is a multiple of the unit operator,

$$[A+(d/dt)B, B]=-i[\lambda y^\dagger+\mu y, uy^\dagger-vy] =-i(\mu u+\lambda v)\rightarrow -2iv(i(d/dt)-\omega)u.$$

The last form involves discarding a total time derivative that will not contribute to the final result. To evaluate $[A,B]$ we must refer to the meaning of K_f and $K_f^\dagger$ that is supplied by the actual equations of motion,

$$K_f(t)=Q(t)+(\omega_0-\omega_+)y(t)$$
$$K_f^\dagger(t)=Q(t)+(\omega_0-\omega_-)y^\dagger(t),$$

for then

$$[A(t),B(t)]=-i[u(\omega_0-\omega_-)y^\dagger+v(\omega_0-\omega_+)y, uy^\dagger-vy] =2ivu(\omega-\omega_0),$$

which is also proportional to the unit operator. Accordingly,

$$\left(\exp\left[-i\int dt(\lambda y^\dagger+\mu y)\right]\right)_+ =\left(\exp\left[-i\int dt[u(K^*+K_f^\dagger)+v(K+K_f)]\right]\right)_+ \times\exp\left[i(\omega-\omega_0)\int dtvu-i\int dtv\left(i\frac{d}{dt}-\omega\right)u\right],$$

and complex conjugation yields the analogous result for negatively time-ordered products.

With the aid of the differential equation obeyed by G, we now get

$$\left\langle t_2\left|\left(\exp\left[i\int dt(uK_f^\dagger+vK_f)\right]\right)_- \times\left(\exp\left[-i\int dt(uK_f^\dagger+vK_f)\right]\right)_+\right|t_2\right\rangle_{\vartheta_0\vartheta}^{K} =\exp\left[-i\int_{t_2}^{t_1}dtv(t)\kappa u(t)\right],$$

where

$$\kappa=\gamma\begin{pmatrix} n+\frac{1}{2}, & -n \\ -n-1, & n+\frac{1}{2}\end{pmatrix}+i(\omega-\omega_0)\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}.$$

The elements of this matrix are also expressed by

$$\kappa\delta(t-t')=\begin{pmatrix}\langle(K_f(t)K_f^\dagger(t'))_+\rangle, & -\langle K_f^\dagger(t')K_f(t)\rangle \\ -\langle K_f(t)K_f^\dagger(t')\rangle, & \langle(K_f(t)K_f^\dagger(t'))_-\rangle\end{pmatrix}.$$

Such expectation values are to be understood as effective evaluations that serve to describe the properties of the oscillator under the circumstances that validate the various approximations that have been used.

It will be observed that when n is sufficiently large to permit the neglect of all other terms,

$$\kappa\backsimeq\tfrac{1}{2}a\begin{pmatrix}1 & -1\\ -1 & 1\end{pmatrix}\quad[\tfrac{1}{2}a=\gamma(n+\tfrac{1}{2})],$$

and the sense of operator multiplication is no longer significant. This is the classical limit, for which

$$\left\langle\exp\left[-i\int dt(uK_f^\dagger+vK_f)\right]\right\rangle_\vartheta =\exp\left[-\int dt\tfrac{1}{2}av(t)u(t)\right],$$

where we have placed $u_+-u_-=u$, $v_+-v_-=v$. On introducing real components of the random force

$$K_f=2^{-\frac{1}{2}}(K_1+iK_2),\quad K_f^\dagger=2^{-\frac{1}{2}}(K_1-iK_2),$$

the classical limiting result reads

$$\left\langle\exp\left[-i\int dt(u_1K_1+u_2K_2)\right]\right\rangle_\vartheta =\exp\left[-\int dt\tfrac{1}{4}a(u_1^2+u_2^2)\right].$$

The fluctuations at different times are independent. If we consider time-averaged forces,

$$\bar{K}=\frac{1}{\Delta t}\int_t^{t+\Delta t}dt'K(t'),$$

we find by Fourier transformation that

$$\langle\delta(\bar{K}_1-K_1')\delta(\bar{K}_2-K_2')\rangle_\vartheta=\frac{\Delta t}{\pi a}\exp\left[-\frac{\Delta t}{a}(K_1'^2+K_2'^2)\right],$$

which is the Gaussian distribution giving the probability that the force averaged over a time interval Δt will have a value within a small neighborhood of the point K'. In this classical limit the fluctuation constant a is related to the damping or dissipation constant γ and the macroscopic temperature ϑ by

$$a=(2\gamma/\omega)\vartheta.$$

Our simplified equations can also be applied to situations in which the external system is not at thermal equilibrium. To see this possibility let us return to the real positive functions $A_{-+}(\omega)$, $A_{+-}(\omega)$ that describe the external system and remark that, generally,

$$\frac{A_{-+}(-\omega)}{A_{+-}(-\omega)}=\left[\frac{A_{-+}(\omega)}{A_{+-}(\omega)}\right]^{-1}\geq 0.$$

These properties can be expressed by writing

$$A_{-+}(\omega)/A_{+-}(\omega)=e^{\omega\beta(\omega)},$$

where $\beta(\omega)$ is a real even function that can range from $-\infty$ to $+\infty$. When only one value of ω is of interest, all conceivable situations for the external system can be described by the single parameter β, the reciprocal of which appears as an effective temperature of the macroscopic system. A new physical domain that appears in this way is characterized by negative temperature, $\beta<0$. Since a is an intrinsically positive constant, it is γ that will reverse sign

$$-\gamma=a(1-e^{-|\beta|\omega})/(1+e^{-|\beta|\omega})>0,$$

and the effect of the external system on the oscillator changes from damping to amplification.

We shall discuss the following physical sequence. At time t_2 the oscillator, in a thermal mixture of states at temperature ϑ_0, is acted on by external forces which are present for a time, short in comparison with $1/|\gamma|$. After a sufficiently extended interval $\sim(t_1-t_2)$ such that the amplification factor or gain is very large,

$$k=e^{\frac{1}{2}|\gamma|(t_1-t_2)}\gg 1,$$

measurements are made in the neighborhood of time t_1. A prediction of all such measurements is contained in the general expectation value formula. Approximations that convey the physical situation under consideration are given by

$$\int dtdt'(\mu_+-\mu_-)(t)e^{-i\omega+(t-t')}\eta_+(t-t')K(t')$$
$$\simeq k\int dt(\mu_+-\mu_-)(t)e^{-i\omega t}\int dt'e^{i\omega t'}K(t'),$$

$$\int dtdt'K^*(t)e^{-i\omega-(t-t')}\eta_-(t-t')(\lambda_+-\lambda_-)(t')$$
$$\simeq k\int dtK^*(t)e^{-i\omega t}\int dt'e^{i\omega t'}(\lambda_+-\lambda_-)(t'),$$

and

$$i\int dtdt'\mu(t)G(t-t_2,t'-t_2)\lambda(t')$$
$$\simeq k^2(\langle n\rangle_{\vartheta_0}+(1-e^{-|\beta|\omega})^{-1})\int dt(\mu_+-\mu_-)(t)e^{-i\omega t}$$
$$\times\int dt'(\lambda_+-\lambda_-)(t')e^{i\omega t'}.$$

From the appearance of the combinations $\mu_+-\mu_-=\mu$, $\lambda_+-\lambda_-=\lambda$ only, we recognize that noncommutativity of operator multiplication is no longer significant, and thus the motion of the oscillator has been amplified to the classical level. To express the consequences most simply, we write

$$y(t)=ke^{-i\omega t}(y_s+y_n)$$
$$y^\dagger(t)=ke^{i\omega t}(y_s{}^*+y_n{}^*),$$

with

$$y_s=-i\int_{t_2}^{\infty}dt'e^{i\omega t'}K(t'),$$

and, on defining

$$u=k\int dte^{i\omega t}\lambda(t),\quad v=k\int dte^{-i\omega t}\mu(t),$$

we obtain the time-independent result

$$\langle\exp[-i(uy_n{}^*+vy_n)]\rangle$$
$$=\exp[-(\langle n\rangle_{\vartheta_0}+(1-e^{-|\beta|\omega})^{-1})vu],$$

which implies that

$$\langle y_n\rangle=\langle y_n{}^*\rangle=0$$
$$\langle|y_n|^2\rangle=\langle n\rangle_{\vartheta_0}+(1-e^{-|\beta|\omega})^{-1}\geq\langle n\rangle_{\vartheta_0}+1.$$

Thus, the oscillator coordinate $y(t)$ is the amplified superposition of two harmonic terms, one of definite amplitude and phase (signal), the other with random amplitude and phase (noise), governed by a two-dimensional Gaussian probability distribution.

These considerations with regard to amplification can be viewed as a primitive model of a maser device,[5] with the oscillator corresponding to a single mode of a resonant electromagnetic cavity, and the external system to an atomic ensemble wherein, for a selected pair of levels, the thermal population inequality is reversed by some means such as physical separation or electromagnetic pumping.

AN IMPROVED TREATMENT

In this section we seek to remove some of the limitations of the preceding discussion. To aid in dealing successfully with the nonlocal time behavior of the oscillator, it is convenient to replace the non-Hermitian operator description with one employing Hermitian operators. Accordingly, we begin the development again, now using the Lagrangian operator

$$L=p(dq/dt)-\tfrac{1}{2}(p^2+\omega_0{}^2q^2)+qF(t)+qQ+L_{\text{ext}},$$

where Q has altered its meaning by a constant factor. One could also include an external prescribed force that is coupled to p. We repeat the previous approximate construction of the transformation function $\langle t_2|t_2\rangle_{\vartheta_0}{}^{F\pm}$ which proceeds by the introduction of an effective action operator that retains only the simplest correlation aspects of the external system, as comprised in

$$A(t-t')=\begin{pmatrix}\langle(Q(t)Q(t'))_+\rangle_\vartheta & \langle Q(t')Q(t)\rangle_\vartheta\\ \langle Q(t)Q(t')\rangle_\vartheta & \langle(Q(t)Q(t'))_-\rangle_\vartheta\end{pmatrix}.$$

[5] A similar model has been discussed recently by R. Serber and C. H. Townes, *Symposium on Quantum Electronics* (Columbia University Press, New York, 1960).

The action operator, with no other approximations, is

$$W=\int_{t_2}^{t_1}dt\left[p\frac{dq}{dt}-\tfrac{1}{2}(p^2+\omega_0^2q^2)+qF(t)|_+-|_-\right]$$
$$+\tfrac{1}{2}i\int_{t_2}^{t_1}dtdt'[(q(t)q(t'))_+A_{++}(t-t')$$
$$-2q_-(t)q_+(t')A_{-+}(t-t')+(q(t)q(t'))_-A_{--}(t-t')],$$

and the implied equations of motion, presented as second-order differential equations after eliminating

$$p=dq/dt,$$

are

$$\left(\frac{d^2}{dt^2}+\omega_0^2\right)q_+(t)$$
$$-i\int_{t_2}^{t_1}dt'[A_{++}(t-t')q_+(t')-A_{+-}(t-t')q_-(t')]=F_+(t)$$

and

$$\left(\frac{d^2}{dt'}+\omega_0^2\right)q_-(t)$$
$$+i\int_{t_2}^{t_1}dt[A_{--}(t-t')q_-(t')-A_{-+}(t-t')q_+(t')]=F_-(t).$$

It will be seen that the adjoint operation is equivalent to the interchange of the $\pm$labels.

We define

$$-iA_r(t-t')=\langle[Q(t),Q(t')]\rangle_s\eta_+(t-t')$$
$$=A_{++}-A_{+-}=A_{-+}-A_{--}$$

and

$$-iA_a(t-t')=-\langle[Q(t),Q(t')]\rangle_s\eta_-(t-t')$$
$$=A_{++}-A_{-+}=A_{+-}-A_{--},$$

together with

$$a(t-t')=\langle\{Q(t),Q(t')\}\rangle_s$$
$$=A_{+-}+A_{-+},$$

which enables us to present the integro-differential equations as

$$\left(\frac{d^2}{dt^2}+\omega_0^2\right)(q_--q_+)(t)-\int_{t_2}^{t_1}dt'A_a(t-t')(q_--q_+)(t')$$
$$=(F_--F_+)(t)$$

and

$$\left(\frac{d^2}{dt^2}+\omega_0^2\right)(q_++q_-)(t)-\int_{t_2}^{t_1}dt'A_r(t-t')(q_++q_-)(t')$$
$$+i\int_{t_2}^{t_1}dt'a(t-t')(q_--q_+)(t')=(F_++F_-)(t).$$

The accompanying boundary conditions are

$$(q_--q_+)(t_1)=0,\quad (d/dt)(q_--q_+)(t_1)=0$$

and

$$q_-(t_2)=q_+(t_2)\cosh\beta_0\omega_0+\frac{i}{\omega_0}\frac{d}{dt}q_+(t_2)\sinh\beta_0\omega_0$$
$$\frac{d}{dt}q_-(t_2)=-i\omega_0q_+(t_2)\sinh\beta_0\omega_0+\frac{d}{dt}q_+(t_2)\cosh\beta_0\omega_0,$$

or, more conveniently expressed,

$$(q_++q_-)(t_2)=\frac{i}{\omega_0}\coth(\tfrac{1}{2}\beta_0\omega_0)\frac{d}{dt}(q_--q_+)(t_2)$$
$$\frac{d}{dt}(q_++q_-)(t_2)=-i\omega_0\coth(\tfrac{1}{2}\beta_0\omega_0)(q_--q_+)(t_2),$$

which replace the non-Hermitian relations

$$y_-(t_2)=e^{\beta_0\omega_0}y_+(t_2),\quad y_-^\dagger(t_2)=e^{-\beta_0\omega_0}y_+^\dagger(t_2).$$

Note that it is the intrinsic oscillator frequency ω_0 that appears here since the initial condition refers to a thermal mixture of unperturbed oscillator states.

The required solution of the equation for q_--q_+ can be written as

$$(q_--q_+)(t)=\int_{-\infty}^{\infty}dt'G_a(t-t')(F_--F_+)(t'),$$

where $G_a(t-t')$ is the real Green's function defined by

$$\left(\frac{d^2}{dt^2}+\omega_0^2\right)G_a(t-t')-\int_{-\infty}^{\infty}d\tau A_a(t-\tau)G_a(\tau-t')=\delta(t-t')$$

and

$$G_a(t-t')=0,\quad t>t'.$$

Implicit is the time t_1 as one beyond which F_--F_+ equals zero. The initial conditions for the second equation, which this solution supplies, are

$$(q_++q_-)(t_2)=\frac{i}{\omega_0}\coth(\tfrac{1}{2}\beta_0\omega_0)\int_{t_2}^{\infty}dt'\frac{\partial}{\partial t_2}$$
$$\times G_a(t_2-t')(F_--F_+)(t')$$

and

$$\frac{d}{dt}(q_++q_-)(t_2)$$
$$=-i\omega_0\coth(\tfrac{1}{2}\beta_0\omega_0)\int_{t_2}^{\infty}dt'G_a(t_2-t')(F_--F_+)(t').$$

The Green's function that is appropriate for the equation obeyed by q_++q_- is defined by

$$\left(\frac{d^2}{dt^2}+\omega_0^2\right)G_r(t-t')-\int_{-\infty}^{\infty}d\tau A_r(t-\tau)G_r(\tau-t')=\delta(t-t'),$$
$$G_r(t-t')=0,\quad t<t',$$

and the two real functions are related by

$$G_a(t-t')=G_r(t'-t).$$

The desired solution of the second differential equation is

$$(q_+ + q_-)(t)=\int_{t_2}^{\infty} dt' G_r(t-t')(F_+ + F_-)(t') - i\int_{t_2}^{\infty} dt' w(t-t_2, t'-t_2)(F_- - F_+)(t'),$$

where

$$w(t-t_2, t'-t_2) = \int_{t_2}^{\infty} d\tau d\tau' G_r(t-\tau)a(\tau-\tau')G_a(\tau'-t') + \frac{1}{\omega_0}\coth(\tfrac{1}{2}\beta_0\omega_0)\left[\frac{\partial}{\partial t_2}G_r(t-t_2)\frac{\partial}{\partial t_2}G_a(t_2-t') + \omega_0^2 G_r(t-t_2)G_a(t_2-t')\right]$$

is a real symmetrical function of its two arguments.

The differential description of the transformation function that these solutions imply is indicated by

$$\delta_{F\pm}\langle t_2|t_2\rangle^{F\pm} = i\left\langle t_2\left|\int dt(\delta F_+ q_+ - \delta F_- q_-)\right|t_2\right\rangle = -\tfrac{1}{2}i\left\langle t_2\left|\int dt[\delta(F_- - F_+)(q_+ + q_-) + \delta(F_+ + F_-)(q_- - q_+)]\right|t_2\right\rangle,$$

and the result of integration is

$$\langle t_2|t_2\rangle_{\vartheta_0\vartheta}{}^{F\pm} = \exp\left\{-\tfrac{1}{2}i\int dtdt'(F_- - F_+)(t)G_r(t-t')(F_+ + F_-)(t') - \tfrac{1}{4}\int dtdt'(F_- - F_+)(t)w(t-t_2, t'-t_2)(F_- - F_+)(t')\right\}.$$

This can also be displayed in the matrix form

$$\langle t_2|t_2\rangle_{\vartheta_0\vartheta}{}^{F\pm} = \exp\left\{\tfrac{1}{2}i\int dtdt' F(t)G_{\vartheta_0\vartheta}(t-t_2, t'-t_2)F(t')\right\},$$

with

$$G_{\vartheta_0\vartheta}(t-t_2, t'-t_2) = \tfrac{1}{2}G_r(t-t')\begin{pmatrix}1 & 1\\ -1 & -1\end{pmatrix} + \tfrac{1}{2}G_a(t-t')\begin{pmatrix}1 & -1\\ 1 & -1\end{pmatrix} + \tfrac{1}{2}iw(t-t_2, t'-t_2)\begin{pmatrix}1 & -1\\ -1 & 1\end{pmatrix}.$$

The latter obeys

$$G(t',t)^T = G(t,t')$$

$$-\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}G(t,t')^*\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix} = G(t,t'),$$

and its elements are given by

$$G = i\begin{pmatrix}\langle(q(t)q(t'))_+\rangle_{\vartheta_0\vartheta}, & -\langle q(t')q(t)\rangle_{\vartheta_0\vartheta}\\ -\langle q(t)q(t')\rangle_{\vartheta_0\vartheta}, & \langle(q(t)q(t'))_-\rangle_{\vartheta_0\vartheta}\end{pmatrix}.$$

We note the identifications

$$G_r(t-t') = i\langle[q(t),q(t')]\rangle\eta_+(t-t')$$
$$G_a(t-t') = -i\langle[q(t),q(t')]\rangle\eta_-(t-t')$$
$$w(t-t_2, t'-t_2) = \langle\{q(t),q(t')\}\rangle.$$

It is also seen that the sum of the columns of G is proportional to $G_r(t-t')$, while the sum of the rows contains only $G_a(t-t')$.

We shall suppose that $G_r(t-t')$ can have no more than exponential growth, $\sim e^{\alpha(t-t')}$, as $t-t'\to\infty$. Then the complex Fourier transform

$$G(\zeta)=\int_{-\infty}^{\infty} d(t-t')e^{i\zeta(t-t')}G_r(t-t')$$

exists in the upper half-plane

$$\mathrm{Im}\zeta > \alpha$$

and is given explicitly by

$$G(\zeta)=[\omega_0^2-\zeta^2-A(\zeta)]^{-1}.$$

Here

$$A(\zeta)=\int_{-\infty}^{\infty} d(t-t')e^{i\zeta(t-t')}A_r(t-t') = i\int_0^{\infty} d\tau e^{i\zeta\tau}\int_{-\infty}^{\infty}\frac{d\omega}{2\pi}e^{-i\omega\tau}(A_{-+}-A_{+-})(\omega) = \int_{-\infty}^{\infty}\frac{d\omega}{2\pi}\frac{(A_{-+}-A_{+-})(\omega)}{\omega-\zeta}$$

or, since $(A_{-+}-A_{+-})(\omega)$ is an odd function of ω,

$$A(\zeta)=\int_0^{\infty}\frac{d\omega}{\pi}\frac{\omega(A_{-+}-A_{+-})(\omega)}{\omega^2-\zeta^2}.$$

We have already remarked on the generality of the representation

$$A_{-+}(\omega)/A_{+-}(\omega)=e^{\omega\beta(\omega)},\quad \beta(-\omega)=\beta(\omega),$$

and thus we shall write

$$(A_{-+}-A_{+-})(\omega)=a(\omega)\tanh[\tfrac{1}{2}\omega\beta(\omega)]$$
$$(A_{-+}+A_{+-})(\omega)=a(\omega)=a(-\omega)\geq 0,$$

which gives

$$G(\zeta)^{-1}=\omega_0^2-\zeta^2-\int_0^\infty \frac{d\omega}{\pi}\frac{\omega a(\omega)\tanh[\frac{1}{2}\omega\beta(\omega)]}{\omega^2-\zeta^2}.$$

Since this is an even function of ζ, it also represents the Fourier transform of G_a in the lower half-plane $\mathrm{Im}\zeta<-\alpha$.

If the effective temperature is positive and finite at all frequencies, $\beta(\omega)>0$, $G(\zeta)$ can have no complex poles as a function of the variable ζ^2. A complex pole at $\zeta^2=x+iy$, $y\neq 0$, is a zero of $G(\zeta)^{-1}$ and requires that

$$y\left[1+\int_0^\infty \frac{d\omega}{\pi}\frac{\omega a(\omega)\tanh[\frac{1}{2}\omega\beta(\omega)]}{(\omega^2-x)^2+y^2}\right]=0,$$

which is impossible since the quantity in brackets exceeds unity. On letting y approach zero, we see that a pole of $G(\zeta)$ can occur at a point $x=\omega'^2>0$ only if $a(\omega')=0$. If the external system responds through the oscillator coupling to any impressed frequency, $a(\omega)>0$ for all ω and no pole can appear on the positive real axis of ζ^2. As to the negative real axis, $G(\zeta)^{-1}$ is a monotonically decreasing function of $\zeta^2=x$ that begins at $+\infty$ for $x=-\infty$ and will therefore have no zero on the negative real axis if it is still positive at $x=0$. The corresponding condition is

$$\omega_0^2>\int_0^\infty \frac{d\omega}{\pi}a(\omega)\frac{\tanh[\frac{1}{2}\omega\beta(\omega)]}{\omega}.$$

Under these circumstances $\alpha=0$, for $G(\zeta)$, *qua* function of ζ^2, has no singularity other than the branch line on the positive real axis, and the ζ singularities are therefore confined entirely to the real axis. This is indicated by

$$G(\zeta)=\int_0^\infty d\omega^2\frac{B(\omega^2)}{\omega^2-\zeta^2}=\int_{-\infty}^\infty d\omega\epsilon(\omega)\frac{B(\omega^2)}{\omega-\zeta},$$

and $B(\omega^2)$ is the positive quantity

$$B(\omega^2)=\frac{(2\pi)^{-1}a(\omega)\tanh[\frac{1}{2}|\omega|\beta(\omega)]}{\left[\omega_0^2-\omega^2-P\int_0^\infty \frac{d\omega'^2}{2\pi}\frac{\tanh(\frac{1}{2}\omega'\beta(\omega'))}{\omega'^2-\omega^2}a(\omega')\right]^2+[\frac{1}{2}a(\omega)\tanh\frac{1}{2}\omega\beta(\omega)]^2}.$$

Some integral relations are easily obtained by comparison of asymptotic forms. Thus

$$\int_0^\infty d\omega^2 B(\omega^2)=1,$$

$$\int_0^\infty d\omega^2\omega^2 B(\omega^2)=\omega_0^2,$$

and

$$\int_0^\infty d\omega^2\omega^4 B(\omega^2)=\omega_0^4+\int_0^\infty \frac{d\omega^2}{2\pi}a(\omega)\tanh[\tfrac{1}{2}\omega\beta(\omega)]$$
$$=\omega_0^4+\langle[i\dot{Q},Q]\rangle_\beta,$$

while setting $\zeta=0$ yields

$$\int_0^\infty d\omega^2\frac{B(\omega^2)}{\omega^2}=\left[\omega_0^2-\int_0^\infty \frac{d\omega}{\pi}a(\omega)\frac{\tanh\frac{1}{2}\omega\beta(\omega)}{\omega}\right]^{-1}$$

The Green's functions are recovered on using the inverse Fourier transformation

$$G(t-t')=\int_{-\infty}^\infty \frac{d\zeta}{2\pi}e^{-i\zeta(t-t')}G(\zeta),$$

where the path of integration is drawn in the half-plane of regularity. Accordingly,

$$G_r(t-t')=\int_0^\infty d\omega^2 B(\omega^2)\frac{\sin\omega(t-t')}{\omega}\eta_+(t-t')$$

and

$$G_a(t-t')=-\int_0^\infty d\omega^2 B(\omega^2)\frac{\sin\omega(t-t')}{\omega}\eta_-(t-t').$$

The integral relations mentioned previously can be expressed in terms of these Green's functions. Thus,

$$\int_0^\infty d\tau G_r(\tau)=\left[\omega_0^2-\int_0^\infty \frac{d\omega}{\pi}a\frac{\tanh(\frac{1}{2}\omega\beta)}{\omega}\right]^{-1},$$

while, in the limit of small positive τ,

$$G_r(\tau)-(1/\omega_0)\sin\omega_0\tau\sim(\tau^5/5!)\langle[i\dot{Q},Q]\rangle_\beta,$$

which indicates the initial effect of the coupling to the external system.

The function $B(\omega^2)$ is bounded, and the Green's functions must therefore approach zero as $|t-t'|\to\infty$. Accordingly, all reference to the initial oscillator condition and to the time t_2 must eventually disappear. For sufficiently large $t-t_2$, $t'-t_2$, the function $w(t-t_2,$

$t'-t_2$) reduces to

$$w(t-t')=\int_{-\infty}^{\infty}d\tau d\tau' G_r(t-\tau)a(\tau-\tau')G_a(\tau'-t')$$

$$=\int_{-\infty}^{\infty}\frac{d\omega}{2\pi}e^{-i\omega(t-t')}G(\omega+i\epsilon)a(\omega)G(\omega-i\epsilon)\big|_{\epsilon\to 0}.$$

But

$$(1/2\pi)a(\omega)|G(\omega+i\epsilon)|^2=B(\omega^2)\coth[\tfrac{1}{2}|\omega|\beta(\omega)],$$

and, therefore,

$$w(t-t')=\int_0^{\infty}d\omega^2 B(\omega^2)\coth[\tfrac{1}{2}\omega\beta(\omega)]\frac{1}{\omega}\cos\omega(t-t').$$

The corresponding asymptotic form of the matrix $G(t-t_2, t'-t_2)$ is given by

$$G(t-t')=i\int_{-\infty}^{\infty}d\omega B(\omega^2)e^{-i\omega(t-t')}\begin{pmatrix}\eta_+(\omega)\eta_+(t-t')+\eta_-(\omega)\eta_-(t-t')+n, & -\eta_-(\omega)-n\\ -\eta_+(\omega)-n, & \eta_+(\omega)\eta_-(t-t')+\eta_-(\omega)\eta_+(t-t')+n\end{pmatrix},$$

with

$$n(\omega)=(e^{|\omega|\beta(\omega)}-1)^{-1},$$

which describes the oscillator in equilibrium at each frequency with the external system. When the temperature is frequency independent, this is thermal equilibrium. Note also that at zero temperature $n(\omega)=0$, and $G(t-t')_{++}$ is characterized by the temporal outgoing wave boundary condition—positive (negative) frequencies for positive (negative) time difference. The situation is similar for $G(t-t')_{--}$ as a function of $t'-t$.

It can no longer be maintained that placing $\beta_0=\beta$ removes all reference to the initial time. An interval must elapse before thermal equilibrium is established at the common temperature. This can be seen by evaluating the t_2 derivative of $w(t-t_2, t'-t_2)$:

$$\frac{\partial}{\partial t_2}w=-G_r(t-t_2)\int_{t_2}^{\infty}d\tau' a(t_2-\tau')G_a(\tau'-t')$$

$$-\int_{t_2}^{\infty}d\tau G_r(t-\tau)a(\tau-t_2)G_a(t_2-t')$$

$$+\frac{1}{\omega_0}\coth(\tfrac{1}{2}\omega_0\beta_0)$$

$$\times\left\{\frac{\partial}{\partial t_2}G_r(t-t_2)\int_{-\infty}^{\infty}d\tau' A_a(t_2-\tau')G_a(\tau'-t')\right.$$

$$\left.+\int_{-\infty}^{\infty}d\tau A_r(t-\tau)G_r(\tau-t_2)\frac{\partial}{\partial t_2}G_a(t_2-t')\right\},$$

for if this is to vanish, the integrals involving G_r, say, must be expressible as linear combinations of $G_r(t-t_2)$ and its time derivative, which returns us to the approximate treatment of the preceding section, including the approximate identification of ω_0 with the effective oscillator frequency. Hence $\vartheta_0=\vartheta$ does not represent the initial condition of thermal equilibrium between oscillator and external system. While it is perfectly clear that the latter situation is described by the matrix $G_3(t-t')$, a derivation that employs thermal equilibrium as an initial condition would be desirable.

The required derivation is produced by the device of computing the trace of the transformation function $\langle t_2'|t_2\rangle^{F\pm}$, in which the return path terminates at the different time $t_2'=t_2-T$, and the external force $F_-(t)$ is zero in the interval between t_2 and t_2'. The particular significance of the trace appears on varying the parameter λ that measures the coupling between oscillator and external system:

$$\frac{\partial}{\partial\lambda}\langle t_2'|t_2\rangle^{F\pm}=i\left\langle t_2'\left|\left[\int_{t_2}^{t_1}dtq_+Q_+(t)-\int_{t_2'}^{t_1}dtq_-Q_-(t)\right.\right.\right.$$

$$\left.\left.\left.+G_\lambda(t_2')-G_\lambda(t_2)\right]\right|t_2\right\rangle^{F\pm}.$$

The operators G_λ are needed to generate infinitesimal transformations of the individual states at the corresponding times, if these states are defined by physical quantities that depend upon λ, such as the total energy. There is no analogous contribution to the trace, however, for the trace is independent of the representation, which is understood to be defined similarly at t_2 and t_2', and one could use a complete set that does not refer to λ. More generally, we observe that $G_\lambda(t_2')$ bears the same relation to the $\langle t_2'|$ states as does $G_\lambda(t_2)$ to the states at time t_2, and therefore

$$\mathrm{tr}\langle t_2'|G_\lambda(t_2')|t_2\rangle-\mathrm{tr}\langle t_2'|G_\lambda(t_2)|t_2\rangle=0.$$

Accordingly, the construction of an effective action operator can proceed as before, with appropriately modified ranges of time integration, and, for the external system, with

$$\langle Q(t)Q(t')\rangle=\frac{\mathrm{tr}\langle t_2'|Q(t)Q(t')|t_2\rangle}{\mathrm{tr}\langle t_2'|t_2\rangle}.$$

This trace structure implies that

$$\langle Q(t)Q(t_2)\rangle=\langle Q(t_2')Q(t)\rangle$$

or, since these correlation functions depend only on

time differences, that

$$A_{-+}(t-t_2)=A_{+-}(t-t_2'),$$

which is also expressed by

$$A_{-+}(\omega)=e^{-i\omega T}A_{+-}(\omega).$$

The equations of motion for $t>t_2$ are given by

$$\left(\frac{d^2}{dt^2}+\omega_0^2\right)(q_--q_+)(t)-\int_{-\infty}^{\infty}dt'A_a(t-t')(q_--q_+)(t')=(F_--F_+)(t)$$

and

$$\left(\frac{d^2}{dt^2}+\omega_0^2\right)(q_++q_-)(t)-\int_{t_2}^{\infty}dt'A_r(t-t')(q_++q_-)(t')+i\int_{t_2}^{\infty}dt'a(t-t')(q_--q_+)(t')=(F_++F_-)(t)-2i\int_{t_2'}^{t_2}dt'A_{+-}(t-t')q_-(t').$$

These are supplemented by the equation for $q_-(t)$ in the interval from t_2' to t_2:

$$\left(\frac{d^2}{dt^2}+\omega_0^2\right)q_-(t)+i\int_{t_2'}^{t_2}dt'A_{--}(t-t')q_-(t')=-i\int_{t_2}^{\infty}dt'A_{-+}(t-t')(q_--q_+)(t'),$$

and the effective boundary condition

$$q_-(t_2')=q_+(t_2).$$

The equation for q_--q_+ is solved as before,

$$(q_--q_+)(t)=\int_{-\infty}^{\infty}dt'G_a(t-t')(F_--F_+)(t'),$$

whereas

$$\begin{aligned}(q_++q_-)(t)&=\int_{-\infty}^{\infty}dt'G_r(t-t')(F_++F_-)(t')\\&\quad-i\int_{t_2}^{\infty}d\tau G_r(t-\tau)\int_{t_2}^{\infty}dt'a(\tau-t')(q_--q_+)(t')\\&\quad-2i\int_{t_2}^{\infty}d\tau G_r(t-\tau)\int_{t_2-T}^{t_2}dt'A_{+-}(\tau-t')q_-(t')\\&\quad+G_r(t-t_2)\frac{\partial}{\partial t_2}(q_++q_-)(t_2)\\&\quad-\frac{\partial}{\partial t_2}G_r(t-t_2)(q_++q_-)(t_2),\end{aligned}$$

which has been written for external forces that are zero until the moment t_2 has passed.

Perhaps the simplest procedure at this point is to ask for the dependence of the latter solution upon t_2, for fixed T. We find that

$$\begin{aligned}\frac{\partial}{\partial t_2}(q_++q_-)(t)&=-\int_{t_2}^{\infty}dt'G_r(t-t')A_r(t'-t_2)(q_++q_-)(t_2)\\&\quad+i\int_{t_2}^{\infty}dt'G_r(t-t')a(t'-t_2)(q_--q_+)(t_2)\\&\quad-2i\int_{t_2}^{\infty}dt'G_r(t-t')[A_{+-}(t'-t_2)q_-(t_2)\\&\quad-A_{-+}(t'-t_2)q_+(t_2)],\end{aligned}$$

on using the relations

$$\int_{-\infty}^{\infty}d\tau A_r(t-\tau)G_r(\tau-t')=\int_{-\infty}^{\infty}d\tau G_r(t-\tau)A_r(\tau-t'),$$

$$A_{+-}(t-t_2')q_-(t_2')=A_{-+}(t-t_2)q_+(t_2).$$

Therefore,

$$(\partial/\partial t_2)(q_++q_-)(t)=0,$$

since, with positive time argument,

$$a-iA_r=2A_{-+}$$
$$a+iA_r=2A_{+-}.$$

The utility of this result depends upon the approach of the Green's functions to zero with increasing magnitude of the time argument, which is assured, after making the substitution $T\rightarrow i\beta$, under the circumstances we have indicated. Then we can let $t_2\rightarrow-\infty$ and obtain

$$(q_++q_-)(t)=\int_{-\infty}^{\infty}dt'G_r(t-t')(F_++F_-)(t')-i\int_{-\infty}^{\infty}dt'w(t-t')(F_--F_+)(t')$$

with

$$w(t-t')=\int_{-\infty}^{\infty}d\tau d\tau'G_r(t-\tau)a(\tau-\tau')G_a(\tau'-t'),$$

as anticipated.

Our results determine the trace ratio

$$\frac{\mathrm{tr}\langle t_2'|t_2\rangle^{F\pm}}{\mathrm{tr}\langle t_2'|t_2\rangle}=\frac{\mathrm{tr}\langle t_2|e^{iTH}|t_2\rangle^{F\pm}}{\mathrm{tr}e^{iTH}},$$

where H is the Hamiltonian operator of the complete system, and the substitution $T\rightarrow i\beta$ yields the transformation function

$$\langle t_2|t_2\rangle_\beta^{F\pm}=\exp\left[\tfrac{1}{2}i\int dtdt'F(t)G_\beta(t-t')F(t')\right]$$

with

$$G_\vartheta(t-t')=\tfrac{1}{2}G_r(t-t')\begin{pmatrix}1 & 1\\ -1 & -1\end{pmatrix}+\tfrac{1}{2}G_a(t-t')\begin{pmatrix}1 & -1\\ 1 & -1\end{pmatrix}$$
$$+\tfrac{1}{2}iw(t-t')\begin{pmatrix}1 & -1\\ -1 & 1\end{pmatrix}$$

and

$$w(t-t')=\int_{-\infty}^{\infty}d\omega B(\omega^2)\coth(\tfrac{1}{2}|\omega|\beta)e^{-i\omega(t-t')}.$$

We can also write

$$w(t-t')=\int_{-\infty}^{\infty}d\tau C(t-\tau)(G_a-G_r)(\tau-t'),$$

where

$$C(t-t')=\frac{i}{2\pi}P\int_{-\infty}^{\infty}d\omega\coth(\tfrac{1}{2}\omega\beta)e^{-i\omega(t-t')}$$
$$=\frac{1}{\beta}\coth\left[\frac{\pi}{\beta}(t-t')\right].$$

What is asserted here about expectation values in the presence of an external field $F(t)$ becomes explicit on writing

$$F_\pm(t)=f_\pm(t)+F(t)$$

and indicating the effect of $f_\pm(t)$ by equivalent time-ordered operators,

$$\left\langle\left(\exp\left[-i\int dtf_-(t)q(t)\right]\right)_-\right.$$
$$\left.\times\left(\exp\left[i\int dtf_+(t)q(t)\right]\right)_+\right\rangle_\vartheta^F$$
$$=\exp\left\{\tfrac{1}{2}i\int dtdt'f(t)G_\vartheta(t-t')f(t')\right.$$
$$\left.+i\int dtdt'(f_+-f_-)(t)G_r(t-t')F(t')\right\}.$$

Thus

$$\langle q(t)\rangle_\vartheta^F=\int_{-\infty}^{\infty}dt'G_r(t-t')F(t')$$

and the properties of $q-\langle q\rangle_\vartheta^F$, which are independent of F, are given by setting $F=0$ in the general result. In particular, we recover the matrix identity

$$G_\vartheta(t-t')=i\begin{pmatrix}\langle(q(t)q(t'))_+\rangle_\vartheta, & -\langle q(t')q(t)\rangle_\vartheta\\ \langle q(t)q(t')\rangle_\vartheta, & \langle(q(t)q(t'))_-\rangle_\vartheta\end{pmatrix}.$$

The relation between w and G_a-G_r can then be displayed as a connection between symmetrical product and commutator expectation values

$$\langle\{q(t),q(t')\}\rangle_\vartheta=\int_{-\infty}^{\infty}d\tau C(t-\tau)\left\langle\frac{1}{i}[q(\tau),q(t')]\right\rangle_\vartheta$$

In addition to the trace ratio, which determines the thermal average transformation function $\langle t_2|t_2\rangle_\vartheta^{F\pm}$ with its attendant physical information, it is possible to compute the trace

$$\mathrm{tr}\langle t_2'|t_2\rangle=\mathrm{tr}e^{iTH}\longrightarrow\mathrm{tr}e^{-\beta H}$$

which describes the complete energy spectrum and thereby the thermostatic properties of the oscillator in equilibrium with the external system. For this purpose we set $F_\pm=0$ for $t>t_2$ and apply an arbitrary external force $F_-(t)$ in the interval from t_2' to t_2. Moreover, the coupling term between oscillator and external system in the effective action operator is supplied with the variable factor λ (formerly λ^2). Then we have

$$\frac{\partial}{\partial\lambda}\mathrm{tr}\langle t_2'|t_2\rangle^{F-}$$
$$=-\tfrac{1}{2}\mathrm{tr}\left\langle t_2'\left|\int_{t_2'}^{t_2}dtdt'A_{--}(t-t')(q(t)q(t'))_-\right|t_2\right\rangle^{F-}$$
$$=-\tfrac{1}{2}i\int_{t_2'}^{t_2}dtdt'A_{--}(t-t')\frac{\delta}{\delta F_-(t')}\mathrm{tr}\langle t_2'|q_-(t)|t_2\rangle^{F-},$$

where $q_-(t)$ obeys the equation of motion

$$\left(\frac{d^2}{dt^2}+\omega_0^2\right)q_-(t)+i\lambda\int_{t_2'}^{t_2}dt'A_{--}(t-t')q_-(t')=F_-(t)$$

with the accompanying boundary condition

$$q_-(t_2')=q_+(t_2)=q_-(t_2),$$

which is a statement of periodicity for the interval $T=t_2'-t_2$. The solution of this equation is

$$q_-(t)=\int_{t_2'}^{t_2}dt'G(t-t')F_-(t'),$$

where the Green's function obeys

$$\left(\frac{d^2}{dt^2}+\omega_0^2\right)G(t-t')+i\lambda\int_{t_2'}^{t_2}d\tau A_{--}(t-\tau)G(\tau-t')=\delta(t-t')$$

and the requirement of periodicity. We can now place $F_-=0$ in the differential equation for the trace, and obtain

$$\frac{\partial}{\partial\lambda}\log\mathrm{tr}\langle t_2'|t_2\rangle=-\tfrac{1}{2}i\int_{t_2'}^{t_2}dtdt'A_{--}(t-t')G(t-t').$$

The periodic Green's function is given by the Fourier series

$$G(t-t')=\frac{1}{T}\sum_{n=-\infty}^{\infty}\exp\left[-\frac{2\pi in}{T}(t-t')\right]G(n)$$

with

$$G(n)=\left[\omega_0^2-\left(\frac{2\pi n}{T}\right)^2-\lambda A(n)\right]^{-1}=G(-n)$$

and

$$A(n)=-\int_{t_2'}^{t_2}dt\,\exp\left[\frac{2\pi in}{T}(t-t')\right]iA_{--}(t-t')$$

$$=\int_0^\infty\frac{\omega d\omega}{\pi}\frac{(A_{-+}-A_{+-})(\omega)}{\omega^2-(2\pi n/T)^2},$$

where, it is to be recalled,

$$A_{-+}(\omega)=e^{-i\omega T}A_{+-}(\omega),$$

so that the integrand has no singularities at $\omega T=2\pi|n|$. Now we have

$$\frac{\partial}{\partial\lambda}\log\mathrm{tr}=\tfrac{1}{2}\sum_{-\infty}^{\infty}A(n)G(n)$$

$$=-\frac{1}{2}\frac{\partial}{\partial\lambda}\sum_{-\infty}^{\infty}\log\left[\omega_0^2-\left(\frac{2\pi n}{T}\right)^2-\lambda A(n)\right]$$

which, together with the initial condition

$$\lambda=0:\quad \mathrm{tr}e^{iTH}=(\mathrm{tr}_e e^{iTH_{\mathrm{ext}}})\sum_{n=0}^{\infty}e^{i(n+\frac{1}{2})\omega_0 T}$$

$$=(\mathrm{tr}_e)(i/2\sin\tfrac{1}{2}\omega_0 T),$$

yields

$$\mathrm{tr}e^{iTH}=(\mathrm{tr}_e)(i/2\sin\tfrac{1}{2}\omega_0 T)$$

$$\times\exp\left\{-\tfrac{1}{2}\sum_{-\infty}^{\infty}\log\left[\frac{\omega_0^2-(2\pi n/T)^2-A(n)}{\omega_0^2-(2\pi n/T)^2}\right]\right\}.$$

We have already introduced the function

$$G^{-1}(\zeta)=\omega_0^2-\zeta^2-\int_0^\infty\frac{\omega d\omega}{\pi}\frac{(A_{-+}-A_{+-})(\omega)}{\omega^2-\zeta^2}$$

and examined some of its properties for real and positive $A_{-+}(\omega)$, $A_{+-}(\omega)$. This situation is recovered on making the substitution $T\rightarrow i\beta$, and thus

$$Z=\mathrm{tr}e^{-\beta H}=Z_e(1/2\sinh\tfrac{1}{2}\beta\omega_0)$$

$$\times\exp\left\{-\tfrac{1}{2}\sum_{-\infty}^{\infty}\log\left[\frac{G^{-1}(i2\pi n/\beta)}{\omega_0^2+(2\pi n/\beta)^2}\right]\right\},$$

the existence of which for all $\beta>0$ requires that $G^{-1}(\zeta)$ remain positive at every value comprised in $\zeta^2=-(2\pi n/\beta)^2$, which is to say the entire negative ζ^2 axis including the origin. The condition

$$G^{-1}(0)>0$$

is thereby identified as a stability criterion. To evaluate the summation over n most conveniently we shall give an alternative construction for the function $\log(G^{-1}(\zeta)/-\zeta^2)$, which, as a function of ζ^2, has all its singularities located on the branch line extending from 0 to ∞ and vanishes at infinity in this cut plane. Hence

$$\log(G^{-1}(\zeta)/-\zeta^2)=\frac{1}{\pi}\int_0^\infty d\omega^2\frac{\varphi(\omega)}{\omega^2-\zeta^2},$$

where the value

$$\varphi(0)=\pi$$

reproduces the pole of $G^{-1}(\zeta)/(-\zeta^2)$ at $\zeta^2=0$. We also recognize, on relating the two forms,

$$G^{-1}(\zeta)=(-\zeta^2)\exp\left[\frac{1}{\pi}\int_0^\infty d\omega^2\frac{\varphi(\omega)}{\omega^2-\zeta^2}\right]$$

$$=\omega_0^2-\zeta^2-\int_0^\infty\frac{d\omega^2}{2\pi}\frac{a(\omega)\tanh(\frac{1}{2}\omega\beta)}{\omega^2-\zeta^2},$$

that

$$-\tfrac{1}{2}a(\omega)\tanh(\tfrac{1}{2}\omega\beta)\cot\varphi(\omega)$$

$$=\omega_0^2-\omega^2-P\int_0^\infty\frac{d\omega'}{\pi}\frac{\omega' a(\omega')\tanh(\frac{1}{2}\omega'\beta)}{\omega'^2-\omega^2}.$$

The positive value of the right-hand side as $\omega\rightarrow 0$ shows that $\varphi(\omega)$ approaches the zero frequency limiting value of π from below, and the assumption that $a(\omega)>0$ for all ω implies

$$\pi\geq\varphi(\omega)>0,$$

where the lower limit is approached as $\omega\rightarrow\infty$.

A comparison of asymptotic forms for $G^{-1}(\zeta)$ shows that

$$\omega_0^2=\frac{1}{\pi}\int_0^\infty d\omega^2\varphi(\omega)=\int_0^\infty d\omega\left(-\frac{1}{\pi}\frac{d\varphi(\omega)}{d\omega}\right)\omega^2,$$

while

$$\int_0^\infty d\omega\left(-\frac{1}{\pi}\frac{d\varphi(\omega)}{d\omega}\right)\omega^4=\omega_0^4+2\langle[i\dot{Q},Q]\rangle_\beta.$$

The introduction of the phase derivative can also be performed directly in the structure of $G^{-1}(\zeta)$,

$$G^{-1}(\zeta)=\exp\left[\int_0^\infty d\omega\left(-\frac{1}{\pi}\frac{d\varphi(\omega)}{d\omega}\right)\log(\omega^2-\zeta^2)\right],$$

and equating the two values for $G^{-1}(0)$ gives

$$\int_0^\infty d\omega\left(-\frac{1}{\pi}\frac{d\varphi(\omega)}{d\omega}\right)\log\omega^2$$

$$=\log\left[\omega_0^2-\int_0^\infty\frac{d\omega}{\pi}a(\omega)\frac{\tanh\frac{1}{2}\omega\beta}{\omega}\right].$$

We now have the representation

$$\log\left[\frac{G^{-1}(i2\pi n/\beta)}{\omega_0^2+(2\pi n/\beta)^2}\right]$$

$$=\int_0^\infty d\omega\left(-\frac{1}{\pi}\frac{d\varphi(\omega)}{d\omega}\right)\log\frac{\omega^2+(2\pi n/\beta)^2}{\omega_0^2+(2\pi n/\beta)^2},$$

and the summation formula derived from the product form of the hyperbolic sine function,

$$\tfrac{1}{2}\sum_{-\infty}^{\infty}\log\frac{\omega^2+(2\pi n/\beta)^2}{\omega_0^2+(2\pi n/\beta)^2}=\log\left[\frac{\sinh\frac{1}{2}\omega\beta}{\sinh\frac{1}{2}\omega_0\beta}\right],$$

gives us the desired result

$$Z=Z_e\exp\left[-\int_0^\infty d\omega\left(-\frac{1}{\pi}\frac{d\varphi(\omega)}{d\omega}\right)\log 2\sinh\tfrac{1}{2}\omega\beta\right].$$

The second factor can be ascribed to the oscillator, with its properties modified by interaction with the external system. The average energy of the oscillator at temperature $\vartheta=\beta^{-1}$ is therefore given by

$$E=\frac{\partial}{\partial\beta}\int_0^\infty d\omega\left(-\frac{1}{\pi}\frac{d\varphi}{d\omega}\right)\log 2\sinh\tfrac{1}{2}\omega\beta$$

in which the temperature dependence of the phase $\varphi(\omega)$ is not to be overlooked. In an extreme high-temperature limit, such that $\omega\beta\ll 1$ for all significant frequencies, we have

$$E\simeq\frac{\partial}{\partial\beta}[\log\beta+\tfrac{1}{2}\log(\omega_0^2-\beta\langle Q^2\rangle_\vartheta)],$$

and the simple classical result $E=\vartheta$ appears when $\langle Q^2\rangle_\vartheta$ is proportional to ϑ. The oscillator energy at zero temperature is given by

$$E_0=\tfrac{1}{2}\int_0^\infty d\omega\left(-\frac{1}{\pi}\frac{d\varphi}{d\omega}\right)_{\vartheta=0}\omega,$$

and the oscillator contribution to the specific heat vanishes.

The following physical situation has consequences that resemble the simple model of the previous section. For values of $\omega\lesssim\omega_0$, $a(\omega)\tanh(\frac{1}{2}\omega\beta)\ll\omega_0^2$, and $a(\omega)$ differs significantly from zero until one attains frequencies that are large in comparison with ω_0. The magnitudes that $a(\omega)$ can assume at frequencies greater than ω_0 is limited only by the assumed absence of rapid variations and by the requirement of stability. The latter is generally assured if

$$\frac{1}{\pi}\int_{\sim\omega_0}^\infty\frac{d\omega}{\omega}a(\omega)<\omega_0^2.$$

We shall suppose that the stability requirement is comfortably satisfied, so that the right-hand side of the equation for $\cot\varphi(\omega)$ is an appreciable fraction of ω_0^2 at sufficiently low frequencies. Then $\tan\varphi$ is very small at such frequencies, or $\varphi(\omega)\sim\pi$, and this persists until we reach the immediate neighborhood of the frequency $\omega_1<\omega_0$ such that

$$\omega_0^2-\omega_1^2-P\int_0^\infty\frac{d\omega}{\pi}\frac{\omega a(\omega)\tanh(\frac{1}{2}\omega\beta)}{\omega^2-\omega_1^2}=0.$$

That the function in question, $\mathrm{Re}G^{-1}(\omega+i0)$, has a zero, follows from its positive value at $\omega=0$ and its asymptotic approach to $-\infty$ with indefinitely increasing frequency. Under the conditions we have described, with the major contribution to the integral coming from high frequencies, the zero point is given approximately as

$$\omega_1^2\simeq\omega_0^2-\int_0^\infty\frac{d\omega}{\pi}a(\omega)\frac{\tanh(\frac{1}{2}\omega\beta)}{\omega},$$

and somewhat more accurately by

$$\omega_1^2=B\left[\omega_0^2-\int_0^\infty\frac{d\omega}{\pi}a(\omega)\frac{\tanh(\frac{1}{2}\omega\beta)}{\omega}\right],$$

where

$$B^{-1}=1+P\int_0^\infty\frac{d\omega}{2\pi}\frac{1}{\omega^2-\omega_1^2}\frac{d}{d\omega}[a(\omega)\tanh(\tfrac{1}{2}\omega\beta)].$$

As we shall see, B is less than unity, but only slightly so under the circumstances assumed.

In the neighborhood of the frequency ω_1, the equation that determines $\varphi(\omega)$ can be approximated by

$$-\tfrac{1}{2}a(\omega_1)\tanh(\tfrac{1}{2}\omega_1\beta)\cot\varphi(\omega)=B^{-1}(\omega_1^2-\omega^2)$$

or

$$\cot\varphi(\omega)=(\omega^2-\omega_1^2)/\gamma\omega_1\simeq(\omega-\omega_1)/\tfrac{1}{2}\gamma,$$

with the definition

$$\gamma=\tfrac{1}{2}Ba(\omega_1)[\tanh(\tfrac{1}{2}\omega_1\beta)/\omega_1]\ll\omega_1.$$

Hence, as ω rises through the frequency ω_1, φ decreases abruptly from a value close to π to one near zero. The subsequent variations of the phase are comparatively gradual, and φ eventually approaches zero as $\omega\to\infty$. A simple evaluation of the average oscillator energy can be given when the frequency range $\omega>\omega_1$ over which $a(\omega)$ is appreciable in magnitude is such that $\beta\omega\gg 1$. There will be no significant temperature variation in the latter domain and in particular ω_1 should be essentially temperature independent. Then, since $-(1/\pi)(d\varphi/d\omega)$ in the neighborhood of ω_1 closely resembles $\delta(\omega-\omega_1)$, we have approximately

$$E=\frac{\partial}{\partial\beta}\left[\log(2\sinh\tfrac{1}{2}\omega_1\beta)+\beta\int_{>\omega_1}^\infty d\omega\left(-\frac{1}{\pi}\frac{d\varphi}{d\omega}\right)\tfrac{1}{2}\omega\right]$$

$$=\omega_1\left(\frac{1}{e^{\beta\omega_1}-1}+\tfrac{1}{2}\right)+\int_{>\omega_1}^\infty d\omega\left(-\frac{1}{\pi}\frac{d\varphi}{d\omega}\right)\tfrac{1}{2}\omega,$$

which describes a simple oscillator of frequency ω_1, with a displaced origin of energy.

Note that with $\varphi(\omega)$ very small at a frequency slightly greater than ω_1 and zero at infinite frequency, we have

$$\int_{>\omega_1}^\infty d\omega\left(-\frac{1}{\pi}\frac{d\varphi}{d\omega}\right)\tfrac{1}{2}\omega\simeq\frac{1}{2\pi}\int_{>\omega_1}^\infty d\omega\,\varphi(\omega)>0.$$

Related integrals are

$$\omega_0^2-\omega_1^2\simeq\frac{2}{\pi}\int_{>\omega_1}^{\infty}d\omega\omega\varphi(\omega)>0$$

and

$$\log B^{-1}\simeq\frac{2}{\pi}\int_{>\omega_1}^{\infty}\frac{d\omega}{\omega}\varphi(\omega)>0.$$

The latter result confirms that $B<1$. A somewhat more accurate formula for B is

$$B=\exp\left[-\int_{>\omega_1}^{\infty}d\omega\left(-\frac{1}{\pi}\frac{d\varphi(\omega)}{d\omega}\right)\log(\omega^2-\omega_1^2)\right].$$

If the major contributions to all these integrals come from the general vicinity of a frequency $\bar{\omega}\gg\omega_0$, we can make the crude estimates

$$\frac{1}{2\pi}\int_{>\omega_1}^{\infty}d\omega\varphi(\omega)\sim\frac{\omega_0^2}{\bar{\omega}}\ll\omega_1,\ \log B^{-1}\sim\left(\frac{\omega_0}{\bar{\omega}}\right)^2\ll 1.$$

Then neither the energy shift nor the deviation of the factor B from unity are particularly significant effects.

The approximation of $\mathrm{Re}G(\omega+i0)$ as $B^{-1}(\omega_1^2-\omega^2)$ evidently holds from zero frequency up to a frequency considerably in excess of ω_1. Throughout this frequency range we have

$$-\tfrac{1}{2}a(\omega)\tanh(\tfrac{1}{2}\omega\beta)\cot\varphi(\omega)=B^{-1}(\omega_1^2-\omega^2)$$

or

$$\cot\varphi(\omega)=(\omega_1^2-\omega^2)/\gamma\omega$$

with

$$\gamma(\omega)=\tfrac{1}{2}Ba(\omega)\tanh(\tfrac{1}{2}\omega\beta)/\omega.$$

If in particular $\beta\omega_1\ll 1$, the frequencies under consideration are in the classical domain and γ is the frequency independent constant

$$\gamma=\tfrac{1}{4}Ba(0)\beta.$$

To regard γ as constant for a quantum oscillator requires a suitable frequency restriction to the vicinity of ω_1. The function $B(\omega^2)$ can be computed from

$$B(\omega^2)=\frac{2}{\pi}\frac{\sin^2\varphi(\omega)}{a(\omega)\tanh(\frac{1}{2}\omega\beta)}$$

$$=\frac{1}{\pi}\frac{B}{\gamma\omega}\frac{1}{\cot^2\varphi+1},$$

and accordingly is given by

$$B(\omega^2)=B\frac{1}{\pi}\frac{\gamma\omega}{(\omega^2-\omega_1^2)^2+(\gamma\omega)^2}$$

$$=B\frac{\gamma\omega}{\pi}\left|\frac{1}{\omega^2+i\gamma\omega-\omega_1^2}\right|^2.$$

The further concentration on the immediate vicinity of ω_1, $|\omega-\omega_1|\sim\gamma$, gives

$$B(\omega^2)=\frac{B}{2\omega_1}\frac{1}{\pi}\frac{\frac{1}{2}\gamma}{(\omega-\omega_1)^2+(\frac{1}{2}\gamma)^2}$$

which clearly identifies $B<1$ with the contribution to the integral $\int d\omega^2 B(\omega^2)$ that comes from the vicinity of this resonance of width γ at frequency ω_1, although the same result is obtained without the last approximation. The remainder of the integral, $1-B$, arises from frequencies considerably higher than ω_1 according to our assumptions.

There is a similar decomposition of the expressions for the Green's functions. Thus, with $t>t'$,

$$G_r(t-t')\simeq B\int_{\infty}^{\infty}\frac{d\omega}{2\pi}e^{-i\omega(t-t')}\frac{1}{-\omega^2-i\gamma\omega+\omega_1^2}$$

$$+\int_{\gg\omega_1^2}^{\infty}d\omega^2B(\omega^2)\frac{\sin\omega(t-t')}{\omega}.$$

The second high-frequency term will decrease very quickly on the time scale set by $1/\omega_1$. Accordingly, in using this Green's function, say in the evaluation of

$$\langle q(t)\rangle_{\mathcal{F}}=\int_{-\infty}^{\infty}dt'G_r(t-t')F(t')$$

for an external force that does not vary rapidly in relation to ω_1, the contribution of the high-frequency term is essentially given by

$$F(t)\int_0^{\infty}d(t-t')\int_{\gg\omega_1^2}^{\infty}d\omega^2B(\omega^2)\frac{\sin\omega(t-t')}{\omega}$$

$$=F(t)\int_{\gg\omega_1^2}^{\infty}d\omega^2\frac{B(\omega^2)}{\omega^2}.$$

But

$$\int_{\gg\omega_1^2}^{\infty}d\omega^2\frac{B(\omega^2)}{\omega^2}\simeq\left[\omega_0^2-\int_0^{\infty}\frac{d\omega}{\pi}a(\omega)\frac{\tanh\frac{1}{2}\omega\beta}{\omega}\right]^{-1}-\frac{B}{\omega_1^2},$$

$$\simeq 0,$$

and the response to such an external force is adequately described by the low-frequency part of the Green's function. We can represent this situation by an equivalent differential equation

$$\left(\frac{d^2}{dt^2}+\gamma\frac{d}{dt}+\omega_1^2\right)\langle q(t)\rangle_{\mathcal{F}}=BF(t)$$

which needs no further qualification when the oscillations are classical but implies a restriction to a frequency interval within which γ is constant, for quantum oscillations. We note the reduction in the effectiveness of the external force by the factor B. Under the circumstances

outlined this effect is not important and we shall place B equal to unity.

One can make a general replacement of the Green's functions by their low-frequency parts:

$$G_r(t-t') \to e^{-\frac{1}{2}\gamma(t-t')}\frac{1}{\omega_1}\sin(\omega_1(t-t'))\eta_+(t-t')$$

$$G_a(t-t') \to -e^{-\frac{1}{2}\gamma|t-t'|}\frac{1}{\omega_1}\sin(\omega_1(t-t'))\eta_-(t-t'),$$

if one limits the time localizability of measurements so that only time averages of $q(t)$ are of physical interest. This is represented in the expectation value formula by considering only functions $f_\pm(t)$ that do not vary too quickly. The corresponding replacement for $w(t-t')$ is

$$w(t-t') \to \coth(\tfrac{1}{2}\omega_1\beta)e^{-\frac{1}{2}\gamma|t-t'|}\frac{1}{\omega_1}\cos\omega_1(t-t'),$$

and the entire matrix $G_\vartheta(t-t')$ obtained in this way obeys the differential equation

$$\left(\frac{d^2}{dt^2}+\gamma\frac{d}{dt}+\omega_1^2\right)\left(\frac{d^2}{dt'^2}+\gamma\frac{d}{dt'}+\omega_1^2\right)G_\vartheta(t-t')$$

$$=\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\left(\frac{d^2}{dt^2}+\omega_1^2\right)\delta(t-t')$$

$$+\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}\gamma\frac{d}{dt}\delta(t-t')+\begin{pmatrix}1 & -1\\ -1 & 1\end{pmatrix}\tfrac{1}{2}ia\delta(t-t'),$$

where $a=a(\omega_1)$.

The simplest presentation of results is again to be found in the Langevin viewpoint, which directs the emphasis from the coordinate operator $q(t)$ to the fluctuating force defined by

$$\left(\frac{d^2}{dt^2}+\gamma\frac{d}{dt}+\omega_1^2\right)q(t)=F(t)+F_f(t),$$

which is to say

$$F_f(t)=Q(t)+\gamma\frac{d}{dt}q(t)+(\omega_1^2-\omega_0^2)q(t).$$

This change is introduced by the substitution

$$f_\pm(t)=\left(\frac{d^2}{dt^2}-\gamma\frac{d}{dt}+\omega_1^2\right)k_\pm(t),$$

and the necessary partial integrations involve the previously established lemma on time-ordered operators, which here asserts that

$$\left(\exp\left[i\int dtfq\right]\right)_+=\left(\exp\left[i\int dtk(F+F_f)\right]\right)_+$$

$$\times\exp\left\{\tfrac{1}{2}i\int dt[(\omega_1^2-\omega_0^2)k^2+\omega_1^2k^2-(dk/dt)^2]\right\}.$$

We now find

$$\left\langle\left(\exp\left[-i\int dtk_-F_f\right]\right)_-\left(\exp\left[i\int dtk_+F_f\right]\right)_+\right\rangle_\vartheta$$

$$=\exp\left[-\tfrac{1}{2}\int dtdt'k(t)\zeta(t-t')k(t')\right]$$

with

$$\zeta(t-t')=\tfrac{1}{2}a\begin{pmatrix}1 & -1\\ -1 & 1\end{pmatrix}\delta(t-t')$$

$$-i(\omega_0^2-\omega_1^2)\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\delta(t-t')$$

$$+i\gamma\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}\frac{d}{dt}\delta(t-t').$$

The latter matrix can also be identified as

$$\zeta(t-t')=\begin{pmatrix}\langle(F_f(t)F_f(t'))_+\rangle_\vartheta, & -\langle F_f(t')F_f(t)\rangle_\vartheta\\ -\langle F_f(t)F_f(t')\rangle_\vartheta, & \langle(F_f(t)F_f(t'))_-\rangle_\vartheta\end{pmatrix}.$$

In the classical limit

$$\left\langle\exp\left[i\int dtkF_f\right]\right\rangle_\vartheta=\exp\left[-\tfrac{1}{4}a\int dtk^2\right]$$

and

$$\tfrac{1}{4}a=\gamma\vartheta.$$

If a comparison is made with the similar results of the previous section it can be appreciated that the frequency range has been extended and the restriction $\omega_1\simeq\omega_0$ removed.

We return from these extended considerations on thermal equilibrium and consider one extreme example of negative temperature for the external system. This is described by

$$a(\omega)=a\delta(\omega-\omega_1),\quad \omega>0$$

and

$$-\beta(\omega_1)=|\beta|>0.$$

With the definition

$$(1/\pi)\omega_1 a\tanh(\tfrac{1}{2}\omega_1|\beta|)=(\omega_1\mu)^2,$$

we have

$$G(\zeta)=\left[\omega_0^2-\zeta^2+\frac{(\omega_1\mu)^2}{\omega_1^2-\zeta^2}\right]^{-1}$$
$$=\frac{\omega_1^2-\zeta^2}{[\zeta^2-\frac{1}{2}(\omega_0^2+\omega_1^2)]^2+(\omega_1\mu)^2-[\frac{1}{2}(\omega_0^2-\omega_1^2)]^2}.$$

As a function of ζ^2, $G(\zeta)$ now has complex poles if

$$\tfrac{1}{2}|\omega_0^2-\omega_1^2|<\omega_1\mu.$$

We shall suppose, for simplicity, that $\omega_1=\omega_0$ and $\mu\ll\omega_0$. Then the poles of

$$G(\zeta)=\tfrac{1}{2}[(\omega_0^2+i\omega_0\mu-\zeta^2)^{-1}+(\omega_0^2-i\omega_0\mu-\zeta^2)^{-1}]$$

are located at $\zeta=\pm(\omega_0+\frac{1}{2}i\mu)$ and $\zeta=\pm(\omega-\frac{1}{2}i\mu)$. Accordingly, $G(\zeta)$ is regular outside a strip of width $2\alpha=\mu$. The associated Green's functions are given by

$$G_r(t-t')=\cosh(\tfrac{1}{2}\mu(t-t'))(\omega_0)^{-1}\sin(\omega_0(t-t'))\eta_+(t-t'),$$
$$G_a(t-t')=-\cosh(\tfrac{1}{2}\mu(t-t'))(\omega_0)^{-1}\sin(\omega_0(t-t'))\eta_-(t-t'),$$

and the function $w(t-t_2, t'-t_2)$, computed for $\omega_0(t-t_2)$, $\omega_0(t'-t_2)\gg1$, is

$$w(t-t_2, t'-t_2)$$
$$\simeq(\omega_0)^{-1}\cos(\omega_0(t-t'))[\coth(\tfrac{1}{2}\omega_0|\beta|)\sinh(\tfrac{1}{2}\mu(t-t_2))$$
$$\times\sinh(\tfrac{1}{2}\mu(t'-t_2))+\coth(\tfrac{1}{2}\omega_0\beta_0)$$
$$\times\cosh(\tfrac{1}{2}\mu(t-t_2))\cosh(\tfrac{1}{2}\mu(t'-t_2))].$$

After the larger time intervals $\mu(t-t_2)$, $\mu(t'-t_2)\gg1$, we have

$$w(t-t_2, t'-t_2)\sim(2\omega_0)^{-1}e^{+\frac{1}{2}\mu(t-t_2)}e^{+\frac{1}{2}\mu(t'-t_2)}\cos\omega_0(t-t')$$
$$\times[n_0+(1-e^{-\omega_0|\beta|})^{-1}],$$

with

$$n_0=(e^{\omega_0\beta_0}-1)^{-1}.$$

When t is in the vicinity of a time t_1, such that the amplification factor

$$k\simeq\tfrac{1}{2}e^{\frac{1}{2}\mu(t_1-t_2)}\gg1,$$

the oscillator is described by the classical coordinate

$$q(t)=k[q_s(t)+q_n(t)].$$

Here

$$q_s(t)=\int_{t_2}^{\infty}dt'\frac{1}{\omega_0}\sin(\omega_0(t-t'))F(t')e^{-\frac{1}{2}\mu(t'-t_2)}$$

and

$$q_n(t)=q_1\cos\omega t+q_2\sin\omega t,$$

where q_1 and q_2 are characterized by the expectation value formula

$$\langle e^{i(q_1f_1+q_2f_2)}\rangle=\exp[-(\nu/\omega_0)\tfrac{1}{2}(f_1^2+f_2^2)],$$

in which

$$\nu=n_0+(1-e^{-\omega_0|\beta|})^{-1}.$$

Accordingly, the probability of observing q_1 and q_2 within the range dq_1, dq_2 is

$$p(q_1q_2)dq_1dq_2=\frac{1}{2\pi}\frac{\omega_0}{\nu}\exp\left[-\frac{\omega_0}{\nu}\frac{1}{2}(q_1^2+q_2^2)\right]dq_1dq_2$$
$$=\frac{\omega_0}{\nu}\exp\left(-\frac{1}{2}\frac{\omega_0}{\nu}q_n^2\right)q_n dq_n\frac{1}{2\pi}d\varphi,$$

where q_n and φ are the amplitude and phase of $q_n(t)$. Despite rather different assumptions about the external system, these are the same conclusions as before, apart from a factor of $\frac{1}{2}$ in the formula for the gain.

GENERAL THEORY

The whole of the preceding discussion assumes an external system that is only slightly influenced by the presence of the oscillator. Now we must attempt to place this simplification within the framework of a general formulation. A more thorough treatment is also a practical necessity in situations such as those producing amplification of the oscillator motion, for a sizeable reaction in the external system must eventually appear, unless a counter mechanism is provided.

It is useful to supplement the previous Lagrangian operator with the term $q'(t)Q$, in which $q'(t)$ is an arbitrary numerical function of time, and also, to imagine the coupling term qQ supplied with a variable factor λ. Then

$$\frac{\partial}{\partial\lambda}\langle t_2|t_2\rangle^{F\pm q\pm'}$$
$$=i\left\langle\left|\int dt(q_+Q_+-q_-Q_-)\right|\right\rangle,$$
$$=-i\int_{t_2}^{t_1}dt\left(\frac{\delta}{\delta F_+(t)}\frac{\delta}{\delta q_+'(t)}-\frac{\delta}{\delta F_-(t)}\frac{\delta}{\delta q_-'(t)}\right)$$
$$\times\langle t_2|t_2\rangle^{F\pm q\pm'},$$

provided that the states to which the transformation function refers do not depend upon the coupling between the systems, or that the trace of the transformation function is being evaluated. A similar statement would apply to a transformation function with different terminal times. This differential equation implies an integrated form, in which the transformation function for the fully coupled system ($\lambda=1$) is expressed in terms of the transformation function for the uncoupled system ($\lambda=0$). The latter is the product of transformation functions for the independent oscillator and external system. The relation is

$$\langle t_2|t_2\rangle^{F\pm}=\exp\left[-i\int_{t_2}^{t_1}dt\left(\frac{\delta}{\delta F_+}\frac{\delta}{\delta q_+'}-\frac{\delta}{\delta F_-}\frac{\delta}{\delta q_-'}\right)\right]$$
$$\times\langle t_2|t_2\rangle_{\rm osc}^{F\pm}\langle t_2|t_2\rangle_{\rm ext}^{q\pm'}|_{q_\pm'=0},$$

and we have indicated that $q_\pm'$ is finally set equal to zero if we are concerned only with measurements on the oscillator.

Let us consider for the moment just the external system with the perturbation $q'Q$, the effect of which is indicated by[6]

$$\langle t_2|t_2\rangle^{q\pm'}=\left\langle t_2\left|\left(\exp\left[-i\int dtq_-'Q\right]\right)_-\right.\right.$$
$$\times\left.\left(\exp\left[i\int dtq_+'Q\right]\right)_+\right|t_2\Big\rangle.$$

We shall define

$$Q_+(t,q_\pm')=\frac{\langle t_2|Q_+(t)|t_2\rangle^{q\pm'}}{\langle t_2|t_2\rangle^{q\pm'}}$$
$$=\frac{1}{i}\frac{\delta}{\delta q_+'(t)}\log\langle t_2|t_2\rangle^{q\pm'}$$

and similarly

$$Q_-(t,q_\pm')=\frac{\langle t_2|Q_-(t)|t_2\rangle^{q\pm'}}{\langle t_2|t_2\rangle^{q\pm'}}$$
$$=-\frac{1}{i}\frac{\delta}{\delta q_-'(t)}\log\langle t_2|t_2\rangle^{q\pm'}.$$

When $q_\pm'(t)=q'(t)$, we have

$$Q_+(t,q')=Q_-(t,q')=\langle t_2|Q(t)|t_2\rangle^{q'},$$

which is the expectation value of $Q(t)$ in the presence of the perturbation described by $q'(t)$. This is assumed to be zero for $q'(t)=0$ and depends generally upon the history of $q'(t)$ between t_2 and the given time.

The operators $q_\pm(t)$ are produced within the transformation function by the functional differential operators $(\pm 1/i)\,\delta/\delta F_\pm(t)$, and since the equation of motion for the uncoupled oscillator is

$$\left(\frac{d^2}{dt^2}+\omega_0^2\right)q(t)=F(t),$$

we have

$$\left(\frac{\partial^2}{\partial t^2}+\omega_0^2\right)\left(\pm\frac{1}{i}\right)\frac{\delta}{\delta F_\pm(t)}\langle t_2|t_2\rangle^{F\pm}$$
$$=\exp\left[-i\int dt\left(\frac{\delta}{\delta F_+}\frac{\delta}{\delta q_+'}-\frac{\delta}{\delta F_-}\frac{\delta}{\delta q_-'}\right)\right]$$
$$\times F_\pm(t)\langle t_2|t_2\rangle_{\rm osc}^{F\pm}\langle t_2|t_2\rangle_{\rm ext}^{q\pm'}|_{q_\pm'=0}.$$

On moving $F_\pm(t)$ to the left of the exponential, this becomes

$$F_\pm(t)\langle t_2|t_2\rangle^{F\pm}+\exp[\ ]\langle t_2|t_2\rangle_{\rm osc}^{F\pm}$$
$$\times\left(\pm\frac{1}{i}\right)\frac{\delta}{\delta q_\pm'(t)}\langle t_2|t_2\rangle_{\rm ext}^{q\pm'}|_{q_\pm'=0}.$$

But

$$\left(\pm\frac{1}{i}\right)\frac{\delta}{\delta q_\pm'(t)}\langle t_2|t_2\rangle_{\rm ext}^{q\pm'}=Q_\pm(t,q_\pm')\langle t_2|t_2\rangle_{\rm ext}^{q\pm'},$$

and furthermore,

$$\exp[\ ]Q(t,q_\pm')\langle t_2|t_2\rangle_{\rm osc}^{F\pm}\langle t_2|t_2\rangle_{\rm ext}^{q\pm'}|_{q_\pm'=0}$$
$$=Q\left(t,\pm\frac{1}{i}\frac{\delta}{\delta F_\pm}\right)\langle t_2|t_2\rangle^{F\pm},$$

which leads us to the following functional differential equation for the transformation function $\langle t_2|t_2\rangle^{F\pm}$, in which a knowledge is assumed of the external system's reaction to the perturbation $q_\pm'(t)$:

$$\left[\left(\frac{\partial^2}{\partial t^2}+\omega_0^2\right)\left(\pm\frac{1}{i}\right)\frac{\delta}{\delta F_\pm(t)}-Q_\pm\left(t,\pm\frac{1}{i}\frac{\delta}{\delta F_\pm}\right)-F_\pm(t)\right]$$
$$\times\langle t_2|t_2\rangle^{F\pm}=0.$$

Throughout this discussion one must distinguish between the $\pm$ signs attached to particular components and those involved in the listing of complete sets of variables.

The differential equations for time development are supplemented by boundary conditions which assert, at a time t_1 beyond which $F_+(t)=F_-(t)$, that

$$\left(\frac{\delta}{\delta F_+(t_1)}+\frac{\delta}{\delta F_-(t_1)}\right)\langle t_2|t_2\rangle^{F\pm}=0$$

while, for the example of the transformation function $\langle t_2|t_2\rangle_{\vartheta_0\vartheta}{}^{F\pm}$, we have the initial conditions

$$\left[\left(\frac{\delta}{\delta F_+}-\frac{\delta}{\delta F_-}\right)(t_2)\right.$$
$$\left.+\frac{i}{\omega_0}\coth(\tfrac{1}{2}\omega_0\beta_0)\frac{\partial}{\partial t_2}\left(\frac{\delta}{\delta F_+}+\frac{\delta}{\delta F_-}\right)(t_2)\right]\langle t_2|t_2\rangle^{F\pm}=0$$

and

$$\left[\frac{\partial}{\partial t_2}\left(\frac{\delta}{\delta F_+}-\frac{\delta}{\delta F_-}\right)(t_2)\right.$$
$$\left.-i\omega_0\coth(\tfrac{1}{2}\omega_0\beta_0)\left(\frac{\delta}{\delta F_+}+\frac{\delta}{\delta F_-}\right)(t_2)\right]\langle t_2|t_2\rangle^{F\pm}=0.$$

The previous treatment can now be identified as the

[6] Such positive and negative time-ordered products occur in a recent paper by K. Symanzik [J. Math. Phys. 1, 249 (1960)], which appeared after this paper had been written and its contents used as a basis for lectures delivered at the Brandeis Summer School, July, 1960.

approximation of the $Q_\pm(t,q_\pm')$ by linear functions of $q_\pm'$,

$$Q_+(t,q_\pm')=i\int dt'[A_{++}(t-t')q_+'(t')-A_{+-}(t-t')q_-'(t')]$$

$$Q_-(t,q_\pm')=i\int dt'[A_{-+}(t-t')q_+'(t')-A_{--}(t-t')q_-'(t')],$$

wherein the linear equations for the operators $q_\pm(t)$ and their meaning in terms of variations of the $F_\pm$ have been united in one pair of functional differential equations. This relation becomes clearer if one writes

$$\langle t_2|t_2\rangle^{F_\pm}=\exp[iW(F_\pm)]$$

and, with the definition

$$q_\pm(t,F_\pm)=(\pm)\frac{\delta}{\delta F_\pm(t)}W(F_\pm)$$

$$=\left(\pm\frac{1}{i}\right)\frac{\delta}{\delta F_\pm(t)}\log\langle t_2|t_2\rangle^{F_\pm},$$

converts the functional differential equations into

$$\left(\frac{\partial^2}{\partial t^2}+\omega_0^2\right)q_\pm(t,F_\pm)-Q_\pm\left(t,\ q_\pm\pm\frac{1}{i}\frac{\delta}{\delta F_\pm}\right)-F_\pm(t)=0.$$

The boundary conditions now appear as

$$(q_+-q_-)(t_1,F_\pm)=0$$

and

$$(q_++q_-)(t_2,F_\pm)+\frac{i}{\omega_0}\coth(\tfrac{1}{2}\omega_0\beta_0)\frac{\partial}{\partial t_2}(q_+-q_-)(t_2,F_\pm)=0$$

$$\frac{\partial}{\partial t_2}(q_++q_-)(t_2,F_\pm)-i\omega_0\coth(\tfrac{1}{2}\omega_0\beta_0)(q_+-q_-)(t_2,F_\pm)=0.$$

When the $Q_\pm$ are linear functions of $q_\pm$, the functional differential operators disappear[7] and we regain the linear equations for $q_\pm(t)$, which in turn imply the quadratic form of $W(F_\pm)$ that characterizes the preceding discussion.

[7] The degeneration of the functional equations into ordinary differential equations also occurs when the motion of the oscillator is classical and free of fluctuation.

Wigner Coefficients for the R_4 Group and Some Applications*

L. C. BIEDENHARN
Department of Physics, Duke University, Durham, North Carolina
(Received October 14, 1960)

The local isomorphism of the R_4 group to the group $R_3 \times R_3$ is utilized to obtain R_4 Wigner coefficients for those representations in which the subgroup R_3 is diagonal. The R_4 Wigner coefficients so defined are then used to obtain recursion relations and differential equations for the representation coefficients, when the group is parametrized appropriately. The R_4 spherical harmonics, and their properties, are explicitly obtained as specializations of the general formulas. Physical application to the problem of geometrizing the Coulomb field is briefly discussed.

1. INTRODUCTION AND SUMMARY

ALTHOUGH the four-dimensional rotation group (R_4) is of intrinsic general interest for applications in physics, this group is of special interest for applications to (nonrelativistic) physical problems concerned with the Coulomb field. It has long been known that the symmetry group of the nonrelativistic Coulomb central field is that of four-dimensional rotations in projective momentum space.[1,2] From this point of view, one may justifiably say that problems involving the (nonrelativistic) Coulomb field are, using Fock's phrase, exercises in the "geometry of the Coulomb field." It is an enticing prospect to be able to treat the innumerable applications of the Coulomb field from such a general and unifying point of view,[1] and, in particular, the possible applications to the treatment of Coulomb excitation[3] furnished original motivation behind the present work.

The present work represents a first step in this program, and is concerned solely with constructing and obtaining relations for the representations of the R_4 group. This is accomplished by exploiting the well-known fact that R_4 is locally isomorphic to $R_3 \times R_3$ and thus homomorphic to $u_2 \times u_2$. One may immediately carry over the equally well-known results for the representations of u_2 to obtain representations of R_4. These are, however, *not* the representations desired, for in applications to the Coulomb field one seeks representations in which the subgroup R_3 is diagonal. This change of basis can be very economically accomplished by first obtaining the Wigner coefficients for reducing the Kronecker product of $R_4 \cdot R_4$ (Sec. 3), and then using these coefficients to obtain recursion formulas [Sec. 4(b)] and differential equations [Sec. 4(e)] for the representation coefficients. Section 4 also discusses the Wigner-Eckart theorem, matrices of the group generators, the R_4 spherical harmonics, and their addition law. A concluding section discusses implications of these results, and possible further extensions.

2. SUMMARY OF THE PROPERTIES OF THE R_4 GROUP

The rotation group in four dimensions has been treated in many places,[4] primarily as a particular instance of the general treatment of the characters of the n-dimensional rotation group. We shall, for convenience to the reader, summarize the features relevant for the results to be obtained in subsequent sections.

The group R_4 is the group of all linear substitutions, with positive determinant, that leave the quadratic form

$$\sum_{\alpha=1}^{4} x_\alpha^2$$

invariant; the group is generated by the six infinitesimal rotation operators $D_{\alpha\beta} = -D_{\beta\alpha}$ defined by

$$D_{\alpha\beta} = -i[x_\alpha(\partial/\partial x_\beta) - x_\beta(\partial/\partial x_\alpha)]. \tag{1}$$

For these operators, one has the commutation relations

$$\begin{aligned}[D_{\alpha\beta}, D_{\gamma\delta}] &\equiv D_{\alpha\beta}D_{\gamma\delta} - D_{\gamma\delta}D_{\alpha\beta} \\ &= i\delta_{\alpha\gamma}D_{\beta\delta} + i\delta_{\beta\delta}D_{\alpha\gamma} + i\delta_{\alpha\delta}D_{\gamma\beta} + i\delta_{\beta\gamma}D_{\delta\alpha}.\end{aligned} \tag{2}$$

These relations may be more readily understood by introducing two operators $\mathbf{L} = \hat{x}D_{23} + \hat{y}D_{31} + \hat{z}D_{12}$ and $\mathbf{A} = \hat{x}D_{14} + \hat{y}D_{24} + \hat{z}D_{34}$. The commutation relations (2) then read

$$\begin{aligned}[L_i, L_j] &= i\epsilon_{ijk}L_k, \\ [L_i, A_j] &= i\epsilon_{ijk}A_k, \\ [A_i, A_j] &= i\epsilon_{ijk}L_k.\end{aligned} \tag{3}$$

The linear combinations

$$\mathbf{M} \equiv \tfrac{1}{2}(\mathbf{L} + \mathbf{A}) \quad \text{and} \quad \mathbf{N} \equiv \tfrac{1}{2}(\mathbf{L} - \mathbf{A})$$

are of particular interest for they obey the commutation

* Supported in part by the U. S. Atomic Energy Commission, the National Aeronautics and Space Administration, and the Research Corporation.

[1] V. Fock, Z. Physik **98**, 145 (1935).

[2] V. Bargmann, Z. Physik **99**, 576 (1936).

[3] To be precise, the Coulomb excitation problem involves the Lorentz group and not R_4, since the states lie in the continuum, but by considering imaginary angles one can obtain suitable (algebraic) representations.

[4] F. D. Murnaghan, *The Theory of Group Representations* (The Johns Hopkins Press, Baltimore, Maryland, 1938); Hermann Weyl, *The Classical Groups* (Princeton University Press, Princeton, New Jersey, 1946); D. E. Littlewood, *The Theory of Group Characters* (Clarendon Press, Oxford, England, 1950).

rules[5]

$$\begin{aligned}[\mathbf{M},\mathbf{N}]&=0,\\ \mathbf{M}\times\mathbf{M}&=i\mathbf{M},\\ \mathbf{N}\times\mathbf{N}&=i\mathbf{N}.\end{aligned}\tag{4}$$

Since these are simply the commutation relations for two commuting three-dimensional angular momentum operators, one concludes that the group generated by the $D_{\alpha\beta}$ is isomorphic to the direct product of two three-dimensional rotation groups.

The commutation relations (2) show that the group is characterized by two invariants

$$I_1=\tfrac{1}{2}\sum_{\alpha\beta}D_{\alpha\beta}^2 \quad\text{and}\quad I_2=\tfrac{1}{2}\sum_{\alpha\beta\gamma\delta}\epsilon_{\alpha\beta\gamma\delta}D_{\alpha\beta}D_{\gamma\delta}.\tag{5}$$

These may be equivalently written as

$$\begin{aligned}I_1&=2(\mathbf{M}^2+\mathbf{N}^2)=\mathbf{L}^2+\mathbf{A}^2,\\ I_2&=\mathbf{M}^2-\mathbf{N}^2=\mathbf{L}\cdot\mathbf{A}.\end{aligned}\tag{6}$$

For the operators defined by (1), the invariant I_2 is always zero. The commutation relations include, however, more general (not necessarily two-valued) representations. By utilizing the fact that $R_4=R_3\times R_3$, we may write out the general group element as

$$g(R_4)=\exp(i\mathbf{a}\cdot\mathbf{M})\exp(i\mathbf{b}\cdot\mathbf{N}),$$

from which the general irreducible characters can be quickly obtained. Since the characters of R_3 are given by

$$\mathrm{X}^{(j)}(R_3)=\sin[\tfrac{1}{2}(2j+1)\vartheta]/\sin(\tfrac{1}{2}\vartheta),\quad \mathbf{M}^2\to j(j+1),$$

one obtains the general irreducible character of the R_4 group as[6]

$$\mathrm{X}^{[p,q]}(R_4)=\frac{\sin[\tfrac{1}{2}(2j_1+1)\vartheta_1]\sin[\tfrac{1}{2}(2j_2+1)\vartheta_2]}{\sin(\tfrac{1}{2}\vartheta_1)\sin(\tfrac{1}{2}\vartheta_2)},$$

where

$$\mathbf{M}^2\to j_1(j_1+1),\quad \mathbf{N}^2\to j_2(j_2+1),$$

and

$$p\equiv j_1+j_2,\quad q\equiv j_1-j_2,\quad p\geqslant q.\tag{7}$$

The irreducible characters of R_4 are Schur functions of the two angles—$\varphi_1\equiv\frac{1}{2}(\vartheta_1+\vartheta_2)$, $\varphi_2\equiv\frac{1}{2}(\vartheta_1-\vartheta_2)$—and are characterized by the two (partition) numbers p and q, where p is *positive* and q may be *positive or negative.* [If p, q are *integral* we have a true character; if p,q are *half-integral,* we have a two-valued (spin) representation.]

The characters denoted by $\mathrm{X}^{[p,q]}$ and $\mathrm{X}^{[p,-q]}$ are designated as *conjugate* characters. If $q=0$ (which implies that $\mathbf{L}\cdot\mathbf{A}=0$), the character is self-conjugate; solutions of the Laplace equation in four dimensions thus belong to self-conjugate representations.

Since the characters (7) are explicitly real and the representations unitary, it follows that the classes of R_4 are all *ambivalent,* i.e., every element is equivalent to its inverse.

Consider now the Kronecker product of two of these irreducible representations.[7] Since we have $R_4=R_3^{(1)}\times R_3^{(2)}$, one sees that

$$R_4\cdot R_4=(R_3^{(1)}\cdot R_3^{(1)})\times(R_3^{(2)}\cdot R_3^{(2)}).$$

By using the well-known result that

$$\mathrm{X}^{(j)}(R_3)\mathrm{X}^{(j'')}(R_3)=\sum_{J=|j_1-j_2|}^{j_1+j_2}\mathrm{X}^{(J)}(R_3),$$

one sees now that the Kronecker product $\mathrm{X}^{[p,q]}\mathrm{X}^{[p'',q'']}$ contains every irreducible representation $[P,Q]$ either *once or not at all.* This result, and the result that the classes of R_4 are ambivalent, are the two conditions required in order that the group R_4 be *simple reducible,* using the definition given by Wigner.[8] [One notes, incidentally, that if reflections are included (so that the group becomes O_4), the irreducible characters are now $\mathrm{X}^{[p,0]}$ and $\mathrm{X}^{[p,q]}+\mathrm{X}^{[p,-q]}$; *the group is no longer simply reducible.* Moreover, if p,q are half-integral, then the representation $\mathrm{X}^{[p,q]}+\mathrm{X}^{[p,-q]}$ now requires two independent spinors, one from each space $(\mathbf{M},\mathbf{N})$—a point very familiar now from neutrino theory!]

3. DETERMINATION OF THE R_4 WIGNER COEFFICIENTS

The Wigner coefficients were originally defined as the coefficients which reduce the Kronecker product of two irreducible representations of the R_3 (and u_2) group.[9] It is reasonable to extend this definition of the Wigner coefficients *to include the reduction of the Kronecker product of any simply reducible group.*

Let us consider first the ancillary task of parametrizing the R_4 group. There are two independent ways to do this, corresponding to the two independent ways to parametrize four dimensions with angular coordinates.

The first way utilizes

$$\begin{aligned}x_1&=r\cos\theta\cos\alpha, & x_3&=r\sin\theta\cos\beta,\\ x_2&=r\cos\theta\cos\alpha, & x_4&=r\sin\theta\sin\beta,\end{aligned}\tag{8}$$

which correspond to plane rotations in perpendicular spaces. The group element of R_4 is then conveniently given in the form

$$g=\exp(i\mathbf{a}\cdot\mathbf{M})\exp(i\mathbf{b}\cdot\mathbf{N}).\tag{9}$$

The parametrization most frequently encountered in physical applications, however, is the polar para-

[5] W. Pauli, CERN-56-31, Geneva, 1956 (unpublished).

[6] E. P. Wigner, *Group Theory* (Academic Press, Inc., New York, 1959); see theorems in Chap. 16, Secs. 2–4.

[7] We distinguish the two different types of direct product by a center dot · and a multiplication sign ×.

[8] E. P. Wigner, Am. J. Math. 63, 57 (1941).

[9] This is to be distinguished from the Clebsch-Gordan series, which treats the characters of the direct product. The so-called Clebsch-Gordan coefficients appear not to have been obtained by these authors.

metrization:

$$\begin{aligned} x_1&=r\cos\chi,\\ x_2&=r\sin\chi\cos\theta,\\ x_3&=r\sin\chi\sin\theta\sin\phi,\\ x_4&=r\sin\chi\sin\theta\cos\phi. \end{aligned} \tag{10}$$

Corresponding to this, one seeks to put the group elements in a form in which the rotations in the $(x_2x_3x_4)$ space are diagonal. The canonical form of the R_4 group elements are the two angles that determine the class; thus

$$\begin{aligned} g'&=\exp(i\vartheta_1M_z)\exp(i\vartheta_2N_z)\\ &=\exp\{\tfrac{1}{2}i[\vartheta_1(L_z+A_z)+\vartheta_2(L_z-A_z)]\}\\ &=\exp(i\varphi_2A_z)\exp(i\varphi_1L_z). \end{aligned} \tag{11}$$

The *general* element then is obtained as

$$g=R(e^{i\varphi_2A_z}e^{i\varphi_1L_z})S, \tag{12}$$

where R and S are arbitrary rotations that leave, say, the four-axis unchanged. On using the Euler form for the R and S, one finds that the general element of R_4 may be parametrized as

$$g(\alpha_1\cdots\alpha_6)=e^{i\alpha_1L_z}e^{i\alpha_2L_y}e^{i\alpha_3L_z}e^{i\alpha_4A_z}e^{i\alpha_5L_y}e^{i\alpha_6L_z}. \tag{13}$$

[The two redundant angles involved for arbitrary R, S have been eliminated in Eq. (13).]

This form is particularly advantageous in that only the matrices for the operator $e^{i\alpha_4A_z}$ need be investigated, the other matrices being well known.

Let us now turn to the problem of the Wigner coefficients. In the form given by Eq. (9), an irreducible basis is given by the state vectors $|j_1m_1;j_2m_2\rangle$. To obtain the basis corresponding to Eq. (13), one notes first that besides the two invariants $\mathbf{L}^2+\mathbf{A}^2$ and $\mathbf{L}\cdot\mathbf{A}$, the operators $\mathbf{L}^2$ and L_z are to be diagonal. Clearly then the required basis is given by

$$\begin{aligned} |pqlm\rangle&=|j_1+j_2,\ j_1-j_2lm\rangle\\ &=\sum_{m_1m_2}\langle j_1j_2m_1m_2|j_1j_2lm\rangle|j_1m_1;j_2m_2\rangle. \end{aligned} \tag{14}$$

[The matrix $\langle j_1j_2\cdots\rangle$ in this equation is the usual (R_3) Wigner coefficient.]

The representation matrices for the R_4 group are the matrices of Eq. (13) evaluated in the basis, Eq. (14). Thus, one has explicitly

$$\begin{aligned} &D_{l''m'';\,lm}{}^{[p,q]}(\alpha_1\cdots\alpha_6)\\ &=\langle pql''m''|e^{i\alpha_1L_z}e^{i\alpha_2L_y}e^{i\alpha_3L_z}e^{i\alpha_4A_z}\\ &\qquad\times e^{i\alpha_5L_y}e^{i\alpha_6L_z}|pqlm\rangle. \end{aligned} \tag{15}$$

To determine the Wigner coefficients one has to determine the matrices that reduce the Kronecker product; that is, one seeks the matrices $\mathbf{U}$ in the equation,

$$\mathbf{D}^{[p',q']}(\alpha)\cdot\mathbf{D}^{[p,q]}(\alpha)=\sum_{PQ}\mathbf{U}^{\dagger}D^{[P,Q]}(\alpha)\mathbf{U}. \tag{16}$$

Since, however, we already know the desired result in the basis corresponding to Eq. (9)—this is the standard Wigner result—it is clear that Eq. (16) presents an equally standard problem in the recoupling of angular momenta,[10,11] and the result is found to be

$$\begin{aligned} &\langle p'q'l'm';\,pqlm|\,\mathbf{U}\,|p'q'pq;\,PQLM\rangle\\ &=[(P+Q+1)(P-Q+1)(2l+1)(2l'+1)]^{\frac{1}{2}}\\ &\quad\cdot\langle l'lm'm|l'lLM\rangle\\ &\qquad\cdot X\begin{pmatrix}\frac{1}{2}(p'+q') & \frac{1}{2}(p+q) & \frac{1}{2}(P+Q)\\ \frac{1}{2}(p'-q') & \frac{1}{2}(p-q) & \frac{1}{2}(P-Q)\\ l' & l & L\end{pmatrix}. \end{aligned} \tag{17}$$

[In Eq. (17) the $X(\cdots)$ represents Wigner's $(9-j)$ coefficient, or, as it is also called, Fano's X coefficient; compare reference 11.]

4. APPLICATIONS OF THE R_4 WIGNER COEFFICIENT

The Wigner coefficients furnish the key to most of the group theoretical results that can be obtained. We shall in the present section briefly give some of these results; since we have reversed the usual procedure by obtaining the Wigner coefficients *before* we found the representation coefficients, one task will be to find various relations for the representations themselves.

(a) Wigner-Eckart Theorem

Let us suppose that the physical problem possesses the symmetry of R_4 so that the state vectors are of the form $|\alpha,pqlm\rangle$, where α represents the various other quantum numbers. Consider the matrix elements of the operator $\mathcal{O}$:

$$M\equiv\langle\alpha'';\,p''q''l''m''|\,\mathcal{O}_{\lambda\mu}{}^{[k,k']}|\alpha;\,pqlm\rangle. \tag{18}$$

This matrix element is unchanged upon inserting the unit operator $R^{-1}R$, where R is a rotation with arbitrary angles $(\alpha_1\cdots\alpha_6)$. On noting first that, by definition,

$$\begin{aligned} &R|\alpha;\,pqlm\rangle\\ &\quad=\sum_{l''m''}D_{l''m'';\,lm}{}^{[p,q]}(\alpha_1\cdots\alpha_6)|\alpha;\,pql''m''\rangle, \end{aligned} \tag{19}$$

and, secondly, the definition of an R_4 tensor operator,

$$R\mathcal{O}_{\lambda\mu}{}^{[kk']}R^{-1}=\sum_{\lambda''\mu''}D_{\lambda''\mu'';\,\lambda\mu}{}^{[k,k']}\mathcal{O}_{\lambda''\mu''}{}^{[k,k']}, \tag{20}$$

one finds the result [using Eqs. (16) and (17)] that

$$\begin{aligned} &\langle\alpha'';\,p''q''l''m''|\,\mathcal{O}_{\lambda\mu}{}^{[kk']}|\alpha;\,pqlm\rangle\\ &\quad=\langle pqlm;\,kk'\lambda\mu|\,U\,|pqkk';\,p''q''l''m''\rangle\\ &\qquad\times\langle\alpha''p''q''\|kk'\|\alpha pq\rangle. \end{aligned} \tag{21}$$

[10] L. C. Biedenharn, J. M. Blatt, and M. E. Rose, Revs. Modern Phys. 24, 249 (1952).

[11] A. R. Edmonds, *Angular Momentum in Quantum Mechanics* (Princeton University Press, Princeton, New Jersey, 1957); G. Racah and U. Fano, *Irreducible Tensorial Sets* (Academic Press, Inc., New York, 1959); M. E. Rose, *Elementary Theory of Angular Momentum* (John Wiley & Sons, Inc., New York, 1958).

[The matrix $\langle\cdots\|\cdots\|\cdots\rangle$ in Eq. (21) is the reduced matrix element.]

It is clear that this result is merely a direct transcription of Wigner's original result. One caution has to be noted in this proof; *it has been implicitly assumed that the generators of the symmetry group have the form of rotation operators in the coordinates of the physical problem.* This is necessary in order that the assumed result, $R(fg)=(Rf)g+f(Rg)$, holds.

(b) General Addition Theorem for Gegenbauer Polynomials

The composition law for group elements has the form of a general addition theorem.[12] Thus upon taking matrix elements of the composition law for the general group element in the form of Eq. (13), one obtains

$$D_{l'm',lm}{}^{[p,q]}(R_{12}) = \sum_{l''m''} D_{l'm';l''m''}{}^{[p,q]}(R_1)D_{l''m'';lm}{}^{[p,q]}(R_2), \quad (22)$$

with R_{12} being the product rotation, R_1R_2.

On specializing this result to the case where $l,m=0$, and taking the complex conjugate, one obtains

$$D_{l'm';00}{}^{[p,q]*}(R) = \sum_{l''m''} D_{l'm';l''m''}{}^{[p,q]*}(RS^{-1})D_{l''m'',00}{}^{[p,q]*}(S) = \sum_{l''m''} D_{l''m'';l'm'}{}^{[p,q]}(SR^{-1})D_{l''m'';00}{}^{[p,q]*}(S). \quad (23)$$

This result is in the standard form for the transformation of basis vectors, Eq. (19); thus we can identify

$$D_{lm;00}{}^{[p,q]*}(S)=(\text{const})\cdot|pqlm\rangle. \quad (24)$$

In order that $l'=0$ with a nonvanishing result, it is necessary that $p=q$. By analogy to the spherical harmonics one then defines the R_4 spherical harmonic:

$$Y_{nlm}(\chi\vartheta\varphi)\equiv\frac{n}{[2\pi^2]^{\frac{1}{2}}}D_{lm;00}{}^{[n-1,0]*}(\chi\vartheta\varphi). \quad (25)$$

[The use of $n=p+1$ stems from Fock.[1] The R_4 spherical harmonics are orthonormal over the interval $0\leqslant\chi,\ \vartheta\leqslant\pi,\ 0\leqslant\varphi\leqslant2\pi$ with $d\Omega_4=\sin^2\chi\sin\vartheta d\chi d\vartheta d\varphi$. These functions are (within a normalization) the Gegenbauer polynomials ($\nu=1$).]

Equation (23) has the form of a general addition theorem for the Y_{nlm}. An even more general addition theorem is contained in Eq. (22).

(c) Matrix Elements of the Group Generators

Matrix elements for the operators **L** and **A** in the basis $|pqlm\rangle$ are readily obtained by the same recoupling techniques used earlier. We summarize these results here.

The general result is

$$\left\langle p''q''l''m''\left|\binom{L_\mu}{A_\mu}\right|pqlm\right\rangle = \delta_{pp''}\delta_{qq''}\cdot\tfrac{1}{2}(2l+1)^{\frac{1}{2}}\langle l1m\mu|l1l''m''\rangle \times\Big\{[(p-q)(p+1-q)(p+2-q)]^{\frac{1}{2}} \times W\left(\frac{p-q}{2}1\frac{p+q}{2}l;\frac{p-q}{2}l''\right) +(\pm)[(p+q)(p+1+q)(p+2+q)]^{\frac{1}{2}} \times W\left(\frac{p+q}{2}1\frac{p-q}{2}l'';\frac{p+q}{2}l\right)\Big\}. \quad (26)$$

[The $W(\cdots)$ function in Eq. (26) is the Racah coefficient.[6,11]]

For L_μ these results are as expected. For A_μ more explicit algebraic results may be of value. Suppressing the indexes p, q, one has

$$\langle l''m''|A_\mu|lm\rangle=\langle l1m\mu|l1l''m''\rangle\cdot\langle l''\|A\|l\rangle$$

with

$$\langle l''\|A\|l\rangle = \begin{cases} \left[\dfrac{(p+l+2)(p-l)(l+1-q)(l+1+q)}{(l+1)(2l+3)}\right]^{\frac{1}{2}}, & l''=l+1 \\ \dfrac{q(p+1)}{[l(l+1)]^{\frac{1}{2}}}, & l''=l \\ -\left[\dfrac{(p+l+1)(p+1-l)(l+q)(l-q)}{l(2l-1)}\right]^{\frac{1}{2}}, & l''=l-1. \end{cases} \quad (27)$$

(d) Recursion Relations for the Representation Coefficients

Once we have the R_4 Wigner coefficients, it is a straightforward matter to obtain various recursion relations based upon the Kronecker product law, Eqs. (16) and (17). The general group element has the form of Eq. (13), and only matrix elements of the operator $e^{i\alpha_4A_3}$ are unknown. Thus, the general form is

$$\langle pql''m''|R(\alpha_1\cdots\alpha_6)|pqlm\rangle = \sum_{l'''m'''} D_{m'',m'''}{}^{(l'')}(\alpha_1\alpha_2\alpha_3) \times\langle pql''m'''|e^{i\alpha_4A_3}|pql'''m'''\rangle D_{m''',m}{}^{(l''')}(0\alpha_5\alpha_6), \quad (28)$$

where $D_{\cdots}{}^{(l)}(\cdots)$ are R_3 rotation matrices.[6]

We seek to determine recursion relations for the

[12] E. P. Wigner, "Application of group theory to the special functions of mathematical physics," Princeton lecture notes, 1955 (unpublished).

particular matrix element (diagonal in m):

$$\langle pql''m''|e^{i\chi A_z}|pqlm\rangle \equiv A_{l''m;lm}{}^{[p,q]}(\chi). \quad (29)$$

In order to begin the recursion process, we must calculate explicitly the simplest cases:

$p=q=\frac{1}{2}$:

$$\begin{aligned} &A_{l''m;lm}{}^{[\frac{1}{2},\pm\frac{1}{2}]}(\chi) \\ &\quad=\textstyle\sum_n[(i\chi)^n/n!]\langle\tfrac{1}{2}\pm\tfrac{1}{2}l''m|A_z{}^n|\tfrac{1}{2}\pm\tfrac{1}{2}lm\rangle \\ &\quad=(\pm)\delta_{l}{}^{l''}\delta_{l}{}^{1}e^{im\chi}, \quad m=\pm\tfrac{1}{2}. \end{aligned} \quad (30)$$

$p=1$, $q=0$:

$$A_{l''m,lm}{}^{[1,0]}(\chi)=\delta_m{}^0\cdot\begin{cases}\cos\chi, & l=l'' \\ i\sin\chi, & l\neq l'',\end{cases} \quad (31)$$

with $(l,l'')=(0,1)$.

Consider now the functions appropriate to the R_4 spherical harmonics $A_{lm;00}{}^{[p,0]}(\chi)$. If we apply Eq. (16) to these, we find

$$\begin{aligned} &A_{lm;00}{}^{[p,0]}(\chi)A_{l'm';00}{}^{[1,0]}(\chi) \\ &\quad=\sum_{PQLM}\langle p0lm;10l'm'|p010PQLM\rangle \\ &\qquad\times D_{LM;00}{}^{[P,Q]}(\chi)\cdot\langle p000;1000|p010PQ00\rangle. \end{aligned} \quad (32)$$

Since the R_4 Wigner coefficients are unitary (and also explicitly real), we can invert Eq. (16), and hence Eq. (32), to obtain

$$\begin{aligned} &\sum_{lm,l'm'}\langle p0lm;10l'm'|p010PQLM\rangle \\ &\qquad\times A_{lm;00}{}^{[P,0]}(\chi)A_{l'm';00}{}^{[1,0]}(\chi) \\ &\quad=A_{LM;00}{}^{[PQ]}(\chi)\cdot\langle p000;1000|p010PQ00\rangle. \end{aligned} \quad (33)$$

If we now choose the values $P=p$, $Q=1$ the right-hand side of Eq. (33) vanishes identically. On introducing the algebraic forms for the coefficients given in Eq. (17), one finds the desired recursion relation from Eq. (33). This result reads (noting $A_{lm;00}{}^{[p,0]}=\delta_m{}^0A_{l0;00}{}^{[p,0]}$)

$$\begin{aligned} &\cos\chi\cdot A_{l0;00}{}^{[p,0]}(\chi) \\ &\quad=-i\sin\chi\cdot A_{l+1,0;00}{}^{[p,0]}(\chi)\left[\frac{(p+1)^2-(l+1)^2}{(2l+3)(2l-1)}\right]^{\frac{1}{2}} \\ &\qquad+i\sin\chi\cdot A_{l-1,0;00}{}^{[p,0]}(\chi)\left[\frac{(p+1)^2-l^2}{(2l+1)(2l-1)}\right]^{\frac{1}{2}}. \end{aligned} \quad (34)$$

(Analogous functions to the $A^{[p,0]}$ were given by Fock[1]; the phase and normalization conventions differ from the above. Just as in the case of the spherical harmonics, the "natural" phase, i.e., the Condon-Shortley phase, is obtained directly from the generators of the group.)

[It is of interest to note that the recursion relation given in Eq. (34), where extended to include *imaginary* angles χ, then contains the recursion formula for the radial Coulomb function, $p+1\rightarrow i\eta$. Explicit discussion of this, and related questions, will be given elsewhere.]

The recursion relation given by Eq. (34) keeps p fixed and recurs on the index l. A second type of recursion relation would keep l fixed and recur on p. Such a relation can be easily obtained directly from Eq. (32). Thus, one finds that

$$\begin{aligned} &[(p+2)(p+2+l)(p+1-l)]^{\frac{1}{2}}A_{l0;00}{}^{[p+1,0]}(\chi) \\ &\quad-2(p+1)^{\frac{1}{2}}\cdot\cos\chi\cdot A_{l0;00}{}^{[p,0]}(\chi) \\ &\quad+[p(p+1+l)(p-l)]^{\frac{1}{2}}A_{l0;00}{}^{[p-1,0]}(\chi)=0. \end{aligned} \quad (35)$$

The final recursion relation which it is necessary to obtain is a recursion relation for the general representation coefficient, $A_{lm;l'm}{}^{[p,q]}(\chi)$. The technique for generating such a relation from the Wigner coefficients is clear; one uses the unitarity of these coefficients to invert the Kronecker product law, Eq. (16), preserving the values of p,q in the representation generated. The simplest such relation for the general representation coefficient is—unlike the R_3 case—now a *five-term* recursion relation.

This recursion relation is

$$\begin{aligned} 0=&A_{l+2,m;l'm}{}^{[p,q]}(\chi)\cdot\frac{\sin^2\chi}{[(2l+1)(2l+5)]^{\frac{1}{2}}}\cdot\frac{[(l+2)^2-m^2]^{\frac{1}{2}}[(l+1)^2-m^2]^{\frac{1}{2}}}{(l+1)(l+2)(2l+3)}\cdot \\ &\cdot[(p+1)^2-(l+2)^2]^{\frac{1}{2}}\cdot[(p+1)^2-(l+1)^2]^{\frac{1}{2}}[(l+2)^2-q^2]^{\frac{1}{2}}[(l+1)^2-q^2]^{\frac{1}{2}} \\ &+A_{l+1,m;l'm}{}^{[p,q]}(\chi)\cdot\frac{\sin\chi}{[(2l+1)(2l+3)]^{\frac{1}{2}}}\cdot[(l+1)^2-m^2]^{\frac{1}{2}}[(p+1)^2-(l+1)^2]^{\frac{1}{2}}[(l+1)^2-q^2]^{\frac{1}{2}}\cdot \\ &\cdot\left[-2i\cos\chi-2m\sin\chi\cdot\frac{q(p+1)}{l(l+1)(l+2)}\right] \\ &+A_{l,m;l'm}{}^{[p,q]}(\chi)\cdot\left[(l''-l)(l''+l+1)-\frac{2\sin^2\chi}{(2l-1)(2l+3)}\cdot\right. \\ &\left.\cdot\left\{l(l+1)[m^2+2-3l(l+1)]+\frac{[l(l+1)-3m^2]q^2(p+1)^2}{l(l+1)}+[(p+1)^2+q^2-1][l(l+1)+m^2-1]\right\}\right] \\ &+A_{l-1,m;l'm}{}^{[p,q]}(\chi)\cdot\frac{\sin\chi}{[(2l-1)(2l+1)]^{\frac{1}{2}}}\cdot[l^2-m^2]^{\frac{1}{2}}[(p+1)^2-l^2]^{\frac{1}{2}}[l^2-q^2]^{\frac{1}{2}}\cdot\left[2i\cos\chi+2m\sin\chi\cdot\frac{q(p+1)}{(l-1)(l)(l+1)}\right] \\ &+A_{l-2,m;l'm}{}^{[p,q]}(\chi)\cdot\frac{\sin^2\chi}{[(2l+1)(2l-3)]^{\frac{1}{2}}}\cdot\frac{[l^2-m^2]^{\frac{1}{2}}[(l-1)^2-m^2]^{\frac{1}{2}}}{l(l-1)(2l-1)} \\ &\qquad\cdot[(p+1)^2-l^2]^{\frac{1}{2}}[(p+1)^2-(l-1)^2]^{\frac{1}{2}}[l^2-q^2]^{\frac{1}{2}}[(l-1)^2-q^2]^{\frac{1}{2}}. \end{aligned} \quad (36)$$

[The representation $A_{l0;00}{}^{[2,0]}(\chi)$ is utilized in obtaining this result, and the phase conventions are important. These coefficients are

$$A_{l0;00}{}^{[2,0]}(\chi)=\begin{cases}1/3(4\cos^2\chi-1) & l=0\\ i(8/3)^{\frac{1}{2}}\sin\chi\cos\chi & l=1\\ -(8^{\frac{1}{2}}/3)\sin^2\chi & l=2.\end{cases}]$$

The square roots in Eq. (36) may be eliminated by redefining the elements $A^{[p,q]}$ to include appropriate factorials.

For completeness, two further properties of the representation coefficients are recorded here. The first is the relation between the A::: and the R_4 spherical harmonics. This is implicit in the definition, but explicitly one has

$$Y_{nlm}(\chi\vartheta\varphi)=\left[\frac{2n^2}{\pi(2l+1)}\right]^{\frac{1}{2}}A_{l0;00}{}^{[n-1,0]*}Y_l{}^m(\vartheta\varphi). \quad (37)$$

The second relation is the behavior of the representation coefficients under complex conjugation. Since the generators, according to the commutation rules, change sign under complex conjugation, only the behavior of the basis vectors under time reversal enters. With the usual phase convention this is

$$T|pqlm\rangle=(-)^{l-m}|pql-m\rangle.$$

Thus, one obtains

$$D_{l''m'';lm}{}^{[p,q]}(\alpha)^*=(-)^{l-l''+m-m''}D_{l'',-m'';l,-m}{}^{[p,q]}(\alpha). \quad (38)$$

As a particular case of this, the R_4 spherical harmonics satisfy the rule

$$Y_{nlm}(\chi\vartheta\varphi)^*=(-)^{l-m}Y_{nl,-m}(\chi\vartheta\varphi). \quad (39)$$

(e) Differential Equations for the Representation Coefficients

The basic differential relation is implicit in the definition of the representation coefficient $A^{[p,q]}(\chi)$, Eq. (29). Differentiation with respect to χ yields the result

$$\frac{d}{d\chi}[A_{lm;l''m}{}^{[p,q]}(\chi)]=i\sum_{l'''}\langle pqlm|A_z|pql'''m\rangle A_{l'''m;l''m}{}^{[p,q]}(\chi). \quad (40)$$

On introducing the explicit values for the matrix elements, one obtains

$$\frac{d}{d\chi}[A_{lm;l''m}{}^{[p,q]}(\chi)]=i\left[\frac{[(p+1)^2-l^2][l^2-q^2][l^2-m^2]}{l^2(2l-1)(2l+1)}\right]^{\frac{1}{2}}A_{l-1,m;l''m}{}^{[p,q]}(\chi)+\frac{imq(p+1)}{l(l+1)}\cdot A_{lm;l''m}{}^{[p,q]}(\chi)$$
$$+i\left[\frac{[(p+1)^2-(l+1)^2][(l+1)^2-m^2][(l+1)^2-q^2]}{(l+1)^2(2l+1)(2l+3)}\right]^{\frac{1}{2}}A_{l+1,m;l''m}{}^{[p,q]}(\chi). \quad (41)$$

Consider now the special case of the R_4 spherical harmonics. The differential relation Eq. (41) takes then a very simple form:

$$\frac{d}{d\chi}[A_{l0;00}{}^{[p,0]}]=il\left[\frac{(p+1)^2-l^2}{(2l-1)(2l+1)}\right]^{\frac{1}{2}}A_{l-1,0;00}{}^{[p,0]}$$
$$+i(l+1)\left[\frac{(p+1)^2-(l+1)^2}{(2l+1)(2l+3)}\right]^{\frac{1}{2}}A_{l+1,0;00}{}^{[p,0]}. \quad (42)$$

This result, and the three-term recursion relations, Eq. (34), are of the standard form for the hypergeometric functions. It is convenient to combine these into the usual form of "raising" and "lowering" operators:

$$\frac{d}{d\chi}(A_{l0;00}{}^{[p,0]})-l\cdot\cot\chi\cdot A_{l0;00}{}^{[p,0]}=i\left[\frac{[(p+1)^2-(l+1)^2][2l+1]}{2l+3}\right]^{\frac{1}{2}}A_{l+1,0;00}{}^{[p,0]}; \quad (43a)$$

$$\frac{d}{d\chi}(A_{l,0;00}{}^{[p,0]})+(l+1)\cdot\cot\chi\cdot A_{l0;00}{}^{[p,0]}=i\left\{\frac{[(p+1)^2-l^2][2l+1]}{2l-1}\right\}^{\frac{1}{2}}A_{l-1,0;00}{}^{[p,0]}. \quad (43b)$$

These equations imply that the $A^{[p,0]}$ satisfy the differential equation

$$\left[\frac{d^2}{d\chi^2}+2\cot\chi\frac{d}{d\chi}+p(p+2)-\frac{l(l+1)}{\sin^2\chi}\right]\times A_{l0;00}{}^{[p,0]}(\chi)=0. \quad (44)$$

The general solution then follows, and is found to be

$$A_{l0;00}{}^{[p,0]}(\chi)=i^l[2l+1]^{\frac{1}{2}}\left[\frac{\Gamma(l+p+2)}{(p+1)\Gamma(p+1-l)}\right]^{\frac{1}{2}}\times\frac{\Gamma(\frac{3}{2})}{2^l\Gamma(l+\frac{3}{2})}\cdot[\sin\chi]^l\times{}_2F_1(l-p,l+2+p,l+\tfrac{3}{2};\tfrac{1}{2}(1-\cos\chi)). \quad (45)$$

A useful form of this, completely analogous to that of the usual spherical harmonics, is the relation

$$A_{l0;00}{}^{[p,0]}(\chi)=i^l[2l+1]^{\frac{1}{2}}\left[\frac{\Gamma(p+1-l)}{(p+1)\Gamma(p+2+l)}\right]^{\frac{1}{2}}\cdot[\sin\chi]^l\frac{d^l}{d(\cos\chi)^l}\left[\frac{\sin(p+1)\chi}{\sin\chi}\right]. \quad (46)$$

To obtain the differential equation—and the explicit solution—for the general representation coefficients is a much more involved task than for the special case involved for the (R_4) spherical harmonics. Although the differential relation, Eq. (40), is not appreciably more difficult, the fact that a five-term recursion relation is necessary, implies that the desired differential equation is of the fourth order.

It is fortunate, however, that—unlike the general case for a three-term derivative relation and a five-term recursion formula—it is still possible to define raising and lowering operators for the general representation coefficient. Upon using Eqs. (41) and (36), one finds for the raising operator (after considerable manipulation!) the result

$$\Theta_{pq;ll''m}{}^{(+)}A_{lm;l''m}{}^{[p,q]}=-A_{l+1,m;l''m}{}^{[p,q]}, \quad (47a)$$

with the operator $\Theta^{(+)}$ defined as

$$\Theta_{pq;ll''m}{}^{(+)}\equiv\tfrac{1}{2}(l+1)[(p+2+l)(p-l)(l+1-q)\times(l+1+q)(l+1+m)(l+1-m)]^{-\frac{1}{2}}\cdot\left[\frac{2l+1}{2l+3}\right]^{-\frac{1}{2}}\left\{i\cot\chi+\frac{2mq(p+1)}{l(l+2)(2l+1)}\right\}^{-1}\cdot\left\{\frac{d^2}{d\chi^2}-2l\cot\chi\frac{d}{d\chi}+m^2+q^2+p(p+2)-2l(l+1)+\frac{2imq(p+1)}{l+1}\cot\chi+\frac{(l-l'')(l+l''+1)}{\sin^2\chi}\right\}. \quad (47b)$$

The lowering operator is found to be

$$\Theta_{pq;ll''m}{}^{(-)}A_{lm;l''m}{}^{[pq]}=A_{l-1,m;l''m}{}^{[p,q]}, \quad (48a)$$

with the operator $\Theta^{(-)}$ defined as

$$\Theta_{pq;ll''m}{}^{(-)}\equiv\tfrac{1}{2}l\left[(p+1+l)(p+1-l)(l^2-q^2)(l^2-m^2)\left(\frac{2l-1}{2l+1}\right)\right]^{-\frac{1}{2}}\cdot\left\{i\cot\chi+\frac{2mq(p+1)}{l(l+2)(2l+1)}\right\}^{-1}\cdot\left\{\frac{d^2}{d\chi^2}+2(l+1)\left[\cot\chi-\frac{2imq(p+1)}{l(l+1)(l+2)}\right]\frac{d}{d\chi}+m^2+q^2+p(p+2)-2l(l+1)-\frac{m^2q^2(p+1)^2}{l(l+1)(l+2)}+\frac{(l-l'')(l''+l+1)}{\sin^2\chi}-\frac{2imq(p+1)}{l}\cot\chi\right\}. \quad (48b)$$

These results then lead to the fourth-order differential equation which the general representation coefficient satisfies, i.e.,

$$[\Theta_{pq;l+1,l''m}{}^{(-)}\Theta_{pq;ll''m}{}^{(+)}+1]A_{lm;l''m}{}^{[pq]}(\chi)=0. \quad (49)$$

This differential equation, as one might expect, is quite complicated when written out explicitly, and since the analysis of such equations is not well developed, it will not be treated further.

Since the application of the raising operator to the coefficient $A_{pm;l''m}{}^{[p,q]}$ must give zero, one obtains a second-order differential equation for this special case. Specializing further to the case where $q=0$ (which is of interest for the transformation of the R_4 spherical harmonics), one finds the result that

$$A_{pm;l''m}{}^{[p0]}(\chi)=(2i)^{p-l''}\frac{p!}{(l'')!}\left[\frac{(p+m)!(p-m)!(2l'')!(2l''+1)!}{(l''-m)!(l''+m)!(2p)!(p-l'')!(p+l''+1)!}\right]^{\frac{1}{2}}\cdot[\sin\chi]^{p-l''}\,{}_2F_1[m-l'',-m-l'',\tfrac{1}{2}-l'';\tfrac{1}{2}(1-\cos\chi)]. \quad (50)$$

(The second solution to the differential equation is eliminated by the required behavior at $\chi=0$.) It may be noted that the coefficient $A_{p0;l''0}{}^{[p,q]}$ satisfies the same differential equation with $q\leftrightarrow m$.) For the representation coefficients with $l\neq p$, the lowering operator may be used successively.

One may obtain explicit solutions for the case $l=q$, by exploiting the vanishing of the lowering operator. Thus for the special case where $m=0$, one obtains

$$A_{q0;l''0}{}^{[p,q]}(\chi)=i^{l''-q}\cdot\left[\frac{(q+\frac{1}{2})!(l'')!(l''+q)!(p-q)!(p+l''+1)!}{q!(l''+\frac{1}{2})!(2l'')!(l''-q)!(p+1+q)!(p-l'')!}\right]^{\frac{1}{2}}\cdot[\sin\chi]^{l''-q}\,{}_2F_1[l''-p,l''+p+2,l''+\tfrac{3}{2};\tfrac{1}{2}(1-\cos\chi)]. \quad (51)$$

The results given in Eqs. (50) and (51) are seen to include Eq. (45) as a special case.

To conclude this section it would be desirable to exhibit the explicit general solution to the fourth-order differential equation, Eq. (49). This can be done, and the result is given in Eq. (52). It should be noted, however, that this result has not been obtained directly from the differential equations, as the previous results were, but from an exploitation of the isomorphism used earlier in Sec. 3. Consequently, this general result is not so useful for explicit evaluation as, for example, Eq. (50) or (51).

The general representation coefficient has the explicit form

$$A_{lm;l''m}{}^{[pq]}(\chi)=e^{-im\chi}\cdot[(2l+1)(2l''+1)]^{\frac{1}{2}}\cdot[(l+m)!(l-m)!(l''+m)!(l''-m)!]^{\frac{1}{2}}$$
$$\cdot\left[\frac{(p-l)!(l+q)!(l-q)!(p-l'')!(l''+q)!(l''-q)!}{(p+1+l)!(p+1+l'')!}\right]^{\frac{1}{2}}\cdot\sum_{\kappa\lambda\mu}(-)^{\kappa+\lambda}e^{2i\mu\chi}\frac{[\frac{1}{2}(p+q)-\mu]![\frac{1}{2}(p+q)+\mu]!}{\kappa!\lambda!(p-l-\kappa)!(p-l''-\lambda)!}$$
$$\times\frac{[\frac{1}{2}(p-q)-\mu]![\frac{1}{2}(p-q)+\mu]!}{[\frac{1}{2}(p+q)-\mu-\kappa]![\frac{1}{2}(p+q)-\mu-\lambda]![l-\frac{1}{2}(p-q)+\mu+\kappa]![l''-\frac{1}{2}(p-q)+\mu+\lambda]!}$$
$$\times\frac{1}{[\frac{1}{2}(p-q)+m-\mu-\kappa]![\frac{1}{2}(p-q)+m-\mu-\lambda]![l-\frac{1}{2}(p+q)-m+\mu+\kappa]![l''-\frac{1}{2}(p+q)-m+\mu+\lambda]!}. \quad (52)$$

The summations in Eq. (52) are over all integers, positive and negative, including zero. This result is necessarily quite complicated, but it reduces properly, after some manipulation, to the special cases given earlier.

5. CONCLUDING REMARKS

We have in the preceding sections exploited the homomorphism of the R_4 group to the $u_2\times u_2$ group to obtain the R_4 Wigner coefficients. The application of these coefficients to effect a complete determination of the representation matrix elements and their properties, is essentially only a detailed elaboration of the results of group theory for the special functions of mathematical physics[12]; the sole novelty of the present work is the central role played by the Wigner coefficients.

The real interest in the results obtained lies in their physical applications. As mentioned in the Introduction, the original motivation of the present work came from the possibility of "geometrizing" the Coulomb field. In his pioneer paper,[13] Pauli showed that the nonrelativistic central Coulomb field problem could be characterized by two three-vectors: the angular momentum, $\mathbf{L}=\hbar^{-1}\mathbf{r}\times\mathbf{p}$, and the Runge-Lenz vector,

$$\mathfrak{A}\equiv(2Ze^2m\hbar)^{-1}[\mathbf{L}\times\mathbf{p}-\mathbf{p}\times\mathbf{L}]+\hat{r},$$

both of which commuted with the Hamiltonian $H=(p^2/2m)-(Ze^2/r)$. If one defines a new vector $\mathbf{A}$ by the equation: $\mathbf{A}\equiv(-2H/Z^2e^2m)^{-\frac{1}{2}}\mathfrak{A}$, the commutation relations for $\mathbf{L}$ and $\mathbf{A}$ are precisely those of Eq. (3); the Hamiltonian then takes the form:

$$\mathbf{A}^2+\mathbf{L}^2+1=(-Z^2e^4m/2\hbar^2H)^{-1}\rightarrow(p+1)^2.$$

The difficulty in applying these results to the complete geometrizing of Coulomb field—which otherwise would be already contained in the Wigner-Eckart theorem of Sec. 4(a)—lies in the fact that the realization of the operators $\mathbf{L}$ and $\mathbf{A}$ in either configuration- or momentum-space does *not* satisfy the necessary operator rule $\mathcal{O}(fg)=f\mathcal{O}g+g\mathcal{O}f$ which underlies our development. Bargmann[2] did achieve a partial result along these lines, but only at the expense of replacing the operator H by a c-number operator. His results, like Fock, employed a projective momentum space, but the radius of the projective sphere is then a function of H. Only within a subspace of given energy can the results obtained in previous sections be directly applied to evaluate matrix elements of Coulomb field problems. We shall not detail such results here, but it is interesting to note that this is apparently the group theoretical basis for the often observed fact that the "no-energy loss" Coulomb excitation process is of striking simplicity.[14] There seems to be no basic reason why an appropriate canonical transformation cannot be found for the desired operator realization, but no such transformation has been found as yet.

Besides the true representations, the results of the previous sections also include (two-component) spinor representations. These are of interest in that they provide a natural extension of the Coulomb field problem to particles with spin. Now spin $\frac{1}{2}$ enters naturally only through relativistic considerations, and, as is well known, the symmetry of the Coulomb field problem is destroyed by both the relativistic change of mass with energy, and the spin precessional effects. Nonetheless, the extension of the Coulomb field spinor results allows a form of *approximate* relativistic Coulomb field problem of high symmetry. An example of such a

[13] W. Pauli, Z. Physik **36**, 336 (1926).

[14] L. C. Biedenharn and C. M. Class, Phys. Rev. **98**, 691 (1955). This simplicity extends even to the relativistic problem [compare Mary E. Young and L. C. Biedenharn, Bull. Am. Phys. Soc. Ser. II, **5**, 112 (1960)].

result is the well-known Sommerfeld-Maue approximate solution. The R_4 Wigner coefficients are useful here, since they allow the introduction R_4 coupled spinors, analogous to the Pauli central field spinors χ_κ^μ introduced earlier.[15] A detailed discussion of the extensions mentioned here will be given elsewhere.

ACKNOWLEDGMENTS

A part of the work reported here was carried out while the author was a Fulbright Fellow at the University of Manchester, and a Guggenheim Fellow at the Institute for Theoretical Physics, Copenhagen; the aid of these organizations is gratefully acknowledged.

[15] M. E. Rose, L. C. Biedenharn, and G. B. Arfken, Phys. Rev. 85, 5 (1952).

Note added in proof. Dr. M. Hamermesh has kindly called my attention to the papers of A. Z. Dolginov, Soviet Phys.—JETP 3, 589 (1956) and A. Z. Dolginov and I. N. Toptygin, Soviet Phys.—JETP 8, 550 (1959), in which these authors have earlier discussed some of the results obtained above. In particular, these authors have defined the Wigner coefficient for the R_4 group [although, to be precise, the result they gave (their Eq. 27) is not quite as general as Eq. 17, above]. The applicability of the R_3 Wigner coefficients to the group R_4 has been noted, independently, by several other authors recently: D. Park, Z. Physik **159**, 155 (1960); James D. Louck, Los Alamos Scientific Laboratory Rept. LA-2451 (October, 1960); and by W. T. Sharp in his comprehensive treatment of Racah algebras (thesis, Princeton University, 1960) issued as A.E.C.L. 1098, Chalk River, Ontario, September, 1960.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 2, NUMBER 3 MAY–JUNE, 1961

Ellipsoidal Distributions of Charge or Mass*

B. C. CARLSON
Institute for Atomic Research and Department of Physics, Iowa State University, Ames, Iowa
(Received September 19, 1960)

The Coulomb (or gravitational) energy is calculated for a distribution of charge (or mass) in which the surfaces of constant density are a family of similar concentric ellipsoids. The density can vary in any manner from one surface to another, and the ellipsoids need not have an axis of symmetry. Two examples are discussed: the charge distribution of a deformed atomic nucleus having a diffuse surface, and the mass distribution of a stellar galaxy. The energy is shown to be a product of two factors. One is the energy of the spherical distribution from which the ellipsoidal distribution can be obtained by a volume-preserving deformation. The other is an anisotropy factor that has a simple geometrical significance and depends only on the two eccentricities of the ellipsoids. Its values range from unity to zero and are tabulated numerically.

I. INTRODUCTION

AMONG the largest as well as the smallest physical systems, there occur distributions of matter that have an ellipsoidal shape and are composed of elements that exert inverse-square-law forces on one another. Approximately a quarter of the brightest galaxies belong to a type known as elliptical galaxies, in which the density of stars decreases smoothly in all directions from the center. They generally contain very little interstellar gas or dust, show almost no internal structure, and have nearly ellipsoidal shapes with axial symmetry and various degrees of flattening.[1] Another quarter consists of lenticular galaxies, many of which are similar to the ellipticals but show a little differentiation between internal regions. The remaining lenticulars and most of the spiral galaxies show too much internal structure to be approximated, except in a crude fashion, by a mass distribution with ellipsoidal surfaces of constant density.

Near the other extreme of size, the atomic nuclei of the rare-earth and actinide regions, as well as some of the light nuclei, have equilibrium deformations of ellipsoidal shape. The departures from spherical symmetry are never as large as they sometimes are in the elliptical galaxies, but both prolate and oblate deformations occur and there is substantial evidence for shapes without axial symmetry (triaxial ellipsoids).[2] Unlike the gravitational attractions between the stars of a galaxy, the repulsive Coulomb forces between the protons in a nucleus do not control the dynamics of the system. Nevertheless, the Coulomb energy is an important term in the semiempirical mass formula; it supplies the energy released in fission and plays an important part in the energetics of any disintegration into charged fragments.

To a first approximation, the charge distribution of a deformed nucleus can be treated as a homogeneous ellipsoid (a uniform charge density bounded by a sharp ellipsoidal surface). Since the potential in the interior of a homogeneous ellipsoid is a quadratic function of the Cartesian position coordinates,[3] it is easy to obtain

* Work performed in the Ames Laboratory of the U. S. Atomic Energy Commission.

[1] G. de Vaucouleurs, *Handbuch der Physik* (Springer-Verlag, Berlin, Germany, 1959), Vol. LIII, p. 305.

[2] The asymmetric rotor theory of Davydov and Filippov is compared with experimental data by D. M. Van Patter, Nuclear Phys. 14, 42 (1959).

[3] O. D. Kellogg, *Foundations of Potential Theory* (Springer-Verlag, Berlin, Germany, 1929), p. 194.

result is the well-known Sommerfeld-Maue approximate solution. The R_4 Wigner coefficients are useful here, since they allow the introduction R_4 coupled spinors, analogous to the Pauli central field spinors χ_κ^μ introduced earlier.[15] A detailed discussion of the extensions mentioned here will be given elsewhere.

ACKNOWLEDGMENTS

A part of the work reported here was carried out while the author was a Fulbright Fellow at the University of Manchester, and a Guggenheim Fellow at the Institute for Theoretical Physics, Copenhagen; the aid of these organizations is gratefully acknowledged.

[15] M. E. Rose, L. C. Biedenharn, and G. B. Arfken, Phys. Rev. 85, 5 (1952).

Note added in proof. Dr. M. Hamermesh has kindly called my attention to the papers of A. Z. Dolginov, Soviet Phys.—JETP 3, 589 (1956) and A. Z. Dolginov and I. N. Toptygin, Soviet Phys.—JETP 8, 550 (1959), in which these authors have earlier discussed some of the results obtained above. In particular, these authors have defined the Wigner coefficient for the R_4 group [although, to be precise, the result they gave (their Eq. 27) is not quite as general as Eq. 17, above]. The applicability of the R_3 Wigner coefficients to the group R_4 has been noted, independently, by several other authors recently: D. Park, Z. Physik **159**, 155 (1960); James D. Louck, Los Alamos Scientific Laboratory Rept. LA-2451 (October, 1960); and by W. T. Sharp in his comprehensive treatment of Racah algebras (thesis, Princeton University, 1960) issued as A.E.C.L. 1098, Chalk River, Ontario, September, 1960.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 2, NUMBER 3 MAY–JUNE, 1961

Ellipsoidal Distributions of Charge or Mass*

B. C. CARLSON
Institute for Atomic Research and Department of Physics, Iowa State University, Ames, Iowa
(Received September 19, 1960)

The Coulomb (or gravitational) energy is calculated for a distribution of charge (or mass) in which the surfaces of constant density are a family of similar concentric ellipsoids. The density can vary in any manner from one surface to another, and the ellipsoids need not have an axis of symmetry. Two examples are discussed: the charge distribution of a deformed atomic nucleus having a diffuse surface, and the mass distribution of a stellar galaxy. The energy is shown to be a product of two factors. One is the energy of the spherical distribution from which the ellipsoidal distribution can be obtained by a volume-preserving deformation. The other is an anisotropy factor that has a simple geometrical significance and depends only on the two eccentricities of the ellipsoids. Its values range from unity to zero and are tabulated numerically.

I. INTRODUCTION

AMONG the largest as well as the smallest physical systems, there occur distributions of matter that have an ellipsoidal shape and are composed of elements that exert inverse-square-law forces on one another. Approximately a quarter of the brightest galaxies belong to a type known as elliptical galaxies, in which the density of stars decreases smoothly in all directions from the center. They generally contain very little interstellar gas or dust, show almost no internal structure, and have nearly ellipsoidal shapes with axial symmetry and various degrees of flattening.[1] Another quarter consists of lenticular galaxies, many of which are similar to the ellipticals but show a little differentiation between internal regions. The remaining lenticulars and most of the spiral galaxies show too much internal structure to be approximated, except in a crude fashion, by a mass distribution with ellipsoidal surfaces of constant density.

Near the other extreme of size, the atomic nuclei of the rare-earth and actinide regions, as well as some of the light nuclei, have equilibrium deformations of ellipsoidal shape. The departures from spherical symmetry are never as large as they sometimes are in the elliptical galaxies, but both prolate and oblate deformations occur and there is substantial evidence for shapes without axial symmetry (triaxial ellipsoids).[2] Unlike the gravitational attractions between the stars of a galaxy, the repulsive Coulomb forces between the protons in a nucleus do not control the dynamics of the system. Nevertheless, the Coulomb energy is an important term in the semiempirical mass formula; it supplies the energy released in fission and plays an important part in the energetics of any disintegration into charged fragments.

To a first approximation, the charge distribution of a deformed nucleus can be treated as a homogeneous ellipsoid (a uniform charge density bounded by a sharp ellipsoidal surface). Since the potential in the interior of a homogeneous ellipsoid is a quadratic function of the Cartesian position coordinates,[3] it is easy to obtain

* Work performed in the Ames Laboratory of the U. S. Atomic Energy Commission.

[1] G. de Vaucouleurs, *Handbuch der Physik* (Springer-Verlag, Berlin, Germany, 1959), Vol. LIII, p. 305.

[2] The asymmetric rotor theory of Davydov and Filippov is compared with experimental data by D. M. Van Patter, Nuclear Phys. 14, 42 (1959).

[3] O. D. Kellogg, *Foundations of Potential Theory* (Springer-Verlag, Berlin, Germany, 1929), p. 194.

the Coulomb energy by integrating the product of the potential and the constant charge density over the interior of the ellipsoid. (Alternatively, for small deformations, one can expand the equation of the surface in spherical harmonics of degree two and calculate the Coulomb energy to lowest order in the expansion coefficients.) For a better approximation, one would like to take into account the diffuseness of the nuclear surface; furthermore, the mass distribution in a galaxy tapers off gradually and does not resemble a homogeneous ellipsoid.

In order to discuss the energy of such distributions, which are not known in exact detail but are neither spherical nor homogeneous, we shall assume that the surfaces of constant density are a family of similar concentric ellipsoids.[4] This assumption seems reasonable both for nuclei and for elliptical galaxies, although some reservations about its accuracy will be mentioned in Sec. V. (It also applies moderately well to the mass distribution of the earth, a slightly flattened ellipsoid with a sharp surface but nonuniform density, although the high-density contours near the earth's center are thought to be more nearly spherical than the low-density contours near the surface.[5])

The Coulomb or gravitational energy of a distribution that satisfies this assumption, but is otherwise arbitrary, will be shown in Sec. II to reduce exactly to a product of two factors. One of these, to be called the anisotropy factor, depends only on the eccentricities of the ellipsoids and has a simple geometrical significance; the other depends only on the variation of density from one surface to another and has a simple physical significance.

Since the energy is not changed by permuting the three axes, it seems appropriate to replace the eccentricities by two parameters that are elementary symmetric functions of the axis lengths. In Sec. IV the anisotropy factor, which is essentially an incomplete elliptic integral of the first kind, is tabulated as a function of two such symmetric parameters. As in the unified model of the nucleus, one parameter is a measure of the departure from spherical symmetry while the other specifies the shape of the deformation (prolate, oblate, or triaxial). The anisotropy factor is insensitive to the shape parameter for small and even for moderately large deformations.

Section V contains a discussion of two examples, the Coulomb energy of a Gd^{160} nucleus and the gravitational energy of the Large Magellanic Cloud. An Appendix gives the recursion relations that were used to calculate the incomplete elliptic integrals of the first kind.

[4] A series expansion of the potential of such a distribution will be given in a later paper.

[5] K. Jung, *Handbuch der Physik* (Springer-Verlag, Berlin, Germany, 1956), Vol. XLVII, p. 604.

II. FACTORIZATION OF THE ENERGY

In a family of similar concentric ellipsoids, each member corresponds to a constant value of

$$R^2=(x^2/a^2)+(y^2/b^2)+(z^2/c^2), \tag{2.1}$$

where x, y, and z are Cartesian coordinates and the corresponding semiaxes have lengths aR, bR, and cR. It is convenient to regard R as a length and to let a, b, and c be dimensionless numbers that are normalized according to

$$abc=1 \tag{2.2}$$

and are common to all ellipsoids in the family. (Similar ellipsoids, unlike confocal ellipsoids, are simple enlargements of one another.) When a, b, and c are separately equal to unity, the ellipsoids degenerate to spheres and R reduces to the radius r.

We shall want to compare a spherically symmetric distribution having density $\rho(r)$ with a corresponding ellipsoidal distribution having the same total charge but density $\rho(R)$, the functional dependence of ρ on its argument being exactly the same in the two cases. The ellipsoidal distribution can be obtained from the spherical one by a volume-preserving deformation that moves each charge element from its original location P to a final location P':

$$P=(x,y,z)\rightarrow P'=(ax,by,cz). \tag{2.3}$$

Each surface of constant density is deformed from a sphere into an ellipsoid of the same volume, the initial value of r and the final value of R being numerically equal.

The Coulomb energy of the ellipsoidal distribution is given (in rationalized mks units) by

$$U(a,b,c)=\frac{1}{2}\frac{1}{4\pi\epsilon_0}\int\int d\mathbf{r}d\mathbf{r}'\frac{\rho(R)\rho(R')}{|\mathbf{r}-\mathbf{r}'|}. \tag{2.4}$$

In the gravitational case $(4\pi\epsilon_0)^{-1}$ is replaced by the negative of the gravitational constant G. The energy $U(1,1,1)$ of the corresponding spherical distribution differs from (2.4) only by the occurrence of r and r' in place of R and R'.

The result to be proved in this section is that the right-hand side of Eq. (2.4) can be factored:

$$U(a,b,c)=\langle r_1^2\rangle U(1,1,1). \tag{2.5}$$

The second factor is the energy of the spherical distribution, which depends only on the functional form of ρ. The first factor is an anisotropy factor that depends only on a, b, and c and has a simple geometrical meaning. The deformation (2.3) changes the unit sphere into a "unit ellipsoid" that has the same volume. The radius vector from its center to a point on its surface is $\mathbf{r}_1=(x_1,y_1,z_1)$, where the components are dimensionless numbers satisfying

$$(x_1^2/a^2)+(y_1^2/b^2)+(z_1^2/c^2)=1. \tag{2.6}$$

In spherical polar coordinates, the radius vector has components (r_1,θ,φ) and its square is found from Eq. (2.6) to be

$$r_1^2=\frac{1}{S(\theta,\varphi)}=\left(\frac{\sin^2\theta\cos^2\varphi}{a^2}+\frac{\sin^2\theta\sin^2\varphi}{b^2}+\frac{\cos^2\theta}{c^2}\right)^{-1}. \quad (2.7)$$

The anisotropy factor is the mean square radius of the unit ellipsoid, where the mean is an average over angles:

$$\langle r_1^2\rangle=\int r_1^2\frac{d\Omega}{4\pi}=\frac{1}{4\pi}\int_0^{2\pi}\int_0^{\pi}\frac{\sin\theta d\theta d\varphi}{S(\theta,\varphi)}. \quad (2.8)$$

We shall give two proofs of Eq. (2.5). The first is formal and self-contained; the second is more physical but makes use of a theorem from potential theory. The formal proof begins by inserting a Fourier integral for the reciprocal distance in Eq. (2.4):

$$U(a,b,c)=\frac{1}{2\epsilon_0}\frac{1}{(2\pi)^3}\int d\mathbf{r}d\mathbf{r}'d\mathbf{k}\,\rho(R)\frac{e^{i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')}}{k^2}\rho(R'). \quad (2.9)$$

If we define $\mathbf{R}=(x/a,y/b,z/c)$ and $\mathbf{K}=(ak_x,bk_y,ck_z)$, then the Fourier transform of $\rho(R)$ can be written in the form

$$\begin{aligned}\int d\mathbf{r}e^{i\mathbf{k}\cdot\mathbf{r}}\rho(R)&=\int d\mathbf{R}e^{i\mathbf{K}\cdot\mathbf{R}}\rho(R)\\&=\frac{4\pi}{K}\int_0^{\infty}\rho(R)R\sin(KR)dR\\&=4\pi\epsilon_0\int_0^{\infty}\tau(R)\cos(KR)dR, \quad (2.10)\end{aligned}$$

where

$$\tau(R)=\frac{1}{\epsilon_0}\int_R^{\infty}\rho(R')R'dR'. \quad (2.11)$$

The last step in Eq. (2.10) follows from an integration by parts in which the boundary terms vanish.[6] Equation (2.9) now becomes

$$U(a,b,c)=\frac{\epsilon_0}{\pi}\int\frac{d\mathbf{K}}{k^2}\left(\int_0^{\infty}\tau(R)\cos(KR)dR\right)^2. \quad (2.12)$$

If we choose spherical polar coordinates (K,θ,φ) in K space, then $k^2=K^2S(\theta,\varphi)$, where $S(\theta,\varphi)$ is defined by Eq. (2.7). The integral over angles is just the one in Eq. (2.8):

$$\begin{aligned}U(a,b,c)&=4\epsilon_0\langle r_1^2\rangle\int_0^{\infty}dK\left(\int_0^{\infty}\tau(R)\cos(KR)dR\right)^2\\&=\langle r_1^2\rangle 2\pi\epsilon_0\int_0^{\infty}[\tau(R)]^2dR, \quad (2.13)\end{aligned}$$

where Parseval's formula has been used in the last step. For the particular case of spherical symmetry, Eq. (2.13) reduces to

$$U(1,1,1)=2\pi\epsilon_0\int_0^{\infty}[\tau(r)]^2dr, \quad (2.14)$$

an expression for the energy of a spherical distribution that has been derived more directly elsewhere.[7] Since the variable of integration is immaterial, Eq. (2.14) can be substituted back into Eq. (2.13) to obtain Eq. (2.5).

The second proof starts from Newton's theorem: if a thin shell between two similar concentric ellipsoids is filled with a uniform charge density, the potential is constant in the region enclosed by the shell.[8] In order to evaluate the constant, it suffices to find the potential at the center. An element of solid angle $d\Omega$ about the center cuts from the shell a small cylinder with slant height dr and projected base area $r^2d\Omega$, where r is determined as a function of angle by $r^2=R^2/S(\theta,\varphi)$ in accordance with Eqs. (2.1) and (2.7). The charge contained in this cylinder contributes to the potential at the center an amount

$$\rho r^2drd\Omega/4\pi\epsilon_0 r=\rho RdRd\Omega/4\pi\epsilon_0 S(\theta,\varphi).$$

Integration over angles shows that the total potential at the center is

$$\langle r_1^2\rangle\rho RdR/\epsilon_0.$$

Hence, when a spherical shell is subjected to the deformation (2.3), the constant potential inside the shell is changed only by a factor $\langle r_1^2\rangle$.

We now assemble the complete ellipsoidal distribution by bringing up shells of successively smaller size. When all shells larger than the one labeled by R have been assembled, the potential in the empty space enclosed by them has the constant value

$$\frac{\langle r_1^2\rangle}{\epsilon_0}\int_R^{\infty}\rho(R')R'dR'=\langle r_1^2\rangle\tau(R).$$

[We note in passing that the potential at the center of an ellipsoidal distribution is $\langle r_1^2\rangle\tau(0)$.] The work required to bring up the next shell of thickness dR and volume $4\pi R^2dR$ is

$$\rho(R)4\pi R^2dR\langle r_1^2\rangle\tau(R).$$

The total work required to assemble the distribution is

$$U(a,b,c)=\langle r_1^2\rangle 4\pi\int_0^{\infty}\tau(R)\rho(R)R^2dR. \quad (2.15)$$

As in the case of Eq. (2.13), the coefficient of $\langle r_1^2\rangle$ on the right-hand side must be the energy of a spherical

[6] Both the total charge and the energy of the spherical distribution are assumed to exist. Then $r\tau(r)$ necessarily tends to zero at the origin and also at large distances.

[7] B. C. Carlson, Iowa State J. Sci. **35**, 319 (1961).

[8] W. D. MacMillan, *The Theory of the Potential* (Dover Publications, New York, 1958), p. 10.

distribution; this particular expression for it is again obtained more directly in footnote 7.

III. ANISOTROPY FACTOR AS A FUNCTION OF THE ECCENTRICITIES

The mean square radius of the unit ellipsoid, given by Eq. (2.8), can be written as

$$\langle r_1^2\rangle=\frac{2}{\pi}\int_0^{\pi/2}\int_0^{\pi/2}\left(\frac{\sin^2\theta\cos^2\varphi}{a^2}+\frac{\sin^2\theta\sin^2\varphi}{b^2}+\frac{\cos^2\theta}{c^2}\right)^{-1}\times\sin\theta d\theta d\varphi, \quad (3.1)$$

where the integration now extends over one octant of the ellipsoid. By taking $\tan\varphi$ as a new variable, the integration over φ can be carried out[9]:

$$\langle r_1^2\rangle=\int_0^{\pi/2}\left(\frac{\sin^2\theta}{a^2}+\frac{\cos^2\theta}{c^2}\right)^{-\frac{1}{2}}\left(\frac{\sin^2\theta}{b^2}+\frac{\cos^2\theta}{c^2}\right)^{-\frac{1}{2}}\times\sin\theta d\theta. \quad (3.2)$$

Because of its geometrical meaning, the integral is a symmetric function of a, b, and c; therefore, no generality is lost by assuming

$$a\leq b\leq c. \quad (3.3)$$

We shall retain this assumption throughout Sec. III and define the eccentricities

$$\epsilon=(1-a^2/c^2)^{\frac{1}{2}},\qquad \epsilon'=(1-b^2/c^2)^{\frac{1}{2}}\leq\epsilon. \quad (3.4)$$

The inverse relations are found from Eq. (2.2) to be

$$a=(1-\epsilon^2)^{1/3}(1-\epsilon'^2)^{-1/6},\quad b=(1-\epsilon^2)^{-1/6}(1-\epsilon'^2)^{1/3},\quad c=(1-\epsilon^2)^{-1/6}(1-\epsilon'^2)^{-1/6}. \quad (3.5)$$

Substitution of $\cos\theta=x$ in Eq. (3.2) gives

$$\langle r_1^2\rangle=\frac{1}{c}\int_0^1(1-\epsilon^2x^2)^{-\frac{1}{2}}(1-\epsilon'^2x^2)^{-\frac{1}{2}}dx. \quad (3.6)$$

The further substitution of $\epsilon x=\sin\psi$ puts the integral in Legendre's standard form for the incomplete elliptic integral of the first kind:

$$\langle r_1^2\rangle=\frac{1}{c\epsilon}\int_0^{\varphi}\frac{d\psi}{(1-k^2\sin^2\psi)^{\frac{1}{2}}}=\frac{1}{c\epsilon}F(\varphi,k), \quad (3.7)$$

where

$$\sin\varphi=\epsilon,\qquad k=\epsilon'/\epsilon. \quad (3.8)$$

When the unit ellipsoid has an axis of symmetry, the elliptic integral reduces to an elementary integral. In the prolate case inverse hyperbolic functions occur:

$$a=b\leq c,\quad \epsilon'=\epsilon,\quad k=1,$$
$$\langle r_1^2\rangle=\frac{1}{2c\epsilon}\ln\frac{1+\epsilon}{1-\epsilon}=\frac{1}{c\epsilon}\tanh^{-1}\epsilon=\frac{(1-\epsilon^2)^{\frac{1}{3}}}{\epsilon}\tanh^{-1}\epsilon=1-\frac{1}{45}\epsilon^4-\frac{64}{2835}\epsilon^6-\frac{58}{2835}\epsilon^8-\cdots. \quad (3.9)$$

In the oblate case, inverse circular functions occur:

$$a\leq b=c,\quad \epsilon'=0=k,$$
$$\langle r_1^2\rangle=\frac{1}{c\epsilon}\sin^{-1}\epsilon=\frac{(1-\epsilon^2)^{1/6}}{\epsilon}\sin^{-1}\epsilon=1-\frac{1}{45}\epsilon^4-\frac{62}{2835}\epsilon^6-\frac{55}{2835}\epsilon^8-\cdots. \quad (3.10)$$

The leading coefficients of the power series in the prolate and oblate cases are nearly equal. However, this does not imply that $\langle r_1^2\rangle$ is nearly independent of ϵ' for a general ellipsoid with small eccentricities; on the contrary, the right-hand side of Eq. (3.6) is formally symmetric in ϵ and ϵ'. Change of variable to $u=x^2$ gives the integral form of Appell's first hypergeometric function of two variables[10]:

$$\langle r_1^2\rangle=\frac{1}{2c}\int_0^1 u^{-\frac{1}{2}}(1-\epsilon^2u)^{-\frac{1}{2}}(1-\epsilon'^2u)^{-\frac{1}{2}}du=\frac{1}{c}F_1(\tfrac{1}{2};\tfrac{1}{2},\tfrac{1}{2};\tfrac{3}{2};\epsilon^2,\epsilon'^2)=\frac{1}{c}\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\frac{(\frac{1}{2})_{m+n}(\frac{1}{2})_m(\frac{1}{2})_n}{(\frac{3}{2})_{m+n}m!n!}\epsilon^{2m}\epsilon'^{2n}, \quad (3.11)$$

where

$$(\alpha)_m=\alpha(\alpha+1)(\alpha+2)\cdots(\alpha+m-1).$$

By making a binomial expansion of $1/c$ from Eq. (3.5) and multiplying the two series, we obtain the leading

[9] Most of the changes of variable used in Sec. III and the first paragraph of Sec. IV can be found in texts that discuss the potential of a homogeneous ellipsoid. Several well-known relations will be essentially reproduced here because of minor differences arising from our normalization of the semiaxes.

[10] W. N. Bailey, *Generalized Hypergeometric Series* (Cambridge University Press, New York, 1935), pp. 73, 77.

terms

$$\langle r_1^2\rangle = 1-\frac{1}{45}(\epsilon^4-\epsilon^2\epsilon'^2+\epsilon'^4) - \frac{2}{2835}(\epsilon^2+\epsilon'^2)(31\epsilon^4-46\epsilon^2\epsilon'^2+31\epsilon'^4)+\cdots. \quad (3.12)$$

IV. ANISOTROPY FACTOR AS A FUNCTION OF SYMMETRIC PARAMETERS

If it should be inconvenient to make the assumption (3.3), the definitions of the eccentricities in terms of a, b, and c could be made to depend on the relative sizes of these numbers. For instance $(1-\epsilon^2)^{\frac{1}{2}}$ would be defined as the ratio of the smallest to the largest. The smallest of three numbers can be regarded as a symmetric function of the three numbers; similarly, the eccentricities could then be regarded as symmetric functions of a, b, and c, but not at the same time as elementary functions in the usual sense of the word. To find a more satisfactory choice of symmetric parameters, we consider an integral representation of $\langle r_1^2\rangle$ which makes its symmetry conspicuous: substitution of $\lambda=c^2\tan^2\theta$ in Eq. (3.2) gives

$$\langle r_1^2\rangle = \tfrac{1}{2}\int_0^\infty [(\lambda+a^2)(\lambda+b^2)(\lambda+c^2)]^{-\frac{1}{2}}d\lambda. \quad (4.1)$$

We shall now introduce symmetric parameters ξ and η in the course of a proof that $\langle r_1^2\rangle$ cannot exceed unity. The expression in square brackets can be rewritten as

$$(\lambda+a^2)(\lambda+b^2)(\lambda+c^2)=(\lambda+1)^3+\xi\lambda^2+\eta\lambda, \quad (4.2)$$

where

$$\xi=a^2+b^2+c^2-3, \qquad \eta=a^{-2}+b^{-2}+c^{-2}-3. \quad (4.3)$$

The quantities $\xi+3$ and $\eta+3$ are both of the form $u+v+w$, where u, v, and w are positive and their product is $uvw=1$ by Eq. (2.2). Under these conditions, we have

$$(u+v+w)^2=\sum_{u,v,w}\left(u^2+\frac{2}{u}\right)\geq 9,$$

where the equality holds only when u, v, and w are all unity. This result implies

$$\xi\geq 0, \quad \eta\geq 0,$$

and

$$\langle r_1^2\rangle \leq \tfrac{1}{2}\int_0^\infty (\lambda+1)^{-\frac{3}{2}}d\lambda = 1, \quad (4.4)$$

where the equalities hold only when the unit ellipsoid is a sphere. Any nonspherical ellipsoidal distribution has a lower self-energy than the corresponding spherical distribution. We note in passing that Eqs. (4.4) and (3.7) give an upper bound for the incomplete elliptic integral of the first kind:

$$F(\varphi,k)\leq \frac{\sin\varphi}{(\cos\varphi)^{1/3}(1-k^2\sin^2\varphi)^{1/6}} \quad \left(\varphi<\frac{\pi}{2}\right). \quad (4.5)$$

The inequality is a close one for small φ.

If ξ and η are sufficiently small, we can substitute Eq. (4.2) in Eq. (4.1) and make a binomial expansion of the integrand:

$$\langle r_1^2\rangle = \tfrac{1}{2}\int_0^\infty [(\lambda+1)^3+\xi\lambda^2+\eta\lambda]^{-\frac{1}{2}}d\lambda = \sum_{s=0}^\infty (-1)^s(2s-1)!!K_s(\xi,\eta), \quad (4.6)$$

where

$$K_s(\xi,\eta)=\frac{1}{2^{s+1}s!}\int_0^\infty \frac{(\xi\lambda+\eta)^s\lambda^s}{(\lambda+1)^{3s+\frac{1}{2}}}d\lambda \quad (4.7)$$

is a homogeneous polynomial in ξ and η of degree s. The polynomial can be written[11] as a terminating hypergeometric series:

$$K_s(\xi,\eta)=\frac{(2s-1)!!}{(4s+1)!!}\eta^s\,{}_2F_1\left(-s,s+1;2s+\tfrac{3}{2};1-\frac{\xi}{\eta}\right) \quad (4.8)$$

$$=\frac{(4s-1)!!}{(6s+1)!!}\eta^s\,{}_2F_1\left(-s,s+1;\tfrac{1}{2}-2s;\frac{\xi}{\eta}\right). \quad (4.9)$$

Equation (4.8) is useful in relation to the special case $\xi=\eta$ because the ${}_2F_1$ function then reduces to unity. From Eqs. (4.3) and (2.2) one finds that the condition $\xi=\eta$ means that one semiaxis has unit length while the other two are reciprocals. In the notation of Sec. III we have

$$b=1, \quad c=1/a, \qquad \epsilon^2=\epsilon'^2(2-\epsilon'^2), \qquad 1/k^2=1+\cos\varphi, \qquad \eta=\xi=a^2+c^2-2=\cos\varphi+\sec\varphi-2. \quad (4.10)$$

The mean square radius in this case turns out to be a generalized hypergeometric series:

$$\langle r_1^2\rangle=\sum_{s=0}^\infty (-1)^s\frac{[(2s-1)!!]^2}{(4s+1)!!}\xi^s = 1-\frac{1}{3\cdot 5}\xi+\frac{1\cdot 3}{5\cdot 7\cdot 9}\xi^2-\frac{1\cdot 3\cdot 5}{7\cdot 9\cdot 11\cdot 13}\xi^3+\cdots = {}_3F_2\left(\frac{1}{2},\frac{1}{2},1;\frac{3}{4},\frac{5}{4};-\frac{\xi}{4}\right). \quad (4.11)$$

[11] A. Erdélyi, W. Magnus, F. Oberhettinger, and F. G. Tricomi, *Higher Transcendental Functions* (McGraw-Hill Book Company, New York, 1953), Vol. I, pp. 60, 108.

In the general case, Eq. (4.9) is more convenient for determining $\langle r_1^2\rangle$ as a power series in ξ and η:

$$\langle r_1^2\rangle = 1-\frac{1}{105}(4\xi+3\eta)+\frac{1}{15\,015}\times(48\xi^2+60\xi\eta+35\eta^2)+\cdots. \quad (4.12)$$

Although the parameters ξ and η have the advantage of being symmetric polynomials in the three semiaxis lengths, they have some drawbacks, too. One of the foremost is their awkwardness for describing ellipsoids of revolution. We shall turn now to a pair of parameters used by Hill and Wheeler[12] in the collective model of the nucleus:

$$a=\exp\left[\sigma\cos\left(\gamma-\frac{2\pi}{3}\right)\right],\quad \frac{b}{c}=\exp\left[\sqrt{3}\sigma\sin\left(\gamma-\frac{2\pi}{3}\right)\right],$$
$$b=\exp\left[\sigma\cos\left(\gamma+\frac{2\pi}{3}\right)\right],\quad \frac{c}{a}=\exp\left[\sqrt{3}\sigma\sin\left(\gamma+\frac{2\pi}{3}\right)\right],$$
$$c=\exp(\sigma\cos\gamma),\qquad \frac{a}{b}=\exp(\sqrt{3}\sigma\sin\gamma). \quad (4.13)$$

There is a one-to-one correspondence between the ordered triples of positive numbers (a,b,c) satisfying $abc=1$ and the points of the $\sigma\gamma$ plane $(0\leq\sigma,\ 0\leq\gamma<2\pi)$. The inverse relations are

$$\begin{aligned}\sigma^2&=\tfrac{2}{3}[(\ln a)^2+(\ln b)^2+(\ln c)^2],\\ \cos\gamma&=(1/\sigma)\ln c,\\ \sin\gamma&=(1/\sqrt{3}\sigma)\ln(a/b),\\ \cos3\gamma&=(4/\sigma^3)(\ln a)(\ln b)(\ln c).\end{aligned} \quad (4.14)$$

Permutations of a, b, and c carry the point (σ,γ) into the six points $(\sigma,\pm\gamma)$, $(\sigma,120°\pm\gamma)$, and $(\sigma,240°\pm\gamma)$ without changing the values of the symmetric functions σ and $\cos3\gamma$. The prolate ellipsoids of revolution are described by $\cos3\gamma=1$, the oblate ones by $\cos3\gamma=-1$, and the unit ellipsoids with one semiaxis of unit length by $\cos3\gamma=0$.

A nearly spherical ellipsoid has a radius $r(\theta,\varphi)$ that can be written, to first order in σ, as a constant plus spherical harmonics of degree two. More briefly, a small ellipsoidal deformation is a quadrupole deformation. To establish the connection in this case between the notation of Eqs. (4.13) for ellipsoids and the now widespread notation introduced by Bohr[13] to describe nuclear quadrupole deformations, it suffices to compare the expressions for the semiaxis lengths; one finds that Bohr's shape parameter γ is the same as the one used here and that his measure β of deformation magnitude is a constant times σ. To second order in σ, however, these statements are no longer true. By Eq. (2.7), the quadrupole part of

$$r_1(\theta,\varphi)=[S(\theta,\varphi)]^{-\frac{1}{2}}=\sum_{l,m}a_{lm}Y_l^m(\theta,\varphi) \quad (4.15)$$

is given to second order in σ by

$$\begin{aligned}a_{00}&=(4\pi)^{\frac{1}{2}}(1-\tfrac{1}{5}\sigma^2),\\ a_{20}&=(4\pi/5)^{\frac{1}{2}}[\sigma\cos\gamma-(1/14)\sigma^2\cos2\gamma],\\ a_{2-1}&=a_{21}=0,\\ a_{2-2}&=a_{22}=(2\pi/5)^{\frac{1}{2}}[\sigma\sin\gamma+(1/14)\sigma^2\sin2\gamma].\end{aligned} \quad (4.16)$$

To this order, the two notations are related by

$$\begin{aligned}\beta^2&=(a_{20})^2+2(a_{22})^2=\frac{4\pi}{5}\sigma^2\left(1-\frac{1}{7}\sigma\cos3\gamma\right),\\ \tan\gamma_{\rm Bohr}&=\frac{\sqrt{2}a_{22}}{a_{20}}=\tan\gamma\left(1+\frac{\sigma\sin3\gamma}{7\sin2\gamma}\right),\\ \cos3\gamma_{\rm Bohr}&=\cos3\gamma-\frac{3}{14}\sigma(\sin3\gamma)^2.\end{aligned} \quad (4.17)$$

The deformations found in nuclei are sufficiently small that the second term of

$$\sigma=0.631\beta(1+0.045\beta\cos3\gamma+\cdots) \quad (4.18)$$

may ordinarily be neglected.

It is clear from Eqs. (4.13) that any function of a, b, and c is a periodic function of γ with period 2π and, if square integrable, can be represented by a Fourier series in γ; furthermore, if the function is symmetric in a, b, and c, Eqs. (4.14) show that its Fourier series can contain only terms of the type $\cos3n\gamma$. Since increasing γ by π is formally equivalent to changing the sign of σ, the coefficient of $\cos3n\gamma$ must be an even function of σ if n is even and an odd function if n is odd. Finally, we note that a, b, and c are expressible in terms of $\sigma\cos\gamma$ and $\sigma\sin\gamma$; any function of these variables that is proportional to an integral power m of σ cannot contain Fourier harmonics of order higher than m. Therefore, if the coefficient of $\cos3n\gamma$ can be represented by a power series in σ, the lowest power that occurs will be σ^{3n}. These arguments show that a symmetric function such as $\langle r_1^2\rangle$, ξ, or η can be developed in a series of the form

$$\begin{aligned}&\alpha_0+\beta_0\sigma^2+\gamma_0\sigma^4+\delta_0\sigma^6+\cdots\\ &\quad+(\alpha_3\sigma^3+\beta_3\sigma^5+\cdots)\cos3\gamma\\ &\quad+(\alpha_6\sigma^6+\cdots)\cos6\gamma\\ &\quad+\cdots.\end{aligned} \quad (4.19)$$

Herein lies a principal advantage of the present choice of parameters; $\langle r_1^2\rangle$ is independent of γ to order σ^2, is a linear function of $\cos3\gamma$ to order σ^5, and is a quadratic function of $\cos3\gamma$ to order σ^8.

[12] D. L. Hill and J. A. Wheeler, Phys. Rev. **89**, 1102 (1953), Fig. 13. The condition $a\leq b\leq c$ of Sec. III corresponds to $-60°\leq\gamma\leq0°$. It would be natural in the present context to reverse the sign of γ in Eqs. (4.13) and thereby to interchange the definitions of a and b, but we shall retain the choice of sign that has become conventional in the unified model of the nucleus.

[13] A. Bohr, Kgl. Danske Videnskab. Selskab, Mat.-fys. Medd. **26**, No. 14 (1952).

In order to find the coefficients in the series (4.19) for $\langle r_1^2\rangle$, we first consider the series

$$\xi+3=a^2+b^2+c^2=\sum_{n=0}^{\infty} g_n(\sigma)\cos 3n\gamma,$$
$$g_n(\sigma)=(2-\delta_{n0})\frac{1}{2\pi}\int_{-\pi}^{\pi}(a^2+b^2+c^2)\cos 3n\gamma d\gamma. \tag{4.20}$$

The three terms in parentheses make equal contributions to the integral, for they differ from one another only by a phase angle in γ of $2\pi/3$, which is a period of $\cos 3n\gamma$. Replacing the terms in parentheses by $3c^2$, we obtain[14]

$$g_n(\sigma)=(2-\delta_{n0})\frac{3}{\pi}\int_0^{\pi} e^{2\sigma\cos\gamma}\cos 3n\gamma d\gamma$$
$$=(2-\delta_{n0})3I_{3n}(2\sigma),$$

where

$$I_{3n}(2\sigma)=\sum_{m=0}^{\infty}\frac{\sigma^{3n+2m}}{m!(3n+m)!} \tag{4.21}$$

is a Bessel function of imaginary argument. The Fourier expansion of ξ is, therefore,

$$\xi=3I_0(2\sigma)-3+6\sum_{n=1}^{\infty} I_{3n}(2\sigma)\cos 3n\gamma \tag{4.22}$$
$$=3\sigma^2+\sigma^3\cos 3\gamma+\tfrac{3}{4}\sigma^4+(\sigma^5/4)\cos 3\gamma$$
$$+(\sigma^6/120)(10+\cos 6\gamma)+\cdots. \tag{4.23}$$

Since η differs from ξ only by a phase angle of π in γ, we have

$$\eta=3I_0(2\sigma)-3+6\sum_{n=1}^{\infty}(-1)^n I_{3n}(2\sigma)\cos 3n\gamma. \tag{4.24}$$

Finally we substitute these expansions in Eq. (4.12) and obtain, to terms of order σ^6,

$$\langle r_1^2\rangle=1-\frac{1}{5}\sigma^2-\frac{1}{105}\sigma^3\cos 3\gamma+\frac{1}{28}\sigma^4+\frac{13}{4620}\sigma^5\cos 3\gamma$$
$$-\frac{203}{28\,600}\sigma^6+\frac{379}{900\,900}\sigma^6(\cos 3\gamma)^2+\cdots. \tag{4.25}$$

Inspection of the coefficients shows that $\langle r_1^2\rangle$ should be very nearly a linear function of $\cos 3\gamma$ for $\sigma\lesssim 1$. This expectation is borne out by Table I, which was com-

TABLE I. The mean square radius $\langle r_1^2\rangle$ of the ellipsoid $(x_1/a)^2+(y_1/b)^2+(z_1/c)^2=1$, where $abc=1$. The mean is an average over angles, defined by Eq. (2.8). The parameters σ and $\cos 3\gamma$, defined by Eqs. (4.14), are symmetric functions of a, b, and c; the magnitude of the deformation from sphericity is measured by σ and the shape of the deformation by $\cos 3\gamma$. The value of $\cos 3\gamma$ is 1 for a prolate ellipsoid of revolution, -1 for an oblate one, and 0 for a triaxial ellipsoid with one semiaxis of unit length. The ratio of axis lengths for an ellipsoid of revolution is $\exp(3\sigma/2)$. The relation of $\langle r_1^2\rangle$ to the incomplete elliptic integral of the first kind is given by Eq. (3.7). The ratio of the Coulomb energy of an ellipsoidal charge distribution to the energy of the corresponding spherical distribution is equal to $\langle r_1^2\rangle$.

σ \ $\cos 3\gamma$	−1.0	−0.8	−0.6	−0.4	−0.2	0.0	0.2	0.4	0.6	0.8	1.0
0	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000
0.05	0.99950	0.99950	0.99950	0.99950	0.99950	0.99950	0.99950	0.99950	0.99950	0.99950	0.99950
0.10	0.99801	0.99801	0.99801	0.99801	0.99801	0.99800	0.99800	0.99800	0.99800	0.99800	0.99799
0.15	0.99555	0.99554	0.99554	0.99553	0.99552	0.99552	0.99551	0.99551	0.99550	0.99549	0.99549
0.20	0.99213	0.99212	0.99210	0.99209	0.99207	0.99206	0.99204	0.99203	0.99201	0.99200	0.99198
0.25	0.98778	0.98775	0.98773	0.98770	0.98767	0.98764	0.98761	0.98758	0.98755	0.98752	0.98749
0.30	0.98254	0.98248	0.98243	0.98238	0.98233	0.98228	0.98223	0.98218	0.98213	0.98208	0.98203
0.35	0.97642	0.97634	0.97626	0.97618	0.97610	0.97602	0.97594	0.97587	0.97579	0.97571	0.97563
0.40	0.96947	0.96935	0.96924	0.96912	0.96900	0.96889	0.96877	0.96865	0.96854	0.96842	0.96831
0.45	0.96173	0.96156	0.96140	0.96124	0.96107	0.96091	0.96074	0.96058	0.96042	0.96026	0.96009
0.50	0.95324	0.95302	0.95279	0.95257	0.95235	0.95213	0.95191	0.95169	0.95146	0.95124	0.95103
0.6	0.93419	0.93382	0.93344	0.93307	0.93269	0.93232	0.93195	0.93158	0.93121	0.93085	0.93048
0.7	0.91270	0.91212	0.91154	0.91096	0.91039	0.90982	0.90925	0.90869	0.90813	0.90757	0.90702
0.8	0.88914	0.88830	0.88747	0.88664	0.88581	0.88500	0.88419	0.88338	0.88258	0.88179	0.88100
0.9	0.86391	0.86275	0.86160	0.86047	0.85934	0.85823	0.85713	0.85603	0.85495	0.85389	0.85283
1.0	0.83736	0.83583	0.83431	0.83282	0.83135	0.82989	0.82845	0.82703	0.82564	0.82425	0.82289
1.2	0.78169	0.77924	0.77684	0.77448	0.77218	0.76991	0.76770	0.76553	0.76341	0.76132	0.75929
1.4	0.72452	0.72096	0.71749	0.71412	0.71084	0.70766	0.70457	0.70156	0.69864	0.69581	0.69307
1.6	0.66774	0.66290	0.65824	0.65374	0.64941	0.64524	0.64123	0.63737	0.63366	0.63009	0.62666
1.8	0.61268	0.60646	0.60052	0.59485	0.58945	0.58429	0.57937	0.57469	0.57023	0.56599	0.56197
2.0	0.56024	0.55258	0.54535	0.53852	0.53206	0.52597	0.52021	0.51478	0.50967	0.50486	0.50035
2.5	0.44343	0.43228	0.42205	0.41265	0.40399	0.39601	0.38866	0.38189	0.37567	0.36998	0.36481
3.0	0.34804	0.33400	0.32155	0.31043	0.30045	0.29147	0.28338	0.27609	0.26955	0.26371	0.25857
3.5	0.27206	0.25605	0.24236	0.23050	0.22014	0.21102	0.20297	0.19585	0.18960	0.18414	0.17947
4.0	0.21225	0.19521	0.18122	0.16949	0.15949	0.15089	0.14343	0.13696	0.13136	0.12657	0.12259
4.5	0.16544	0.14820	0.13467	0.12368	0.11455	0.10686	0.10031	0.09471	0.08994	0.08593	0.08269
5.0	0.12889	0.11212	0.09956	0.08968	0.08169	0.07509	0.06957	0.06492	0.06100	0.05776	0.05520

[14] G. N. Watson, *A Treatise on the Theory of Bessel Functions* (Cambridge University Press, New York, 1948), p. 181.

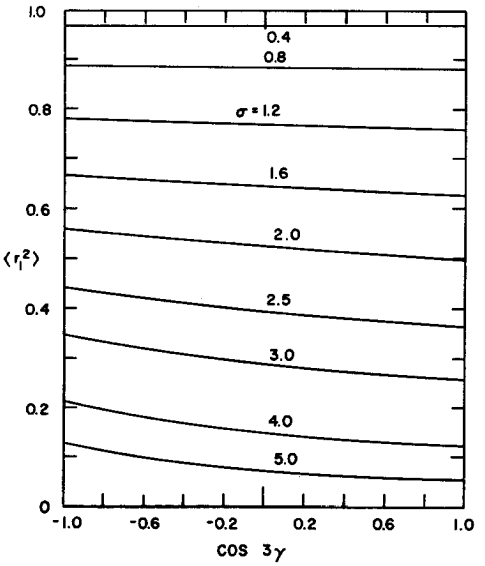

FIG. 1. The anisotropy factor $\langle r_1^2 \rangle$ as a function of the shape parameter for fixed values of the deformation magnitude. The symbols are explained in the caption of Table I.

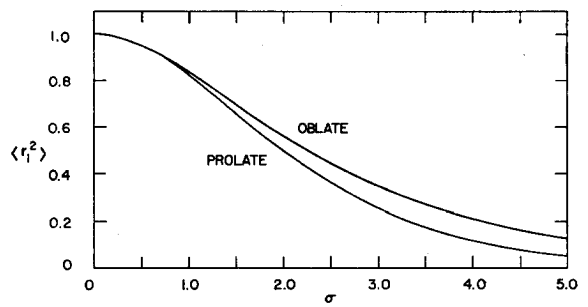

FIG. 2. The anisotropy factor $\langle r_1^2 \rangle$ as a function of the deformation magnitude for ellipsoids of revolution. The symbols are explained in the caption of Table I.

puted by use of Eqs. (3.7), (3.8), (3.4), and (4.13). Even for σ as large as 2, the value of $\langle r_1^2 \rangle$ at $\cos 3\gamma = 0$ differs by less than 1% from the average of its values at $\cos 3\gamma = \pm 1$. This nearness to linearity, which is shown graphically in Fig. 1, is impressive when one realizes that $\sigma = 2$ corresponds to elliptic integrals with argument $\varphi > 87°$ and to prolate or oblate ellipsoids of revolution with one axis 20 times longer than another.

The dependence of $\langle r_1^2 \rangle$ on σ is illustrated in Fig. 2; curves for other values of γ would lie between the prolate and oblate curves. The behavior at large σ can be obtained from Eq. (3.9) in the prolate case:

$$\cos 3\gamma = 1, \quad e^{-\sigma/2} = a = b < c = e^{\sigma},$$
$$\langle r_1^2 \rangle = e^{-\sigma}(1 - e^{-3\sigma})^{-\frac{1}{2}} \cosh^{-1} e^{3\sigma/2}$$
$$= e^{-\sigma}(\tfrac{3}{2}\sigma + \ln 2) + e^{-4\sigma}\tfrac{1}{2}(\tfrac{3}{2}\sigma + \ln 2 - \tfrac{1}{2}) + \cdots. \quad (4.27)$$

From Eq. (3.10), the corresponding relations for the oblate case are

$$\cos 3\gamma = -1, \quad e^{-\sigma} = a < b = c = e^{\sigma/2},$$
$$\langle r_1^2 \rangle = e^{-\sigma/2}(1 - e^{-3\sigma})^{-\frac{1}{2}} \cos^{-1} e^{-3\sigma/2}$$
$$= (\pi/2)e^{-\sigma/2} - e^{-2\sigma} + (\pi/4)e^{-7\sigma/2} - \cdots. \quad (4.28)$$

One would like to have a simple formula giving a good approximation to $\langle r_1^2 \rangle$ for all values of $\cos 3\gamma$ when σ is large; Eq. (4.30) below falls just a little short of this mark. From an expansion of $F(\varphi, k)$ in the neighborhood of its singularity,[15] we find

$$F(\varphi, k) \approx \ln 4 - \ln[\cos\varphi + (1 - k^2 \sin^2\varphi)^{\frac{1}{2}}], \quad (\cos\varphi \ll 1,\ 1 - k \ll 1). \quad (4.29)$$

As σ tends to infinity, $\cos\varphi$ tends to zero uniformly in γ, but k tends to unity (nonuniformly in γ) for $\cos 3\gamma \neq -1$ while remaining always zero in the oblate case. Excluding the oblate case, we substitute Eq. (4.29) in Eq. (3.7) and obtain

$$\langle r_1^2 \rangle \approx \frac{1}{c} \ln \frac{4c}{a+b}$$

$$\langle r_1^2 \rangle \approx \exp(-\sigma\cos\gamma)\left\{\ln 4 + \sqrt{3}\sigma \sin\left(\frac{\pi}{3} - \gamma\right) - \ln[1 + \exp(-\sqrt{3}\sigma \sin\gamma)]\right\}, (0 \leq \gamma < \pi/3). \quad (4.30)$$

For $\sigma = 2$, the error of this approximation increases from 0.2% to 4% as $\cos 3\gamma$ decreases from 0.8 to -0.8; for $\sigma = 5$ the corresponding errors are 0.02% and 0.5%.

V. TWO EXAMPLES

The deformations found in atomic nuclei ($\sigma \lesssim 0.3$) are small enough so that values of Coulomb deformation energy obtained from Eq. (2.5) do not differ significantly from those obtained from the assumption of a uniform charge distribution with quadrupole deformation. The point here is not the value but the assumption on which it is based; we have shown that the results are valid for a nucleus with a diffuse surface if the charge distribution is ellipsoidal in the sense of Sec. I.

Perhaps as reasonable an assumption as one can make at present about the charge distribution in a heavy nucleus is that it has a deformed Fermi shape:

$$\rho(R) = \rho_0\left[1 + \exp\frac{R - 1.07A^{\frac{1}{3}}f}{2.40f/4.39}\right]^{-1}, \quad (5.1)$$

where R is defined by Eq. (2.1) and ρ_0 is determined by the total charge. We shall now use this model to calculate the Coulomb deformation energy of Gd^{160}. The corresponding spherical Fermi distribution, with R replaced by r, has a Coulomb energy and a mean square

[15] E. L. Kaplan, J. Math. and Phys. 27, 11 (1948).

radius given by[16]

$$U=560 \text{ Mev},$$
$$\langle r^2\rangle=24.4f^2. \tag{5.2}$$

The amount of deformation is determined by the intrinsic quadrupole moment, found experimentally from the Coulomb excitation cross section to have magnitude[17]

$$|Q_0|=10\times10^{-24} \text{ cm}^2; \tag{5.3}$$

the sign of Q_0 is positive if the nucleus is a prolate ellipsoid of revolution. The relation between Q_0 and σ is easily found from the equation

$$\int(2z^2-x^2-y^2)\rho(R)d\mathbf{r} = (2c^2-a^2-b^2)\tfrac{1}{5}\int\rho(R)R^2d\mathbf{R}. \tag{5.4}$$

The integral on the right-hand side is the total charge times the mean square radius of the corresponding spherical distribution. For a prolate ellipsoid of revolution, we thus have

$$\begin{aligned}Q_0&=\tfrac{2}{5}(e^{2\sigma}-e^{-\sigma})Z\langle r^2\rangle_{\text{spher}}\\&=Z\langle r^2\rangle_{\text{spher}}2\sigma(1+\tfrac{1}{2}\sigma+\cdots).\end{aligned} \tag{5.5}$$

In the oblate case the sign of σ is reversed in Eq. (5.5). Substitution of Eqs. (5.2) and (5.3) leads to

$$\sigma=0.27,$$
$$\langle r_1^2\rangle=0.985, \tag{5.6}$$

where the latter value is obtained from Eq. (4.25). Thus, the energy of the deformed distribution (5.1) is lower than 560 Mev by 1.5% or 8.3 Mev. Since the quantum-mechanical exchange corrections amount to approximately 5% of the classical Coulomb energy, according to the statistical estimate, a better figure for the Coulomb deformation energy would be 7.9 Mev.

Two remarks may be made in support of the assumption (5.1). First, an ellipsoidal distribution satisfies the reasonable condition[18] that the dependence of density on radius should be the same in all directions aside from an angle-dependent radial scale factor, for this condition is equivalent to requiring that the surfaces of constant density be simple enlargements of one another. Second, a widely used nuclear model is that of nucleons moving in an anisotropic harmonic oscillator potential, which may be written in the form $V=\frac{1}{2}sR^2$; the ellipsoidal surfaces of constant nuclear potential are, according to the Thomas-Fermi statistical approximation, also surfaces of constant density.

Assumptions about the mass distribution in stellar galaxies must be even more tentative. For elliptical galaxies, the observed luminosity contours are indeed approximately elliptical, although their eccentricity tends to increase and then go through a broad maximum as one proceeds outwards from the center.[19] Furthermore, there are marked departures from elliptical shape in galaxies with a high degree of flattening. However, the ratio of mass to luminosity may not be constant throughout a galaxy.

A more direct source of information about mass distribution is the measurement of rotational velocity as a function of distance from the galactic center. If the motion is assumed to be circular, the gravitational force can then be deduced as a function of distance. A detailed analysis of this kind has been carried out for radio observations of the Large Magellanic Cloud by Kerr and de Vaucouleurs.[20] Although this is not an elliptical galaxy, they have fitted the data with an oblate ellipsoidal distribution of the form

$$\rho=\begin{cases}\rho_0[1-(R^2/A^2)] & R\leq A,\\ 0 & A\leq R,\end{cases} \tag{5.7}$$
$$a=b=5c,$$
$$A=5^{-\frac{1}{2}}\times2.8 \text{ kiloparsec}=1.64 \text{ kiloparsec},$$
$$\rho_0=0.30 \text{ suns/parsec}^3.$$

The total mass, $M=2.2\times10^9$ suns, includes their estimated correction for random motions, but not an additional correction for mass in the outer fringe of the galaxy, beyond $R=A$. From Eqs. (2.11) and (4.13) and Table I, we find

$$\sigma=1.07,\quad \gamma=\pi,$$
$$\langle r_1^2\rangle=0.82,$$
$$\tau(R)=\begin{cases}-\pi G\rho_0A^2[1-(R^2/A^2)]^2 & R\leq A\\ 0 & A\leq R.\end{cases} \tag{5.8}$$

At the center of the galaxy, the potential is $\langle r_1^2\rangle\tau(0)$, according to Sec. II, and the escape velocity is found from this to be 120 km/sec. The spherical distribution corresponding to (5.7) has, by Eq. (2.14), a gravitational energy

$$U(1,1,1)=-(5/7)(GM^2/A); \tag{5.9}$$

on dividing by the squared velocity of light, the mass equivalent of this energy turns out to be 81 suns. It should be noted that the gravitational energy of the galaxy, as defined here, does not include the gravitational self-energies of the individual stars. Since the energy of the deformed distribution is $\langle r_1^2\rangle U(1,1,1)$, the gravitational energy of deformation has a mass equivalent of 15 suns.

[16] Reference 7: Eqs. (4.11), (4.12), and (4.24).
[17] K. Alder, A. Bohr, T. Huus, B. Mottelson, and A. Winther, Revs. Modern Phys. 28, 432 (1956), Table V.2.
[18] D. L. Hill, *Handbuch der Physik* (Springer-Verlag, Berlin, Germany, 1957), Vol. XXXIX, p. 181.
[19] G. de Vaucouleurs, *Handbuch der Physik* (Springer-Verlag, Berlin, Germany, 1959), Vol. LIII, p. 322.
[20] F. J. Kerr and G. de Vaucouleurs, Australian J. Phys. 9, 90 (1956).

ACKNOWLEDGMENTS

I should like to thank Professor Joseph M. Keller for valuable suggestions, Ruben G. Babayan for efficient assistance with numerical work, and Dr. Donald R. Fitzwater for coding the calculation of Table I. The calculation was performed on an IBM 704 by courtesy of the Midwestern Universities Research Association.

APPENDIX

The incomplete elliptic integrals $F(\varphi,k)$ required for Table I have values of k that are often very close to unity. Except for the elementary case $k=0$, they were computed by recursion relations derived from King's "hyperbolic scale of increasing amplitudes."[21] The recursion process terminates with the first step when k is unity and converges rapidly unless $1-\sin\varphi \ll 1-k$.

One calculates first a succession of arithmetic and geometric means:

$$\begin{aligned} &\alpha_0=1, \quad \gamma_0=k\neq 0, \\ &\alpha_{n+1}=\tfrac{1}{2}(\alpha_n+\gamma_n), \quad \gamma_{n+1}=(\alpha_n\gamma_n)^{\frac{1}{2}}, \\ &\omega_n=\tfrac{1}{4}(\alpha_{n+1}/\alpha_{n+2})^2. \end{aligned} \tag{A.1}$$

Four interations always sufficed to determine α_∞. The main recursion relations are then

$$\begin{aligned} &H=(1+\sin\varphi)/\cos\varphi, \\ &P_0=\frac{1-k}{1+k}H^{-2}, \quad R_0=\frac{1-k}{1+k}H^2, \\ &Q_n=\frac{1+R_n}{1+P_n}, \\ &P_{n+1}=\omega_n P_n{}^2 Q_n, \quad R_{n+1}=\omega_n R_n{}^2/Q_n. \end{aligned} \tag{A.2}$$

The elliptic integral is given by

$$F(\varphi,k)=(\ln H-\sum_{n=0}^{\infty} 2^{-n-2}\ln Q_n)/\alpha_\infty. \tag{A.3}$$

The value of n at which the terms of the series became less than 10^{-8} was approximately 6 as an average but increased to 12 for $\sigma=5$, $\cos 3\gamma=-0.8$.

The connection between King's recursion relations and ours is given by

$$\begin{aligned} &\ln H=\varphi_0, \\ &\ln Q_n=4\varphi_n-2\varphi_{n+1}. \end{aligned} \tag{A.4}$$

[21] L. V. King, *On the Direct Numerical Calculation of Elliptic Functions and Integrals* (Cambridge University Press, New York, 1924).

Group-Theoretical Treatment of Time- and Energy-Dependent Multiple Scattering with Application to the Slowing-Down of Neutrons

E. GUTH AND E. INÖNÜ
Oak Ridge National Laboratory, Oak Ridge, Tennessee*
(Received August 31, 1960)

Wigner developed a group-theoretical method for those problems of multiple scattering in which the elementary scattering law is invariant under a group of transformations. The integral transforms, used in more standard treatments for the reduction of convolutions, come in here more naturally through the representations of the groups of transformations. Wigner's work is extended here to include the time-dependent slowing-down of neutrons. In this case a group of linear transformations comes in, which does not yield orthogonality. Nevertheless, it is possible to determine all positive time moments of the distribution function and from them the distribution function itself. The conditions for the existence of the group are that the total scattering cross section is proportional to v^γ (γ: any real number) and that either the ratio of absorption to scattering cross section is constant (including zero) or that the absorption cross section varies as $1/v$. Moreover, it is assumed that the angular dependence can be anisotropic, but does not depend on energy in the center of mass system. For the special case of no absorption and spherically symmetric elastic scattering in center of mass system, our solution reduces to Waller's recent exact expression. As a further generalization, we discuss the group which, with the same assumptions about the cross sections, exists for the case of time-energy-space-direction dependence. Here also, the group-theoretical method yields naturally all positive moments of the distribution functions.

I. INTRODUCTION

WIGNER has given a group-theoretical method for the exact evaluation of appropriate moments in those problems of multiple scattering in which the elementary scattering law is invariant under a group of transformations.[1] He has shown in detail how this method can be applied to the scattering of neutrons without energy change in two and three dimensions and indicated the generalization to scattering with energy change in the usual special case encountered in neutron slowing-down theory. The majority of such problems have already been solved by more standard methods usually based on integral transforms.[2] However, the new method affords a unifying point of view based on the symmetry of the problem and enables one to select immediately those cases in which an exact solution is (at least formally) possible. In addition, to quote Wigner, "the method of computation seems somewhat more transparent."

We would like to discuss here the application of this method to the time-dependent slowing-down of neutrons which was not considered by Wigner.[1] The standard conditions of slowing-down theory will be assumed; i.e., we shall consider only elastic collisions with free moderator nuclei which are at rest. After sketching the method briefly, we will first discuss as simple illustrations the purely energy-dependent and purely time-dependent problems and then treat the case of time-energy dependence, assuming that the total scattering cross section is proportional to v^γ, where γ may be any real number and that either the ratio of absorption cross section to scattering cross section is constant (including zero) or that the absorption cross section varies as $1/v$. The differential scattering cross section will not be assumed to be isotropic in the center of mass system. More generally, we shall assume that the angular dependence is independent of energy in the c.m. system. It will be shown that Wigner's method gives here directly the positive time moments of the distribution functions for arbitrary source distributions. The distribution functions can be determined through these moments. The solution will be seen to coincide with Waller's exact expression[3] in the special case of spherically symmetric elastic scattering in c.m. with no absorption. Finally, we will indicate the group which, with the same assumptions for the cross sections, exists for the general case of time-energy-space-direction dependence and outline the method of solution which yields all positive moments of the distribution functions for such problems.

A. Sketch of the Method

One starts from the integral equation which defines the problem of multiple scattering:

$$f_{n+1}(t)=\int f_n(s)P(s,t)ds, \tag{1}$$

where $f_n(s)$ is the probability distribution function after n "elementary events" per unit volume element of s, s standing for all the variables which define the state of the system, and $P(s,t)$ is the probability that an elementary event changes the state s into a unit volume element at t. The integration is over all possible states s. In general, the elementary event will consist of a collision plus the subsequent traversal of a free path (or of a free path plus a subsequent collision). The total distribution function is given by

$$f(s)=\sum_{n=0}^{\infty} f_n(s).$$

* Operated for the U. S. Atomic Energy Commission by the Union Carbide Nuclear Company.

[1] E. P. Wigner, Phys. Rev. 94, 17 (1954).

[2] Cf. e.g., C. C. Grosjean, dissertation, Columbia University, New York, New York, 1951.

[3] I. Waller, *Proceedings of the Second United Nations International Conference on the Peaceful Uses of Atomic Energy* (United Nations, New York, 1958), Vol. 16, P/153.

Equation (1) expresses the characteristic symmetry of any multiple scattering problem. Because of the identical character of each successive elementary event, $f_{n+1}(t)$ can be expressed by an $(n+1)$-fold integral, as follows from (1):

$$f_{n+1}(t)=\int\cdots\int f_0(s_0)P(s_0,s_1)P(s_1,s_2)\cdots P(s_n,t)ds_0ds_1\cdots ds_n, \quad (1a)$$

where $f_0(s_0)$ describes the given source distribution. Solution of the problem implies the reduction of this $(n+1)$-fold integral to one or at most a few integrals. Now if the problem is such that: (1) the transition probability $P(s,t)$ is invariant under a group of transformations G, and (2) every state of the system can be obtained from a fixed standard state by operations of this group G, a solution can be derived in the following way. By using the second condition, one establishes a correspondence between the states and the group elements. In the simplest case of one-to-one correspondence, we simply identify the states with the group elements. Assuming this is the case, we interpret s and t as the group elements and obtain from the first condition:

$$P(s,t)=P(e,s^{-1}t)=P(s^{-1}t), \quad (2)$$

where in the last term we have omitted to write the unit element e which corresponds to the standard state. Consequently, Eq. (1) may be written as

$$f_{n+1}(t)=\int f_n(s)P(s^{-1}t)ds, \quad (3)$$

where the integration may be interpreted as an invariant integration over the group space. We shall always consider left-invariant integration and define the volume element accordingly. Note that $f_n(s)$ and $f(s)$ are distribution functions per unit volume element of group space and, in general, differ from the usual distribution functions by some weight factors.

One considers now a representation $D^{(k)}(t)$ of the group G which is characterized by some variable (or variables) k, and defines the matrices:

$$\Phi_{\mu\nu}{}^{(n)}(k)=\int f_n(t)D_{\mu\nu}{}^{(k)}(t)dt, \quad (4)$$

$$\Pi_{\mu\nu}(k)=\int P(t)D_{\mu\nu}{}^{(k)}(t)dt, \quad (5)$$

where the integrations are over the group space. On multiplying both sides of Eq. (3) by $D_{\mu\nu}{}^{(k)}$, integrating over the group space and using the invariance property of the group integral, one obtains directly:

$$\Phi_{\mu\nu}{}^{(n+1)}(k)=\sum_{\mu'}\Phi_{\mu\mu'}{}^{(n)}(k)\Pi_{\mu'\nu}(k). \quad (6)$$

From (6) follows the solution for $\Phi^{(n)}(k)$:

$$\Phi_{\mu\nu}{}^{(n)}(k)=\sum_{\mu'}\Phi_{\mu\mu'}{}^{(0)}(k)\Pi_{\mu'\nu}{}^{n}(k), \quad (7)$$

where $\Phi^{(0)}$ is obtained from $f_0(s)$. The corresponding "transform" for the total distribution function is given formally by

$$\Phi(k)=(1-\Pi)^{-1}\Phi^{(0)}(k). \quad (8)$$

If the representations are chosen properly, the matrices $\Phi_{\mu\nu}(k)$ which are obtained in this way turn out to be nothing else than the appropriate moments of the problem. In all cases considered so far, it has also been possible to invert the relations (4) and (8) and obtain the distributions themselves. The main practical problem consists in finding a simple expression for the matrices Π^n and $1/(1-\Pi)$ it can be tackled easily provided one chooses a convenient form for the representations.

B. Simple Examples

1. Purely Energy-Dependent Slowing Down

As an illustration of the method, we consider the stationary slowing down of neutrons in an infinite homogeneous medium. On neglecting the space and direction dependence, we define the state of the system by the energy E of the neutrons. Then, assuming pure scattering, we have simply:

$$P(E'\to E)=\frac{\sigma_s(E'\to E)}{\sigma_s(E')}\text{ per unit }dE, \quad (9)$$

where $\sigma_s(E'\to E)$ is the differential cross section for a scattering from E' to within a unit interval at E, the energies being measured in the laboratory system, and $\sigma_s(E')$ is the total scattering cross section.

Now, if the probability for a fractional energy change $(E'-E)/E'$ is independent of the energies, as it is usually the case for the elastic collisions of neutrons with moderator nuclei which are free and at rest, we have for any positive real number κ,

$$P(\kappa E'\to\kappa E)\kappa=P(E'\to E). \quad (10)$$

Here the problem admits the multiplicative group of positive real numbers. The volume element in the group space is $d\kappa/\kappa$. The irreducible representations of this group are one dimensional and given simply by $D^{(s)}(\kappa)=\kappa^s$, where s may have any complex value. We obtain

$$\Pi(s)=\int_0^\infty\frac{\sigma_s(E_0\to\kappa E_0)}{\sigma_s(E_0)}\kappa^s d\kappa, \quad (11)$$

$$\Phi^{(n)}(s)=\int_0^\infty f_n(\kappa E_0)\kappa^{s-1}d\kappa, \quad (12)$$

where E_0 is a fixed energy. Thus, in this case the

appropriate moment is just the Mellin transform of the distribution. The exact solution for $f(E)$ can be formally expressed as

$$f(\kappa E_0)=\frac{1}{2\pi i}\int_{s_1-i\infty}^{s_1+i\infty}\frac{1}{1-\Pi(s)}\Phi^{(0)}(s)\kappa^{-s}ds, \quad (13)$$

where $\Phi^{(0)}(s)$ is determined from $f_0(\kappa E_0)$.

If, instead of the energy E, one uses the lethargy variable $u=\log(E_0/E)$, the appropriate group is the additive group of real numbers; its representations may be taken as e^{isu}(unitary) or e^{-su}(nonunitary) and correspondingly, one will get Fourier or Laplace transforms. For the special case of isotopic scattering in the c.m. system, all these transforms have been used and the final inversion integrals carried out in various approximations by many authors, among others by Waller[4] (who used Fourier transforms), by Davison[5] (who used Mellin transforms) and by Adler,[6] Marshak,[7] Teichmann,[8] and Boffi[9] (who used Laplace transforms).

2. *Purely Time-Dependent Case*

We specify the state of the system only by the time variable t; i.e., we neglect space and direction dependence and assume that the energy does not change in the collisions. $f_n(t)$ gives then the number of neutrons which have made n collisions before the time t. For the transition probability, we have

$$P(t'\to t)=\exp[-\sigma_t v(t-t')]\sigma_s v\epsilon(t-t'), \quad (14)$$

where σ_s and σ_t are the macroscopic scattering and total cross sections, respectively; v is the constant velocity and $\epsilon(t)$ is the step function defined by

$$\begin{aligned}\epsilon(t)&=1 \quad \text{for} \quad t\geq 0\\ \epsilon(t)&=0 \quad \text{for} \quad t<0.\end{aligned} \quad (15)$$

Clearly, for any T we have,

$$P(t'+T\to t+T)=P(t'\to t). \quad (16)$$

The problem admits the group of displacements in time. Since the source is usually assumed to be zero for $t<0$, it is more convenient to take the representations in the form $D^{(k)}(t)=e^{-kt}$, where k is a positive real number. Then $\Pi(k)$ and $\Phi(k)$ become the Laplace transforms of $P(t)$ and $f(t)$, respectively:

$$\Pi(k)=\int_{-\infty}^{+\infty}P(t)e^{-kt}dt=\frac{\sigma_s v}{\sigma_t v+k}, \quad (17)$$

$$\Phi(k)=\int_0^{\infty}f(t)e^{-kt}dt, \quad (18)$$

where we have assumed $f(t)=0$ for $t<0$. As in the previous example, the representation is one dimensional, Π^n is simply the nth power of Π, and the complete solution can be written as

$$f(t)=\frac{1}{2\pi i}\int_{k_1-i\infty}^{k_1+i\infty}\frac{1}{1-\Pi(k)}\Phi^{(n)}(k)e^{kt}dk. \quad (19)$$

The inversion in (19) can be readily carried out. On introducing the absorption cross section $\sigma_a=\sigma_t-\sigma_s$, one finds

$$f(t)=f_0(t)+\sigma_s v\exp(-\sigma_a vt)\int_0^t f_0(t)\exp(\sigma_a vt)dt. \quad (20)$$

This result could, of course, have been obtained by direct integration of the differential equation satisfied by $f(t)$.

C. Time- and Energy-Dependent Slowing Down

We consider now the more realistic case of time and energy (or lethargy) dependence. By taking as the elementary event a collision at energy E' and time t' (which reduces the energy to E) and the subsequent free traversal with velocity v until the next collision at time t, one can write the transition probability per unit intervals in E and t as

$$P(E',t'\to E,t)=\sigma_s(E'\to E)\times\exp[-\sigma_t v(t-t')]v\epsilon(t-t'). \quad (21)$$

The group of time displacements still leaves P invariant, but since it does not involve the energy, it is not sufficient to solve the problem. If one makes a Laplace transformation in time, the result is still an integral equation in E. Whether there exists a more general group involving both E and t depends on the energy variation of the cross sections. Let us again assume that $[\sigma_s(E'\to E)/\sigma_s(E')]dE$ is invariant under multiplication of energies by a positive real number. The remaining conditions on σ_t and σ_s/σ_t can then be enumerated as follows:

I. If $\sigma_t=a/v$ and $\sigma_s/\sigma_t=c$, where a and c are constants, there is a symmetry group for P which is the direct product of the displacement group in time and the multiplicative group in energy; or in terms of lethargy, it is the direct product of the two displacement groups in time and in lethargy. Therefore, in this case, the representations are again one dimensional and, using lethargy, may be taken as $D^{(k,w)}(t,u)=e^{-kt}e^{-uw}$. The solution will be obtained by means of a double Laplace transform:

$$\Pi(k,w)=[ca/(a+k)]Q(w), \quad (22)$$

where

$$Q(w)=\int_0^{\infty}\frac{\sigma_s(0\to u)}{\sigma_s(0)}e^{-uw}du. \quad (23)$$

Π^n is again a simple product and $f(u,t)$ can be

[4] I. Waller, Arkiv Mat. Astron. Fysik **34A**, No. 4 (1947).
[5] B. Davison, *Neutron Transport Theory* (Oxford University Press, New York, 1957), Chap. XXII.
[6] F. Adler, Phys. Rev. **60**, 279 (1941).
[7] R. E. Marshak, Revs. Modern Phys. **19**, 185 (1947).
[8] T. Teichmann, Nuclear Sci. and Eng. **7**, 292 (1960).
[9] V. C. Boffi, Ann. Phys. **9**, 435 (1960).

expressed formally as

$$f(u,t)=\frac{1}{2\pi i}\frac{1}{2\pi i}\int_{w_1-i\infty}^{w_1+i\infty}e^{uw}dw\int_{k_1-i\infty}^{k_1+i\infty}e^{kt}dk \times\frac{1}{1-\Pi(k,w)}\Phi^{(0)}(k,w). \quad (24)$$

The evaluation of this double integral has been considered by Olsson[10] for the case of isotropic elastic scattering, no absorption, and with the source $f_0(u,t)=\delta(u)e^{-\alpha t}\epsilon(t)$. He obtained an asymptotic expression valid for large u and arbitrary A, where A is the mass number of moderator nuclei.

II. In the more interesting case, again with $\sigma_s/\sigma_t=$constant, there may be a noncommutative group in (E,t) which leaves the product $\sigma_t v(t-t')$ invariant. Now, $(t-t')$ is left invariant up to a factor by a general linear transformation in t, say $t_1=\alpha t+\beta$. If $\sigma_t v=f(v)$, the product $f(v)(t-t')$ will remain invariant under multiplication of the velocity (or energy) by a positive real number provided one has

$$\alpha=1/f(v) \quad \text{and} \quad f(vv')=f(v)f(v').$$

In other words, $f(v)$ must be a representation of the multiplicative group, i.e., have the form $f(v)=v^\gamma$, where γ may have any real value. In short, if

$$\sigma_t\alpha v^\gamma \quad \text{and} \quad \sigma_s/\sigma_t=c\ (c\text{: constant}),$$

the problem admits the symmetry group in (E,t) defined by

$$\begin{aligned}(E_1,t_1)(E_2,t_2)&=(E_1E_2,\ t_1+t_2E_1^{-\kappa}) \qquad 0<E<\infty\\ (E,t)^{-1}&=(E^{-1},\ -tE^\kappa), \quad -\infty<t<+\infty\\ (1,0)&=e\ \text{(the unit element)};\end{aligned} \quad (25)$$

or in (u,t) by

$$\begin{aligned}(u_1,t_1)(u_2,t_2)&=(u_1+u_2,\ t_1+t_2e^{\kappa u_1}), \qquad -\infty<u<+\infty\\ (u,t)^{-1}&=(-u,\ -te^{-\kappa u}), \quad -\infty<t<+\infty\\ (0,0)&=e,\end{aligned} \quad (26)$$

where we have put $\kappa=(\gamma+1)/2$.

In one respect, this group is different from all the others considered in Wigner's article. Left- and right-invariant group integrations are not (cf. Appendix I) identical here.[11] In fact, in terms of (u,t), the weight factor for left-invariant integration is $e^{-\kappa u}$, while for right-invariant integration it is simply 1. One might think that this feature would cause difficulties in solving the inversion problem, but it is not so here. The positive time moments which are obtained by a direct application of the method suffice to determine the distribution functions themselves.

For simplicity, we shall carry out the calculation only for the total distribution function $f(u,t)$. Clearly the moments of $f_n(u,t)$ may be obtained in the same way. The equation for the total $f(u,t)$ can be written in terms of (u,t) as

$$f(u,t)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}f(u',t')P(u',t'\to u,t) \times e^{-\kappa u'}du'dt'+f_0(u,t), \quad (27)$$

with the conditions $f(u,t)=0$, $f_0(u,t)=0$ for $u<0$ or $t<0$. In this equation, we have

$$P(u',t'\to u,t)=p(u-u')\exp[-ae^{-\kappa u}(t-t')]ac \times\exp[-\kappa(u'-u)]\epsilon(t-t'). \quad (28)$$

Here $p(u-u')$ is the probability for the change of lethargy from u' to within a unit interval at u and is assumed to depend only on the difference $u-u'$; v_0 is the highest value of the initial velocities with respect to which the lethargy is defined by $u=2\log(v_0/v)$, $a=\sigma_{t0}v_0$, denoting by σ_{t0} the value of σ_t for v_0. Note that $f(u,t)$ is the distribution function per unit volume of group space and is related to the usual distribution function $\varphi(u,t)$ per unit $dudt$ by

$$f(u,t)e^{-\kappa u}=\varphi(u,t). \quad (29)$$

The transition probability (28) remains invariant under the operations of the group of linear transformations given by (26). Furthermore, every state (u,t) of the system can be obtained from the initial state $(u=0,\ t=0)$ by operating on (0,0) with the group element (u,t).

It is convenient to take the representations in the form (cf. Appendix II)

$$\begin{aligned}D_{\mu\nu}(u,t)&=[t^{\nu-\mu}/(\nu-\mu)!]\exp[-(\eta+\kappa\nu)u] \qquad \text{for } \nu\geq\mu\\ &=0 \qquad \text{for } \nu<\mu.\end{aligned} \quad (30)$$

We have then

$$\begin{aligned}\Pi_{\mu\nu}(\eta)&=(c/a^{\nu-\mu})Q[\eta+\kappa(\mu-1)] \quad \text{for } \nu\geq\mu\\ &=0 \qquad \text{for } \nu<\mu,\end{aligned} \quad (31)$$

where

$$Q(\eta)=\int_0^\infty e^{-\eta u}p(u)du, \quad (32)$$

and

$$\Phi_{\mu\nu}(\eta)=[1/(\nu-\mu)!]N_{\nu-\mu}[\eta+\kappa(\nu+1)]. \quad (33)$$

Here $N_\nu(\eta)$ is proportional to the Laplace transform of the νth time moment of $f(u,t)$:

$$N_\nu(\eta)=\frac{1}{\nu!}\int_0^\infty e^{-\eta u}du\int_0^\infty f(u,t)t^\nu dt. \quad (34)$$

On multiplying Eq. (27) by representation (30),

[10] O. Olsson, Arkiv Fysik **10**, No. 12 (1956).

[11] We thank Dr. V. Bargmann for kindly pointing out this fact.

integrating over the group space, defining $U_\nu(\eta)=a^{\nu+1}\times N_\nu(\eta)$, and taking $\mu=0$, we obtain finally

$$U_\nu[\eta+\kappa(\nu+1)]=c\sum_{\nu'=0}^{\nu}U_{\nu'}[\eta+\kappa(\nu'+1)]\times Q[\eta+\kappa(\nu'-1)]+V_\nu[\eta+\kappa(\nu+1)], \quad (35)$$

where V_ν corresponds to f_0 in the same way that U_ν corresponds to f.

The solution of Eq. (35) is easily found to be

$$U_\nu(\eta)=\frac{V_0(\eta-\nu\kappa)}{\prod_{\lambda=0}^{\nu}\{1-cQ[\eta-(\lambda+2)\kappa]\}}+\sum_{\mu=1}^{\nu}\frac{V_\mu[\eta-(\nu-\mu)\kappa]-V_{\mu-1}[\eta-(\nu-\mu+1)\kappa]}{\prod_{\lambda=0}^{\nu-\mu}\{1-cQ[\eta-(\lambda+2)\kappa]\}}. \quad (36)$$

The positive time moments of $f(u,t)$ can be expressed essentially as the inverse Laplace transform of this function $U_\nu(\eta)$. On multiplying with the factor $e^{-\kappa u}$ we obtain, per unit du,

$$M_\nu(u)=\int_0^\infty\varphi(u,t)t^\nu dt=\frac{\nu!e^{-\kappa u}}{a^{\nu+1}}\frac{1}{2\pi i}\times\int_{\eta_1-i\infty}^{\eta_1+i\infty}e^{\eta u}U_\nu(\eta)d\eta. \quad (37)$$

For the special case of no absorption ($c=1$), isotropic elastic scattering in c.m., and an initial monoenergetic burst $f_0(u,t)=\epsilon(t)\delta(u)e^{-at}$, the expression (37) reduces to

$$M_\nu(u)=\frac{\nu!}{a^{\nu+1}}e^{\kappa u}\frac{1}{2\pi i}\int_{\eta_1-i\infty}^{\eta_1+i\infty}e^{\eta u}\prod_{\lambda=0}^{\nu}G^{-1}(\eta-\kappa\lambda)d\eta, \quad (38)$$

with

$$G(\eta)=1-\frac{1}{\alpha}\frac{1-e^{-q(1+\eta)}}{1+\eta}, \quad (39)$$

and

$$\alpha=4A/(A+1)^2,\quad q=\log(A+1/A-1)^2,$$

where A is the mass number of the moderator nuclei.

Expression (38) was first obtained by Waller[12] from his exact solution[3] for $\varphi(u,t)$. Eriksson[13] has evaluated the integral for large values of u.

The distribution function itself is determined by the moments. It is possible to show, for instance, that Waller's expression for the solution in the above-mentioned special case may be derived from the M_ν's given by (38). On considering the Laplace transform of $\varphi(u,t)$ with respect to time, we have

$$\psi(u,\lambda)=\int_0^\infty e^{-\lambda t}\varphi(u,t)dt=\sum_{\nu=0}^{\infty}\frac{(-\lambda)^\nu}{\nu!}M_\nu(u), \quad (40)$$

or substituting for M_ν from (38),

$$=e^{\kappa u}\sum_{\nu=0}^{\infty}\frac{1}{a}\left(\frac{-\lambda}{a}\right)^\nu\frac{1}{2\pi i}\int_{\eta_1-i\infty}^{\eta_1+i\infty}e^{\eta u}\prod_{\lambda=0}^{\nu}G^{-1}(\eta-\kappa\lambda)d\eta. \quad (40a)$$

On using Waller's function $P(\eta)$ defined by the difference equation

$$P(\eta)=P(\eta+\kappa)G(\eta+\kappa), \quad (41)$$

the product in (40a) may be expressed as

$$\prod_{\lambda=0}^{\nu}G^{-1}(\eta-\lambda\kappa)=P(\eta)P^{-1}[\eta-(\nu+1)\kappa]. \quad (41a)$$

On introducing further the inverse Laplace transform of $P^{-1}(\eta)$,

$$S(\omega)=\frac{1}{2\pi i}\int_{\eta_2-i\infty}^{\eta_2+i\infty}e^{\eta\omega}P^{-1}(\eta)d\eta,$$

for which we have

$$P^{-1}(\eta)=\int_0^\infty S(\omega)e^{-\eta\omega}d\omega,$$

the summation in (40a) can be carried out to obtain,

$$\psi(u,\lambda)=e^{\kappa u}\frac{1}{2\pi i}\int_{\eta_1-i\infty}^{\eta_1+i\infty}e^{\eta u}P(\eta)d\eta\times\int_0^\infty d\omega e^{-\eta\omega}\frac{1}{\lambda+ae^{-\omega\kappa}}S(\omega). \quad (42)$$

The inverse Laplace transform of (42) with respect to λ yields Waller's solution:

$$\varphi(u,t)=e^{\kappa u}\frac{1}{2\pi i}\int_{\eta_1-i\infty}^{\eta_1+i\infty}e^{\eta u}d\eta P(\eta)\int_0^\infty\exp(-\eta\omega)\times\exp(-ate^{-\kappa\omega})d\omega\frac{1}{2\pi i}\int_{\eta_2-i\infty}^{\eta_2+i\infty}P^{-1}(\eta')e^{\eta'\omega}d\eta'. \quad (43)$$

The distribution function in the case of nonisotropic scattering and $c\neq1$ can be expressed in a similar form by defining a function $R(\eta)$ through the difference equation

$$R(\eta)=R(\eta+\kappa)\{1-cQ(\eta+\kappa)\}. \quad (44)$$

The corresponding solution (for the same source) is obtained by simply replacing $P(\eta)$ by $R(\eta)$ everywhere in (43).

To derive the solution for an arbitrary source distribution, one has to evaluate the expression corresponding

[12] I. Waller, Arkiv Fysik (to be published).
[13] K. E. Eriksson, Arkiv Fysik **16**, No. 1 (1959).

to (40), where the M_ν's are now taken from (37). An alternative way is to express the general distribution by means of the solution for a monoenergetic burst. For instance, if the initial burst has a distribution in energy given by $g(u)$, the solution may be immediately written as

$$\varphi(u,t)=\int_0^\infty du_0 g(u_0)\,\varphi(u-u_0,\,t;\,ae^{-\kappa u_0}), \qquad (45)$$

where $\varphi(u-u_0,t;ae^{-\kappa u_0})$ denotes the expression (43) in which the argument u is replaced by $u-u_0$ and the parameter a replaced by $ae^{-\kappa u_0}$.

III. *If* $\sigma_s \propto v^\gamma$ *and* $\sigma_a \propto (1/v)$, the preceding solution again applies with only a slight modification. By defining

$$e^{\sigma_a v t} f(u,t)=F(u,t),$$

one obtains for $F(u,t)$ an equation of the same form as (27) in which, however, $\sigma_s/\sigma_t=1$ and $F_0(u,t)=e^{\sigma_a v t}\times f_0(u,t)$. Consequently, as already pointed out by Waller,[3] one can write the solution for this case as

$$f(u,t)=e^{-\sigma_a v t}F(u,t)_{c=1}. \qquad (46)$$

It may be worth mentioning that the results of the three preceding sections can also be applied almost directly to a special problem of space-energy dependence in which the scattering is assumed to be spherically symmetric in the latoratory system. Then the transition probability $P(E',\mathbf{r}'\to E,\mathbf{r})$ has a similar form to the $P(E',t'\to E,t)$ of (21) with $v(t-t')$ replaced by $|\mathbf{r}-\mathbf{r}'|$. In determining the symmetry group, the exponent γ now plays the same role as did $(\gamma+1)$ for the (E,t) case; e.g., for $\sigma_t=$constant (other conditions being the same) we have the commutative group in $(E,\mathbf{r})$. If spherical symmetry is not assumed in the laboratory system, the space dependence cannnot be separated from the velocity-direction dependence.

D. Time-Energy-Space-Direction Dependent Distribution

The general slowing-down problem also admits a group which involves all the seven variables if the cross sections satisfy the conditions of Sec. C.II or C.III. The state of the system is now defined by four variables $(\mathbf{r},t)$ for space time, one variable E for the energy, and two variables $(\boldsymbol{\Omega})$ for the velocity direction. By taking the elementary event again as a collision plus the subsequent free traversal, one can write the elementary transition probability as[14]

$$P(\mathbf{r}',t',E',\boldsymbol{\Omega}'\to\mathbf{r},t,E,\boldsymbol{\Omega})$$
$$=\sigma_s(E',\boldsymbol{\Omega}'\to E,\boldsymbol{\Omega})\exp[-\sigma_t(E)|\mathbf{r}-\mathbf{r}'|]$$
$$\times\delta\left(\frac{\mathbf{r}-\mathbf{r}'}{|\mathbf{r}-\mathbf{r}'|}-\boldsymbol{\Omega}\right)\delta\left(\frac{|\mathbf{r}-\mathbf{r}'|}{v}-(t-t')\right), \qquad (47)$$

[14] E. Guth and E. Inönü, Phys. Rev. **118**, 899 (1960).

where $\sigma_s(E',\boldsymbol{\Omega}'\to E,\boldsymbol{\Omega})$ is the probability that at $\mathbf{r}'$, the neutron will undergo a scattering from E', $\boldsymbol{\Omega}'$ to within a unit interval at E, $\boldsymbol{\Omega}$. Now consider the group of transformations defined by the reducible matrix

$$\begin{pmatrix} \epsilon^{-\gamma/2}R & \mathbf{a} & & & & \\ 0 & 1 & & & & \\ & & R & & & \\ & & & \epsilon & & \\ & & & & \epsilon^{-(\gamma+1)/2} & \tau \\ & & & & 0 & 1 \end{pmatrix}, \qquad (48)$$

which operates on the vector

$$\begin{pmatrix} \mathbf{r} \\ 1 \\ \boldsymbol{\Omega} \\ E \\ t \\ 1 \end{pmatrix}$$

that characterizes the state of the system. In (48), ϵ is a positive real number, R is a rotation matrix, $\mathbf{a}$ is a space displacement vector, τ is a time displacement parameter, and γ is a real number. The group space has eight dimensions, three of which are for the displacements, three for the rotations (one can use e.g., the three Euler parameters as described in Wigner's article), and two for linear transformations in time and energy. The volume element for left-invariant integration contains a factor $\epsilon^{(\gamma+1)/2}$, so that the transition probability per unit volume element in group space is (47) multiplied by $\epsilon^{-(\gamma+1)/2}$. One can easily check that this probability is invariant under the group of transformations (48) provided the cross sections satisfy the conditions of Sec. C.II or C.III.

The representations of the group (48) can be obtained by forming the direct product of the representations of the two groups represented by the following matrices:

$$\begin{pmatrix} e^{(\gamma/2)u}R & \mathbf{a} \\ 0 & 1 \end{pmatrix} \qquad (49a)$$

and

$$\begin{pmatrix} e^{[(\gamma+1)/2]u} & t \\ 0 & 1 \end{pmatrix}, \qquad (49b)$$

where we have again used lethargy as variable. The representations of the group (49b) which lead to the positive time moments have already been discussed. The corresponding representations of the group (49a) may be taken in the following form (cf. Appendix III):

$$D^{(\eta)}{}_{\mu' l' m',\mu l m}(\mathbf{r},u,R)$$
$$=\frac{\exp[-(\tfrac{1}{2}\gamma\mu+\eta)u]}{(\mu-\mu')!}\sum_{m''=-l}^{l} D_{m'' m}{}^{(l)}(R)$$
$$\times\int(\mathbf{p}\cdot\mathbf{r})^{\mu-\mu'}Y_{lm''}(\mathbf{p})Y_{l'm'}{}^*(\mathbf{p})d\mathbf{p} \qquad (50)$$

where η may be any complex number and characterizes the representation; $D^{(l)}(R)$ is the usual $(2l+1)$-dimensional unitary representation of the rotation group, Y_{lm}'s are the normalized spherical harmonics, $\mathbf{p}$ is the radius vector for a point on the unit sphere, and the integration is over the unit sphere. R is the rotation which takes the standard velocity direction $\mathbf{\Omega}_0$ into the velocity direction $\mathbf{\Omega}$ of the state considered. Similarly, $\mathbf{r}$ and u are the position and lethargy corresponding to the state considered.

The significance of operating with the representation (50) becomes clearer when one considers simple source distributions. For instance, if we have an isotropic point source at the origin, the distribution function can only depend on r and the angle between $\mathbf{r}$ and $\mathbf{\Omega}$. In this case, writing the rotation R as $R=ST$ where T brings Ω_0 into the direction of $\mathbf{r}$ and S brings the direction of $\mathbf{r}$ into $\mathbf{\Omega}$, using Euler angles to express these rotations and carrying out all the integrations, remembering the independence of $f(r,\mathbf{\Omega},u)$ from rotations around $\mathbf{\Omega}$ and $\mathbf{r}$, one sees that the matrix element $\Phi_{000,\mu l0}$ reduces essentially to

$$\int_0^\infty\int_{-1}^{+1}\int_0^\infty f(r,\theta,u)r^2r^\mu dr P_l(\cos\theta)d(\cos\theta) \times\exp\left\{-\left[\left(\frac{\gamma}{2}\mu+1\right)+\eta\right]u\right\}du.$$

It is clear that using the representations of the full group one would be able to express in a closed form the Laplace transform with respect to lethargy of the μth space and νth time moments of the lth Legendre component of the distribution function $f(r,\theta,u,t)$. We shall not carry out this calculation here. The results in the stationary case and for spherically symmetric elastic scattering in c.m. have already been given by Placzek.[15] We would like to point out only that more complicated cases with, e.g., anisotropic sources can also be treated using the representations (50) in their general form.

ACKNOWLEDGMENT

We would like to thank Dr. E. P. Wigner for several illuminating discussions.

APPENDIX

We want to discuss here briefly the integration over the group (26) and the determination of the representations (30) and (50).

I. The weight factors for invariant integration may be obtained either by using the standard formula[16] or by direct calculation in the following way. For left-invariant group integration one should have

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(u,t)g_l(u,t)dudt = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f[(u',t')(u,t)]g_l(u,t)dudt \quad (A.1)$$

for any function $f(u,t)$ which makes the integrals converge; g_l is the weight factor to be determined. On writing out the product $(u',t')(u,t)$ according to (26) and making a change of variables, one derives from (1) the relation

$$g_l(u,t)=g_l[u-u',\ (t-t')e^{-\kappa u'}]e^{-\kappa u'} \quad (A.2)$$

which must be valid for any (u',t'). By putting in particular $u'=u$, $t'=t$, we find, defining $g(0,0)=1$,

$$g_l(u,t)=e^{-\kappa u}. \quad (A.3)$$

On the other hand, the equation for right-invariant integration,

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(u,t)g_r(u,t)dudt = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f[(u,t)(u',t')]g_r(u,t)dudt \quad (A.4)$$

gives, in the same way,

$$g_r(u,t)=1. \quad (A.5)$$

II. Consider the transformations induced in a space (ξ,τ) by the group elements $O(u,t)$. Let us define for any function $f(\xi,\tau)$ the operation $P(O)f(\xi,\tau)$ by

$$P(O)f(\xi,\tau)=f(\xi',\tau')=f[\xi-u,\ (\tau-t)e^{-\kappa u}]. \quad (A.6)$$

The operators $P(O)$ form a representation of the group (26):

$$P(O_1)P(O_2)=P(O_1O_2). \quad (A.7)$$

It follows that if one can find a set of functions $\phi_\mu(\xi,\tau)$ for which

$$P(O)\phi_\mu(\xi,\tau)=\textstyle\sum_{\mu'}D_{\mu'\mu}(u,t)\phi_{\mu'}(\xi,\tau), \quad (A.8)$$

then the coefficients $D_{\mu'\mu}(u,t)$ would also form a representation of the same group. It is easy to guess the form of the proper set for the group (26). In fact for

$$\phi_\mu(\xi,\tau)=e^{\eta\xi}[(-\tau)^\mu/\mu\ !], \quad (A.9)$$

where η is any complex number and μ any positive integer or zero, we have

$$\phi_\mu[\xi-\mu,\ (\tau-t)e^{-\kappa u}]=\sum_{\mu'=0}^{\mu} e^{-(\eta+\kappa\mu)u}\frac{t^{\mu-\mu'}}{(\mu-\mu')\,!}\phi_{\mu'}(\xi,\tau). \quad (A.10)$$

Therefore,

$$D_{\mu'\mu}^{(\eta)}(u,t)=e^{-(\eta+\kappa\mu)u}[t^{\mu-\mu'}/(\mu-\mu')\,!] \quad \text{for } \mu'\leq\mu,\ \mu=0,1,\cdots,\nu,$$
$$=0, \quad \text{for } \mu'>\mu, \quad (A.11)$$

[15] G. Placzek, declassified Rept. MDDC-2 (1946).

[16] E. P. Wigner, *Group Theory and Its Applications to the Quantum Mechanics of Atomic Spectra* (Academic Press, Inc., New York, 1959), p. 99.

in a ν-dimensional representation of the group (26) characterized by the complex variable η. All its elements below the main diagonal are equal to zero; however, the representation matrix cannot be completely reduced, as it may be checked directly. We also note that the representation is neither unitary, nor orthogonal, but has the convenient form for the evaluation of positive time moments of the distribution function $f(u,t)$.

III. The representation (50) may be obtained by using the method of "little groups."[17,18] In this method one builds up the representations of the full group out of the representations of an invariant subgroup by means of representations of the so-called little groups. The group defined by the matrix (49a) contains linear transformations $(\mathbf{a},u)$ and rotations (R). Let us denote the corresponding operators by $L(\mathbf{a},u)$ and $M(R)$. To the general group element $(\mathbf{a},u,R)$, there corresponds the operator $O(\mathbf{a},u,R)=L(\mathbf{a},u)M(R)$. The operators satisfy the following group relations:

$$\begin{aligned} L(\mathbf{a}_1,u_1)L(\mathbf{a}_2,u_2)&=L(\mathbf{a}_1+e^{wu_1}\mathbf{a}_2,\ u_1+u_2)\\ M(R_1)M(R_2)&=M(R_1R_2)\\ M(R)L(\mathbf{a},u)&=L(R\mathbf{a},u)M(R), \end{aligned} \tag{A.12}$$

where we have put $\gamma/2=w$. Consequently, we have

$$\begin{aligned} &O(\mathbf{a}_1,u_1,R_1)O(\mathbf{a}_2,u_2,R_2)\\ &\quad=O(\mathbf{a}_1+e^{wu_1}R_1\mathbf{a}_2,\ u_1+u_2,\ R_1R_2). \end{aligned} \tag{A.13}$$

The relations (A.12) indicate that the operators $L(\mathbf{a},u)$ and the corresponding transformations $(\mathbf{a},u)$ form an invariant subgroup. The representations of this subgroup (which will lead to space moments) may be written down by an immediate generalization of the representations (30). In terms of a set of functions $\phi_\mu(\mathbf{p})$, where $\mathbf{p}$ is the radius vector of a point on the unit sphere, one can write

$$L(\mathbf{a},u)\phi_\mu(\mathbf{p})=\sum_{\mu'=0}^{\mu} D_{\mu'\mu}{}^{(\eta)}(\mathbf{a},u;\ \mathbf{p})\phi_{\mu'}(\mathbf{p}), \tag{A.14}$$

where $D_{\mu'\mu}{}^{(\eta)}$ is the representation matrix defined for all values of μ, μ' as

$$\begin{aligned} D_{\mu'\mu}{}^{(\eta)}(\mathbf{a},u;\ \mathbf{p})&=e^{-(\eta+w\mu)u}[(\mathbf{p}\cdot\mathbf{a})^{\mu-\mu'}/(u-u')!] \qquad \text{for } \mu'\leq\mu\\ &=0 \qquad \text{for } \mu'>\mu. \end{aligned} \tag{A.15}$$

We select now a vector $\mathbf{p}_0$ (e.g., in the positive z direction) and keep it fixed. The two dimensional rotations P_φ and $\mathbf{p}_0$ from the little group in our problem. To associate a P_φ to every R and $\mathbf{p}$, we define for every $\mathbf{p}$ a rotation $S(\mathbf{p})$ such that

$$S(\mathbf{p})\mathbf{p}_0=\mathbf{p}. \tag{A.16}$$

Then, the rotation defined by

$$P_\varphi=S(\mathbf{p})^{-1}RS(R^{-1}\mathbf{p}) \tag{A.17}$$

leaves $\mathbf{p}_0$ invariant and hence, is a rotation around $\mathbf{p}_0$ whose angle of rotation depends on R and $\mathbf{p}$. The irreducible representations of the little group are one dimensional and have the form $e^{is\varphi}$. It then follows that all the representations of the group (49a) which are induced by the representations (A.14) can be reduced to the form

$$\begin{aligned} &O(\mathbf{a},u,R)\phi_\mu(\mathbf{p})\\ &\quad=e^{is\varphi(R,\mathbf{p})}\sum_{\mu'=0}^{\mu} D_{\mu'\mu}{}^{(\eta)}(\mathbf{a},u;\ \mathbf{p})\phi_{\mu'}(R^{-1}\mathbf{p}). \end{aligned} \tag{A.18}$$

These representations are characterized by the values of η and s. For our purposes, it is enough to consider the representations with $s=0$; these have a definite parity, while those with $s\neq0$ have not.[18] Since the transition probabilities considered in our problem are invariant under space reflexion, the representations without a definite parity cannot be used in these calculations.

To obtain the form (50), we expand $\phi_\mu(\mathbf{p})$ in spherical harmonics $Y_{lm}(\mathbf{p})$ (which are normalized over the unit sphere):

$$\phi_\mu(\mathbf{p})=\sum_{lm} A_{\mu,lm}Y_{lm}(\mathbf{p}). \tag{A.19}$$

We have then, letting $s=0$,

$$\begin{aligned} &O(\mathbf{a},u,R)\phi_\mu(\mathbf{p})\\ &\quad=\sum_{lm\mu'} A_{\mu',lm}D_{\mu'\mu}{}^{(\eta)}(\mathbf{a},u;\ \mathbf{p})Y_{lm}(R^{-1}\mathbf{p}), \end{aligned} \tag{A.20a}$$

which can be written in terms of $Y_{l'm'}(\mathbf{p})$ as

$$=\sum_{lm,\mu'} A_{\mu',lm}\sum_{l'm'} D^{(\eta)}{}_{\mu'l'm',\mu lm}(\mathbf{a},u,R)Y_{l'm'}(\mathbf{p}). \tag{A.20b}$$

The expansion coefficients $D_{\mu'l'm',\mu lm}$ form the required representation of the group (49a). On comparing (A.20a) and (A.20b), we obtain

$$\begin{aligned} &D^{(\eta)}{}_{\mu'l'm',\mu lm}(\mathbf{a},u,R)\\ &\quad=\int D_{\mu'\mu}{}^{(\eta)}(\mathbf{a},u;\ \mathbf{p})Y_{l'm'}{}^{*}(\mathbf{p})Y_{lm}(R^{-1}\mathbf{p})d\mathbf{p}, \end{aligned} \tag{A.21}$$

where the integration is over the unit sphere. The form (50) is obtained from (A.21) by expressing $Y_{lm}(R^{-1}\mathbf{p})$ in terms of the representations $D^{(l)}(R)_{m'm}$ of the rotation group.

[17] See, e.g., J. S. Lomont, *Applications of Finite Groups* (Academic Press, Inc., New York, 1959), p. 230.

[18] E. Inönü and E. P. Wigner, Nuovo cimento **9**, 706 (1952).