Structure of the Generalized Veneziano Amplitudes.* I. Radon Transforms

Alfred C. T. Wu
Department of Physics, University of Michigan, Ann Arbor, Michigan
(Received 31 August 1970)

The (tree-graph) generalized Veneziano amplitudes are shown to be boundary values of a new class of generalized hypergeometric functions having the property that they are Radon transforms of products of *linear* forms.

1. INTRODUCTION

The Veneziano model has attracted considerable interest.[1-12] We address ourselves here to the question of the functional structure of the generalized Veneziano amplitudes.

As is well known, the original Veneziano amplitude for a four-body process was written down as a beta function:

$$V_4(-\alpha_s,-\alpha_t)=\frac{\Gamma(-\alpha_s)\Gamma(-\alpha_t)}{\Gamma(-\alpha_s-\alpha_t)}=\int_0^1 dx x^{-\alpha_s-1}(1-x)^{-\alpha_t-1}. \tag{1}$$

V_4 can obviously be written as a special case of the Gauss hypergeometric function ${}_2F_1(a,b;c;w)$:

$$V_4(-\alpha_s,-\alpha_t)=\Gamma(-\alpha_s)\,\Gamma(1+\alpha_s)\times {}_2F_1(-\alpha_s,\alpha_s+\alpha_t+1;1;1). \tag{2}$$

Furthermore, Bialas and Pokorski[8] pointed out that the five-particle amplitude as generalized by Bardakci and Ruegg[2] is a special case of the generalized hypergeometric function ${}_3F_2(a,b,c;d,e;w)$ at $w=1$:

$$V_5(-\alpha_{ij})=\frac{\Gamma(-\alpha_{12})\Gamma(-\alpha_{23})}{\Gamma(-\alpha_{12}-\alpha_{23})}\frac{\Gamma(-\alpha_{34})\Gamma(-\alpha_{45})}{\Gamma(-\alpha_{34}-\alpha_{45})}\times {}_3F_2(\alpha_{15}-\alpha_{23}-\alpha_{34},-\alpha_{45},-\alpha_{12};-\alpha_{34}-\alpha_{45},-\alpha_{12}-\alpha_{23};1). \tag{3}$$

The question immediately arises as to what the corresponding statements for the higher amplitudes are. We wish to make the following observations:

(a) The N-particle Veneziano amplitudes[4-6,9-11] $V_N, N\geq 6$, evidently do not belong to any familiar class of generalized hypergeometric functions. Thus a naive extrapolation that V_6 might be related to some function like ${}_pF_q(a_1\ldots a_p;b_1,\ldots,b_q;w)$ at $w=1$ is obviously false. (b) Radon structure: The class of generalized Veneziano amplitudes[4-6,9-11] corresponding to the so-called tree graphs may be regarded as the boundary values of a new class of generalized hypergeometric functions in several variables. This class of functions satisfies the criterion that they are Radon transforms of products of *linear* forms.[13] On the other hand, this simply property apparently breaks down when one considers the generalized amplitudes (for $N>4$) corresponding to the nonplanar graphs, such as those discussed by Virasoro[14] and Mandelstam.[15]

2. GENERALIZED VENEZIANO AMPLITUDES AS BOUNDARY VALUES OF A CLASS OF GENERALIZED HYPERGEOMETRIC FUNCTIONS

We define a class of generalized hypergeometric functions $F^{(n)}(a_i,b_i,c_{ij};w_{ij})$ of $\frac{1}{2}n(n-1)$ variables w_{ij} as follows:

$$F^{(n)}(a_i,b_i,c_{ij},w_{ij})=\int_0^1\cdots\int\prod_{k=1}^n dx_k\prod_{k=1}^n[x_k^{a_k}(1-x_k)^{b_k}]\times\prod_{i<j}(x_i-w_{ij}x_j)^{c_{ij}} \tag{4}$$

These functions justify the name of generalized hypergeometric functions in the sense of Gel'fand,[13] namely, they are Radon transforms of products of *linear* forms in an $(n+1)$-dimensional space. Obviously, one can write

$$F^{(n)}(a_i,b_i,c_{ij};w_{ij})=\int\cdots\int\prod_{k=1}^{n+1}dx_k\,\delta(1-\Sigma x_i)\prod_{k=1}^{N}(\xi^{(k)},x)^{s_k}, \tag{5}$$

where $N=\frac{1}{2}n(n+3)$ and

$$(\xi^{(k)},x)=\sum_{i=1}^{n+1}\xi_i^{(k)}x_i$$

is a linear form in x with coefficients ξ_i^k. The w's enter through the ξ's.

Special cases of $F^{(n)}$ are easily recognizable, for example,

$$F^{(1)}(a,b)=\Gamma(a+1)\Gamma(-a)\times {}_2F_1(a+1,-(a+b+1);1;1) \tag{6}$$

$$F^{(2)}(a_i,b_i,c_{12};w_{12})=\frac{\Gamma(a_2+1)\Gamma(b_2+1)}{\Gamma(a_2+b_2+2)}\frac{\Gamma(a_1+c_{12}+1)\Gamma(b_1+1)}{\Gamma(a_1+b_1+c_{12}+2)}\times {}_3F_2(-c_{12},a_2+1,-(a_1+b_1+c_{12}+1);a_2+b_2+2,-(a_1+c_{12});w_{12}). \tag{7}$$

As far as the author is aware, $F^{(n)}$ for $n>3$ does not seem to correspond to any known function in the mathematical literature.

It can be easily verified that by simple change of variables, the $(n+3)$-point generalized Veneziano function[4-6,9-11] can be brought to the form of Eq. (4) with all $w_{ij}=1$, namely,

$$V_{n+3}=\text{const}\,F^{(n)}(a_i,b_i,c_{ij};w_{ij}=1), \tag{8}$$

where the parameters [analogs of the α_{ij} in Eq. (3)] entering in V_{n+3} are combinations of the a's, b's, and c's.

To be precise, the suggested change of variables for V_{n+3} consists of letting

$$x_k = \prod_{l=1}^{k} u_l, \qquad k = 1, 2, \ldots, n, \tag{9}$$

where the u's are the usual integration variables [4-6, 9-11] successively associated with the internal lines of a multiperipheral graph. This prescription can be easily checked with the aid of Chan's explicit formulas[4-6] for $n = 3$ and $n = 4$, or with the form written down by Bardakci and Ruegg[11].

In a following paper, we shall study the structure of the generalized Veneziano amplitudes from the point of view of group representations.

ACKNOWLEDGEMENT

This work owes its origin to the author's unavoidable exposure to the Veneziano disease during his stay at CERN in 1969.

* Work supported in part by the U.S. Office of Naval Research, under Contract No. NONR-1224(59).
1 G. Veneziano, Nuovo Cimento **57A**, 190 (1968).
2 K. Bardakci and H. Ruegg, Phys. Letters **28B**, 242 (1968).
3 M. A. Virasoro, Phys. Letters **22**, 37 (1969).
4 H.-M. Chan, Phys. Letters **28B**, 425 (1969).
5 H.-M. Chan and S. T. Tsou, CERN Preprint TH 969, 1968 (unpublished).
6 H.-M. Chan and S. T. Tsou, Phys. Letters **28B**, 485 (1969).
7 J. L. Hopkinson and E. Plahte, Phys. Letters **28B**, 489 (1969).
8 A. Bialas and S. Pokorski, Nucl. Phys. B10, 399 (1969).
9 C. L. Goebel and B. Sakita, Phys. Rev. Letters **22**, 259 (1969).
10 Z. Koba and H. B. Nielson, Nucl. Phys. **B12**, 517 (1969).
11 K. Bardakci and H. Ruegg, Phys. Rev. **181**, 1884 (1969).
12 See, e.g., H.-M. Chan, 'The Generalized Veneziano Model' review paper presented at the Royal Society Meeting on Duality Reggeons and Resonances in Elementary Particle Processes, 1969, CERN Preprint TH 1057, 1969.
13 I. M. Gel'fand, M. I. Graev, and N. Ya. Vilenkin, *Generalized Functions*, (Academic, New York, 1966), Vol. V.
14 M. A. Virasoro, Phys. Rev. **177**, 2309 (1969).
15 S. Mandelstam, Phys. Rev. **183**, 1374 (1969); Phys. Rev. D **1**, 1720 (1970).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 10 OCTOBER 1971

Structure of the Generalized Veneziano Amplitudes.* II. Possible Group Theoretic Content

Alfred C. T. Wu
Department of Physics, University of Michigan, Ann Arbor, Michigan
(Received 31 August 1970)

An attempt is made to generalize a result of Vilenkin that the beta function (being adopted as Veneziano formula for four-particle processes) appears as the kernel (in the integral) of the irreducible representation (in the Mellin-transformed space) of the group of 2×2 unimodular triangular matrices. It is shown that with a modified multiplier in the Gel'fand–Naimark prescription for the representation of the group of $(n+1) \times (n+1)$ unimodular triangular matrices, the (tree-graph) generalized Veneziano amplitude for $(n+3)$-particle processes is recovered by a limiting procedure.

1. INTRODUCTION

In a previous paper,[1] we discussed the functional structure of the $(n+3)$-point (tree-graph) Veneziano amplitudes viewed as the boundary values of a new class of generalized hypergeometric functions having the property that they are the Radon transforms of products of *linear* forms in an $(n+1)$-dimensional space.

In this paper, we shall investigate possible group theoretic content of the generalized Veneziano amplitudes.

Many special functions have acquired a new level of respectability as well as a deeper *raison d'être* when it can be shown that a given function occurs naturally in the representation theory of certain groups, even though this connection is in general not one to one. While it is by now more or less a standard textbook exercise[2-5] in going from the representations of known groups to known functions,[6] the converse route (namely, given the function, finding the group) is obviously much more hazardous. Nevertheless, we venture to point out that, in the mathematical literature, certain results exist for the connection between the representations of the group $SL(2, R)$ and the Gauss hypergeometric function ${}_2F_1(a, b; c; x)$ and, in particular, the special case of triangular matrices of $SL(2, R)$ yields the beta function as the kernel.[7]

The apparent interest in the Veneziano model[8] perhaps justifies an attempt to generalize the Vilenkin result for the beta function.

A straightforward application of the Gel'fand–Naimark scheme[9] for the irreducible representations of $(n+1) \times (n+1)$ unimodular triangular matrices will yield in general a class of functions different from the class of functions $F^{(n)}$ discussed in Paper I. The differences generally are twofold. (i) In one aspect, the multiplier in the Gel'fand–Naimark scheme, which consists of the products of n principal minors[10] $\Delta_{p\,p+1\ldots n+1}$ $(p = 2, \ldots, n+1)$, is found to be insufficient in generating all the desired tree-graph links in the Veneziano formula. One can remedy this if one modifies the

where the parameters [analogs of the α_{ij} in Eq. (3)] entering in V_{n+3} are combinations of the a's, b's, and c's.

To be precise, the suggested change of variables for V_{n+3} consists of letting

$$x_k = \prod_{l=1}^{k} u_l, \qquad k = 1, 2, \ldots, n, \tag{9}$$

where the u's are the usual integration variables [4-6, 9-11] successively associated with the internal lines of a multiperipheral graph. This prescription can be easily checked with the aid of Chan's explicit formulas[4-6] for $n = 3$ and $n = 4$, or with the form written down by Bardakci and Ruegg[11].

In a following paper, we shall study the structure of the generalized Veneziano amplitudes from the point of view of group representations.

ACKNOWLEDGEMENT

This work owes its origin to the author's unavoidable exposure to the Veneziano disease during his stay at CERN in 1969.

* Work supported in part by the U.S. Office of Naval Research, under Contract No. NONR-1224(59).
1 G. Veneziano, Nuovo Cimento **57A**, 190 (1968).
2 K. Bardakci and H. Ruegg, Phys. Letters **28B**, 242 (1968).
3 M. A. Virasoro, Phys. Letters **22**, 37 (1969).
4 H.-M. Chan, Phys. Letters **28B**, 425 (1969).
5 H.-M. Chan and S. T. Tsou, CERN Preprint TH 969, 1968 (unpublished).
6 H.-M. Chan and S. T. Tsou, Phys. Letters **28B**, 485 (1969).
7 J. L. Hopkinson and E. Plahte, Phys. Letters **28B**, 489 (1969).
8 A. Bialas and S. Pokorski, Nucl. Phys. B10, 399 (1969).
9 C. L. Goebel and B. Sakita, Phys. Rev. Letters **22**, 259 (1969).
10 Z. Koba and H. B. Nielson, Nucl. Phys. **B12**, 517 (1969).
11 K. Bardakci and H. Ruegg, Phys. Rev. **181**, 1884 (1969).
12 See, e.g., H.-M. Chan, 'The Generalized Veneziano Model' review paper presented at the Royal Society Meeting on Duality Reggeons and Resonances in Elementary Particle Processes, 1969, CERN Preprint TH 1057, 1969.
13 I. M. Gel'fand, M. I. Graev, and N. Ya. Vilenkin, *Generalized Functions*, (Academic, New York, 1966), Vol. V.
14 M. A. Virasoro, Phys. Rev. **177**, 2309 (1969).
15 S. Mandelstam, Phys. Rev. **183**, 1374 (1969); Phys. Rev. D **1**, 1720 (1970).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 10 OCTOBER 1971

Structure of the Generalized Veneziano Amplitudes.* II. Possible Group Theoretic Content

Alfred C. T. Wu
Department of Physics, University of Michigan, Ann Arbor, Michigan
(Received 31 August 1970)

An attempt is made to generalize a result of Vilenkin that the beta function (being adopted as Veneziano formula for four-particle processes) appears as the kernel (in the integral) of the irreducible representation (in the Mellin-transformed space) of the group of 2×2 unimodular triangular matrices. It is shown that with a modified multiplier in the Gel'fand–Naimark prescription for the representation of the group of $(n+1) \times (n+1)$ unimodular triangular matrices, the (tree-graph) generalized Veneziano amplitude for $(n+3)$-particle processes is recovered by a limiting procedure.

1. INTRODUCTION

In a previous paper,[1] we discussed the functional structure of the $(n+3)$-point (tree-graph) Veneziano amplitudes viewed as the boundary values of a new class of generalized hypergeometric functions having the property that they are the Radon transforms of products of *linear* forms in an $(n+1)$-dimensional space.

In this paper, we shall investigate possible group theoretic content of the generalized Veneziano amplitudes.

Many special functions have acquired a new level of respectability as well as a deeper *raison d'être* when it can be shown that a given function occurs naturally in the representation theory of certain groups, even though this connection is in general not one to one. While it is by now more or less a standard textbook exercise[2-5] in going from the representations of known groups to known functions,[6] the converse route (namely, given the function, finding the group) is obviously much more hazardous. Nevertheless, we venture to point out that, in the mathematical literature, certain results exist for the connection between the representations of the group $SL(2, R)$ and the Gauss hypergeometric function ${}_2F_1(a, b; c; x)$ and, in particular, the special case of triangular matrices of $SL(2, R)$ yields the beta function as the kernel.[7]

The apparent interest in the Veneziano model[8] perhaps justifies an attempt to generalize the Vilenkin result for the beta function.

A straightforward application of the Gel'fand–Naimark scheme[9] for the irreducible representations of $(n+1) \times (n+1)$ unimodular triangular matrices will yield in general a class of functions different from the class of functions $F^{(n)}$ discussed in Paper I. The differences generally are twofold. (i) In one aspect, the multiplier in the Gel'fand–Naimark scheme, which consists of the products of n principal minors[10] $\Delta_{p\,p+1\ldots n+1}$ $(p = 2, \ldots, n+1)$, is found to be insufficient in generating all the desired tree-graph links in the Veneziano formula. One can remedy this if one modifies the

multiplier by a product of $\frac{1}{2}(n-1)(n-2)$ uniquely defined cofactors. As an example, the case $n=4$ will call for factors of Δ_{23}, Δ_{34}, and Δ_{234} in addition to those Δ_{2345}, Δ_{345}, Δ_{45}, and Δ_5 inherent in the Gel'fand–Naimark scheme. (ii) On the other hand, the group of $(n+1)\times(n+1)$ unimodular triangular matrices for $n>2$ are actually richer than necessary for generating the Veneziano functions. To specialize, we only need the non-vanishing entries in the elements $g_{i\,i+1(i=1,\ldots,n)}$ besides the diagonal elements. In order to preserve the group property, we cannot arbitrarily set those unwanted elements directly equal to zero since this special subset of triangular matrices is not closed under group multiplication. So, one can only approach the desired result by letting some parameters vanish in the limit sense.

Under these two provisos, the tree-graph generalized Veneziano functions are recovered from representations of $(n+1)\times(n+1)$ unimodular triangular matrices.

For the sake of readability, the essential steps in the establishment[7] of the connection between the beta function and the representation of the 2×2 unimodular triangular matrices are summarized in Sec. 2. This will serve as a prototype upon which our generalization will be based. We hope that by going over this simplest case, our generalization to the higher-rank case will not be obscured by the mounting algebraic complexities. To keep the notation straight, we devoted Sec. 3 to setting up the Gel'fand–Naimark machinery for the representations of $(n+1)\times(n+1)$ unimodular matrices. A modified multiplier will be defined and the representation operator will be carried into the Mellin-transformed space. A limiting procedure will be stated to get to the generalized Veneziano function.

2. BETA FUNCTION AND GROUP OF 2×2 UNIMODULAR TRIANGULAR MATRICES

The essential steps in the derivation are as follows:[7]

A. Irreducible Representation of $SL(2,\mathrm{R})$

Step 1: The representation is to be realized in the space of (infinitely differentiable, square-integrable) homogeneous functions of degree ρ. A signature factor $\epsilon(x)$ is included. On account of homogeneity, we have, for

$$g\equiv\begin{pmatrix} g_{11} & g_{12}\\ g_{21} & g_{22}\end{pmatrix},$$

$$T_g f(z)=|g_{12}z+g_{22}|^{\rho}[\epsilon(g_{12}z+g_{22})]^m \times f\left(\frac{g_{11}z+g_{21}}{g_{12}z+g_{22}}\right), \tag{1}$$

where $\epsilon(x)$ is a short-hand notation for $x/|x|$.

Note that the representation is unitary for pure imaginary ρ.

Step 2: (a) Carry the representation intc the Mellin-transformed space. It will be convenient to separate out the support properties over the positive and negative half-spaces:

$$\tilde{T}_g\tilde{f}_{\pm}(s)=\int_{-\infty}^{\infty} dz z_{\pm}^{s-1}|g_{12}z+g_{22}|^{\rho} \times[\epsilon(g_{12}z+g_{22})]^m f\left(\frac{g_{11}z+g_{21}}{g_{12}z+g_{22}}\right), \tag{2}$$

where

$$\tilde{f}_{\pm}(s)\equiv\int_{-\infty}^{\infty} dz\, z_{\pm}^{s-1}f(z), \tag{3}$$

$$z_+^s\equiv\theta(z)z^s,\ [\theta(z)\text{ is the usual step function}], \tag{4a}$$

$$z_-^s\equiv\theta(-z)(-z)^s. \tag{4b}$$

(b) Re-expressing $f(x)$ in terms of $\tilde{f}_{\pm}(t)$ by making an inverse Mellin transform and interchanging the order of integration which can be justified, we get, by writing in the matrix form,

$$\tilde{T}_g\begin{pmatrix}\tilde{f}_+(s)\\ \tilde{f}_-(s)\end{pmatrix}=\int_{c-i\infty}^{c+i\infty} dt\begin{pmatrix}K_{++} & K_{+-}\\ K_{-+} & K_{--}\end{pmatrix}\begin{pmatrix}\tilde{f}_+(t)\\ \tilde{f}_-(t)\end{pmatrix}, \tag{5}$$

where the kernel is given by

$$K_{\pm\pm}(s,t;\rho,m;g)=\frac{1}{2\pi i}\int_{-\infty}^{\infty} dz\, z_{\pm}^{s-1} \times\left(\frac{g_{11}z+g_{21}}{g_{12}z+g_{22}}\right)_{\pm}^{-t}|g_{12}z+g_{22}|^{\rho}[\epsilon(g_{12}z+g_{22})]^m. \tag{6}$$

Step 3: Identification: For the case of 2×2 Hermitian or unitary matrices, the kernel is readily recognized as the Gauss hypergeometric function ${}_2F_1$.

B. The Special Case of Triangular Matrices

In this case, the kernel is reduced to the beta function. Take

$$g=\begin{pmatrix}1 & a\\ 0 & 1\end{pmatrix}, \tag{7}$$

$$K_{++}(s,t;\rho;g)=(2\pi i)^{-1}\int_0^{\infty} dz z^{s-t-1}(1+az)_+^{\rho+t} =(2\pi i)^{-1}a^{t-s}\beta(s-t,-\rho-s). \tag{8}$$

3. REPRESENTATIONS OF $(n+1)\times(n+1)$ UNIMODULAR TRIANGULAR MATRICES

We now generalize the discussion of Sec. 2 to the case of general n.

Step 1: The representation will be realized

in the space of (infinitely differentiable, square-integrable) functions $f(z_{ij})$ such that

$$T_g f(z_{ij}) = \alpha(\hat{g}) f(\hat{z}_{ij}), \tag{9}$$

where T_g is the representation operator corresponding to the group element g. The set of variables $z_{ij}, i > j$, in general are $\frac{1}{2}n(n+1)$ in number and are cast in the form of a triangular matrix with $z_{ii} = 1$ along the diagonal:

$$z = \begin{pmatrix} 1 & & & & & \\ z_{21} & 1 & & & & \\ z_{31} & z_{32} & 1 & & & \\ \cdot & & \cdot & \cdot & & \\ \cdot & & & \cdot & \cdot & \\ \cdot & & & & \cdot & \\ z_{n+1} & & & & z_{n+1,n} & 1 \end{pmatrix}. \tag{10}$$

The matrix $\hat{g}$ in (9) is defined as

$$\hat{g} \equiv zg. \tag{11}$$

The variables $\hat{z}_{ij}$ on the right-hand side of (9) are given as follows:

$$\hat{z}_{ij} = \Delta^{-1}_{i\,i+1\ldots n+1} \begin{vmatrix} \hat{g}_{ij} & \hat{g}_{i\,i+1} & \cdots & \hat{g}_{i\,n+1} \\ \hat{g}_{i+1\,j} & \hat{g}_{i+1,i+1} & \cdots & \hat{g}_{i+1,n+1} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \hat{g}_{n+1\,j} & \hat{g}_{n+1,i+1} & \cdots & \hat{g}_{n+1,n+1} \end{vmatrix}, \tag{12}$$

where

$$\Delta_{i\,i+1\ldots s} \equiv \begin{vmatrix} \hat{g}_{ii} & \hat{g}_{i\,i+1} & \cdots & \hat{g}_{is} \\ \hat{g}_{i+1\,i} & \hat{g}_{i+1,i+1} & \cdots & \hat{g}_{i+1,s} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \hat{g}_{s\,i} & \hat{g}_{s,i+1} & \cdots & \hat{g}_{s,s} \end{vmatrix}; \tag{13}$$

i.e., $\Delta_{i\,i+1\ldots s}$ is a short-hand notation for the principal minor of matrix $\hat{g}$ starting from the (ii)th element and ending with the (ss)th element.[10]

For the multiplier $\alpha(\hat{g})$ in (9), we choose the following expression:

$$\alpha(\hat{g}) = \prod_{j=2}^{n} \prod_{i=2}^{j+1} |\Delta_{i\,i+1\ldots j+1}|^{\rho_{ij}} [\epsilon(\Delta_{i\,i+1\ldots j+1})]^{m_{ij}}, \tag{14}$$

with the understanding that

$$\Delta_{n+1\,n+1} \equiv \hat{g}_{n+1\,n+1}. \tag{15}$$

The block of terms with $j = n$ corresponds to the multiplier of Gel'fand–Naimark,[9] while the presence of the remainder factors from $j = 2$ to $j = n - 1$ may be interpreted as due to the process of cyclic completion.

Step 2: With nearly the same procedure as carried out in Sec. 2, we cast now the representation in the Mellin-transformed space. For the eventual purpose we have in mind, we shall only consider the special case by setting

$$z_{ij} = 0 \quad \text{if} \quad i > j + 1. \tag{16}$$

Thus there are altogether n variables $z_{i+1\,i}$.

We get the analog of Eq. (2):

$$\tilde{T}_g \tilde{f}_{\pm\ldots\pm}(s_{i+1,i}) = \int_{-\infty}^{\infty} \cdots \int \prod_{i=1}^{n} dz_{i+1\,i} \times \prod_{i=1}^{n} (z_{i+1,i})_{\pm}^{s_{i+1,i}-1} \alpha(\hat{g}) f(\hat{z}_{j+1\,j}). \tag{17}$$

Expressing f in terms of the $\tilde{f}$ gives

$$\tilde{T}_g \tilde{f}_{\pm\ldots\pm}(s_{i+1,i}) = \int_{c-i\infty}^{c+i\infty} \cdots \int \prod_{k=1}^{n} dt_{k+1\,k} \times \sum K^{(n)}_{(\pm\pm)\ldots(\pm\pm)}(s,t;g) \tilde{f}_{\pm\ldots\pm}(t_{j+1,j}), \tag{18}$$

where

$$K^{(n)}_{(\pm\pm)\ldots(\pm\pm)}(s,t;g) \equiv \frac{1}{(2\pi i)^n} \int_{-\infty}^{\infty} \cdots \int \prod_{i=1}^{n} dz_{i+1,i} \prod_{i} (z_{i+1,i})_{\pm}^{s_{i+1,i}-1} \times \prod_{j} (\hat{z}_{j+1,j})_{\pm}^{t_{j+1,j}} \alpha(\hat{g}). \tag{19}$$

Step 3: Identification: To specialize, we take g to be triangular:

$$g_{ij} = 0 \quad \text{for} \quad i > j, \quad g_{ii} = 1. \tag{20}$$

Our main proposition will be the following:

Lemma: In the limit of vanishing g_{ij} for $j > i + 1$, the kernel $K^{(n)}_{+\ldots+}$ given by (19) reduces to the $(n+3)$-point Veneziano function of the form[1]

$$V_{n+3} = \text{const} \int_0^1 \cdots \int \prod_{i=1}^{n} dx_i \times \prod_{k} [x_k^{a_k} (1-x_k)^{b_k}] \prod_{i<j} (x_i - x_j)^{c_{ij}}. \tag{21}$$

Proof: We shall make the technical assumption that the limit $g_{ij} \to 0$ for $j > i + 1$ can be taken inside the integral (19). When that is the case, the matrix $\hat{g}$ of (11) takes the form of (22), where a superscript 0 denotes the limit when appropriate:

$$\hat{g}^0 = zg^0 = \begin{pmatrix} 1 & & & & \\ z_{21} & 1 & & & \\ 0 & z_{32} & 1 & & \\ \cdot & \cdot & \cdot & \cdot & \\ \cdot & \cdot & & \cdot & \\ \cdot & \cdot & & \cdot & \cdot \\ 0 & 0 & & z_{n+1n} & 1 \end{pmatrix} \begin{pmatrix} 1 & g_{12} & 0 & \cdots & 0 \\ & 1 & g_{13} & & \\ & & 1 & & \\ & & & \cdot & \\ & & & & \cdot \\ & & & & \cdot\, g_{nn+1} \\ & & & & 1 \end{pmatrix}. \tag{22}$$

All the minors $\Delta^0_{j\ldots s}$ of $\hat{g}^0$ can be evaluated in a straightforward manner. The essential step involved is showing that, by appropriate change of variables, all the factors in (21) are recovered from (19).

We find the following sequence of variable transformations convenient: (i) First scale $z_{i+1\,i}$ by $g_{i\,i+1}$, i.e., let

$$z'_{i+1\,i} = z_{i+1\,i}\, g_{i\,i+1}, \qquad i = 1, \cdots, n, \tag{23}$$

(ii) Let

$$z'_{i+1\,i} = v_i/(1 - v_i), \qquad i = 1, \cdots, n. \tag{24}$$

The range for v_i is obviously $(0, 1)$. (iii) Let

$$\begin{aligned} x_1 &= 1 - v_2(1 - v_1), \\ x_2 &= 1 - v_2, \\ x_3 &= x_2 v_3, \\ x_4 &= x_3[1 - v_4(1 - v_3)]^{-1}, \\ &\cdot \\ &\cdot \\ &\cdot \\ x_m &= x_3 \hat{g}^0_{44}\, \Delta^0_{56\ldots m+1}/\Delta^0_{456\ldots m+1}, \\ & 5 \leq m \leq n. \end{aligned} \tag{25}$$

We note in passing that the inverse transformations from x to v can all be expressed in terms of the anharmonic ratio of four points:

$$\begin{aligned} v_1 &= R(x_1, 1, x_2, \infty), \\ v_2 &= R(x_2, 0, 1, \infty), \\ v_3 &= R(x_3, x_2, 0, \infty), \\ v_4 &= R(x_4, x_2, x_3, 0), \\ v_m &= R(x_m, x_2, x_{m-1}, x_{m-2}) \quad \text{for} \quad 5 \leq m \leq n, \end{aligned} \tag{26}$$

where

$$R(\xi_1, \xi_2, \xi_3, \xi_4) \equiv \frac{(\xi_1 - \xi_3)(\xi_2 - \xi_4)}{(\xi_1 - \xi_4)(\xi_2 - \xi_3)} \tag{27}$$

is the so-called anharmonic ratio of four points.[11]

It is then a simple matter to verify that the set of factors in (19) is mapped into the set of factors in (21). The coefficients a's, b's, c's in (21) are expressible as linear combination of the s's, t's and ρ's in (19).

* Work supported in part by the U.S. Office of Naval Research, under Contract No. NONR-1224 (59).

1 A. C. T. Wu, J. Math. Phys. 12, 2035 (1971).

2 E. P. Wigner "Applications of Group Theory to the Special Functions of Mathematical Physics," unpublished lecture notes, Princeton University, Princeton, N.J., 1955.

3 N. J. Vilenkin, *Special Functions and the Theory of Group Representations*, (Am. Math. Soc., Providence, R.I., 1968).

4 W. Miller, Jr., *Lie Theory and Special Functions* (Academic, New York, 1968).

5 J. D. Talman, *Special Functions—A Group Theoretic Approach* (Benjamin, New York, 1968).

6 Such as (i) from the group of rotations in three dimensions $O(3)$ to the Jacobi polynomials which include the Legendre polynomials as special cases, (ii) from the group of Euclidean motions in two dimensions $E(2)$ to the Bessel functions, (iii) from the Poincare group in two dimensions to the Hankel functions, etc.

7 Ref. 3, Chap. 7, especially p. 367.

8 See, e.g., Paper I, Ref. 12.

9 I. M. Gel'fand and M. A. Naimark, *Unitäre Darstellungen der Klassischen Gruppen* (Akademie-Verlag, Berlin, 1957).

10 Our $\Delta_{p\,p+1\ldots n}$ corresponds to the g^1_p of Ref. 9.

11 Anharmonic ratios appeared in an entirely different approach in Z. Koba and H. B. Nielson, Nucl. Phys. B12, 517 (1969).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 10 OCTOBER 1971

Exact Solution of a Time-Dependent Quantal Harmonic Oscillator with a Singular Perturbation

P. Camiz
Istituto di Fisica, Universitá di Roma,
Istituto Nazionale di Fisica Nucleare, Sezione di Roma, Rome, Italy

and

A. Gerardi* and C. Marchioro*
Istituto Matematico dell'Universitá di Roma, Rome, Italy

and

E. Presutti and E. Scacciatelli
Istituto di Fisica, Facoltá di Ingegneria, Universitá di Roma, Rome, Italy
Gruppo Nazionale di Struttura della Materia del Consiglio Nazionale delle Ricerche, Rome, Italy
(Received 4 January 1971)

The quantal problem of a particle interacting in one dimension with an external time-dependent quadratic potential and a constant inverse square potential is exactly solved. The solutions are found both in the Schrödinger representation, by using a generating function or a time-dependent raising operator, and in the Heisenberg picture. They depend only on the solution of the classical harmonic oscillator. The generalizations to the n-dimensional problem and to the problem of N particles in one dimension, interacting pairwise via a quadratic time-dependent potential and a constant inverse square potential, are finally sketched.

1. INTRODUCTION

It is well known that the time-dependent Schrödinger equation only in few cases can be exactly solved, so that usually approximation methods are needed. An exact solution can then provide, as well as a possible model for physical phenomena, a significiant test for these approximation methods.

The harmonic oscillator with variable frequency is one of these exactly solvable problems largely studied for its great physical interest and its relative simplicity.[1-3].

In this paper we consider a perturbed harmonic oscillator with a time-dependent frequency that is the one-dimensional problem of a particle interacting with an external quadratic time-dependent potential and a constant inversely quadratic potential. We give the explicit solutions of the time-dependent Schrödinger equation that at $t = 0$ reproduce the eigenfunctions of the stationary problem (i.e., the case of a constant strength of the force). So the general solution is found. It is shown that it depends only on the solution of the classical imperturbed oscillator.

In Sec. 2 the problem is solved by introducing a time dependent generating function. In Sec. 3 a raising operator is found for the stationary problem, and its generalization to the time-dependent case is given. In Sec. 4 the Heisenberg representation is studied. Finally in Sec. 5 the previous results are used to sketch the solutions of the n-dimensional problem and of the more complicated problem of N particles in one dimension interacting pairwise via the sum of a time-dependent quadratic potential and a constant inversely quadratic potential.

2. SOLUTION VIA THE GENERATING FUNCTION METHOD

Let the Hamiltonian of the system be

$$H(t) = P^2/2m + \tfrac{1}{2}m\omega^2(t)x^2 + g/x^2, \qquad (2.1)$$

where $\omega(t)$ is a regular function of the time and g is a constant $g > -\hbar^2/8m$ to prevent collapse.[4]

The time-dependent Schrödinger equation is

$$\left(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + \tfrac{1}{2}m\omega^2(t)x^2 + \frac{g}{x^2}\right)\psi(x,t) = i\hbar\frac{\partial}{\partial t}\psi(x,t). \qquad (2.2)$$

We shall confine ourselves in the sector $x \geq 0$; in fact, the singular nature of the g/x^2 term forbids any transition between the sectors $x > 0$ and $x < 0$, so that the solution can be extended to the $x < 0$ region without any further condition of continuity.

We want to find the solutions $\psi_n(x,t)$ of Eq. (2.2) which reproduce at $t = 0$ solutions of the stationary Schrödinger equation [$\omega(0) = \omega_0$]:

$$\left(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + \tfrac{1}{2}m\omega_0^2x^2 + \frac{g}{x^2}\right)\Phi(x) = E\Phi(x). \qquad (2.3)$$

The solutions of Eq. (2.3) are[5]

$$\Phi_n(x) = \left(\frac{4m\omega_0}{\hbar}\right)^{1/4}\left(\frac{\Gamma(n+1)}{\Gamma(a+n+1)}\right)^{1/2}\left(\frac{m\omega_0}{\hbar}x^2\right)^{(2a+1)/4} \times \exp\left(-\frac{m\omega_0}{2\hbar}x^2\right)L_n^a\left(\frac{m\omega_0x^2}{\hbar}\right), \qquad (2.4)$$

$$E_n = \hbar\omega_0(2n + a + 1), \qquad (2.5)$$

where

$$a = \tfrac{1}{2}(1 + 8mg/\hbar^2)^{1/2}. \qquad (2.6)$$

We introduce the generating function

$$G(z,x,t) = \sum_{n=0}^{\infty}\left(\frac{\Gamma(a+n+i)}{\Gamma(n+1)}\right)^{1/2}\Psi_n(x,t)z^n, \qquad (2.7)$$

where z is a subsidiary variable.

$G(z,x,t)$ satisfies Eq. (2.2) with initial conditions [from Eq. (2.4)]

$$G(z,x,0) = \left(\frac{4m\omega_0}{\hbar}\right)^{1/4}\left(\frac{m\omega_0}{\hbar}x^2\right)^{(2a+1)/4}$$

$$\times \exp\left(\frac{m\omega_0}{2\hbar} x^2 \frac{(z+1)}{(z-1)}\right)(1-z)^{-(a+1)}. \quad (2.8)$$

We make the ansatz

$$G(z,x,t) = A\alpha(z,t)x^{a+1/2} \exp[\beta(z,t)x^2], \quad (2.9)$$

where

$$A = (2)^{1/2}(m\omega_0/\hbar)^{(a+1)/2}. \quad (2.10)$$

Inserting Eq. (2.9) in (2.2) and in (2.8), we obtain

$$\frac{\partial\beta(z,t)}{\partial t} = i\frac{2\hbar}{m}\beta^2 - \frac{im}{2\hbar}\omega^2(t), \quad \beta(z,0) = \frac{m\omega_0}{2\hbar}\frac{z+1}{z-1}, \quad (2.11)$$

$$\frac{\partial\alpha(t,z)}{\partial t} = i\frac{2\hbar}{m}(a+1)\alpha(z,t)\beta(z,t), \quad \alpha(z,0) = (1-z)^{-(a+1)}. \quad (2.12)$$

If we define $\eta(t) = |\eta(t)| \exp[i\gamma(t)]$ as the solution of the classical equation of motion

$$\ddot{\eta}(t) + \omega^2(t)\eta(t) = 0, \quad \eta(0) = 1, \ \dot{\eta}(0) = -i\omega_0, \quad (2.13)$$

from Eqs (2.11) and (2.12) we obtain

$$\alpha(z,t) = [\eta(t)]^{-(a+1)} \exp[2i\gamma(t)(a+1)]\{1 - z\exp[2i\gamma(t)]\}^{-(a+1)}, \quad (2.14)$$

$$\beta(t) = \frac{im}{2\hbar}\left(\frac{\dot{\eta}(t)}{\eta(t)} - 2i\dot{\gamma}(t)\{1 - z\exp[2i\gamma(t)]\}^{-1}\right). \quad (2.15)$$

Thus

$$G(z,x,t) = Ax^{(2a+1/2)}[\eta(t)]^{-(a+1)}(1-s)^{-(a+1)} \times \exp\left(\frac{im}{2\hbar}\frac{\dot{\eta}(t)x^2}{\eta(t)} + 2i\gamma(t)(a+1) + \frac{m}{\hbar}\dot{\gamma}(t)x^2\right)\exp\left(-\frac{m}{\hbar}\dot{\gamma}(t)x^2\frac{s}{s-1}\right), \quad (2.16)$$

where $s = z\exp[2i\gamma(t)]$.

We have thereby obtained an expression of $G(z,x,t)$ formally similar to Eq. (2.8); hence from Eq. (2.7) we have

$$\psi_n(x,t) = \left[2\left(\frac{m\omega_0}{\hbar}\right)^{a+1}\frac{\Gamma(n+1)}{\Gamma(a+n+1)}\right]^{1/2}[\eta(t)]^{-(a+1)} \times x^{(2a+1)/2}\exp\left[2i\gamma(t)(a+n+1) \times \frac{m}{\hbar}x^2\left(\frac{i\dot{\eta}(t)}{2\eta(t)} + \dot{\gamma}(t)\right)\right]L_n^a\left(-\frac{mx^2}{\hbar}\dot{\gamma}(t)\right). \quad (2.17)$$

Using the relation $|\eta(t)|^2\dot{\gamma}(t) = -\omega_0$, one can verify that the solutions $\psi_n(x,t)$ of Eq. (2.17) reduce, for $g = 0$, to the solutions of the time-dependent harmonic oscillator $\Psi_{2n+1}(x,t)$ given in Ref. 2.

3. RAISING OPERATOR

The linearity of the energy spectrum of the stationary case suggests the existence of a raising operator A^+, obeying the equation

$$[H(0), A^+] = (\Delta E)A^+, \quad [H(0), A] = -(\Delta E)A. \quad (3.1)$$

A solution of this equation is given by

$$A^+ = a^+a^+ + g/\hbar\omega_0 x^2, \quad (3.2)$$

where a^+ is the raising operator of the harmonic oscillator:

$$a^+ = i(2m\hbar\omega_0)^{-1/2}\left(-\hbar\frac{\partial}{\partial x} + m\omega_0 x\right). \quad (3.3)$$

The corresponding variation in energy is given by

$$\Delta E = 2\hbar\omega_0. \quad (3.4)$$

The energy spectrum is lower bounded, and so we must find the eigenfunctions $\psi_k(x)$ of the stationary problem for which the equation

$$A\psi_k(x) = 0 \quad (3.5)$$

holds.

The corresponding eigenvalues are easily found by the relation

$$(A\psi_K(x), A\psi_K(x)) = (\psi_K(x), A^+A\psi_K(x)) = \left(\psi_K(x), \left[\left(\frac{H(0)}{\hbar\omega_0} - \frac{1}{2}\right)^2 - \frac{H(0)}{\hbar\omega_0} - \frac{2mg}{\hbar^2} + \frac{1}{2}\right] \times \psi_K(x)\right) = \frac{E_K^2}{\hbar^2\omega_0^2} - \frac{2E_K}{\hbar\omega_0} + \frac{3}{4} - \frac{2gm}{\hbar^2} = 0, \quad (3.6)$$

from which the ground-state energy results:

$$E_0 = \hbar\omega_0[1 \pm \tfrac{1}{2}(1 + 8gm/\hbar^2)^{1/2}], \quad (3.7)$$

where only the upper sign is allowed, as we can see at glance from Eq. (3.7).

Once $\psi_0(x)$ is determined, the whole set of eigenfunctions can be built up by means of the operator A^+.

To solve now the time-dependent problem, we seek, in analogy with the method of Ref. 3 concerning the pure harmonic oscillator, a time-dependent operator $\tilde{A}^+(x, \partial/\partial x, t)$ characterized by the following properties:

(a) Operating on a solution $\Psi(x,t)$ of the time-dependent Schrödinger equation, it gives another solution $\tilde{\Psi}(x,t)$ of the same equation.

(b) At $t = 0$ it reduces to the operator A^+ defined by Eq. (3.2).

It is easily proved that $\tilde{A}^+(x, \partial/\partial x, t)$ must satisfy the equation[3]

$$[H(t), \tilde{A}^+] = i\hbar \frac{\partial \tilde{A}^+}{\partial t}, \tag{3.8}$$

with the initial conditions

$$\tilde{A}^+\left(x, \frac{\partial}{\partial x}, 0\right) = A^+. \tag{3.9}$$

The operator obtained by the following ansatz,

$$\tilde{A}^+(t) = \left(\mathrm{E}_1(t)\frac{\partial}{\partial x} + \mathrm{E}_2(t)x\right)^2 + \frac{\mathrm{E}_3(t)}{x^2}, \tag{3.10}$$

satisfies Eqs. (3.8) and (3.9), if $\mathrm{E}_1(t)$, $\mathrm{E}_2(t)$, and $\mathrm{E}_3(t)$ are solutions of the simple equations

$$\begin{aligned}
&\ddot{\mathrm{E}}_1(t) + \omega^2(t)\mathrm{E}_1(t) = 0, \quad \mathrm{E}_1(0) = -i(\hbar/2m\omega_0)^{1/2},\\
&\mathrm{E}_2(t) = -i(m/\hbar)\dot{\mathrm{E}}_1(t), \quad \mathrm{E}_2(0) = i(m\omega_0/2\hbar)^{1/2},\\
&\mathrm{E}_3(t) = -(2m/\hbar^2)g\mathrm{E}_1(t)^2.
\end{aligned} \tag{3.11}$$

The time dependence of any given initial state can be obtained once we know the solution $\psi_0(x,t)$, which for $t = 0$ reduces to the ground-state eigenfunction. It can be observed that

$$\tilde{A}\left(x, \frac{\partial}{\partial x}, t\right)\psi_0(x,t) = 0 \tag{3.12}$$

holds, which is an ordinary differential equation in the variable x with time-dependent coefficients. The solution $\psi_0(x,t)$ is then found by solving Eq. (3.12), and a complete set of solutions of Eq. (2.2) can be obtained by multiple applications of the time-dependent raising operator $\tilde{A}^+(x, \partial/\partial x, t)$.

4. HEISENBERG REPRESENTATION

In the Heisenberg representation we are interested at the time evolution of the operators $\hat{x}(t)$ and $\hat{P}(t)$; it is convenient to study first the evolution of three operators $\hat{B}(t)$, $\hat{C}(t)$, and $\hat{D}(t)$ characterized by the initial conditions

$$\hat{B}(0) = \frac{\hat{P}^2}{2m} + \frac{g}{\hat{x}^2}, \tag{4.1}$$

$$\hat{C}(0) = \hat{x}^2, \tag{4.2}$$

$$\hat{D}(0) = \hat{x}\hat{p} + \hat{p}\hat{x}. \tag{4.3}$$

Starting from the Heisenberg equations

$$\begin{aligned}
-i\hbar\frac{d\hat{B}(t)}{dt} &= [\hat{H}(t), \hat{B}(t)],\\
-i\hbar\frac{d\hat{C}(t)}{dt} &= [\hat{H}(t), \hat{C}(t)],\\
-i\hbar\frac{d\hat{D}(t)}{dt} &= [\hat{H}(t), \hat{D}(t)],
\end{aligned} \tag{4.4}$$

we obtain the system of differential linear equations:

$$\begin{aligned}
&\frac{d\hat{B}(t)}{dt} = -\frac{\omega^2(t)}{2}\hat{D}(t), \quad \frac{d\hat{C}(t)}{dt} = \frac{\hat{D}(t)}{m},\\
&\frac{d}{dt}\hat{D}(t) = 4\hat{B}(t) - 2m\omega^2(t)\hat{C}(t).
\end{aligned} \tag{4.5}$$

We make the ansatz that $\hat{B}(t)$, $\hat{C}(t)$, and $\hat{D}(t)$ can be expressed as a linear combination (with time-dependent coefficients) of the operators $\hat{B}(0)$, $\hat{C}(0)$, $\hat{D}(0)$; comparing then separately in Eq. (4.5) the coefficients of the three operators $\hat{B}(0)$, $\hat{C}(0)$, and $\hat{D}(0)$, we obtain a set of nine ordinary linear differential equations, which give a link between the nine coefficients. In conclusion we can write the solutions of Eq. (4.5) in the form

$$\begin{aligned}
\hat{B}(t) &= (1/\omega_0^2)\dot{f}_1^2(t)\hat{B}(0) + \tfrac{1}{2}m\dot{f}_2^2(t)\hat{C}(0)\\
&\quad + (1/\omega_0)\,\mathrm{Im}[\dot{f}_3^2(t)]\hat{D}(0),\\
\hat{C}(t) &= (2/\omega_0^2 m)\, f_1^2(t)\hat{B}(0) + f_2^2(t)\hat{C}(0)\\
&\quad + (2/m\omega_0)\,\mathrm{Im}[f_3^2(t)]\hat{D}(0),\\
\hat{D}(t) &= (4/\omega_0^2)\,\dot{f}_1(t)f_1(t)\hat{B}(0)\\
&\quad + 2m\dot{f}_2(t)f_2(t)\hat{C}(0) + (4/\omega_0)\,\mathrm{Im}[f_3(t)\dot{f}_3(t)]\hat{D}(0),
\end{aligned} \tag{4.6}$$

where the $f_J(t)$s satisfy the classical oscillator equation

$$\ddot{f}_J(t) + \omega^2(t)f_J(t) = 0, \tag{4.7}$$

with the initial conditions

$$\begin{aligned}
&f_1(0) = 0, \quad f_2(0) = 1, \quad f_3(0) = 1, \quad \dot{f}_1(0) = \omega_0,\\
&\dot{f}_2(0) = 0, \quad \dot{f}_3(0) = \tfrac{1}{4}i\omega_0.
\end{aligned} \tag{4.8}$$

Now it is a trivial matter to express $\hat{x}(t)$ and $\hat{p}(t)$ in terms of $\hat{B}(t)$ and $\hat{D}(t)$:

$$\hat{x}^2(t) = \hat{C}(t), \quad \hat{p}(t) = [\hat{D}(t) - i\hbar][2\hat{x}(t)]^{-1}, \tag{4.9}$$

and this achieves our task.

5. GENERALIZATIONS

(a) Let us consider a particle moving in a potential

$$V(\mathbf{r}) = \tfrac{1}{2}m\omega^2(t)|\mathbf{r}|^2 + g/|\mathbf{r}|^2 \tag{5.1}$$

in an n-dimensional space. By expanding the wavefunction in terms of hyperspherical harmonics, the problem is immediately reduced to the angular equation of a free particle and to a radial equation of the form of (2.2), where the effective coupling constant is now the sum of g plus $1/2m$ times the eigenvalue of the squared total hyperangular momentum.

(b) Let us consider N identical particles in one dimension interacting pairwise via an inverse square potential and a quadratic time-dependent potential:

$$V_{ij} = \tfrac{1}{2}m\omega^2(t)(x_i - x_J)^2 + g(x_i - x_J)^{-2}. \tag{5.2}$$

The stationary problem has been studied by Calogero,[6] who gave the whole set of eigenvalues and the characterization of the eigenfunctions. Introducing the mean square radius r,

$$r^2 = \frac{1}{N}\sum_{i=2}^{N}\sum_{j=1}^{i-1}(x_i - x_j)^2, \tag{5.3}$$

we separate the problem into an angular part independent on $\omega(t)$ and a radial part of the form of Eq. (2.2), where the effective coupling constant, as in the previous case, depends on the eigenvalues of an operator acting only on the "angular" variables.

The problems (a) and (b) can then be solved by using the generating function method.

The raising operator method can be, in principle, applied to both problems, but the stationary problems are not yet solved by this method, owing to their intrinsic complications. The research for the complete algebra of the raising operators in case (b), in particular, for the three-body problem is in progress.

* Research partially supported by the Consiglio Nazionale delle Ricerche.

1 R. P. Feynman, Rev. Mod. Phys. **20**, 367 (1948); Phys. Rev. **84**, 108 (1951); J. Schwinger, *ibid.* **91**, 728 (1953); K. Husimi, Progr. Theoret. Phys. (Kyoto) **9**, 381 (1953); A. M. Dykhne, Zh. Eksp. Teor. Fiz. **38**, 570 (1960) [Sov. Phys. JETP **11**, 411 (1960)]; M. Kolsrud, Phys. Rev. **104**, 1186 (1956); R. M. Kulsrud, *ibid.* **106**, 205 (1957); F. Hetwek and A. Schluter, Z. Naturforschr. **12a**, 844 (1957); G. Backus, A. Lenard, and R. Zulsud, *ibid.* **15a**, 1007 (1960); S. Chandrasekhar, in *The Plasma in a Magnetic Field*, R. K. M. Landshoff, Ed. (Stanford U.P., Palo Alto, Calif., (1958), p. 3; M. Kolsrud, Kgl. Norske Videnskab. Selskabs, Forh., **31**, No. 5 (1958); A. Lenard, Ann. Phys. (N. Y.) **6**, 261 (1959); L. M. Garrido, Progr. Theoret. Phys. (Kyoto) **26**, 577 (1961); J. E. Littlewood, Ann. Phys. (N.Y.) **21**, 233 (1963); **29**, 13 (1964); H. R. Lewis, Phys. Rev. Letters **18**, 510 (1967); J. Math. Phys. **9**, 1976 (1968); H. R. Lewis and W. B. Reisenfeld, J. Math. Phys. **10**, 1458 (1969); B. Crosignani, P. Di Porto, and S. Solimeno, *ibid.* **10**, 1922 (1969); W. H. Louisell, A. Yariv, and E. Sigman, Phys. Rev. **124**, 1646 (1961); B. R. Mollow and R. J. Glauber *ibid.* **160**, 1076, 1097 (1967); B. R. Mollow, ibid. **162**, 1256 (1967); C. Marchioro, E. Presutti, and E. Scacciatelli, "Breaking of Coherence in a Monochromatic Photon Beam by an External Quadratic Time-Dependent Interaction" (unpublished).

2 V. S. Popov, A. M. Perelomov, Zh. Eksp. Theor. Fiz. **56**, 1375 (1969); **57**, 1684 (1969) [Sov. Phys. JETP **29**, 738 (1969); **30**, 911 (1970)].

3 E. Salusti and F. Zirilli, Nuovo Cimento Letters **4**, 999 (1970).

4 L. D. Landau and E. M. Lifshitz, *Quantum Mechanics* (Pergamon, New York, 1958), Sec. 35.

5 F. Calogero, J. Math. Phys. **10**, 2191 (1969).

6 F. Calogero, J. Math. Phys. **12**, 419 (1971).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 10 OCTOBER 1971

Relaxation of Local Thermal Deviations from Equilibrium

Gérard G. Emch* and Charles Radin
(Instituut voor Theoretische Fysica, Universiteit Nijmegen, The Netherlands)
(Received 31 March 1971)

We give precise conditions sufficient to guarantee that an infinite system will act as a thermal reservoir for any of its finite parts. In particular we show that these conditions are satisfied for the X-Y model. Further results on the ergodic behavior and general relaxation properties of the systems considered are also obtained directly from the C^*-algebraic methods used in the main body of the paper.

1. HEURISTIC STATEMENT OF THE PROBLEM

A physical system is ordinarily defined by the attribution of a Hamiltonian $H(\Omega)$ to every finite piece $\Sigma(\Omega)$ of it; here Ω denotes the spatial extension of $\Sigma(\Omega)$.

Given $\Sigma(\Omega)$, we cut out of it a piece $\Sigma(\Omega_0)$ and denote by $\Sigma(\Omega\backslash\Omega_0)$ the remaining part of $\Sigma(\Omega)$. One would presume, according to the phenomenological laws of thermodynamics, that if Ω is very large compared to Ω_0, then $\Sigma(\Omega\backslash\Omega_0)$ serves as a *thermal reservoir* for $\Sigma(\Omega_0)$. The traditional approach of statistical mechanics suggests that the mechanical description of this situation will be mathematically simpler (and thus more efficient) if the limit of large systems is taken. To this effect, we shall first suppose that $\Sigma(\Omega_0)$ is of *finite* extent and that $\Sigma(\Omega)$ is of *infinite* extent in all directions around Ω_0. We shall then discuss briefly the thermodynamical consequences of letting Ω_0 become very large.

To formulate more specifically the problem to which we want to address ourselves, we consider an initial situation where (i) the system $\Sigma(\Omega_0)$, cut off from the rest of $\Sigma(\Omega)$, is in the canonical equilibrium state ϕ_{Ω_0,β_0} corresponding to $H(\Omega_0)$ and the natural temperature β_0 and (ii) the system $\Sigma(\Omega)$ extends over the entire physical space and $\Sigma(\Omega\backslash\Omega_0)$, cut off from $\Sigma(\Omega_0)$, is in the state $\phi_{R,\beta}$ obtained as the thermodynamical limit (as Ω' becomes infinite) of the states $\phi_{\Omega'\backslash\Omega_0,\beta}$, which are defined as the canonical equilibrium states computed from $H(\Omega'\backslash\Omega_0)$ for the natural temperature $\beta \neq \beta_0$. One would then expect that this state of the composite system "relaxes" to the state $\phi_{S+R,\beta}$ obtained as the thermodynamical limit of the canonical equilibrium states $\phi_{\Omega',\beta}$ computed from $H(\Omega')$, these Hamiltonians taking into account the interactions between $\Sigma(\Omega_0)$ and $\Sigma(\Omega\backslash\Omega_0)$.

On a more modest level, one would at least expect that the time average (as well as the space average)

The stationary problem has been studied by Calogero,[6] who gave the whole set of eigenvalues and the characterization of the eigenfunctions. Introducing the mean square radius r,

$$r^2 = \frac{1}{N}\sum_{i=2}^{N}\sum_{j=1}^{i-1}(x_i - x_j)^2, \tag{5.3}$$

we separate the problem into an angular part independent on $\omega(t)$ and a radial part of the form of Eq. (2.2), where the effective coupling constant, as in the previous case, depends on the eigenvalues of an operator acting only on the "angular" variables.

The problems (a) and (b) can then be solved by using the generating function method.

The raising operator method can be, in principle, applied to both problems, but the stationary problems are not yet solved by this method, owing to their intrinsic complications. The research for the complete algebra of the raising operators in case (b), in particular, for the three-body problem is in progress.

* Research partially supported by the Consiglio Nazionale delle Ricerche.

1 R. P. Feynman, Rev. Mod. Phys. **20**, 367 (1948); Phys. Rev. **84**, 108 (1951); J. Schwinger, *ibid.* **91**, 728 (1953); K. Husimi, Progr. Theoret. Phys. (Kyoto) **9**, 381 (1953); A. M. Dykhne, Zh. Eksp. Teor. Fiz. **38**, 570 (1960) [Sov. Phys. JETP **11**, 411 (1960)]; M. Kolsrud, Phys. Rev. **104**, 1186 (1956); R. M. Kulsrud, *ibid.* **106**, 205 (1957); F. Hetwek and A. Schluter, Z. Naturforschr. **12a**, 844 (1957); G. Backus, A. Lenard, and R. Zulsud, *ibid.* **15a**, 1007 (1960); S. Chandrasekhar, in *The Plasma in a Magnetic Field*, R. K. M. Landshoff, Ed. (Stanford U.P., Palo Alto, Calif., (1958), p. 3; M. Kolsrud, Kgl. Norske Videnskab. Selskabs, Forh., **31**, No. 5 (1958); A. Lenard, Ann. Phys. (N. Y.) **6**, 261 (1959); L. M. Garrido, Progr. Theoret. Phys. (Kyoto) **26**, 577 (1961); J. E. Littlewood, Ann. Phys. (N.Y.) **21**, 233 (1963); **29**, 13 (1964); H. R. Lewis, Phys. Rev. Letters **18**, 510 (1967); J. Math. Phys. **9**, 1976 (1968); H. R. Lewis and W. B. Reisenfeld, J. Math. Phys. **10**, 1458 (1969); B. Crosignani, P. Di Porto, and S. Solimeno, *ibid.* **10**, 1922 (1969); W. H. Louisell, A. Yariv, and E. Sigman, Phys. Rev. **124**, 1646 (1961); B. R. Mollow and R. J. Glauber *ibid.* **160**, 1076, 1097 (1967); B. R. Mollow, ibid. **162**, 1256 (1967); C. Marchioro, E. Presutti, and E. Scacciatelli, "Breaking of Coherence in a Monochromatic Photon Beam by an External Quadratic Time-Dependent Interaction" (unpublished).

2 V. S. Popov, A. M. Perelomov, Zh. Eksp. Theor. Fiz. **56**, 1375 (1969); **57**, 1684 (1969) [Sov. Phys. JETP **29**, 738 (1969); **30**, 911 (1970)].

3 E. Salusti and F. Zirilli, Nuovo Cimento Letters **4**, 999 (1970).

4 L. D. Landau and E. M. Lifshitz, *Quantum Mechanics* (Pergamon, New York, 1958), Sec. 35.

5 F. Calogero, J. Math. Phys. **10**, 2191 (1969).

6 F. Calogero, J. Math. Phys. **12**, 419 (1971).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 10 OCTOBER 1971

Relaxation of Local Thermal Deviations from Equilibrium

Gérard G. Emch* and Charles Radin
(Instituut voor Theoretische Fysica, Universiteit Nijmegen, The Netherlands)
(Received 31 March 1971)

We give precise conditions sufficient to guarantee that an infinite system will act as a thermal reservoir for any of its finite parts. In particular we show that these conditions are satisfied for the X-Y model. Further results on the ergodic behavior and general relaxation properties of the systems considered are also obtained directly from the C^*-algebraic methods used in the main body of the paper.

1. HEURISTIC STATEMENT OF THE PROBLEM

A physical system is ordinarily defined by the attribution of a Hamiltonian $H(\Omega)$ to every finite piece $\Sigma(\Omega)$ of it; here Ω denotes the spatial extension of $\Sigma(\Omega)$.

Given $\Sigma(\Omega)$, we cut out of it a piece $\Sigma(\Omega_0)$ and denote by $\Sigma(\Omega\backslash\Omega_0)$ the remaining part of $\Sigma(\Omega)$. One would presume, according to the phenomenological laws of thermodynamics, that if Ω is very large compared to Ω_0, then $\Sigma(\Omega\backslash\Omega_0)$ serves as a *thermal reservoir* for $\Sigma(\Omega_0)$. The traditional approach of statistical mechanics suggests that the mechanical description of this situation will be mathematically simpler (and thus more efficient) if the limit of large systems is taken. To this effect, we shall first suppose that $\Sigma(\Omega_0)$ is of *finite* extent and that $\Sigma(\Omega)$ is of *infinite* extent in all directions around Ω_0. We shall then discuss briefly the thermodynamical consequences of letting Ω_0 become very large.

To formulate more specifically the problem to which we want to address ourselves, we consider an initial situation where (i) the system $\Sigma(\Omega_0)$, cut off from the rest of $\Sigma(\Omega)$, is in the canonical equilibrium state ϕ_{Ω_0,β_0} corresponding to $H(\Omega_0)$ and the natural temperature β_0 and (ii) the system $\Sigma(\Omega)$ extends over the entire physical space and $\Sigma(\Omega\backslash\Omega_0)$, cut off from $\Sigma(\Omega_0)$, is in the state $\phi_{R,\beta}$ obtained as the thermodynamical limit (as Ω' becomes infinite) of the states $\phi_{\Omega'\backslash\Omega_0,\beta}$, which are defined as the canonical equilibrium states computed from $H(\Omega'\backslash\Omega_0)$ for the natural temperature $\beta \neq \beta_0$. One would then expect that this state of the composite system "relaxes" to the state $\phi_{S+R,\beta}$ obtained as the thermodynamical limit of the canonical equilibrium states $\phi_{\Omega',\beta}$ computed from $H(\Omega')$, these Hamiltonians taking into account the interactions between $\Sigma(\Omega_0)$ and $\Sigma(\Omega\backslash\Omega_0)$.

On a more modest level, one would at least expect that the time average (as well as the space average)

of the initial state described above is equal to $\phi_{S+R,\beta}$.

If, furthermore, Ω_0 is taken to be large enough, one would expect that, *well inside* Ω_0, the system $\Sigma(\Omega_0)$ would, in the same sense as above, approach *its own* thermal equilibrium for the natural temperature β.

These expected properties are clearly of an ergodic character. Pathological cases where they would not be satisfied are likely to occur as well. The question to which we want to address ourselves is rather whether one can *prove* explicitly (i.e., without uncontrollable approximations or unwarranted statistical assumptions) that there are mechanical models that *do* exhibit these properties.

We shall present such a proof. Some related ergodic properties will be proven as well, along the principal line of argument. We shall then touch upon the question of an actual (i.e., pointwise in time) relaxation to equilibrium. The latter problem is that of ultimate physical interest. Some results are already known in this connection.[1,2] However, information on the latter type of problem is usually obtained through detailed computations which are restricted to the models analyzed (e.g., one makes use of the precise behavior of the excitation spectrum as the thermodynamical limit is taken). In contrast, the results presented in this paper are obtained by methods of a more general character, and as such, stand a better chance to extend further than the details specific to the models treated.

2. THE X-Y MODEL

We consider a one-dimensional lattice-spin system. To every site i in $\mathbb{Z}$ is attached a spin-$\frac{1}{2}$ particle σ_i, and hence a copy $\mathfrak{A}_i$ of the C^*-algebra of the 2×2 matrices. The observables attached to each finite region Ω in $\mathbb{Z}$ are therefore the Hermitian elements of the C^*-algebra $\mathfrak{A}_\Omega = \bigotimes_{i \in \Omega} \mathfrak{A}_i$, and the C^*-algebra of quasilocal observables on the infinite lattice is the C^*-inductive limit:

$$\mathfrak{A} = \overline{\bigcup_{\Omega \subset \mathbb{Z}} \mathfrak{A}_\Omega}\,.$$

To every interval $[a,b] = \{i \in \mathbb{Z} \mid a \leq i \leq b\}$ of $\mathbb{Z}$ we attribute the Hamiltonian

$$H_{[a,b]} = -\sum_{a=1}^{b-1} (1+\zeta)\,\sigma_i^x \sigma_{i+1}^x + (1-\zeta)\,\sigma_i^y \sigma_{i+1}^y\,,$$

the time evolution $\alpha_{[a,b]}(t)$ defined by

$$\alpha_{[a,b]}(t)\,[A] = \exp(-iH_{[a,b]}t)A \exp(iH_{[a,b]}t)$$
$$\text{for all } A \in \mathfrak{A}_{[a,b]},$$

and the canonical equilibrium state $\phi_{[a,b]}$ defined by

$$\langle \phi_{[a,b]}; A\rangle = \mathrm{Tr}[\exp(-\beta H_{[a,b]})A]/\mathrm{Tr}[\exp(-\beta H_{[a,b]})]$$
$$\text{for all } A \in \mathfrak{A}_{[a,b]}.$$

From the fact that the interaction from which the Hamiltonian $H_{[a,b]}$ is built is of finite range, we know[3,4] that a time evolution and a Gibbs state can be naturally defined for the infinite system described by $\mathfrak{A}$. Specifically there exists a one-parameter, strongly continuous group of automorphisms $\alpha(t)$ of $\mathfrak{A}$ and a state ϕ on $\mathfrak{A}$ such that, for every finite Ω_0 in $\mathbb{Z}$ and $A \in \mathfrak{A}_{\Omega_0}$,

$$\lim_{a \to -\infty,\, b \to +\infty} \|\alpha(t)\,[A] - \alpha_{[a,b]}(t)\,[A]\| = 0,$$
$$\lim_{a \to -\infty,\, b \to +\infty} |\langle \phi; A\rangle - \langle \phi_{[a,b]}; A\rangle| = 0.$$

Furthermore, ϕ is uniformly clustering, i.e., for every $A \in \mathfrak{A}$ and $\epsilon > 0$ there exists a finite N such that

$$|\langle \phi; AB\rangle - \langle \phi; A\rangle\langle \phi; B\rangle| \leq \epsilon \|B\|$$

for all B in $\mathfrak{A}_{(\mathbb{Z}\setminus[-N,N])}$. Moreover, ϕ satisfies the KMS boundary condition with respect to $\alpha(t)$ (for definitions see Ref. 5). As a consequence[6] of the uniform clustering of ϕ, ϕ is[4] extremal with respect to the KMS condition; so ψ KMS and $\psi \leq \lambda\phi$ together imply $\psi = \phi$.

We now define the automorphism γ of $\mathfrak{A}$ by means of

$$\gamma[\sigma_i^z] = \sigma_i^z;\ \gamma[\sigma_i^x] = -\sigma_i^x;\ \gamma[\sigma_i^y] = -\sigma_i^y$$

for all i in $\mathbb{Z}$. Let us denote by $\mathfrak{A}_e$ the C^*-subalgebra of $\mathfrak{A}$ consisting of its "even" elements, i.e.,

$$\mathfrak{A}_e = \{A \in \mathfrak{A} \mid \gamma[A] = A\}.$$

Clearly $H_{[a,b]}$ belongs to $\mathfrak{A}_e$. As a consequence, γ commutes with each $\alpha_\Omega(t)$ and hence with $\{\alpha(t) \mid t \in \mathbb{R}\}$; also, ϕ is even (i.e., $\langle \phi; \gamma[A]\rangle = \langle \phi; A\rangle$ for all $A \in \mathfrak{A}$). Hence ϕ is determined by its restriction ϕ_e to $\mathfrak{A}_e$, and $\alpha(t)$ maps $\mathfrak{A}_e$ onto itself. Let $\alpha_e(t)$ denote the restriction of $\alpha(t)$ to $\mathfrak{A}_e$. From the corresponding properties of ϕ, one concludes immediately that ϕ_e is KMS with respect to $\alpha_e(t)$ and is uniformly clustering on $\mathfrak{A}_e$. One then checks that the arguments of Ref. 6, Properties 2.2 and 2.3, go through for $\mathfrak{A}_e$, and hence one concludes that ϕ_e is extremal KMS on $\mathfrak{A}_e$.

The arguments developed so far apply to any one-dimensional lattice-spin system with even, finite-range, lattice-invariant interaction. We now use a specific ergodic property of the X-Y model. As is well known, the Jordan–Wigner transformation[7] brings $H_{[a,b]}$ into a form which is quadratic in Fermi operators. It can be seen that in this form the interaction satisfies the assumptions of theorem II in Ref. 8, so that $\{\alpha_e(t) \mid t \in \mathbb{R}\}$ acts as a strongly asymptotically Abelian group of automorphisms of $\mathfrak{A}_e$, i.e.,

$$\lim_{|t| \to \infty} \|[A, \alpha_e(t)\,[B]]\| = 0 \quad \text{for all} \quad A, B \text{ in } \mathfrak{A}_e.$$

Therefore, if η is any invariant mean on $\mathbb{R}$, we have

$$\eta\langle \phi_e; C^*[A, \alpha_e(\cdot)[B]]C\rangle = 0$$

for all A, B and C in $\mathfrak{A}_e$. Consequently,[9] the state ϕ_e on $\mathfrak{A}_e$ is not only extremal KMS, but also extremal time invariant. Hence $\psi_e \leq \lambda\phi_e$ and ψ_e time invariant together imply $\psi_e = \phi_e$. In particular, if ψ is even on $\mathfrak{A}$, time invariant, and satisfies $\psi \leq \lambda\phi$, we can conclude that $\psi = \phi$, a fact that we shall use repeatedly in the next section.

3. ERGODIC BEHAVIOR

Let $[c,d]$ be a finite interval in $\mathbb{Z}$. For each finite interval $[a,b]$ in $\mathbb{Z}$ such that $a < c-1$ and $b > d+1$, we define

$$H^c_{[a,b]} = H_{[a,c-1]} + H_{[c,d]} + H_{[d+1,b]} = H_{[a,b]} + V,$$

with V independent of a and b. We then define the state $\phi^c_{[a,b]}$ on $\mathfrak{A}_{[a,b]}$ by

$$\langle \phi^c_{[a,b]}; A\rangle = \mathrm{Tr}[\exp(-\beta H^c_{[a,b]})A]/\mathrm{Tr}(-\beta H^c_{[a,b]})$$

Using Araki's proofs,[4] it has been shown[10] that a state ϕ^c exists on $\mathfrak{A}$ such that

$$\lim_{a\to-\infty,\ b\to+\infty} |\langle \phi^c; A\rangle - \langle \phi^c_{[a,b]}; A\rangle| = 0$$

for all A in $\mathfrak{A}_{\Omega_0}$ and all finite Ω_0. Hence ϕ^c can be interpreted as the Gibbs state corresponding to the modified infinite chain obtained by cutting off the interaction between $c-1$ and c and between d and $d+1$. Moreover, there exists[10] a real constant λ such that $\phi^c \leq \lambda\phi$. Let us now write

$$\mathfrak{A}_S = \mathfrak{A}_{[c,d]}, \qquad \mathfrak{A}_R = \mathfrak{A}_{\mathbb{Z}\setminus[c,d]}.$$

Since $\mathfrak{A}$ is isomorphic[11] to $\mathfrak{A}_S \otimes \mathfrak{A}_R$ and since ϕ^c_R (resp. ϕ^c_S), defined as the restriction of ϕ^c to $\mathfrak{A}_R$ (resp. $\mathfrak{A}_S$), is the Gibbs state of the system $\Sigma(\mathbb{Z}\setminus[c,d])$ (resp. $\Sigma([c,d])$) for the temperature β, we have $\phi^c = \phi^c_S \otimes \phi^c_R$. Since ϕ^c_S is faithful on $\mathfrak{A}_S$ (i.e., $\langle \phi^c_S; A^*A\rangle = 0$ implies $A = 0$) and since $\mathfrak{A}_S$ is finite dimensional, there exists a real constant λ_S depending only on ϕ^c_S such that $\psi_S \leq \lambda_S\phi^c_S$ for all states ψ_S on $\mathfrak{A}_S$. It then follows easily that

$$\psi \equiv \psi_S \otimes \phi^c_R \leq \lambda_S\phi^c \leq \lambda_S\lambda\phi$$

for all ψ_S on $\mathfrak{A}_S$. Since ϕ is invariant with respect to $\{\alpha(t) \mid t \in \mathbb{R}\}$, we have further that, for any invariant mean η on $\mathbb{R}$, the state $\eta\psi$ defined by

$$\langle \eta\psi; A\rangle = \eta\langle \psi; \alpha(\cdot)[A]\rangle$$

also satisfies $\eta\psi \leq \lambda_S\lambda\phi$. In particular, $(\eta\psi)_e \leq \lambda_S\lambda\phi_e$. Since ϕ_e is extremal time invariant, $(\eta\psi)_e = \phi_e$. Now suppose that ψ_S is even on $\mathfrak{A}_S$ and therefore that ψ is even on $\mathfrak{A}$. Since γ commutes with $\{\alpha(t) \mid t \in \mathbb{R}\}$, $\eta\psi$ is also even. We have thus proven that, for any even state ψ_S on $\mathfrak{A}_S$ and any invariant mean η on $\mathbb{R}$,

$$\eta(\psi_S \otimes \phi^c_R) = \phi$$

as states on the whole algebra $\mathfrak{A}$. Incidentally, $\langle \eta(\psi_S \otimes \phi^c_R); A\rangle$ can be computed to be

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T \langle \psi_S \otimes \phi^c_R; \alpha(t)[A]\rangle dt.$$

We emphasize in particular that for any natural temperature β_0, the corresponding canonical equilibrium state ϕ^c_{S,β_0} on $\mathfrak{A}_S$ is even. Consequently, for any natural temperatures β_0 and β, the time average of the initial state $\phi^c_{\beta_0,\beta} \equiv \phi^c_{S,\beta_0} \otimes \phi^c_{R,\beta}$ on the cut system $S+R$ is equal to the canonical equilibrium state ϕ_β of the joint system for the temperature β, independently of β_0. (Since ϕ_β is[4] extremal lattice invariant, the same result using space averages is easy to prove.)

Hence the first ergodic property mentioned in Sec. 1 is given a precise meaning, and is proven to hold for the model considered: The time average (and space average) of $\phi^c_{\beta_0,\beta}$ coincides with ϕ_β for all finite temperatures β_0 and β.

It has been shown[10] that the Gibbs state ψ, corresponding to a situation where $H_{[a,b]}$ is modified by a certain type of perturbation, exists and satisfies $\psi \leq \lambda\phi$. This is in particular the case for the X-Y model and any local, fixed perturbation. If these perturbations are, moreover, even (such as

$$V = -\sum_{i\in[c,d]} B_i\sigma^z_i$$

or

$$\sum_{i\in[c,d]} [(1+\zeta_i)\sigma^x_i\sigma^x_{i+1} + (1-\zeta_i)\sigma^y_i\sigma^y_{i+1}]),$$

and switched off at $t = 0$, then we notice that an argument analogous to that developed above shows that the time average over positive times (as well as the space average) of the corresponding Gibbs states ψ coincides with ϕ.

We now turn to the case where Ω_0 can be taken to be very large (but still finite). Let Ω_1 be any finite interval in $\mathbb{Z}$. We know that on $\mathfrak{A}$, and hence *a fortiori* on $\mathfrak{A}_{\Omega_1}$, ϕ_Ω converges to ϕ in the weak*-topology as Ω tends to infinity. Since Ω_1 is finite, $\mathfrak{A}^*_{\Omega_1}$ is a finite-dimensional linear space, and so its weak*- and norm-topologies coincide. Consequently, given $\epsilon > 0$ and Ω_1, there exists a finite $\Omega_0 \supseteq \Omega_1$ such that for all $\Omega \supseteq \Omega_0$

$$|\langle \eta\psi; A\rangle - \langle \phi_\Omega; A\rangle| \leq \epsilon\,\|A\|$$

for all A in $\mathfrak{A}_{\Omega_1}$ and all states ψ considered above. Hence we can conclude that well inside Ω_0 (namely in Ω_1) $\eta\psi$ is as close as one wants to the canonical equilibrium of $\Sigma(\Omega_0)$ for the temperature β.

4. RELAXATION PROPERTIES

To prove $\eta\psi = \phi$, we used a very much weakened form of the strong asymptotic Abelian character of the action of $\{\alpha_e(t) \mid t \in \mathbf{R}\}$ on $\mathfrak{A}_e$, namely

$$\eta\langle \phi_e; C^*[A, \alpha_e(\cdot)[B]]C\rangle = 0$$

for all A, B and C in $\mathfrak{A}_e$, and for the restriction

ϕ_e of the Gibbs state ϕ to $\mathfrak{A}_e$. The strong asymptotic Abelian character of the evolution is evidently much more stringent a condition (and, as such, is less likely to hold in general). It nevertheless does hold for the X-Y model, so that we can considerably strengthen the results of the previous section. These further results will be obtained by means of a generalization of an argument published by Kastler,[12] which we now describe.

For any representation π of $\mathfrak{A}_e$ in some Hilbert space $\mathfrak{H}$, we denote by $\mathfrak{F}_\pi$ the von Neumann algebra $\pi(\mathfrak{A}_e)'' \cap \pi(\mathfrak{A}_e)'$. Its commutant $\mathfrak{F}_\pi'$ is the von Neumann algebra generated by $\pi(\mathfrak{A}_e)''$ and $\pi(\mathfrak{A}_e)'$, and hence by $\pi(\mathfrak{A}_e)$ and $\pi(\mathfrak{A}_e)'$. Consequently, given any element B in $\mathfrak{F}_\pi'$, any $\epsilon > 0$, and any Ψ_i $(i = 1, 2)$ in $\mathfrak{H}$, there exists $B_0 = \sum_{i=1}^{n} \pi(A_k)B_k$, with A_k in $\mathfrak{A}_e$, B_k in $\pi(\mathfrak{A}_e)'$ and n finite, such that

$$\|(B - B_0)\Psi_i\| \leq \epsilon, \quad i = 1, 2.$$

Using now the strong asymptotic Abelian character of $\{\alpha_e(t) \mid t \in \mathbb{R}\}$ on $\mathfrak{A}_e$, we see that for each A in $\mathfrak{A}_e$ there exists a positive number T such that

$$\|[B_0, \pi(\alpha_e(t)[A])]\| \leq \epsilon$$

for all t with $|t| \geq T$. From these inequalities, we conclude that

$$|(\Psi_1, [B, \pi(\alpha_e(t)[A])]\Psi_2)| \leq \epsilon\,(2\|A\| + 1)\,\|\Psi_1\|\,\|\Psi_2\|,$$

which is to say that, for any B in $\mathfrak{F}_\pi'$ and A in $\mathfrak{A}_e$, $[B, \pi(\alpha_e(t)[A])]$ tends to zero in the weak operator topology as $|t|$ tends to infinity. Let now π be primary and $\|\Psi_1\| = \|\Psi_2\| = 1$. We form the states Ψ_i, $i = 1, 2$, defined on $\mathfrak{A}_e$ by

$$\langle\psi_i; A\rangle = (\Psi_i, \pi(A)\Psi_i)$$

and notice that there exists a unitary operator U in $\mathfrak{B}(\mathfrak{H}) = \mathfrak{F}_\pi'$ such that $\Psi_2 = U\Psi_1$ and hence

$$\langle\psi_1; \alpha_e(t)[A]\rangle - \langle\psi_2; \alpha_e(t)[A]\rangle = (\Psi_2, [U, \pi(\alpha_e(t)[A])]\Psi_1).$$

We conclude from this that for any two vector states ψ_1 and ψ_2 on $\pi(\mathfrak{A}_e)$, which is assumed to be primary, we have

$$\lim_{|t|\to\infty} |\langle\psi_1; \alpha_e(t)[A]\rangle - \langle\psi_2; \alpha_e(t)[A]\rangle| = 0$$

for all A in $\mathfrak{A}_e$.

In particular, if ϕ_e is extremal KMS, if π is the representation of $\mathfrak{A}_e$ associated to ϕ_e (and is thus primary), and if $\psi_e \leq \lambda\phi_e$, then

$$\lim_{|t|\to\infty} \langle\psi_e; \alpha_e(t)[A]\rangle = \langle\phi_e; A\rangle$$

for all A in $\mathfrak{A}_e$.

Thus for the states ψ considered in the preceding section, which are even, we have not only $\eta\psi = \phi$, but actually

$$\lim_{|t|\to\infty} \langle\psi; \alpha(t)[A]\rangle = \langle\phi; A\rangle$$

for all A in $\mathfrak{A}$, i.e., these states actually relax to the canonical equilibrium state ϕ.

5. CONCLUSION

We have provided a positive answer to the question to which we addressed ourselves: Thermal baths can indeed work in the sense set forth in Sec. 1. We also mention that the methods used to prove this result apply to some other situations. Indeed, the essential ingredients are (a) the time-evolution acts in an asymptotically Abelian manner on the even subalgebra $\mathfrak{A}_e$ and (b) the initial state $\phi_{B_0,\beta}$ satisfies the following two properties: (i) $\phi_{B_0,\beta}$ is even, (ii) $\phi^{\varepsilon}_{B_0,\beta} \leq c\phi_\beta$ for some positive number c. These conditions also hold[10] when (a′) a uniform magnetic field B in the z direction is added to the X-Y Hamiltonian and (b′) the initial state $\phi_\beta^{B'}$ is the canonical equilibrium state corresponding to β and $B'(i) = B(i)$ for all but a finite number of sites $i \in \mathbb{Z}$. Hence $\phi_\beta^{B'}(t)$ relaxes to the equilibrium state ϕ_B^β, in agreement with results obtained by Abraham *et al.*[1] and Tjon.[2] The exact solubility of the X-Y model also enabled these authors to analyze in detail the excitation spectrum as the thermodynamic limit is taken; they then use this analysis to compute the *rate* at which equilibrium is reached. The purpose of the present paper was, rather, to emphasize some immediate consequences of general ergodic properties, which we illustrate with the X-Y model.

ACKNOWLEDGMENTS

The authors would like to thank Dr. G. Gallavotti for discussions on his work[1] and for showing to us the preprint of Dr. H. Narnhofer.[8]

* On leave of absence from the Department of Physics and Astronomy of the University of Rochester, Rochester, New York.

1 D. B. Abraham, E. Baruch, G. Gallavotti, and A. Martin-Löf, Phys. Rev. Letters **25**, 1449 (1970).
2 J. A. Tjon, Phys. Rev. B **2**, 2411 (1970).
3 D. R. Robinson, Commun. Math. Phys. **7**, 337 (1968).
4 H. Araki, Commun. Math. Phys. **14**, 120 (1969).
5 R. Haag, N. M. Hugenholtz, and M. Winnink, Commun. Math. Phys. **5**, 215 (1967).
6 O. E. Lanford and D. Ruelle, Commun. Math. **13**, 194 (1969).
7 E. Lieb, T. Schultz, and D. Mattis, Ann. Phys. (N.Y.) **16**, 407 (1961).
8 H. Narnhofer, Preprint IHES, 1970 (unpublished).
9 H. Araki and H. Miyata, Publ. Res. Inst. Math. Sci. Kyoto Univ., Ser. A, **4**, 373 (1968).
10 C. Radin, preprint, Nijmegen, 1971 (unpublished).
11 A. Guichardet, Ann. Sci. Ecole. Norm. Sup. **83**, 1 (1966).
12 D. Kastler, in *Cargèse Lectures in Theoretical Physics*, edited by F. Lurçat (Gordon and Breach, New York, 1967).

On the Martin–Newton Condition for Finite-Dimensional Unitary Matrices

C. Eftimiu
Department of Physics, University of Missouri, St. Louis, Missouri
(Received 11 May, 1971)

A prescription is formulated by means of which one can write conditions for the existence of solutions of a nonlinear system expressing the unitarity of $n \times n$ symmetric matrices. These conditions are less restrictive than those analog to the Martin–Newton condition for the nonlinear integral equation representing the unitarity condition in the elastic region. At least for $n \leq 4$ these conditions are also necessary.

This paper is concerned with the following problem: Given a $n \times n$ unitary and symmetric matrix

$$U = 1 + iT, \tag{1}$$

the elements of the T matrix having the form

$$T_{ij} = T_{ji} = a_{ij} e^{i\varphi_{ij}}, \qquad a_{ij} > 0, \tag{2}$$

under what conditions are the phases ϕ_{ij} determined when the moduli a_{ij} are known? Equivalently, the problem is to find out under what conditions the nonlinear system of equations for the phases ϕ

$$2a_{ij} \sin\phi_{ij} = \sum_{k=1}^{n} a_{ik}a_{jk} e^{i(\varphi_{ik} - \varphi_{jk})},$$
$$j \geq i = 1, 2, \ldots, n, \tag{3}$$

which is an expression of the unitarity condition, has a solution.

This problem may be considered interesting as it stands, and one might think of a number of physical applications in which it is relevant. Our own interest stemmed from a study of the nonlinear integral equation which represents the unitarity condition in the elastic region.[1] Indeed, the system (3) can be viewed as emerging from a discretization of this integral equation.

We shall begin by approaching the problem in a pedestrian manner and consider first the simplest nontrivial case of a 2×2 matrix. Many interesting features of the problem are obscured by the simplicity of this case, but we shall still examine it for the sake of completeness.

The system (3) is then

$$\sin\phi_{11} = (1/2a_{11})(a_{11}^2 + a_{12}^2), \tag{4a}$$

$$\sin\phi_{22} = (1/2a_{22})(a_{12}^2 + a_{22}^2), \tag{4b}$$

$$\sin\phi_{12} = \tfrac{1}{2}\left[a_{11}\cos(\phi_{11} - \phi_{12}) + a_{22}\cos(\phi_{12} - \phi_{22})\right], \tag{4c}$$

$$0 = a_{11}\sin(\phi_{11} - \phi_{12}) + a_{22}\sin(\phi_{12} - \phi_{22}) \tag{4d}$$

The first two equations obviously yield the phases of the diagonal elements, provided the conditions

$$2a_{11} \geq a_{11}^2 + a_{12}^2 \quad \text{or} \quad 1 - (1 - a_{12}^2)^{1/2} \leq a_{11} \leq 1 + (1 - a_{12}^2)^{1/2},$$
$$2a_{22} \geq a_{12}^2 + a_{22}^2 \quad \text{or} \quad 1 - (1 - a_{12}^2)^{1/2} \leq a_{22} \leq 1 + (1 - a_{12}^2)^{1/2}, \tag{5}$$

which are meaningful only if

$$a_{12} \leq 1, \tag{6}$$

are satisfied. The third unknown ϕ_{12} follows from either Eq. (4c),

$$\tan\phi_{12} = \frac{a_{11}\sin\phi_{11} - a_{22}\sin\phi_{22}}{a_{11}\cos\phi_{11} - a_{22}\cos\phi_{22}}, \tag{7}$$

or (4d),

$$\tan\phi_{12} = \frac{a_{11}\cos\phi_{11} + a_{22}\cos\phi_{22}}{2 - a_{11}\sin\phi_{11} - a_{22}\sin\phi_{22}}. \tag{8}$$

The expressions (7) and (8) are in fact identical.

The inequalities (5) and (6) are therefore the necessary and sufficient conditions for the existence of at least one solution of the system (4) (note the trivial ambiguity: if ϕ is a solution set, so is $\pi - \phi$).

Although we will not be concerned specifically with the question of uniqueness, let us remark in passing that if condition (6) were "saturated",

$$a_{12} = 1$$

so that

$$a_{11} = a_{22} = 1$$

as well, there would be an infinity of solutions:

$$\phi_{11} = \phi_{22} = \tfrac{1}{2}\pi, \qquad \phi_{12} = \text{arbitrary},$$

yielding matrices

$$U = \begin{pmatrix} 0 & ie^{i\phi_{12}} \\ ie^{i\phi_{12}} & 0 \end{pmatrix},$$

which are unitary for any real ϕ_{12}.

Let us consider now the more interesting case of 3×3 matrices. The unitary system (3) reads

$$\sin\phi_{11} = (1/2a_{11})(a_{11}^2 + a_{12}^2 + a_{13}^2), \tag{9}$$

$$2a_{12}\sin\phi_{12} = a_{11}a_{12}\cos(\phi_{11} - \phi_{12}) + a_{12}a_{22}\cos(\phi_{12} - \phi_{22}) + a_{13}a_{23}\cos(\phi_{13} - \phi_{23}), \tag{10}$$

$$0 = a_{11}a_{12}\sin(\phi_{11} - \phi_{12}) + a_{12}a_{22}\sin(\phi_{12} - \phi_{22}) + a_{13}a_{23}\sin(\phi_{13} - \phi_{23}), \tag{11}$$

and to each of these three equations we should add two more, obtained through circular permutations of the indices.

Again, we find equations [(9) and its circular permutations] which provide directly the phases of the diagonal elements, if the inequality

$$2a_{12} \geq a_{11}^2 + a_{12}^2 + a_{13}^2 \quad \text{or}$$
$$1 - (1 - a_{12}^2 - a_{13}^2)^{1/2} \leq a_{11} \leq 1 + (1 - a_{12}^2 - a_{13}^2)^{1/2} \tag{12}$$

and its circular permutations are fulfilled. The inequality (12) also implies that the inequality

$$a_{12}^2 + a_{13}^2 \leq 1 \tag{13}$$

and its circular permutations should also be fulfilled, but now these conditions are no longer necessary and sufficient as well for the existence of real phases of nondiagonal elements as solutions of either the system of three equations (10) or the system of three equations (11).

In order to find these conditions, the most direct approach is of course to actually solve either the system (10) or the system (11) (indeed, either of them should yield the same solution) with respect to ϕ_{12}, ϕ_{23}, ϕ_{13}—assuming that ϕ_{11}, ϕ_{22}, ϕ_{33} have already been calculated from (9), subject to conditions (12) and (13)—and read the conditions directly on the explicit expressions of the solution.

This solving can still be done relatively easy for 3×3 matrices if, rather than considering the systems (10) or (11) separately, one chooses to combine them. For instance, if one eliminates the difference $\phi_{13} - \phi_{23}$ between Eqs. (10) and (11), one gets the equation

$$A \cos 2\phi_{12} + B \sin 2\phi_{12} = C, \tag{14}$$

where

$$\begin{aligned} A &= a_{11}a_{22}\cos(\phi_{11} + \phi_{22}) + a_{11}\sin\phi_{11} \\ &\quad + a_{22}\sin\phi_{22} - 1, \\ B &= a_{11}a_{22}\sin(\phi_{11} + \phi_{22}) - a_{11}\cos\phi_{11} \\ &\quad - a_{22}\cos\phi_{22}, \\ C &= (a_{13}^2 a_{23}^2 a_{12}^{-2} - a_{11}^2 - a_{22}^2)/2 \\ &\quad + a_{11}\sin\phi_{11} + a_{22}\sin\phi_{22} - 1. \end{aligned} \tag{15}$$

Equation (14) has the solution

$$\tan\phi_{12} = [B \pm (A^2 + B^2 - C^2)^{1/2}]/(A + C), \tag{16}$$

which is real if, and only if,

$$A^2 + B^2 \geq C^2. \tag{17}$$

This explicitly means that

$$2a_{12}a_{23}a_{13} \geq a_{12}^2 a_{13}^2 + a_{23}^2 a_{12}^2 + a_{13}^2 a_{23}^2. \tag{18}$$

Similar expressions can be obtained for the remaining two phases ϕ_{23} and ϕ_{13} through circular permutations, but it is readily apparent that condition (18) is invariant under such permutations. One should also note that the conditions of the type (13) are all fulfilled if (18) is fulfilled. Indeed, if one writes (18) as

$$a_{13}^2 + a_{23}^2 \leq 2x - x^2,$$

where $x = a_{13}a_{23}/a_{12}$, it is clear that the right-hand side cannot exceed unity for any x.

We conclude, therefore, that the inequalities (12) and (18) are the necessary and sufficient conditions for the existence of a solution of the unitary system for 3×3 matrices.

Finding the existence conditions by actually solving the unitary system is not, however, a procedure that could be followed in general. Already for 4×4 matrices, the algebraic effort is quite serious and, as n increases further, it becomes practically impossible. It is important, therefore, to devise a method by which the existence conditions can be stated without actually solving the system (3). Such a method will be described in the following, and its correctness will be checked against the results obtained for 2×2 and 3×3 matrices.

There are, in fact, quite a number of possible indirect approaches to the problem, but none seems to lead faster and to a better result than the fixed-point theorem approach which was also used by Martin and Newton. For the 3×3 matrices and the system of equations (10), we consider a three-dimensional linear space of vectors $x = (x_1, x_2, x_3)$, endowed with the norm

$$\| x \| \equiv \max_i |x_i|. \tag{19}$$

The system (10) can then be written as

$$x = \mathcal{G}(x) \tag{20}$$

($x_1 = \sin\phi_{12}$, $x_2 = \sin\phi_{23}$, $x_3 = \sin\phi_{13}$) and one can see without difficulty that the nonlinear operator $\mathcal{G}$ maps the ball

$$\| x \| \leq 1 \tag{21}$$

continuously into itself, provided the inequality

$$2a_{12} \geq a_{11}a_{12} + a_{12}a_{22} + a_{13}a_{23}, \tag{22}$$

together with two other inequalities obtained through circular permutations, is fulfilled. It follows then from Brower fixed point theorem that the operator $\mathcal{G}$ has a fixed point in the ball (21), i.e., that the system (10) has a solution.

The inequalities (12) and (22) are the analog of the existence condition derived by Martin and Newton for the nonlinear integral equation. They are clearly not the necessary and sufficient conditions

for the existence of a solution of the unitary system in this case, for we already know that these are (12) and (18) and the latter is by no means equivalent to the conditions (22). Just to illustrate the extent to which they may be different, let us consider the following numerical example:

$$a_{11} = a_{22} = a_{33} = 9/5,$$
$$a_{12} = a_{23} = a_{13} = 3/5\sqrt{2}.$$

The conditions (12) reduce then each to the inequality $1 \geq 1$, but the conditions (22) each read,

$$2 \geq 18/5 + 3/5\sqrt{2} = 4.02. \quad (?)$$

Condition (18) is fulfilled:

$$2 \geq 9/5\sqrt{2} = 1.27$$

so that a solution must exist. In fact, it is not very hard to find it:

$$\phi_{11} = \phi_{22} = \phi_{33} = \pi/2,$$
$$\phi_{12} = \phi_{23} = \phi_{13} + \pi, \quad \sin\phi_{12} = 3/8\sqrt{2}$$

In spite of this inequivalence, there is a very simple way leading from condition (22) to condition (18), and this is based on the observation that condition (18), which is part of the system of necessary and sufficient conditions, is independent of the moduli a_{11}, a_{22}, a_{33}. As a matter of fact, it can be proved quite easily that, if one assumes that conditions (12) are fulfilled and concentrates just on the system (10), the *necessary* and sufficient conditions for the existence of a solution of this system cannot possibly involve the quantities a_{11}, a_{22}, a_{33}. This is simply because if follows from Eqs. (10) that

$$\begin{aligned}
&\frac{\partial\phi_{12}}{\partial a_{11}} = \frac{\partial\phi_{13}}{\partial a_{11}} = \frac{1}{2}\cos\phi_{11}, \quad \frac{\partial\phi_{12}}{\partial a_{33}} = 0,\\
&\frac{\partial\phi_{12}}{\partial a_{22}} = \frac{\partial\phi_{23}}{\partial a_{22}} = \frac{1}{2}\cos\phi_{22}, \quad \frac{\partial\phi_{23}}{\partial a_{11}} = 0,\\
&\frac{\partial\phi_{13}}{\partial a_{33}} = \frac{\partial\phi_{23}}{\partial a_{33}} = \frac{1}{2}\cos\phi_{33}, \quad \frac{\partial\phi_{13}}{\partial a_{22}} = 0.
\end{aligned} \quad (23)$$

Hence, if the system (10) had a solution for fixed a_{12}, a_{23}, a_{13} and some a_{11}, a_{22}, a_{33} with values within the ranges indicated in (12), it should also have a solution for the same values of a_{12}, a_{23}, a_{13}, but other values of a_{11}, a_{22}, a_{33} from some open neighborhood of the former values which is still within the intervals (12). By continuity, one can thus cover the whole range of values allowed by (12) for a_{11}, a_{22}, a_{33} for fixed a_{12}, a_{23}, a_{13}.

Consequently, even if we did not know what the necessary and sufficient conditions really are, the mere fact that (22) involves the quantities a_{11}, a_{22}, a_{33} would be a sure indication that they can only be a sufficient condition for the existence of a solution of (10). However, if for fixed a_{12}, a_{23}, a_{12} one lets a_{11}, a_{22}, a_{33} in (22) vary within the limits given in (12), one obtains a class of sufficient conditions for a corresponding class of unitary systems, the least restrictive of which being those obtained by giving a_{11}, a_{22}, a_{33} their minimum values allowed by (12). When one makes these substitutions, each of the conditions (22) becomes identical to condition (18). According to our observation then, by guaranteeing the existence of a solution of the unitary system in which, for given a_{12}, a_{23}, a_{13}, the moduli a_{11}, a_{22}, a_{33} have the minimum values compatible with (12), condition (18)—together with (12)—guarantees at the same time the existence of a solution of any other unitary system involving the same a_{12}, a_{23}, a_{13} but any other choice of values of a_{11}, a_{22}, a_{33} allowed by (12).

It appears, therefore, that we can formulate a prescription by means of which one could write the necessary and sufficient conditions for the existence of a solution of the unitary system (3) and which does not require the knowledge of the solution. As such, it can be stated for any n.

One starts by establishing the Martin–Newton conditions for the system (3), by using essentially the same argument based on Brower's fixed-point theorem:

$$2a_{ij} \geq \sum_{k=1}^{n} a_{ik}a_{jk}, \quad i,j = 1,2,\ldots,n, \quad (24)$$

and one splits it into two parts

$$2a_{ii} \geq \sum_{k=1}^{n} a_{ik}^2, \quad i = 1,2,\ldots,n, \quad (25)$$

$$2a_{ij} \geq \sum_{k=1}^{n} a_{ik}a_{jk}, \quad i < j. \quad (26)$$

The n conditions (25) are already a component of the system of necessary and sufficient conditions for the system (3), and they can be used to establish the minimum values of a_{ii} for given $a_{ik}, i \neq k$:

$$a_{ii} \geq 1 - \Big(1 - \sum_{k}{}' a_{ik}^2\Big)^{1/2}, \quad (27)$$

where the prime indicates that the term with $k = i$ should be omitted.

It can now again be checked that, for fixed a_{ik} $(i \neq k)$, a solution exists in any neighborhood of those values of a_{ii} for which a solution is known to exist, provided conditions (25) are fulfilled. One substitutes then the minimum values of a_{ii} as given by (27) into the $\frac{1}{2}n(n-1)$ conditions (26), and one gets $(\alpha \neq \beta)$

$$\begin{aligned}
&4a_{\alpha\beta}^{-2}\Big(\sum_{k}{}'' a_{\alpha k}a_{\beta k}\Big)^2 + 4\Big(\sum_{j}{}' a_{\alpha j}^2\Big)\Big(\sum_{l}{}' a_{\beta l}^2\Big)\\
&\quad \geq \Big[a_{\alpha\beta}^{-2}\Big(\sum_{k} a_{\alpha k}a_{\beta k}\Big)^2 + \sum_{j} a_{\alpha j}^2 + \sum_{l} a_{\beta l}^2\Big]^2,
\end{aligned} \quad (28)$$

where the double prime indicates that the terms in the sum over k with $k = \alpha$ or $k = \beta$ should be omitted.

It is straightforward and easy to check that for 2×2 and 3×3 matrices conditions (28) coincide with (6) and (18), respectively. It is also straightforward but less than easy to check that (28), and also (25), are also necessary and sufficient conditions for the existence of a solution of the system (3) for 4×4 matrices. (Because it involves extremely tedious calculations, the case of 4×4 matrices is not discussed here.) We are aware though that there is no substitute for rigorous proof and that we failed to produce such a proof for $n > 4$. But even if (25) and (28) were not necessary conditions for the existence of a solution of (3), they clearly are sufficient and significantly less restrictive than the Martin–Newton condition. Because of this, and particularly in view of the fact that it has been pointed out[2] that the Martin–Newton condition is not fulfilled in a number of physical cases, it would be interesting to extend the results given in this paper to infinite matrices in general and specifically to the non-linear integral equation of Ref. 1.

---

[1] R. G. Newton, J. Math. Phys. **9**, 2050 (1968); A. Martin, Nuovo Cimento **59A**, 131 (1969).

[2] I. A. Sakmar, Lett. Nuovo Cimento **2**, 256 (1969); H. Goldberg, Phys. Rev. D **1**, 1242 (1970).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 10 OCTOBER 1971

A Convergent Expansion for the Resolvent of: $\varphi^4 :_{1+1}$*

Paul Federbush
Department of Mathematics, University of Michigan, Ann Arbor, Michigan 48104
(Received 12 April 1971)

For the model theory, $N + \lambda \int :\varphi^4 : dx$ in a box, a convergent expansion of the resolvent is exhibited. This also provides another proof of boundedness below for the model.

We consider the field theory model with Hamiltonian $N + \int : \varphi^4 : dx$ in a box of length 1. We obtain a convergent expansion for the resolvent of this model and at the same time another proof of boundedness below of the Hamiltonian. The main idea is to consider the 'subblocks' of the Hamiltonian obtained by restricting to states with particle number spectrum lying between N and $2N$. The resolvent of a subblock is shown to have very small matrix elements connecting states with a large difference in particle number. The extension of the present results to the more general model $N_\tau + \lambda \int : \varphi^{2s} : dx$ has not yet been achieved.

We begin with the basic theorem to be used.

Theorem: Let A be a positive self-adjoint operator of norm $\leq M$, and $|\alpha\rangle$ and $|\beta\rangle$ be two vectors of unit length. Suppose $\langle \alpha | A^k | \beta \rangle = 0, 0 \leq k \leq N$. Then, for any $\lambda > 0$, a real number,

$$|\left\langle \alpha \left| \frac{1}{\lambda + A} \right| \beta \right\rangle| \leq \frac{4\sqrt{M}}{\lambda\sqrt{2\lambda}} \cdot \left(\frac{1}{1 + \sqrt{2\lambda/M}} \right)^N \tag{1}$$

$$\sim \exp(-\sqrt{2\lambda/M}\, N), \tag{2}$$

where in (2) it is assumed M and N are large.

Proof: $(\lambda + A)^{-1}$ and $(\lambda + A)^{-1} - P_N(A)$ have the same matrix elements between $|\alpha\rangle$ and $|\beta\rangle$, where $P_N(x)$ is any polynomial of degree N. This implies the matrix element is smaller than the supremum of $|(\lambda + x)^{-1} - P_N(x)|$ for values of x in the spectrum of A. We make a linear change of variables moving the spectrum from $[0, M]$ to $[-1, 1]$. Now one has the function $[\lambda + (x + 1)\frac{1}{2}M]^{-1}$ on the interval $[-1, 1]$. There is a basic theorem[1] in the theory of polynomial approximation stating that if f is analytic in an ellipse with foci at -1 and 1 and major and minor radii a and b, then it may be approximated on $[-1, 1]$ by a polynomial of degree N within

$$\frac{2f \max}{(a + b - 1)} \left(\frac{1}{a + b} \right)^N \tag{3}$$

in the uniform norm. Here f max is the supremum of the absolute value of f in the ellipse. The theorem is obtained applying this result to the ellipse with $a = 1 + \lambda/M$ and $f = [\lambda + (x + 1)\frac{1}{2}M]^{-1}$.

We now come to the Hamiltonian

$$H = N + \int_0^1 : \varphi^4 : dx = N + V. \tag{4}$$

We define P_i as the projection operator onto states with numbers of particles lying in the range

$$2^i \leq N < 2^{i+2}, \quad i = -1, 0, \cdots, \tag{5}$$

and P_e and P_d as the projection operator onto states with numbers of particles in the ranges

$$\bigcup_{i \text{ even}} (2^i - 4 \leq N \leq 2^i + 4) \tag{6}$$

and

$$\bigcup_{i \text{ odd}} (2^i - 4 \leq N \leq 2^i + 4), \tag{7}$$

respectively. We define

where the double prime indicates that the terms in the sum over k with $k = \alpha$ or $k = \beta$ should be omitted.

It is straightforward and easy to check that for 2×2 and 3×3 matrices conditions (28) coincide with (6) and (18), respectively. It is also straightforward but less than easy to check that (28), and also (25), are also necessary and sufficient conditions for the existence of a solution of the system (3) for 4×4 matrices. (Because it involves extremely tedious calculations, the case of 4×4 matrices is not discussed here.) We are aware though that there is no substitute for rigorous proof and that we failed to produce such a proof for $n > 4$. But even if (25) and (28) were not necessary conditions for the existence of a solution of (3), they clearly are sufficient and significantly less restrictive than the Martin–Newton condition. Because of this, and particularly in view of the fact that it has been pointed out[2] that the Martin–Newton condition is not fulfilled in a number of physical cases, it would be interesting to extend the results given in this paper to infinite matrices in general and specifically to the non-linear integral equation of Ref. 1.

---

[1] R. G. Newton, J. Math. Phys. **9**, 2050 (1968); A. Martin, Nuovo Cimento **59A**, 131 (1969).

[2] I. A. Sakmar, Lett. Nuovo Cimento **2**, 256 (1969); H. Goldberg, Phys. Rev. D **1**, 1242 (1970).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 10 OCTOBER 1971

A Convergent Expansion for the Resolvent of: $\varphi^4 :_{1+1}$*

Paul Federbush
Department of Mathematics, University of Michigan, Ann Arbor, Michigan 48104
(Received 12 April 1971)

For the model theory, $N + \lambda \int :\varphi^4 : dx$ in a box, a convergent expansion of the resolvent is exhibited. This also provides another proof of boundedness below for the model.

We consider the field theory model with Hamiltonian $N + \int : \varphi^4 : dx$ in a box of length 1. We obtain a convergent expansion for the resolvent of this model and at the same time another proof of boundedness below of the Hamiltonian. The main idea is to consider the 'subblocks' of the Hamiltonian obtained by restricting to states with particle number spectrum lying between N and $2N$. The resolvent of a subblock is shown to have very small matrix elements connecting states with a large difference in particle number. The extension of the present results to the more general model $N_\tau + \lambda \int : \varphi^{2s} : dx$ has not yet been achieved.

We begin with the basic theorem to be used.

Theorem: Let A be a positive self-adjoint operator of norm $\leq M$, and $|\alpha\rangle$ and $|\beta\rangle$ be two vectors of unit length. Suppose $\langle \alpha | A^k | \beta \rangle = 0, 0 \leq k \leq N$. Then, for any $\lambda > 0$, a real number,

$$|\left\langle \alpha \left| \frac{1}{\lambda + A} \right| \beta \right\rangle| \leq \frac{4\sqrt{M}}{\lambda\sqrt{2\lambda}} \cdot \left(\frac{1}{1 + \sqrt{2\lambda/M}} \right)^N \tag{1}$$

$$\sim \exp(-\sqrt{2\lambda/M}\, N), \tag{2}$$

where in (2) it is assumed M and N are large.

Proof: $(\lambda + A)^{-1}$ and $(\lambda + A)^{-1} - P_N(A)$ have the same matrix elements between $|\alpha\rangle$ and $|\beta\rangle$, where $P_N(x)$ is any polynomial of degree N. This implies the matrix element is smaller than the supremum of $|(\lambda + x)^{-1} - P_N(x)|$ for values of x in the spectrum of A. We make a linear change of variables moving the spectrum from $[0, M]$ to $[-1, 1]$. Now one has the function $[\lambda + (x + 1)\frac{1}{2}M]^{-1}$ on the interval $[-1, 1]$. There is a basic theorem[1] in the theory of polynomial approximation stating that if f is analytic in an ellipse with foci at -1 and 1 and major and minor radii a and b, then it may be approximated on $[-1, 1]$ by a polynomial of degree N within

$$\frac{2f \max}{(a + b - 1)} \left(\frac{1}{a + b} \right)^N \tag{3}$$

in the uniform norm. Here f max is the supremum of the absolute value of f in the ellipse. The theorem is obtained applying this result to the ellipse with $a = 1 + \lambda/M$ and $f = [\lambda + (x + 1)\frac{1}{2}M]^{-1}$.

We now come to the Hamiltonian

$$H = N + \int_0^1 : \varphi^4 : dx = N + V. \tag{4}$$

We define P_i as the projection operator onto states with numbers of particles lying in the range

$$2^i \leq N < 2^{i+2}, \quad i = -1, 0, \cdots, \tag{5}$$

and P_e and P_d as the projection operator onto states with numbers of particles in the ranges

$$\bigcup_{i \text{ even}} (2^i - 4 \leq N \leq 2^i + 4) \tag{6}$$

and

$$\bigcup_{i \text{ odd}} (2^i - 4 \leq N \leq 2^i + 4), \tag{7}$$

respectively. We define

$$H_i = P_i H P_i, \tag{8}$$

$$H_e = \sum_{i \text{ even}} H_i, \tag{9}$$

$$H_d = \sum_{i \text{ odd}} H_i, \tag{10}$$

$$H = H_e + L_e = H_d + L_d. \tag{11}$$

We note

$$P_e L_e P_e = L_e P_e = P_e L_e = L_e, \tag{12}$$

$$P_d L_d P_d = L_d P_d = P_d L_d = L_d. \tag{13}$$

The expansion of the resolvent we are after is the following:

$$\begin{aligned}\frac{1}{E+H} &= \frac{1}{E+H_e} - \frac{1}{E+H_d} L_e \frac{1}{E+H_e} \\ &\quad + \frac{1}{E+H_e} L_d \frac{1}{E+H_d} L_e \frac{1}{E+H_e} - \cdots \\ &= \frac{1}{E+H_e} - \frac{1}{E+H_d} P_e L_e P_e \frac{1}{E+H_e} \\ &\quad + \frac{1}{E+H_e} P_d L_d P_d \frac{1}{E+H_d} P_e L_e P_e \frac{1}{E+H_e} \cdots .\end{aligned} \tag{14}$$

This expansion converges for E large enough, as we will show; for E large enough $(E + H_e)^{-1}$ and $(E + H_d)^{-1}$ exist and are bounded, $|L_d P_d (E + H_d)^{-1} P_e| < \frac{1}{2}$ and $|L_e P_e (E + H_e)^{-1} P_d| < \frac{1}{2}$, and $(E + H_d)^{-1} L_e (E + H_e)^{-1}$ and $L_d (E + H_d)^{-1} L_e (E + H_e)^{-1}$ are bounded.

The following two estimates easily yield the required relations above.

Estimate 1:

$$H_i \geqslant 2^{i-1} P_i, \quad i \text{ large} \tag{15}$$

Estimate 2:

$$\left| P_e \frac{P_i}{E + H_i} P_d \right| \leqslant c_1 \exp(-c_2 2^{i/2}) \quad \text{for some } c_1, c_2 > 0. \tag{16}$$

Proof of estimate 1: We write V as the sum $V_k + R_k$, where as usual V_k contains those terms in the expansion of V all of whose momenta are less than or equal k in absolute value

$$H_i = P_i N P_i + P_i V_{k_i} P_i + P_i R_{k_i} P_i. \tag{17}$$

We first note

$$P_i N P_i \geqslant 2^i P_i. \tag{18}$$

As in Ref. 2, one has

$$V_k \geqslant -c(\ln k)^2. \tag{19}$$

Picking

$$k_i = \exp[(1/\sqrt{c}) 2^{(i-1)/2}], \tag{20}$$

we get

$$P_i V_{k_i} P_i \geqslant -2^{i-1} P_i. \tag{21}$$

From

$$|P_i R_{k_i} P_i| \leqslant d(2^{i+2}) 2 1/\sqrt{k_i}, \tag{22}$$

a standard N_τ estimate, see Ref. 3, we quickly get

$$|P_i R_{k_i} P_i| \leqslant 2^{i-1}, \quad i \text{ large}. \tag{23}$$

And thus, using (23), (21), and (18), we obtain estimate 1 from (17).

Proof of Estimate 2: Clearly

$$|E + H_i| \leqslant e 2^{2i} \tag{24}$$

for some e (by an N_τ estimate again), and

$$E + H_i \geqslant 2^{i-1} \tag{25}$$

for E large enough. We apply the theorem with $|\alpha\rangle = P_i P_d |a\rangle$, $|\beta\rangle = P_i P_e |b\rangle$ ($|a\rangle$ and $|b\rangle$ normalized vectors, $\lambda = 2^{i-1}$, $M = e 2^{2i}$, $A = E + H_i - 2^{i-1}$, and $N < [(2^{i+1} - 4) - (2^i + 4)]/4$. This approach is easily generalized to $N_\tau + \int :\varphi^4: dx$.

The subject of obtaining convergent expansions for the resolvents of other field theory models seems interesting, as is the question of whether this is a way of obtaining lower bound estimates for other models.

---

* This work was supported in part by NSF Grant GP-17523.

1 A. F. Timan, *Theory of Approximation of Functions of a Real Variable* (MacMillan, New York, 1963), p. 281.

2 E. Nelson, "A Quartic Interaction in Two Dimensions," in *Mathematical Theory of Elementary Particles*, edited by R. Goodman and I. Segal (MIT Press, Cambridge, Mass., 1966).

3 J. Glimm, Commun. Math. Phys. 5, 343 (1967).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 10 OCTOBER 1971

Nearly Symmetric Advection

A. M. Soward*
Cooperative Institute for Research in Enviromental Sciences, University of Colorado Boulder, Colorado 80302
(Received 27 January 1971)

A class of exact solutions of the heat conduction equation and the magnetic induction equation are constructed for the special case where there is no diffusivity or time dependence. It is shown how these elementary solutions may facilitate the solution of a more general class of problems in which diffusivity and time dependence may also be present. Moreover, the existence of effective variables, first introduced by Braginskii for the kinematic dynamo problem in the lowest-order approximation, is a direct consequence of these elementary solutions.

I. INTRODUCTION

During recent years considerable interest has been shown in nearly symmetric dynamos.[1-6] Braginskii[1] was able to show that, in an asymptotic analysis, a remarkably elegant pair of coupled equations [essentially (63) and (64)] could be obtained to describe the axisymmetric part of the magnetic field by the introduction of effective variables. The resulting equations may be solved relatively easily for certain boundary conditions, and Braginskii[3] was thus able to demonstrate the existence of nearly symmetric kinematic dynamos. Tough[4] extended Braginskii's analysis and showed that, by modifying the definition of effective variables, Braginskii's equations for the symmetric part of the magnetic field are still valid in the second approximation. In a study of the hydromagnetic dynamo, Tough and Roberts[6] showed that in the first approximation these quantities make similar simplifications to the equation of motion. In the previous paper,[7] henceforth referred to as Paper I, it was established that the equations for the axisymmetric motion obtained by Tough and Roberts[6] are still valid in the second approximation, provided that the effective variables obtained by Tough[4] are used. Since effective variables play such an important role in nearly symmetric dynamo theory, it is of considerable interest to know precisely how they come about. Tough[4] suggested that effective variables may be relevant to all orders of approximation, i.e., an effective velocity, magnetic field, and Γ [see (64)] can be defined such that Eqs. (63) and (64) are correct to arbitrary order. By considering the heat conduction equation where effective variables are also relevant in the lowest two approximations, it was shown in Paper I that effective variables do not exist which make simplifications of the above type. Consequently, since the problem had several points in common with the magnetic induction equation, it was concluded that effective variables would not be relevant in the third approximation and that the suggestion made by Tough is unlikely. In this paper a considerably weaker result than the one proposed by Tough is obtained. However, the result does indicate why considerable simplifications can be made in the low-order approximations by the introduction of effective variables.

The primary concern of this paper is in obtaining a particular class of solutions of the heat conduction equation

$$R\mathbf{u}\cdot\nabla\theta = 0, \quad \nabla\cdot\mathbf{u} = 0, \tag{1}$$

and the magnetic induction equation

$$O = R\nabla\times(\mathbf{u}\times\mathbf{b}), \quad \nabla\cdot\mathbf{b} = 0, \quad \nabla\cdot\mathbf{u} = 0, \tag{2}$$

where the velocity $\mathbf{u}$ is given by

$$\mathbf{u} = U(\rho, z)\mathbf{i}_\phi + R^{-1/2}\mathbf{u}'(\rho, \phi, z) + R^{-1}\mathbf{u}_p(R, \rho, z) \tag{3}$$

and the constant R is large,

$$R \gg 1. \tag{4}$$

Here ρ, ϕ, z are cylindrical polar coordinates: ρ is the distance from the axis, ϕ is the azimuthal angle ($\mathbf{i}_\phi$ is the unit vector in the ϕ direction), and z is the distance along the axis. It is assumed throughout that primed quantities have zero ϕ average:

$$\langle \mathbf{f}' \rangle = 0, \tag{5}$$

where

$$\langle \mathbf{f} \rangle = \langle f_\rho \rangle \mathbf{i}_\rho + \langle f_\phi \rangle \mathbf{i}_\phi + \langle f_z \rangle \mathbf{i}_z, \langle f \rangle = \frac{1}{2\pi}\int_0^{2\pi} f\, d\phi, \tag{6}$$

and $\mathbf{i}_\rho$ and $\mathbf{i}_z$ are the unit vectors in the ρ and z directions. Solutions of (1) and (2) are sought in the form

$$\theta = \Theta(\rho, z) + R^{-1/2}\theta'(R, \rho, \phi, z), \tag{7}$$

$$\mathbf{b} = B(\rho, z)\mathbf{i}_\phi + R^{-1/2}\mathbf{b}'(R, \rho, \phi, z) + R^{-1}\mathbf{b}_p(R, \rho, z). \tag{8}$$

Unlike most previous work in this field, there is no restriction on the vectors $\mathbf{u}_p$ and $\mathbf{b}_p$ lying in meridional planes.

Much of the notation and procedure is similar to Paper I. However, in order that this paper should be self-contained, the relevant definitions are restated. The operator $\partial_1/\partial\phi$ is defined not to differentiate unit vectors:

$$\frac{\partial_1}{\partial\phi}\mathbf{f} = \frac{\partial f_\rho}{\partial\phi}\mathbf{i}_\rho + \frac{\partial f_\phi}{\partial\phi}\mathbf{i}_\phi + \frac{\partial f_z}{\partial z}\mathbf{i}_z. \tag{9}$$

The operator $\hat{}$ is defined by

$$\frac{\partial_1}{\partial\phi}\mathbf{f}' = \mathbf{f}' \quad \text{where} \quad \langle \mathbf{f}' \rangle = 0, \tag{10}$$

and results in the identity

$$\sum_{j=1}^{N}\langle \hat{f}_1' \cdots \hat{f}_{j-1}' f_j' \hat{f}_{j+1}' \cdots \hat{f}_N' \rangle = 0. \tag{11}$$

For quantities which do not have a vanishing ϕ average, the bar operator is introduced where

$$\bar{f} = \widehat{(f - \langle f \rangle)}. \tag{12}$$

Some new notation is introduced. The operators and variables $\partial_i, \chi^{(i)}, u_i'$ are defined by

$$\begin{aligned} \{\partial_i\} &= \left\{\frac{1}{\rho}\frac{\partial}{\partial\rho}, \frac{\partial}{\partial\phi}, \frac{\partial}{\partial z}\right\}, \\ \{\chi^{(i)}\} &= \left\{\rho, \frac{1}{\rho}, 1\right\}, \\ \{u_i'\} &= \{u_\rho', u_\phi', u_z'\}. \end{aligned} \tag{13}$$

Since $\nabla \cdot \mathbf{u}' = 0$, it follows that

$$\partial_i \chi^{(i)} u_i' = 0, \tag{14}$$

where the summation is over repeated *suffices*. Of course, the bracketed superscript takes part in the summation, but a product of the type $\chi^{(i)} u_i'$ is just one term and not the sum of three. With this summation convention, the operator $\mathbf{u}' \cdot \nabla$ is alternatively

$$\mathbf{u}' \cdot \nabla = u_i' \chi^{(i)} \partial_i. \tag{15}$$

The operator ∂_i is such that products commute:

$$\partial_i \partial_j = \partial_j \partial_i. \tag{16}$$

In order to construct a certain class of solutions of (1) and (2), two sets of vectors $\mathbf{v}^{(n)}, \mathbf{f}^{(n)}$ and a set of scalars $\psi^{(n)}$ must be introduced. Their significance will become apparent later in the section. The set of vectors $\mathbf{v}^{(n)}$, $n = 1, 2, \cdots$, is defined by their components in cylindrical polar coordinates, namely

$$\begin{aligned} \chi^{(i)} v_i^{(1)} &= \mathbf{u}' \cdot \nabla \left(\frac{\rho}{U} \chi^{(i)} \hat{u}_i' \right), \\ \chi^{(i)} v_i^{(n)} &= \mathbf{u}' \cdot \nabla \left(\frac{\rho}{U} \chi^{(i)} \overline{v_i^{(n-1)}} \right) \\ &\quad - \sum_{r=1}^{n-2} \mathbf{V}^{(r)} \cdot \nabla \left(\frac{\rho}{U} \chi^{(i)} \overline{v_i^{(n-r-1)}} \right) - \mathbf{V}^{(n-1)} \cdot \nabla \left(\frac{\rho}{U} \chi^{(i)} \hat{u}_i' \right), \\ &\quad n \geq 2, \end{aligned} \tag{17}$$

where

$$\mathbf{V}^{(n)} = \langle \mathbf{v}^{(n)} \rangle,$$

and for the case $n = 2$ the meaningless summation $\sum_{r=1}^{0}$ is omitted. The omission is also made in the next two definitions. The set of scalars $\psi^{(n)}, n = 1, 2, \cdots$, is defined by

$$\begin{aligned} \psi^{(1)} &= \mathbf{u}' \cdot \nabla \Theta, \\ \psi^{(n)} &= \mathbf{u}' \cdot \nabla \frac{\rho}{U} \overline{\psi^{(n-1)}} - \sum_{r=1}^{n-2} \mathbf{V}^{(r)} \cdot \nabla \frac{\rho}{U} \overline{\psi^{(n-r-1)}} \\ &\quad - \mathbf{V}^{(n-1)} \cdot \nabla \Theta, \quad n \geq 2. \end{aligned} \tag{18}$$

The set of vectors $\mathbf{f}^{(n)}$, $n = 0, 1, 2, \cdots$, is defined by

$$\begin{aligned} \mathbf{f}^{(o)} &= \mathbf{u}' \Theta, \\ \mathbf{f}^{(n)} &= \mathbf{u}' \frac{\rho}{U} \overline{\psi^{(n)}} - \sum_{r=1}^{n-1} \mathbf{V}^{(r)} \frac{\rho}{U} \overline{\psi^{(n-r)}} - \mathbf{V}^{(n)} \Theta, \quad n \geq 1, \\ \mathbf{F}^{(n)} &= \langle \mathbf{f}^{(n)} \rangle. \end{aligned} \tag{19}$$

A simple result of the definitions (18) and (19) is now established. From (19) it follows that

$$\mathbf{u}' \times \mathbf{f}^{(n)} = -\sum_{r=1}^{n-1} \mathbf{u}' \times \mathbf{V}^{(r)} \frac{\rho}{U} \overline{\psi^{(n-r)}} - \mathbf{u}' \times \mathbf{V}^{(n)} \Theta$$

and

$$\begin{aligned} \mathbf{V}^{(r)} \times \mathbf{f}^{(n-r)} &= -\mathbf{u}' \times \mathbf{V}^{(r)} \frac{\rho}{U} \overline{\psi^{(n-r)}} \\ &\quad - \sum_{s=1}^{n-r-1} \mathbf{V}^{(r)} \times \mathbf{V}^{(s)} \frac{\rho}{U} \overline{\psi^{(n-r-s)}} - \mathbf{V}^{(r)} \times \mathbf{V}^{(n-r)} \Theta. \end{aligned}$$

Combining these results leads to the identity

$$\mathbf{u}' \times \mathbf{f}^{(n)} = \sum_{r=1}^{n} \mathbf{V}^{(r)} \times \mathbf{f}^{(n-r)}. \tag{20}$$

Finally, the vectors $\mathbf{V}, \mathbf{f}, \mathbf{F}$ and the scalar ψ are defined by

$$\mathbf{V} = \sum_{n=1}^{\infty} (-1)^{n-1} R^{1/2-n/2} \mathbf{V}^{(n)} \tag{21a}$$

$$\mathbf{f} = \sum_{n=0}^{\infty} (-1)^n R^{1/2-n/2} \mathbf{f}^{(n)}, \quad \mathbf{F} = \langle \mathbf{f} \rangle, \tag{21b}$$

$$\psi = \frac{\rho}{U} \sum_{n=1}^{\infty} (-1)^n R^{1/2-n/2} \overline{\psi^{(n)}}. \tag{21c}$$

Of course, these quantities may not exist as the series may diverge. In the subsequent analysis, the existence of the quantities is assumed, and no attempt is made to determine the precise conditions for existence. However, it is to be expected that, provided that U is bounded away from zero ($|U| \geq \delta > 0$) and $\Theta, U, \mathbf{u}'$ have continuous derivatives to all orders, $\mathbf{V}, \mathbf{f}, \psi$ are well defined for R sufficiently large. Even if the series (21) diverge, the summation of a finite number of terms is likely to have a useful meaning as an asymptotic expansion.

The principal results of this paper are now summarized as three theorems as follows:

Theorem 1: The vector $\mathbf{V}^{(n)}$ has zero divergence,

$$\nabla\cdot\mathbf{V}^{(n)} = 0 \tag{22}$$

and the ϕ average of $\psi^{(n)}$ is zero,

$$\langle\psi^{(n)}\rangle = 0. \tag{23}$$

This theorem has immediate consequences. Evidently, from the definitions of $\mathbf{f}^{(n)}$ and $\mathbf{F}^{(n)}$, it follows that

$$\nabla\cdot\mathbf{f}^{(n)} = \psi^{(n+1)} \tag{24}$$

and

$$\nabla\cdot\mathbf{F}^{(n)} = 0. \tag{25}$$

Despite the innocent appearance of Theorem 1, its proof is somewhat lengthy and so has been relegated to the Appendix. However, Theorem 1 is important since it is required in the proofs of Theorems 2 and 3 and also shows that both V and F have zero divergence. The two main theorems are the following.

Theorem 2: For the particular case where

$$\mathbf{u}_p = \mathbf{V},$$

the general solution of (1), taking the form (7), is

$$\theta = \Theta + R^{-1/2}\psi, \tag{26}$$

where Θ is an arbitrary function of ρ and z.

Theorem 3: For the particular case where

$$\mathbf{u}_p = \mathbf{V},$$

a class of solutions of (2), taking the form (8), is

$$\mathbf{b} = (\Theta + R^{-1/2}\psi)\mathbf{u}, \tag{27}$$

where

$$B = \Theta U$$

and Θ is an arbitrary function of ρ and z.

Noting that $\nabla\cdot\mathbf{b} = 0$, we see that Theorem 3 is a trivial consequence of Theorem 2. Moreover, when $\mathbf{u}_p = \mathbf{V}$, (27) gives the general solution of (2) for magnetic fields aligned to the flow. Of course, there may be other solutions for which u and b are not parallel. Finally, with the help of the definitions (18), (19) and (21), it can be shown that (27) is equivalent to

$$\mathbf{b} = (B + R^{-1/2}U\psi)\mathbf{i}_\phi + R^{-1}\mathbf{f}. \tag{28}$$

The solutions of the heat conduction and magnetic induction equations given by Theorems 2 and 3, namely (26) and (27), are called elementary solutions.

Evidently (1) and (2) may be solved by elementary methods, e.g., (1) has the solution θ is constant on streamlines. However, the significant feature of Theorem 2 is that, with the choice $\mathbf{u}_p = \mathbf{V}$, the function $\Theta(\rho, z)$ which generates the solution (26) is arbitrary. By a simple argument it is now shown that this property implies that the velocity $\mathbf{u}_0 = U\mathbf{i}_\phi + R^{-1/2}\mathbf{u}' + R^{-1}\mathbf{V}$ describes closed streamlines which tend to the circles $\rho = \text{const}$, $z = \text{const}$ as $R \to \infty$. Consider a plane $\phi = \phi_0 (= \text{const})$. Suppose a streamline intersects the plane at $\mathbf{x}_0$ and intersects the plane again at $\mathbf{x}_1$ ($\phi = \phi_0 + 2\pi$). Since the meridional velocities are at most order $R^{-1/2}$, the distance between the points is also order $R^{-1/2}$. Thus the position vector $\mathbf{x}_1$ is defined by $\mathbf{x}_1 = \mathbf{x}_0 + R^{-1/2}\eta$. It is now assumed that $\theta(\mathbf{x}_1)$ may be expanded as a Taylor series at the point $\mathbf{x}_0$. This assumption is in accord with the rather general assumptions necessary for the validity of Theorems 1–3. Now since θ is constant on streamlines, it follows by Theorem 2 that

$$\Theta(\mathbf{x}_1) - \Theta(\mathbf{x}_0) = -R^{-1/2}[\psi(\mathbf{x}_1) - \psi(\mathbf{x}_0)]$$

and hence that

$$[\boldsymbol{\eta}\cdot\nabla\Theta](\mathbf{x}_0) + O(R^{-1/2}\,|\,\eta\,|^2) = R^{-1/2}[\boldsymbol{\eta}\cdot\nabla\psi](\mathbf{x}_0) + O(R^{-1}\,|\,\eta\,|^2).$$

Clearly, since Θ is an arbitrary function of ρ and z, and $R \gg 1$, the only value of η (small compared to $R^{1/2}$) satisfying the above equation is

$$\eta = 0$$

and consequently the points $\mathbf{x}_0$ and $\mathbf{x}_1$ are coincident. In other words, according to Theorem 2, given an order 1 axisymmetric azimuthal velocity $U\mathbf{i}_\phi$ and an order $R^{-1/2}$ asymmetric velocity $R^{-1/2}\mathbf{u}'$, the order R^{-1} axisymmetric velocity $R^{-1}\mathbf{V}$ constructed has the property that the streamlines of $\mathbf{u}_0$ are closed: All velocities have zero divergence. Moreover, the magnetic field $\mathbf{b}_0 = (B + R^{-1/2}U\psi)\mathbf{i} + R^{-1}\mathbf{f}$, given by (28), describes closed field lines since it is aligned to the flow $\mathbf{u}_0$.

Theorem 2 is proved in Sec. II. Then, with the help of the theorem, a procedure for solving the full heat conduction equation (37) by successive approximation to all orders is outlined. Since it is natural to introduce the "effective" velocity vector

$$\mathbf{u}_{ep} = \mathbf{u}_p - \mathbf{V} \tag{29}$$

into the approximation scheme, the procedure provides a framework for discussing the significance of an effective velocity. In a similar way, Theorem 3 may be used to help solve the magnetic induction equation (55) by successive approximation—the method is outlined in Sec. III. As well as introducing the velocity $\mathbf{u}_{ep}$ into the approximation

scheme, it is natural to introduce the effective magnetic field vector

$$\mathbf{b}_{ep} = \mathbf{b}_p - \mathbf{F}. \tag{30}$$

To lowest order the meridional components of the vectors $\mathbf{u}_{ep}$ and $\mathbf{b}_{ep}$ are the effective velocity and magnetic field vectors first introduced by Braginskii[1]. Again, as in the case of heat conduction equation, the analysis provides a framework for discussing the significance of an effective magnetic field. Briefly, the velocity vector $\mathbf{u}_0$ and the magnetic vector $\mathbf{b}_0$ are introduced so that a systematic procedure for solving the heat conduction and magnetic induction equations can be formulated. To this end, the closed streamline property of $\mathbf{u}_0$ is only incidental. However, it is this physical property that clearly gives rise to its mathematical usefulness.

The approach adopted in this paper is suggested by the work in Sec. III of Paper I on the heat conduction equation. For this reason it is an advantage if the reader is acquainted with the work described in Paper I. First Theorem 1 and its proof are suggested by the material between (I.54) and (I.63).[8] Second, the possibility of Theorems 2 and 3 is suggested by (I.63).

II. THE HEAT CONDUCTION EQUATION

In this section Theorem 2 is established, and it is shown how the elementary solutions may help to solve the heat conduction equation.

The subsequent analysis is made clearer by introducing the linear operators $A, Q, D^{(n)}$, and D defined by

$$\begin{aligned} A\mathbf{z} &= \mathbf{u}'\cdot\nabla\mathbf{z}, \qquad Q\mathbf{z} = (\rho/U)\bar{\mathbf{z}}, \\ D^{(n)}\mathbf{z} &= \mathbf{V}^{(n)}\cdot\nabla\mathbf{z}, \qquad D\mathbf{z} = \mathbf{V}\cdot\nabla\mathbf{z}. \end{aligned} \tag{31}$$

In terms of these operators the fluctuating part of (1) leads to the equation

$$(1 + R^{-1/2}QA + R^{-1}QD)\theta' = -QA\Theta \tag{32}$$

[see (I.51)], where it is supposed that $\mathbf{u}_p = \mathbf{V}$. Formally the equation has the solution

$$\theta' = -(1 + R^{-1/2}QA + R^{-1}QD)^{-1}QA\Theta, \tag{33}$$

where the inverse is defined to be the formal binomial expansion of the expression in brackets. It is then possible to show that θ' is given uniquely by ψ and the existence of ψ justifies the inversion. Instead of carrying out the inversion, which is a tedious procedure, it is verified that θ' is ψ.

By setting $\theta' = \psi$ in the left-hand side of (32), it follows that

$$\begin{aligned} &\Big(1 + R^{-1/2}QA + Q\sum_{n=1}^{\infty}(-1)^{n-1}R^{-1/2-n/2}D^{(n)}\Big) \\ &\quad\times\Big(\sum_{n=1}^{\infty}(-1)^n R^{1/2-n/2}Q\psi^{(n)}\Big) \\ &= Q\sum_{n=1}^{\infty}(-1)^n R^{1/2-n/2}\psi^{(n)} \\ &\quad + Q\sum_{n=2}^{\infty}(-1)^{n-1}R^{1/2-n/2}AQ\psi^{(n-1)} \\ &\quad + Q\sum_{n=3}^{\infty}\sum_{r=1}^{n-2}(-1)^n R^{1/2-n/2}D^{(r)}Q\psi^{(n-r-1)}. \end{aligned}$$

After regrouping terms, it follows that

$$\begin{aligned} &(1 + R^{-1/2}QA + R^{-1}QD)\psi \\ &\quad = -QA\Theta + Q\sum_{n=2}^{\infty}(-1)^n R^{1/2-n/2}\,(\psi^{(n)} - AQ\psi^{(n-1)} \\ &\quad + \sum_{r=1}^{n-2}D^{(r)}Q\psi^{(n-r-1)}). \end{aligned} \tag{34}$$

Since $QD^{(n)}\Theta = 0$, the terms in the summation vanish by (18), and hence (32) is verified.

It is now necessary to show that the ϕ average of (1) vanishes. Evidently

$$\langle\mathbf{u}'\cdot\nabla\theta'\rangle = \Big\langle\sum_{n=1}^{\infty}(-1)^n R^{1/2-n/2}\mathbf{u}'\cdot\nabla\Big(\frac{\rho}{U}\overline{\psi^{(n)}}\Big)\Big\rangle,$$

so that using the definition (18) and Theorem 1 leads to

$$\langle\mathbf{u}'\cdot\nabla\theta'\rangle = \sum_{n=1}^{\infty}(-1)^n R^{1/2-n/2}\mathbf{V}^{(n)}\cdot\nabla\Theta = -\mathbf{V}\cdot\nabla\Theta. \tag{35}$$

Hence the ϕ average of (1) is

$$R\langle\mathbf{u}\cdot\nabla\theta\rangle = \mathbf{V}\cdot\nabla\Theta + \langle\mathbf{u}'\cdot\nabla\theta'\rangle = 0. \tag{36}$$

Thus Theorem 2 is established.

The full heat conduction equation

$$\frac{\partial\theta}{\partial t} + R\mathbf{u}\cdot\nabla\theta = \nabla^2\theta, \tag{37}$$

where t is the time, is considered. The mean part of (37) may be solved formally by considering the equation

$$\begin{aligned} L\Theta &= -\langle\mathbf{u}'\cdot\nabla\theta'\rangle \\ &= \langle A(1 + R^{-1/2}QA + R^{-1}QL)^{-1}\,QA\Theta\rangle, \end{aligned} \tag{38}$$

where

$$L \equiv \frac{\partial}{\partial t} + \mathbf{u}_p\cdot\nabla - \nabla^2.$$

(Note that $\mathbf{u}_p$ is not necessarily equal to $\mathbf{V}$.) However, in the proof of Theorem 2 it is established that

$$\mathbf{V}\cdot\nabla\Theta = \langle A(1 + R^{-1/2}QA + R^{-1}QD)^{-1}QA\Theta\rangle. \tag{39}$$

An effective operator L_e is now defined by

$$L_e \equiv \frac{\partial}{\partial t} + (\mathbf{u}_p - \mathbf{V})\cdot\nabla - \nabla^2. \tag{40}$$

Hence combining (38) and (39) leads to

$$L_e\Theta = \langle A[(1 + R^{-1/2}QA + R^{-1}QL)^{-1} - (1 + R^{-1/2}QA + R^{-1}QD)^{-1}]QA\Theta\rangle,$$

or

$$L_e\Theta = -R^{-1}\langle AQ(1 + R^{-1/2}AQ + R^{-1}DQ)^{-1} \times L_e(1 + R^{-1/2}QA + R^{-1}QL)^{-1}QA\Theta\rangle. \quad (41)$$

The inverses are defined by their formal binomial expansions. Of course, the expansions may not be valid for reasons similar to the criticism following (21). If we assume the expansions are correct, the advantage of (41) over (38) is clear. Since the expression on the right-hand side is now order R^{-1}, rather than order 1, the equation is in a convenient form for solution by successive approximation.

Finally, correct to order R^{-1}, Eq. (41) is

$$L_e\Theta = -R^{-1}\langle AQL_eQA\Theta\rangle, \quad (42)$$

which is in agreement with (I.63).

III. THE MAGNETIC INDUCTION EQUATION

It was shown in Sec. I that Theorem 3 is a trivial consequence of Theorem 2. Since the full magnetic induction equation is to be considered later in this section, an alternative proof of Theorem 3 is given which leads directly (without the help of Theorem 2) to the expression (28) for the magnetic field. Moreover, the approximate method of solving the magnetic induction equation (55) follows naturally from this alternative proof.

The fluctuating part of (2) leads to

$$\mathbf{b}' + \rho\mathbf{i}_\phi\frac{U}{\rho}\hat{\mathbf{b}}'\cdot\nabla\frac{\rho}{U} - R^{-1/2}\frac{\rho}{U}\nabla\times\overline{(\mathbf{u}'\times\mathbf{b}')} - R^{-1}\frac{\rho}{U}\nabla\times(\mathbf{u}_p\times\hat{\mathbf{b}}') = \frac{B}{U}\mathbf{u}' - \rho\mathbf{i}_\phi\frac{\rho}{U}\hat{\mathbf{u}}'\cdot\nabla\frac{B}{\rho} + R^{-1}\frac{\rho}{U}\nabla\times(\hat{\mathbf{u}}'\times\mathbf{b}_p). \quad (43)$$

The second term in (43) may be re-expressed by substituting the value of $\mathbf{b}'$ defined by (43) itself. This leads to a cumbersome expression which is most conveniently expressed by the operators S, C, E, H and T defined by

$$S\mathbf{z} = \mathbf{u}'\times\mathbf{z}, \quad C\mathbf{z} = \mathbf{V}\times\mathbf{z}, \quad E\mathbf{z} = \mathbf{F}\times\mathbf{z},$$
$$H\mathbf{z}' = \Theta\mathbf{z}' - \rho\mathbf{i}_\phi\nabla\Theta\cdot\hat{\mathbf{z}}', \quad (44)$$
$$T\mathbf{z} = -\frac{\rho}{U}\nabla\times\bar{\mathbf{z}} + \rho\mathbf{i}_\phi\nabla\frac{\rho}{U}\cdot\nabla\times\hat{\bar{\mathbf{z}}},$$

where it is supposed that $\mathbf{u}_p = \mathbf{V}$ and $\mathbf{b}_p = \mathbf{F}$.[9] Equation (43) relating $\mathbf{u}'$ and $\mathbf{b}'$ becomes

$$(1 + R^{-1/2}TS + R^{-1}TC)\mathbf{b}' = (H + R^{-1}TE)\mathbf{u}'. \quad (45)$$

As in Sec. II, $\mathbf{b}'$ may be determined by inverting the operator on the left-hand side. Instead it is verified that

$$\mathbf{b}' = U\psi\mathbf{i}_\phi + R^{-1/2}(\mathbf{f} - \mathbf{F}) \quad (46)$$

solves (45).

With $\mathbf{b}'$ given by (46), $S\mathbf{b}'$ and $R^{-1/2}C\mathbf{b}'$ are

$$S\mathbf{b}' = \sum_{n=1}^{\infty}(-1)^nR^{1/2-n/2}\mathbf{u}'\overline{\psi^{(n)}}\times\rho\mathbf{i}_\phi + \sum_{n=0}^{\infty}(-1)^nR^{-n/2}\mathbf{u}'\times(\mathbf{f}^{(n)} - \mathbf{F}^{(n)}) \quad (47)$$

and

$$R^{-1/2}C\mathbf{b}' = -\sum_{n=2}^{\infty}(-1)^nR^{1/2-n/2}\left(\sum_{r=1}^{n-1}\mathbf{V}^{(r)}\overline{\psi^{(n-r)}}\right)\times\rho\mathbf{i}_\phi - \sum_{n=1}^{\infty}(-1)^nR^{-n/2}\left(\sum_{r=1}^{n}\mathbf{V}^{(r)}\times(\mathbf{f}^{(n-r)} - \mathbf{F}^{(n-r)}\right). \quad (48)$$

Combining these results and making use of the identity (20) leads to

$$(S + R^{-1/2}C)\mathbf{b}' = \sum_{n=1}^{\infty}(-1)^nR^{1/2-n/2}(\mathbf{f}^{(n)} + \mathbf{V}^{(n)}\Theta)\times U\mathbf{i}_\phi - R^{-1/2}\mathbf{u}'\times\mathbf{F} + \sum_{n=1}^{\infty}(-1)^nR^{-n/2}\left(\sum_{r=1}^{n-1}\mathbf{V}^{(r)}\times\mathbf{F}^{(n-r)}\right). \quad (49)$$

After some straightforward manipulation, operating on (49) by T leads to

$$(TS + R^{-1/2}TC)\mathbf{b}' = R^{-1/2}TE\mathbf{u}' + \sum_{n=1}^{\infty}(-1)^nR^{1/2-n/2}\left(-\mathbf{f}^{(n)} + \mathbf{F}^{(n)} + \rho\mathbf{i}_\phi\overline{\psi^{(n+1)}}\right). \quad (50)$$

It follows that

$$(1 + R^{-1/2}TS + R^{-1}TC)\mathbf{b}' = \mathbf{f}^{(0)} - \rho\mathbf{i}_\phi\overline{\psi^{(1)}} + R^{-1}TE\mathbf{u}'. \quad (51)$$

However,

$$H\mathbf{u}' = \mathbf{f}^{(0)} - \rho\mathbf{i}_\phi\overline{\psi^{(1)}},$$

and hence Eq. (45) is satisfied by (46).

It is now necessary to show that the ϕ average of (2) vanishes. Evidently

$$\langle \mathbf{u}' \times \mathbf{b}' \rangle = \langle S\mathbf{b}' \rangle, \tag{52}$$

and with the aid of (19), (20) and (47) this leads to

$$\begin{aligned}\langle \mathbf{u}' \times \mathbf{b}' \rangle &= \sum_{n=1}^{\infty} (-1)^n R^{1/2-n/2} (\mathbf{F}^{(n)} + \mathbf{V}^{(n)}\Theta) \times U\mathbf{i}_\phi + \sum_{n=0}^{\infty} (-1)^n R^{-n/2} \Big(\sum_{r=1}^{n-1} \mathbf{V}^{(r)} \times \mathbf{F}^{(n-r)} \Big), \\ &= -U\mathbf{i}_\phi \times \mathbf{F} - \mathbf{V} \times B\mathbf{i}_\phi - R^{-1}\mathbf{V} \times \mathbf{F}. \end{aligned} \tag{53}$$

Hence the ϕ average of (2) is

$$R\langle \nabla \times (\mathbf{u} \times \mathbf{b}) \rangle = R\nabla \times [(U\mathbf{i}_\phi + R^{-1}\mathbf{V}) \times (B\mathbf{i}_\phi + R^{-1}\mathbf{F})] + \nabla \times \langle \mathbf{u}' \times \mathbf{b}' \rangle = 0. \tag{54}$$

It is a simple matter to verify that $\nabla \cdot \mathbf{b}' = 0$, and thus Theorem 3 is established.

The full magnetic induction equation

$$\frac{\partial \mathbf{b}}{\partial t} = R\nabla \times (\mathbf{u} \times \mathbf{b}) + \nabla^2 \mathbf{b} \tag{55}$$

is considered. The mean part of the equation is

$$\frac{\partial}{\partial t}(B\mathbf{i}_\phi + R^{-1}\mathbf{b}_p) = R\nabla \times [(U\mathbf{i}_\phi + R^{-1}\mathbf{u}_p) \times (B\mathbf{i}_\phi + R^{-1}\mathbf{b}_p)] + \nabla \times \langle \mathbf{u}' \times \mathbf{b}' \rangle + \nabla^2 (B\mathbf{i}_\phi + R^{-1}\mathbf{b}_p), \tag{56}$$

where formally

$$\langle \mathbf{u}' \times \mathbf{b}' \rangle = -\langle S(1 + R^{-1/2}TS + R^{-1}GM)^{-1}(H + R^{-1}TN)\mathbf{u}' \rangle \tag{57}$$

and

$$G\mathbf{z} = \frac{\rho}{U}\bar{\mathbf{z}} - \rho \mathbf{i}_\phi \nabla \frac{\rho}{U} \cdot \hat{\bar{\mathbf{z}}}, \qquad M\mathbf{z} = \frac{\partial}{\partial t}\mathbf{z} - \nabla \times (\mathbf{u}_p \times \mathbf{z}) - \nabla^2 \mathbf{z}, \quad N\mathbf{z} = \mathbf{b}_p \times \mathbf{z}. \tag{58}$$

As in the case of the heat conduction equation, some simplifications to (56) can be made by using the identity

$$R(U\mathbf{i}_\phi + R^{-1}\mathbf{V}) \times (B\mathbf{i}_\phi + R^{-1}\mathbf{F}) = -\langle S(1 + R^{-1/2}TS + R^{-1}TC)^{-1}(H + \mathrm{R}^{-1}TE)\mathbf{u}' \rangle, \tag{59}$$

and this leads to

$$\begin{aligned}&\frac{\partial}{\partial t}(B\mathbf{i}_\phi + R^{-1}\mathbf{b}_p) - R\nabla \times [(U\mathbf{i}_\phi + R^{-1}\mathbf{u}_p) \times (B\mathbf{i}_\phi + R^{-1}\mathbf{b}_p) - (U\mathbf{i}_\phi + R^{-1}\mathbf{V}) \times (B\mathbf{i}_\phi + R^{-1}\mathbf{F}) - \nabla^2 (B\mathbf{i}_\phi + R^{-1}\mathbf{b}_p)] \\ &\quad = -\nabla \times \langle S[(1 + R^{-1/2}TS + R^{-1}GM)^{-1}(H + R^{-1}TN) - (1 + R^{-1/2}TS + R^{-1}TC)^{-1}(H + R^{-1}TE)]\mathbf{u}' \rangle. \end{aligned} \tag{60}$$

Further manipulation leads to the form

$$\begin{aligned}&\Big(\frac{\partial}{\partial t}(B\mathbf{i}_\phi) - \nabla \times (\mathbf{u}_{ep} \times B\mathbf{i}_\phi) - \nabla \times (U\mathbf{i}_\phi \times \mathbf{b}_{ep}) - \nabla^2 (B\mathbf{i}_\phi) \Big) + R^{-1} \Big(\frac{\partial}{\partial t}\mathbf{b}_{ep} - \nabla \times (\mathbf{u}_{ep} \times \mathbf{b}_{ep}) - \nabla^2 \mathbf{b}_{ep} \Big) \\ &\quad = -R^{-1} \Big(\frac{\partial}{\partial t}\mathbf{F} - \nabla \times (\mathbf{u}_{ep} \times \mathbf{F}) - \nabla \times (\mathbf{V} \times \mathbf{b}_{ep}) - \nabla^2 \mathbf{F} \Big) \\ &\quad + R^{-1} \nabla \times \langle SK^{-1}[GM_e(K + R^{-1}GM_e)^{-1}(H + R^{-1}TN) - TN_e]\mathbf{u}' \rangle, \end{aligned} \tag{61}$$

where

$$\begin{aligned}&\mathbf{u}_{ep} = \mathbf{u}_p - \mathbf{V}, \quad \mathbf{b}_{ep} = \mathbf{b}_p - \mathbf{F}, \quad M_e\mathbf{z} = \frac{\partial}{\partial t}\mathbf{z} - \nabla \times (\mathbf{u}_{ep} \times \mathbf{z}) - \nabla^2 \mathbf{z}, \quad N_e\mathbf{z} = \mathbf{b}_{ep} \times \mathbf{z}, \\ &\quad K = 1 + R^{-1/2}TS + R^{-1}TC. \end{aligned} \tag{62}$$

For the purpose of iteration it is probably convenient to restrict $\mathbf{u}_{ep}$ and $\mathbf{b}_{ep}$ to being meridional vectors, but this does not appear to be necessary. The azimuthal and meridional components of (61) are considered separately. The azimuthal component yields the equation

$$\frac{\partial}{\partial t}(B\mathbf{i}_\phi) - \nabla \times (\mathbf{u}_{ep} \times B\mathbf{i}_\phi) - \nabla \times (U\mathbf{i}_\phi \times \mathbf{b}_{ep}) - \nabla^2 (B\mathbf{i}_\phi) = O(R^{-1}). \tag{63}$$

This equation is well known and corresponds to (I.22a). However, in order to solve this equation, further information about $\mathbf{b}_{ep}$ is needed. Hence the order R^{-1} meridional terms of (61) must be considered. Since the terms on the right-hand side of (61) are also order R^{-1}, the procedure becomes more involved, and there appears to be no simple way of making further reductions. Now correct to order $R^{-1/2}$ an equation for $\mathbf{b}_{ep}$ can be obtained from the meridional component of (61) with the aid of (63), namely

$$\frac{\partial}{\partial t}\mathbf{b}_{ep} - \nabla \times (\mathbf{u}_{ep} \times \mathbf{b}_{ep}) - \nabla^2 \mathbf{b}_{ep} = \nabla \times (\Gamma B \mathbf{i}_\phi), \tag{64}$$

where Γ is a function of U and $\mathbf{u}'$ and is given correct to order $R^{-1/2}$ by Tough.[4] It should be emphasised that (64) is obtained from (61) after very lengthy calculations. To this order, the advection terms on the right-hand side cancel, and the $\nabla \times (\Gamma B \mathbf{i}_\phi)$ term results from the difference of the diffusion terms. To lowest-order, Eq. (63) and (64) were first obtained by Braginskii.[1]

ACKNOWLEDGMENTS

The author wishes to thank Professor S. Childress for comments on a preliminary draft of this paper and Professor P. H. Roberts for a critical reading of the final manuscript. The work reported here was begun during the tenure of a Post Doctoral Visiting Membership awarded by the Courant Institute of Mathematical Sciences, New York University, New York, and was finished during the tenure of a Visiting Fellowship awarded by the Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, Colorado.

APPENDIX:

Theorem 1 is established by induction. It is assumed that

$$\langle \psi^{(n)} \rangle = 0, \qquad \nabla \cdot \mathbf{V}^{(n-1)} = 0, \tag{A1}$$

for $n \le N-1$, and it will be shown that the result is true for $n = N$.

Subsequently, when a subscripted suffix is also repeated, an abbreviated suffix notation is adopted, e.g.,

$$\partial_{i_1}\partial_{i_2}\{\chi^{(i_1)} f_{i_1}^{(q_1)} \chi^{(i_2)} f_{i_2}^{(q_2)}\}$$

will be denoted by

$$\partial_1\partial_2\{\chi^{(1)} f_1^{(q_1)} \chi^{(2)} f_2^{(q_2)}\}.$$

Note that since the i_1 and i_2 in the example are repeated, the expression represents a sum of nine terms. Using the abbreviated notation, we introduce the following functions:

$$F_{n,r,s} = \binom{n-r-1}{s} \sum_{(q_i:r)} \left\langle \partial_1 \cdots \partial_{n-r} \left(\frac{e}{U} \overline{\psi^{(q_0)}} \frac{e}{U} \chi^{(1)} \overline{v_1^{(q_1)}} \cdots \frac{e}{U} \chi^{(s)} \overline{v_s^{(q_s)}} \chi^{(s+1)} u'_{s+1} \frac{e}{U} \chi^{(s+2)} \hat{u}'_{s+2} \cdots \frac{e}{U} \chi^{(n-r)} \hat{u}'_{n-r} \right) \right\rangle,$$

$$G_{n,r,s} = \binom{n-r-1}{s-1} \sum_{(q_i:r)} \left\langle \partial_1 \cdots \partial_{n-r} \left(\frac{e}{U} \overline{\psi^{(q_0)}} \frac{e}{U} \chi^{(1)} \overline{v_1^{(q_1)}} \cdots \frac{e}{U} \chi^{(s-1)} \overline{v_{s-1}^{(q_{s-1})}} \chi^{(s)} V_s^{(q_s)} \frac{e}{U} \chi^{(s+1)} \hat{u}'_{s+1} \cdots \frac{e}{U} \chi^{(n-r)} \hat{u}'_{n-r} \right) \right\rangle,$$

$$f_{n,r,s} = \binom{n-r-1}{s} \sum_{(q_i:r)} \left\langle \partial_1 \cdots \partial_{n-r} \left(\Theta \frac{e}{U} \chi^{(1)} \overline{v_1^{(q_1)}} \cdots \frac{e}{U} \chi^{(s)} \overline{v_s^{(q_s)}} \chi^{(s+1)} u'_{s+1} \frac{e}{U} \chi^{(s+2)} \hat{u}'_{s+2} \cdots \frac{e}{U} \chi^{(n-r)} \hat{u}'_{n-r} \right) \right\rangle,$$

$$g_{n,r,s} = \binom{n-r-1}{s-1} \sum_{(q_i:r)} \left\langle \partial_1 \cdots \partial_{n-r} \left(\Theta \frac{e}{U} \chi^{(1)} \overline{v_1^{(q_1)}} \cdots \frac{e}{U} \chi^{(s-1)} v_{s-1}^{(q_{s-1})} \chi^{(s)} V_s^{(q_s)} \frac{e}{U} \chi^{(s+1)} \hat{u}'_{s+1} \cdots \frac{e}{U} \chi^{(n-r)} \hat{u}'_{n-r} \right) \right\rangle, \tag{A2}$$

where

$$\binom{n}{s} = \frac{n!}{s!\,(n-s)!}$$

and where the summation $(q_i : r)$ is over all distinct vectors $(q_0, q_1, \cdots, q_s)$ such that $\sum_{i=0}^{s} q_i = r$, where the q_i are nonzero, positive integers. For $f_{n,r,s}$ and $g_{n,r,s}$, the q_0 is absent.

A second induction proof is now used to show that[10]

$$\frac{1}{(N-1)!} \left\langle \partial_1 \cdots \partial_N \left(\Theta \chi^{(1)} u'_1 \frac{e}{U} \chi^{(2)} \hat{u}'_2 \cdots \frac{e}{U} \chi^{(N)} \hat{u}'_N \right) \right\rangle$$

$$= \begin{cases} -\dfrac{1}{(N-2)!}\,(F_{N,1,0} + f_{N,1,1} - g_{N,1,1}), & r = 1, \\ \dfrac{(-1)^r}{(N-r-1)!} \left(\displaystyle\sum_{s=0}^{N-r-1} F_{N,r,s} - \sum_{s=1}^{N-r} G_{N,r,s} + \sum_{s=1}^{N-r-1} f_{N,r,s} - \sum_{s=1}^{N-r} g_{N,r,s} \right), & 2 \le r \le N-2, \\ (-1)^{N-1}\,(F_{N,N-1,0} - G_{N,N-1,1} - g_{N,N-1,1}), & r = N-1, \end{cases} \tag{A3}$$

for $N \ge 2$.

By noting that $\nabla\cdot\mathbf{V}^{(q)} = 0, q = 1, 2, \cdots, N-2$, one of the differentiations in the definition (A2) is carried out and gives

$$F_{N,r,s} = F^{(1)}_{N,r,s} + F^{(2)}_{N,r,s} + F^{(3)}_{N,r,s}, \quad G_{N,r,s} = G^{(1)}_{N,r,s} + G^{(2)}_{N,r,s} + G^{(3)}_{N,r,s}, \quad f_{N,r,s} = f^{(1)}_{N,r,s} + f^{(2)}_{N,r,s} + f^{(3)}_{N,r,s},$$
$$g_{N,r,s} = g^{(1)}_{N,r,s} + g^{(2)}_{N,r,s} + g^{(3)}_{N,r,s}, \tag{A4}$$

where

$$F^{(1)}_{N,r,s} = \binom{N-r-1}{s} \sum_{(q_i:r)} \left\langle \partial_1 \cdots \partial_{N-r-1} \left[\left(\mathbf{u}'\cdot\nabla \frac{e}{U}\overline{\psi^{(q_0)}} \right) \frac{e}{U}\chi^{(1)}\overline{v_1^{(q_1)}} \cdots \frac{e}{U}\chi^{(s)}\overline{v_s^{(q_s)}} \frac{e}{U}\chi^{(s+1)}\hat{u}'_{s+1} \cdots \frac{e}{U}\chi^{(N-r-1)} \right.\right.$$
$$\left.\left. \times \hat{u}'_{N-r-1} \right] \right\rangle, \quad 0 \le s \le N-r-1,$$

$$F^{(2)}_{N,r,s} = \begin{cases} s\binom{N-r-1}{s} \sum_{(q_i:r)} \left\langle \partial_1 \cdots \partial_{N-r-1} \left[\frac{e}{U}\overline{\psi^{(q_0)}} \frac{e}{U}\chi^{(1)}\overline{v_1^{(q_1)}} \cdots \frac{e}{U}\chi^{(s-1)}\overline{v_{s-1}^{(q_{s-1})}} \left(\mathbf{u}'\cdot\nabla \frac{e}{U}\chi^{(s)}\overline{v_s^{(q_s)}}\right) \frac{e}{U}\chi^{(s+1)} \right.\right. \\ \qquad \left.\left. \times \hat{u}'_{s+1} \cdots \frac{e}{U}\chi^{(N-r-1)}\hat{u}'_{N-r-1} \right] \right\rangle, \quad 1 \le s \le N-r-1, \\ 0, \quad s = 0, \end{cases}$$

$$F^{(3)}_{N,r,s} = \begin{cases} (N-r-s-1)\binom{N-r-1}{s} \sum_{(q_i:r)} \left\langle \partial_1 \cdots \partial_{N-r-1} \left[\frac{e}{U}\overline{\psi^{(q_0)}} \frac{e}{U}\chi^{(1)}\overline{v_1^{(q_1)}} \cdots \frac{e}{U}\chi^{(s)}\overline{v_s^{(q_s)}} \left(\mathbf{u}'\cdot\nabla \frac{e}{U}\chi^{(s+1)} \right.\right.\right. \\ \qquad \left.\left.\left. \times \hat{u}'_{s+1}\right) \frac{e}{U}\chi^{(s+2)}\hat{u}'_{s+2} \cdots \frac{e}{U}\chi^{(N-r-1)}\hat{u}'_{N-r-1} \right] \right\rangle, \quad 0 \le s \le N-r-2, \\ 0, \quad s = N-r-1, \end{cases}$$

and

$$G^{(1)}_{N,r,s} = \binom{N-r-1}{s-1} \sum_{(q_i:r)} \left\langle \partial_1 \cdots \partial_{N-r-1} \left[\left(\mathbf{V}^{(q)}\cdot\nabla \frac{e}{U}\overline{\psi^{(q_0)}} \right) \frac{e}{U}\chi^{(1)}\overline{v_1^{(q_1)}} \cdots \frac{e}{U}\chi^{(s-1)}\overline{v_{s-1}^{(q_{s-1})}} \frac{e}{U}\chi^{(s)}\hat{u}'_s \cdots \frac{e}{U} \right.\right.$$
$$\left.\left. \times \chi^{(N-r-1)}\hat{u}'_{N-r-1} \right] \right\rangle, \quad 1 \le s \le N-r,$$

$$G^{(2)}_{N,r,s} = \begin{cases} (s-1)\binom{N-r-1}{s-1} \sum_{(q_i:r)} \left\langle \partial_1 \cdots \partial_{N-r-1} \left[\frac{e}{U}\overline{\psi^{(q_0)}} \frac{e}{U}\chi^{(1)}\overline{v_1^{(q_1)}} \cdots \frac{e}{U}\chi^{(s-2)}\overline{v_{s-2}^{(q_{s-2})}} \right.\right. \\ \qquad \left.\left. \times \left(\mathbf{V}^{(q)}\cdot\nabla \frac{e}{U}\chi^{(s-1)}\overline{v_{s-1}^{(q_{s-1})}}\right) \frac{e}{U}\chi^{(s)}\hat{u}'_s \cdots \frac{e}{U}\chi^{(N-r-1)}\hat{u}'_{N-r-1} \right] \right\rangle, \quad 2 \le s \le N-r, \\ 0, \quad s = 1, \end{cases}$$

$$G^{(3)}_{N,r,s} = \begin{cases} (N-r-s)\binom{N-r-1}{s-1} \sum_{(q_i:r)} \left\langle \partial_1 \cdots \partial_{N-r-1} \left[\frac{e}{U}\overline{\psi^{(q_0)}} \frac{e}{U}\chi^{(1)}\overline{v_1^{(q_1)}} \cdots \frac{e}{U}\chi^{(s-1)}\overline{v_{s-1}^{(q_{s-1})}} \right.\right. \\ \qquad \left.\left. \times \left(\mathbf{V}^{(q)}\cdot\nabla \frac{e}{U}\chi^{(s)}\hat{u}'_s\right) \frac{e}{U}\chi^{(s+1)}\hat{u}'_{s+1} \cdots \frac{e}{U}\chi^{(N-r-1)}\hat{u}'_{N-r-1} \right] \right\rangle, \quad 1 \le s \le N-r-1, \\ 0, \quad s = N-r, \end{cases}$$

and

$$f^{(1)}_{N,r,s} = \begin{cases} \binom{N-r-1}{s} \sum_{(q_i:r)} \left\langle \partial_1 \cdots \partial_{N-r-1} \left[\left(\mathbf{u}'\cdot\nabla\Theta\right) \frac{e}{U}\chi^{(1)}\overline{v_1^{(q_1)}} \cdots \frac{e}{U}\chi^{(s)}\overline{v_s^{(q_s)}} \frac{e}{U}\chi^{(s+1)}\hat{u}'_{s+1} \cdots \frac{e}{U}\chi^{(N-r-1)}\hat{u}'_{N-r-1} \right] \right\rangle, \\ \quad 1 \le s \le N-r-1, \end{cases}$$

$$f^{(2)}_{N,r,s} = s\binom{N-r-1}{s} \sum_{(q_i:r)} \left\langle \partial_1 \cdots \partial_{N-r-1} \left[\Theta \frac{e}{U}\chi^{(1)}\overline{v_1^{(q_1)}} \cdots \frac{e}{U}\chi^{(s-1)}\overline{v_{s-1}^{(q_{s-1})}} \left(\mathbf{u}'\cdot\nabla \frac{e}{U}\chi^{(s)}\overline{v_s^{(q_s)}}\right) \frac{e}{U}\chi^{(s+1)} \right.\right.$$
$$\left.\left. \times \hat{u}'_{s+1} \cdots \frac{e}{U}\chi^{(N-r-1)}\hat{u}'_{N-r-1} \right] \right\rangle, \quad 1 \le s \le N-r-1,$$

$$f^{(3)}_{N,r,s} = \begin{cases} (N-r-s-1)\binom{N-r-1}{s}\sum_{(q_i:r)}\left\langle \partial_1\cdots\partial_{N-r-1}\left[\Theta\,\frac{e}{U}\,\chi^{(1)}\overline{v_1^{(q_1)}}\cdots\frac{e}{U}\,\chi^{(s)}\overline{v_s^{(q_s)}}\left(\mathbf{u}'\cdot\nabla\,\frac{e}{U}\,\chi^{(s+1)}\hat{u}'^{(q_{s+1})}_{s+1}\right)\right.\right. \\ \qquad \left.\left.\times\,\frac{e}{U}\,\chi^{(s+2)}\hat{u}'_{s+2}\cdots\frac{e}{U}\,\chi^{(N-r-1)}\hat{u}'_{N-r-1}\right]\right\rangle, \quad 1\le s\le N-r-2, \\ 0, \quad s=N-r-1, \end{cases}$$

and

$$g^{(1)}_{N,r,s} = \begin{cases} \binom{N-r-1}{s-1}\sum_{(q_i:r)}\left\langle \partial_1\cdots\partial_{N-r-1}\left[\left(\mathbf{V}^{(q)}\cdot\nabla\Theta\right)\frac{e}{U}\,\chi^{(1)}\overline{v_1^{(q_1)}}\cdots\frac{e}{U}\,\chi^{(s-1)}\overline{v_{s-1}^{(q_{s-1})}}\,\frac{e}{U}\,\chi^{(s)}\hat{u}'_s\cdots\frac{e}{U}\right.\right. \\ \qquad \left.\left.\times\,\chi^{(N-r-1)}\hat{u}'_{N-r-1}\right]\right\rangle, \quad 1\le s\le N-r, \end{cases}$$

$$g^{(2)}_{N,r,s} = \begin{cases} (s-1)\binom{N-r-1}{s-1}\sum_{(q_i:r)}\left\langle \partial_1\cdots\partial_{N-r-1}\left[\Theta\,\frac{e}{U}\,\chi^{(1)}\overline{v_1^{(q_1)}}\cdots\frac{e}{U}\,\chi^{(s-2)}\overline{v_{s-2}^{(q_{s-2})}}\left(\mathbf{V}^{(q)}\cdot\nabla\,\frac{e}{U}\,\chi^{(s-1)}\overline{v_{s-1}^{(q_{s-1})}}\right)\right.\right. \\ \qquad \left.\left.\times\,\frac{e}{U}\,\chi^{(s)}\hat{u}'_s\cdots\frac{e}{U}\,\chi^{(N-r-1)}\hat{u}'_{N-r-1}\right]\right\rangle, \quad 2\le s\le N-r, \\ 0, \quad s=1, \end{cases}$$

$$g^{(3)}_{N,r,s} = \begin{cases} (N-r-s)\binom{N-r-1}{s-1}\sum_{(q_i:r)}\left\langle \partial_1\cdots\partial_{N-r-1}\left[\Theta\,\frac{e}{U}\,\chi^{(1)}\overline{v_1^{(q_1)}}\cdots\frac{e}{U}\,\chi^{(s-1)}\overline{v_{s-1}^{(q_{s-1})}}\left(\mathbf{V}^{(q)}\cdot\nabla\,\frac{e}{U}\,\chi^{(s)}\hat{u}'_s\right)\right.\right. \\ \qquad \left.\left.\times\,\frac{e}{U}\,\chi^{(s+1)}\hat{u}'_{s+1}\cdots\frac{e}{U}\,\chi^{(N-r-1)}\hat{u}'_{N-r-1}\right]\right\rangle, \quad 1\le s\le N-r-1, \\ 0, \quad s=N-r. \end{cases}$$

The above terms are collected together and with the help of the definitions (17) and (18) lead to

$$F^{(1)}_{N,r,s} - G^{(1)}_{N,r,s+1} - g^{(1)}_{N,r,s+1}$$
$$= \binom{N-r-1}{s}\sum_{(q_i:r)}\left\langle \partial_1\cdots\partial_{N-r-1}\left(\psi^{(q_0+1)}\,\frac{e}{U}\,\chi^{(1)}\overline{v_1^{(q_1)}}\cdots\frac{e}{U}\,\chi^{(s)}\overline{v_s^{(q_s)}}\,\frac{e}{U}\,\chi^{(s+1)}\hat{u}'_{s+1}\cdots\frac{e}{U}\,\chi^{(N-r-1)}\right.\right.$$
$$\left.\left.\times\,\hat{u}'_{N-r-1}\right)\right\rangle, \quad 0\le s\le N-r-1,$$

$$f^{(1)}_{N,r,s} = \binom{N-r-1}{s}\sum_{(q_i:r)}\left\langle \partial_1\cdots\partial_{N-r-1}\left(\psi^{(1)}\,\frac{e}{U}\,\chi^{(1)}\overline{v_1^{(q_1)}}\cdots\frac{e}{U}\,\chi^{(s)}\overline{v_s^{(q_s)}}\,\frac{e}{U}\,\chi^{(s+1)}\hat{u}'_{s+1}\cdots\frac{e}{U}\,\chi^{(N-r-1)}\right.\right.$$
$$\left.\left.\times\,\hat{u}'_{N-r-1}\right)\right\rangle, \quad 1\le s\le N-r-1,$$

and

$$F^{(2)}_{N,r,s} - G^{(2)}_{N,r,s+1} - G^{(3)}_{N,r,s}$$
$$= \begin{cases} s\binom{N-r-1}{s}\sum_{(q_i:r)}\left\langle \partial_1\cdots\partial_{N-r-1}\left(\overline{\psi^{(q_0)}}\,\frac{e}{U}\,\chi^{(1)}\overline{v_1^{(q_1)}}\cdots\frac{e}{U}\,\chi^{(s-1)}\overline{v_{s-1}^{(q_{s-1})}}\,\frac{e}{U}\,\chi^{(s)}v_s^{(q_s+1)}\,\frac{e}{U}\,\chi^{(s+1)}\right.\right. \\ \qquad \left.\left.\times\,\hat{u}'_{s+1}\cdots\frac{e}{U}\,\chi^{(N-r-1)}\hat{u}'_{N-r-1}\right)\right\rangle, \quad 1\le s\le N-r-1, \\ 0, \quad s=0, N-r, \end{cases}$$

$$F^{(3)}_{N,r,s-1} = \begin{cases} s\binom{N-r-1}{s}\sum_{(q_i:r)}\left\langle \partial_1\cdots\partial_{N-r-1}\left(\overline{\psi^{(q_0)}}\,\frac{e}{U}\,\chi^{(1)}\overline{v_1^{(q_1)}}\cdots\frac{e}{U}\,\chi^{(s-1)}\overline{v_{s-1}^{(q_{s-1})}}\,\frac{e}{U}\,\chi^{(s)}v_s^{(1)}\,\frac{e}{U}\,\chi^{(s+1)}\right.\right. \\ \qquad \left.\left.\times\,\hat{u}'_{s+1}\cdots\frac{e}{U}\,\chi^{(N-r-1)}\hat{u}'_{N-r-1}\right)\right\rangle, \quad 1\le s\le N-r-1, \\ 0, \quad s=N-r, \end{cases}$$

and

$$f^{(2)}_{N,r,s} - g^{(2)}_{N,r,s+1} - g^{(3)}_{N,r,s}$$

$$= \begin{cases} s\binom{N-r-1}{s} \sum\limits_{(q_i:r)} \left\langle \partial_1 \cdots \partial_{N-r-1} \left(\Theta \frac{e}{U} \chi^{(1)} \overline{v_1^{(q_1)}} \cdots \frac{e}{U} \chi^{(s-1)} \overline{v_{s-1}^{(q_{s-1})}} \frac{e}{U} \chi^{(s)} v_s^{(q_s+1)} \frac{e}{U} \chi^{(s+1)} \right.\right. \\ \qquad \left.\left. \times \hat{u}'_{s+1} \cdots \frac{e}{U} \chi^{(N-r-1)} \hat{u}'_{N-r-1} \right) \right\rangle, \quad 1 \le s \le N-r-1, \\ 0, \quad s = 0, N-r, \end{cases}$$

$$f^{(3)}_{N,r,s-1} = \begin{cases} s\binom{N-r-1}{s} \sum\limits_{(q_i:r)} \left\langle \partial_1 \cdots \partial_{N-r-1} \left(\Theta \frac{e}{U} \chi^{(1)} \overline{v_1^{(q_1)}} \cdots \frac{e}{U} \chi^{(s-1)} \overline{v_{s-1}^{(q_{s-1})}} \frac{e}{U} \chi^{(s)} v_s^{(1)} \frac{e}{U} \chi^{(s+1)} \right.\right. \\ \qquad \left.\left. \times \hat{u}'_{s+1} \cdots \frac{e}{U} \chi^{(N-r-1)} \hat{u}'_{N-r-1} \right) \right\rangle, \quad 2 \le s \le N-r-1, \\ 0, \quad s = 1, N-r. \end{cases}$$

Collecting together the above terms leads to

$$F^{(1)}_{N,r,s} - G^{(1)}_{N,r,s+1} + f^{(1)}_{N,r,s} - g^{(1)}_{N,r,s+1}$$

$$= \binom{N-r-1}{s} \sum_{(q_i:r+1)} \left\langle \partial_1 \cdots \partial_{N-r-1} \left(\psi^{(q_0)} \frac{e}{U} \chi^{(1)} \overline{v_1^{(q_1)}} \cdots \frac{e}{U} \chi^{(s)} \overline{v_s^{(q_s)}} \frac{e}{U} \chi^{(s+1)} \hat{u}'_{s+1} \cdots \frac{e}{U} \chi^{(N-r-1)} \hat{u}'_{N-r-1} \right) \right\rangle,$$

$$0 \le s \le N-r-1,$$

and

$$F^{(2)}_{N,r,s} + F^{(3)}_{N,r,s-1} - G^{(2)}_{N,r,s+1} - G^{(3)}_{N,r,s}$$

$$= \begin{cases} s\binom{N-r-1}{s} \sum\limits_{(q_i:r+1)} \left\langle \partial_1 \cdots \partial_{N-r-1} \left(\overline{\psi^{(q_0)}} \frac{e}{U} \chi^{(1)} \overline{v_1^{(q_1)}} \cdots \frac{e}{U} \chi^{(s-1)} \overline{v_{s-1}^{(q_{s-1})}} \frac{e}{U} \chi^{(s)} v_s^{(q_s)} \frac{e}{U} \chi^{(s+1)} \hat{u}'_{s+1} \cdots \frac{e}{U} \chi^{(N-r-1)} \right.\right. \\ \qquad \left.\left. \times u'_{N-r-1} \right) \right\rangle, \quad 1 \le s \le N-r-1, \\ 0, \quad s = 0, N-r, \end{cases}$$

and

$$f^{(2)}_{N,r,s} + f^{(3)}_{N,r,s-1} - g^{(2)}_{N,r,s+1} - g^{(3)}_{N,r,s}$$

$$= \begin{cases} s\binom{N-r-1}{s} \sum\limits_{(q_i:r+1)} \left\langle \partial_1 \cdots \partial_{N-r-1} \left(\Theta \frac{e}{U} \chi^{(1)} \overline{v_1^{(q_1)}} \cdots \frac{e}{U} \chi^{(s-1)} \overline{v_{s-1}^{(q_{s-1})}} \frac{e}{U} \chi^{(s)} v_s^{(q_s)} \frac{e}{U} \chi^{(s+1)} \hat{u}'_{s+1} \cdots \frac{e}{U} \chi^{(N-r-1)} \hat{u}'_{N-r-1} \right) \right\rangle, \\ \qquad 1 \le s \le N-r-1, \\ 0, \quad s = 0, N-r. \end{cases}$$

Further manipulation using the identity (11) and noting that $\langle \psi^{(n)} \rangle = 0$, $n = 1, \cdots, N-1$, leads to

$$F^{(1)}_{N,r,s} + F^{(2)}_{N,r,s} + F^{(3)}_{N,r,s-1} - G^{(1)}_{N,r,s+1} - G^{(2)}_{N,r,s+1} - G^{(3)}_{N,r,s} + f^{(1)}_{N,r,s} - g^{(1)}_{N,r,s+1}$$

$$= -(N-r-1) \left\{ \binom{N-r-2}{s} \sum_{(q_i:r+1)} \left\langle \partial_1 \cdots \partial_{N-r-1} \left(\frac{e}{U} \overline{\psi^{(q_0)}} \frac{e}{U} \chi^{(1)} \overline{v_1^{(q_1)}} \cdots \frac{e}{U} \chi^{(s)} \overline{v_s^{(q_s)}} \chi^{(s+1)} u'_{s+1} \frac{e}{U} \chi^{(s+2)} \right.\right.\right.$$

$$\left.\left.\left. \times \hat{u}'_{s+2} \cdots \frac{e}{U} \chi^{(N-r-1)} \hat{u}'_{N-r-1} \right) \right\rangle - \binom{N-r-2}{s-1} \sum_{(q_i:r+1)} \left\langle \partial_1 \cdots \partial_{N-r-1} \frac{e}{U} \overline{\psi^{(q_0)}} \frac{e}{U} \chi^{(1)} \overline{v_1^{(q_1)}} \cdots \frac{e}{U} \chi^{(s-1)} \overline{v_{s-1}^{(q_{s-1})}} \right.$$

$$\left. \times \chi^{(s)} V_s^{(q_s)} \frac{e}{U} \chi^{(s+1)} \hat{u}'_{s+1} \cdots \frac{e}{U} \chi^{(N-r-1)} \hat{u}'_{N-r-1} \right) \Big\rangle \Big\},$$

$$= \begin{cases} -(N-r-1) \ (F_{N,r+1,s} - G_{N,r+1,s}), & 0 \le s \le N-r-1, \quad r \le N-2, \\ 0, \quad s = N-r. \end{cases} \tag{A5}$$

Similarly it can be shown that

$$f^{(2)}_{N,r,s} + f^{(3)}_{N,r,s-1} - g^{(2)}_{N,r,s+1} - g^{(3)}_{N,r,s}$$
$$= \begin{cases} -(N-r-1)(f_{N,r+1,s} - g_{N,r+1,s}), & 1 \le s \le N-r-1, \quad r \le N-2, \\ 0, & s = 0, N-r. \end{cases} \tag{A6}$$

It is a simple matter to show that (A3) is correct for $r = 1$, and induction using (A4), (A5), and (A6) establishes (A3) for $1 \le r \le N-1$.

Now, by the identity (11) and the commutation property of the operator ∂_i [Eq.(16)], the left-hand side of (A3) is zero. Hence for the case $r = N-1$ we have the identity

$$\left\langle \partial_1 \left(\frac{e}{U} \overline{\psi^{(N-1)}} \chi^{(1)} u'_1 \right) \right\rangle - \sum_{(q_i:N-1)} \left\langle \partial_1 \left(\frac{e}{U} \psi^{(q_0)} \chi^{(1)} V_1^{(q_1)} \right) \right\rangle - \left\langle \partial_1 \left(\Theta \chi^{(1)} V_1^{(N-1)} \right) \right\rangle = 0. \tag{A7}$$

Setting $\Theta = 1$, we conclude that

$$\nabla \cdot \mathbf{V}^{(N-1)} = 0, \tag{A8}$$

and consequently by (A7)

$$\langle \psi^{(N)} \rangle = 0. \tag{A9}$$

The results (A8) and (A9) are easily established for $N = 1$ and $N = 2$. This completes the induction proof.

* Present address: School of Mathematics, University of Newcastle upon Tyne, NE1 7 IU, England.

1 S. I. Braginskii, Zh. Eksp. Teor. Fiz, **47**, 1084 (1964) [Sov. Phys. JETP **20**, 726 (1965)].
2 S. I. Braginskii, Zh. Eksp. Teor. Fiz. **47**, 2178 (1964) [Sov. Phys. JETP **20**, 1462 (1965)].
3 S. I. Braginskii, Geomagnetizm i Aeronomiya **4**, 732 (1964) [Geomagnetism and Aeronomy **4**, 572 (1964)].
4 J. G. Tough, Geophys. J. **13**, 393 (1967).
5 J. G. Tough and R. D. Gibson, *The Application of Modern Physics to the Earth and Planetary Interiors* (Interscience, New York, 1969), p. 555.
6 J. G. Tough and P. H. Roberts, Phys. Earth Planet. Interiors **1**, 288 (1968).
7 A. M. Soward, J. Math. Phys. **12**, 1900 (1971).

8 The abbreviation (I*n*) is used to denote Eq. (*n*) in the previous paper.[7]
9 Since only a particular class of solutions of (2) is sought, the supposition $\mathbf{b}_p = \mathbf{F}$ is justified.
10 For the *s* summation, not all the terms indicated by the summation necessarily exist. There is a further restriction on the terms F, G, f, g given by the definition (A2) concerning the summation over $(q_i : r)$. Clearly in (A3) the cases $r = 1$ and $N-1$ are special cases of the case $2 \le r \le N-2$.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 10 OCTOBER 1971

Two-Time Spin-Pair Correlation Function of the Heisenberg Magnet at Infinite Temperature

Tohru Morita

Department of Applied Science, Faculty of Engineering, Tohoku University, Sendai, Japan

and

*Department of Physics, Ohio University, Athens, Ohio 45701**

(Received 20 April 1971)

The expansion coefficients in powers of time of the two-time spin-pair correlation function are obtained up to terms of order t^8 for the Heisenberg magnet of spin $\frac{1}{2}$ at infinite temperature for the square and cubic lattices and up to terms of order t^{10} for the linear chain. The result is applicable to the isotropic as well as anisotropic Heisenberg magnet where the exchange integrals in the z direction and in the orthogonal plane are assumed to be different. Analysis of the result is left to future work.

1. INTRODUCTION

The two-time spin-pair correlation function is the quantity of primary importance in the theory of neutron scattering from magnetic materials.[1,2] A number of attempts to obtain it have been presented for the Heisenberg model. The approximate methods[3] always involve an unclear error. The exact calculation for a finite linear chain at infinite temperature has been given,[4] but its generalization to the more interesting three-dimensional case seems impossible at the present time. The expansion formula for the two-time correlation function in powers of time has been written by Kubo.[5] The two leading terms of the expansion of $\langle S_i^z(t) S_f^z(0) \rangle$ were given for the Heisenberg magnet at infinite temperature by de Gennes[6] and Collins and Marshall.[7] The corresponding terms are given by McFadden and Tahir-Kheli,[8] by assuming different values for the exchange integrals $J_{\parallel}$ and $J_{\perp}$ in the z–z direction and in the orthogonal plane. The third term of order t^6 has been reported by McFadden and Tahir-Kheli[9] for general spin and Morita *et al.*[10] for spin $\frac{1}{2}$.

Similarly it can be shown that

$$f^{(2)}_{N,r,s} + f^{(3)}_{N,r,s-1} - g^{(2)}_{N,r,s+1} - g^{(3)}_{N,r,s}$$
$$= \begin{cases} -(N-r-1)(f_{N,r+1,s} - g_{N,r+1,s}), & 1 \le s \le N-r-1, \quad r \le N-2, \\ 0, & s = 0, N-r. \end{cases} \tag{A6}$$

It is a simple matter to show that (A3) is correct for $r = 1$, and induction using (A4), (A5), and (A6) establishes (A3) for $1 \le r \le N-1$.

Now, by the identity (11) and the commutation property of the operator ∂_i [Eq.(16)], the left-hand side of (A3) is zero. Hence for the case $r = N-1$ we have the identity

$$\left\langle \partial_1 \left(\frac{e}{U} \overline{\psi^{(N-1)}} \chi^{(1)} u'_1 \right) \right\rangle - \sum_{(q_i:N-1)} \left\langle \partial_1 \left(\frac{e}{U} \psi^{(q_0)} \chi^{(1)} V_1^{(q_1)} \right) \right\rangle - \left\langle \partial_1 \left(\Theta \chi^{(1)} V_1^{(N-1)} \right) \right\rangle = 0. \tag{A7}$$

Setting $\Theta = 1$, we conclude that

$$\nabla \cdot \mathbf{V}^{(N-1)} = 0, \tag{A8}$$

and consequently by (A7)

$$\langle \psi^{(N)} \rangle = 0. \tag{A9}$$

The results (A8) and (A9) are easily established for $N = 1$ and $N = 2$. This completes the induction proof.

* Present address: School of Mathematics, University of Newcastle upon Tyne, NE1 7 IU, England.

1 S. I. Braginskii, Zh. Eksp. Teor. Fiz, **47**, 1084 (1964) [Sov. Phys. JETP **20**, 726 (1965)].
2 S. I. Braginskii, Zh. Eksp. Teor. Fiz. **47**, 2178 (1964) [Sov. Phys. JETP **20**, 1462 (1965)].
3 S. I. Braginskii, Geomagnetizm i Aeronomiya **4**, 732 (1964) [Geomagnetism and Aeronomy **4**, 572 (1964)].
4 J. G. Tough, Geophys. J. **13**, 393 (1967).
5 J. G. Tough and R. D. Gibson, *The Application of Modern Physics to the Earth and Planetary Interiors* (Interscience, New York, 1969), p. 555.
6 J. G. Tough and P. H. Roberts, Phys. Earth Planet. Interiors **1**, 288 (1968).
7 A. M. Soward, J. Math. Phys. **12**, 1900 (1971).

8 The abbreviation (I*n*) is used to denote Eq. (*n*) in the previous paper.[7]
9 Since only a particular class of solutions of (2) is sought, the supposition $\mathbf{b}_p = \mathbf{F}$ is justified.
10 For the *s* summation, not all the terms indicated by the summation necessarily exist. There is a further restriction on the terms F, G, f, g given by the definition (A2) concerning the summation over $(q_i : r)$. Clearly in (A3) the cases $r = 1$ and $N-1$ are special cases of the case $2 \le r \le N-2$.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 10 OCTOBER 1971

Two-Time Spin-Pair Correlation Function of the Heisenberg Magnet at Infinite Temperature

Tohru Morita

Department of Applied Science, Faculty of Engineering, Tohoku University, Sendai, Japan

and

*Department of Physics, Ohio University, Athens, Ohio 45701**

(Received 20 April 1971)

The expansion coefficients in powers of time of the two-time spin-pair correlation function are obtained up to terms of order t^8 for the Heisenberg magnet of spin $\frac{1}{2}$ at infinite temperature for the square and cubic lattices and up to terms of order t^{10} for the linear chain. The result is applicable to the isotropic as well as anisotropic Heisenberg magnet where the exchange integrals in the z direction and in the orthogonal plane are assumed to be different. Analysis of the result is left to future work.

1. INTRODUCTION

The two-time spin-pair correlation function is the quantity of primary importance in the theory of neutron scattering from magnetic materials.[1,2] A number of attempts to obtain it have been presented for the Heisenberg model. The approximate methods[3] always involve an unclear error. The exact calculation for a finite linear chain at infinite temperature has been given,[4] but its generalization to the more interesting three-dimensional case seems impossible at the present time. The expansion formula for the two-time correlation function in powers of time has been written by Kubo.[5] The two leading terms of the expansion of $\langle S_i^z(t) S_f^z(0) \rangle$ were given for the Heisenberg magnet at infinite temperature by de Gennes[6] and Collins and Marshall.[7] The corresponding terms are given by McFadden and Tahir-Kheli,[8] by assuming different values for the exchange integrals $J_{\parallel}$ and $J_{\perp}$ in the z–z direction and in the orthogonal plane. The third term of order t^6 has been reported by McFadden and Tahir-Kheli[9] for general spin and Morita *et al.*[10] for spin $\frac{1}{2}$.

Nakamura[11] gave the fourth term of order t^8 in the expansion of the autocorrelation function $\langle S_i^z(t)S_i^z(0)\rangle$ for the one-dimensional Heisenberg magnet at infinite temperature. The fifth term of that expansion is given in Ref. 10. In the present paper, we report the calculation of the coefficients of the expansion of $\langle S_i^z(t)S_f^z(0)\rangle$ as far as the fourth term of order t^8 for the square and cubic Heisenberg magnets and the fifth term of order t^{10} for the linear Heisenberg magnet. The present calculation is restricted to spin $\frac{1}{2}$ and infinite temperature.

2. BASIC FORMULAS

We consider the Heisenberg magnet. The Hamiltonian of the system is given by

$$H = -\sum_f \sum_g [J_\perp(f,g)S_f^- S_g^+ + J_\parallel(f,g)S_f^z S_g^z], \quad (2.1)$$

where $J_\perp(f,g)$ and $J_\parallel(f,g)$ are equal to $J_\perp$ and $J_\parallel$, respectively, when f and g are nearest-neighbor lattice sites and zero otherwise. This Hamiltonian reduces to the one for the Heisenberg model when $J_\parallel = J_\perp$. If $J_\parallel = 0$ and the system is one dimensional, one obtains the X-Y model for which we know the analytic expression for the two-time spin-pair correlation function.

We calculate the two-time spin-pair correlation function $\sigma(R,t)$, which is defined by

$$\sigma(R_{if},t) = \langle S_i^z(t)S_f^z(0)\rangle - \langle S_i^z\rangle\langle S_f^z\rangle. \quad (2.2)$$

Here $S_i^z(t)$ is the Heisenberg time-shifted operator:

$$S_i^z(t) = e^{iHt}S_i^z e^{-iHt}. \quad (2.3)$$

In the present paper, we assume that the system is at infinite temperature. Hence the average $\langle A\rangle$ of an arbitrary operator A is calculated by

$$\langle A\rangle = \mathrm{tr}A/\mathrm{tr}1. \quad (2.4)$$

We evaluate the trace in the representation in which the z components of N spins are given. Then (2.4) is written as

$$\langle A\rangle = \frac{1}{2^N}\sum_{\{S_i^z\}}\langle\psi\{S_i^z\}|A|\psi\{S_i^z\}\rangle, \quad (2.5)$$

where $\psi\{S_i^z\}$ is unity when all the z components of spins are equal to the given respective value in the set $\{S_i^z\}$ and zero otherwise.

Because of the well-known property of the trace, $\sigma(R_{if},t) = \sigma(R_{fi},-t)$ at infinite temperature. For a regular lattice, there is an inversion symmetry about the center of the ith and fth lattice site. By this property, $\sigma(R_{if},t) = \sigma(R_{fi},t)$. Combining these two properties, one sees that $\sigma(R_{fi},t) = \sigma(R_{fi},-t)$; namely, the function $\sigma(R_{if},t)$ is an even function of time t.

Expanding $S_i^z(t)$ defined by (2.3) in powers of time t and substituting the result into (2.2), one obtains

$$\sigma(R_{if},t) = \sigma^{(0)}(R_{if}) + \sum_{n=1}^{\infty}\frac{(-1)^n}{(2n)!}\sigma^{(2n)}(R_{if})t^{2n}, \quad (2.6)$$

where

$$\begin{aligned}\sigma^{(0)}(R_{if}) &= \langle S_i^z S_f^z\rangle = \tfrac{1}{4}\delta_{if},\\ \sigma^{(2)}(R_{if}) &= \langle[H,[H,S_i^z]]S_f^z\rangle,\\ \sigma^{(4)}(R_{if}) &= \langle[H,[H,[H,[H,S_i^z]]]]S_f^z\rangle,\end{aligned} \quad (2.7)$$

and so on.

When (2.1) is substituted in (2.7), $\sigma^{(2n)}(R_{if})$ is expressed as a sum of terms with factors $J_\parallel^{2l}J_\perp^{2n-2l}$, where $2l = 0,2,4,\cdots,2n$. We shall denote the coefficients as $\sigma_{2l}{}^{(2n)}(R_{if})$:

$$\sigma^{(2n)}(R_{if}) = \sum_{l=0}^{n} J_\parallel{}^{2l}J_\perp{}^{2n-2l}\sigma_{2l}{}^{(2n)}(R_{if}). \quad (2.8)$$

For the isotropic case $J_\parallel = J_\perp \equiv J$, one has

$$\sigma^{(2n)}(R_{if}) = J^{2n}\sigma_t{}^{(2n)}(R_{if}), \quad (2.9)$$

where

$$\sigma_t{}^{(2n)}(R_{if}) = \sum_{l=0}^{n}\sigma_{2l}{}^{(2n)}(R_{if}). \quad (2.10)$$

As the Hamiltonian H commutes with the sum of S^z over all the spins, $\sum_i S_i^z = \sum_f S_f^z$, one obtains the sum rule:

$$\sum_i \sigma^{(2n)}(R_{if}) = 0, \quad n \geq 1, \quad (2.11)$$

and

$$\sum_f \sigma^{(2n)}(R_{if}) = 0, \quad n \geq 1. \quad (2.12)$$

Substituting (2.8) into (2.11) or (2.12) and noticing that the equation obtained must be valid for arbitrary values of $J_\parallel$ and $J_\perp$, one gets the sum rules for $J_{2l}{}^{(2n)}(R_{if})$ as follows:

$$\sum_i \sigma_{2l}{}^{(2n)}(R_{if}) = 0, \quad n \geq 1, \quad (2.13)$$

$$\sum_f \sigma_{2l}{}^{(2n)}(R_{if}) = 0, \quad n \geq 1. \quad (2.14)$$

Summing over l, the corresponding relations for $\sigma_t{}^{(2n)}(R_{if})$ are obtained. The relation (2.14) and the corresponding relation for $\sigma_t{}^{(2n)}(R_{if})$ will be used to check the final results.

3. COMPUTATION

The computation consists of commutations of spin operators for pairs of neighboring sites. The commutation at each stage is taken between the spin operators for the pair of sites (say ith and jth sites) and one of the three terms ($S_i^+S_j^-$, $S_i^-S_j^+$ and $2S_i^zS_j^z$) referred to that pair in the Hamiltonian. The commutation relations for all the possible cases are given in the form shown in Table I as a memory in the computer.

The calculations are performed for ten finite diagrams shown in Fig. 1. The lattice sites and

TABLE I. Commutation relations for a pair of lattice sites, 1 and 2. The first column in each result is the number of nonzero terms to be summed to give the result. The second and third columns denote the operators for the sites 1 and 2, respectively. $-$, $+$, and z represent S^-, S^+, and S^z, respectively, and 1 for operator 1. The last column is the numerical factor for the product of the operators.

Spin operators at sites 1 & 2	Result of commutation with $S_1^-S_2^+$	Result of commutation with $S_1^+S_2^-$	Result of commutation with $2S_1^zS_2^z$
1 1	0	0	0
1 $-$	1 $-$ z 2.0	0	0 z $-$ -2.0
1 z	1 $-$ $+$ -1.0	1 $+$ $-$ 1.0	0
1 $+$	0	1 $+$ z -2.0	1 z $+$ 2.0
$-$ 1	0	1 z $-$ 2.0	1 $-$ z -2.0
$-$ $-$	0	0	0
$-$ z	0	1 1 $-$ 0.5	1 $-$ 1 -0.5
$-$ $+$	0	2 $\{z\ 1\ 1.0\}$ $\{1\ z\ -1.0\}$	0
z 1	1 $-$ $+$ 1.0	1 $+$ $-$ -1.0	0
z $-$	1 $-$ 1 0.5	0	1 1 $-$ -0.5
z z	0	0	0
z $+$	0	1 $+$ 1 -0.5	1 1 $+$ 0.5
$+$ 1	1 z $+$ -2.0	0	1 $+$ z 2.0
$+$ $-$	2 $\{z\ 1\ -1.0\}$ $\{1\ z\ 1.0\}$	0	0
$+$ z	1 1 $+$ -0.5	0	1 $+$ 1 0.5
$+$ $+$	0	0	0

bonds are labeled, and which lattice sites are connected by a bond is put in the memory.

The following are the cases to be covered in the calculation of the term of order t^n; where the number given in parenthesis is the possible number of the cases at that stage:

diagram (number of diagrams with less than $\frac{1}{2}n + 1$ bonds, given in Fig. 1);

initial site I (number of sites in the diagram);

{ first bond (number of bonds in the diagram),
term in H (three terms),
result of commutation (0, 1, or 2);

{ second bond (number of bonds in the diagram),
term in H (three terms),
result of commutation (0, 1, or 2);

.
.
.

{ nth bond (number of bonds in the diagram),
term in H (three terms),
result of commutation (0, 1 or 2);

final site F (number of sites in the diagram).

After the nth commutation, it is checked whether all the bonds of the diagram are used and, if such is the case, the trace is taken and the contribution is accumulated separately according to the number of times in which the term involving $J_{||}$ in the Hamiltonian is used for the commutation among n commutations. The result of each commutation is zero or is expressed by a product of spins or a sum of two terms as shown in Table I. If it is zero, we go to another term of H or another bond or another site.

Except for the initial site I, when the commutation is taken for the first time with respect to one of the bonds connected to a site the resulting state involves one of the operators S^+, S^-, or S^z for that site and also one of S^+, S^-, or S^z for the other site of the bond. If that site is not chosen to be F, another commutation must be involved to that site to give a nonzero result; by means of this commutation the other site involved in the commutation enters one of the states S^+, S^-, or S^z. Hence commutations must be taken at least $2(m-1)$ times

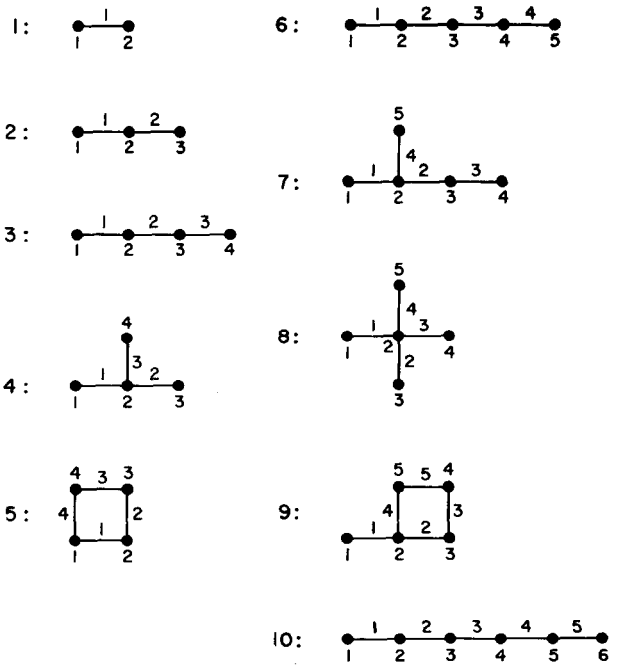


Fig. 1

TABLE II. The coefficients $\gamma_l^{(2n)}$(diagram, IF) and $\gamma_l^{(2n)}$(diagram, IF) for the terms of order t^{2n} for the nine diagrams given in **Fig. 1**, where IF denote the initial and final sites.

$2n = 2$

diagram	IF	$\gamma_t^{(2)}$	$\gamma_0^{(2)}$
1	11	0.5	0.5
1	12	−0.5	−0.5

$2n = 4$

diagram	IF	$\gamma_t^{(4)}$	$\gamma_0^{(4)}$	$\gamma_2^{(4)}$
1	11	2.0	2.0	0.0
1	12	−2.0	−2.0	0.0
2	11	1.0	0.5	0.5
2	12	−2.5	−2.0	−0.5
2	13	1.5	1.5	0.0
2	22	5.0	4.0	1.0

$2n = 6$

diagram	IF	$\gamma_t^{(6)}$	$\gamma_0^{(6)}$	$\gamma_2^{(6)}$	$\gamma_4^{(6)}$
1	11	8.0	8.0	0.0	0.0
1	12	−8.0	−8.0	0.0	0.0
2	11	15.0	9.0	5.5	0.5
2	12	−32.5	−24.0	−8.0	−0.5
2	13	17.5	15.0	2.5	0.0
2	22	65.0	48.0	16.0	1.0
3	11	2.0	0.5	1.5	0.0
3	12	−7.0	−3.0	−4.0	0.0
3	13	10.0	7.5	2.5	0.0
3	14	−5.0	−5.0	0.0	0.0
3	22	24.0	8.5	12.5	3.0
3	23	−27.0	−13.0	−11.0	−3.0
4	11	10.0	4.0	3.0	3.0
4	12	−35.0	−24.0	−8.0	−3.0
4	13	12.5	10.0	2.5	0.0
4	22	105.0	72.0	24.0	9.0
5	11	−7.0	0.0	−7.0	0.0
5	12	11.0	0.0	11.0	0.0
5	13	−15.0	0.0	−15.0	0.0

$2n = 8$

diagram	IF	$\gamma_t^{(8)}$	$\gamma_0^{(8)}$	$\gamma_2^{(8)}$	$\gamma_4^{(8)}$	$\gamma_6^{(8)}$
1	11	32.0	32.0	0.0	0.0	0.0
1	12	−32.0	−32.0	0.0	0.0	0.0
2	11	161.0	98.0	54.0	8.5	0.5
2	12	−332.5	−224.0	−96.0	−12.0	−0.5
2	13	171.5	126.0	42.0	3.5	0.0
2	22	665.0	448.0	192.0	24.0	1.0
3	11	54.0	16.5	33.5	4.0	0.0
3	12	−190.5	−76.0	−96.5	−18.0	0.0
3	13	241.5	143.5	84.0	14.0	0.0
3	14	−105.0	−84.0	−21.0	0.0	0.0
3	22	677.0	208.5	325.5	128.0	15.0
3	23	−728.0	−276.0	−313.0	−124.0	−15.0
4	11	315.0	157.0	72.0	71.0	15.0
4	12	−1015.0	−640.0	−240.0	−120.0	−15.0
4	13	350.0	241.5	84.0	24.5	0.0
4	22	3045.0	1920.0	720.0	360.0	45.0
5	11	−97.0	138.0	−250.0	15.0	0.0
5	12	199.0	−160.0	335.0	24.0	0.0
5	13	−301.0	182.0	−420.0	−63.0	0.0
6	11	4.0	0.5	2.5	1.0	0.0
6	12	−18.0	−4.0	−13.0	−1.0	0.0
6	13	35.0	14.0	21.0	0.0	0.0
6	14	−38.5	−28.0	−10.5	0.0	0.0
6	15	17.5	17.5	0.0	0.0	0.0
6	22	92.0	15.0	53.5	23.5	0.0
6	23	−137.0	−32.0	−72.0	−33.0	0.0
6	24	101.5	49.0	42.0	10.5	0.0
6	33	204.0	36.0	102.0	66.0	0.0
7	11	48.0	11.0	20.5	16.5	0.0
7	12	−195.0	−74.0	−94.0	−27.0	0.0
7	13	157.5	84.0	63.0	10.5	0.0
7	14	−52.5	−42.0	−10.5	0.0	0.0
7	15	42.0	21.0	21.0	0.0	0.0
7	22	900.0	292.0	398.0	165.0	45.0
7	23	−685.0	−256.0	−252.0	−132.0	−45.0
7	24	175.0	112.0	42.0	21.0	0.0
7	33	460.0	120.0	158.0	137.0	45.0
7	34	−90.0	−32.0	−32.0	−26.0	0.0
7	44	20.0	4.0	11.0	5.0	0.0
8	11	210.0	72.0	48.0	45.0	45.0
8	12	−945.0	−576.0	−216.0	−108.0	−45.0
8	13	245.0	168.0	56.0	21.0	0.0
8	22	3780.0	2304.0	864.0	432.0	180.0
9	11	−14.0	0.0	−14.0	0.0	0.0
9	12	126.0	0.0	126.0	0.0	0.0
9	13	−77.0	0.0	−91.0	14.0	0.0
9	14	42.0	0.0	70.0	−28.0	0.0
9	22	−324.0	0.0	−336.0	12.0	0.0
9	23	218.0	0.0	196.0	22.0	0.0
9	24	−238.0	0.0	−182.0	−56.0	0.0
9	33	−98.0	0.0	−56.0	−42.0	0.0
9	34	125.0	0.0	77.0	48.0	0.0
9	35	−168.0	0.0	−126.0	−42.0	0.0
9	44	−54.0	0.0	−42.0	−12.0	0.0

TABLE III. The number of different ways $n(R_{if}, \text{diagram}, IF)$ by which diagrams with the initial and final sites IF occur in the lattices with the difference of the final and initial sites R_{if}. (l), (lm) and (l,m,n) for R_{if} correspond to the linear, square, and simple cubic lattices, respectively.

$R_{if} = (0), (0,0), (0,0,0)$

diagram	IF	(0)	(0,0)	(0,0,0)
1	11	2	4	6
2	11	2	12	30
2	22	1	6	15
3	11	2	36	150
3	22	2	36	150
4	11		12	60
4	22		4	20
5	11		4	12
6	11	2	100	726
6	22	2	100	726
6	33	1	50	363
7	11		72	600
7	22		36	300
7	33		36	300
7	44		36	300
8	11		4	60
8	22		1	15
9	11		8	48
9	22		8	48
9	33		16	96
9	44		8	48

$R_{if} = (1), (1,0), (1,0,0)$

diagram	IF	(1)	(1,0)	(1,0,0)
1	12	1	1	1
2	12	1	3	5
2	21	1	3	5
3	12	1	9	25
3	14		2	4
3	21	1	9	25
3	23	1	9	25
4	12		3	10
4	21		3	10
5	12		2	4
6	12	1	25	121
6	14		4	16
6	21	1	25	121
6	23	1	25	121
6	32	1	25	121
6	41		4	16
7	12		18	100
7	14		4	16
7	21		18	100
7	23		9	50
7	32		9	50
7	34		9	50
7	41		4	16
7	43		9	50
8	12		1	10
8	21		1	10
9	12		2	8
9	21		2	8
9	23		4	16
9	32		4	16
9	34		4	16
9	43		4	16

$R_{if} = (2), (2,0), (2,0,0)$

diagram	IF	(2)	(2,0)	(2,0,0)
2	13	1	1	1
3	13	1	3	5
3	31	1	3	5
4	13		2	4
6	13	1	9	25
6	15		6	12
6	24	1	9	25
6	31	1	9	25
7	13		6	20
7	15		6	20
7	24		3	10
7	31		6	20
7	42		3	10
8	13		1	6
9	13		2	4
9	31		2	4

$R_{if} = (3), (3,0), (3,0,0)$

diagram	IF	(3)	(3,0)	(3,0,0)
3	14	1	1	1
6	14	1	3	5
6	41	1	3	5
7	14		2	4
7	41		2	4

$R_{if} = (4), (4,0), (4,0,0)$

diagram	IF	(4)	(4,0)	(4,0,0)
6	15	1	1	1

$R_{if} = (1,1), (1,1,0)$

diagram	IF	(1,1)	(1,1,0)
2	13	2	2
3	13	6	10
3	31	6	10
4	13	4	8
5	13	1	1
6	13	16	48
6	15	4	16
6	24	16	48
6	31	16	48
7	13	12	40
7	15	12	40
7	24	6	20
7	31	12	40
7	42	6	20
8	13	2	12
9	13	2	6
9	24	2	4
9	31	2	6
9	35	4	8
9	42	2	4

TABLE III–(continued)

R_{if} = (2, 1), (2, 1, 0)

diagram	IF	(2, 1)	(2, 1, 0)
3	14	3	3
6	14	9	15
6	41	9	15
7	14	6	12
7	41	6	12
9	14	1	1
9	41	1	1

R_{if} = (1, 1, 1)

diagram	IF	(1, 1, 1)
3	14	6
6	14	30
6	41	30
7	14	24
7	41	24
9	14	3
9	41	3

R_{if} = (3, 1), (3, 1, 0)

diagram	IF	(3, 1)	(3, 1, 0)
6	15	4	4

R_{if} = (2, 2), (2, 2, 0)

diagram	IF	(2, 2)	(2, 2, 0)
6	15	6	6

R_{if} = (2, 1, 1)

diagram	IF	(2, 1, 1)
6	15	12

TABLE IV. Expansion coefficients, $\sigma_t^{(2n)}(R_{if})$ and $\sigma_{2l}^{(2n)}(R_{if})$ for the linear chain.

R_{if}	$2n$	$\sigma_t^{(2n)}$	$\sigma_0^{(2n)}$	$\sigma_2^{(2)}$	$\sigma_4^{(2n)}$	$\sigma_6^{(2n)}$	$\sigma_8^{(2n)}$
(0)	0	0.25	0.25	0.0	0.0	0.0	0.0
	2	1.0	1.0	0.0	0.0	0.0	0.0
	4	11.0	9.0	2.0	0.0	0.0	0.0
	6	163.0	100.0	55.0	8.0	0.0	0.0
	8	2 909.0	1 225.0	1 232.0	420.0	32.0	0.0
	10	60 704.0	15 876.0	25 704.0	16 260.0	2736.0	128.0
(1)	2	−0.5	−0.5	0.0	0.0	0.0	0.0
	4	−7.0	−6.0	−1.0	0.0	0.0	0.0
	6	−114.0	−75.0	−35.0	−4.0	0.0	0.0
	8	−2 116.0	−980.0	−868.0	−252.0	−16.0	0.0
	10	−44 356.0	−13 230.0	−19 026.0	−10 452.0	−1584.0	−64.0
(2)	4	1.5	1.5	0.0	0.0	0.0	0.0
	6	37.5	30.0	7.5	0.0	0.0	0.0
	8	826.0	490.0	294.0	42.0	0.0	0.0
	10	18 300.0	7 560.0	7 896.0	2 628.0	216.0	0.0
(3)	6	−5.0	−5.0	0.0	0.0	0.0	0.0
	8	−182.0	−140.0	−42.0	0.0	0.0	0.0
	10	−5 073.0	−2 835.0	−1 932.0	−306.0	0.0	0.0
(4)	8	17.5	17.5	0.0	0.0	0.0	0.0
	10	840.0	630.0	210.0			
(5)	10	−63.0	−63.0	0.0	0.0	0.0	0.0

in order to get a contribution from a diagram of m lattice sites. If the term of order t^n is of interest, the number of sites m which are involved in the commutation must be given by $m \leq \frac{1}{2}n + 1$. For $n = 2, 4, 6, 8, m \leq 2, 3, 4, 5$. For this reason, we need the diagrams with $m \leq 5$ lattice sites for the calculation of the terms up to t^8. They are listed as diagrams 1–9 in Fig. 1. In calculating the term of t^{10} for the linear chain, we need the linear diagram with six lattice sites; that is listed as diagram 10 in Fig. 1.

Table II shows the average of commutations for those diagrams. For these diagrams, we used the notations $\gamma_t^{(2n)}$(diagram, IF), where I and F mean initial and final sites, respectively, and $\gamma_{2l}^{(2n)}$(diagram, IF) in place of $\sigma_t^{(2n)}(R_{if})$ and $\sigma_{2l}^{(2n)}(R_{if})$, respectively. As the $\gamma_t^{(2n)}$(diagram, IF) and $\gamma_{2l}^{(2n)}$(diagram, IF) are symmetric with respect to I and F, the values are not listed for $F < I$. The results are given for $2n \leq 8$. As the Hamiltonian commutes with the sum of S^z for all the sites, e.g., $\Sigma_I S_I{}^z = \Sigma_F S_F{}^z$, we have identities:

$$\begin{aligned} &\Sigma_I \gamma_{2l}^{(2n)}(\text{diagram}, IF) = 0, \quad n \geq 1 \\ &\Sigma_F \gamma_{2l}^{(2n)}(\text{diagram}, IF) = 0, \quad n \geq 1, \end{aligned} \tag{3.1}$$

and the corresponding relations for $\gamma_t^{(2n)}$. These relations provide a method of checking our results.

In order to calculate the space–time correlation functions, one needs the number of how many ways those diagrams with the initial and final sites (IF)

TABLE V. Expansion coefficients, $\sigma_t^{(2n)}(R_{if})$ and $\sigma_{2l}^{(2n)}(R_{if})$ for the square lattice.

R_{if}	$2n$	$\sigma_t^{(2n)}$	$\sigma_0^{(2n)}$	$\sigma_2^{(2n)}$	$\sigma_4^{(2n)}$	$\sigma_6^{(2n)}$
(0,0)	0	0.25	0.25	0.0	0.0	0.0
	2	2.0	2.0	0.0	0.0	0.0
	4	50.0	38.0	12.0	0.0	0.0
	6	2 050.0	1 088.0	770.0	192.0	0.0
	8	120 790.0	43 918.0	47 080.0	25 280.0	4512.0
(1,0)	2	−0.5	−0.5	0.0	0.0	0.0
	4	−17.0	−14.0	−3.0	0.0	0.0
	6	−770.0	−477.0	−245.0	−48.0	0.0
	8	−46 000.0	−20 916.0	−16 600.0	−7 356.0	−1128.0
(2,0)	4	1.5	1.5	0.0	0.0	0.0
	6	102.5	80.0	22.5	0.0	0.0
	8	7 098.0	4 242.0	2 296.0	560.0	0.0
(3,0)	6	−5.0	−5.0	0.0	0.0	0.0
	8	−546.0	−420.0	−126.0	0.0	0.0
(4,0)	8	17.5	17.5	0.0	0.0	0.0
(1,1)	4	3.0	3.0	0.0	0.0	0.0
	6	190.0	160.0	30.0	0.0	0.0
	8	12 096.0	8 372.0	3 136.0	588.0	0.0
(2,1)	6	−15.0	−15.0	0.0	0.0	0.0
	8	−1 554.0	−1 260.0	−238.0	−56.0	0.0
(3,1)	8	70.0	70.0	0.0	0.0	0.0
(2,2)	8	105.0	105.0	0.0	0.0	0.0

TABLE VI. Expansion coefficients $\sigma_t^{(2n)}(R_{if})$ and $\sigma_{2l}^{(2n)}(R_{if})$ for the simple cubic lattice.

R_{if}	$2n$	$\sigma_t^{(2n)}$	$\sigma_0^{(2n)}$	$\sigma_2^{(2n)}$	$\sigma_4^{(2n)}$	$\sigma_6^{(2n)}$
(0,0,0)	0	0.25	0.25	0.0	0.0	0.0
	2	3.0	3.0	0.0	0.0	0.0
	4	117.0	87.0	30.0	0.0	0.0
	6	7 989.0	4 068.0	3 081.0	840.0	0.0
	8	830 907.0	287 679.0	325 800.0	180 948.0	36 480.0
(1,0,0)	2	−0.5	−0.5	0.0	0.0	0.0
	4	−27.0	−22.0	−5.0	0.0	0.0
	6	−2 034.0	−1 223.0	−671.0	−140.0	0.0
	8	−213 836.0	−92 820.0	−79 164.0	−35 772.0	−6 080.0
(2,0,0)	4	1.5	1.5	0.0	0.0	0.0
	6	167.5	130.0	37.5	0.0	0.0
	8	19 978.0	11 690.0	6 706.0	1 582.0	0.0
(3,0,0)	6	−5.0	−5.0	0.0	0.0	0.0
	8	−910.0	−700.0	−210.0	0.0	0.0
(4,0,0)	8	17.5	17.5	17.5	0.0	0.0
(1,1,0)	4	3.0	3.0	0.0	0.0	0.0
	6	320.0	260.0	60.0	0.0	0.0
	8	36 232.0	23 268.0	10 724.0	2 240.0	0.0
(2,1,0)	6	−15.0	−15.0	0.0	0.0	0.0
	8	−2 646.0	−2 100.0	−490.0	−56.0	0.0
(3,1,0)	8	70.0	70.0	0.0	0.0	0.0
(2,2,0)	8	105.0	105.0	0.0	0.0	0.0
(1,1,1)	6	−30.0	−30.0	0.0	0.0	0.0
	8	−5 208.0	−4 200.0	−840.0	−168.0	0.0
(2,1,1)	8	210.0	210.0	0.0	0.0	0.0

appear when the initial and final lattice sites of which the difference is R_{if}, are given in the lattice. Those numbers $n(R_{if};\text{diagram},IF)$ are listed in Table III for the linear, square, and cubic lattices. The results for those lattices are obtained as a sum of products of the average for each figure and the number of ways that figure appears:

$$\sigma_{2l}^{(2n)}(R_{if}) = \sum_{\text{diagram}} \sum_{IF} n(R_{if};\text{diagram},IF)\gamma_{2l}^{(2n)}(\text{diagram},IF) \tag{3.2}$$

and

$$\sigma_t^{(2n)}(R_{if}) = \sum_{\text{diagram}} \sum_{IF} n(R_{if};\text{diagram},IF)\gamma_t^{(2n)}(\text{diagram},IF). \tag{3.3}$$

By using Tables II and III, we obtain our final results: Tables IV–VI for $\sigma_{2l}^{(2n)}(R_{if})$ and $\sigma_t^{(2n)}(R_{if})$. Table IV includes the values for $2n = 10$, but the data used in their calculation are not given in Tables II and III. The results have been confirmed

to satisfy the sum rules (2.14) and the corresponding relation for $\sigma_t^{(2n)}(R_{if})$.

$\sigma_0^{(2n)}(R_{if})$ in Table IV are the expansion coefficients for the $X-Y$ model, for which Katsura *et al.*[12] gave the following closed expression:

$$\langle S_i^z(0)S_f^z(t)\rangle = \tfrac{1}{4}[J_{i-f}(2J_\perp t)]^2. \quad (3.4)$$

Our coefficients $\sigma_t^{(2n)}(R_{if})$ in Table IV are confirmed to be correct with the aid of the formula

$$J_n(2t)^2 = t^{2n}\sum_{k=0}^{\infty}\frac{(-1)^k[(2n+2k)!]^2}{[(n+2k)!]^2(2n+k)!\,k!}\,t^{2k} \quad (3.5)$$

4. CONCLUDING REMARKS

This paper is devoted to describe the computation of obtaining the coefficients of the series expansion in powers of time of the two-time spin-pair correlation function of the Heisenberg model of spin $\frac{1}{2}$ at infinite temperature, and to report the results of the computation. The main results are given in Tables IV–VI. Because of the limitation in the computer used, the coefficients of higher orders are obtained after a number of runs of the computer. In order to save the computer time, it is devised that the trial of choosing sites and bonds is terminated when choice of the bonds after that stage will result in zero contribution. Such a device is used by Kobayashi[13] in his calculation for the one-dimensional case. Details of these procedures are not described in the text.

The results obtained for the term of order t^4 agree with those give in Ref. 6–8. The results for the order t^6 is not in agreement with Ref. 9. The results obtained by the analytic expression for the X–Y model[12] is used to check a part of the present computation.

Analysis of our coefficients are in consideration. The computation of the higher-order terms and the computation for finite temperatures and for higher spins are now the problems under consideration.

Cook and Richards[14] announced a future paper on the term of order t^8 of the autocorrelation function. It is hoped that they will give a check for that part of the present result.

ACKNOWLEDGMENTS

The author is grateful to Professor S. Katsura for valuable discussions which stimulated him to take up this computation. It is acknowledged that he is indebted to K. Kobayashi and Y. Abe for information about their programs on the one-dimensional case. The numerical computation was made with the aid of computer NEAC 2200 (model 500) of Computer Center, Tohoku University and computer IBM 360/44 of the Computer Center, Ohio University.

---

* Present address.

1 L. Van Hove, Phys. Rev. **95**, 1374 (1954).
2 P. G. de Gennes, in Magnetism, edited by G. T. Rado and H. Suhl (Academic, New York, 1963), Vol. III, p. 115.
3 E.g., P. Resibois and M. De Leener, Phys. Rev. **152**, 305, 318 (1966); **178**, 806, 819 (1969); K. Kawasaki, J. Chem. Phys. Solids **28**, 1277 (1967); Progr. Theoret. Phys. (Kyoto) **39**, 285 (1968); T. Kawasaki, *ibid.* **39**, 331 (1968); H. Mori and H. Okamoto, *ibid.* 40, 1287 (1968), M. Blume and J. Hubbard, Phys. Rev. B 1, 3815 (1970).
4 F. Carboni and P. M. Richards, Phys. Rev. **177**, 889 (1969).
5 R. Kubo, in *Fluctuation, Relaxation and Resonance in Magnetic Systems*, edited by D. ter Haar (Oliver and Boyd, Edinburgh, 1962), p. 23.
6 P. G. de Gennes, J. Phys. Chem. Solids **4**, 223 (1967).
7 M. F. Collins and W. Marshall, Proc. Phys. Soc. (London) 92, 390 (1967); see also R. A. Tahir-Kheli and D. G. McFadden, Phys. Rev. 182, 604 (1969).
8 D. G. McFadden and R. A. Tahir-Kheli, Phys. Rev. B **1**, 3649 (1970).
9 D. G. McFadden and R. A. Tahir-Kheli, Phys. Rev. B 1, 3671 (1970).
10 T. Morita, K. Kobayashi, S. Katsura, and Y. Abe, Phys. Letters **32A**, 367 (1970).
11 T. Nakamura (unpublished); see Ref. 10.
12 S. Katsura, T. Horiguchi, and M. Suzuki, Physica **46**, 69 (1970).
13 K. Kobayashi, thesis, 1970 (in Japanese).
14 A. C. Cook and P. M. Richards, Bull. Am. Phys. Soc. **16**, 350 (1971).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 10 OCTOBER 1971

On the Decomposition $SO(p, 1) \supset SO(p-1, 1)$ for Most Degenerate Representations

Charles P. Boyer and Farhad Ardalan*
Department of Physics, The Pennsylvania State University, University Park, Pennsylvania 16802
(Received 26 March 1971)

We discuss the problem of decomposing the matrix elements for degenerate representations of $SO(p, 1)$ according to the noncompact subgroup $SO(p-1, 1)$. The overlap functions which obtain this decomposition are calculated explicitly and their properties are used to discuss the $SO(p, 1)$ generalized Regge expansion.

1. INTRODUCTION

During the past few years, there has been a rapid development in the representation theory for noncompact Lie groups.[1] A considerable amount of interest has been focused on the so-called pseudo-orthogonal groups of the type[2] $SO(p, q)$. In addition, if either p or q is equal to unity, the analysis of such representation theory becomes much simpler. Indeed in this case, $SO(p, 1)$, all the unitary irreducible representations have been classified.[3] In most of this literature, the representations have been constructed in a basis made up of the finite-dimensional unitary irreducible representations of the maximal compact subgroups. More recently, however, for various reasons, attentions have been focused on the decomposition according to noncompact subgroups.[4-6] In these cases the results indicate that a unitary irreducible representation of the continuous class decomposes according to its noncompact subgroup into a direct integral of unitary irreducible representations of the continuous class taken twice and a direct sum of unitary irreducible representations of the discrete class taken once (if they occur at all).

In the present paper, we discuss this result for the decomposition $SO(p, 1) \supset SO(p-1, 1)$ in the case of the most-degenerate,[6] continuous, principal series. Although Limic and Niederle previously gave this result,[7] they confined themselves to the decomposition in their representation space. On the other hand, we give, as well, the decomposition of the matrix elements and calculate explicitly the so-called overlap functions between the canonical basis and pseudobasis. Moreover, our representation of these functions is in a form which reveals their singularity structure explicitly and leads to a generalized Regge expansion as well as factorization of the residues.

2. CONSTRUCTION OF THE UNITARY IRREDUCIBLE REPRESENTATIONS OF $SO(p, 1)$ ON THE HYPERBOLOID $H^{(p-1)}$

A single sheet $H_+^{(p-1)} = \{\eta_\mu : \eta^\mu \eta_\mu = \eta_0^2 - \eta_1^2 - \cdots - \eta_{p-1}^2 = 1, \eta_0 > 0\}$ of the two-sheeted hyperboloid is a transitive manifold for the proper group $SO_0(p-1, 1)$. In fact, $H_+^{(p-1)}$ is a rank -1 homogeneous space[2,8] of $SO_0(p-1, 1)$ which is homeomorphic to the coset space $SO_0(p-1, 1)/SO(p-1)$. We can construct the so-called quasiregular representation of $SO_0(p-1, 1)$ on $H_+^{(p-1)}$ as follows: Take the continuously differentiable functions on $H_+^{(p-1)}$ and make a Hilbert space out of them with the inner product

$$(f_1, f_2) = \int_{H_+^{(p-1)}} d\Omega_p \bar{f}_1(\Omega_p) f_2(\Omega_p) \tag{2.1}$$

and norm

$$\|f\| = \int_{H_+^{(p-1)}} d\Omega_p |f(\Omega_p)|^2 < \infty,$$

where $d\Omega_p$ is the usual invariant measure on $H_+^{(p-1)}$; denote this space by $\mathfrak{L}^2(H_+^{(p-1)})$; then let $L(g)$ be an operator on $\mathfrak{L}^2(H_+^{(p-1)})$ such that for each $g \in SO_0(p-1, 1)$ and every $f \in \mathfrak{L}^2(H_+^{(p-1)})$ we have

$$L(g) f(\Omega_p) = f(g^{-1}\Omega_p). \tag{2.2}$$

It is easily verified that $L(g)$ is a representation of $SO_0(p-1, 1)$ and it follows from the invariance of the measure under $SO_0(p-1, 1)$ transformations that $L(g)$ is unitary. However, $L(g)$ is not irreducible, but decomposes into a direct integral of irreducible representations by the Gel'fand-Graev theorem.[9] The irreducible content of $\mathfrak{L}^2(H_+^{(p-1)})$ is obtained by considering the eigenvectors of the only nonvanishing Casimir operator of $SO_0(p-1, 1)$ on $H_+^{(p-1)}$, the Laplace-Beltrami[2,8] operator $\Delta(H_+^{(p-1)})$.

Now we try to extend the representation (2.2) with the aid of multiplier representations to a representation of $SO_0(p, 1)$ in analogy with the representations of $SO(p)$ on the $(p-1)$-dimensional sphere,[10] $S^{(p-1)}$. To this end we consider the quasiregular representation of $SO_0(p, 1)$ on the upper sheet of the p-dimensional cone, $\xi_0^2 - \xi_1^2 - \cdots - \xi_p^2 = 0$, $\xi_0 > 0$, with continuously differentiable homogeneous functions of degree σ of the homogeneous variables $\eta^\mu = \xi^\mu/\xi^p$, $\mu = 0, 1, \cdots, p-1$. First notice that since ξ^p varies between $-\infty$ and $+\infty$, η^μ is not well defined when ξ^p takes on the value zero. Secondly, since $\xi_0 > 0$, when $\xi^p < 0$ we have $\eta_0 < 0$ and we cannot identify this with the single sheet $H_+^{(p-1)}$. In fact, we can handle both of these problems by considering both the sheets $H_+^{(p-1)}$ and $H_-^{(p-1)}$ and taking those functions in $\mathfrak{L}^2(H_+^{(p-1)})$ and $\mathfrak{L}^2(H_-^{(p-1)})$ that are related in a continuous fashion at infinity. The relevant Hilbert space then becomes the direct sum $\mathfrak{L}^2(H^{(p-1)}) = \mathfrak{L}^2(H_+^{(p-1)}) \oplus \mathfrak{L}^2(H_-^{(p-1)})$, and the continuity problem can be made manifest by mapping $\mathfrak{L}^2(H^{(p-1)})$ unitarily[11] onto $L^2(S^{(p-1)})$. As a result, we have two manifolds, $H_+^{(p-1)}$ and $H_-^{(p-1)}$, on which the group $SO_0(p-1, 1)$ acts transitively and hence two quasiregular representations of $SO_0(p-1, 1)$. And it is for just this reason that the decomposition $SO_0(p, 1) \supset SO_0(p-1, 1)$ occurs with multiplicity 2. We will show this explicitly in the next section.

We are now in a position to extend the representation (2.2) of $SO_0(p-1, 1)$ to a representation of

$SO_0(p,1)$: Let $T^{\sigma\rho}(g)$ be an operator on $\mathfrak{L}^2(H^{(p-1)})$ such that for each $g \in SO_0(p,1)$ and every $f \in \mathfrak{L}^2(H^{(p-1)})$ we have

$$T^{\sigma\rho}(g)f(\eta) = |(g^{-1})^{\nu}_{p}\eta_{\nu} + (g^{-1})^{p}_{p}|^{\sigma} \times f\left(\frac{(g^{-1})^{\nu}_{\mu}\eta_{\nu} + (g^{-1})^{p}_{\mu}}{(g^{-1})^{\nu}_{p}\eta_{\nu} + (g^{-1})^{p}_{p}}\right). \tag{2.3}$$

Again it is easily verified that $T^{\sigma\rho}(g)$ is a representation. By the transformation properties of the measure we see that $T^{\sigma\rho}(g)$ is unitary with respect to the inner product (2.1) extended to both sheets $H_{+}^{(p-1)}$ and $H_{-}^{(p-1)}$ if and only if

$$\sigma = -\tfrac{1}{2}(p-1) + i\rho, \qquad \rho \text{ real}.$$

We can, however, restrict ρ to be nonnegative since $T^{-\rho}$ and T^{ρ} are unitarily equivalent. It can also be verified that $T^{\sigma\rho}(g)$ is an irreducible representation of $SO_0(p,1)$. Thus Eq. (2.3) describes a unitary irreducible representation of the most degenerate continuous class of the group[12] $SO_0(p,1)$.

Before proceeding with the actual decomposition, a few words are in order concerning the infinitesimal generators of the group representation (2.3). As is known,[8] given a continuous unitary representation of a group, one can always obtain self-adjoint operators (infinitesimal generators) via the one-parameter subgroups. Performing this formally for the representation (2.3), we find

$$S_{\mu\nu} = i\left(\eta_{\mu}\frac{\partial}{\partial\eta^{\nu}} - \eta_{\nu}\frac{\partial}{\partial\eta^{\mu}}\right) \tag{2.4}$$

for the $SO(p-1,1)$ subgroup and

$$\Gamma_{\mu} = i\left(\frac{\partial}{\partial\eta^{\mu}} - \eta_{\mu}\eta^{\nu}\frac{\partial}{\partial\eta^{\nu}} + \sigma\eta_{\mu}\right) \tag{2.5}$$

for the remaining generators. Easily we arrive at the Lie algebra of $SO(p,1)$:

$$\begin{aligned}[S_{\mu\nu}, S_{\sigma\lambda}] &= i(g_{\nu\sigma}S_{\mu\lambda} - g_{\mu\sigma}S_{\nu\lambda} - g_{\nu\lambda}S_{\mu\sigma} + g_{\mu\lambda}S_{\nu\sigma}),\\ [S_{\mu\nu}, \Gamma_{\lambda}] &= i(g_{\nu\lambda}\Gamma_{\mu} - g_{\mu\lambda}\Gamma_{\nu}),\\ [\Gamma_{\mu}, \Gamma_{\nu}] &= iS_{\mu\nu}.\end{aligned} \tag{2.6}$$

However, it is not enough simply to write down the infinitesimal generators as differential operators on a Hilbert space in order to have a representation of a Lie algebra. One must know precisely the domain of definition of these operators. As an example of this, we see that the generators Γ_{μ} given above along with the $S_{\mu\nu}$ appear to leave the subspaces $\mathfrak{L}^2(H_{\pm}^{(p-1)})$ invariant.[13] Thus it appears that one can represent the generators of $SO(p,1)$ on the subspace $L^2(H_{+}^{(p-1)})$ alone. That this is only apparent can be seen readily from the global viewpoint as discussed above; but from the local viewpoint this information is buried within the domain of definition of the generators (2.4) and (2.5). The constraints mentioned previously for the functions $f^{\pm} \in \mathfrak{L}^2(H_{\pm}^{(p-1)})$, i.e., continuity at infinity, can be understood by breaking them up into even and odd functions. As long as the generators do not mix up the even and odd functions, the continuity problem remains the same and the subspaces $\mathfrak{L}^2(H_{\pm}^{(p-1)})$ remain separately invariant; this is the case for the generators $S_{\mu\nu}$. However, if the even and odd functions are mixed by the generators, as is the case for Γ_{μ}, the functions in the other subspace must also change in order to keep the proper continuity. Hence the infinitesimal generators Γ_{μ} do not leave $\mathfrak{L}^2(H_{\pm}^{(p-1)})$ separately invariant, as we have seen must be the case from the global considerations.

3. THE DECOMPOSITION $SO(p,1) \supset SO(p-1,1)$

We begin by introducing a system of coordinates on $H^{(p-1)}$ for which the Casimir operators of both the subgroups $SO(p)$ and $SO(p-1,1)$ separate. They are just the spherical coordinates on $H_{\pm}^{(p-1)}$ given by

$$\begin{aligned}\eta_0 &= \pm\cosh a,\\ \eta^1 &= \sinh a \sin\theta_{p-2}\cdots\sin\theta_1, \quad 0 \le \theta_1 < 2\pi,\\ \eta^2 &= \sinh a \sin\theta_{p-2}\cdots\sin\theta_2\cos\theta_1, \quad 0 \le \theta_i \le \pi,\\ &\cdots \quad i = 2,\cdots,p-2\\ \eta^{p-1} &= \sinh a \sin\theta_{p-2}\cos\theta_{p-3},\\ \eta^{p} &= \sinh a \cos\theta_{p-2}, \quad 0 \le a < \infty.\end{aligned} \tag{3.1}$$

The $SO(p-1,1)$ Casimir operator is just the Laplace–Beltrami operator on $H_{\pm}^{(p-1)}$:

$$\begin{aligned}-\Delta(H_{\pm}^{(p-1)}) &= \frac{1}{\sinh^{p-2}a}\frac{\partial}{\partial a}\sinh^{p-2}a\frac{\partial}{\partial a}\\ &+ \frac{1}{\sinh^2 a\sin^{p-3}\theta_{p-2}}\frac{\partial}{\partial\theta_{p-2}}\sin^{p-3}\theta_{p-2}\frac{\partial}{\partial\theta_{p-2}}\\ &+ \cdots\frac{1}{\sinh^2 a\sin^2\theta_{p-2}\cdots\sin^2\theta_2}\frac{\partial^2}{\partial\theta_1^2},\end{aligned} \tag{3.2}$$

or perhaps in the more convenient form

$$\begin{aligned}-\Delta(H_{\pm}^{(p-1)}) &= \frac{\partial^2}{\partial a^2} + (p-2)\coth a\frac{\partial}{\partial a}\\ &- \frac{1}{\sinh^2 a}\Delta(S^{(p-2)}),\end{aligned} \tag{3.3}$$

where $\Delta(S^{(p-2)})$ is the Laplace-Beltrami operator on the $(p-2)$-dimensional sphere. Now, to find the Casimir operator of the compact subgroup $SO(p)$, we use explicit expressions for $SO(p-1,1)$ Casimir operator,

$$\tfrac{1}{2}S_{\mu\nu}S^{\mu\nu} = \Gamma^2 + \sigma(\sigma+p-1) = -\Delta(H^{(p-1)}),$$

and the $SO(p)$ Casimir operator,

$C_{SO(p)} = \frac{1}{2}S_{ij}S^{ij} + \Gamma_i\Gamma_i$, with $i = 1, \cdots, p-2$,

to obtain

$$C_{SO(p)} = -\Delta(S^{(p-2)}) + \Delta(H^{(p-1)}) + \Gamma_0^2 - \sigma(\sigma + p - 1). \tag{3.4}$$

But we can find Γ_0 in the spherical system easily:

$$\Gamma_0 = \pm i\left(-\sinh a\frac{\partial}{\partial a} + \sigma\cosh a\right);$$

thus

$$-C_{SO(p)} = \cosh^2 a\frac{\partial^2}{\partial a^2} + \left((p-2)\coth a + (1-2\sigma)\sinh a\cosh a\right)\frac{\partial}{\partial a} + \sigma(\sigma-1)\sinh^2 a - (p-1)\sigma + \coth^2 a\,\Delta(S^{(p-2)}). \tag{3.5}$$

Of course, the eigenvectors of $\Delta(S^{(p-2)})$ are just the $(p-2)$-dimensional spherical harmonics[14] with eigenvalue $n_{p-2}(n_{p-2}+p-3)$. Hence, the eigenvalue problem for Eq. (3.5) can be solved yielding, aside from the multiplier factor, the polynomial solutions in $1/\cosh a$,

$$\cosh^{\sigma} a\tanh^{n_{p-2}} a\, C^{n_{p-2}+(p-2)/2}_{n_{p-1}-n_{p-2}}(\pm 1/\cosh a).$$

The factor $\cosh^{\sigma} a$ is a reflection of the fact that the measure on $H^{(p-1)}$ is not invariant under $SO(p)$ transformations, and the multiplier factor is just what is necessary to make the representation unitary.

The eigenvalue problem for $\Delta(H_{\pm}^{(p-1)})$ has a purely continuous spectrum; for it to be handled rigorously, one must venture outside the standard Hilbert space techniques. This can be done using Gel'fand's "rigged Hilbert space" approach[15]; however, nothing more will be said about this approach in the present paper except that we will consider the eigenfunctions of $\Delta(H_{\pm}^{(p-1)})$ to be δ-function normalizable. They are, aside from the normalization constant, given by

$$(\sinh a)^{-(p-3)/2}\, P^{-[n_{p-2}+(p-3)/2]}_{-1/2+i\nu}(\cosh a),$$

where $-[(p-2)/2] - \nu^2$ is the eigenvalue of $\Delta(H^{(p-1)})$. Also, for convenience in what follows we put $n_{p-1} = n$ and $n_{p-2} = l$. We can now write two complete orthonormal bases for $\mathcal{L}^2(H^{(p-1)})$: the canonical basis

$$f_{n_{p-1},N}(a, \theta_{p-2}, \cdots, \theta_1) \equiv \langle a, \theta_{p-2}, \cdots, \theta_1 | n_{p-1}, N\rangle = N_n \cosh^{\sigma-1} a\,\sinh^l a\, C^{l+(p-2)/2}_{n-l}(\pm 1/\cosh a) \times Y_N(\theta_{p-2}, \cdots, \theta_1) \tag{3.6}$$

and the pseudobasis,

$$\phi^{(\pm)}_{\nu,N}(a, \theta_{p-2}, \cdots, \theta_1) \equiv \langle a, \theta_{p-2}, \cdots, \theta_1 | \nu_{p-1}, N\rangle = N_\nu(\sinh a)^{-(p-3)/2}\, P^{-[l+(p-3)/2]}_{-1/2+i\nu}(\cosh a) \times Y_N(\theta_{p-2}, \cdots, \theta_1), \tag{3.7}$$

where N denotes $(n_{p-2}, \cdots, n_1)$ and $Y_N(\theta_{p-2}, \cdots, \theta_1)$ are the $(p-2)$-dimensional spherical harmonics. The normalization constants are given by

$$N_{nN} = N_n N_{n_{p-2}} = \frac{1}{\sqrt{2\pi}}\prod_{j=1}^{p-2}\Gamma\left(n_j + \frac{j}{2}\right)2^{n_j+(j-1)/2}\left[\frac{(n_{j+1}+j/2)\,\Gamma(n_{j+1}-n_j+1)}{\pi\Gamma(n_{j+1}+n_j+j)}\right]^{1/2}$$

and

$$\left[N_{\nu N} = N_\nu N_{n_{p-2}} = \left|\Gamma\left(i\nu + l + \frac{p-2}{2}\right)\right|^2\frac{\nu\sinh\pi\nu}{\pi}\right]^{1/2} N_{n_{p-2}}. \tag{3.8}$$

The orthonormality and completeness properties of these basis functions are well known[11]:

$$\int_{H_+ + H_-} d\Omega_p\, \bar{f}_{n',N'}(\Omega_p)\, f_{n,N}(\Omega_p) = \delta_{n'_{p-1},n_{p-1}}\cdots\delta_{n'_1,n_1}, \qquad \sum_{n_{p-1},\cdots,n_1} f_{n,N}(\Omega_p)\,\bar{f}_{n,N}(\Omega'_p) = \delta(\Omega_p - \Omega'_p)$$

$$\int_{H_+ + H_-} d\Omega_p\, \bar{\phi}^{(\tau)}_{\nu',N'}(\Omega_p)\,\phi^{(\tau)}_{\nu,N}(\Omega_p) = \delta_{\tau,\tau'}\,\delta(\nu-\nu')\delta_{n'_{p-2},n_{p-2}}\cdots\delta_{n'_1,n_1},$$

$$\sum_{\tau,n_{p-2},\cdots,n_1}\int_0^\infty d\nu\,\phi^{(\tau)}_{\nu,n}(\Omega_p)\,\bar{\phi}^{(\tau)}_{\nu,N}(\Omega'_p) = \delta(\Omega_p - \Omega'_p), \tag{3.9}$$

where the range of summation over the n_i is $n_1 = -n_2, \cdots n_2; n_{p-1} = 0, 1, \cdots$; and $n_i = 0, 1, \cdots$, $n_{i+1}, p = 2, \cdots, p-2$ if n_{i+1} is an integer; otherwise, the sequence extends to infinity. Using these relations, we can expand any function in $\mathfrak{L}^2(H^{(p-1)})$ in terms of either the canonical basis or pseudobasis

$$f(a, \theta_{p-2}, \cdots, \theta_1) = \sum_{n,N} a_{n,N} f_{n,N}(a, \theta_{p-2}, \cdots, \theta_1)$$

$$= \sum_N \sum_{\tau=-}^{+} \int_0^\infty d\nu\, b^{\tau}_{\nu,N}\, \phi^{(\tau)}_{\nu,N}(a, \theta_{p-2}, \cdots, \theta_1), \quad (3.10)$$

and the overlap functions

$$K^\rho_l(\nu, \pm, n) = \int_0^\infty \sinh^{p-2} a\, da\, (\sinh a)^{-(p-3)/2} \times P^{-[l+(p-3)/2]}_{-1/2+i\nu} \cosh^\sigma a \tanh^l a \times C^{l+(p-2)/2}_{n-l}(\pm 1/\cosh a) \quad (3.11)$$

allow us to go from one expansion to the other via

$$b^{(\tau)}_{\nu,N} = \sum_{n=0}^{\infty} a_{n,N} K^\sigma_l(\nu, \tau, n). \quad (3.12)$$

Moreover, by using the "overlap" functions (3.11), the transition between the canonical basis and pseudobasis for the matrix elements of group transformations can be accomplished. Denoting the matrix elements[16] of $SO(p, 1)$ in the canonical and pseudobasis, respectively,[17] by $T^{\rho\rho}_{n'N',nN}(g)$ and $T^{\rho\rho}_{\nu'N',\nu N}(g)$, we find, using the relations (3.9) and (3.11),

$$T^\rho_{\nu'N',\nu N}(g) = \sum_{n,n'} K^\rho_{l'}(\nu', \tau, n')\, T^{\rho\rho}_{n'N',nN}(g) \times \overline{K}^\rho_l(\nu, \tau, n) \quad (3.13)$$

and

$$T^{\rho\rho}_{n'N',nN}(g) = \sum_{\tau',\tau} \int d\nu' \int d\nu\, \overline{K}^\rho_{l'}(\nu', \tau', n') \times T^{\rho\rho}_{\nu'N',\nu N}(g) K^\rho_l(\nu, \tau, n). \quad (3.14)$$

By restricting g in Eq. (3.14) to the subgroup $SO(p-1, 1)$, we arrive at the basic formula for the decomposition $SO(p, 1) \supset SO(p-1, 1)$:

$$T^{\rho\rho}_{n'N',nN}(h) = \sum_\tau \int d\nu\, \overline{K}^\rho_{l'}(\nu, \tau, n')\, T^{\nu\rho-1}_{N'N}(h)\, K^\rho_l(\nu, \tau, n) \quad (3.15)$$

for $h \in SO(p-1, 1)$. In addition, taking g to be the identity in Eqs. (3 13) and (3.14), the orthonormality and completeness conditions for the "overlap" functions can be obtained:

$$\sum_\tau \int d\nu\, \overline{K}^\rho_l(\nu, \tau, n') K^\rho_l(\nu, \tau, n) = \delta_{n',n},$$
$$\sum_n K^\rho_l(\nu', \tau', n) \overline{K}^\rho_l(\nu, \tau, n) = \delta_{\tau',\tau}\delta(\nu' - \nu). \quad (3.16)$$

Note that in order to be able to write down equations like Eqs. (3. 13)–(3. 16), we must convince ourselves that the K functions have the proper convergence properties.

4. CALCULATION OF THE OVERLAP FUNCTIONS

First, it is easily seen from Eq. (3. 11) that

$$K^\rho_l(\nu, -, n) = (-1)^{n-l} K^\rho_l(\nu, +, n) \quad (4.1)$$

and

$$\overline{K}^\rho_l(\nu, \tau, n) = K^{-\rho}_l(-\nu, \tau, n) = K^{-\rho}_l(\nu, \tau, n). \quad (4.2)$$

To perform the integral (3. 11), we make the simple change of variables $x = \cosh a$; then

$$K^\rho_l(\nu, +, n) = N_\nu N_N \int_1^\infty dx (1-x^2)^{[l+(p-3)/2]/2} x^{\sigma-l} P^{-[l+(p-3)/2]}_{-1/2+i\nu}(x) C^{l+(p-2)/2}_{n-l}(1/x) \quad (4.3)$$

and expand the Gegenbauer polynomials as

$$C^{l+(p-2)/2}_{n-l}\left(\frac{1}{x}\right) = \frac{1}{\Gamma[l + \frac{1}{2}(p-2)]} \sum_{j=0}^{\{(n-l)/2\}} (-1)^j \frac{\Gamma(n+1-j)}{\Gamma(n+1-l-2j)j!}\left(\frac{2}{x}\right)^{n-l-2j}, \quad (4.4)$$

which yields

$$\frac{N_\nu N_n}{\Gamma(l + \frac{1}{2}(p-2))} \sum_{j=0}^{\{(n-l)/2\}} \frac{(-1)^j}{j!} \frac{\Gamma(n+1-j)}{\Gamma(n-l-2j)} 2^{n-l-2j} \int_1^\infty dx\, x^{\sigma-n+2j}(1-x^2)^{[l+(p-3)/2]/2} P^{-[l+(p-3)/2]}_{-1/2+i\nu}(x), \quad (4.5)$$

where $\{\alpha\}$ = largest integer $\leq \alpha$. The integral (4. 5) can be performed with the aid of formula 18. 2. 3 of Ref. 18, yielding[19]

$$K^\rho_l(\nu, +, n) = \sqrt{\pi}\, \frac{N_\nu N_n 2^{-l-(p-2)/2}}{\Gamma(l + \frac{1}{2}(p-2))} \sum_{j=0}^{\{(n-l)/2\}} \frac{(-1)^j}{j!} \times \frac{\Gamma(n+1-j)\Gamma(\frac{1}{2}(n-l+\frac{1}{2}-i\rho-i\nu)-j)\Gamma(\frac{1}{2}(n-l+\frac{1}{2}-i\rho+i\nu)-j)}{\Gamma(\frac{1}{2}(n-l+1)-j)\Gamma(\frac{1}{2}(n-l+2)-j)\Gamma(\frac{1}{2}(n+\frac{1}{2}(p-1)-i\rho)-j)\Gamma(\frac{1}{2}(n+\frac{1}{2}(p+1)-i\rho)-j)}. \quad (4.6)$$

Equation (4.6) can be expressed in terms of the generalized hypergeometric functions[20]

$$ {}_pF_q\left[\begin{matrix}\alpha_1, \cdots, \alpha_p \\ \beta_1, \cdots, \beta_q\end{matrix}; z\right] = \sum_{n=0}^{\infty} \frac{(\alpha_1)_n \cdots (\alpha_p)_n}{(\beta_1)_n \cdots (\beta_q)_n} \frac{z^n}{n!}, \qquad (a)_n = \frac{\Gamma(a+n)}{\Gamma(a)} $$

by using the relation

$$ \Gamma(a-n) = \frac{(-1)^n \Gamma(a)}{(1-a)_n} $$

to obtain

$$ K_l^{\rho}(\nu, +, n) = \frac{\sqrt{\pi} N_\nu N_n}{2^{l+(p-3)/2}\Gamma(l+\frac{1}{2}(p-2))} $$

$$ \times \frac{\Gamma(n+1)\Gamma(\frac{1}{2}(n-l+\frac{1}{2}-i\rho-i\nu))\,\Gamma(\frac{1}{2}(n-l+\frac{1}{2}-i\rho-i\nu))}{\Gamma(\frac{1}{2}(n-l+1))\Gamma(\frac{1}{2}(n-l+2))\Gamma(\frac{1}{2}[n+\frac{1}{2}(p-1)-i\rho])\Gamma(\frac{1}{2}[n+\frac{1}{2}(p+1)-i\rho])} $$

$$ \times {}_4F_3\left\{\begin{matrix}-\frac{1}{2}(n-l-1), -\frac{1}{2}(n-l), -\frac{1}{2}[n+\frac{1}{2}(p-5)-i\rho], -\frac{1}{2}[n+\frac{1}{2}(p-3)-i\rho] \\ -n, -\frac{1}{2}(n-l+\frac{3}{2}-i\rho-i\nu), -\frac{1}{2}(n-l+\frac{3}{2}-i\rho+i\nu)\end{matrix}; 1\right\}. \tag{4.7} $$

Although the integral (4.5) does not converge for all complex ν, either Eq. (4.6) or (4.7) gives an analytic continuation of the overlap functions into the complex ν plane. And indeed Eq. (4.6) provides a very convenient form in which to read off the analytic structure. We see that, aside from the square root branch points contained in the N_ν factor, there are the following sequences of poles in the variable $\sigma(p-1) = -\frac{1}{2}(p-2) + i\nu$:

$$ \text{for } n-l \text{ even:} \begin{cases}\sigma(p-1) = -i\rho + 2k - \frac{1}{2}(p-3), \\ \sigma(p-1) = i\rho - 2k - \frac{1}{2}(p-1);\end{cases} $$

$$ \text{for } n-l \text{ odd:} \begin{cases}\sigma(p-1) = -i\rho + (2k+1) - \frac{1}{2}(p-3), \\ \sigma(p-1) = i\rho - (2k+1) - \frac{1}{2}(p-1);\end{cases} $$

$$ k = 0, 1, \cdots. \tag{4.8} $$

Thus, in the case $p = 3$, we have essentially one-half the poles of Sciarrino and Toller. This corresponds to the fact that, for Toller $M = 0$ poles, the Sciarrino–Toller residues vanish for those poles omitted in Eq. (4.8). In fact, for the degenerate continuous representations ($M = 0$) this is the total content of Eq. (5.24) of Ref. 4. Hence, we expect the factorization of generalized Regge residues. More specifically, an $SO(p,1)$ Regge pole with factorizable residue decomposes into a series of $SO(p-1,1)$ Regge poles, each with factorizable residue.

It is easy to cast the overlap functions into a form closely analogous to that of Sciarrino and Toller by using a different expansion for the Gegenbauer polynomials in Eq. (4.3). By using the expansion

$$ C_{n-l}^{l+(p-2/2)}(z) = \frac{1}{\Gamma(l+\frac{1}{2}(p-2))} \times \sum_{j=0}^{n-l} \frac{\Gamma(l+\frac{1}{2}(p-2)+j)}{\Gamma(2l+p-2+2j)} \times \frac{\Gamma(n+l+\frac{1}{2}p+j)\,2^{-j}(1-z)^j}{\Gamma(n-l+1-j)j!}, $$

performing the integration with the aid of formula 18.2.10 of Ref. 18, and expressing the result in terms of Meijer's G functions,[20] we can write

$$ K_l^{\sigma}(\nu, +, n) = \frac{\cos\pi i\nu N_\nu N_n}{\pi\Gamma[l+\frac{1}{2}(p-2)]} \sum_{j=0}^{n-l} (-1)^j \frac{\Gamma[l+\frac{1}{2}(p-2)+j]\Gamma(n+l+\frac{1}{2}p+j)2^{-j}}{\Gamma(2l+p-2-2j)\Gamma(n-l+1-j)j!\Gamma(l+j-\sigma)} $$

$$ \times G_{33}^{23}\left\{\tfrac{1}{2}\middle|\begin{matrix}-l-\frac{1}{2}(p-3)-j, & \frac{1}{2}+i\nu, & \frac{1}{2}-i\nu \\ -\sigma-\frac{1}{2}(p-1), & 0, & -l-\frac{1}{2}(p-3)\end{matrix}\right\}. \tag{4.9} $$

The poles are obtained easily by using the Mellin–Barnes integral representation yielding poles at

$$ \sigma(p-1) = -i\rho + k - \tfrac{1}{2}(p-3), $$

$$ \sigma(p-1) = i\rho - k - \tfrac{1}{2}(p-1). $$

Comparing this with Eq. (4.8), we see that the residues at the poles in Eq. (4.9) must vanish when $n-l+k$ is odd.

5. ASYMPTOTIC BEHAVIOR

As mentioned previously, in order that the decomposition (3.15) be meaningful, we must assure ourselves that the integral in this decomposition converges.

The asymptotic behavior of the K functions in the variable ν can be obtained from Eqs. (4.6) and (3.8). The Γ functions in Eq. (4.6) have the following asymptotic behavior[19]:

$$\Gamma(x+iy) < |\Gamma(x+iy)| \xrightarrow{|y|\to\infty} (2\pi)^{1/2} \times |y|^{x-1/2} e^{-(\pi/2)|y|},$$

while $N_\nu \to |\nu|^{l+(p-3)/2}$. Thus, for fixed n, l, and ρ

$$N_\nu \Gamma\left(\tfrac{1}{2}(n-l+\tfrac{1}{2}-i\rho-i\nu)\right)\Gamma\left(\tfrac{1}{2}(n-l+\tfrac{1}{2}-i\rho+i\nu)\right) < (2\pi)|\nu|^{n+p-3} e^{-\pi|\nu|}.$$

Hence, $K_l^\rho(\nu,\tau,n)$ dies off as ν goes to infinity faster than any negative power. But, from Eq. (2.13) of Ref. 16, the only dependence of $T^{\nu\rho-1}_{N'N}(\alpha)$ on ν aside from the oscillating factor $e^{-i\alpha\nu}$ is through the hypergeometric function:

$${}_2F_1(n'_{p-2}+\tfrac{1}{2}(p-2)-i\nu; n_{p-3}+\tfrac{1}{2}(p-2)+k+k'; n'_{p-2}+n_{p-2}+p-2; 1-e^{-2\alpha}),$$

which at worst dies off like a negative power as $\nu\to\infty$. Thus, the integral in Eq. (2.13) converges, and the decomposition is well defined.

Finally, we can analytically continue $T^{\rho\rho}_{N'N,nN}(g)$ in the complex plane of the variable ρ to obtain a generalization for $SO(p,1)$ of the usual Regge pole procedure.[4,21] When this is done, any of the poles of the integrand in Eq. (3.15) that cross the integration path become the Regge poles and a Regge expansion is obtained. Of course, by examining the asymptotic behavior of $T^{\nu}_{N'N}(\alpha)$ as $\alpha\to+\infty$, we can never hope to get the integrand of the "background" integral any better than

$$e^{-\alpha[(p-2)/2]}$$

for any complex ν. To see this, we write, using one of Kummer's identities[20] and Eq. (2.14) of Ref. 16,

$$T^{\nu}_{NN'} \sim \lambda_1 e^{-[(p-2)/2]\alpha+i\nu\alpha} + \lambda_2 e^{-[(p-2)/2]\alpha-i\nu\alpha} \quad \text{as} \quad \alpha\to\infty.$$

In order to get a complete Regge asymptotic expansion, one must deal with the second-kind functions. A general procedure which can be applied in our case for finding these functions is given in Ruhl's book.[1]

ACKNOWLEDGMENTS

One of us (C.P.B.) would like to acknowledge Dr. G. N. Fleming for valuable discussion and criticism and Dr. S. Ström for a communication pointing out Ref. 7.

---

* Present address: Arya-Mehr University of Technology, Tehran, Iran.

1 Books on the subject include I. M. Gel'fand, R. A. Minlos, and Z. Ya. Shapiro, *Representations on the Rotation and Lorentz Groups and Their Applications* (Pergamon, New York, 1963); I. M. Gel'Fand and M. A. Naimark, *Unitäire Darstellungen der Klassischen Gruppen* (Akademie-Verlag, Berlin 1957); M. A. Naimark, *Linear Representations of the Lorentz Group* (Pergamon, New York, 1964); W. Rühl, *The Lorentz Group and Harmonic Analysis* (Benjamin, New York, 1970); G. W. Mackey, *Induced Representations of Groups and Quantum Mechanics* (Benjamin, New York, 1967); N. Ya. Vilenkin, *Special Functions and the Theory of Group Representations* (Am. Math. Soc., Providence, R.I., 1968); well-known articles are V. Bargmann, Ann. Math **48**, 568 (1947); Harish-Chandra, Proc. Roy, Soc. (London) **A189**, 272 (1947); L. H. Thomas, Ann. Math. **42**, 113 (1941); T. D. Newton, *ibid.* **51**, 73 (1950); J. Dixmier, Bull. Soc. Math. (France) **89**, 9 (1961).

2 R. Raczka, N. Limic, and J. Niederle, J. Math. Phys. **7**, 1861 (1966); **7**, 2026 (1966); **8**, 1079 (1967). For the case $SO(4,2)$ A. O. Barut and A. Bohm, University of Texas Report (unpublished).

3 U. Ottoson, Commun. Math. Phys. **8**, 228 (1968); J. Dixmier, Compt. Rend. **250**, 3263 (1960).

4 A. Sciarrino and M. Toller, J. Math. Phys. **8**, 1252 (1967).

5 W. Rühl, Ref. 1; N. Mukunda, J. Math. Phys. **8**, 2210 (1967); **9**, 50 (1968); S. Ström, Arkiv Fysik **34**, 215 (1967); Ann. Inst. Henri Poincaré **13A**, 77 (1970); N. W. MacFadyen, Carnegie-Mellon University Report (unpublished).

6 Representations of the most degenerate type have been discussed in N. Ya Vilenkin, Ref. 1, Ref. 2, and M. Bander and C. Itzykson, Rev. Mod. Phys. **38**, 330, 346 (1966).

7 N. Limic and J. Niederle, Ann. Inst. Henri Poincaré **9A**, 327 (1968).

8 S. Helgason, *Differential Geometry and Symmetric Spaces* (Academic, New York, 1967).

9 I. M. Gel'fand, M. I. Graev, and N. Ya. Vilenkin, *Generalized Functions* (Academic, New York, 1966), Vol. 5; N. Ya. Vilenkin and Ya. A. Smorodinski, Zh. Eksp. Teor. Fiz. **46**, 1793 (1964) [Sov. Phys. JETP **19**, 1209 (1964)].

10 N. Ya. Vilenkin, Ref. 1.

11 M. Bander and C. Itzykson, Ref. 6.

12 In fact, we have constructed unitary irreducible representations of the complete group $O(p,1)$, but for our purposes we need only consider the connected component $SO(p,1)$.

13 In this regard see also N. Mukunda, Ref. 5.

14 A Erdélyi *et al.*, *Higher Transcendental Functions* (McGraw-Hill, New York, 1953), Vol. 2.

15 I. M. Gel'fand and N. Ya. Vilenkin, *Generalized Functions* (Academic, New York, 1964), Vol. 4. For an application of these techniques in the case $SU(1,1)$ see G. Lindblad and B. Nagel, Ann. Inst. Henri Poincaré **13A**, 27 (1970).

16 C. P. Boyer, J. Math. Phys. **12**, No. 8 (1971).

17 We will use the notation for the matrix elements $T^{\sigma\rho}_{n'N',nN}$ and $T^{\rho\rho}_{n'N',nN}$ interchangably, where $\sigma = -\frac{1}{2}(p-1)+i\rho$ is understood. The same should be understood for the functions $K_l^\sigma(\nu,\tau,n)$.

18 A. Erdélyi *et al.*, *Tables of Integral Transforms* (McGraw-Hill, New York, 1954), Vol. 2.

19 This expression is similar to the one obtained for $SO(3,1)$ in D. A. Akyeampong, J. F. Boyce, and M. A. Rashid, Nuovo Cimento **53**, 737 (1968).

20 Reference 14, Vol. 1.

21 R. Delbourgo, K. Koller, and R. M. Williams, J. Math. Phys. **10**, 957 (1969). K. B. Wolf, Tel Aviv University Report (unpublished).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 10 OCTOBER 1971

The T Matrix for the Hulthén Potential

O. P. Bahethi and M. G. Fuda
Department of Physics and Astronomy, State University of New York at Buffalo, Buffalo, New York 14214
(Received 23 April 1971)

A closed-form expression for the s wave part of the Hulthén potential's T matrix is derived. This expression is used to obtain a simpler result for the terms in the Weinberg series for the T matrix than the results given previously by other authors.

1. INTRODUCTION

Recently[1] one of us (M.G.F) was able to derive the completely off-shell, two-particle T matrix for the exponential potential. The derivation was based on the differential equation approach used by Van Leeuwen and Reiner[2] to obtain the T-matrix for a potential consisting of a chain of rectangular wells. The inhomogeneous differential equation that arises in this approach is similar to the well-known Bethe–Goldstone[3] equation.

In this paper we show that with some modification the differential equation approach can be used to find the fully off-shell, two-particle T matrix for the Hulthén potential.[4] The modification amounts to obtaining an expression for the T matrix which does not involve the potential explicitly, thereby avoiding certain integrals that are difficult to do because of the $1/r$ singularity in the Hulthén potential. Using our closed-form expression for the Hulthén T matrix we are also able to find a simpler expression for the Weinberg series[5] for the T matrix than the expressions given previously by other authors[6]. The Weinberg series[5] is a separable expansion of the two-particle T matrix which has been shown to be of practical value in calculations on three-particle systems[6,7].

2. THE T MATRIX FOR THE HULTHEN POTENTIAL

The two-particle T matrix is the solution of either of the equations

$$T(s) = V + VG_0(s)T(s), \tag{2.1}$$

$$T(s) = V + T(s)G_0(s)V, \tag{2.2}$$

where V is the two-particle potential, s is a complex parameter, and $G_0(s)$, the free particle resolvent, is defined formally in terms of the kinetic energy operator H_0 by

$$G_0(s) = (s - H_0)^{-1}. \tag{2.3}$$

Unless stated otherwise, we will assume that s has a small positive imaginary part, thereby guaranteeing the correct outgoing wave boundary condition, i.e.,

$$s = E + i\epsilon, \quad 0 < \epsilon \ll 1. \tag{2.4}$$

Following Van Leeuwen and Reiner[2], we define

$$\Omega(s) = 1 + G_0(s)T(s). \tag{2.5}$$

Using (2.1) and (2.5) we obtain

$$T(s) = V\Omega(s) \tag{2.6}$$

and

$$(s - H_0 - V)\Omega(s) = s - H_0, \tag{2.7}$$

which we write out in a mixed representation, i.e.,

$$[s + \nabla^2 - V(r)]\langle \mathbf{r}|\Omega(s)|qlm\rangle = (s - q^2) \times \langle \mathbf{r}|qlm\rangle, \tag{2.8}$$

where

$$\langle \mathbf{r}|qlm\rangle = (2\pi^2)^{-1/2} j_l(qr) Y_{lm}(\hat{r}). \tag{2.9}$$

$j_l(qr)$ is the usual spherical Bessel function and $Y_{lm}(\hat{r})$ is a spherical harmonic. We are working in units in which $\hbar^2/2m$ is 1. Since the potential is central, we can write

$$\langle \mathbf{r}|\Omega(s)|qlm\rangle = (2\pi^2)^{-1/2}\Omega_l(r, q; s)Y_{lm}(\hat{r}), \tag{2.10}$$

which upon substitution into (2.8) gives us

$$\left(s + \frac{1}{r}\frac{d^2}{dr^2}r - \frac{l(l+1)}{r^2} - V(r)\right)\Omega_l(r, q; s) = (s - q^2)j_l(qr). \tag{2.11}$$

The solution of (2.8) or (2.11) will be referred to in this paper as the off-shell wavefunction. If one has an expression for the off-shell wavefunction, it can be inserted in (2.6) in order to obtain the T matrix. When we attempted to do this for the Hulthén potential,[4] we found that the integrals encountered were difficult to do because of the $1/r$ singularity in the potential. An expression for the T matrix which does not involve the potential explicitly can be obtained from (2.5).

If we write out (2.5) in the representation given by (2.9) and use (2.9) and (2.10), we obtain

$$(s - p^2)^{-1}\langle plm|T(s)|qlm\rangle = (2\pi^2)^{-1}\int_0^\infty r^2 dr j_l(pr) \times [\Omega_l(r, q; s) - j_l(qr)]. \tag{2.12}$$

From (2.11) it follows that outside the range of the force,

$$\Omega_l(r, q; s) = j_l(qr) + A_l(q, k)h_l^{(+)}(kr), \tag{2.13}$$

where

$$s = k^2, \quad \mathrm{Im}k > 0, \tag{2.14}$$

and $h_l^{(+)}$ is a spherical Hankel function as defined

in Messiah.[8] This Hankel function satisfies the outgoing wave boundary condition implied by (2.4). Using (2.14) it is easy to show that

$$\int_0^\infty r^2 dr j_l(pr) h_l^{(+)}(kr) = -k^{-1}(s-p^2)^{-1}(p/k)^l. \tag{2.15}$$

Combining (2.12) and (2.15) we arrive at

$$\langle plm|T(s)|qlm\rangle = -(2\pi^2 k)^{-1}(p/k)^l A_l(q,k) + (2\pi^2)^{-1}(k^2-p^2)\int_0^\infty r^2 dr j_l(pr)[\Omega_l(r,q;s) - j_l(qr) - A_l(q,k)h_l^{(+)}(kr)]. \tag{2.16}$$

By setting $p=k$ in (2.16), it follows that A_l is the half-off-shell T matrix. Thus (2.13) and (2.16) become

$$\Omega_l(r,q;s) = j_l(qr) - 2\pi^2 k\langle klm|T(s)|qlm\rangle h_l^{(+)}(kr), \tag{2.17}$$

$$\langle plm|T(s)|qlm\rangle = (p/k)^l\langle klm|T(s)|qlm\rangle + (2\pi^2)^{-1}(k^2-p^2)\int_0^\infty r^2 dr j_l(pr)[\Omega_l(r,q;s) - j_l(qr) + 2\pi^2 k\langle klm|T(s)|qlm\rangle h_l^{(+)}(kr)]. \tag{2.18}$$

Clearly (2.17) is a generalization of the well-known result that the coefficient of the outgoing wave in the ordinary Schrödinger wavefunction is proportional to the partial wave scattering amplitude. (2.18) is a generalization of the relation[9] between the half-off-shell T matrix ($q=k$), the on-shell T matrix ($p=q=k$) and the ordinary Schrödinger wavefunction.

We now turn our attention to finding the $l=0$ component of the off-shell wavefunction for the Hulthén potential,[4] which is given by

$$V(r) = V_0 e^{-r/a}/(1-e^{-r/a}). \tag{2.19}$$

If we insert (2.19) in (2.11), set $l=0$, let

$$\Omega_0(r,q;s) = u(r,q;s)/qr, \tag{2.20}$$

$$z = e^{-r/a}, \tag{2.21}$$

and

$$u = e^{ikr}w, \tag{2.22}$$

we arrive at

$$z(1-z)\frac{d^2w}{dz^2} + [C-(A+B+1)z]\frac{dw}{dz} - ABw = \frac{k^2a^2-q^2a^2}{2i}\left[z^{i(ka-qa)-1} - z^{i(ka+qa)-1} - z^{i(ka-qa)} + z^{i(ka+qa)}\right], \tag{2.23}$$

where

$$A(k) = -ika + ia(k^2+V_0)^{1/2},$$
$$B(k) = -ika - ia(k^2+V_0)^{1/2},$$
$$C(k) = 1 - 2ika. \tag{2.24}$$

The solution of (2.23) can be obtained from an inhomogeneous hypergeometric differential equation studied by Babister,[10] namely,

$$z(1-z)\frac{d^2y}{dz^2} + [C-(A+B+1)z]\frac{dy}{dz} - ABy = z^{\sigma-1}. \tag{2.25}$$

A particular integral of (2.25) is the function[10]

$$f_\sigma(A,B;C;z) = \frac{z}{\sigma(\sigma+C-1)} \times {}_3F_2(1,\sigma+A,\sigma+B;\sigma+1,\sigma+C;z), \tag{2.26}$$

where ${}_3F_2$ is a special case of the generalized hypergeometric function defined by

$${}_mF_n(\alpha_1,\alpha_2,\cdots,\alpha_m;\beta_1,\beta_2,\cdots,\beta_n;z) = \sum_{k=0}^{\infty}\frac{(\alpha_1)_k(\alpha_2)_k\cdots(\alpha_m)_k}{(\beta_1)_k(\beta_2)_k\cdots(\beta_n)_k}\frac{z^k}{k!}. \tag{2.27}$$

The series in (2.26) converges when $|z|<1$; it converges when $|z|=1$ provided that $\mathrm{Re}(C-A-B)>0$, which from (2.24) is true in our case.

If one uses the relation[11]

$$(\sigma+A)(\sigma+B)f_{\sigma+1}(A,B;C;z) = \sigma(\sigma+C-1)\times f_\sigma(A,B;C;z) - z^\sigma, \tag{2.28}$$

it follows from (2.21)–(2.23), (2.25), and (2.26) that the reduced off-shell wavefunction is given by

$$u(r,q;s) = \sin qr - 2\pi^2 qT(k,q;s)e^{ikr}F(A,B;C;e^{-r/a}) + \frac{e^{ikr}V_0a^2}{2i}[f_{1+i(ka-qa)}(A,B;C;e^{-r/a}) - f_{1+i(ka+qa)}(A,B;C;e^{-r/a})], \tag{2.29}$$

where F is the hypergeometric function [(2.27) with $m=2$ and $n=1$] and

$$T(k,q;s) = \langle k00|T(s)|q00\rangle = \frac{V_0a^2}{4\pi^2 qi} \times \frac{f_{1+i(ka-qa)}(A,B;C;1) - f_{1+i(ka+qa)}(A,B;C;1)}{F(A,B;C;1)}. \tag{2.30}$$

The off-shell function u satisfies the boundary condition $u(o,q;s)=0$ and the boundary condition at large r implied by (2.17) and (2.20).

It is of interest to see how the off-shell wavefunction u goes over into the ordinary Schrödinger wavefunction when $q=k$. From (2.26)–(2.28) it follows that

$$ABf_1(A,B;C;z) = F(A,B;C;z) - 1. \tag{2.31}$$

If one uses (2.31), the relation [12]

$$f_{\sigma+1-C}(A,B;C;z) = z^{1-C} f_\sigma(A+1-C, B+1-C; 2-C; z), \tag{2.32}$$

with $\sigma = 1$, (2.24), and (2.30), it is straightforward to show that

$$u(r,k;s) = [2if(k)]^{-1}[f(-k)f(k,r) - f(k)f(-k,r)], \tag{2.33}$$

where the Jost function[13]

$$f(k) = F(A,B;C;1), \tag{2.34}$$

and the irregular solution of the Schrödinger equation[13]

$$f(k,r) = e^{ikr} F(A,B;C;e^{-r/a}). \tag{2.35}$$

Using[13]

$$f(k) = |f(k)| e^{-i\delta(k)} \tag{2.36}$$

and

$$f(-k) = |f(k)| e^{i\delta(k)}, \quad k \text{ real}, \tag{2.37}$$

we see that it follows from (2.30)–(2.32), and (2.24) that our on-shell T matrix is related to the phase shift δ by

$$T(k,k;s) = -(2\pi^2 k)^{-1} e^{i\delta(k)} \sin\delta(k). \tag{2.38}$$

We now show that it is possible to write the half-off-shell T matrix given by (2.30) in closed form in terms of Γ functions. Using the results of Secs. 3.5 and 3.6 of Ref. 14 one can show that

$$[\Gamma(s)\Gamma(e)\Gamma(f)]^{-1} {}_3F_2(a,b,c;e,f;1) = [\Gamma(f-a)\Gamma(s+a)\Gamma(e)]^{-1} \times {}_3F_2(a, e-b, e-c; s+a, e; 1),$$

where

$$s = e + f - a - b - c. \tag{2.39}$$

From the series (2.27) for ${}_3F_2$ it follows that

$${}_3F_2(1,a,b;2,c;z) = \frac{(c-1)}{z(a-1)(b-1)} [F(a-1, b-1; c-1; z) - 1]. \tag{2.40}$$

If now we use (2.24), (2.39), (2.40), and the well-known result[15]

$$F(a,b;c;1) = \Gamma(c)\Gamma(c-a-b)/\Gamma(c-a)\Gamma(c-b), \tag{2.41}$$

it is straightforward to show from (2.30) that

$$T(k,q;s) = I(k,q) + I(k,-q), \tag{2.42}$$

where

$$I(k,q) = -\frac{1}{4\pi^2 iq f(k)} \times \frac{\Gamma(1+iqa-ika)\Gamma(1+iqa+ika)}{\Gamma(1+iqa-ika-A)\Gamma(1+iqa-ika-B)}. \tag{2.43}$$

This result for the half-off-shell T matrix agrees with the result obtained from a different approach by Ford.[16]

In order to obtain the fully off-shell T matrix it is necessary to combine (2.29) and (2.20), and insert the result into (2.18), with $l = 0$. By using (2.31) it is easy to show that

$$\begin{aligned} u(r,q;s) - \sin qr + 2\pi^2 q T(k,q;s) e^{ikr} &= V_0 a^2 e^{ikr}\{ -2\pi^2 q T(k,q;s) f_1(A,B;C;e^{-r/a}) \\ &+ (2i)^{-1}[f_{1+i(ka-qa)}(A,B;C;e^{-r/a}) \\ &- f_{1+i(ka+qa)}(A,B;C;e^{-r/a})]\}. \end{aligned} \tag{2.44}$$

If now one makes the change of variable given by (2.21) in the integral arising in (2.18), all of the integrations can be carried out by using the result

$$\begin{aligned} \int_0^1 dz\, z^{\rho-1} f_\sigma(a,b;c;z)dz &= [\sigma(\sigma+c-1)(\rho+\sigma)]^{-1} \\ &\times {}_4F_3(1, \sigma+a, \sigma+b, \rho+\sigma; \sigma+1, \sigma+c, \rho+\sigma+1; 1). \end{aligned} \tag{2.45}$$

This result is obtained by integrating the series for f_σ [see (2.26) and (2.27)] term by term. To get our final result for the T matrix we also use

$${}_4F_3(1, 1+a, 1+b, 1+c; 2, 1+e, 1+f; z) = (ef/abcz)[{}_3F_2(a,b,c;e,f;z) - 1], \tag{2.46}$$

which follows directly from (2.27). The final result for the Hulthén T matrix is given by the relations

$$\rho = ia(p-k), \tag{2.47}$$

$$\sigma = 1 + ia(k-q), \tag{2.48}$$

$$X(p,k) = \rho^{-1} {}_3F_2(A,B,\rho;C,\rho+1;1), \tag{2.49}$$

$$Y(p,q;k) = \sigma^{-1}(\sigma+C-1)^{-1}(\rho+\sigma)^{-1} \times {}_4F_3(1, \sigma+A, \sigma+B, \sigma+\rho; \sigma+1, \sigma+C, \rho+\sigma+1; 1), \tag{2.50}$$

$$\begin{aligned} T(p,q;s) = \frac{(k^2-p^2)a}{4\pi^2 ipq}\{ &2\pi^2 q T(k,q;s) \\ &\times [X(p,k) - X(-p,k)] - (V_0 a^2/2i) \\ &\times [Y(p,q,k) - Y(p,-q,k) - Y(-p,q,k) \\ &+ Y(-p,-q,k)]\}. \end{aligned} \tag{2.51}$$

We have performed a couple of checks on this fairly complicated result. If one sets $p = k$ in (2.51) and uses (2.49) and (2.27), it is easy to see that one gets the correct limit. The other check we have carried out is to set $q = k$ and see if we got the result obtained from (2.42) and (2.43), using the property

$$T(p, q; s) = T(q, p; s). \tag{2.52}$$

In order to carry out this check, we use (2.24), (2.34), (2.36), (2.37), (2.38), (2.46), and (2.47) to show that

$$T(p, k; s) = -\frac{(k^2 - p^2)a}{8\pi^2 pkf(k)} \{ -f(-k)[X(p, k) - X(-p, k)] + f(k)[X(p, -k) - X(-p, -k)]\}. \tag{2.53}$$

From Eq. (3) in Sec. 3.7 of Ref. 13, (2.24), (2.34), and (2.47), it follows that

$$f(-k)X(p, k) - f(k)X(p, -k) = \frac{2k}{ia(p^2 - k^2)} \frac{\Gamma(1 + \rho)\Gamma(2 + \rho - C)}{\Gamma(1 + \rho - B)\Gamma(1 + \rho - A)}. \tag{2.54}$$

Substituting (2.54) in (2.53), using (2.52), and comparing with (2.42) and (2.43), we find that (2.51) has the right limit when $q = k$.

3. THE WEINBERG SERIES

In this section we show how a simple expression for the terms in the Weinberg series[5] for the Hulthén T matrix can be obtained from our expression (2.51). The Weinberg series[5] for the T matrix $T(s; \lambda)$ arising from the potential λV is given by

$$T(s; \lambda) = \sum_\nu \frac{V|\psi_\nu(s)\rangle\langle\psi_\nu(s^*)|V}{\langle\psi_\nu(s^*)|V|\psi(s)\rangle} \frac{\lambda}{1 - \lambda\eta_\nu(s)}, \tag{3.1}$$

where the $|\psi_\nu\rangle$ and η_ν are the eigenfunctions and eigenvalues of the kernel of the Lippmann–Schwinger equation; i.e.,

$$G_0(s)V|\psi_\nu(s)\rangle = \eta_\nu(s)|\psi_\nu(s)\rangle. \tag{3.2}$$

From (3.1) it is obvious that

$$\frac{V|\psi_\nu(s)\rangle\langle\psi_\nu(s^*)|V}{\langle\psi_\nu(s^*)|V|\psi_\nu(s)\rangle} = \eta_\nu^2(s)\lim_{\lambda\to\eta_\nu^{-1}}[\eta_\nu^{-1}(s) - \lambda]T(s; \lambda). \tag{3.3}$$

Thus we can obtain the terms in the separable expansion (3.1) directly from the T matrix.

The s-wave eigenvalues for the Hulthén potential (2.19) are given by[13]

$$\eta_\nu(s) = V_0a^2/\nu(2iak - \nu), \qquad \nu = 1, 2, 3, \cdots. \tag{3.4}$$

In our formula (2.51) for the T matrix we replace V_0 everywhere by λV_0, so, for example, A [see (2.24)] becomes

$$A(k, \lambda) = -ika + ia(k^2 + \lambda V_0)^{1/2}. \tag{3.5}$$

From (3.3)–(3.5) it follows that for the Hulthén potential

$$\frac{V|\psi_\nu(s)\rangle\langle\psi_\nu(s^*)|V}{\langle\psi_\nu(s^*)|V|\psi_\nu(s)\rangle} = \frac{(2ika - 2\nu)}{V_0a^2}\eta_\nu^2(s)\lim_{A\to -\nu}[A(k, \lambda) + \nu]T(s; \lambda). \tag{3.6}$$

Using (2.34), (2.41)–(2.43), (2.49)–(2.51), and the relation

$$\lim_{z\to n}(z + n)\Gamma(z) = (-1)^n/n!, \qquad n = 0, 1, 2, \cdots, \tag{3.7}$$

we can show that

$$\lim_{A\to\nu}[A(k, \lambda) + \nu]\, T(p, q; s; \lambda) = \frac{-(k^2 - p^2)a(C)_\nu}{8\pi^2 pq(\nu - 1)!}[W_\nu(q, k) - W_\nu(-q, k)] \times [X(p, k; \lambda) - X(-p, k; \lambda)]_{A=-\nu}, \tag{3.8}$$

where

$$W_\nu(q, k) = \frac{(-iqa - ika)_\nu}{(1 + iqa - ika)_\nu}. \tag{3.9}$$

From the results of Secs. 3.5 and 3.6 of Ref. 13, it follows that

$$X(p, k; \lambda)|_{A=-\nu} = \rho^{-1}(\rho + 1)^{-1}{}_3F_2(1 - \nu, 1 + \nu - 2ika, \rho; 1 - 2ika, \rho + 2; 1). \tag{3.10}$$

According to Saalschütz' s theorem (see Sec. 2.2 of Ref. 13)

$${}_3F_2(-n, a, b; c, 1 + a + b - c - n; 1) = \frac{(c - a)_n(c - b)_n}{(c)_n(c - a - b)_n}, \qquad n = 0, 1, 2, \cdots. \tag{3.11}$$

Combining (3.6), (3.8), (3.10), and (3.11), we arrive at the following expression for the terms in the Weinberg series:

$$\frac{V|\psi_\nu(s)\rangle\langle\psi_\nu(s^*)|V}{\langle\psi_\nu(s^*)|V|\psi_\nu(s)\rangle} = \frac{(\nu - ika)\eta_\nu(s)}{4\pi^2 pqa} \times [W_\nu(p, k) - W_\nu(-p, k)][W_\nu(q, k) - W_\nu(-q, k)]. \tag{3.12}$$

As mentioned before, this result is much simpler than the results given previously by other authors.[6]

[1] M. G. Fuda, J. Math. Phys. 12, 1163 (1971).
[2] J. M. J. Van Leeuwen and A. S. Reiner, Physica 27, 99 (1961).
[3] H. A. Bethe and J. Goldstone, Proc. Roy. Soc. (London) A238, 551 (1957).
[4] L. Hulthén, Ark. Mat. Astron. Fysik 28A, No. 5 (1942); 29B, No. 1 (1942).
[5] S. Weinberg, Phys. Rev. 131, 440 (1963).
[6] A. G. Sitenko, V. F. Kharchenko, and N. M. Petrov, Phys. Letters 28B, 308 (1968); V. F. Kharchenko and N. M. Petrov, Nucl. Phys. A137, 417 (1969); L. P. Kok, thesis, Rijksuniversiteit te Groningen, 1969.
[7] J. S. Ball and D. Y. Wong, Phys. Rev. 169, 1362 (1968); M. G. Fuda, *ibid.* 174, 1134 (1968); 178, 1682 (1969); 186, 1078 (1969).
[8] A. Messiah, *Quantum Mechanics*, (Wiley-Interscience, New York, 1961–2).
[9] T. Fulton and P. Schwed, Phys. Rev. 115, 973 (1959); H. P. Noyes, Phys. Rev. Letters 15, 538 (1965).
[10] A. W. Babister, *Transcendental Functions Satisfying Nonhomogeneous Linear Differential Equations* (Macmillan, New York, 1967), Chap. 6.
[11] Equation (6.213) of Ref. 10.
[12] Equation (6.189) of Ref. 10.
[13] R. G. Newton, *Scattering Theory of Waves and Particles* (McGraw-Hill, New York, 1966), in particular, Chap. 12 and Sec. 14.4.
[14] W. N. Bailey, *Generalized Hypergeometric Series* (Cambridge U. P., London, 1935).
[15] See, for example, Sec. 1.3 of Ref. 14.
[16] W. F. Ford, Phys. Rev. 133, B1616 (1964).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 10 OCTOBER 1971

Scalar Diffraction by a Curved Edge*

Paul B. Bailey and George E. Barr
Sandia Laboratories, Albuquerque, New Mexico 87115
(Received 15 October 1970)

The Green's function is constructed for the problem of time-dependent scalar diffraction by a planar curved edge in a form which includes contributions to the field from nonlocal parts of the diffracting edge. The Green's function involves an elementary solution of the wave equation corresponding to the diffracted wave which is obtained in the form of a series connected with the geometric properties of the wave fronts. The half-plane, slit, and circular aperture are treated as particular cases.

1. INTRODUCTION

In diffraction by obstacles with edges, the geometrically next most simple problem after diffraction by a straight edge is diffraction by any planar edge not necessarily straight. Such problems have been treated by Keller's "geometrical theory of diffraction,"[1–3] which provides a mechanism for obtaining asymptotic solutions of diffraction problems, and by its suitable generalizations.[4,5] These methods either generate asymptotic expansions for the fields for small values of the wave length or provide asymptotic solutions where only the local geometry of the diffracting screen is of concern, and contributions of waves coming from distant parts of the screen are ignored. Exactly such problems for time-dependent diffraction by regular plane edges shall be considered here for the case of zero initial conditions. (Nonzero initial conditions can be treated in the same manner, with trivial modifications.) The present interest, however, lies in determining the solution in a form which includes contributions to the field from nonlocal parts of the screen.

The problem is to determine the function $u(t, P)$, which satisfies the wave equation

$$\Box u_\nu \equiv \left(\nabla^2 - \frac{\partial^2}{\partial t^2}\right) u_\nu(t, P) = f(t, P) \tag{1}$$

(with wave velocity taken to be unity) everywhere except possibly on the screen π, the initial conditions

$$u_\nu(0, P) = \frac{\partial u_\nu}{\partial t}(0, P) = 0 \tag{2}$$

for all points P of 3-space, and one of the following boundary conditions:

$$\text{(i)}\quad u_1(t, P) = U_1(t, P) \qquad \text{for } P \text{ on } \pi,$$
$$\text{(ii)}\quad \frac{\partial u_2}{\partial n}(t, P) = U_2(t, P) \qquad \text{for } P \text{ on } \pi. \tag{3}$$

The function $u_\nu(t, P)$ must also satisfy the edge condition, $u_\nu(t, P) < +\infty$ at the edge. (The same condition is assumed for the time-dependent equation and the reduced wave equation.) Here t represents the time, which is nonnegative, P is an arbitrary point in 3-space, and the functions f, U_1, U_2 are given arbitrary functions which are assumed to be as smooth as necessary and are compatible with (2) and (3).

The objective is the determination of the Green's function $g(t - T, P, Q)$ for the problem (1)–(3), so that $u(t, P)$ can be obtained in the form

$$u(t, P) = \iint f(T, Q) g(t - T, P, Q) dT dQ,$$

where dQ signifies an element of volume in 3-space, and the integral is over all points (T, Q) of the four-dimensional space–time for which $g(t - T, P, Q)$ is not zero.

In the special case of a straight edge and $U_\nu = 0$, this problem has been solved. Since the method of solution used here for a curved edge is different from that employed by Wait,[6] it is somewhat more illuminating to first review the solution for the case of a straight edge.

The basic tool to be employed in the construction of the solution is Green's theorem for the wave operator $\Box$ for the scalar problem. It will be referred to as the fundamental formula (F). That is, if $u(t, P)$ satisfies Eq. (1) and if $v(t, P)$ satisfies the homogeneous wave equation

$$\Box v(t, P) = 0, \tag{4}$$

[1] M. G. Fuda, J. Math. Phys. 12, 1163 (1971).
[2] J. M. J. Van Leeuwen and A. S. Reiner, Physica 27, 99 (1961).
[3] H. A. Bethe and J. Goldstone, Proc. Roy. Soc. (London) A238, 551 (1957).
[4] L. Hulthén, Ark. Mat. Astron. Fysik 28A, No. 5 (1942); 29B, No. 1 (1942).
[5] S. Weinberg, Phys. Rev. 131, 440 (1963).
[6] A. G. Sitenko, V. F. Kharchenko, and N. M. Petrov, Phys. Letters 28B, 308 (1968); V. F. Kharchenko and N. M. Petrov, Nucl. Phys. A137, 417 (1969); L. P. Kok, thesis, Rijksuniversiteit te Groningen, 1969.
[7] J. S. Ball and D. Y. Wong, Phys. Rev. 169, 1362 (1968); M. G. Fuda, *ibid.* 174, 1134 (1968); 178, 1682 (1969); 186, 1078 (1969).
[8] A. Messiah, *Quantum Mechanics*, (Wiley-Interscience, New York, 1961–2).
[9] T. Fulton and P. Schwed, Phys. Rev. 115, 973 (1959); H. P. Noyes, Phys. Rev. Letters 15, 538 (1965).
[10] A. W. Babister, *Transcendental Functions Satisfying Nonhomogeneous Linear Differential Equations* (Macmillan, New York, 1967), Chap. 6.
[11] Equation (6.213) of Ref. 10.
[12] Equation (6.189) of Ref. 10.
[13] R. G. Newton, *Scattering Theory of Waves and Particles* (McGraw-Hill, New York, 1966), in particular, Chap. 12 and Sec. 14.4.
[14] W. N. Bailey, *Generalized Hypergeometric Series* (Cambridge U. P., London, 1935).
[15] See, for example, Sec. 1.3 of Ref. 14.
[16] W. F. Ford, Phys. Rev. 133, B1616 (1964).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 10 OCTOBER 1971

Scalar Diffraction by a Curved Edge*

Paul B. Bailey and George E. Barr
Sandia Laboratories, Albuquerque, New Mexico 87115
(Received 15 October 1970)

The Green's function is constructed for the problem of time-dependent scalar diffraction by a planar curved edge in a form which includes contributions to the field from nonlocal parts of the diffracting edge. The Green's function involves an elementary solution of the wave equation corresponding to the diffracted wave which is obtained in the form of a series connected with the geometric properties of the wave fronts. The half-plane, slit, and circular aperture are treated as particular cases.

1. INTRODUCTION

In diffraction by obstacles with edges, the geometrically next most simple problem after diffraction by a straight edge is diffraction by any planar edge not necessarily straight. Such problems have been treated by Keller's "geometrical theory of diffraction,"[1–3] which provides a mechanism for obtaining asymptotic solutions of diffraction problems, and by its suitable generalizations.[4,5] These methods either generate asymptotic expansions for the fields for small values of the wave length or provide asymptotic solutions where only the local geometry of the diffracting screen is of concern, and contributions of waves coming from distant parts of the screen are ignored. Exactly such problems for time-dependent diffraction by regular plane edges shall be considered here for the case of zero initial conditions. (Nonzero initial conditions can be treated in the same manner, with trivial modifications.) The present interest, however, lies in determining the solution in a form which includes contributions to the field from nonlocal parts of the screen.

The problem is to determine the function $u(t, P)$, which satisfies the wave equation

$$\Box u_\nu \equiv \left(\nabla^2 - \frac{\partial^2}{\partial t^2}\right) u_\nu(t, P) = f(t, P) \tag{1}$$

(with wave velocity taken to be unity) everywhere except possibly on the screen π, the initial conditions

$$u_\nu(0, P) = \frac{\partial u_\nu}{\partial t}(0, P) = 0 \tag{2}$$

for all points P of 3-space, and one of the following boundary conditions:

$$\text{(i)}\quad u_1(t, P) = U_1(t, P) \qquad \text{for } P \text{ on } \pi,$$
$$\text{(ii)}\quad \frac{\partial u_2}{\partial n}(t, P) = U_2(t, P) \qquad \text{for } P \text{ on } \pi. \tag{3}$$

The function $u_\nu(t, P)$ must also satisfy the edge condition, $u_\nu(t, P) < +\infty$ at the edge. (The same condition is assumed for the time-dependent equation and the reduced wave equation.) Here t represents the time, which is nonnegative, P is an arbitrary point in 3-space, and the functions f, U_1, U_2 are given arbitrary functions which are assumed to be as smooth as necessary and are compatible with (2) and (3).

The objective is the determination of the Green's function $g(t - T, P, Q)$ for the problem (1)–(3), so that $u(t, P)$ can be obtained in the form

$$u(t, P) = \iint f(T, Q) g(t - T, P, Q) dT dQ,$$

where dQ signifies an element of volume in 3-space, and the integral is over all points (T, Q) of the four-dimensional space–time for which $g(t - T, P, Q)$ is not zero.

In the special case of a straight edge and $U_\nu = 0$, this problem has been solved. Since the method of solution used here for a curved edge is different from that employed by Wait,[6] it is somewhat more illuminating to first review the solution for the case of a straight edge.

The basic tool to be employed in the construction of the solution is Green's theorem for the wave operator $\Box$ for the scalar problem. It will be referred to as the fundamental formula (F). That is, if $u(t, P)$ satisfies Eq. (1) and if $v(t, P)$ satisfies the homogeneous wave equation

$$\Box v(t, P) = 0, \tag{4}$$

then we obtain (F):

$$\int_D f(T, Q)v(T, Q)\,dTdQ + \int_{\partial D}\left(u\frac{dv}{d\nu} - v\frac{du}{d\nu}\right)dS = 0,$$

where D is any regular region in the four-dimensional space–time throughout which u and v have no singularities of any kind, ∂D is the boundary of D, and dS is an element of surface area on ∂D. The directional derivative $d/d\nu$, called the "transversal derivative," is defined by

$$\frac{du}{d\nu} \equiv l_0\frac{\partial u}{\partial T} - l_1\frac{\partial u}{\partial X} - l_2\frac{\partial u}{\partial Y} - l_3\frac{\partial u}{\partial Z}, \qquad (5)$$

where l_0, l_1, l_2, l_3 are the direction cosines of the inward normal to the surface. This formula may be proved in exactly the same way as Gauss's divergence theorem.[7] For functions u and v, which may have singularities within or on the boundary of the region, one can either apply (F) to a neighboring region where no singularities exist and then take limits, or one can interpret the resulting integrals in a generalized sense.[8]

2. THE STRAIGHT EDGE

If t is so small that no radiation could yet be received at P from any point Q on the edge, i.e., $0 < t <$ distance from P to the edge, the problem does not differ from that for a whole plane, in which case the solution is known to be

$$u_\nu(t, P) = -\frac{1}{4\pi}\int_{D_i} f(T, Q)\frac{\delta(t - T - R)}{R}\,dTdQ + \frac{(-1)^{\nu+1}}{4\pi}\int_{D_r} f(T, Q)\frac{\delta(t - T - R^*)}{R^*}\,dTdQ, \qquad (6)$$

where R is the distance $\overline{PQ}$; D_i is the region (space–time) bounded by the portion of the cone $T = t - R$ which extends from its vertex (t, P) to the half-hyperplane π, and the initial hyperplane $T = 0$; R^* is the distance $\overline{P^*Q}$, where P^* is the image of P in the extension of π; D_r is the region bounded by the portion of the cone $T = t - R^*$ which lies on the same side of π as does P, the half-hyperplane π, and the hyperplane $T = 0$.

Since δ is the Dirac δ function, (6) could be written in the form

$$u_\nu(t, P) = -\frac{1}{4\pi}\int_{R\le t} f(t - R, Q)\frac{1}{R}\,dQ + \frac{(-1)^{\nu+1}}{4\pi} \times \int_{R^*\le t} f(t - R^*, Q)\frac{1}{R^*}\,dQ.$$

Since the form in (6) shows the Green's function is

$$g_\nu(t - T, P, Q) = -\frac{1}{4\pi}\begin{cases}\dfrac{1}{R}\,\delta(t - T - R) & \text{in } D_i - D_r,\\[2ex] \dfrac{1}{R}\,\delta(t - T - R) + \dfrac{(-1)^\nu}{R^*}\,\delta(t - T - R^*) & \text{in } D_r,\end{cases}$$

one need only consider the case t is greater than the distance from P to the edge. First consider the case with P located above the half-plane π, for which case P will receive direct, reflected, and diffracted radiation.

Since the cone $T = t - R$ now will contain a portion of the edge of the half-plane, the boundary of the region D_i, which consists of all those points (T, Q) from which direct radiation can be received at (t, P), is made up of a portion of the cone $T = t - R$, the plane $T = 0$, the half-plane π, together with a portion of the "shadow" plane. By "shadow" plane is meant the plane determined by the edge of the half-plane π and the point (t, P). Denote this shadow plane (actually a hyperplane) by p. Similarly the region D_r consists of all points (T, Q) from which reflected radiation can be received at (t, P) and has its boundary made up of a portion of the cone $T = t - R^*$, the plane $T = 0$, the half-plane π, together with the shadow plane p^*, which is the reflection of p in the extension of π. Finally there is the region D_d, which consists of all points (T, Q) from which diffracted radiation can be received at (t, P). Its boundary is a "conelike" surface C_d, with equation of the form $T = t - a$, where a is the shortest distance from P to Q by a path which touches the edge. That is, by a broken line $\overline{PE} + \overline{EQ}$, where E is the point of the edge which minimizes the sum. If $P = (x, y, z)$ and $Q = (X, Y, Z)$ and if the edge is the line $x = 0, y = 0$, then

$$a = [(r + \rho)^2 + (z - Z)^2]^{1/2}, \qquad (7)$$

where $r = (x^2 + y^2)^{1/2}$ and $\rho = (X^2 + Y^2)^{1/2}$.

The idea is to apply the fundamental formula (F) to each of the regions D_i, D_r, D_d using the (unknown) function $u(t, P)$ and appropriate solutions v of Eq. (4). First apply (F) to D_i with $v = \delta(t - T - R)/R$, yielding

$$0 = \int_{D_i} f\frac{\delta}{R}\,dTdQ + \int_{c_i+\pi_i+p_+}\left(\frac{\delta}{R}\frac{du}{d\nu} - u\frac{d}{d\nu}\frac{\delta}{R}\right)dS, \qquad (8)$$

where c_i, π_i, p_+ denote the respective portions of the surfaces of the cone $T = t - R$, and planes π and p which make up the surface ∂D_i. (The integral over $T = 0$ is zero since u and $\partial u/\partial T$ vanish there.) One finds that the only part of c_i giving a

nonzero contribution is the neighborhood of the vertex (t, P) and this contribution equals $4\pi\, u(t, P)$ so that (8) can be rewritten as

$$-4\pi\, u(t,P) = \int_{D_i} f\frac{\delta}{R}\,dTdQ + \int_{\pi_i+p_+}\left(\frac{\delta}{R}\frac{du}{d\nu} - u\frac{d}{d\nu}\frac{\delta}{R}\right)dS.$$

On p_+ the transversal derivative is just the outward normal, and

$$\frac{d}{dn}\frac{\delta(t-T-R)}{R} = -\frac{1}{R}\left(\delta'(t-T-R) + \frac{\delta(t-T-R)}{R}\right)\frac{dR}{dn},$$

which vanishes on p because $dR/dn = 0$ there. Thus,

$$-4\pi\, u(t,P) = \int_{D_i} f\frac{\delta}{R}\,dTdQ + \int_{\pi_i}\left(\frac{\delta}{R}\frac{du}{d\nu} - u\frac{d}{d\nu}\frac{\delta}{R}\right)dS + \int_{p_+}\frac{\delta}{R}\frac{du}{d\nu}\,dS, \tag{9}$$

where one notices that the first term on the right-hand side is known, as is the integral of either $u(d/d\nu)\,(\delta/R)$ or $(\delta/R)\,(du/d\nu)$, since either u or $du/d\nu$ is given on π.

Next apply (F) to the region D_d picking for v an appropriate function w and obtain

$$0 = \int_{D_d} fw\,dTdQ + \int_{C_d+\pi_{d+}+\pi_{d-}}\left(w\frac{du}{d\nu} - u\frac{dw}{d\nu}\right)dS. \tag{10}$$

The only term involving values of u other than on π is the surface integral over C_d; so if w is the appropriate solution of (4) which agrees with $\delta(t-T-R)/R$, then one can have

$$\int_{p_+}\frac{\delta}{R}\frac{du}{d\nu}\,dS = \int_{C_d}\left(u\frac{dw}{d\nu} - w\frac{du}{d\nu}\right)dS. \tag{11}$$

If this is true then (10) can be added to (9) to yield

$$-4\pi\, u(t,P) = \int_{D_i} f\frac{\delta}{R}\,dTdQ + \int_{D_d} fw\,dTdQ + \int_{\pi_i}\left(\frac{\delta}{R}\frac{du}{d\nu} - u\frac{d}{d\nu}\frac{\delta}{R}\right)dS + \int_{\pi_{d+}+\pi_{d-}}\left(w\frac{du}{d\nu} - u\frac{dw}{d\nu}\right)dS. \tag{12}$$

If this argument is repeated with regions D_γ and D_d and functions $\delta(t-T-R^*)/R^*$ and w^*, which satisfies the condition corresponding to (11), namely

$$\int_{p_+^*}\frac{\delta}{R^*}\frac{du}{d\nu}\,dS = \int_{C_d}\left(u\frac{dw^*}{d\nu} - w^*\frac{du}{d\nu}\right)dS, \tag{13}$$

one obtains

$$0 = \int_{D_\gamma} f\frac{\delta^*}{R^*}\,dTdQ + \int_{D_d} fw^*\,dTdQ + \int_{\pi_i}\left(\frac{\delta^*}{R^*}\frac{du}{d\nu} - u\frac{d}{d\nu}\frac{\delta^*}{R^*}\right)dS + \int_{\pi_{d+}+\pi_{d-}}\left(w^*\frac{du}{d\nu} - u\frac{dw^*}{d\nu}\right)dS. \tag{14}$$

Finally, subtracting and adding (12) to (14) yields, respectively,

$$-4\pi\, u_1(t,P) = \int_{D_i} f\frac{\delta}{R}\,dTdQ - \int_{D_\gamma} f\frac{\delta^*}{R^*}\,dTdQ + \int_{D_d} f(w-w^*)\,dTdQ - 2\int_{\pi_i} U_1\frac{d}{d\nu}\frac{\delta}{R}\,dS - 4\int_{\pi_{d+}} U_1\frac{dw}{d\nu}\,dS \tag{15a}$$

and

$$-4\pi\, u_2(t,P) = \int_{D_i} f\frac{\delta}{R}\,dTdQ + \int_{D_\gamma} f\frac{\delta^*}{R^*}\,dTdQ + \int_{D_d} f(w+w^*)\,dTdQ + 2\int_{\pi_i} U_2\frac{\delta}{R}\,dS + 4\int_{\pi_{d+}} wU_2\,dS, \tag{15b}$$

upon taking into account the facts that $R = R^*$, $\delta = \delta^*$, $w = w^*$,

$$\frac{d}{d\nu}\frac{\delta^*}{R^*} = -\frac{d}{d\nu}\frac{\delta}{R},$$

hence $dw^*/d\nu = -\,dw/d\nu$ on π, and the boundary conditions. Thus, if a solution w of (4) can be found which satisfies (11) and a w^* satisfying (13), Eq. (15) give the solutions to the problems (1)–(3).

3. CONSTRUCTION OF W FOR THE STRAIGHT EDGE

A construction of elementary solutions of quite general hyperbolic differential equations has been explained by Hadamard[7] and is adapted for use here. The main difference is that instead of constructing the solution of (4) corresponding to the region bounded by the conoid $\Gamma = 0, \Gamma \equiv (t-T)^2 - R^2$, with the form $U\Gamma^{-q}$, with U regular and taking on a specified value at the vertex (t, P) of the conoid, one must construct the solution of (4) corresponding to the region bounded by $\gamma = 0, \gamma \equiv (t-T)^2 - a^2$, which has the form $V\gamma^{-q}$, with V regular most everywhere and taking such values along the hyperplane p that condition (11) is satisfied.

On the surface $T = t - a$ (the only portion of $\gamma = 0$ the argument is concerned with), where a is given by (7), a convenient coordinate system can be chosen as follows: If $Q(X, Y, Z)$ is any point on this surface, and if E is the point of the edge such that $\overline{PE} + \overline{EQ}$ is a minimum, then **PE** and **QE** make equal angles φ with the edge, so if E, P have coordinates $(0, 0, e)$ and (x, y, z), respectively, and the edge is $x = 0 = y$, then

$$r = s_0 \sin\varphi, \quad \rho = \sigma \sin\varphi,$$
$$e - z = s_0 \cos\varphi, \quad Z - e = \sigma \cos\varphi,$$

where s_0 denotes the distance $\overline{PE}$ and s is the distance $PE + EQ$, so

$$\rho = s \sin\varphi - r, \quad Z - z = s \cos\varphi.$$

The rectangular coordinates of a point Q are then seen to be

$$Q = \langle (s \sin\varphi - r)\cos\Theta, (s \sin\varphi - r)\sin\Theta, z + s\cos\varphi \rangle.$$

Thus a natural coordinate system for the surface integrals in (15) is (s, Θ, φ), where

$$T = t - s, \quad X = (s \sin\varphi - r)\cos\Theta,$$
$$Y = (s \sin\varphi - r)\sin\Theta, \quad Z = z + s\cos\varphi,$$

where s ranges from $r \csc\varphi$ to t.

Since w is expected to be infinite on $\gamma = 0$, one must apply (F) to a neighboring surface and approach $\gamma = 0$ by a limiting process. For this purpose it is convenient to use the surface $T = t - a - \epsilon$, $\epsilon > 0$. On this surface,

$$dS = \sqrt{2}\, dX dY dZ$$

and

$$dT + \frac{r+\rho}{a}\frac{X}{\rho}dX + \frac{r+\rho}{a}dY + \frac{Z-z}{a}dZ = 0;$$

the unit inward normal is

$$\langle l_0, l_1, l_2, l_3 \rangle = -\frac{1}{\sqrt{2}}\left\langle 1, \frac{r+\rho}{a}\cdot\frac{X}{\rho}, \frac{r+\rho}{a}\cdot\frac{Y}{\rho}, \frac{Z-z}{a}\right\rangle,$$

the directional derivative multiplied by dS is

$$dS\frac{d}{d\nu} = -\left(\frac{\partial}{\partial T} - \frac{r+\rho}{a}\frac{X}{\rho}\frac{\partial}{\partial X} - \frac{r+\rho}{a}\frac{Y}{\rho}\frac{\partial}{\partial Y} - \frac{Z-z}{a}\frac{\partial}{\partial Z}\right),$$

which in the coordinates (s, Θ, φ) of Q (with $T = t - s - \epsilon$) becomes

$$dS\cdot\frac{d}{d\nu} = (s\sin\varphi - r)s\, ds\, d\Theta\, d\varphi\,\frac{\partial}{\partial s}.$$

For the surface p_+, a convenient choice of coordinates is (T, ρ, Θ), where $X = \rho\cos\Theta$, $Y = \rho\sin\Theta$, and p has the equation $\Theta = \theta + \pi$, with θ defined by $x = r\cos\theta$, $y = r\sin\theta$. Hence, if one continues to write u, w indifferently of the coordinate system being used and remembering which surface one is integrating over, condition (11) takes the form

$$\iiint\left(\frac{1}{\rho}\frac{\partial u}{\partial\Theta}\frac{\delta(t-T-R)}{R}\right)_{\Theta=\theta+\pi} dT d\rho dZ = \iiint\left(u\frac{\partial w}{\partial\sigma} - w\frac{\partial u}{\partial\sigma}\right)s(s\sin\varphi - r)ds\, d\Theta\, d\varphi, \tag{16}$$

since on p_+,

$$\frac{du}{d\nu} = \frac{1}{\rho}\frac{\partial u}{\partial\Theta}$$

and $dS = dT d\rho dZ$.

The term on the left-hand side (lhs) is equivalent to

$$\iint\frac{1}{\rho a}\frac{\partial u}{\partial\Theta}(t - a - \epsilon, \theta + \pi, z)\, d\rho dz,$$

while the term on the rhs can be integrated by parts with respect to s, giving

$$-\iint[s(s\sin\varphi - r)uw]_{r\csc\varphi}^{t-\epsilon} d\Theta d\varphi + \iiint u\left(2s(s\sin\varphi - r)\frac{\partial w}{\partial s} + (2s\sin\varphi - rw)\right) \times ds\, d\Theta\, d\varphi.$$

The first term here vanishes, since $u(t - s - \epsilon, s, \Theta, \varphi) = 0$ when $s = t - \epsilon$, so that (16) becomes

$$\lim_{\epsilon\to 0}\iint\frac{1}{\rho a}\frac{\partial u}{\partial\Theta}(t - a - \epsilon, \theta + \pi, z)\, d\rho dz = \lim_{\epsilon\to 0}\iiint_{r\csc\varphi}^{t-\epsilon} u\left(2s(s\sin\varphi - r)\frac{\partial w}{\partial s} + (2s\sin\varphi - r)w\right)ds\, d\Theta\, d\varphi.$$

In this expression, the change of variables from ρ, z to s, φ in the lhs ($s\sin\varphi = r + \rho$, $s\cos\varphi = Z - z$), would make the lhs and rhs agree if and only if

$$\lim_{\epsilon\to 0}\int_0^{2\pi} u(t - s - \epsilon, s, \Theta, \varphi)\left[2(s\sin\varphi - r)\frac{\partial}{\partial s} \times w(t - s - \epsilon, s, \Theta, \varphi) + \left(2\sin\varphi - \frac{r}{s}\right) \times w(t - s - \epsilon, s, \Theta, \varphi)\right]s\, d\Theta = \frac{1}{\rho}\frac{\partial u}{\partial\Theta}(t - s, s, \theta + \pi, \varphi), \tag{17}$$

(since a equals the arclength s).

The quantity in brackets must tend to zero with ϵ for $\Theta \neq \theta + \pi$ and tend to infinity for $\Theta = \theta + \pi$. In fact, it must behave like $\delta'(\Theta = \theta + \pi)$. Hence if $w = V\gamma^{-q}$, with V regular inside $\gamma = 0$, then V must still have a singularity of some sort where $\gamma = 0$ meets p (i.e., where $\Theta = \theta + \pi$), so it may be expected that V has, the form $V = W\Gamma^{-\alpha}$, with $\alpha > 0$ or

$$w \equiv W\Gamma^{-\alpha}\gamma^{-q}, \tag{18}$$

where W remains finite within and on $\gamma = 0$.

Substituting (18) for w in the large parentheses in (17) gives the expression

$$[2r\rho(1+\cos\overline{\theta-\Theta})+2s\epsilon+\epsilon^2]^{-\alpha-1}(2s+\epsilon)^{-q-1}$$
$$\times(2\rho(2s+\epsilon)[2r\rho(1+\cos\overline{\theta-\Theta})$$
$$+2s\epsilon+\epsilon^2]\frac{\partial\tilde{W}}{\partial s}+\tilde{W}[4r\rho(1+\cos\overline{\theta-\Theta})$$
$$\times(2\rho+r-\alpha r-\alpha\rho-2q\rho)+\epsilon R]), \quad (19)$$

where $\tilde{W}$ is $W(t-s-\epsilon, s, \Theta, \varphi)$, $\rho = s\sin\varphi - r$, and R indicates the remaining terms which have not been written explicitly. In order to satisfy (17), the expression in large parentheses in (19) must go to zero with ϵ. Hence, $\overline{W} \equiv W(t-s, s, \Theta, \varphi)$ satisfies the differential equation

$$2\rho s\frac{\partial\overline{W}}{\partial s}+(2\rho+r-2\alpha r-2\alpha\rho-2\rho q)\overline{W}=0, \quad (20)$$

which is readily integrated to give

$$\overline{W}=D(\Theta,\varphi)s^{q-1/2}\rho^{\alpha-1/2}, \quad (21)$$

where $D(\Theta, \varphi)$ is independent of s and remains finite when $\Theta = \theta + \pi$.

The singular behavior of (20) arises from the factor $[2r\rho(1+\cos\overline{\theta-\Theta})+2s\epsilon+\epsilon^2]^{-\alpha-1}$, which, for $\Theta = \theta + \pi$, tends to infinity as $\epsilon \to 0$; since it is only the singularity which determines the validity of (17), it suffices to look at the most singular term in (19). Since W is to have no singularities, $\tilde{W}$ in (19) may be replaced by $\overline{W}$ from (21). The most singular term in the resulting expression can be simplified to

$$\tfrac{1}{2}\alpha(4r)^{-\alpha-q+1}\rho^{-q-1/2}s^{q-3/2}\lambda^{-q+1}$$
$$\times(\cos^2\tfrac{1}{2}\overline{\theta-\Theta}+\lambda)^{-\alpha-1}D(\Theta,\varphi), \quad (22)$$

with

$$\lambda = s\epsilon/2r\rho.$$

Near $\Theta = \theta + \pi$, $\cos\frac{1}{2}(\theta - \Theta)$ is near zero, which poses difficulties for the choice of α, q, and $D(\Theta, \varphi)$. Consider for the moment the similar but simpler problem of determining α, q, and $\beta > 0$, such that for an arbitrary but nice enough function $\Psi(x)$ the following relation holds:

$$\lim_{\lambda\downarrow 0}\lambda^{-q+1}\int_{-1}^{1}\psi(x)x^{\beta}(x^2+\lambda)^{-\alpha-1}dx = \text{const}\cdot\psi'(0). \quad (23)$$

Clearly $-q + 1 > 0$, and since $\alpha > 0$, integration by parts gives

$$\lim_{\lambda\downarrow 0}\lambda^{-q+1}\Big(-\psi(x)x^{\beta-1}(x^2+\lambda)^{-\alpha}/2\alpha\Big|_{-1}^{1}$$
$$+(1/2\alpha)\int_{-1}^{1}(x^2+\lambda)^{-\alpha}[(\beta-1)x^{\beta-2}\psi(x)$$
$$+x^{\beta-1}\psi'(x)]dx = C\psi'(0). \quad (24)$$

The first term has limit zero and since the first term of the integrand is more singular than the second, with the second giving rise to the finite limit $\psi'(0)$, necessarily the coefficient of the first term must vanish. So $\beta = 1$, and (24) reduces to

$$\lim_{\lambda\downarrow 0}\lambda^{-q+1}\int_{-1}^{1}(x^2+\lambda)^{-\alpha}\psi'(x)dx = 2\alpha C\psi'(0).$$

Now the lhs can be written

$$\lim_{\lambda\downarrow 0}\lambda^{-q+1}\Big(\psi'(0)\int_{-1}^{1}(x^2+\lambda)^{-\alpha}dx+\int_{-1}^{1}(x^2+\lambda)^{-\alpha}$$
$$\times[\psi'(x)-\psi'(0)]dx\Big),$$

where the second term is less singular than the first and the first has asymptotic behavior:

$$\psi'(0)\Big(\frac{1}{\frac{1}{2}-\alpha}\lambda^{-q+1}+\frac{\Gamma(\frac{1}{2})\Gamma(\alpha-\frac{1}{2})}{\Gamma(\alpha)}\lambda^{3/2-\alpha-q}\Big).$$

Accordingly, it follows that

$$\tfrac{3}{2}-\alpha = q \quad (25)$$

and that the constant in (23) is

$$C=\Gamma(\tfrac{1}{2})\Gamma(\alpha-\tfrac{1}{2})(\Gamma(\alpha))^{-1}.$$

So for (17) to hold with w of the form (18), the function $D(\Theta, \varphi)$ in (21) must have a simple factor of $\cos\frac{1}{2}(\theta - \Theta)$, the constant q is given in terms of α by (25), and the limit of the lhs of (17) for $\epsilon \to 0$ is then equal to [using (25)]

$$\alpha(4r)^{-1/2}\rho^{\alpha-2}s^{-\alpha}\Gamma(\tfrac{1}{2})\Gamma(\alpha-\tfrac{1}{2})(\Gamma(\alpha))^{-1}$$
$$\times\{D(\Theta,\varphi)[\cos\tfrac{1}{2}(\theta-\Theta)]^{-1}\}_{\Theta=\theta+\pi}\frac{\partial u}{\partial\Theta}(t-s,s,\theta$$
$$+\pi,\varphi).$$

Hence, for equality with the rhs of (17) one must have $\alpha = 1$, giving $C = 1$ and

$$D(\Theta,\varphi)=(2/\pi)(\sqrt{r})\cos\tfrac{1}{2}(\theta-\Theta),$$

and accordingly

$$W(t-s,s,\Theta,\varphi)\equiv\overline{W}=(2/\pi)(r\rho)^{1/2}\cos\tfrac{1}{2}(\theta-\Theta). \quad (26)$$

Expression (26) is the value of W only on the surface $\gamma = 0$, hence one can write

$$W=\overline{W}+V_1\gamma\Gamma,$$

where V_1 is a function having no singularities on $\gamma = 0$. Since $w \equiv W\Gamma^{-1}\gamma^{-1/2}$ is to be a solution of (4), the equation for V_1 can be obtained by substituting $w = \overline{W}\Gamma^{-1}\gamma^{-1/2} + V_1\gamma^{1/2}$ in (4). Since one finds that

$$\Box(\overline{W}\Gamma^{-1}\gamma^{-1/2}) = 0,$$

$V_1 = 0$ gives a solution (for the straight edge)

$$w = \overline{W}\Gamma^{-1}\gamma^{-1/2}. \tag{27}$$

That this is the only solution of the form (18) which remains finite on the edge will be shown later (Sec. 7).

With this function for w, Eqs. (15) give the solution of the diffraction problem corresponding to the boundary condition (3). One sees that for zero boundary conditions these solutions agree with those obtained by Wait.[6] It is merely a matter of differentiating Wait's solutions with respect to t.

4. CONSTRUCTION OF AN ELEMENTARY SOLUTION FOR DIFFRACTION BY A CURVED EDGE

For simplicity assume first that the point P and time t are such that there is only one minimal path $\overline{PE} + \overline{EQ}$ between P and any point Q via the given curved edge, for any point Q of three-space such that $\overline{PE} + \overline{EQ} < t$. Suppose also that the distance from P to the edge is less than t, so that diffracted radiation can be received at P at time t. Let a again denote this minimal path length from any arbitrary point Q to P via the edge. [Note that a is not given by (7) for any but a straight edge.] Proceeding exactly as in Sec. 2, we must again construct a solution w of (4) which takes on such values on the surface $\gamma = 0$ that Eq. (11) holds.

To carry out this construction in general one needs to define a suitable coordinate system, which is done in the following way. Let an arbitrary point E of the edge by $(\alpha(e), 0, \beta(e))$, where e denotes arc length along the edge. Hence

$$(\alpha'(e))^2 + (\beta'(e))^2 = 1.$$

Given $P(x, y, z)$, $E(\alpha(e), 0, \beta(e))$ and denoting the distance $|\overline{PE}|$ by σ_0,

$$\sigma_0 = [(\alpha - x)^2 + y^2 + (\beta - z)^2]^{1/2},$$

then the angle φ between **EP** and the tangent line at E is given by

$$\cos\varphi = \sigma_0^{-1}\,[(x - \alpha)\alpha' + (z - \beta)\beta'\,]. \tag{28}$$

Since the path $\overline{PE} + \overline{EQ}$ is the minimal path from P to Q via the edge if and only if the line $\overline{QE}$ also makes an angle φ with the tangent line at E, then Q necessarily lies on the cone which has vertex E, axis the tangent line at E, and semivertical angle φ. It follows that if Θ denotes the angle between the plane π and the plane containing $\overline{QE}$ and the tangent line at E, and if σ denotes the length $|\overline{EQ}|$, then the vector EQ is

$$\mathbf{EQ} = \sigma\langle -\alpha'\cos\varphi + \beta'\sin\varphi\cos\Theta,\ \sin\varphi\sin\Theta, -\beta'\times\cos\varphi - \alpha'\sin\varphi\cos\Theta\rangle,$$

and similarly the vector **EP** is

$$\mathbf{EP} = \sigma_0\langle \alpha'\cos\varphi + \beta'\sin\varphi\cos\theta,\ \sin\varphi\sin\theta, \beta'\times\cos\varphi - \alpha'\sin\varphi\cos\theta\rangle,$$

where θ denotes the angle between the plane π and the plane containing $\overline{EP}$ and the tangent line at E (see Fig. 1).

One finds then that the coordinates (X, Y, Z) of Q are given in terms of the coordinates (σ, Θ, e) by the relations

$$\begin{aligned} X &= \alpha - \sigma(\alpha'\cos\varphi - \beta'\sin\varphi\cos\Theta),\\ Y &= \sigma\sin\varphi\sin\Theta,\\ Z &= \beta - \sigma(\beta'\cos\varphi + \alpha'\sin\varphi\cos\Theta), \end{aligned} \tag{29}$$

where φ is given by (28), and one finds the angle Ω between the lines $\overline{PE}$ and $\overline{QE}$ is given by

$$\begin{aligned} \cos\Omega &= (\mathbf{EQ}, \mathbf{EP})\,|\overline{EQ}|^{-1}\,|EP|^{-1}\\ &= \sin^2\varphi\cos(\theta - \Theta) - \cos^2\varphi, \end{aligned}$$

where

$$\sin\theta = y\sigma_0^{-1}\sin^{-1}\varphi.$$

Since one now has

$$\cos\tfrac{1}{2}\Omega = \sin\varphi\cos\tfrac{1}{2}(\theta - \Theta),$$

one finds

$$a^2 - R^2 = 4\sigma_0\sigma\sin^2\varphi\cos^2\tfrac{1}{2}(\theta - \Theta),$$

and in addition one notes that in these coordinates $a = \sigma + \sigma_0$.

To carry out this construction in general, one notices that since a is the minimum distance from P to Q via the edge (or, equivalently, since $T = t - a$ is a characteristic surface for the wave operator, i.e., a wave front), it follows that $|\nabla a|^2 = 1$. To

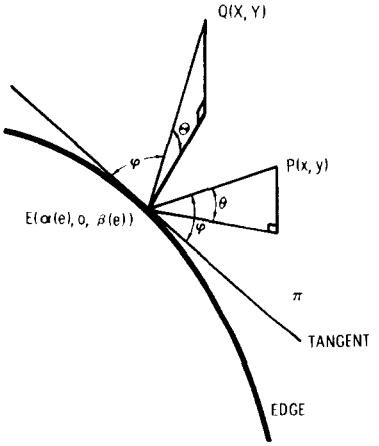


FIG. 1. $E(\alpha(e), 0, \beta(e))$ is an arbitrary point of the edge. The lines **QE** and **PE** make equal angles φ with the tangent to the edge at E to satisfy the minimal path requirements. The angle Θ denotes the angle between the plane of the edge π and the plane containing QE and the tangent line at E. The angle θ is defined similarly.

show this, suppose the minimum path from P to Q meets the edge at the point $E(\alpha(e), 0, \beta(e))$. The distance from P to Q along the minimal path is then

$$a = \{[x - \alpha(e)]^2 + y^2 + [z - \beta(e)]^2\}^{1/2} + \{[X - \alpha(e)]^2 + Y^2 + [Z - \beta(e)]^2\}^{1/2}, \quad (30)$$

where e is such that $\partial a/\partial e = 0$. But then

$$\nabla a = \{[X - \alpha(e)]^2 + Y^2 + [Z - \beta(e)]^2\}^{-1/2} \times \langle X - \alpha(e), Y, Z - \beta(e)\rangle,$$

from which the observation is obvious. Similarly one has $|\nabla R|^2 = 1$.

On the surface $T = t - a - \epsilon$, $\epsilon > 0$,

$$dS = \sqrt{2}\, dX dY dZ$$

and

$$dT + \frac{\partial a}{\partial X} dX + \frac{\partial a}{\partial Y} dY + \frac{\partial a}{\partial Z} aZ = 0,$$

the unit inward normal is

$$\langle l_0, l_1, l_2, l_3\rangle = -\frac{1}{\sqrt{2}}\left\langle 1, \frac{\partial a}{\partial X}, \frac{\partial a}{\partial Y}, \frac{\partial a}{\partial Z}\right\rangle,$$

the directional derivative multiplied by dS is

$$dS\,\frac{d}{d\nu} = -\left(\frac{\partial}{\partial T} - \frac{\partial a}{\partial X}\frac{\partial}{\partial X} - \frac{\partial a}{\partial Y}\frac{\partial}{\partial Y} - \frac{\partial a}{\partial Z}\frac{\partial}{\partial Z}\right),$$

which in the coordinate system (σ, Θ, e) is (with $T = t - a - \epsilon$)

$$dS\,\frac{d}{d\nu} = J d\sigma d\Theta de\,\frac{\partial}{\partial \sigma},$$

where J is the Jacobian of the transformation from (X, Y, Z) to (σ, Θ, e). Thus in place of (16) one now gets

$$\iiint\left(\frac{1}{\sigma \sin\varphi}\frac{\partial u}{\partial \Theta}\frac{\sigma(t - T - R)}{R}\right)_{\Theta = \theta + \pi}\frac{1}{\sigma \sin\varphi} J dT d\sigma de = \iiint\left(u\frac{\partial w}{\partial \sigma} - w\frac{\partial u}{\partial \sigma}\right) J d\sigma d\Theta de. \quad (31)$$

Use has been made of the fact that on p_+,

$$\frac{du}{d\nu} = \frac{1}{\sigma \sin\varphi}\frac{\partial u}{\partial \Theta}$$

and

$$dS = \frac{1}{\sigma \sin\varphi} J dT d\sigma de.$$

Integration by parts of the rhs of (31) gives

$$\lim_{\epsilon \downarrow 0}\int_0^{2\pi} u(t - a - \epsilon, \sigma, \Theta, e)\left(2J\frac{\partial w}{\partial \sigma}(t - a - \epsilon, \sigma, \Theta, e) + \frac{\partial J}{\partial \sigma} W(t - a - \epsilon, \sigma, \Theta, e)\right) d\Theta$$

$$= \sigma^{-2} a^{-1} \sin^{-2}\varphi\, J(\sigma, \theta + \pi, e)\frac{\partial u}{\partial \Theta} \times (t - a, \sigma, \theta + \pi, e) \cdots \quad (32)$$

in place of (17).

It is shown in the Appendix that

$$\frac{\partial J}{\partial \sigma} = J\nabla^2 a\ .$$

Hence (29) becomes

$$\lim_{\epsilon \downarrow 0}\int_0^{2\pi} u(t - a - \epsilon, \sigma, \Theta, e)\left(2\frac{\partial w}{\partial \sigma}(t - a - \epsilon, \sigma, \Theta, e) + w\nabla^2 a\right) J d\Theta = \frac{1}{a\sigma^2} J(\sigma, \theta + \pi, e)\frac{\partial u}{\partial \Theta}(t - a, \sigma, \theta + \pi, e). \quad (33)$$

Assume that w has the form

$$w \equiv W\Gamma^{-1}\gamma^{-1/2}. \quad (34)$$

Then

$$w(t - a - \epsilon, \sigma, \Theta, e) = W(t - a - \epsilon, \sigma, \Theta, e) \times [a^2 - R^2 + 2a\epsilon + \epsilon^2]^{-1}(2a\epsilon + \epsilon^2)^{-1/2}. \quad (35)$$

Since Ω is the angle $\overline{PE}$ makes with the plane π and $\overline{PE} = \sigma_0$, and hence

$$a^2 - R^2 = 2\sigma\sigma_0(1 + \cos\Omega),$$

one has

$$\frac{\partial w}{\partial \sigma}(t - a - \epsilon, \sigma, \Theta, e) = w\left(\frac{1}{W}\frac{\partial W}{\partial \sigma}(t - a - \epsilon, \sigma, \Theta, e) - \frac{2\sigma_0(1 + \cos\Omega) + 2\epsilon}{a^2 - R^2 + 2a\epsilon + \epsilon^2} - \frac{1}{2a + \epsilon}\right).$$

To satisfy (34) one must have $\overline{W} = W(t - a, \sigma, \Theta, e)$ satisfy the equation

$$2a\sigma\frac{\partial \overline{W}}{\partial \sigma} + (A\sigma - 2a)\overline{W} = 0, \quad (36)$$

in place of (20), where

$$A = a\nabla^2 a - 1.$$

Integrating (36) gives

$$\overline{W} = D(\Theta, e)\sigma a^{1/2} J^{-1/2}. \quad (37)$$

Upon substituting

$$w(t - a - \epsilon, \sigma, \Theta, e) \doteq \overline{W}(a^2 - R^2 + 2a\epsilon + \epsilon^2)^{-1} \times (2a\epsilon + \epsilon^2)^{-1/2}$$

into (33), one finds for the most singular term in large parentheses

$$4D(\Theta, e)a^{-1/2}\sigma_0^{-1/2}\lambda^{1/2}\sigma^{-3/2} \times \sin^{-1}\varphi[\cos^2\tfrac{1}{2}(\theta - \Theta) + \lambda]^{-2}J^{-1/2},$$

where $\lambda = \epsilon a(2\sigma_0\sigma \sin^2\varphi)^{-1}$. The limit of the lhs of (33) is then equal to

$$[D(\Theta, e)(\cos\tfrac{1}{2}\Omega)^{-1}]_{\Theta=\theta+\pi} \cdot \tfrac{1}{2}\pi a^{-1}\sigma^{-2} \sin^{-2}\varphi \frac{\partial u}{\partial \Theta} \times (t - a, \sigma, \theta + \pi, e);$$

so for equality with the rhs one must have

$$D(\Theta, e) = \cos\tfrac{1}{2}\Omega \cdot 2\pi^{-1} \sin\varphi = \pm \pi^{-1}(a^2 - R^2)^{1/2}\sigma_0^{-1/2}\sigma^{-1/2} \sin\varphi,$$

where the plus sign is taken above the shadow surface and the minus sign below. Hence,

$$\overline{W} = \pm \pi^{-1}(a^2 - R^2)^{1/2}a^{1/2}\sigma^{1/2}\sigma_0^{-1/2}J^{-1/2} \sin\varphi. \tag{38}$$

Assume that

$$w = [\pm \pi^{-1}(a^2 - R^2)^{1/2}\Gamma^{-1} + v_0 + v_1\gamma + \cdots] \times \gamma^{-1/2}, \tag{39}$$

with v_n independent of t, T. Then one must have

$$v_0 = \overline{W}(a^2 - R^2)^{-1} \mp \pi^{-1}(a^2 - R^2)^{-1/2}. \tag{40}$$

To determine the coefficients v_n, $n \geq 1$, it is first necessary to make several observations. First, note that

$$\Box((a^2 - R^2)^{1/2}\Gamma^{-1}\gamma^{-1/2}) = -\gamma^{-3/2}\nabla^2 \times (a^2 - R^2)^{1/2},$$

which may be verified by computation. Second, if v is any function independent of t, T then (using $|\nabla a|^2 = 1$, $|\nabla R|^2 = 1$)

$$\Box v\gamma^{n-1/2} = \gamma^{n-1/2}\nabla^2 v - (2n - 1)\gamma^{n-3/2} \times [(\nabla a^2, \nabla v) + (A + 2n)v], \tag{41}$$

and in the coordinate system (σ, Θ, e) one has

$$(\nabla a^2, \nabla v) = 2a \frac{\partial v}{\partial \sigma}(\sigma, \Theta, e).$$

Further, by expanding in the obvious way one has

$$(a^2 - R^2)^{1/2}\Gamma^{-1} = (a^2 - R^2)^{-1/2} - (a^2 - R^2)^{-3/2}\gamma + (a^2 - R^2)^{-5/2}\gamma^2 - \cdots, \tag{42}$$

and upon applying the wave operator term by term to both sides of (42) [using (41)] one obtains

$$\begin{aligned} &-\gamma^{-3/2}\nabla^2(a^2 - R^2)^{1/2} \\ &= \gamma^{-3/2}\left(2a\frac{\partial}{\partial\sigma}(a^2 - R^2)^{-1/2} + A(a^2 - R^2)^{-1/2}\right) \\ &+ \gamma^{-1/2}\left(2a\frac{\partial}{\partial\sigma}(a^2 - R^2)^{-3/2}\right. \\ &\left. + (A + 2)(a^2 - R^2)^{-3/2} + \nabla^2(a^2 - R^2)^{-1/2}\right) \\ &- 3\gamma^{1/2}\left(2a\frac{\partial}{\partial\sigma}(a^2 - R^2)^{-5/2}\right. \\ &\left. + (A + 4)(a^2 - R^2)^{-5/2} + \tfrac{1}{3}\nabla^2(a^2 - R^2)^{-3/2}\right) \\ &+ 5\gamma^{3/2}\left(2a\frac{\partial}{\partial\sigma}(a^2 - R^2)^{-7/2}\right. \\ &\left. + (A + 6)(a^2 - R^2)^{-7/2} + \tfrac{1}{5}\nabla^2(a^2 - R^2)^{-5/2}\right) \\ &+ \cdots. \end{aligned} \tag{43}$$

Comparing the coefficients of both sides of (43) gives

$$2a\frac{\partial}{\partial\sigma}(a^2 - R^2)^{-1/2} + A(a^2 - R^2)^{-1/2} = -\nabla^2(a^2 - R^2)^{1/2}, \text{ etc.} \tag{44}$$

With the above observations one is in a position to determine the coefficients $v_n, n \geq 1$, in (39). First,

$$\begin{aligned} &\Box[\pm \pi^{-1}(a^2 - R^2)^{1/2}\Gamma^{-1} + v_0]\gamma^{-1/2} \\ &= \gamma^{-1/2}\nabla^2 v_0 + \gamma^{-3/2}\left(2a\frac{\partial v_0}{\partial\sigma} + Av_0\right. \\ &\left. \pm \pi^{-1}\nabla^2(a^2 - R^2)^{1/2}\right) = \gamma^{-1/2}\nabla^2 v_0, \end{aligned}$$

since the term in large parentheses vanishes because of (44). It follows that

$$\begin{aligned} &\Box[\pm \pi^{-1}(a^2 - R^2)^{1/2}\Gamma^{-1} + v_0 + v_1\gamma]\gamma^{-1/2} \\ &= \gamma^{1/2}\nabla^2 v_1 - \gamma^{-1/2}\left(2a\frac{\partial v_1}{\partial\sigma} + (A + 2)v_1\right. \\ &\left. - \nabla^2 v_0\right). \end{aligned}$$

Now choose v_1 so as to make the last term vanish, that is

$$2a\frac{\partial v_1}{\partial\sigma} + (A + 2)v_1 = \nabla^2 v_0. \tag{45}$$

Equation (45) may be integrated by using the integrating factor $\frac{1}{2}a^{-1/2}J^{1/2}$, so that

$$v_1 = \tfrac{1}{2}a^{-1/2}J^{-1/2}\left(\int^\sigma a^{-1/2}J^{1/2}\nabla^2 v_0 d\sigma + C_1\right), \tag{46}$$

where C_1 denotes an arbitrary constant of integration–possibly a function of Θ and e but independent of σ.

Assuming for the moment that the C_n are determined, the above procedure for v_1 can be repeated for each of the v_n in turn, leading to the differential equation for v_n:

$$\left(2a\frac{\partial}{\partial\sigma} + A + 2n\right) v_n = (2n-1)^{-1}\nabla^2 v_{n-1},$$

which, upon using the integrating factor $\frac{1}{2}a^{n-3/2}J^{1/2}$, yields

$$v_n = [2(2n-1)]^{-1}a^{-n+1/2}J^{-1/2} \times \left(\int^{\sigma} a^{n-3/2}J^{1/2}\nabla^2 v_{n-1}d\sigma + C_n\right), \quad (47)$$

for $n \geq 1$.

With such functions v_n, whatever the integration constants C_n, w given by (39) will, by construction, be at least a formal solution of the homogeneous wave equation (4), and also take on such values on $\gamma = 0$ that Eq. (11) holds. The determination of the constants C_n will be taken up in Sec. 7, after the remaining steps in the formal construction have been completed. Before attacking the general problem, it will be worthwhile to construct the solution of the comparatively simpler problem of the slit, in order to gain insight into the effect of the edges on waves already diffracted.

5. THE SLIT

The construction of the solution of the wave equation (1) in the presence of a plane with an infinite slit will be carried out much as was done in the two-dimensional case.[9] Namely, one first constructs the function ${}_0u_\nu(t, P)$, which is the solution for small enough time t such that the two half-planes have not had time to interact, that is, for $t \leq d$, where d is the width of the slit. For such t, the solution is the same as for the two half-planes separately. Thus, if a_1 and a_2 denote the minimum distances from P to Q via the edges (say, left and right), respectively; if D_i and D_r are the regions (in space–time) from which, respectively, direct and reflected radiation can be received at (t, P); if D_1 and D_2 are the regions from which diffracted radiation can be received via the respective edges, so that

$$D_j = \{(T, Q) : a_j < t - T\};$$

and if one defines

$${}_0g_\nu(t-T, P, Q) = -\frac{1}{4\pi}\begin{cases}(1/R)\delta(t-T-R) & \text{for } (T,Q) \in D_i,\\ (-1)^\nu(1/R^*)\delta(t-T-R^*) & \text{for } (T,Q)\in D_r,\\ -[w_j + (-1)^\nu w_j^*] & \text{for } (T,Q)\in D_j, \quad j = 1 \text{ and } 2,\end{cases}$$

where w_j is defined by (27) with $a = a_j, j = 1$ and 2; then one has

$${}_0u_\nu(t, P) = \int f(T, Q)\ {}_0g_\nu(t-T, P, Q)dTdQ, \quad (48)$$

where the integration is over D_i, D_r, D_1, and D_2 separately and the results are added. Verification that (48) satisfies the differential equation (1) involves the same computations as for the case of the half-plane and are accordingly omitted here. To verify that the boundary value of ${}_0u_1$ (or $(d/dn)_0u_2$) is zero for $0 \leq t < d$, but not for $t \geq d$, one notes that on the plane, $R = R^*$, $\delta = \delta^*$, $(d/dn)\delta = -(d/dn)\delta^*$, while $w_1 = -w_1^*$, $(d/dn)w_1 = (d/dn)w_1^*$, $w_2 = w_2^*$, $(d/dn)w_2 = (d/dn)w_2^*$ to the right of the left-hand edge, and $w_2 = -w_2^*$, $(d/dn)w_2 = (d/dn)w_2^*$, $w_1 = w_1^*$, $(d/dn)w_1 = -d/dn\ w_1^*$ to the left of the right-hand edge. Thus the boundary condition is satisfied until such time as D_1 meets the right-hand edge or D_2 meets the left-hand edge, that is, at least until $t \geq d$.

It follows that one needs to construct a function ${}_1u_\nu(t, P)$ which satisfies the homogeneous wave equation (4) and which takes on the negative of the values of ${}_0u_1$, if $\nu = 1$ [or has normal derivative equal to $-(d/dn)_0u_2$ if $\nu = 2$] on the slit plane, in, which cases the sum

$$u_\nu = {}_0u_\nu + {}_1u_\nu$$

will be the solution of the problem. Such a function ${}_1u_\nu$ can indeed be constructed by using the terms in (15) corresponding to a nonzero boundary condition. (One needs a pair of such boundary integrals, one for each half-plane.)

Suppose one uses rectangular coordinates, with the slit being the set of points $y = 0, -\frac{1}{2}d \leq x \leq \frac{1}{2}d$. Then the value of ${}_0u_1$ on the slit plane will be

$${}_0u_1(t, x, 0, z) = \frac{1}{2\pi}\int f(T, X, Y, Z)w_i(t-T, x, 0, z; X, Y, Z)dTdXdYdZ$$

over $\begin{cases}0 \leq T < t - (\{x + \frac{1}{2}d + [(X+\frac{1}{2}d)^2 + Y^2]^{1/2}\}^2 + (z-Z)^2)^{1/2}, & x \geq \frac{1}{2}d, \quad i = 1,\\ 0 \leq T < t - (\{-x + \frac{1}{2}d + [(X-\frac{1}{2}d)^2 + Y^2]^{1/2}\}^2 + (z-Z)^2)^{1/2}, & x \leq -\frac{1}{2}d, \quad i = 2,\end{cases}$

and

$${}_1u_1(t, x, y, z) = -\frac{1}{2\pi}\frac{\partial}{\partial y}\int {}_0u_1(t', x', 0, z')[(x-x')^2 + y^2 + (z-z')^2]^{-1/2} \times \delta[t - t' - [(x-x')^2 + y^2 + (z-z')^2]^{1/2}\}\, dt'dx'dz', \quad (49)$$

over $0 \le t' < t - [(x-x')^2 + y^2 + (z-z')^2]^{1/2}$, $x' \ge \frac{1}{2}d$, and $x' \le -\frac{1}{2}d$,

$$-\frac{1}{\pi}\frac{\partial}{\partial y}\int {}_0u_1(t',x',0,z')w_j(t-t',x,y,z;x',0,z')dt'dx'dz'$$

over $\begin{cases} 0 \le t' < t - (\{x' - \frac{1}{2}d + [(x-\frac{1}{2}d)^2 + y^2]^{1/2}\}^2 + (z'-z)^2)^{1/2}, & x' \ge \frac{1}{2}d, \quad j = 2, \\ 0 \le t' < t - (\{-x' - \frac{1}{2}d + [(x+\frac{1}{2}d)^2 + y^2]^{1/2}\}^2 + (z'-z)^2)^{1/2}, & x' \le -\frac{1}{2}d, \quad j = 1. \end{cases}$

Hence one can write

$${}_1u_1(t,x,y,z) = \int f(T,X,Y,Z)\,{}_1g_1(t-T,x,y,z;X,Y,Z)dTdXdYdZ,$$

where

$${}_1g_1(t-T,x,y,z;X,Y,Z) \equiv -\frac{1}{4\pi^2}\frac{\partial}{\partial y}\int w_i(t'-T,x',0,z';X,Y,Z)[(x-x')^2 + y^2 + (z-z')^2]^{-1/2}$$
$$\times \delta(t-t' - [(x-x')^2 + y^2 + (z-z')^2]^{1/2})dt'dx'dz'$$

over $\begin{cases} T + (\{x' + \frac{1}{2}d + [(X+\frac{1}{2}d)^2 + Y^2]^{1/2}\}^2 + (z'-Z)^2)^{1/2} \le t' < t - [(x-x')^2 + y^2 + (z-z')^2]^{1/2}, \\ \quad x' \ge \frac{1}{2}d, \quad i = 1 \\ T + (\{-x' + \frac{1}{2}d + [(X-\frac{1}{2}d)^2 + Y^2]^{1/2}\}^2 + (z'-Z)^2)^{1/2} \le t' < t - [(x-x')^2 + y^2 + (z-z')^2]^{1/2}, \\ \quad x' \le -\frac{1}{2}d, \quad i = 2, \end{cases}$

$$-\frac{1}{2\pi^2}\frac{\partial}{\partial y}\int w_i(t'-T,x',0,z';X,Y,Z)w_j(t-t',x,y,z;x',0,z')dt'dx'dz'$$

over $\begin{cases} T + (\{x' + \frac{1}{2}d + [(X+\frac{1}{2}d)^2 + Y^2]^{1/2}\}^2 + (z'-Z)^2)^{1/2} \le t' < t - (\{x' - \frac{1}{2}d + [(x-\frac{1}{2}d)^2 \\ \quad + y^2]^{1/2}\}^2 + (z'-z)^2)^{1/2}, x' \ge \frac{1}{2}d, \quad i = 1, \quad j = 2, \\ T + (\{-x' + \frac{1}{2}d + [(X-\frac{1}{2}d)^2 + Y^2]^{1/2}\}^2 + (z'-Z)^2)^{1/2} \le t' < t - (\{-x' - \frac{1}{2}d + [(x+\frac{1}{2}d)^2 \\ \quad + y^2]^{1/2}\}^2 + (z'-z)^2)^{1/2}, x' \le -\frac{1}{2}d, \quad i = 2, \quad j = 1. \end{cases}$

The case of boundary condition (3.2) is quite similar and it turns out that

$${}_1g_1(t-T,x,y,z;X,Y,Z) = \frac{\partial}{\partial y}\,{}_1g_2(t-T,x,y,z;X,Y,Z).$$

Then the Green's function for the slit problem in the case $U_\nu \equiv 0$ is given by

$$g_\nu = {}_0g_\nu + {}_1g_\nu \tag{50}$$

and

$$u_\nu \equiv \int f g_\nu dTdQ \tag{51}$$

is the solution being sought.

Notice that the solution for nonzero boundary conditions may be obtained by adding to the rhs of (51) the quantity

$$-\frac{1}{2\pi}\int_{\pi_i} U_1 \frac{d}{d\nu}\frac{\delta}{R}dS - \frac{1}{\pi}\int_{\pi_{d_1}} U_1 \frac{dw_1}{d\nu}dS - \frac{1}{\pi}$$
$$\times \int_{\pi_{d_2}} U_1 \frac{dw_2}{d\nu}dS \quad \text{if} \quad \nu = 1$$

or

$$\frac{1}{2\pi}\int_{\pi_i}\frac{\delta}{R}U_2 dS + \frac{1}{\pi}\int_{\pi_{d_1}} w_1 U_2 dS + \frac{1}{\pi}\int_{\pi_{d_2}} w_2 U_2 dS$$
$$\text{if} \quad \nu = 2,$$

where π_i denotes the portion of the slit plane π which is included in the boundary of D_i, π_{d_j} denotes the upper side of the portion of π included in the boundary of D_j, $j = 1$ and 2. [Each of these terms is a solution of (4) which vanishes and has vanishing 'normal' derivative in the plane of π, except on that portion indicated as the range of integration.]

Verification that (51), with g_ν defined by (50), is correct can be carried out in manner similar to that done[9] for the corresponding two-dimensional problem and hence will not be done here. One need only be able to observe that the rhs of (49), with an arbitrary function ${}_0u_1(t',x',0,z')$ in the integrand does in fact give a solution of the homogeneous wave equation (4) which has continuous first partial derivatives everywhere save on the plane itself, and there the solution takes on the value of ${}_0u_1(t,x,0,z)$.

6. DIFFRACTION BY PLANE CURVED EDGE

In resuming the discussion of the problem of diffraction for a planar screen having a curved edge,

there are a few observations which will be useful. One should notice that the function ${}_1g_\nu$ above is nonzero inside the region bounded by the secondary diffraction wavefronts. That is to say, the surface of D_1 (and likewise D_2) may be thought of as a primary diffraction wavefront which, upon striking the opposite edge of the slit, gives rise to a secondary diffraction wavefront, and it is within this region that ${}_1g_\nu$ is nonzero. Furthermore ${}_1g_\nu$ is a solution of (4).

If one were to apply the fundamental formula (F) to the region D_1 when the time t is larger than d, the width of the slit, one would find that D_1 would for some points P meet the other edge, and hence its boundary would include portions of the other-half-plane. This would introduce an additional surface integral which would require another equation for its elimination. This additional equation would be obtained by applying (F) to the region bounded by this secondary diffraction wavefront with an 'appropriate' solution of (4). This, it would turn out, is essentially ${}_1g_2$ (or half it, the other half is due to the secondary wave from D_2). A similar thing does not occur when this secondary wavefront reaches the edge. No effect of a tertiary diffraction wavefront or so on *ad infinitum* has to be considered, because the relevant portion of ${}_1g_2$ is zero and has zero normal derivative in the extension of the corresponding half-plane.

This observation is important in the treatment of the case of a planar screen with a curved edge where the primary diffraction wave from one part of the edge can meet another part of the edge. The observation implies that the only essential new feature the curved edge has different from the slit (other than geometrical) occurs in the construction of an elementary solution of the wave equation (4) corresponding to a primary diffraction wave. A method of obtaining this solution was described in Sec. 4. One may now complete the solution of the diffraction problem for a curved edge much as was done for the slit.

Define

$$ {}_0g_\nu(t-T, P, Q) = -(4\pi)^{-1} \times \begin{cases} R^{-1}\delta(t-T-R) & \text{for} \quad (T, Q) \in D_i \\ (-1)^\nu R^{*-1}\delta(t-T-R^*) & \text{for} \quad (T, Q) \in D_r, \\ -(w + (-1)^\nu w^*) & \text{for} \quad (T, Q) \in D_d, \end{cases} $$

where by D_d is meant the set of all points (T, σ, Θ, e) from which diffracted radiation can be received at time t at the point P. Thus

$$ D_d = \{(T, \sigma, \Theta, e) : a < t - T\}. $$

One notices that to any one point Q of Euclidian three-space there may correspond more than one "point" (σ, Θ, e) in the space consisting of all such triads. Such "points" are to be considered distinct, just as the points in D_r are considered distinct, from those in D_i, even though the two regions overlap in four-space. Then

$$ {}_0u_\nu(t, P) = \int f(T, Q)\, {}_0g_\nu(t - T, P, Q)\, dT\, dQ, $$

where the integration is over D_i, D_r, and D_d separately and the results are added and where dQ is the spatial integral, i.e., $Jd\sigma d\Theta de$ for D_d, is a solution of the problems (1), (3), with $U_\nu \equiv 0$, for times t sufficiently small that no secondary diffraction occurs.

For larger t, one needs to define an additional function ${}_1g_\nu$ just as in the case of a slit by

$$ {}_1g_2(t - T, P, Q) \equiv -(4\pi^2)^{-1} \int_{\pi_i} w(T' - T, Q', Q) R'^{-1} \times \delta(t - T' - R')\, dT'\, dS' - (2\pi^2)^{-1} \times \int_{\pi_d} w(T' - T, Q', Q)\, w(t - T', P, Q')\, dT'\, dS', $$

and

$$ {}_1g_1 = \frac{\partial}{\partial y}\, {}_1g_2, $$

where $R' = \overline{Q'Q}$ and where dS' denotes an element of surface on the plane π.

7. THE EDGE CONDITION AND THE INTEGRATION CONSTANTS

The functions $v_n, n \geq 1$, of Sec. 4 were defined by Eq. (47) except that the integration constants C_n and the lower limit of integration have not yet been specified. What should now be done is to show that the "edge condition"

$$ u_\nu(t, P) < +\infty \tag{52} $$

at the edge of π determine the C_n (and hence the v_n) uniquely and that the resulting series for w actually does converge in some region, thereby establishing the existence of a solution $u_\nu(t, P)$ to the stated problem. Unfortunately the existence problem has not yet been resolved, being apparently much more difficult than the uniqueness problem. However, assuming existence [i.e., that for some choice of constants C_n the resulting series converges to a function w satisfying the edge condition (52)] it will be shown that the C_n are uniquely determined. Namely, that all the $C_n = 0$ and the lower limit of integration in (47) is zero, so that one has

$$ v_n = [2(2n-1)]^{-1} a^{-n+1/2} J^{-1/2} \int_0^\sigma a^{n-3/2} J^{1/2} \nabla^2 \times v_{n-1}\, d\sigma, \quad n \geq 1, \tag{53} $$

To see this, note that if w remains finite at the edge, then from (34) so does W, and from (39) one has

$$ W = \pm \pi^{-1}(a^2 - R^2)^{1/2} + \Gamma(v_0 + v_1\gamma + \cdots) $$

for γ small enough,

from which it obviously follows that each of the

v_n must remain finite (since the v_n are independent of t, T.) But since $v_0 = O(\sigma^{1/2})$ as $\sigma \to 0$, is as easy to verify from Eqs. (37) and (40), and $J = O(\sigma)$, and since the effect of the Laplacian can only be to reduce the powers of σ by 0, 1, or 2, it is clear that v_1 cannot remain finite at $\sigma = 0$ unless the quantity in large parentheses in (46) vanishes there, which implies that the integral is actually of order $\sigma^{1/2}$ (or higher) and not of order σ^{-1} or $\ln\sigma$, as at first sight appears possible. This means that the lower limit of integration can actually be taken to be zero and that then $C_1 = 0$ also. The same argument can clearly be repeated for each v_n in turn, yielding (53) as stated.

8. A SPECIAL CASE OF DIFFRACTION BY A CIRCULAR HOLE

In the case of a circular hole, if the observation point P is on the axis of the hole, one can carry out the computations involved in the construction of the function denoted above by w; these have been done in part. For example, if the radius of the hole be denoted by c then in Eqs. (30) one has

$$\alpha(e) = c\cos e/c, \quad \beta(e) = c\sin e/c,$$
$$\varphi(e) = \pi/2 \text{ for all } e,$$

also

$$A = -1 + 2a\sigma^{-1} - ac\sigma^{-1}\rho^{-1}, \quad \rho = c + \sigma\cos\Theta,$$
$$J = \sigma c^{-1}\rho.$$

Hence, with $\overline{W}$ from Eqs. (36) and (39), one finds

$$v_0 = \pm\,\pi^{-1}(a^2 - R^2)^{-1/2}(M - 1),$$
$$M = (-a\rho^{-1}\cos\theta)^{1/2}.$$

From here on the computations become rather tedious, but one can obtain

$$\nabla^2 v_0 = \pm\,\pi^{-1}(a^2 - R^2)^{-1/2}[\tfrac{1}{4}(\rho^{-2} - a^{-2})M - (a^2 - R^2)^{-1}\sigma_0\sigma^{-1}(M-1)^2],$$

which is actually $O(\sigma^{-1/2})$ as $\sigma \to 0$, and not $O(\sigma^{-2/3})$ as appears at first sight. Thus $v_1 = O(\sigma^{1/2})$.

The leading term of $\nabla^2 v_1$ has been computed and found to be $O(\sigma^{-1/2})$; hence $v_2 = O(\sigma^{1/2})$ also.

APPENDIX:

The Jacobian J of the transformation (29) is clearly equal to

$$J = \sigma\sin\varphi \begin{vmatrix} -\alpha'\cos\varphi + \beta'\sin\varphi\cos\Theta & -\beta'\sin\Theta & \alpha' + \sigma\dfrac{\partial}{\partial e}(-\alpha'\cos\varphi + \beta'\sin\varphi\cos\Theta) \\ \sin\varphi\sin\Theta & \cos\Theta & \sigma\dfrac{\partial}{\partial e}(\sin\varphi)\sin\Theta \\ -\beta'\cos\varphi - \alpha'\sin\varphi\cos\Theta & \alpha'\sin\Theta & \beta' - \sigma\dfrac{\partial}{\partial e}(\beta'\cos\varphi + \alpha'\sin\varphi\cos\Theta) \end{vmatrix},$$

which straightforward computation shows (φ depending on e) to be just

$$J = \sigma\sin\varphi(\sin\varphi + \sigma B),$$

where

$$B = \sigma_0^{-1}\sin\varphi - \alpha''(\beta')^{-1}(\cos\theta + \cos\Theta).$$

Notice that on the shadow surface this reduces to just

$$J = a\sigma\sigma_0^{-1}\sin^2\varphi.$$

Proof that $\dfrac{\partial J}{\partial S} = J\nabla^2 a$

First note that from the above expression for J, one has

$$J^{-1}\frac{\partial J}{\partial S} = 2\sigma^{-1} - J^{-1}\sin^2\varphi.$$

On the other hand, differentiation of (30) yields

$$\frac{\partial^2 a}{\partial X^2} = -(X-\alpha)^2\sigma^{-3} + \sigma^{-1} - \{\alpha'\sigma^{-1} + (X-\alpha)\sigma^{-3}[(\alpha - X)\alpha' + (\beta - Z)\beta']\}\frac{\partial e}{\partial X},$$

$$\frac{\partial^2 a}{\partial Y^2} = -Y^2\sigma^{-3} + \sigma^{-1} - \{Y\sigma^{-3}[(\alpha - X)\alpha' + (\beta - Z)\beta']\}\frac{\partial e}{\partial Y},$$

$$\frac{\partial^2 a}{\partial Z^2} = -(Z-\beta)^2\sigma^{-3} + \sigma^{-1} - \{\beta'\sigma^{-1} + (Z-\beta)\sigma^{-3}[(\alpha - X)\alpha' + (\beta - Z)\beta']\}\frac{\partial e}{\partial Z};$$

hence

$$\nabla^2 a = 2\sigma^{-1} - \sigma^{-1}\left(\alpha'\frac{\partial e}{\partial X} + \beta'\frac{\partial e}{\partial Z}\right) - \sigma^{-2}[(\alpha - X)\alpha' + (\beta - Z)\beta']\left((X-\alpha)\sigma^{-1}\frac{\partial e}{\partial X} + Y\sigma^{-1}\frac{\partial e}{\partial Y} + (Z-\beta)\sigma^{-1}\frac{\partial e}{\partial Z}\right).$$

But the quantity in large parenthesis is just $(\partial/\partial\sigma)\,e(\sigma, \Theta, e)$, which is zero since σ, Θ, e are independent. Hence

$$\nabla^2 a = 2\sigma^{-1} - \sigma^{-1}\left(\alpha'\frac{\partial e}{\partial X} + \beta'\frac{\partial e}{\partial Z}\right).$$

But

$$\frac{\partial e}{\partial X} = J^{-1}\frac{\partial(Y,Z)}{\partial(\sigma,\Theta)} = \sin\varphi\,[\alpha'\sin\varphi\sin^2\Theta - \cos\Theta(\beta'\cos\varphi - \alpha'\sin\varphi\cos\Theta)],$$

$$\frac{\partial e}{\partial Z} = J^{-1}\frac{\partial(X,Y)}{\partial(\sigma,\Theta)} = \sin\varphi\,[\beta'\sin\varphi\sin^2\Theta + \cos\Theta(\alpha'\cos\varphi + \beta'\sin\varphi\cos\Theta)],$$

so

$$\alpha'\frac{\partial e}{\partial X} + \beta'\frac{\partial e}{\partial Z} = \sigma J^{-1}\sin^2\varphi,$$

from which the stated result is obvious.

* Work supported by the U. S. Atomic Energy Commission.

1 J. B. Keller, J. Opt. Soc. Am. 52, 116 (1962).
2 J. B. Keller, in *Diffraction by Polygonal Cylinders. Electromagnetic Waves*, (Ed. R. E. Langer) (Univ. Wisconsin Press, Madison, Wisc., 1962).
3 R. M. Lewis and J. B. Keller, "Asymptotic Methods for Partial Differential Equations: The Reduced Wave Equations and Maxwell's Equations." Research Report EM-194 (New York University, 1964) (unpublished).
4 D. S. Ahluwalia and R. M. Lewis, J. Inst. Math. Appl. 5, 113 (1969).
5 R. M. Lewis, "The Progressing Wave Formalism," in *Proceedings of the Symposium on Quasi-Optics* (Polytechnic Press, Polytechnic Institute of Brooklyn, Brooklyn, N.Y., 1964).
6 J. R. Wait, Can. J. Phys. 35, 693 (1957).
7 J. Hadamard, *Lectures on Cauchy's Problem in Linear Partial Differential Equations* (Dover, New York, 1952), pp. 58–64, 92–98.
8 I. M. Gelfand and G. E. Shilov, *Generalized Functions* (Academic, New York, 1964), Vol. I.
9 P. B. Bailey and G. E. Barr, J. Math. Phys. 10, 1906 (1969).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 10 OCTOBER 1971

Another Bound on the Absorptive Part of Elastic Scattering Amplitudes

R. Savit and R. Blankenbecler
Stanford Linear Accelerator Center, Stanford University, Stanford, California 94305

and

M. B. Einhorn
Lawrence Radiation Laboratory, Berkeley, California 94720

(Received 25 January 1971)

Via Lagrange multipliers for inequality constraints, an upper bound on the absorptive part of the elastic scattering amplitudes is derived assuming unitarity, a fixed total and elastic cross section, and the condition that the partial waves decrease monotonically with increasing angular momentum. Numerical results are given.

1. INTRODUCTION

Consider the elastic scattering of equal mass particles of spin zero. Given the total cross section and the elastic cross section, as well as the unitarity requirement on the partial wave amplitudes, how large can the absorptive part of the elastic scattering amplitude become at any given scattering angle? This problem has been solved by Singh and Roy,[1] and the maximum value has been compared with experimental differential cross sections at high energies and for small scattering angles on the basis of several further assumptions: (i) At high energies, the equal mass assumption can be relaxed. (ii) The unpolarized differential cross sections are independent of the spin of the external particles and, hence, the spin-zero bound applies. (iii) The amplitude, in the region of the diffraction peak, is purely imaginary. The comparison[1] with experimental data is rather good for small angles, but for larger angles the data fall far below the calculated bound.

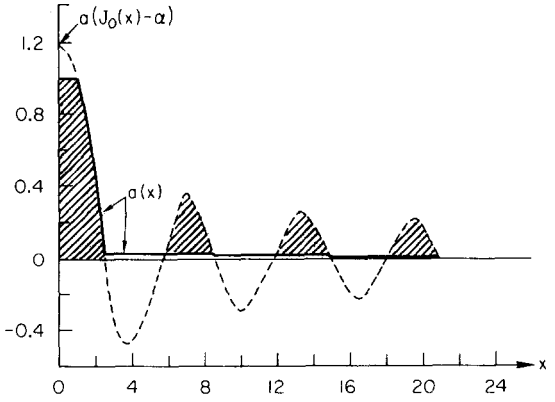


FIG. 1. An example of a local maximum.

The distribution of partial wave amplitudes which achieves this bound looks very much like a Fresnel zone plate, carefully constructed to maximize the scattering in the given direction. The distribution is illustrated as the shaded region of Fig. 1, the details of which will be explained later. The larger the angle, the more zones are required. More conventional models of matter would have a central core surrounded perhaps by successively less absorptive regions. A particularly simple way to implement this intuition is to require the imaginary parts of the partial wave amplitudes to decrease monotonically with increasing angular momentum.[2] This is not unreasonable for energies above resonances. Adding this assumption to those given above should yield a better bound at larger

But

$$\frac{\partial e}{\partial X} = J^{-1}\frac{\partial(Y,Z)}{\partial(\sigma,\Theta)} = \sin\varphi\,[\alpha'\sin\varphi\sin^2\Theta - \cos\Theta(\beta'\cos\varphi - \alpha'\sin\varphi\cos\Theta)],$$

$$\frac{\partial e}{\partial Z} = J^{-1}\frac{\partial(X,Y)}{\partial(\sigma,\Theta)} = \sin\varphi\,[\beta'\sin\varphi\sin^2\Theta + \cos\Theta(\alpha'\cos\varphi + \beta'\sin\varphi\cos\Theta)],$$

so

$$\alpha'\frac{\partial e}{\partial X} + \beta'\frac{\partial e}{\partial Z} = \sigma J^{-1}\sin^2\varphi,$$

from which the stated result is obvious.

* Work supported by the U. S. Atomic Energy Commission.

1 J. B. Keller, J. Opt. Soc. Am. 52, 116 (1962).
2 J. B. Keller, in *Diffraction by Polygonal Cylinders. Electromagnetic Waves*, (Ed. R. E. Langer) (Univ. Wisconsin Press, Madison, Wisc., 1962).
3 R. M. Lewis and J. B. Keller, "Asymptotic Methods for Partial Differential Equations: The Reduced Wave Equations and Maxwell's Equations." Research Report EM-194 (New York University, 1964) (unpublished).
4 D. S. Ahluwalia and R. M. Lewis, J. Inst. Math. Appl. 5, 113 (1969).
5 R. M. Lewis, "The Progressing Wave Formalism," in *Proceedings of the Symposium on Quasi-Optics* (Polytechnic Press, Polytechnic Institute of Brooklyn, Brooklyn, N.Y., 1964).
6 J. R. Wait, Can. J. Phys. 35, 693 (1957).
7 J. Hadamard, *Lectures on Cauchy's Problem in Linear Partial Differential Equations* (Dover, New York, 1952), pp. 58–64, 92–98.
8 I. M. Gelfand and G. E. Shilov, *Generalized Functions* (Academic, New York, 1964), Vol. I.
9 P. B. Bailey and G. E. Barr, J. Math. Phys. 10, 1906 (1969).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 10 OCTOBER 1971

Another Bound on the Absorptive Part of Elastic Scattering Amplitudes

R. Savit and R. Blankenbecler
Stanford Linear Accelerator Center, Stanford University, Stanford, California 94305

and

M. B. Einhorn
Lawrence Radiation Laboratory, Berkeley, California 94720

(Received 25 January 1971)

Via Lagrange multipliers for inequality constraints, an upper bound on the absorptive part of the elastic scattering amplitudes is derived assuming unitarity, a fixed total and elastic cross section, and the condition that the partial waves decrease monotonically with increasing angular momentum. Numerical results are given.

1. INTRODUCTION

Consider the elastic scattering of equal mass particles of spin zero. Given the total cross section and the elastic cross section, as well as the unitarity requirement on the partial wave amplitudes, how large can the absorptive part of the elastic scattering amplitude become at any given scattering angle? This problem has been solved by Singh and Roy,[1] and the maximum value has been compared with experimental differential cross sections at high energies and for small scattering angles on the basis of several further assumptions: (i) At high energies, the equal mass assumption can be relaxed. (ii) The unpolarized differential cross sections are independent of the spin of the external particles and, hence, the spin-zero bound applies. (iii) The amplitude, in the region of the diffraction peak, is purely imaginary. The comparison[1] with experimental data is rather good for small angles, but for larger angles the data fall far below the calculated bound.

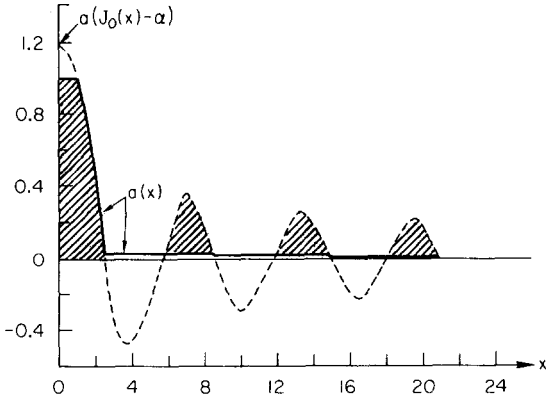


FIG. 1. An example of a local maximum.

The distribution of partial wave amplitudes which achieves this bound looks very much like a Fresnel zone plate, carefully constructed to maximize the scattering in the given direction. The distribution is illustrated as the shaded region of Fig. 1, the details of which will be explained later. The larger the angle, the more zones are required. More conventional models of matter would have a central core surrounded perhaps by successively less absorptive regions. A particularly simple way to implement this intuition is to require the imaginary parts of the partial wave amplitudes to decrease monotonically with increasing angular momentum.[2] This is not unreasonable for energies above resonances. Adding this assumption to those given above should yield a better bound at larger

angles, precisely where the preceding one fails. It is to the solution of this problem that this paper is devoted.

The approach used in the construction of the solution is the method of Lagrange multipliers generalized to include inequality constraints.[3] One of the main purposes of this paper is to illustrate their use in both a discrete and continuous maximization problem of physical interest.

In Sec. 2, the exact problem is formulated and solved by considering the partial wave amplitudes as discrete variables. In order to clarify the solution, the same problem is treated in Sec. 3 by approximating the discrete partial wave series by a continuum and by assuming the scattering angle is small. Section 4 compares the improved bound with experimental data and interprets the results. Finally, in Sec. 5 we summarize our conclusions. In an appendix, a number of sums are tabulated.

2. MATHEMATICAL DETAILS

The mathematical statement of the problem is as follows[4]: Maximize

$$A = \sum(2l + 1)a_l P_l(z)$$

given the total cross section

$$A_0 = \frac{k^2}{4\pi}\sigma_T = \sum(2l + 1)a_l,$$

the elastic cross section

$$\sum_{\text{el}} = \frac{k^2}{4\pi}\sigma_{\text{el}} = \sum(2l + 1)(a_l^2 + r_l^2),$$

and unitarity

$$u_l \equiv a_l - a_l^2 - r_l^2 \geq 0, \qquad l = 0, 1, 2, \cdots .$$

This is essentially the problem solved by Singh and Roy.[1] In addition, we will require that the partial waves decrease monotonically,[2] i.e.,

$$a_l \geq a_{l+1}. \tag{1}$$

However, thinking of the requirements of statistics as well as of most dynamical models, one would like to impose this requirement separately on the even and odd partial waves, i.e.,

$$a_l \geq a_{l+2}. \tag{2}$$

Fortunately, this is a minor complication. To avoid notational confusion, we solve the problem first assuming (i) and then state the result assuming (ii). From a mathematical point of view, these requirements are interesting since they impose a relation between neighboring partial waves, unlike those examples discussed in EB.[3]

To this end, we introduce the auxiliary function

$$\begin{aligned}\mathfrak{L} &= A + \alpha[A_0 - \sum(2l + 1)a_l] \\ &\quad + \frac{1}{2a}\left(\sum_{\text{el}} - \sum(2l + 1)(a_l^2 + r_l^2)\right) \\ &\quad + \sum(2l + 1)\lambda_l u_l + \sum\omega_l(a_l - a_{l+1}),\end{aligned} \tag{3}$$

where, on the basis of the theorem[3] that the multipliers are the rate of change of the maximum with respect to the constraint, we anticipate that

$$0 < \alpha < 1 \quad \text{and} \quad a > 0.$$

Of course, $\lambda_l \geq 0$ and $\omega_l \geq 0$ for all l. Varying with respect to r_l, we find

$$\frac{\delta\mathfrak{L}}{\delta r_l} = (2l + 1)\left(-\frac{1}{a} - 2\lambda_l\right)r_l = 0.$$

This implies $r_l = 0$. Also,

$$\begin{aligned}\frac{\delta\mathfrak{L}}{\delta a_l} &= (2l + 1)\Big(P_l - \alpha - \frac{a_l}{a} + \lambda_l(1 - 2a_l) \\ &\quad + \frac{\omega_l - \omega_{l-1}}{2l + 1}\Big) = 0.\end{aligned} \tag{4}$$

For convenience, we have defined $\omega_{-1} = 0$. First, we must find the *necessary* conditions on a local maximum. The most general form for a local maximum under these assumptions is

$$\begin{aligned}x = (1 = 1 = \cdots = a_{N_0} > a_{N_0+1} \geq \cdots > a_{L_1} = a_{L_1+1} \\ = \cdots = a_{L_1+N_1} > a_{L_1+N_1+1} > \cdots > a_{L_2} = a_{L_2+1} \\ = \cdots = a_{L_2+N_2} > a_{L_2+N_2+1} > \cdots > a_{L_N} = 0 \\ = 0 = \cdots)\end{aligned}$$

In words, the solution is a series of plateaus, on which at least two partial waves have the same value, and regions where successive partial waves are strictly decreasing. Our problem is to determine where the transition points and jumps occur as well as the values of a_l. As in EB, we define three partitions of the partial waves:

$$\begin{aligned}B_1 &= \{l \mid a_l = 1\}, \\ I &= \{l \mid 0 < a_l < 1\}, \\ B_0 &= \{l \mid a_l = 0\}.\end{aligned}$$

A priori, B_1, I, or B_0 could be empty. It follows from our variational equations (4) that, in B_1,

$$\lambda_l = P_l - \alpha - \frac{1}{a} + \frac{\omega_l - \omega_{l-1}}{2l + 1} \geq 0, \tag{5a}$$

in I,

$$\lambda_l = 0, \qquad \frac{\omega_{l-1} - \omega_l}{2l + 1} = P_l - \alpha - \frac{a_l}{a}, \tag{5b}$$

and, in B_0,

$$\lambda_l = \alpha - P_l + \frac{\omega_{l-1} - \omega_l}{2l+1} \geq 0. \tag{5c}$$

Let us first determine the properties of each of the transition points where a_l actually decreases.

(a) $a_{N_0} > a_{N_0+1}$ implies $\omega_{N_0} = 0$; hence, from (5a)

$$\lambda_{N_0} = P_{N_0} - \alpha - \frac{1}{a} - \frac{\omega_{N_0-1}}{2N_0+1} \geq 0;$$

consequently,

$$P_{N_0} - \alpha - \frac{1}{a} \geq \frac{\omega_{N_0-1}}{2N_0+1} \geq 0.$$

In particular, this means that $a(P_{N_0} - \alpha) \geq 1$; so, as in the case without monotonicity,[3] B_1 is empty if $a(1-\alpha) < 1$.

(b) In I, if, for any l, one has $a_{l-1} > a_l > a_{l+1}$, then $\omega_{l-1} = \omega_l = 0$. Hence, from (5b)

$$a_l = a(P_l - \alpha).$$

Of course, by the definition of I, one must have $1 > a(P_l - \alpha) > 0$.

(c) $a_{L_N-1} > a_{L_N} = 0$ implies $\omega_{L_N-1} = 0$; so, from (5c),

$$\lambda_{L_N} = \alpha - P_{L_N} - \frac{\omega_{L_N}}{2L_N+1} \geq 0,$$

and, consequently,

$$\alpha - P_{L_N} \geq \frac{\omega_{L_N}}{2L_N+1} \geq 0.$$

In particular, this means that $\alpha \geq P_{L_N}$.

Now we come to the heart of the problem, how to determine the plateaus of constancy. The three points (a)–(c) immediately above suggest that the multipliers a and α determine where B_1 stops and the value of largest nonzero partial wave amplitude, just as in the case without the monotonicity requirement. In I, we shall find that the transition points are determined *independently* of a and α. In other words, σ_T and σ_{el} determine the size of the core (B_1) and the value of the largest contributing partial wave (B_0). However, in the intermediate region (I), the shape of the distribution is determined solely by the requirement that the strength of the partial wave amplitudes should be monotonically decreasing. To see this, consider any particular plateau

$$a_{L_j-1} > a_{L_j} = \cdots = a_{L_j+N_j} > a_{L_j+N_j+1}.$$

We know that $\omega_{L_j-1} = 0$. Using the difference equation (5b), one can then solve for

$$-\omega_{L_j+n} = \sum_{L_j}^{L_j+n} (2l+1)\left(P_l - \alpha - \frac{a_{L_J}}{a}\right),$$

$$n = 0, \cdots, N_j.$$

Using $\omega_{L_j+N_j} = 0$, we find

$$0 = \sum_{L_j}^{L_j+N_j} (2l+1)\left(P_l - \alpha - \frac{a_{L_j}}{a}\right).$$

It is convenient at this point to define a weighted average of the Legendre polynomials

$${}_N\langle P_l\rangle_M \equiv \sum_{l=M}^{N} (2l+1)P_l \Big/ \sum_{l=M}^{N} (2l+1).$$

Notice that ${}_N\langle P_l\rangle_N = P_N$ and that

$${}_N\langle P_l + c\rangle_M = {}_N\langle P_l\rangle_M + c$$

for any constant c. We have found in Eq. (4) that

$$a_{L_j} = a({}_{L_j+N_j}\langle P_l\rangle_{L_j} - \alpha).$$

Thus the value of the partial wave amplitude is related to the weighted average of the Legendre polynomials across the plateau. The requirement $a_{L_j} > 0$ implies ${}_{L_j+N_J}\langle P_l\rangle_{L_j} > \alpha \geq 0$. Therefore, α determines the largest nonzero partial wave. Inserting this solution for a_{L_j} into the equation for ω_{L_j+n}, we have

$$-\omega_{L_j+n} = \sum_{L_j}^{L_j+n} (2l+1)(P_l - {}_{L_j+N_j}\langle P_l\rangle_{L_j}),$$

$$n = 0, \cdots, N_j. \tag{6}$$

Note that this can also be written as

$$\omega_{L_j+n} = \sum_{L_j+n+1}^{L_j+N_j} (2l+1)(P_l - {}_{L_j+N_j}\langle P_l\rangle_{L_j});$$

$$n = 0, \cdots, N_j - 1, \tag{7}$$

The condition $\omega_{L_j+n} > 0$ therefore implies

$${}_{L_j+N_j}\langle P_l\rangle_{L_j} \geq {}_{L_j+n}\langle P_l\rangle_{L_j} \quad \text{for} \quad n = 0, \cdots, N_j$$

or

$${}_{L_j+N_j}\langle P_l\rangle_{L_j+N_j-n} \geq {}_{L_j+N_j}\langle P_l\rangle_{L_j} \quad \text{for} \quad n = 0, \cdots, N_j.$$

It can be shown that the conditions that $\omega_{L_j-1} = 0$ and $a_{L_j-1} > a_{L_j}$ imply $P_{L_j-1} > {}_{L_j+N_j}\langle P_l\rangle_{L_j}$, while the conditions $\omega_{L_j+N_j} = 0$ and $a_{L_j} > a_{L_j+N_j+1}$ imply

$${}_{L_j+N_j}\langle P_l\rangle_{L_j} > P_{L_j+N_j+1}.$$

We may summarize all of these inequalities as

$$P_{L_j-1} > {}_{L_j+N_j}\langle P_l\rangle_{L_j} > P_{L_j+N_j+1} \tag{8a}$$

and

$${}_{L_j+N_j}\langle P_l\rangle_{L_j+m} \geq {}_{L_j+N_j}\langle P_l\rangle_{L_j} \geq {}_{L_j+n}\langle P_l\rangle_{L_j},$$

$$n, m = 0, 1, \cdots, N_j. \tag{8b}$$

As the derivation shows, either of the two inequalities in (8b) implies the other. Notice that all reference to α and a have disappeared so that the plateau interval $[L_j, L_j + N_j]$ may be determined solely by properties of the Legendre polynomials. Since successive plateaus must be monotonically decreasing, the condition ${}_{L_j+N_j}\langle P_l\rangle_{L_j} \geq \alpha$ is significant only for the last interval, i.e., only for determining B_0. Thus we may determine all possible plateau intervals independently of the multipliers α and a. Since for a maximum α is greater than zero, we may restrict our determination to those for which ${}_{L_j+N_j}\langle P_l\rangle_{L_j} > 0$.

We have now completely characterized the necessary conditions on a local maximum. In summary, the last value N_0 in B_1 must satisfy $a(P_{N_0} - \alpha) \geq 1$. If $a_{l-1} > a_l > a_{l+1}$, then $1 > a_l = a(P_l - \alpha) > 0$. The plateau interval must satisfy the sets of inequalities expressed by Eqs. (8a) and (8b). On a plateau, $1 > a_{L_j} = a({}_{L_j+N_j}\langle P_l\rangle_{L_j} - \alpha) > 0$. Finally, the first partial wave L_N for which a_l vanishes must satisfy $\alpha \geq P_{L_N}$. To determine the local maximum, one must determine the *sufficiency* of these many conditions. We have not been able to show that these conditions uniquely determine the local maximum; indeed, we suspect that one can probably find some angles for which the local maximum is not unique.[5] Given any given scattering angle, one can use the inequalities to determine the plateau intervals. Then, given σ_T and σ_{el}, one can try to determine α and a to satisfy the equality constraints. In practice, it is easier and faster to choose α and a and to then calculate the corresponding σ_T and σ_{el}.

It is a simple matter now to solve the problem where the monotonicity requirement is applied to the even and odd partial waves separately [Eq. (2)]. The general form of the local maximum is

$$\begin{aligned} x^{\text{even}} = (1 = \cdots = a_{2M_0} &> a_{2M_0+2} > \cdots > a_{2K_1} \\ &= a_{2K_1+2} = \cdots = a_{2K_1+2M_1} > a_{2K_1+2M_1+2} \\ &> \cdots > a_{2K_M} = 0 = 0 = \cdots), \end{aligned}$$

$$\begin{aligned} x^{\text{odd}} = (1 = \cdots = a_{2N_0+1} &> a_{2N_0+3} > \cdots > a_{2L_1+1} \\ &= a_{2L_1+3} = \cdots = a_{2L_1+2N_1+1} > a_{2L_1+2N_1+3} \\ &> \cdots > a_{2L_N+1} = 0 = 0 = \cdots). \end{aligned}$$

Correspondingly, we are led to define weighted averages over the even and odd partial waves separately:

$${}_{2N}\langle P_l\rangle_{2M} \equiv \sum_{n=M}^{N} (4n + 1)P_{2n} \Big/ \sum_{n=M}^{N} (4n + 1),$$

$${}_{2N+1}\langle P_l\rangle_{2M+1} \equiv \sum_{n=M}^{N} (4n + 3)P_{2n+1} \Big/ \sum_{n=M}^{N} (4n + 3).$$

The preceding solution then becomes the correct solution to this problem if one everywhere treats the even and odd partial waves separately. For example, for the even case, the last value in B_1 must satisfy $a(P_{2M_0} - \alpha) \geq 1$. The plateau intervals $[2K_j, 2K_j + 2M_j]$ must satisfy

$$P_{2K_j-2} > {}_{2K_j+2M_j}\langle P_l\rangle_{2K_j} > P_{2K_j+2M_j+2},$$

$${}_{2K_j+2M_j}\langle P_l\rangle_{2K_j+2M} > {}_{2K_j+2M_j}\langle P_l\rangle_{2K_j} > {}_{2K_j+2n}\langle P_l\rangle_{2K_j}$$

$$m, n = 0, 1, \cdots, M_j.$$

On the plateau, we find

$$a_{2K_j} = a({}_{2K_j+2M_j}\langle P_l\rangle_{2K_j} - \alpha).$$

Finally, the first partial wave which vanishes satisfies $\alpha \geq P_{2K_M}$. Corresponding statements hold for the odd integers. In an appendix, we record the values for a number of the sums and averages.

3. CONTINUUM APPROXIMATION

It is slightly more convenient in numerical applications to approximate l by a continuous variable since many partial waves contribute in general, and to assume the scattering angle is small. This approximation also makes it possible to carry the solution further analytically and clarifies the result. Since one of our purposes here is to illustrate a somewhat unfamiliar mathematical method, we will solve the problem again in this approximation. Via the standard replacement of the Legendre function by a Bessel function, the auxiliary function $\mathfrak{L}$ of the last section becomes

$$\begin{aligned} 2\sin^2\tfrac{\mu}{2}\,\mathfrak{L} = \int_0^\infty x\,dx J_0(x)a(x) &+ \alpha\left(\frac{\Delta^2}{8\pi}\sigma_T - \int_0^\infty x\,dx a(x)\right) + \frac{1}{2a}\left(\frac{\Delta^2}{8\pi}\sigma_{el} - \int_0^\infty x\,dx a^2(x)\right) \\ &+ \int_0^\infty x\,dx\lambda(x)u(x) - \int_0^\infty dx\,\omega(x)\frac{da(x)}{dx}, \end{aligned}$$

where

$$\Delta^2 \equiv -t = 4K^2\sin^2\tfrac{1}{2}\mu, \qquad x \simeq (2l + 1)\sin\tfrac{1}{2}\mu,$$

$$\omega(x) \equiv 2\sin^2\tfrac{1}{2}\mu\,\omega_l.$$

(For simplicity we have set $r_l = 0$ at the outset.) One consequence of replacing the Legendre polynomials by Bessel functions has been to shift the dependence on the momentum transfer to the

boundary conditions. [Compare Eq. (3).] It is important to remember that this approximation is best for small angles and when many partial waves contribute to the sums. Notice that the monotonicity requirement, Eq. (1) naturally translates into a negativity condition on the slope of the partial wave amplitude. One could, of course, also generalize the monotonicity condition separately applied to the even and odd partial waves by defining partial wave amplitudes of even and odd signature, but for simplicity, we will ignore this alternative. We will also assume that $a(x)$ is a continuous function, so the most general behavior possible has regions where $a(x)$ is strictly decreasing separated by regions of constant $a(x)$.

Formally, $\mathfrak{L}$ is a function of $a(x)$ and its first derivative da/dx; thus the maximum satisfies the usual Euler–Lagrange equations. One finds

$$J_0(x) - \alpha - \frac{a(x)}{a} + \lambda[1 - 2a(x)] + \frac{1}{x}\frac{d\omega}{dx} = 0. \tag{9}$$

As before, we label the three partitions of the solution B_1, I, and B_0. The obvious analogs of Eqs. (5a), (5b), and (5c) are, respectively,

$$B_1: \quad \lambda(x) = J_0(x) - \alpha - \frac{1}{a} + \frac{1}{x}\frac{d\omega}{dx} \geq 0, \tag{10a}$$

$$I: \quad \lambda(x) = 0, \quad \frac{-1}{x}\frac{d\omega}{dx} = J_0(x) - \alpha - \frac{a(x)}{a}, \tag{10b}$$

$$B_0: \quad \lambda(x) = \alpha - J_0(x) - \frac{1}{x}\frac{d\omega}{dx} \geq 0. \tag{10c}$$

On any interval $y_i \leq x \leq x_{i+1}$ on which $a(x)$ is strictly decreasing, $\omega(x) = 0$ and, hence, $d\omega/dx = 0$. Therefore,

$$a(x) = a[J_0(x) - \alpha]. \tag{11}$$

Continuity of $a(x)$ then implies that

$$1 = a[J_0(y_0) - \alpha], \tag{12}$$

where y_0 is the largest value of x in B_1 and

$$0 = J_0(x_N) - \alpha, \tag{13}$$

where x_N is the smallest value in B_0. Thus, the equality multipliers a and α determine the sets B_1 and B_0. Of course, if $1 > a(1 - \alpha)$, then B_1 is empty.

Let us further explore the intervals $x_i \leq x \leq y_i$ in I, on which $a(x)$ is constant. These are surrounded by intervals on which $a(x)$ is strictly decreasing and, by Eq. (11), $a(x) = a[J_0(x) - \alpha]$ on the *surrounding* intervals. However, since $a(x_i) = a(y_i)$, continuity of $a(x)$ then implies $J_0(x_i) = J_0(y_i)$. It follows from (10b) that $\omega(x)$ is continuous in I and, since ω vanishes on the surrounding intervals, we have $\omega(x_i) = \omega(y_i) = 0$. Inside the plateau interval

$$\frac{d\omega}{dx} = x\left(\frac{a(x_i)}{a} + \alpha - J_0(x)\right) = x[J_0(x_i) - J_0(x)].$$

Using $\omega(x_i) = 0$, this may be easily integrated to obtain $\omega(x)$. The condition $\omega(y_i) = 0$ then leads to

$$J_0(x_i) = {}_{y_i}\langle J_0\rangle_{x_i},$$

where, as in the previous section, we define the weighted average

$${}_y\langle J_0\rangle_x \equiv \int_x^y J_0(z)zdz \Big/ \int_x^y zdz.$$

(See the Appendix for the explicit evaluation.) In summary, the end points of the plateau interval satisfy

$$J_0(x_i) = {}_{y_i}\langle J_0\rangle_{x_i} = J_0(y_i), \tag{14a}$$

which may be compared to Eq. (8a) for the discrete case. One may further exploit the condition $\omega(x) \geq 0$ inside the interval to obtain the analog of Eq. (8b),

$${}_{y_i}\langle J_0\rangle_y \geq {}_{y_i}\langle J_0\rangle_{x_i} \geq {}_x\langle J_0\rangle_{x_i}, \quad \text{for } x_i \leq x, y \leq y_i. \tag{14b}$$

One can show, however, that these inequalities in fact *follow* from (14a), provided that the plateau interval extends over only one cycle of $J_0(x)$. This makes finding the plateau intervals easier in the continuous case than in the discrete case. And, as before, the determination of the possible intervals do not involve a and α.

Given the multipliers a and α, the necessary conditions given above are also sufficient to determine the solution. These multipliers, in turn, are to be determined from the equality constraints

$$\frac{\Delta^2}{8\pi}\sigma_T = \tfrac{1}{2}y_0^2 + \int_{y_0}^{x_N} xdxa(x), \tag{15}$$

$$\frac{\Delta^2}{8\pi}\sigma_{el} = \tfrac{1}{2}y_0^2 + \int_{y_0}^{x_N} xdxa^2(x), \tag{16}$$

where

$$a(x) = \begin{cases} a[J_0(x) - \alpha], & y_i \leq x \leq x_{i+1}, \\ a[J_0(x_i) - \alpha], & x_i \leq x \leq y_i, \end{cases}$$

and, if B_1 is empty, $y_0 = 0$, whereas, if B_1 is not empty, $1 = a[J_0(y_0) - \alpha]$. In any case, x_N satisfies $\alpha = J_0(x_N)$. A typical solution is indicated in Fig. 1, where the dashed curve is $a[J_0(x) - \alpha]$ and the solid curve is the solution $a(x)$. The shaded regions indicate the partial wave amplitudes in the problem without the monotonicity constraint described in the Introduction.

It is interesting to compare the expressions for the end of B_1 and the start of B_0:

$$J_0(y_0) = \alpha + 1/a \quad \text{and} \quad J_0(x_N) = \alpha.$$

As a becomes large, $J_0(y_0) \simeq J_0(x_N)$. Since both these points must lie on falling part of the curve $a(x)$, one has $y_0 \simeq x_N$ as a gets large. Thus the ratio σ_{el}/σ_T approaches unity as a gets large. On the other hand, if a is smaller than $(1-\alpha)^{-1}$, then B_1 is empty and one expects a small value of the ratio σ_{el}/σ_T.

Given values of σ_T and σ_{el} from experimental data, one can determine a and α and, subsequently, the maximum value of the absorptive part of the scattering amplitude may be computed. We now turn to the numerical computations.

4. NUMERICAL RESULTS AND CONCLUSIONS

Because it is somewhat easier to evaluate, we discuss in detail the results in the continuum approximation. We have compared this to the evaluation in the discrete case and found little difference for the momentum transfers with which we will be concerned below. We believe that for the entire range of data presented, the continuum approximation gives a bound within a few percent of the actual bound, and, as we shall see, it is of little interest to inquire into the precise discrepancy.

To find numerical values for the upper bound, the candidates for plateau intervals must first be determined. The end points (x_i, y_i) of each interval are all quite close to the odd zeros of $J_0(x)$. For example, the first region of constant $a(x)$ extends from $x_1 \simeq 2.35$ to $y_1 \simeq 8.55$. For comparison, the first and third zeros of J_0 occur at 2.40 and 8.65. Values of x_i and y_i for successive intervals lie even closer to the higher odd zeros of $J_0(x)$.

Physical values of σ_T for processes of interest are around 40 mb, and realistic values of σ_{el}/σ_T are in the range $\frac{1}{6}-\frac{1}{3}$. For values of the momentum transfer $\Delta^2 \lesssim 2(\mathrm{GeV}/c)^2$, these constraints can be accommodated with $y_0 \lesssim 1.6$ and α in a range that allows between two and eight plateaus. Since there are two equality constraints, the maximum value of $d\sigma/dt$ (neglecting the real part of the amplitude) will depend, in general, on two variables. It is convenient to choose these to be $\tau = \Delta^2\sigma_T$ and $R = \sigma_{el}/\sigma_T$. Figure 2 shows the maximum value of

$$16\pi \frac{d\sigma}{dt}\Big/\sigma_T^2 = \frac{d\sigma}{dt}\Big/\left(\frac{d\sigma}{dt}\right)_{t=0}$$

versus τ for two values of the ratio R.

In Fig. 3, the solution to this problem is compared with data and with the solutions of the problems discussed by Singh and Roy[1] and by Ravenhall and Pardee.[2] To facilitate comparison with earlier results, we have used for the abscissa the variable $\rho = (\Delta^2\sigma_T^2/4\pi\sigma_{el}) = \tau R/4\pi$. Data for πp and pp scattering are indicated.[6] Curve A is the upper bound derived by Singh and Roy in the absence of monotonicity constraints. Curve B is the bound given by Ravenhall and Pardee for the same problem as we have discussed.[7] The curves C and C' are the same bounds appearing in Fig. 2, only here plotted as a function of ρ rather than of τ. Notice that, in Fig. 2, C' lies below C, while, in Fig. 3, C' lies above C over most of its range.

Inasmuch as the variable ρ has been ascribed some significance,[1,2] let us comment on this variable. Suppose B_1 were empty, so that $y_0 = 0$. Then since $a(x)/a$ is independent of a for all x, the ratio

$$A/A_0 = \int_0^{x_N} x\,dx J_0(x)a(x)\Big/\int_0^{x_N} x\,dx\,a(x)$$

is independent of a and depends only on α. However, setting $y_0 = 0$ in Eqs. (15) and (16), one sees that α is determined by the ratio

$$\left(\frac{\Delta^2\sigma_T}{8\pi}\right)^2\Big/\frac{\Delta^2\sigma_{el}}{8\pi} = \tfrac{1}{2}\rho.$$

Therefore, if B_1 is empty, A/A_0 is a function of ρ only, a property which has been called "universality."[1,2] This result is independent of the monotonicity assumption. However, the actual values for the experimental data usually require that B_1 *not* be empty. Thus, in general, the upper bound derived here will depend on R as well as on ρ. In fact, one can show that

$$\frac{\partial}{\partial R}\left(\frac{A}{A_0}\right) = \frac{4\pi}{\Delta^2\sigma_{el}}\,[2y_0J_1(y_0) - y_0^2J_0(y_0)],$$

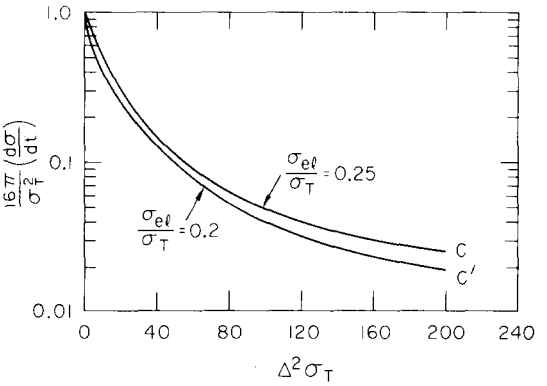


FIG. 2. Upper bounds for $R = 0.20$ and $R = 0.25$.

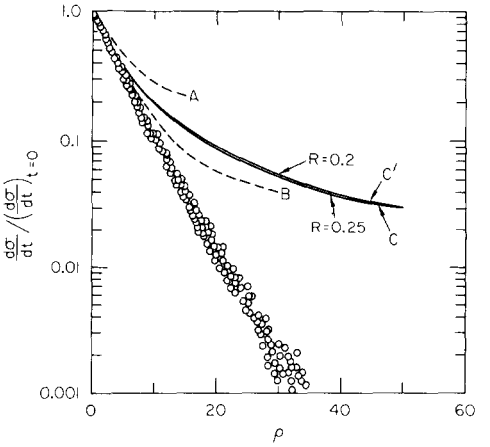


FIG. 3. Comparison of the present solution with previous bounds and with data (see text for full explanation).

where the differentiation is performed for fixed ρ. While this is zero when B_1 is empty, it is not zero in general.

Comparing our upper bound with the data, we see that the addition of the monotonicity requirement has significantly improved the bound of Singh and Roy[1]; however, it still approximates the data only for a very small range of ρ. Already at $\rho = 10$, the bound exceeds the data by a factor of 2; for typical values of σ_T and σ_{el}, this corresponds roughly to $\Delta^2 \lesssim 0.3(\mathrm{GeV}/c)^2$.

5. SUMMARY

One could have hoped that with such general considerations as crude unitarity ($u_l \geq 0$), and the values of the total and elastic cross sections, the shape of the diffraction peak might have been understood. Even assuming the real part of the scattering amplitude is negligible, we found that only for very small values of the momentum transfer does the bound approximate the data. An exponential fit to the data is a good approximation far beyond values for t for which our bound is relevant.

One should note that the values of the a_l which realize the maximum at a particular angle depends on that angle. The upper bound plotted in our graphs is not a reflection of any one set of partial wave amplitudes, but rather, as the angle changes, the values of the a_l also change. Thus, for example, the area under the upper bound could be much larger than σ_{el}, and this turns out to be the case.

We conclude that if the shape of diffraction peaks is to be understood, it is *not* on the basis of the naive considerations discussed here. There is probably a deep dynamical reason both for the small real part (if, indeed, it is small for all t) and for the rapid decrease of the differential cross section with momentum transfer.

ACKNOWLEDGMENTS

Most of this work was performed while one of us (M.B.E.) was still at SLAC. We would like to thank Peete Baer of the SLAC Computer Center for illustrating the efficiency of the APL by finding the plateau intervals in the discrete case. We also wish to acknowledge interesting conversations with Dr. S. Nussinov concerning the statement of this problem.

APPENDIX:

$$\sum_{l=0}^{L}(2l+1)P_l = P'_{L+1} + P'_L,$$

$$\sum_{n=0}^{N}(4n+1)P_{2n} = P'_{2N+1},$$

$$\sum_{n=0}^{N}(4n+1) = (N+1)(2N+1),$$

$$\sum_{n=0}^{N}(4n+3)P_{2n+1} = P'_{2N+2},$$

$$\sum_{n=0}^{N}(4n+3) = (N+1)(2N+3),$$

where

$$P'_l = \frac{d}{dz}P_l(z).$$

Consequently, we have

$${}_{2K+2M}\langle P_l\rangle_{2K} = \frac{P'_{2K+2M+1} - P'_{2K-1}}{(M+1)(4K+2M+1)},$$

$${}_{2L+2N+1}\langle P_l\rangle_{2L+1} = \frac{P'_{2L+2N+2} - P'_{2L}}{(N+1)(4L+2N+3)},$$

$${}_{L+N}\langle P_l\rangle_{L} = \frac{P'_{L+N+1} + P'_{L+N} - P'_L - P'_{L-1}}{(N+2)(2L+N)}.$$

In the continuous case, we have, analogously,

$${}_{y}\langle J_0\rangle_{x} = \frac{2[yJ_1(y) - xJ_1(x)]}{y^2 - x^2}.$$

---

1 V. Singh and S. M. Roy. Phys. Rev. Letters **24**, 28 (1970); Phys. Rev. D **1**, 2638 (1970).

2 A similar generalization has been treated in D. G. Ravenhall and W. J. Pardee, Phys. Rev. D **2**, 589 (1970), although their solution is not the absolute bound.

3 M. B. Einhorn and R. Blankenbecler, Report No. SLAC-PUB-768 (TH), (1970), submitted to Ann. Phys. (N.Y.), hereafter referred to as EB. The present paper can be considered a sequel to Sec. III. B of EB.

4 The notation and conventions are as in EB.

5 However, such angles are probably rare. One could probably show that, among all scattering angles, the set on which the necessary conditions are not also sufficient is of measure zero; we have not tried.

6 Data were taken from the graphs in Ref. 2.

7 There seems to be some confusion in Refs. 1 and 2 on how to treat B_1 and its effect on the solution. In addition, one might wonder why the "upper bound" B determined by Ravenhall and Pardee lies below ours. The answer lies in their identifying the incorrect intervals as plateau intervals, and their curve consequently not being the upper bound. The use of inequality multipliers to impose the monotonicity constraint allows one to uniquely determine the regions of constant $a(x)$.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 10 OCTOBER 1971

Vector Operators and a Polynomial Identity for $SO(n)$

A. J. Bracken and H. S. Green
Department of Mathematical Physics, University of Adelaide, Adelaide, S. Australia 5001*
(Received 5 April 1971)

It is shown that if α denotes an $n \times n$ antisymmetric matrix of operators $\alpha_{pq}, p, q = 1, 2, \ldots, n$, which satisfy the commutation relations characteristic of the Lie algebra of $SO(n)$, then α satisfies an nth degree polynomial identity. A method is presented for determining the form of this polynomial for any value of n. An indication is given of the simple significance of this identity with regard to the problem of resolving an arbitrary n-vector operator into n components, each of which is a vector shift operator for the invariants of the $SO(n)$ Lie algebra.

1. INTRODUCTION

The structure of 3-vector operators in quantum theory was investigated by Dirac,[1] Güttinger, and Pauli,[2] who considered the matrix elements of such operators in an angular momentum basis. Later Wigner[3] indicated the possibility of a systematic treatment of any set of operators transforming according to an irreducible representation of the rotation group, and the calculus of tensor operators was subsequently developed by Racah.[4] Some of the results of these investigations, and their application to calculations in the quantum theory of atomic spectra, can be found in the books by Condon and Shortley[5] and Slater.[6]

Racah[7] and Biedenharn[8] have emphasized the desirability of finding, in the case of other semisimple groups, the generalization of these and other results in the theory of angular momentum, or $SU(2)$. We present here some results in the theory of $SO(n)$, or, more accurately, of its universal covering group [for convenience, this group is subsequently referred to as $SO(n)$], relating in particular to the description of n-vector operators. Even in the much studied case $n = 3$, our approach has, we believe, some novel and attractive features.

We are concerned with the general situation where one is given a set of operators θ_p, α_{qr} $(= -\alpha_{rq})$, $p, q, r = 1, 2, \ldots, n$, satisfying the commutation relations

$$[\alpha_{pq}, \alpha_{rs}] = \delta_{qr}\alpha_{ps} + \delta_{ps}\alpha_{qr} - \delta_{pr}\alpha_{qs} - \delta_{qs}\alpha_{pr}, \tag{1}$$

$$[\theta_p, \alpha_{qr}] = \delta_{pq}\theta_r - \delta_{pr}\theta_q. \tag{2}$$

In particular, the α_{pq} could be anti-Hermitian operators acting in a Hilbert space, in which case they form the generators of a unitary representation, in general reducible, of $SO(n)$. Then $\boldsymbol{\theta}$ $[= (\theta_1, \theta_2, \ldots, \theta_n)]$ is an n-vector operator acting within the corresponding representation space.

In such a case, the Casimir operator $\sigma_2 = \alpha_{pq}\alpha_{qp}$ can be expressed in the form[9]

$$\sigma_2 = 2\Lambda_1(\Lambda_1 + n - 2) + 2\Lambda_2(\Lambda_2 + n - 4) + \cdots + 2\Lambda_m(\Lambda_m + n - 2m), \tag{3}$$

where m is the integral part of $\frac{1}{2}n$ and the eigenvalues λ_i of the operators Λ_i, which serve to label the irreducible components of the representation of $SO(n)$, are either all integers or all half-odd integers and satisfy

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m \geq 0, \quad n = 2m + 1,$$

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{m-1} \geq |\lambda_m| \geq 0, \quad n = 2m. \tag{4}$$

When $n = 3$, an established result[1,2] is that $\boldsymbol{\theta}$ can be resolved into three components, each of which is a 3-vector shift operator for Λ_1, the magnitude of the angular momentum. Thus

$$\boldsymbol{\theta} = \boldsymbol{\theta}^+ + \boldsymbol{\theta}^0 + \boldsymbol{\theta}^-,$$

where

$$\Lambda_1\theta^0 = \theta^0\Lambda_1, \quad \Lambda_1\theta^\pm = \theta^\pm(\Lambda_1 \pm 1).$$

The results obtained by Bhabha,[10] in investigations of 4-vector operators within finite-dimensional representations of $SO(3, 1)$, enable us to deduce that, for $n = 4$,

$$\boldsymbol{\theta} = \boldsymbol{\theta}_1^+ + \boldsymbol{\theta}_1^- + \boldsymbol{\theta}_2^+ + \boldsymbol{\theta}_2^-,$$

where

$$\Lambda_1\theta_1^\pm = \theta_1^\pm(\Lambda_1 \pm 1), \quad \Lambda_2\theta_1^\pm = \theta_1^\pm\Lambda_2, \quad \Lambda_1\theta_2^\pm = \theta_2^\pm\Lambda_1,$$

$$\Lambda_2\theta_2^\pm = \theta_2^\pm(\Lambda_2 \pm 1).$$

It is not difficult (see Appendix B) to deduce the generalization of these results for $n = 3$ and 4. Thus when $n = 2m + 1$, θ can be resolved into components $\theta^0, \theta_i^+, \theta_i^-$, $i = 1, 2, \ldots, m$, where

$$\Lambda_i\theta^0 = \theta^0\Lambda_i, \quad \Lambda_i\theta_j^\pm = \theta_j^\pm(\Lambda_i \pm \delta_{ij}), \tag{5}$$

while, in the case $n = 2m$, the result is the same, except that θ^0 does not occur.

In what follows, it is convenient to think of α_{pq} as the element in the pth row and qth column of an antisymmetric $n \times n$ matrix of operators. We denote this matrix by α and by α^k and $\overline{\alpha^k}$ the matrices whose pqth elements are

$$(\alpha^k)_{pq} = (\alpha^{k-1})_{pr}\alpha_{rq}, \quad k = 2, 3, \ldots,$$

$$(\overline{\alpha^k})_{pq} = (\alpha^k)_{qp}.$$

Furthermore (cf. σ_2), we define σ_k by

$$\sigma_k = \mathrm{tr}\,[\alpha^k] = (\alpha^k)_{pp}.$$

We shall show that, as a consequence of the relations (1), α satisfies an nth-degree polynomial

identity, of the form

$$F_n(\alpha) = \alpha^n + a_1\alpha^{n-1} + \cdots + a_n = 0, \tag{6}$$

where the coefficients a_k, $k = 1, 2, \ldots, n$, are invariants of the $SO(n)$ Lie algebra.

The existence of this identity is in no way dependent on the existence or nonexistence, within the representation space for α, of an n-vector operator θ. However, when such a θ does exist, its resolution into n n-vector shift operators for the $SO(n)$ invariants, as in Eqs. (5) above, can most readily be achieved with the use of this identity, as we shall see.

We are mainly concerned with the determination of the form of the coefficients a_k in Eq. (6), as functions of $\sigma_2, \sigma_3, \ldots$, which are in turn functions, also to be determined, of $\Lambda_1, \Lambda_2, \ldots, \Lambda_m$. Before proceeding, however, we make the following general remarks concerning the identity expressed in Eq. (6).

It is clearly an analog of the Cayley–Hamilton identity for an $n \times n$ matrix of complex numbers. There are, however, some interesting differences.

In particular, suppose we define the determinant of an $n \times n$ matrix A of noncommuting elements, by

$$\det(A) = (1/n!)\epsilon_{ij\ldots m}\epsilon_{pq\ldots t}A_{ip}A_{jq}\cdots A_{mt},$$

where $\epsilon_{ij\ldots m}$ is the alternating tensor, with $\epsilon_{12\ldots n} = 1$. Then in the present context, we find

$$\det(\alpha - \rho I) = \rho^n + a_1'\rho^{n-1} + \cdots + a_n', \tag{7}$$

where ρ is an arbitrary complex number, I is the $n \times n$ unit matrix, and a_k', $k = 1, 2, \ldots, n$, like a_k, is an invariant of $SO(n)$. However, we find that, in general, $a_k' \neq a_k$, in contrast with the case when α is a matrix of complex numbers.

The existence of the polynomial (7) is comparatively well known, having been discussed in studies of Lie algebras by Killing[11] before the turn of the century and, more recently, by Racah[7] and Biedenharn.[8] Much less, it seems, is known of identities of the form in Eq. (6).

This equation actually expresses n^2 identities in the elements of α. Lehrer-Ilamed[12] has shown that n^2 identities of more general form are satisfied by the elements of any $n \times n$ matrix, provided these elements belong to a free associative algebra. Here we are dealing with a special case, where the algebra is in fact a Lie algebra, whose structure constants are such that these identities can be expressed in the simple form of a polynomial identity in α. [One of the authors (H. S. G.) has now determined similar identities for $Sp(n)$ and $SU(n)$.]

In the course of investigating certain identities satisfied by elements of any representation of the Lie algebra of $SU(3)$, Lehrer-Ilamed[13] has utilized similar generalized Cayley–Hamilton identities satisfied by some elements of the algebra. (See also Racah.[14]) However, to our knowledge, the identity (6) for $SO(n)$ has not previously been presented, even for the case $n = 3$.

2. VECTOR SHIFT OPERATORS AND THE POLYNOMIAL IDENTITY

Considering a system of operators $\boldsymbol{\theta}, \alpha$ as in Eqs. (1)–(5), we find that θ^0, θ_i^+, and θ_i^- are eigenvectors of the generator matrix α, in the sense that if θ_τ represents any one of these n operators,

$$\alpha\theta_\tau = d_\tau\theta_\tau,$$

i.e.,

$$\alpha_{pq}\theta_{\tau q} = d_\tau\theta_{\tau p},$$

where d_τ is an invariant of $SO(n)$. This follows from the fact that, for any vector operator θ_τ,

$$[\sigma_2, \theta_\tau] = 2(2\alpha - n + 1)\theta_\tau.$$

For, in view of Eqs. (3) and (5) above, σ_2 commutes with θ^0, so that

$$(\alpha - \tfrac{1}{2}n + \tfrac{1}{2})\theta^0 = 0. \tag{8}$$

Also,

$$[\sigma_2, \theta_i^+] = 2[\Lambda_i(\Lambda_i + n - 2i), \theta_i^+]$$

$$= 2(2\Lambda_i + n - 2i + 1)\theta_i^+,$$

so that

$$(\alpha - \tfrac{1}{2}n + 1)\theta_i^+ = (\Lambda_i + \tfrac{1}{2}n - i)\theta_i^+. \tag{9}$$

Similarly,

$$(\alpha - \tfrac{1}{2}n + 1)\theta_i^- = -(\Lambda_i + \tfrac{1}{2}n - i)\theta_i^-. \tag{10}$$

In Appendix A we show that α satisfies a polynomial identity of the general form of Eq. (6). The results (8)–(10) then allow us to write the identity in more precise form, effectively determining a_k, $k = 1, 2, \ldots, n$ as a function of Λ_i, $i = 1, 2, \ldots, m$. We must have

$$F_n(\alpha) = 0, \tag{11}$$

where

$$F_n(\alpha) = G_m(a^2), \qquad n = 2m, \tag{12}$$

$$F_n(\alpha) = (a - \tfrac{1}{2})G_m(a^2), \qquad n = 2m + 1, \tag{13}$$

and

$$a = \alpha - \tfrac{1}{2}n + 1,$$

$$G_m(a^2) = \prod_{i=1}^{m}[a^2 - (\Lambda_i + \tfrac{1}{2}n - i)^2]. \tag{14}$$

Conversely, once the results (11)–(14) are known, one can see why and how an arbitrary n-vector operator θ can be resolved in the manner indicated by Eqs. (5). Thus, in the case $n = 2m + 1$, we use Eq. (11) in the obvious way to define projection operators $P^0(\alpha), P_i^+(\alpha), P_i^-(\alpha)$, $i = 1, 2, \ldots, m$, which are polynomials of the $(n-1)$th degree in α and which satisfy

$$
\begin{aligned}
&(\alpha - \tfrac{1}{2}n + \tfrac{1}{2})P^0 = 0,\\
&[\alpha - \tfrac{1}{2}n + 1 \mp (\Lambda_i + \tfrac{1}{2}n - i)]P_i^{\pm} = 0,\\
&P^0P_i^{\pm} = P_i^{\pm}P^0 = P_i^+P_j^- = P_i^-P_j^+ = 0,\\
&(P^0)^2 = P^0, \quad P_i^{\pm}P_j^{\pm} = \delta_{ij}P_i^{\pm},\\
&P^0 + \sum_{i=1}^{m}(P_i^+ + P_i^-) = 1. \qquad (15)
\end{aligned}
$$

Then the required resolution is

$$\theta = \theta^0 + \sum_{i=1}^{m}(\theta_i^+ + \theta_i^-),$$

with

$$\theta^0 = P^0\theta, \quad \theta_i^{\pm} = P_i^{\pm}\theta. \qquad (16)$$

The case $n = 2m$ is similar, except that P^0 and θ^0 do not occur.

We do not go into details here, but mention that in order to confirm that Eqs. (5) follow from Eqs. (15) and (6), it is not sufficient to consider the commutators of $\theta^0, \theta_i^{\pm}$ only with σ_2. Rather one needs to calculate the commutators of these vectors with a complete set of invariants, which, like σ_2 but unlike Λ_i, are *explicitly* constructed from the set of α_{pq}. Finally, one must know the expression for each member of this complete set in terms of the Λ_i. We return to this last point in Sec. 5.

3. SYMMETRIC AND ANTISYMMETRIC POLYNOMIALS IN α

If the matrix polynomial $f(\alpha)$ is symmetric, i.e., if $\overline{f(\alpha)} = f(\alpha)$, then

$$g(\alpha) = (\alpha - \tfrac{1}{2}n)f(\alpha) + \tfrac{1}{2}\,\mathrm{tr}\,[f(\alpha)] \qquad (17)$$

is antisymmetric, i.e., $\overline{g(\alpha)} = -g(\alpha)$.

Furthermore, if $g(\alpha)$ is antisymmetric, then

$$h(\alpha) = (\alpha - \tfrac{1}{2}n + 1)g(\alpha) \qquad (18)$$

is symmetric.

The proof is as follows. Since $f_{pq}(\alpha)$ transforms as a tensor under $SO(n)$,

$$
\begin{aligned}
[\alpha_{pq}, f_{rs}(\alpha)] &= \delta_{qr}f_{ps}(\alpha) - \delta_{pr}f_{qs}(\alpha)\\
&\quad + \delta_{qs}f_{rp}(\alpha) - \delta_{ps}f_{rq}(\alpha). \qquad (19)
\end{aligned}
$$

By putting $q = r$ and using $f_{rs}(\alpha) = f_{sr}(\alpha)$, we have

$$\alpha f(\alpha) + \overline{f(\alpha)\alpha} = nf(\alpha) - \mathrm{tr}\,[f(\alpha)],$$

with the help of which the antisymmetry of $g(\alpha)$, defined as in Eq. (17), is readily established. On the other hand, if $f_{rs}(\alpha)$ is replaced in Eq. (19) by $g_{rs}(\alpha)$, where $g_{rs}(\alpha) = -g_{sr}(\alpha)$, we deduce that

$$\alpha g(\alpha) - \overline{g(\alpha)\alpha} = (n-2)g(\alpha),$$

which establishes the symmetry of $h(\alpha)$, defined as in Eq. (18).

Noting that $\alpha^0 = 1$ is symmetric and that α is antisymmetric, we see from these results that any polynomial of degree $2l$ in α, $l = 0, 1, 2, \ldots$, can be expressed as the sum of a symmetric one of degree $2l$ and an antisymmetric one of *lower* degree. Similarly, any polynomial of degree $(2l + 1)$ can be expressed as the sum of an antisymmetric one of that degree and a symmetric one of lower degree.

We then infer from Eqs. (11)–(14) that $F_n(\alpha)$ is symmetric or antisymmetric, according as n is even or odd, i.e.,

$$\overline{F_n(\alpha)} = (-1)^n F_n(\alpha). \qquad (20)$$

4. METHOD FOR THE DETERMINATION OF THE MATRIX POLYNOMIAL

For any given value of n, $F_n(\alpha)$ can be calculated by reduction to polynomial form of the appropriate equation (A3) or (A10), as shown in Appendix A for $n = 3, 4$, and 5. However, as n increases, such a calculation quickly becomes very involved. Here we present, in each of the cases n even and n odd, a method of obtaining $F_n(\alpha)$ quite easily for any given n.

(a) When $n = 2m$ is even, $F_n(\alpha)$ is completely determined by the conditions

(i) $F_n(\alpha) = G_m(a^2)$, where $G_m(a^2)$ is a polynomial in a^2 of degree m and $a = \alpha - m + 1$,

(ii) $\overline{F_n(\alpha)} = F_n(\alpha)$,

(iii) $\mathrm{tr}\,[F_n(\alpha)] = 0$,

which follow from Eqs. (12), (20), and (11). The proof is as follows.

Consider the sequence of polynomials defined by

$$
\begin{aligned}
&f_0 = 1, \quad f_1 = \alpha(\alpha - m + 1) + b_0 f_0,\\
&f_{l+1} = [\alpha(\alpha - 2m + 1) + b_l]f_l\\
&\quad + [\tfrac{1}{2}\alpha + c_l\alpha(\alpha - m + 1) + d_l]\,\mathrm{tr}\,[f_l],\\
&\quad l = 1, 2, \ldots, \qquad (21)
\end{aligned}
$$

where b_0, b_l, c_l, and d_l are arbitrary constants. According to the results of the preceding section, each polynomial in the sequence is symmetric.

Moreover, in view of the conditions (ii) and (iii) above, we see that, for some choice of b_0, b_l, c_l, and d_l, $l = 1, 2, \ldots, m-1$,

$$F_n = f_m - \text{tr}\,[f_m]/(2m). \tag{22}$$

Then f_m, like F_n, must be even in a, and we shall use this to determine b_0, b_l, c_l, and d_l uniquely [except for d_{m-1}, which remains arbitrary, but which does not in any way contribute to F_n, in view of Eq. (22)].

We denote by f_l^0 that part of f_l which is a linear combination of positive powers of α, with numerical coefficients, i.e., not involving $\sigma_2, \sigma_3, \ldots$. It is evident that

$$f_l^0 = \alpha(\alpha - m + 1)p_l[\alpha(\alpha - 2m + 1)]$$
$$= a(a + m - 1)p_l[(a + m - 1)(a - m)],$$

where $p_l(x)$ is a polynomial of the $(l-1)$th degree in x. It is also clear that if f_m is even in a, so are f_m^0 and $(f_m - f_m^0)$. But, if f_m^0 is even in a, we have

$$(a + m - 1)p_m[(a + m - 1)(a - m)]$$
$$= (a - m + 1)p_m[(a - m + 1)(a + m)].$$

We set, in succession, $a = m-1$, $a = m-2, \ldots$, $a = 1$ in this identity and thus obtain (for $n > 2$)

$$p_m[1(2 - 2m)] = p_m[2(3 - 2m)] = \ldots$$
$$= p_m[(m-1)(-m)] = 0,$$

and therefore

$$p_m(x) = (x + 2m - 2)(x + 4m - 6)\cdots$$
$$\times [x + m(m-1)],$$

$$f_m^0 = (a + m - 1)[(a - m + 1)(a + m - 2)]$$
$$\times [(a - m + 2)(a + m - 3)]\cdots[(a-1)(a)]a,$$
$$= \alpha[(\alpha - 2m + 2)(\alpha - 1)][(\alpha - 2m + 3)(\alpha - 2)]$$
$$\times \cdots [(\alpha - m)(\alpha - m + 1)](\alpha - m + 1).$$

Thus $b_l = l(2m - l + 1)$, $l = 0, 1, \ldots, m-1$, and in order that f_m should also be an even function of a, we must take $c_l = -1/(2l)$, $l = 1, 2, \ldots, m-1$ and $d_l = -\frac{1}{2}(m-1)$, $l = 1, 2, \ldots, m-2$, while d_{m-1} is left arbitrary.

Thus, when $n = 2m$,

$$F_n(\alpha) = f_m - \text{tr}\,[f_m]/(2m),$$

where

$$f_1 = \alpha(\alpha - m + 1),$$

$$f_{l+1} = (\alpha - l)(\alpha - 2m + l + 1)f_l - (\alpha - m + 1)$$
$$\times (\alpha - l)\,\text{tr}\,[f_l]/(2l). \tag{23}$$

For example, when $n = 2$,

$$F_2(\alpha) = f_1 - \tfrac{1}{2}\,\text{tr}\,[f_1],$$

$$f_1 = \alpha^2;$$

for $n = 4$ [cf. Eq. (A6)],

$$F_4(\alpha) = f_2 - \tfrac{1}{4}\,\text{tr}\,[f_2],$$

$$f_1 = \alpha(\alpha - 1),$$

$$f_2 = (\alpha - 2)(\alpha - 1)^2 - \tfrac{1}{2}(\alpha - 1)^2\sigma_2;$$

for $n = 6$,

$$F_6(\alpha) = f_3 - \tfrac{1}{6}\,\text{tr}\,[f_3],$$

$$f_1 = \alpha(\alpha - 2),$$

$$f_2 = (\alpha - 4)(\alpha - 2)(\alpha - 1)\alpha - \tfrac{1}{2}(\alpha - 2)(\alpha - 1)\sigma_2,$$

$$f_3 = (\alpha - 4)(\alpha - 3)(\alpha - 2)^2(\alpha - 1)\alpha - \tfrac{1}{2}(\alpha - 3)$$
$$\times (\alpha - 2)^2(\alpha - 1)\sigma_2 - \tfrac{1}{8}(\alpha - 2)^2[2\sigma_4 - 14\sigma_3$$
$$+ 16\sigma_2 - (\sigma_2)^2].$$

(b) When $n = 2m + 1$ is odd, $F_n(\alpha)$ is completely determined by the conditions

(i) $F_n(\alpha) = (a - \frac{1}{2})G_m(a^2)$, where $G_m(a^2)$ is a polynomial in a^2 of degree m and $a = \alpha - m + \frac{1}{2}$,

(ii) $\overline{F_n(\alpha)} = -F_n(\alpha)$,

which follow from Eqs. (13) and (20).

In this case we consider the sequence of antisymmetric polynomials generated by

$$g_1 = \alpha,$$

$$g_{l+1} = [\alpha(\alpha - 2m) + b_l]g_l + [c_l\alpha + \tfrac{1}{2}]\,\text{tr}\,[\alpha g_l],$$
$$l = 1, 2, \ldots.$$

For some choice of b_l and c_l, $l = 1, 2, \ldots, m$, we must have $F_n = g_{m+1}$. We determine these constants uniquely by requiring that g_{m+1} satisfy condition (i) above.

Suppose g_l^0 is obtained from g_l by dropping terms involving σ_k. Then

$$g_{l+1}^0 = \alpha p_l[\alpha(\alpha - 2m)],$$
$$= [a + m - \tfrac{1}{2}]p_l[(a - \tfrac{1}{2})^2 - m^2],$$

where $p_l(x)$ is of degree l in x. From condition (i), we see that g_{m+1}^0 must vanish for $a = \frac{1}{2}$, so that $p_m(-m^2) = 0$ for $m > 0$, and we can write

$$p_m(x) = (x + m^2)q_m(x),$$

$$g_{m+1}^0 = (a + m - \tfrac{1}{2})(a - \tfrac{1}{2})^2 q_m[(a - \tfrac{1}{2})^2 - m^2].$$

Again from condition (i), we see that $(a+m-\frac{1}{2})(a-\frac{1}{2})q_m[(a-\frac{1}{2})^2-m^2]$ must be an even polynomial in a, so that

$$(a+m-\tfrac{1}{2})(a-\tfrac{1}{2})q_m[(a-\tfrac{1}{2})^2-m^2]$$
$$=(a-m+\tfrac{1}{2})(a+\tfrac{1}{2})q_m[(a+\tfrac{1}{2})^2-m^2].$$

We set, in succession, $a=m-\frac{1}{2}$, $a=m-\frac{3}{2}$, $\cdots$, $a=\frac{3}{2}$, and obtain, for $m>1$,

$$q_m[1(1-2m)]=q_m[2(2-2m)]=\cdots$$
$$=q_m[(m-1)(-1-m)]=0,$$

whence

$$q_m(x)=[x+1(2m-1)][x+2(2m-2)]\cdots$$
$$\times[x+(m-1)(m+1)],$$

$$g^0_{m+1}=(a-\tfrac{1}{2})(a+m-\tfrac{1}{2})[(a-m+\tfrac{1}{2})(a+m-\tfrac{3}{2})]$$
$$\times[(a-m+\tfrac{3}{2})(a+m-\tfrac{5}{2})]\cdots$$
$$\times[(a-\tfrac{3}{2})(a+\tfrac{1}{2})](a-\tfrac{1}{2}),$$

$$=\alpha[(\alpha-2m)+1(2m-1)][\alpha(\alpha-2m)+2(2m-2)]$$
$$\times\cdots[\alpha(\alpha-2m)+m^2].$$

For $m=1$, $q_m=1$. Thus $b_l=l(2m-l)$, and, to make $g_{m+1}/(a-\frac{1}{2})$ an even polynomial in a, we must also take $c_l=-1/(2l)$.

Thus, when $n=2m+1$,

$$F_n(\alpha)=g_{m+1}(\alpha),$$

where

$$g_1=\alpha,$$
$$g_{l+1}=(\alpha-l)(\alpha-2m+l)g_l-(\alpha-l)\,\mathrm{tr}[\alpha g_l]/(2l). \tag{24}$$

For example, when $n=3$ [cf. Eq. (A12)],

$$g_2=(\alpha-1)^2\alpha-\tfrac{1}{2}(\alpha-1)\sigma_2;$$

for $n=5$ [cf. Eq. (A13)],

$$g_2=(\alpha-1)(\alpha-3)\alpha-\tfrac{1}{2}(\alpha-1)\sigma_2,$$
$$g_3=(\alpha-2)^2(\alpha-3)(\alpha-1)\alpha-\tfrac{1}{2}(\alpha-1)(\alpha-2)^2\sigma_2$$
$$-\tfrac{1}{4}(\alpha-2)[\sigma_4-4\sigma_3+3\sigma_2-\tfrac{1}{2}(\sigma_2)^2].$$

5. INVARIANTS OF $SO(n)$

Consider the sequence of antisymmetric polynomials $e_l(\alpha)$, $l=1,2,\ldots$, defined by

$$e_1=\alpha,\qquad e_{l+1}=\alpha(\alpha-n+1)e_l+\tfrac{1}{2}\mathrm{tr}[\alpha e_l].$$

The identity $\mathrm{tr}[e_l]=0$ then expresses σ_{2l-1} as a function of $\sigma_2,\sigma_3,\ldots,\sigma_{2l-2}$. Thus

$$\sigma_3=\tfrac{1}{2}(n-2)\sigma_2,$$
$$\sigma_5=\tfrac{1}{2}(3n-4)\sigma_4-\tfrac{1}{2}(n-1)(n-2)\sigma_3-\tfrac{1}{2}(\sigma_2)^2,$$
etc.

We see also, from the fact that $\mathrm{tr}[\alpha\,{}^kF_n(\alpha)]=0$, $k=1,2,\ldots$, that $\sigma_{n+1},\sigma_{n+2},\ldots$ can be expressed as functions of $\sigma_2,\sigma_3,\ldots,\sigma_n$. It follows that all σ_k can be regarded as functions of σ_{2l}, $l=1,2,\ldots,m$, or, alternatively, and more conveniently from our point of view, as functions of τ_l, where

$$\tau_l=\mathrm{tr}[f_l]/(2l),\qquad n=2m,$$
$$\tau_l=\mathrm{tr}[\alpha g_l]/(2l),\qquad n=2m+1,$$

with f_l,g_l as in Eqs. (23) and (24).

The functional dependence of τ_l on Λ_i, $i,l=1,2,\ldots,m$, and hence the eigenvalues of τ_l, can be determined by comparing the two forms $G_m(a^2)$ given, on the one hand in Eq. (14), on the other via Eqs. (23) or (24).

Thus in the case $n=2m+1$, we deduce from Eq. (24) that

$$G_m(a^2)=\{-\tau_m-[a^2-\tfrac{1}{4}]\tau_{m-1}$$
$$-[a^2-\tfrac{1}{4}][a^2-\tfrac{9}{4}]\tau_{m-2}-\cdots$$
$$-[a^2-\tfrac{1}{4}][a^2-\tfrac{9}{4}]\cdots[a^2-(m-\tfrac{3}{2})^2]\tau_1$$
$$+[a^2-\tfrac{1}{4}][a^2-\tfrac{9}{4}]\cdots[a^2-(m-\tfrac{1}{2})^2]\},\tag{25}$$

so that

$$\tau_m=-G_m(\tfrac{1}{4}),\qquad \tau_{m-1}=\tfrac{1}{2}[-\tau_m-G_m(\tfrac{9}{4})],$$
$$\tau_{m-2}=[-\tau_m-6\tau_{m-1}-G_m(\tfrac{25}{4})]/24,\qquad\text{etc.}$$

Then from Eq. (14) we have, writing $\chi_i=(\Lambda_i+\frac{1}{2}n-i)^2$,

$$\tau_m=(-1)^{m+1}\prod_{i=1}^{m}(\chi_i-\tfrac{1}{4}),$$
$$\tau_{m-1}=(-1)^{m+1}\left(\prod_{i=1}^{m}(\chi_i-\tfrac{9}{4})-\prod_{i=1}^{m}(\chi_i-\tfrac{1}{4})\right),\qquad\text{etc.}$$

We note also, from a comparison of Eqs. (25) and (A10), that

$$\beta_p^{(m)}\beta_p^{(m)}=-m\tau_m=m(-1)^m\prod_{i=1}^{m}(\chi_i-\tfrac{1}{4}).$$

In the case $n=2m$, we find from Eq. (23) that

$$G_m(a^2)=\{-\tau_m-a^2\tau_{m-1}-a^2[a^2-1]\tau_{m-2}-\cdots$$
$$-a^2[a^2-1]\cdots[a^2-(m-2)^2]\tau_1$$
$$+a^2[a^2-1]\cdots[a^2-(m-1)^2]\},\tag{26}$$

$$\tau_m = -G_m(0), \quad \tau_{m-1} = -\tau_m - G_m(1),$$

$$\tau_{m-2} = \tfrac{1}{12}[-\tau_m - 4\tau_{m-1} - G_m(4)], \quad \text{etc.}$$

Again writing $\chi_i = (\Lambda_i + \frac{1}{2}n - i)^2$, we see from Eq. (14) that

$$\tau_m = (-1)^{m+1} \prod_{i=1}^{m} \chi_i,$$

$$\tau_{m-1} = (-1)^{m+1}\left[\prod_{i=1}^{m}(\chi_i - 1) - \prod_{i=1}^{m}\chi_i\right], \quad \text{etc.}$$

The set of operators $\tau_1, \tau_2, \ldots, \tau_{m-1}$ and τ_m (or $\beta_p^{(m)}\beta_p^{(m)}$) is a complete set of invariants for $SO(2m+1)$. However, in the case $n = 2m$, as is well known, one pseudoscalar invariant such as $\beta^{(m)}$ (see Appendix A) or Λ_m is needed. We note from a comparison of Eqs. (26) and (A3) that $[\beta^{(m)}]^2/m^2 = -\tau_m$, so that

$$\beta^{(m)} = m(i)^m \prod_{i=1}^{m}(\Lambda_i + \tfrac{1}{2}n - i),$$

where we have taken the sign of the square root which is consistent with that weight vector interpretation of $(\lambda_1, \lambda_2, \ldots, \lambda_m)$ adopted in Appendix B. The set of operators τ_l, $l = 1, 2, \ldots, m-1$, and $\beta^{(m)}$ is a complete set of invariants for $SO(2m)$.

APPENDIX A:

Here we establish the existence of an nth degree polynomial identity for α. Following Bakri,[15] we define the completely antisymmetric tensors $\beta_{pq}^{(k)}{}_{\cdots t}$, $k = 0, 1, \ldots, m$, of rank $n - 2k$, with

$$\beta_{pq\ldots v}^{(0)} = \epsilon_{pq\ldots v},$$

$$\beta_{pq\ldots t}^{(k)} = \tfrac{1}{2}\alpha_{uv}\beta_{pq\ldots tuv}^{(k-1)}. \tag{A1}$$

Then one finds that, for $k = 0, 1, \ldots, m-1$,[15]

$$\beta_{pq\ldots tu}^{(k)}\alpha_{uv} - k\beta_{pq\ldots tv}^{(k)} = (-1)^{n+1}(\delta_{pv}\beta_{qr\ldots t}^{(k+1)} - \delta_{qv}\beta_{pr\ldots t}^{(k+1)} + \cdots + (-1)^n\delta_{tv}\beta_{pq\ldots s}^{(k+1)})/(k+1). \tag{A2}$$

In the case $k = 0$, this identity is proved by inspection, and the proof for general values of k is obtained by induction.

(a) When $n = 2m$, we take $k = m - 1$ in Eq. (A2) to obtain

$$\beta_{pq}^{(m-1)}\alpha_{qr} - (m-1)\beta_{pr}^{(m-1)} = -\delta_{pr}\beta^{(m)}/m,$$

i.e., $\beta^{(m-1)}(\alpha - m + 1) = -\beta^{(m)}/m$. Moreover, $\alpha\beta^{(m-1)} = \beta^{(m-1)}\alpha$, so that

$$(\alpha - m + 1)^2[\beta^{(m-1)}]^2 = [\beta^{(m)}]^2/m^2. \tag{A3}$$

The quantity $[\beta^{(m-1)}]^2$ can be reduced to a polynomial of degree $(2m-2)$ in α, so that Eq. (A3) is in fact the required identity. It is not a simple matter to complete this reduction for a general value of m. However, for small values of m, the desired result can be obtained from the identity

$$\epsilon_{pij\ldots klr}\epsilon_{qst\ldots uvr} = \sum(-1)^{S(P)}\delta_{pq}\delta_{is}\delta_{jt}\cdots\delta_{ku}\delta_{lv}, \tag{A4}$$

where $\Sigma(-1)^{S(P)}$ means a sum over all permutations of $p, i, \ldots, k$, and l, with appropriate signatures, by multiplying with

$$(\alpha_{ij}\cdots\alpha_{kl})(\alpha_{st}\cdots\alpha_{uv}) \tag{A5}$$

and shuffling factors till the required order is reached. For example, when $n = 4$,

$$[\beta^{(1)}]^2 = -\overline{\alpha^2} + \tfrac{1}{2}\sigma_2 = -\alpha^2 + 2\alpha + \tfrac{1}{2}\sigma_2$$

so that

$$(\alpha - 2)(\alpha - 1)^2\alpha - \tfrac{1}{2}(\alpha - 1)^2\sigma_2 + \tfrac{1}{4}[\beta^{(2)}]^2 = 0. \tag{A6}$$

{In the important related case of $SO(3,1)$, with generators $J_{\lambda\mu}, \lambda, \mu = 0, 1, 2, 3$, satisfying

$$[J_{\lambda\mu}, J_{\rho\sigma}] = -i(g_{\lambda\rho}J_{\mu\sigma} + g_{\mu\sigma}J_{\lambda\rho} - g_{\mu\rho}J_{\lambda\sigma} - g_{\lambda\sigma}J_{\mu\rho})$$

where $g_{\lambda\mu}$ is the pseudo-Euclidean metric tensor, the corresponding result is

$$J_\mu{}^\alpha J_\alpha{}^\rho J_\rho{}^\sigma J_\sigma{}^\nu - 4iJ_\mu{}^\alpha J_\alpha{}^\rho J_\rho{}^\nu + (J_1 - 5)J_\mu{}^\alpha J_\alpha{}^\nu - 2i(J_1 - 1)J_\mu{}^\nu = (J_1 + [J_2]^2)\delta_\mu{}^\nu,$$

where $J_1 = \frac{1}{2}J_{\lambda\mu}J^{\lambda\mu}$, and $J_2 = \frac{1}{8}\epsilon_{\lambda\mu\nu\rho}J^{\lambda\mu}J^{\nu\rho}$.}

(b) When $n = 2m + 1$, we take $k = m - 1$ in Eq. (A2) to obtain

$$\beta_{pqr}^{(m-1)}\alpha_{rs} - (m-1)\beta_{pqs}^{(m-1)} = (1/m)(\delta_{sp}\beta_q^{(m)} - \delta_{sq}\beta_p^{(m)}). \tag{A7}$$

Premultiplying this equation, on the one hand by α_{pq}, on the other by $\beta_p^{(m)}$, we obtain

$$\alpha_{pq}\beta_q^{(m)} = \beta_q^{(m)}\alpha_{qp} = m\beta_p^{(m)}, \tag{A8}$$

and

$$\gamma_{qr}\alpha_{rs} + (m-1)\gamma_{sq} = (1/m)(\beta_s^{(m)}\beta_q^{(m)} - \beta_p^{(m)}\beta_p^{(m)}\delta_{qs}), \tag{A9}$$

where $\gamma_{pq} = -\gamma_{qp} = \beta_r^{(m)}\beta_{rpq}^{(m-1)}$. Next, premultiplying Eq. (A9) by $(\alpha_{us} - m\delta_{us})$ and using Eq. (A8), we obtain

$$(\alpha - m)[\overline{\gamma\alpha} + (m-1)\gamma + (1/m)\beta_p^{(m)}\beta_p^{(m)}] = 0$$

or, since γ is antisymmetric,

$$(\alpha - m)[(\alpha - m)\gamma + (1/m)\beta_p^{(m)}\beta_p^{(m)}] = 0. \quad \text{(A10)}$$

This is the required identity, as γ can be reduced to a polynomial of degree $(n-2)$ in α.

In this case, using Eq. (A4), we find

$$\gamma_{pq} = 2^{(2-n)}(\alpha_{ij}\cdots\alpha_{kl})(\alpha_{st}\cdots\alpha_{uv}) \times \sum(-1)^{S(P)}\delta_{pq}\delta_{is}\delta_{jt}\cdots\delta_{kn}\delta_{lv}. \quad \text{(A11)}$$

For example, when $n = 3$, we find $\gamma = \alpha$, so that

$$(\alpha - 1)[(\alpha - 1)\alpha + \beta_p^{(1)}\beta_p^{(1)}] = 0; \quad \text{(A12)}$$

for $n = 5$, after a lengthy calculation, we find

$$\gamma = 2\alpha^3 - 8\alpha^2 + (6 - \sigma_2)\alpha + \sigma_2,$$

so that

$$(\alpha - 2)\{(\alpha - 2)[2\alpha^3 - 8\alpha^2 + (6 - \sigma_2)\alpha + \sigma_2] + \tfrac{1}{2}\beta_p^{(2)}\beta_p^{(2)}\} = 0. \quad \text{(A13)}$$

APPENDIX B:

Here we justify the assertion that an arbitrary n-vector operator θ can be resolved into components satisfying Eqs. (5).

An irreducible representation of $SO(n)$ is characterized by a set of integers or half-odd integers $(\lambda_1, \lambda_2 \ \cdots\ , \lambda_m)$, as in Eq. (4). This set may be interpreted as the weight vector of highest weight for the corresponding representation,[16] when each weight vector is defined as an ordered set of eigenvalues of, in particular, $-i\alpha_{12}, -i\alpha_{34}, \cdots, -i\alpha_{2m-1,2m}$. Accordingly, for such a representation, one can find an element ψ of the representation space, such that

$$-i\alpha_{2i-1,2i}\psi = \lambda_i\psi, \quad i = 1, 2, \cdots, m. \quad \text{(B1)}$$

Moreover, since ψ corresponds to the highest weight and since, for $q > 2i > 2j$,

$$(-i\alpha_{2i-1,2i})(\alpha_{2i-1,q} + i\alpha_{2i,q}) = (\alpha_{2i-1,q} + i\alpha_{2i,q})(-i\alpha_{2i-1,2i} + 1)$$

and

$$[(-i\alpha_{2j-1,2j}), (\alpha_{2i-1,q} + i\alpha_{2i,q})] = 0,$$

it follows that

$$(\alpha_{2i-1,q} + i\alpha_{2i,q})\psi = 0, \quad q > 2i. \quad \text{(B2)}$$

One sees conversely that any ψ satisfying Eqs. (B1) and (B2) belongs to the representation labeled $(\lambda_1, \lambda_2, \cdots, \lambda_m)$. For example, Eq. (B2) implies that (no summation over repeated subscripts here except where indicated)

$$(\alpha_{2i-1,q} - i\alpha_{2i,q})(\alpha_{2i-1,q} + i\alpha_{2i,q})\psi = 0, \quad q > 2i,$$

i.e.,

$$(\alpha_{2i-1,q}\alpha_{q,2i-1} + \alpha_{2i,q}\alpha_{q,2i})\psi = -i\alpha_{2i-1,2i}\psi = \lambda_i\psi.$$

Thus

$$\sum_{q=2i+1}^{n}(\alpha_{2i-1,q}\alpha_{q,2i-1} + \alpha_{2i,q}\alpha_{q,2i}) = (n - 2i)\lambda_i\psi,$$

$$\left[\sum_{q=2i}^{n}(\alpha_{2i-1,q}\alpha_{q,2i-1}) + \sum_{q=2i+1}^{n}(\alpha_{2i,q}\alpha_{q,2i})\right]\psi = \lambda_i(\lambda_i + n - 2i)\psi,$$

and

$$\sigma_2\psi = \sum_{i=1}^{m} 2\lambda_i(\lambda_i + n - 2i)\psi.$$

Now suppose there exists an n-vector operator θ acting on ψ. In the case $n = 2m + 1$, define $\psi_i^{\pm} = (\theta_{2i-1} \pm i\theta_{2i})\psi$, $i = 1, 2, \cdots, m$, and $\psi^0 = \theta_n\psi$. It follows from Eqs. (2), (B1), and (B2) that $\psi_i^{\pm}$ and ψ^0 correspond to weight vectors $(\lambda_1, \lambda_2, \cdots, \lambda_{i-1}, \lambda_i \pm 1, \lambda_{i+1}, \cdots, \lambda_m)$ and $(\lambda_1, \lambda_2, \cdots, \lambda_m)$ and that the set $\psi^0, \psi_i^{\pm}$, $i = 1, 2, \cdots, m$ is invariant under the action of the operators $(\alpha_{2j-1,q} + i\alpha_{2j,q})$, $j = 1, 2, \cdots, m$, $q > 2j$.

We deduce therefore that $\theta\psi$ has components only in respresentations labeled $(\lambda_1, \lambda_2, \cdots, \lambda_{i-1}, \lambda_i \pm 1, \lambda_{i+1}, \cdots, \lambda_m)$ $i = 1, 2, \cdots, m$, or $(\lambda_1, \lambda_2, \cdots, \lambda_m)$. It is easy to see then that the same must be true if ψ is replaced by any element of the representation space labeled $(\lambda_1, \lambda_2, \cdots, \lambda_m)$, so that the required result follows for $n = 2m + 1$. The argument for the case $n = 2m$ is similar except that ψ^0 does not occur, and $\theta\psi$ can have no component in the representation labeled $(\lambda_1, \lambda_2, \cdots, \lambda_m)$.

---

* This work was done in part at the University of Adelaide and in part while the authors were at Michigan State University and the International Centre for Theoretical Physics at Trieste.

1 P. A. M. Dirac, *The Principles of Quantum Mechanics* (Clarendon, Oxford, 1930).

2 P. Güttinger and W. Pauli, Z. Phys. **67**, 743 (1931).

3 E. P. Wigner, *Gruppentheorie und ihre Anwendung auf die Quantenmechanik der Atomspektren* (Braunschweig, 1931; English Ed., Academic, New York, 1959).

4 G. Racah, Phys. Rev. **62**, 438 (1942).

5 E. V. Condon and G. H. Shortley, *The Theory of Atomic Spectra* (Cambridge U.P., Cambridge, 1935).

[6] J. C. Slater, *The Quantum Theory of Atomic Structure* (McGraw-Hill, New York, 1960), Vol. II.
[7] G. Racah, in *Ergebnisse der exacten Naturwissenschaften* (Springer, Berlin, 1965), Vol. 37.
[8] L. C. Biedenharn, in *Lectures on Theoretical Physics, Boulder 1962* (Interscience, New York, 1963), Vol. 5. See also L. C. Biedenharn and J. D. Louck, Commun. Math. Phys. **8**, 89 (1968).
[9] J. D. Louck, Los Alamos Scientific Laboratory Report LA2451 (1960).
[10] H. J. Bhabha, Rev. Mod. Phys. **17**, 200 (1945); **21**, 451 (1949).
[11] W. Killing, Math. Ann. **31**, 252 (1888); **33**, 1 (1889); **34**, 57 (1889); **36**, 161 (1890).
[12] Y. Lehrer-Ilamed, Bull. Res. Council Israel **5A**, 197 (1956).
[13] Y. Lehrer-Ilamed, in *Spectroscopic and Group Theoretical Methods in Physics*, edited by F. Bloch *et al.* (North-Holland, Amsterdam, 1968). See also H. Goldberg and Y. Lehrer-Ilamed, J. Math. Phys. **4**, 501 (1963); B. Gruber and L. O' Raifeartaigh, *ibid.* **5**, 1796 (1964).
[14] G. Racah, in *Group Theoretical Concepts and Methods in Elementary Particle Physics* (Lectures of the 1962 Istanbul Summer School), edited by F. Gürsey, (Gordon and Breach, New York, 1964).
[15] M. M. Bakri, Commun. Math. Phys. **9**, 146 (1968).
[16] F. D. Murnaghan, *The Theory of Group Representations* (Dover, New York, 1963).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 10 OCTOBER 1971

Characteristic Identities for Generators of $GL(n)$, $O(n)$ and $Sp(n)$

H. S. Green
Department of Mathematical Physics, University of Adelaide, Adelaide, S. Australia 5001
(Received 5 April 1971)

A hierarchy of tensor identities, satisfied by the generators of the general linear group $GL(n)$, is obtained in terms of two different sets of invariants. An application to the identification of irreducible representations and the decomposition of reducible representations is described. Similar results are obtained for the generators of orthogonal, pseudo-orthogonal, and symplectic groups.

1. INTRODUCTION

The generators of the group $GL(n)$ satisfy the commutation relations[1]

$$[a^i{}_j, a^k{}_l] = \delta^k{}_j a^i{}_l - \delta^i{}_l a^k{}_j, \quad i, \ldots, l = 1, \ldots, n, \tag{1}$$

and their matrix representations are of some interest, as they also furnish representations for other Lie algebras. Indeed, the commutation relations

$$[z_i, z_j] = C^k{}_{ij} z_k,$$

in which the structure constants necessarily satisfy

$$C^k{}_{ij} + C^k{}_{ji} = 0, C^l{}_{ij} C^m{}_{kl} + C^l{}_{jk} C^m{}_{il} + C^l{}_{ki} C^m{}_{jl} = 0,$$

can be satisfied identically by writing

$$z_i = C^k{}_{ji} a^j{}_k,$$

provided that a^i_j satisfy (1). Different irreducible representations of the a^i_j are, moreover, associated with different sets of eigenvalues of the invariants σ_r defined by (repeated affixes $i, j, k, l, \ldots$ are understood to be summed over values from 1 to n; however, subscripts $r, s, \ldots$ are exempted from this summation convention)

$$\sigma_1 = a^i{}_i, \quad \sigma_2 = a^i{}_j a^j{}_i, \quad \sigma_3 = a^i{}_j a^j{}_k a^k{}_i, \tag{2}$$

etc., which are Casimir operators, i.e., commute with all elements of the algebra. Thus, an irreducible representation of $GL(n)$ can, in principle, be identified by determining the eigenvalues of $\sigma_1, \sigma_2, \ldots, \sigma_n$. Of course, such a representation is not necessarily irreducible for the z_i, and the σ_r, $r \leq n$, are not necessarily independent. The problem of determining a complete set of independent invariants for a semisimple Lie algebra has been considered by Biedenharn[2] and by Gruber and O'Raiffeartaigh.[3]

Another, less explicit but sometimes more convenient way of defining a set of invariants for $GL(n)$ is in terms of the highest weights of the finite-dimensional irreducible representations. Let λ_1 be an operator whose eigenvalue, in a particular representation R of this kind, is the same as the maximum eigenvalue ℓ_1 of $a^1{}_1$ in this representation. Further, let $\lambda_r, r = 2, \ldots, n)$ be an operator whose eigenvalue ℓ_r in R is the same as the maximum eigenvalues of $a^r{}_r$, when $a^1{}_1, \ldots, a^{r-1}{}_{r-1}$ have the eigenvalues $\ell_1, \ldots, \ell_{r-1}$, respectively. Then, if ψ is a vector such that $a^r{}_r \psi = \ell_r \psi$, it must satisfy $a^i{}_j \psi = 0, j > i$, and (as one can verify by computing $\sigma_1 \psi$ and $\sigma_2 \psi$)

$$\sigma_1 = \sum_{r=1}^{n} \lambda_r, \quad \sigma_2 = \sum_{r=1}^{n} \lambda_r (\lambda_r + n + 1 - 2r). \tag{3}$$

The representation R is labeled by $\ell = (\ell_1, \ell_2, \ldots, \ell_n)$, where $\ell_r - \ell_s$ is integral and nonnegative when $r < s$, and can be identified in this way if the dependence of the first n of the σ_r on the λ_s is known. Unfortunately, the complexity of the expressions for the σ_r in terms of the λ_s increases rapidly with r.

One use of the invariants is in the decomposition of a reducible representation into distinct irreducible components, which can be solved by determining the projections on to different eigenvectors of the σ_r (or λ_s). There are, of course, other ways of dealing with this problem, notably the method of character analysis, which has been applied to $U(n)$ and $SU(n)$ by Blaha.[4] Our present interest in the problem arises from its connection with a

[6] J. C. Slater, *The Quantum Theory of Atomic Structure* (McGraw-Hill, New York, 1960), Vol. II.
[7] G. Racah, in *Ergebnisse der exacten Naturwissenschaften* (Springer, Berlin, 1965), Vol. 37.
[8] L. C. Biedenharn, in *Lectures on Theoretical Physics, Boulder 1962* (Interscience, New York, 1963), Vol. 5. See also L. C. Biedenharn and J. D. Louck, Commun. Math. Phys. **8**, 89 (1968).
[9] J. D. Louck, Los Alamos Scientific Laboratory Report LA2451 (1960).
[10] H. J. Bhabha, Rev. Mod. Phys. **17**, 200 (1945); **21**, 451 (1949).
[11] W. Killing, Math. Ann. **31**, 252 (1888); **33**, 1 (1889); **34**, 57 (1889); **36**, 161 (1890).
[12] Y. Lehrer-Ilamed, Bull. Res. Council Israel **5A**, 197 (1956).
[13] Y. Lehrer-Ilamed, in *Spectroscopic and Group Theoretical Methods in Physics*, edited by F. Bloch *et al.* (North-Holland, Amsterdam, 1968). See also H. Goldberg and Y. Lehrer-Ilamed, J. Math. Phys. **4**, 501 (1963); B. Gruber and L. O' Raifeartaigh, *ibid.* **5**, 1796 (1964).
[14] G. Racah, in *Group Theoretical Concepts and Methods in Elementary Particle Physics* (Lectures of the 1962 Istanbul Summer School), edited by F. Gürsey, (Gordon and Breach, New York, 1964).
[15] M. M. Bakri, Commun. Math. Phys. **9**, 146 (1968).
[16] F. D. Murnaghan, *The Theory of Group Representations* (Dover, New York, 1963).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 10 OCTOBER 1971

Characteristic Identities for Generators of $GL(n)$, $O(n)$ and $Sp(n)$

H. S. Green
Department of Mathematical Physics, University of Adelaide, Adelaide, S. Australia 5001
(Received 5 April 1971)

A hierarchy of tensor identities, satisfied by the generators of the general linear group $GL(n)$, is obtained in terms of two different sets of invariants. An application to the identification of irreducible representations and the decomposition of reducible representations is described. Similar results are obtained for the generators of orthogonal, pseudo-orthogonal, and symplectic groups.

1. INTRODUCTION

The generators of the group $GL(n)$ satisfy the commutation relations[1]

$$[a^i{}_j, a^k{}_l] = \delta^k{}_j a^i{}_l - \delta^i{}_l a^k{}_j, \quad i, \ldots, l = 1, \ldots, n, \tag{1}$$

and their matrix representations are of some interest, as they also furnish representations for other Lie algebras. Indeed, the commutation relations

$$[z_i, z_j] = C^k{}_{ij} z_k,$$

in which the structure constants necessarily satisfy

$$C^k{}_{ij} + C^k{}_{ji} = 0, C^l{}_{ij} C^m{}_{kl} + C^l{}_{jk} C^m{}_{il} + C^l{}_{ki} C^m{}_{jl} = 0,$$

can be satisfied identically by writing

$$z_i = C^k{}_{ji} a^j{}_k,$$

provided that a^i_j satisfy (1). Different irreducible representations of the a^i_j are, moreover, associated with different sets of eigenvalues of the invariants σ_r defined by (repeated affixes $i, j, k, l, \ldots$ are understood to be summed over values from 1 to n; however, subscripts $r, s, \ldots$ are exempted from this summation convention)

$$\sigma_1 = a^i{}_i, \quad \sigma_2 = a^i{}_j a^j{}_i, \quad \sigma_3 = a^i{}_j a^j{}_k a^k{}_i, \tag{2}$$

etc., which are Casimir operators, i.e., commute with all elements of the algebra. Thus, an irreducible representation of $GL(n)$ can, in principle, be identified by determining the eigenvalues of $\sigma_1, \sigma_2, \ldots, \sigma_n$. Of course, such a representation is not necessarily irreducible for the z_i, and the σ_r, $r \leq n$, are not necessarily independent. The problem of determining a complete set of independent invariants for a semisimple Lie algebra has been considered by Biedenharn[2] and by Gruber and O'Raiffeartaigh.[3]

Another, less explicit but sometimes more convenient way of defining a set of invariants for $GL(n)$ is in terms of the highest weights of the finite-dimensional irreducible representations. Let λ_1 be an operator whose eigenvalue, in a particular representation R of this kind, is the same as the maximum eigenvalue ℓ_1 of $a^1{}_1$ in this representation. Further, let $\lambda_r, r = 2, \ldots, n)$ be an operator whose eigenvalue ℓ_r in R is the same as the maximum eigenvalues of $a^r{}_r$, when $a^1{}_1, \ldots, a^{r-1}{}_{r-1}$ have the eigenvalues $\ell_1, \ldots, \ell_{r-1}$, respectively. Then, if ψ is a vector such that $a^r{}_r \psi = \ell_r \psi$, it must satisfy $a^i{}_j \psi = 0, j > i$, and (as one can verify by computing $\sigma_1 \psi$ and $\sigma_2 \psi$)

$$\sigma_1 = \sum_{r=1}^{n} \lambda_r, \quad \sigma_2 = \sum_{r=1}^{n} \lambda_r (\lambda_r + n + 1 - 2r). \tag{3}$$

The representation R is labeled by $\ell = (\ell_1, \ell_2, \ldots, \ell_n)$, where $\ell_r - \ell_s$ is integral and nonnegative when $r < s$, and can be identified in this way if the dependence of the first n of the σ_r on the λ_s is known. Unfortunately, the complexity of the expressions for the σ_r in terms of the λ_s increases rapidly with r.

One use of the invariants is in the decomposition of a reducible representation into distinct irreducible components, which can be solved by determining the projections on to different eigenvectors of the σ_r (or λ_s). There are, of course, other ways of dealing with this problem, notably the method of character analysis, which has been applied to $U(n)$ and $SU(n)$ by Blaha.[4] Our present interest in the problem arises from its connection with a

hierarchy of polynomial identities existing between the generators of $GL(n)$ and of certain of its subgroups. Earlier work by Lehrer-Ilamed[5] showed that there are always n^2 identities satisfied by the elements of an n-dimensional matrix which are also elements of an associative algebra. Here we shall show how to determine such identities explicitly for the Lie algebras derived from $GL(n)$. The method is a generalization of that developed previously for $O(n)$ by Bracken and Green.[6] We shall also show how to obtain corresponding identities for tensor operators of arbitrary rank and symmetry type, involving the generators $a^i{}_j$ linearly.

There is, indeed, a hierarchy of characteristic identities for $GL(n)$, $O(n)$, and $Sp(n)$. The simplest and most fundamental type is the characteristic equation satisfied by the generators of $GL(n)$, regarded as a matrix. If we (cf Racah[7]) define series of tensor operators a^r, $r = 0, 1, 2, \ldots$, or powers of a, by

$$(a^0)^i{}_j = \delta^i{}_j, \quad (a^{r+1})^i{}_j = a^i{}_k (a^r)^k{}_j, \tag{4}$$

it is easy to verify that, for $n = 1$ and $n = 2$,

$$\begin{aligned} &a - \sigma_1 = 0 \quad (\text{meaning } a^i{}_j - \sigma_1 \delta^i{}_j = 0), \\ &a^2 - (\sigma_1 + 1)a + \tfrac{1}{2}(\sigma_1^2 + \sigma_1 - \sigma_2) = 0, \end{aligned} \tag{5}$$

respectively. These identities are in some sense analogous to the Cayley–Hamilton identity satisfied by numerical matrices, but, because of the noncommutative nature of the $a^i{}_j$, cannot be obtained by the usual elementary method; the polynomials in a on the left sides are, indeed, different from the Killing determinantal polynomials appearing in the work of Racah[8] and Biedenharn and Louck.[9]

The a^r defined in (4) can be regarded as linear operators on an n-dimensional space of vectors or vector operators $\psi : (a^r\psi)^i = (a^r)^i{}_j \psi^j$. The next in the hierarchy of identities involve linear operators on a space of symmetric or antisymmetric tensors of the second rank. In general, we define

$$\begin{aligned} (a^0)^{pq\ldots s}{}_{jk\ldots m} &= \delta^{pq\ldots s}{}_{jk\ldots m}, \quad (a^{r+1})^{pq\ldots s}{}_{jk\ldots m} \\ &= a^p{}_i (a^r)^{iq\ldots s}{}_{jk\ldots m} + a^q{}_i (a^r)^{pi\ldots s}{}_{jk\ldots m} + \cdots \\ &\quad + a^s{}_i (a^r)^{pq\ldots i}{}_{jk\ldots m}, \end{aligned} \tag{6}$$

where $\delta^{pq\ldots s}{}_{jk\ldots m} \psi^{jk\ldots m} = \psi^{pq\ldots s}$, when $\psi^{pq\ldots s}$ is a tensor or tensor operator of given rank and symmetry. Then, for instance, we have $\delta^{ij}{}_{kl} = \frac{1}{2}(\delta^i{}_k \delta^j{}_l + \delta^i{}_l \delta^j{}_k)$ for symmetric tensors of the second rank, and the characteristic identity for $n = 2$ is

$$(a - \sigma_1 - 2)\,[a^2 - 2(\sigma_1 + 1)a + 2(\sigma_1{}^2 + \sigma_1 - \sigma_2)] = 0. \tag{7}$$

Our principal aim in the following section is to show how to obtain similar identities for arbitrary n and tensors of any rank and symmetry. In Sec. 3 we show informally how such identities are concerned in the decomposition of reducible representations, in particular the direct product of two irreducible representations. Finally, in Sec. 4 we shall outline the extension of these results to generators of the orthogonal, pseudo-orthogonal, and symplectic groups.

2. CONSTRUCTION AND PROPERTIES OF THE IDENTITIES

We shall develop here a method for the systematic determination of the identities for arbitrary n. The actual procedure is summarized in (34), (36), and (39) of this section; we first enumerate the results on which the procedure is based and, where necessary, outline a simple proof.

[1] *The characteristic equation satisfied by the matrix operator a, defined as in (4), may be written*

$$\prod_{r=1}^{n} (a - \lambda_r - n + r) = 0, \tag{8}$$

where the λ_r are the operators appearing in (3).

To prove this, suppose ψ is a vector operator satisfying

$$[a^i{}_j, \psi^k] = \delta^k{}_j \psi^i, \tag{9}$$

e.g., ψ could be the column with elements $a^i{}_{n+1}$ of the matrix $a^i{}_j$ in which $1 \leqslant i, j \leqslant n + 1$. Since

$$a^s{}_s \psi^r = \psi^r (a^s{}_s + \delta_{rs}),$$

there must be a component, ψ_r say, of ψ which increases the eigenvalue of λ_r by 1, leaving the eigenvalues of the other λ_s unchanged:

$$\lambda_s \psi_r = \psi_r (\lambda_s + \delta_{rs}). \tag{10}$$

Now, if $\bar{a}$ is the negative transpose of the matrix a, i.e.,

$$\bar{a}_p{}^q = -a^q{}_p, \tag{11}$$

then we have, for any vector ψ,

$$a\psi = \psi(n - \bar{a}) \tag{12}$$

(meaning $a^i{}_j \psi^j = n\psi^i - \psi^j \bar{a}^i{}_j$), and

$$[\sigma_2, \psi] = a\psi - \psi\bar{a} = (2a - n)\psi. \tag{13}$$

But, according to (3) and (10),

$$\begin{aligned} [\sigma_2, \psi_r] &= [\lambda_r(\lambda_r + n + 1 - 2r), \psi_r] \\ &= (2\lambda_r + n - 2r)\psi_r. \end{aligned}$$

Hence, by substitution of ψ_r for ψ in (11), we have

$$(a - \lambda_r - n + r)\psi_r = 0, \tag{14}$$

and (6) follows. We note, in retrospect, that

$$\psi_r = P_r(a,\lambda)\psi,$$

where $P_r(a,\lambda)$ is the projection operator given by

$$P_r(a,\lambda) = \prod_{s=r} \frac{(a-\lambda_s-n+s)}{(\lambda_r-r-\lambda_s+s)}. \tag{15}$$

[2] *The characteristic equation satisfied by the transpose $\bar{a}$ defined in (11) is*

$$\prod_{r=1}^{n} (\bar{a}-\bar{\lambda}_r-n+r) = 0, \tag{16}$$

where

$$\bar{\lambda}_r = -\lambda_{n+1-r} \tag{17}$$

and $\bar{a}^r$ is defined recursively by

$$(\bar{a}^{r+1})_j{}^i = (\bar{a}^r)_j{}^k a_k{}^i. \tag{18}$$

This result follows directly from (14), which, with the help of (12), can be written

$$\psi_r(\bar{a}+\lambda_r-r+1) = 0. \tag{19}$$

[3] *The characteristic equation satisfied by the operator a, defined in* (6), *depends on the contravariant (or covariant) rank ρ of $a^{pq..s}{}_{jk..m}$ and the symmetry type of this tensor, which corresponds to an irreducible representation $(m_1,m_2,\ldots,m_n)$, with $\sum_r m_r = \rho$, of $GL(n)$. It is*

$$\prod_m (a - \tfrac{1}{2}t_2 - \sum_r m'_r(\lambda_r + \tfrac{1}{2}n + \tfrac{1}{2} - \tfrac{1}{2}m'_r - r)) = 0,$$
$$t_2 = \sum_r m_r(m_r + n + 1 - 2r), \tag{20}$$

where the product $\prod_m$ is over all sets of the non-negative integers m'_r, such that

$$\begin{aligned} &m_n \le m'_r \le m_1, \quad r = 1,2,\ldots,n,\\ &m_{n-1} + m_n \le m'_r + m'_s \le m_1 + m_2, \quad r \ne s,\\ &m_{n-2} + m_{n-1} + m_n \le m'_r + m'_s + m'_t \le m_1\\ &+ m_2 + m_3, \quad r \ne s \ne t,\\ &\qquad\vdots\\ &\sum_r m'_r = \sum_r m_r = 0. \end{aligned} \tag{21}$$

The labeling $(m_1,m_2,\ldots,m_n)$ of the irreducible representation is the usual one and implies that if more than m_1 superscripts of a tensor $\psi^{ij..m}$ of the corresponding symmetry type have the same value r_1, *or* if more than m_c superscripts have the same value r_c when there are already $m_1,\ldots,m_{c-1}$ superscripts with the values $r_1,\ldots,r_{c-1}$, *then* $\psi^{ij..m}$ vanishes identically. This is the origin of the inequalities in (21), as will be seen in the following proof of (20), which is similar to that of (8).

It should be noticed that

$$(a\psi)^{jk..m} = a^j{}_i\psi^{ik..m} + a^k{}_i\psi^{ji..m} + \cdots + a^m{}_i\psi^{jk..i} \tag{22}$$

and that $\bar{a}$ can be defined so that

$$(\psi\bar{a})^{jk..m} = \psi^{ik..m}\bar{a}_i{}^j + \psi^{ji..m}\bar{a}_i{}^k + \cdots + \psi^{jk..i}\bar{a}_i{}^m, \tag{23}$$

with $\bar{a}_p{}^q$ given by (11). Then we have

$$[\sigma_2,\psi] = a\psi - \psi\bar{a}, \psi\bar{a} = (-a + b^i{}_j b^j{}_i)\psi, \tag{24}$$

where $b^i{}_j$ is the multiple substitution operator defined by

$$b^i{}_j\phi^{kl..n} = \delta^k{}_j\phi^{il..n} + \delta^l{}_j\phi^{ki..n} + \ldots + \delta^n{}_j\phi^{kl..n}, \tag{25}$$

for a tensor ϕ of the same rank as ψ. As the $b^i{}_j$ constitute a realization of the generators of $GL(n)$, and $\psi^{jk..m}$ an irreducible representation, it follows from (3) that

$$b^i{}_j b^j{}_i\psi = t_2\psi, \tag{26}$$

with t_2 given by (20). Hence,

$$[\sigma_2,\psi] = (2a - t_2)\psi. \tag{27}$$

But, if $\psi_{st..v}$ is a tensor (with tensor superscripts $jk..m$ omitted) such that

$$\lambda_r\psi_{st..v} = \psi_{st..v}(\lambda_r + \delta_{rs} + \delta_{rt} + \cdots + \delta_{rv}), \tag{28}$$

then it follows from (3) that

$$[\sigma_2,\psi_{st..v}] = \sum_r m'_r(2\lambda_r + n + 1 - m'_r - 2r), \tag{29}$$

where m'_r is the number of subscripts of $\psi_{st..v}$ which are equal to r. Combining (27) and (29), we have

$$[a - \tfrac{1}{2}t_2 - \sum_r m'_r(\lambda_r + \tfrac{1}{2}n + \tfrac{1}{2} - \tfrac{1}{2}m'_r - r)] \times\psi_{st..v} = 0 \tag{30}$$

and our result follows.

[4] *Similarly,*

$$\Pi_m[\bar{a} - \tfrac{1}{2}t_2 - \sum_r m'_r(\bar{\lambda}_r + \tfrac{1}{2}n + \tfrac{1}{2} - \tfrac{1}{2}m'_r - r)] = 0, \tag{31}$$

where $\bar{\lambda}_r$ is again given by (17), and the m'_r again satisfy the inequalities (21).

The remaining results of this section are directed towards the expression of these identities in terms of the σ_r, instead of the λ_r which cannot, in general, be expressed directly in terms of the generators of the group. They will concern only properties of the simple tensor $a^i{}_j$ and identities of the type (8),

though generalizations corresponding to (7) or (20) are certainly possible.

[5] *If* $\sigma_r = (a^r)^i_i$ *and* $\bar{\sigma}_r = (\bar{a}^r)_i{}^i$, and

$$F(a,\sigma) = \prod_{r=1}^{n} (a - \lambda_r - n + r) \tag{32}$$

is the characteristic polynomial in a [which vanishes when the a^i_j *are generators of GL(n)], then*

$$F(a,\sigma) = (-1)^n F(n-1-a,\bar{\sigma}). \tag{33}$$

For, if $\sigma_r = \sigma_r(\lambda)$, then $\bar{\sigma}_r = \sigma_r(\bar{\lambda})$. This identity then follows from (8) and (17).

[6] *We call a polynomial* $g(a,\sigma)$ *in a symmetric if* $g(a,\sigma) = g(\bar{a},\bar{\sigma})$, *and antisymmetric if* $g(a,\sigma) = -g(\bar{a},\bar{\sigma})$. *Then, if* $f_0 = 0, f_1 = a$, *and*

$$f_{r+1} = (a - \tfrac{1}{2}n)f_r + f^{(r)} + \sum_{s=0}^{[\frac{1}{2}r]-1} c^{(s)}_{r-2s-1} f_{r-2s-1},$$

$$f^{(0)} = a,\ f^{(r)} = (\tfrac{1}{2} - a/r)(f_r)^i_i, \tag{34}$$

then f_r *is symmetric or antisymmetric according as* r *is even or odd.*

The proof, by induction, is very simple. If f_r is symmetric (antisymmetric), it follows from

$$[a^i_j, (f_r)^j_k] = n(f_r)^i_k - \delta^i_k (f_r)^j_j$$

that $(a - \frac{1}{2}n)f_r + \frac{1}{2}(f_r)^i_i$ is antisymmetric (symmetric).

[7] *The polynomial* $F(a,\sigma)$ *of (32) is completely determined by the conditions*

$$F(a,\sigma) = (-1)^n F(n-1-a,\bar{\sigma}), \tag{35a}$$

$F(a,\sigma)$ is symmetric when n is even, antisymmetric when n is odd, (35b)

$$[F(a,\sigma)]^i_i = 0. \tag{35c}$$

Equivalently, we could assert that $F(a,\sigma)$ is generated recursively by (34) and

$$F(a,\sigma) = f_n - (f_n)^i_i/n \tag{36}$$

and that the coefficients $c^{(s)}_r$ are uniquely determined by the condition (35a) only. This is, in fact, what we shall prove. Of course, all three conditions are trivial consequences of the vanishing of $F(a,\sigma)$, but the point is that by using only these conditions we can determine $F(a,\sigma)$.

The proof is as follows. From the method of construction, it can be seen that

$$f_r = \sum_{s=0}^{r-1} p^{(s)}_r f^{(s)}, \tag{37}$$

where $p^{(s)}_r = p^{(s)}_r(a - \frac{1}{2}n, \sigma)$ is a polynomial in $a - \frac{1}{2}n$ which is even or odd according as $r - s$ is odd or even, respectively. If the condition (35a) is to be satisfied,

$$(a - \tfrac{1}{2}s)p^{(s)}_n(a - \tfrac{1}{2}n, \sigma) = (a - n + 1 + \tfrac{1}{2}s) \times p^{(s)}_n(a - \tfrac{1}{2}n + 1, \bar{\sigma}).$$

By setting $a - \frac{1}{2}n = r$, for $r = \frac{1}{2}(n-s) - 1, r = \frac{1}{2}(n-s) - 2, \ldots, r = 1$ when $n - s$ is even, or $r = \frac{1}{2}(n-s-1), r = \frac{1}{2}(n-s-3), \ldots r = 1$ when $n - s$ is odd, we see that the polynomial $p^{(s)}_n$ must have zeros for these numerical values of a. Consequently,

$$p^{(s)}_n = (a - \tfrac{1}{2}n) \prod_{r=1}^{\frac{1}{2}(n-s)-1} [(a - \tfrac{1}{2}n)^2 - r^2], \quad n - s \text{ even},$$

$$p^{(s)}_n = \prod_{r=1}^{\frac{1}{2}(n-s-1)} [(a - \tfrac{1}{2}n)^2 - (r - \tfrac{1}{2})^2], \quad n - s \text{ odd}. \tag{38}$$

Thus, f_n has been completely determined.

[8] *The characteristic polynomial is given by (36), where* f_n *is determined from (34), with coefficients* $c^{(s)}_r$ *given by*

$$c^{(s)}_r = \frac{(-1)^{s+1}(2s)!}{2^{4s+3}\, s!(s+1)!} \cdot \frac{(n-r)!}{(n-r-2s-2)!}. \tag{39}$$

The coefficients are determined as follows. By inspection of the coefficients of the polynomials $p^{(r-1)}_n$ we see that

$$c^{(0)}_r + c^{(0)}_{r+1} = -(n-r-1)^2/4,$$

$$c^{(1)}_r + c^{(1)}_{r+1} = c^{(0)}_{r+1} c^{(0)}_{r+2}, \ldots, c^{(s)}_r + c^{(s)}_{r+1} = c^{(0)}_{r+1} c^{(s-1)}_{r+2} + c^{(1)}_{r+1} c^{(s-2)}_{r+4} + \ldots + c^{(s)}_{r+1} c^{(1)}_{r+2s-2}.$$

These difference equations can be solved with the help of

$$c^{(0)}_{n-1} = 0, \quad c^{(s)}_{n-2s-1} = 0,$$

giving

$$c^{(0)}_r = -(n-r)(n-r-1)/2^3,$$

$$c^{(1)}_r = (n-r)(n-r-1)(n-r-2)(n-r-3)/2^7, \tag{40}$$

and, by induction, (39).

[9] *If*

$$\theta_r = (f_r)^j_j/r, \quad L_r = r - \lambda_r - \tfrac{1}{2}(n+1),$$

$$F_s = \tfrac{1}{2}\prod_{r=1}^{n}(L_r + \tfrac{1}{2}s) + \tfrac{1}{2}(-1)^{s-1}\prod_{r=1}^{n}(L_r - \tfrac{1}{2}s), \tag{41}$$

then the θ_r *and the* L_s *are connected by the relations*

$$F_s = -(s-1)!\theta_{n-s+1} - (s-2)!\theta_{n-s+3}/1!$$
$$- (s-3)!\theta_{n-s+5}/2! - \cdots, s!$$
$$\theta_{n-s+1}$$
$$= -sF_s + s\cdot(s-2)F_{s-2}/1!$$
$$- s(s-1)\cdot(s-4)F_{s-4}/2! + \cdots. \quad (42)$$

The first of these relations follows from the identity

$$F(a,\sigma) = -[\theta_n + \theta_{n-1}a' + \theta_{n-2}(a'^2 - \tfrac{1}{4})$$
$$+ \theta_{n-3}a'(a'^2 - 1) + \cdots] = \prod_{r=1}^{n}(L_r + a'), \quad (43)$$

where $a' = a - \frac{1}{2}(n-1)$, which is obtained by equating $F(a,\sigma)$, as given by (36), (37) and (38), with the left side of (8). We simply form F_s by substituting $a' = \frac{1}{2}s$ and $a' = -\frac{1}{2}s$ in (43). The second relation is obtained by solution of the first.

We note also the following associated results, obtained by direct comparison of coefficients of $a'^{n-1}, a'^{n-2}, \ldots$ in (43):

$$S_1 = \sum_r L_r = -\theta_1,$$
$$S_2 = \sum_{r>s} L_r L_s = -\theta_2 - (n+1)n(n-1)/24,$$
$$S_3 = \sum_{r>s>t} L_r L_s L_t = -\theta_3 + n(n-1)(n-2)\theta_1/24,$$
$$S_4 = \sum_{r>s>t>u} L_r L_s L_t L_u = -\theta_4 + (n-1)(n-2)$$
$$\times (n-3)\theta_2/24 + (n+1)n(n-1)(n-2)$$
$$\times (n-3)(5n+7)/560, \quad (44)$$

etc. The eigenvalues of $L_1, L_2, L_3, \ldots$ are in ascending order and are therefore easily identified. The above relations allow symmetric functions of the L_r to be expressed in terms of the σ_r, by substitution of the explicit forms of f_r into $\theta_r = (f_r)^j{}_j/r$. Thus, we have, from (34) and (39),

$$\theta_1 = \sigma_1, \quad 2\theta_2 = \sigma_2 - \sigma_1^2,$$
$$3\theta_3 = \sigma_3 - \tfrac{1}{2}n\sigma_2 - \tfrac{1}{2}(\sigma_1 - \tfrac{1}{2}n)(\sigma_2 - \sigma_1^2)$$
$$- (\sigma_2 - \tfrac{1}{2}\sigma_1 - c_1^{(0)})\sigma_1, \quad \text{etc.} \quad (45)$$

The continuation of this table is most conveniently effected with the help of a LISP computer program.

3. IDENTIFICATION AND DECOMPOSITION OF REPRESENTATIONS

In this section we shall discuss some applications of the foregoing results, which will, incidentally, throw further light on the role of the characteristic identities.

One way of identifying an irreducible representation is to determine the characteristic identity satisfied by the generators. It is, however, necessary to note that although the generators a^i_j in general satisfy an equation of the nth degree, they may also satisfy a reduced equation of lower degree in a particular irreducible representation. For, if the eigenvalues ℓ_s and ℓ_t of λ_s and λ_t are equal in such a representation and $t > s$, then the component of the vector operator ψ satisfying $(a - \lambda_s - n + s)\psi = 0$ must vanish in this representation, and the factor $(a - \ell_s - n + s)$ may therefore be omitted from the characteristic equation

$$\prod_r (a - \ell_r - n + r) = 0. \quad (46)$$

By omitting all such factors, we obtain the reduced equation of the representation. If this reduced equation is of the qth degree in a, the eigenvalues s_r of the σ_r satisfy a qth-order difference equation [obtained by multiplying (46) by a^m and contracting the resulting tensor identity]. The solution of this difference equation is of the form

$$s_r = \sum_t c_t(\ell_t + n - t)^{r-1}, \quad (47)$$

where the c_r are numerical constants. We tabulate for reference the characteristic identities and the eigenvalues of the invariants in some of the simpler irreducible representations:

$(1,0,\ldots)$: $(a-n)a = 0$, $s_r = n^{r-1}$;
$(1,1,0,\ldots)$: $(a-n+1)a = 0$, $s_r = 2(n-1)^{r-1}$;
$(2,0,\ldots)$: $(a-n-1)a = 0$, $s_r = 2(n+1)^{r-1}$;
$(1,1,1,0,\ldots)$: $(a-n+2)a + 0$,
$s_r = 3(n-2)^{r-1}$;
$(2,1,0,\ldots)$: $(a-n-1)(a-n+1)a = 0$,
$s_r = 3[(n+1)^{r-1} + (n-1)^{r-1}]/2$;
$(3,0,\ldots)$: $(a-n-2)a = 0$, $s_r = 3(n+2)^{r-1}$. (48)

We consider now the problem of decomposing a reducible representation into distinct irreducible components. The latter are, of course, distinguished by the eigenvalues of invariants, and can be isolated by applying projection operators corresponding to the different sets of eigenvalues. We intend to apply this procedure to the decomposition of $GL(n) \times GL(n)$. We denote the generators in two irreducible representations by a^i_j and b^i_j, so that those in the product space are

$$c^i_j = a^i_j + b^i_j \quad (49)$$

(or, more precisely, $a^i_j \times 1 + 1 \times b^i_j$). We adopt tensor representations $\psi = \psi^{rs\ldots}_{u\ldots}$ and $\phi = \phi^{kl\ldots}_{n\ldots}$ for vectors in the range of a^i_j and b^i_j, respectively, so that

$$\chi^{kl\ldots}{}_{n\ldots} = \phi^{kl\ldots}{}_{n\ldots}\psi$$

is a vector of the product space which we wish to decompose. Distinct irreducible representations correspond to different eigenvalues of

$$\gamma = \tfrac{1}{2}(a^i{}_j a^j{}_i + b^i{}_j b^j{}_i - c^i{}_j c^j{}_i) = -\,a^i{}_j b^j{}_i\,. \qquad (50)$$

Suppose $c^{(m)}$ are the eigenvalues of γ in the representation, so that

$$\prod_m (\gamma - c^{(m)})\chi^{kl\ldots n} = 0. \qquad (51)$$

According to (25) and (23),

$$\gamma\phi^{kl\ldots n} = -\,(a^k{}_j\phi^{jl\ldots n} + a^l{}_j\phi^{kj\ldots n} + a^n{}_j\phi^{kl\ldots j})$$
$$= (\phi\bar{a})^{kl\ldots n}, \quad \gamma^2\phi^{kl\ldots n} = (\phi\bar{a}^2)^{kl\ldots n},$$

etc. with $\bar{a}$ defined in the generalized way of (6). It follows, then, from (51), that

$$\left(\phi\prod_m(\bar{a} - c^{(m)})\right)^{kl\ldots n}\psi = 0. \qquad (52)$$

This is evidently the reduced counterpart, in the special representation we are considering, of the identity (31), and it follows that

$$c^{(m)} = \tfrac{1}{2}t_2 + \sum_r m_r'(\bar{\ell}_r + \tfrac{1}{2}n + \tfrac{1}{2} - \tfrac{1}{2}m_r' - r), \qquad (53)$$

where $\bar{\ell}_r = -\,\bar{\ell}_{n+1-r}$ is the appropriate eigenvalue of $\bar{\lambda}_r$ and the m_r' must satisfy the inequalities (21). The eigenvalue (53) corresponds to the irreducible representation $(\ell_1 + m_1', \ell_2 + m_2', \ldots)$ contained in the product space we are considering. This shows that the factors of the characteristic equation (8) correspond to different representations generated by the direct product of the adjoint representation $(1, 0, \ldots)$ and that associated with the $a^i{}_j$. The factors of the generalized equation (20) correspond to different representations generated by the direct product of the irreducible representation $(m_1, m_2, \ldots)$ and that associated with the $a^i{}_j$.

Of course, not every set of values of the m_r' satisfying (21) is necessarily represented in the reduced characteristic identity. The vector

$$\left(\phi\prod_{p\neq m}(\bar{a} - c^{(p)})/(c^{(m)} - c^{(p)})\right)^{kl\ldots n}\psi,$$

which is the component of χ corresponding to a proper irreducible representation $(\ell_1 + m_1', \ell_2 + m_2', \ldots)$, cannot vanish, and this requirement in general imposes further conditions on the m_r', analogous to those which restricted the degrees of the characteristic equations shown in (48). These additional conditions can be found by considering in the first instance the possibility that $\phi^{kl\ldots n}$ is completely antisymmetric. Then the conditions (21), and

$$\ell_r + m_r' \geq \ell_{r+1} + m_{r+1}' \qquad (54)$$

are sufficient. The corresponding conditions for tensors of other symmetry types are obtained by expressing them in terms of direct products of completely antisymmetric tensors, e.g., $(2, 0, \ldots) = (1, 0, \ldots) \times (1, 0, \ldots) - (1, 1, 0, \ldots)$. It should be remarked that a given irreducible representation may appear multiply within the product space, so that although (54) is a vector of only one distinct irreducible representation, it may still be decomposable. The decomposition can be effected by projection onto eigenvectors of other independent invariants $a^i{}_j b^j{}_k b^k{}_i$, $a^i{}_j b^j{}_k b^k{}_l b^l{}_i$, etc., of the reducible representations, but we shall not pursue this question here.

4. IDENTITIES FOR SUBGROUPS OF $GL(n)$

It is well known that the introduction of a metric tensor is necessary for the definition of the orthogonal and symplectic subgroups of $GL(n)$. If g_{ij} is symmetric and we define

$$\alpha_{ij} = g_{ik}a^k{}_j - g_{jk}a^k{}_i$$

[where the $a^j{}_k$, as before, are generators of $GL(n)$], then the α_{ij} are generators of the orthogonal transformations which leave the quadratic form $g_{ij}\phi^i\psi^j$, involving two vectors ϕ and ψ, invariant. Similarly, if g_{ij} is antisymmetric and

$$\alpha_{ij} = g_{ik}a^k{}_j + g_{jk}a^k{}_i,$$

then the α_{ij} are generators of the symplectic transformations which leave $g_{ij}\phi^i\psi^j$ invariant, We can accomodate both possibilities by writing

$$\alpha_{ij} = g_{ik}a^k{}_j - g_{kj}a^k{}_i, \quad g_{ij} = \eta g_{ji}, \qquad (55)$$

where $\eta = 1$ for $0(n)$ or -1 for $Sp(n)$. These generators satisfy

$$[\alpha_{ij}, \alpha_{kl}] = g_{kj}\alpha_{il} - g_{il}\alpha_{kj} - g_{ik}\alpha_{jl} + g_{lj}\alpha_{ki}. \qquad (56)$$

If ψ^k is a vector operator,

$$[\alpha_{ij}, \psi^k] = \delta^k{}_j\psi_i - \eta\delta^k{}_i\psi_j, \quad \psi_i = g_{ij}\psi^j. \qquad (57)$$

It is necessary to assume that g_{ij} is nonsingular, which is only possible for $Sp(n)$ when n is even. Then a contravariant g^{ij} exists such that

$$g^{ij}g_{jk} = g_{kj}g^{ji} = \delta^i{}_k, \qquad (58)$$

and we can proceed to define tensor operators α and $\bar{\alpha}$ analogous to a and $\bar{a}$ given by (22) and (23), but constructed from

$$\alpha^i{}_j = g^{ik}\alpha_{kj}, \quad \alpha_j{}^i = \bar{\alpha}_{jk}g^{ki} = -\,\alpha^i{}_j. \qquad (59)$$

The invariants are defined by

$$\sigma_r = (\alpha^r)^i{}_i, \qquad (60)$$

but σ_1 now vanishes identically and, as we shall see, the σ_{2r+1} can be expressed in terms of the σ_{2r}. There are, therefore, h independent invariants, where

$$h = \tfrac{1}{2}n, \quad n \text{ even}, \quad \text{or} \quad h = \tfrac{1}{2}(n-1), \quad n \text{ odd}. \qquad (61)$$

In addition, there is the inversion P, when $\det(-g) = -1$.

We can also introduce the well-known invariants $\lambda_r, r = 1, 2, \ldots h$, implicitly in the following way. First a transformation reducing g_{ij} to the canonical form $g_{ij} = \pm\delta_{ij}$ for $O(n)$; $g_{ij} = \delta_{i+1\,j}$ (i odd) or $-\delta_{ij+1}$ (i even) for $Sp(n)$ is applied. Then, in a particular irreducible representation, suppose that ℓ_1 is the maximum eigenvalue of $\alpha^1{}_2$, and that ℓ_r is the maximum eigenvalue of $\alpha^{2r-1}{}_{2r}$, when the $\alpha^{2s-1}{}_{2s}$ with $s<r$ already have the eigenvalues ℓ_s. When ψ is the eigenvector of this representation corresponding to the eigenvalues ℓ_s, it is readily verified that

$$\sigma_2\psi = 2\sum_{r=1}^{h}\ell_r(\ell_r + n + 1 - \eta - 2r)\psi.$$

We define λ_r as the operator whose eigenvalue is ℓ_r within the irreducible representation described, so that

$$\sigma_2 = 2\sum_{r=1}^{h}\lambda_r(\lambda_r + n + 1 - \eta - 2r) \tag{62}$$

for any representation. If we define λ_r for $h < r \leq n$ by

$$\lambda_r = \eta - \lambda_{n+1-r}, \quad r = n-h+1, \ldots, n,$$
$$\lambda_{h+1} = 1, \quad n \text{ odd}, \tag{63}$$

then we can also write

$$\sigma_2 = \eta h(n-h) + \sum_{r=1}^{n}\lambda_r(\lambda_r - \eta + n + 1 - 2r) \tag{64}$$

From this starting point, it is a straightforward matter to derive the characteristic identities, by the methods of Sec. 2. We shall simply state the principal results. There are, in general, $2h$ vector operators ψ_s (excluding ψ_{h+1}, when n is odd) such that

$$\lambda_r\psi_s = \psi_s(\lambda_r + \delta_{rs}). \tag{65}$$

When n is odd, the additional vector operator ψ_{h+1} commutes with all the λ_r. The vector operators thus specified satisfy

$$(\alpha - \lambda_r - n + \eta + r)\psi_r = 0,$$
$$\psi_r(\bar{\alpha} + \lambda_r + 1 - r) = 0, \quad r = 1, \ldots, n, \tag{66}$$

where $\alpha\psi$ and $\psi\bar{\alpha}$ denote vector operators within components $(\alpha\psi)^j = \alpha^j{}_k\psi^k$, respectively. The characteristic identity is therefore

$$\prod_{r=1}^{n}(\alpha - \lambda_r - n + \eta + r) = 0, \tag{67}$$

and $\bar{\alpha}$ satisfies the *same* equation. An equivalent result was obtained previously, for $O(n)$, by Bracken and Green[6]; the application to $Sp(n)$ is new. Also, the analogy with the corresponding result (8) for $GL(n)$ is interesting.

The result (67) is just the first of a hierarchy. There is a generalization in which α is the tensor operator defined by

$$(\alpha\psi)^{jk\ldots m} = \alpha^j{}_i\psi^{ik\ldots m} + \alpha^k{}_i\psi^{ji\ldots m} + \cdots \alpha^m{}_i\psi^{jk\ldots m} \tag{68}$$

for any irreducible tensor representation of $O(n)$ or $Sp(n)$. If we disregard associated representations (derived from one of lower rank by the use of the alternating tensor) the irreducible tensor representations can be labeled by $(m_1, m_2, \ldots, m_h)$, where $\sum_{r=1}^{h}m_r = \rho$, the rank of the tensor. With such labeling of $\psi^{jk\ldots m}$, the characteristic identity satisfied by α is

$$\prod_{m}\left(\alpha - \tfrac{1}{4}t_2 - \sum_{r=1}^{}m_r'(\lambda_r + \tfrac{1}{2}n - \tfrac{1}{2}\eta + \tfrac{1}{2} - \tfrac{1}{2}m_r' - r)\right) = 0,$$
$$t_2 = 2\sum_{r=1}^{h}m_r(m_r + n + 1 - \eta - 2r),$$
$$0 \leq m_r' \leq m_1,$$
$$0 \leq m_r' + m_s' \leq m_1 + m_2,$$
$$\vdots$$
$$\sum_{r=1}^{n}m_r' \leq \sum_{r=1}^{h}m_r = \rho. \tag{69}$$

Finally, we give the procedure for constructing the characteristic identity (67) in terms of the σ_r, this is somewhat simpler than that given for $GL(n)$, but again depends on the construction of a sequence of self-conjugate or anti-self-conjugate polynomials in α, the conjugate being defined by

$$\bar{f}_j{}^i = g_{jk}f^k{}_lg^{li}. \tag{70}$$

The sequence defined by

$$f_2 = \alpha(\alpha - \tfrac{1}{2}n + \eta),$$
$$f_{2r+2} = (\alpha - r)(\alpha - n + \eta + r)f_{2r} - (\alpha - r)(\alpha - \tfrac{1}{2}n + \eta)(f_{2r})^i{}_i/(2r) \tag{71}$$

is self-conjugate and is unique in leading to a polynomial f_{2h} which is even in $\alpha - h$, when $n = 2h$ is even. The characteristic polynomial is therefore

$$f_{2h} - (f_{2h})^i{}_i/(2h) = 0, \quad n \text{ even}. \tag{72}$$

On the other hand, the sequence defined by

$$f_1 = \alpha, \quad f_{2r+1} = (\alpha - r)(\alpha - n + r)f_{2r-1} - (\alpha - r)(\alpha f_{2r-1})^i{}_i/(2r) \tag{73}$$

is anti-self-conjugate and is unique in leading to a polynomial f_{2h+1} which is of the form $(\alpha - h)$ times a polynomial even in $\alpha - \frac{1}{2}n$, when $n = 2h + 1$ is odd. The characteristic identity for $O(n)$ (given by Bracken and Green[6]) is therefore

$$f_{2h+1} = 0. \tag{74}$$

We noted previously the existence of identical relations between the σ_{2r+1} and the σ_{2r} for $O(n)$ and $Sp(n)$. For instance, $\sigma_3 = (\frac{1}{2}n - \eta)\sigma_2$. These relations are easily obtained from (71), if we note that $(\alpha - \frac{1}{2}n)f_{2r} + \frac{1}{2}(f_{2r})^i{}_i$ is anti-self-conjugate, so that

$$(\alpha f_{2r})^i{}_i = 0. \tag{75}$$

By substituting explicit forms for f_{2r}, we obtain the desired identities.

ACKNOWLEDGMENT

The author wishes to acknowledge several helpful discussions with Professor C. A. Hurst.

[1] See, e.g., M. Hammermesh, *Group Theory* (Addison-Wesley, Reading, Mass., 1962), p. 389.
[2] L. C. Biedenharn, J. Math. Phys. 4, 436 (1963).
[3] B. Gruber and L. O'Raiffeartaigh, J. Math. Phys. 5, 1796 (1964).
[4] S. Blaha, J. Math. Phys. 10, 2156 (1969).
[5] Y. Lehrer-Ilamed, Bull. Res. Counc. Israel **5A**, 197 (1956).
[6] A. J. Bracken and H. S. Green, J. Math. Phys. 12, 2099 (1971).
[7] G. Racah, in *Ergebnisse der exakten Naturwissenschaften* (Springer, Berlin, 1965), Vol. 37, pp. 66–71.
[8] G. Racah, Rend. Lincei **8**, 108 (1950).
[9] L. C. Biedenharn and J. D. Louck, Commun. Math. Phys. 8, 89 (1968); See also Refs. 2 and 3.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 10 OCTOBER 1971

Exact Nearest-Neighbor Statistics for Two-Dimensional Rectangular Lattices

Richmond B. McQuistan
Physics Department and Laboratory for Surface Studies, University of Wisconsin—Milwaukee, Milwaukee, Wisconsin 53201
(Received 1 October 1970)

Expressions are developed which describe exactly the ensemble average of the number of nearest, next-nearest, and third-nearest-neighbor pairs per particle, when indistinguishable particles are arranged on a two-dimensional rectangular lattice.

I. INTRODUCTION

In previous papers[1,2] exact relationships were developed which describe the occupation statistics for one-dimensional arrays of dumbbells and λ-bells. (Here λ refers to the number of contiguous lattice sites occupied by a particle; λ = 2 for dumbbells.) The purpose of the present paper is to point out that these previously obtained results can be utilized to determine the ensemble average of the number of nearest-neighbor pairs, next-nearest-neighbor pairs, etc., for single particles (λ = 1) arranged on a two-dimensional retangular lattice.

II. NEAREST-NEIGHBORS PAIRS, n_1

Consider a rectangular lattice consisting of M columns and N rows on which are arranged a total of p particles. (See Fig. 1.) If there are q particles on one of the rows ($0 \leq q \leq M$) and these particles are arranged in all possible ways, then there are[2]

$$N_r(q, M) = (M - q + 1)\binom{M - r - 1}{q - r} \tag{1}$$

runs along that row which are composed of r contiguous particles. This expression is the result of the following reasoning. If we consider the run of r contiguous particles *and* one of the vacancies which terminates the run as a unit, then there are $M - r - 1$ individuals remaining. $q - r$ of these are particles, and $M - q - 1$ are vacancies, and these may be arranged in all possible ways. However each of the arrangements arising from the permutation of these $M - r - 1$ individuals may be created in $(M - q - 1) + 2 = M - q + 1$ ways because the unit consisting of the unit composed of the r particles and the terminating vacancy can be inserted between each of the $M - q$ vacancies in $M - q - 1$ ways, in addition, it may be inserted between a vacancy and each end of the array. When the q particles on a row are arranged in all possible ways and the positions of the remaining $p - q$ particles

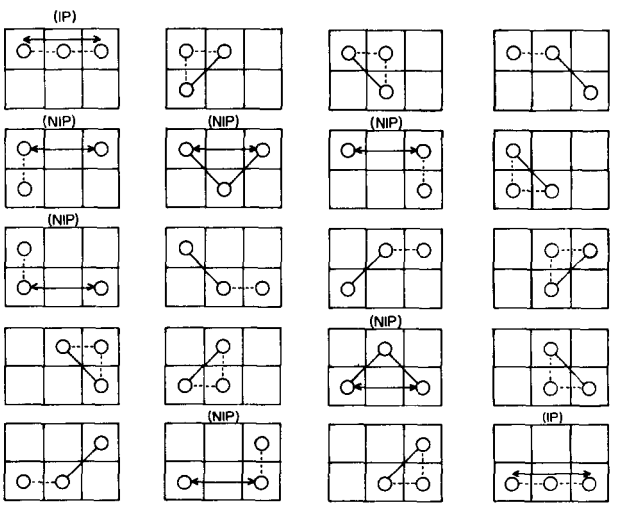


FIG. 1. When three particles are placed on a 2 × 3 array in all possible ways, there are 28 nearest-neighbor pairs (dashed lines); 16 next-nearest-neighbor pairs (solid lines); 8 third-nearest-neighbor pairs (arrows), 2 with an intervening particle (IP) and 6 with no intervening particle (NIP).

We noted previously the existence of identical relations between the σ_{2r+1} and the σ_{2r} for $O(n)$ and $Sp(n)$. For instance, $\sigma_3 = (\frac{1}{2}n - \eta)\sigma_2$. These relations are easily obtained from (71), if we note that $(\alpha - \frac{1}{2}n)f_{2r} + \frac{1}{2}(f_{2r})^i{}_i$ is anti-self-conjugate, so that

$$(\alpha f_{2r})^i{}_i = 0. \tag{75}$$

By substituting explicit forms for f_{2r}, we obtain the desired identities.

ACKNOWLEDGMENT

The author wishes to acknowledge several helpful discussions with Professor C. A. Hurst.

[1] See, e.g., M. Hammermesh, *Group Theory* (Addison-Wesley, Reading, Mass., 1962), p. 389.
[2] L. C. Biedenharn, J. Math. Phys. 4, 436 (1963).
[3] B. Gruber and L. O'Raiffeartaigh, J. Math. Phys. 5, 1796 (1964).
[4] S. Blaha, J. Math. Phys. 10, 2156 (1969).
[5] Y. Lehrer-Ilamed, Bull. Res. Counc. Israel **5A**, 197 (1956).
[6] A. J. Bracken and H. S. Green, J. Math. Phys. 12, 2099 (1971).
[7] G. Racah, in *Ergebnisse der exakten Naturwissenschaften* (Springer, Berlin, 1965), Vol. 37, pp. 66–71.
[8] G. Racah, Rend. Lincei **8**, 108 (1950).
[9] L. C. Biedenharn and J. D. Louck, Commun. Math. Phys. 8, 89 (1968); See also Refs. 2 and 3.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 10 OCTOBER 1971

Exact Nearest-Neighbor Statistics for Two-Dimensional Rectangular Lattices

Richmond B. McQuistan
Physics Department and Laboratory for Surface Studies, University of Wisconsin—Milwaukee, Milwaukee, Wisconsin 53201
(Received 1 October 1970)

Expressions are developed which describe exactly the ensemble average of the number of nearest, next-nearest, and third-nearest-neighbor pairs per particle, when indistinguishable particles are arranged on a two-dimensional rectangular lattice.

I. INTRODUCTION

In previous papers[1,2] exact relationships were developed which describe the occupation statistics for one-dimensional arrays of dumbbells and λ-bells. (Here λ refers to the number of contiguous lattice sites occupied by a particle; λ = 2 for dumbbells.) The purpose of the present paper is to point out that these previously obtained results can be utilized to determine the ensemble average of the number of nearest-neighbor pairs, next-nearest-neighbor pairs, etc., for single particles (λ = 1) arranged on a two-dimensional retangular lattice.

II. NEAREST-NEIGHBORS PAIRS, n_1

Consider a rectangular lattice consisting of M columns and N rows on which are arranged a total of p particles. (See Fig. 1.) If there are q particles on one of the rows ($0 \leq q \leq M$) and these particles are arranged in all possible ways, then there are[2]

$$N_r(q, M) = (M - q + 1)\binom{M - r - 1}{q - r} \tag{1}$$

runs along that row which are composed of r contiguous particles. This expression is the result of the following reasoning. If we consider the run of r contiguous particles *and* one of the vacancies which terminates the run as a unit, then there are $M - r - 1$ individuals remaining. $q - r$ of these are particles, and $M - q - 1$ are vacancies, and these may be arranged in all possible ways. However each of the arrangements arising from the permutation of these $M - r - 1$ individuals may be created in $(M - q - 1) + 2 = M - q + 1$ ways because the unit consisting of the unit composed of the r particles and the terminating vacancy can be inserted between each of the $M - q$ vacancies in $M - q - 1$ ways, in addition, it may be inserted between a vacancy and each end of the array. When the q particles on a row are arranged in all possible ways and the positions of the remaining $p - q$ particles

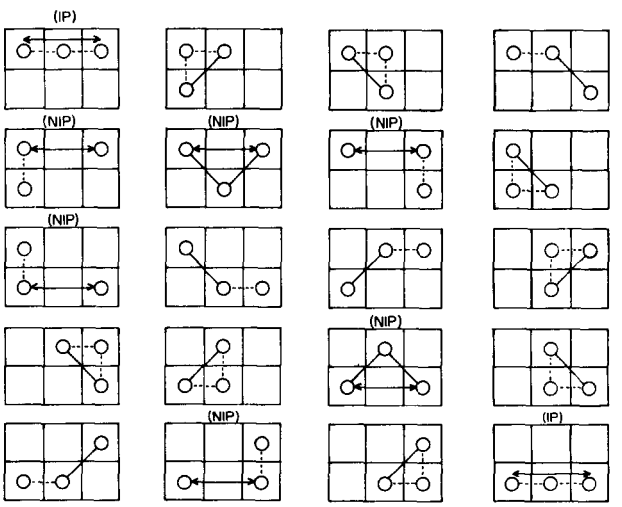


FIG. 1. When three particles are placed on a 2 × 3 array in all possible ways, there are 28 nearest-neighbor pairs (dashed lines); 16 next-nearest-neighbor pairs (solid lines); 8 third-nearest-neighbor pairs (arrows), 2 with an intervening particle (IP) and 6 with no intervening particle (NIP).

are held fixed, n, the number of nearest-neighbors pairs along that row is given by

$$n = \sum_{r=1}^{q} (r-1)N_r(q,M) = \sum_{r=1}^{q} (r-1)(M-q+1)\binom{M-r-1}{q-r} \quad (2)$$

because each run of r contiguous particles along a row contains $r-1$ pairs of filled nearest-neighbor sites. This sum is from the smallest to the largest possible run and yields

$$n = (M-1)\binom{M-2}{q-2}. \quad (3)$$

A row containing a particular arrangement of q particles will occur $\binom{M(N-1)}{p-q}$ times because the rest of the particles, $p-q$, are arranged in all possible ways on the rest of the sites, $M(N-1)$. Then n_{1h}, the total number of horizontal nearest-neighbor pairs which appear on the N rows, is obtained by summing

$$N(M-1)\binom{M-2}{q-2}\binom{M(N-1)}{p-q} \quad (4)$$

over all possible values of q. There are three possible intervals over which the sum is to be taken:

(i) A particular row can be empty but cannot be filled because $p < M$, i.e., $0 \leq q \leq p$;
(ii) a particular row can be empty and can be filled because $M < p < M(N-1)$, i.e., $0 \leq q \leq M$;
(iii) a particular row cannot be empty because $p > M(N-1)$, i.e., $p - M(N-1) \leq q \leq M$.

All three situations lead to the same conclusion:

$$n_{1h} = N(M-1)\binom{MN-2}{p-2}. \quad (5)$$

Similar reasoning leads to the number of "verticle" nearest-neighbor pairs so that n_1, the total number of nearest-neighbor pairs arising when p particles are arranged in all possible ways, is

$$n_1 = (2MN - M - N)\binom{MN-2}{p-2}. \quad (6)$$

Thus the ensemble average number of nearest-neighbor pairs per particle is

$$n_1/p\binom{MN}{p} = (2MN - M - N)\frac{(p-1)}{MN(MN-1)}. \quad (7)$$

In the limit as M and N approach infinity the number of nearest-neighbor pairs per particle is twice the lattice coverage.

III. NEXT-NEAREST-NEIGHBOR PAIRS, n_2

The following calculation of the number of next-nearest-neighbor pairs will be based on a determination of the number of next-nearest neighbors along a diagonal. In other words we will consider two particles occupying adjacent sites along a diagonal as constituting a next-nearest-neighbor pair. Assume that a diagonal d units long ($1 \leq d \leq N$) contains q particles and that these are arranged in all possible ways. Then there are[2] [see Eq. (1)]

$$N_r(q,d) = (d-q+1)\binom{d-r-1}{q-r} \quad (8)$$

runs along that diagonal which are composed of r contiguous particles. The number of occupied adjacent pair sites along the diagonal is then

$$\sum_{r=1}^{q} (r-1)N_r(q,d) = \sum_{r=1}^{q} (r-1)(d-q+1)\binom{d-r-1}{q-r} = (d-1)\binom{d-2}{q-2}. \quad (9)$$

When p particles are arranged in all possible ways, a diagonal containing a particular arrangement of q particles will occur $\binom{MN-d}{p-q}$ times because the $p-q$ remaining particles can be arranged in all possible ways on the remaining $(MN-d)$ sites. Thus the number of occupied adjacent pairs which appear along a diagonal of length d is obtained by summing

$$(d-1)\binom{d-2}{q-2}\binom{MN-d}{p-q}$$

over all possible values of q. There are three cases: (i) $0 \leq q \leq p$, (ii) $0 \leq q \leq d$, (iii) $p - MN + d \leq q \leq d$, and all yield[3]

$$(d-1)\binom{MN-2}{p-2}.$$

On an $M \times N$ rectangular lattice ($M \geq N$) there are (see Fig. 2) four diagonals 1 unit long, four diagonals 2 units long, four diagonals 3 units long, four diagonals $(N-1)$ units long, $2(M-N+1)$ diagonals N units long.

Each diagonal of length d ($1 \leq d \leq N-1$), occurs four times and the diagonal N units long occurs $2(M-N+1)$ times. Thus, n_2, the total number of next-nearest-neighbor pairs occurring when p particles are arranged in all ways, is

$$n_2 = 2(N-1)(M-1)\binom{MN-2}{p-2}. \quad (10)$$

The ensemble average number of next-nearest-neighbor pairs per particle is

$$n_2/p\binom{MN}{p} = (2MN - 2M - 2N - 2)\frac{(p-1)}{MN(MN-1)}. \quad (11)$$

In the limit as M and N approach infinity, the number of nearest-neighbor pairs per particle also increases twice as fast as the coverage.

IV. THIRD-NEAREST-NEIGHBOR PAIRS, n_3

A. With No Intervening Particle (NIP)

If there are q particles on one of the rows and they are arranged in all possible ways, then N_p, the

number of contiguous vacancies of length v, is given by[2]

$$N_v = (q+1)\binom{M-v-1}{q-1}, \tag{12}$$

because the v-tuple vacancy plus one of the particles terminating it can be considered as a unit containing $v+1$ individuals; there are then $M-v-1$ individuals remaining on the row of which $q-1$ are particles and $(M-q-v)$ are vacancies. These can be arranged in

$$\binom{M-v-1}{q-1}$$

ways. Distinguishable arrangements can be created in $(q-1)+2 = q+1$ ways by inserting the v-tuple vacancy and one of its terminating particles between and at the ends of the remaining $q-1$ particles.

To determine the number of third-nearest-neighbor pairs (with no intervening particles, NIP) along a row containing q particles arranged in all possible ways, we set $v=2$ in Eq. (12). Since the remaining $p-q$ particles may be arranged in the remaining $MN-3$ sites in all possible ways, the total number of "horizontal" third-nearest-neighbor pairs (NIP) is

$$N\sum_q (q+1)\binom{M-3}{q-1}\binom{MN-3}{p-q}$$

$$= N(M-2)\binom{MN-3}{p-2}, \tag{13}$$

where again there are the three possible ranges for q discussed in connection with Eq. (5). The number of "vertical" third-nearest-neighbor pairs (NIP) can be obtained by interchanging M and N so that n_3, the number of third-nearest-neighbor pairs (NIP) arising when p particles are arranged in all possible ways is

$$n_3 = (2MN-2M-2N)\binom{MN-3}{p-2}. \tag{14}$$

Thus, the number of third-nearest-neighbor pairs (NIP) per particle is

$$\frac{n_3}{p\binom{MN}{p}} = \frac{(2MN-2M-2N)(p-1)\quad(p-2)}{MN(MN-1)(MN-2)}. \tag{15}$$

For large M and N, the number of third-nearest-neighbor pairs (NIP) per particle varies as $2\theta(1-\theta)$, where θ is the lattice coverage.

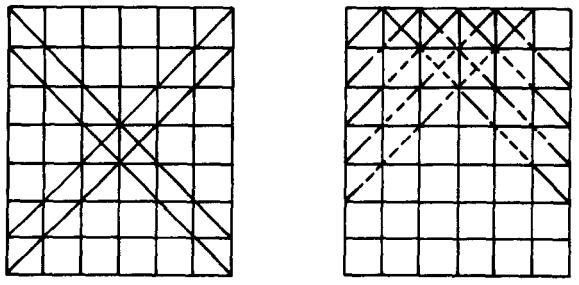


FIG. 2. On an $M=7, N=6$ array there are four diagonals 1, 2, 3, 4, 5 units long (only two of each are shown) and four diagonals 6 units long.

B. With Intervening Particle (IP)

If there are q particles along a row and these are arranged in all possible ways, then there are[2]

$$\sum_{r=2}^{q}(r-2)N_r(q,M)$$

$$= \sum_{r=2}^{q}(r-2)(M-q+1)\binom{M-r-1}{q-r}$$

$$= (M-2) \tag{16}$$

units composed of three particles in a row, i.e., a third-nearest-neighbor pair (IP). The remaining $p-q$ particles may be arranged on the remaining $MN-M$ sites in

$$\binom{MN-M}{p-q}$$

ways so that the total number of horizontal third-nearest-neighbor pairs (IP) is

$$n_{3h} = N(M-2)\binom{MN-3}{p-3}, \tag{17}$$

so that n_3, the total number of third-nearest-neighbor pairs (IP), is

$$n_3 = (2MN-2M-2N)\binom{MN-3}{p-3}, \tag{18}$$

or the ensemble average of the number of third-nearest-neighbor pairs (IP) per particle is

$$\frac{n_3}{p\binom{MN}{p}} = (2MN-2M-2N)\frac{(p-1)(p-2)}{(MN)(MN-1)(MN-2)}. \tag{19}$$

Thus, the total number of third-nearest neighbors per particle is $[(2MN-2M-2N)(p-1)]/(MN)(MN-1)$.

V. SUMMARY

We have calculated the ensemble average of the number of nearest-, next-nearest- and third-nearest-neighbor pairs per particle when particles are arranged on a two-dimensional array. In each case, for large arrays, this number varies as twice the coverage of the array.

The methods described above may be used for exact higher-order nearest-neighbor statistics.

ACKNOWLEDGMENTS

The author wishes to thank Professor D. Lichtman for a critical reading of the manuscript and R. Ristic for the drawings.

[1] D. Lichtman and R. B. McQuistan, J. Math. Phys. 8, 2441 (1967).
[2] R. B. McQuistan, Nuovo Cimento 58B, 86 (1968).
[3] J. Riordan, *An Introduction to Combinatorial Analysis* (Wiley, New York, 1958).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 10 OCTOBER 1971

On the Relation between Master Equations and Random Walks and Their Solutions*

Dick Bedeaux, Katja Lakatos-Lindenberg, and Kurt E. Shuler
Department of Chemistry, University of California San Diego, La Jolla, California 92037
(Received 18 January 1971)

It is shown that there is a simple relation between master equation and random walk solutions. We assume that the random walker takes steps at random times, with the time between steps governed by a probability density $\psi(\Delta t)$. Then, if the random walk transition probability matrix $\mathbf{M}$ and the master equation transition rate matrix $\mathbf{A}$ are related by $\mathbf{A} = (\mathbf{M} - 1)/\tau_1$, where τ_1 is the first moment of $\psi(t)$ and thus the average time between steps, the solutions of the random walk and the master equation approach each other at long times and are essentially equal for times much larger than the maximum of $(\tau_n/n!)^{1/n}$, where τ_n is the nth moment of $\psi(t)$. For a Poisson probability density $\psi(t)$, the solutions are shown to be identical at all times. For the case where $\mathbf{A} \neq (\mathbf{M} - 1)/\tau_1$, the solutions of the master equation and the random walk approach each other at long times and are approximately equal for times much larger than the maximum of $(\tau_n/n!)^{1/n}$ if the eigenvalues and eigenfunctions of $\mathbf{A}$ and $(\mathbf{M} - 1)/\tau_1$ are approximately equal for eigenvalues close to zero.

1. INTRODUCTION

There exists an extensive literature on master equations and random walks and their solutions.[1] We show in this paper that there is a close relation between random walks and master equations and their solutions. We consider random walks[2] in which the walker takes his steps at random times $t_1, t_2, \cdots$ and where the random variables $T_1 = t_1, T_2 = t_2 - t_1, \cdots, T_n = t_n - t_{n-1}$ have a common probability density $\psi(T)$. A random walk with constant time intervals $T_1 = T_2 = \cdots \equiv \tau$ between steps is the special case with $\psi(t) = \delta(t - \tau)$.

In Sec. 2 the random walk and the master equation are formally solved in terms of Green's functions. It is shown that a simple relation between the Green's functions exists if the master operator $\mathbf{A}$ and the random walk transition matrix $\mathbf{M}$ are related by $\mathbf{A} = (\mathbf{M} - 1)/\tau_1$, where τ_1 is the first moment of $\psi(t)$, i.e., the average time between steps.

In Sec. 3 it is shown that for a Poisson process, characterized by $\psi(t) = (1/\tau_1)e^{-t/\tau 1}$, the solutions for a random walk and the corresponding master equation with $\mathbf{A} = (\mathbf{M} - 1)/\tau_1$ are equal *for all times*. This is the only time step distribution for which this is the case. It is also shown that the random walk equation and the corresponding master equation are identical for a Poisson process. It should be stressed that this equivalence is valid independently of the value of the average time between steps, τ_1, and that it is not necessary to go to the limit $\tau_1 \longrightarrow 0$.

In Sec. 4 processes with arbitrary $\psi(t)$ are investigated. We discuss there the implications of the main result of this paper which can be stated here loosely as the following: If $\mathbf{A} = (\mathbf{M} - 1)/\tau_1$, then the random walk and master equation solutions approach each other at long times, and are approximately equal for times much greater than the maximum of $(\tau_n/n!)^{1/n}$, where τ_n is the nth moment of the distribution $\psi(t)$. For $\mathbf{A} \neq (\mathbf{M} - 1)/\tau_1$ some additional conditions must be imposed. These results are stated more precisely in Sec. 4 and are proved in the Appendix. We also present in Sec. 4 a mathematically precise formulation and a rigorous proof of the often stated equivalence of the random walk and master equations in the limit as the time interval between steps tends to zero.

2. FORMAL SOLUTION OF THE RANDOM WALK AND MASTER EQUATIONS

The general equation for a random walk is

$$P(\alpha; n+1) = \sum_{\alpha'} M_{\alpha\alpha'} P(\alpha'; n) = \mathbf{M}\, P(\alpha; n), \tag{2.1}$$

where $P(\alpha; n)$ is the probability that the walker is in state α after the nth step, $M_{\alpha\alpha'}$ is the probability that the walker goes from α to α' in one step, and $\mathbf{M}$ is the transition probability matrix. We impose no restrictions on the number of states between α and α', i.e., random walks with nonnearest neighbor transitions are included in our subsequent analysis. If there is a continuum of states, the sum over α is understood to be an integral over the continuous part and a sum over the discrete part of state space.

To calculate the probability $P(\alpha; t)$ that the random walker is in state α at time t, we must specify the probability that the random walker makes a step in a given time interval. We shall assume[2] that jumps are made at random times $t_1, t_2, t_3, \cdots$, where the random variables $T_i \equiv (t_i - t_{i-1})$, $i = 1, 2, \cdots$, with $t_0 = 0$, have a common probability density $\psi(T)$.

The general form of the master equation is

$$\frac{\partial}{\partial t} Q(\alpha; t) = \sum_{\alpha'} A_{\alpha\alpha'} Q(\alpha'; t) = \mathbf{A}\, Q(\alpha, t), \tag{2.2}$$

where $Q(\alpha; t)$ is the probability that the system is in state α at time t, $A_{\alpha\alpha'}$ is the transition rate from state α' to α, and $\mathbf{A}$ is the transition rate matrix. The transition rate $A_{\alpha\alpha'}$ is related to the more usually employed gain and loss rates $B_{\alpha\alpha'}$ by

$$A_{\alpha\alpha'} = B_{\alpha\alpha'} - \delta_{\alpha\alpha'} \sum_{\alpha''} B_{\alpha''\alpha}, \tag{2.3}$$

in which $\delta_{\alpha\alpha'}$ is the Kronecker δ for discrete states and the Dirac δ function for continuous states.

We consider only processes which are temporally homogeneous, i.e., for which $\mathbf{M}$ and $\psi(T)$ are independent of n and t, and A is independent of t.

The formal solution of the random walk equation, Eq. (2.1), is

$$P(\alpha;n) = \mathbf{M}^n P(\alpha;0). \tag{2.4}$$

The generating function[3] for the random walk is defined by

$$P(\alpha;z) \equiv \sum_{n=0}^{\infty} z^n P(\alpha;n) = (1 - z\mathbf{M})^{-1} P(\alpha;0), \tag{2.5}$$

and the corresponding Green's function is

$$G_{\alpha\alpha'}(z) = (1 - z\mathbf{M})^{-1} \delta_{\alpha,\alpha'}. \tag{2.6}$$

The formal solution of the random walk problem in continuous time is given by[2]

$$P(\alpha;t) = \sum_{n=0}^{\infty} \Phi_t(n) P(\alpha;n), \tag{2.7}$$

where $\Phi_t(n)$ is the probability that the walker has made exactly n steps at time t. This probability is related to the probability density $\psi_n(t)$ that the walker makes his nth step at time t, $t = t_n$, by

$$\Phi_t(n) = \int_0^t \psi_n(\tau) \int_{t-\tau}^{\infty} \psi(\tau')d\tau' d\tau, \tag{2.8}$$

where

$$\psi_n(t) = \int_0^t \psi_{n-1}(\tau)\psi(t-\tau)\,d\tau, \quad n > 1,$$
$$\psi_1(t) = \psi(t), \quad \psi_0(t) = \delta(t). \tag{2.9}$$

The generating function for the random walk as a function of time is defined by the Laplace transform:

$$P(\alpha;s) \equiv \int_0^{\infty} e^{-st} P(\alpha;t)dt. \tag{2.10}$$

Substitution of Eqs. (2.7)–(2.9) into Eq. (2.10) yields

$$P(\alpha;s) = \left(s - \frac{s\tilde{\psi}(s)}{1-\tilde{\psi}(s)}(\mathbf{M}-1)\right)^{-1} P(\alpha;0), \tag{2.11}$$

where

$$\psi(s) \equiv \int_0^{\infty} e^{-st}\psi(t)dt \tag{2.12}$$

and $P(\alpha;0) = P(\alpha;n=0) = P(\alpha;t=0)$. The corresponding Green's function is

$$G_{\alpha\alpha'}(s) = \left(s - \frac{s\tilde{\psi}(s)}{1-\tilde{\psi}(s)}(\mathbf{M}-1)\right)^{-1} \delta_{\alpha,\alpha'}. \tag{2.13}$$

This Green's function is related to the one in Eq. (2.6) by

$$G_{\alpha\alpha'}(s) = \frac{[1-\tilde{\psi}(s)]}{s} G_{\alpha\alpha'}[z = \tilde{\psi}(s)]. \tag{2.14}$$

The formal solution of the master equation, Eq. (2.2), is

$$Q(\alpha;t) = e^{\mathbf{A}t} Q(\alpha;0). \tag{2.15}$$

The generating function for the master equation is defined by the Laplace transform:

$$Q(\alpha;s) \equiv \int_0^{\infty} e^{-st} Q(\alpha;t)dt = (s - \mathbf{A})^{-1} Q(\alpha;0). \tag{2.16}$$

The corresponding Green's function is

$$F_{\alpha\alpha'}(s) = (s - \mathbf{A})^{-1} \delta_{\alpha,\alpha'}. \tag{2.17}$$

It will be shown in the following sections that the solutions of the random walk problem and the master equation are closely related if we make the identification

$$\mathbf{A} = (\mathbf{M} - 1)/\tau_1, \tag{2.18}$$

where

$$\tau_1 \equiv \int_0^{\infty} t\psi(t)dt \tag{2.19}$$

is the average time between steps. $F_{\alpha\alpha'}(s)$ and $G_{\alpha\alpha'}(s)$ are then related by

$$G_{\alpha\alpha'}(s) = \frac{[1-\tilde{\psi}(s)]}{\tau_1 s\tilde{\psi}(s)} F_{\alpha\alpha'}\left(\frac{1-\tilde{\psi}(s)}{\tau_1\tilde{\psi}(s)}\right), \tag{2.20}$$

while $F_{\alpha\alpha'}(s)$ and $G_{\alpha\alpha'}(z)$ are related by

$$G_{\alpha\alpha'}(z) = \tau_1 z\, F_{\alpha\alpha'}\left(\frac{1-z}{\tau_1 z}\right). \tag{2.21}$$

It is thus clear that the solution of any one of the three problems in terms of Green's functions immediately gives the solutions of the other two problems in terms of Green's functions. All three problems can in principle be solved by the diagonalization of the same operator.

In terms of the eigenfunctions and eigenvalues of the operator $\mathbf{A}$, the solutions of the random walk and the master equation can be written as

$$P(\alpha;n) = \sum_a (1 + \tau_1 a)^n f_a(\alpha) P(a;0), \tag{2.22}$$

$$P(\alpha;t) = \sum_a \left[\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} ds e^{st}\left(s - \frac{s\tilde{\psi}(s)\tau_1 a}{1-\tilde{\psi}(s)}\right)^{-1}\right] \times f_a(\alpha) P(a;0), \tag{2.23}$$

$$Q(\alpha;t) = \sum_a e^{at} f_a(\alpha) Q(a;0), \tag{2.24}$$

where

$$P(a;0) = \sum_a g_a(\alpha) P(\alpha;0), \tag{2.25}$$

$$Q(a;0) = \sum_a g_a(\alpha) Q(\alpha;0), \tag{2.26}$$

and $a, g_a(\alpha)$, and $f_a(\alpha)$ are the eigenvalues, left eigenfunctions, and right eigenfunctions of $\mathbf{A}$, respectively. If the spectrum of $\mathbf{A}$ has continuous parts, the sum over a is understood to be an integral over these parts. The quantity c in Eq. (2.23) is any positive constant.

In order that $P(\alpha;n)$ be a well-defined probability, $0 \leq P(\alpha;n) \leq 1$, it is clear that the Euclidean norm of the operator M must be bounded:

$$\|M\| \equiv \max\|Mh\| \leq 1, \qquad \|h\| = 1. \tag{2.27}$$

Because of the relation between **A** and **M** in Eq. (2.18), the bound on **M** gives upper and lower bounds to the spectrum of **A**:

$$-2/\tau_1 \leq a \leq 0. \tag{2.28}$$

Since we have made no restrictive statements on conservation of probability, the results of this paper are also valid for "open" systems, i.e., systems where probability is not preserved in that some or all of the walkers are removed in time (trapped, absorbed, evaporated, etc.).

The existence of a lower bound on the eigenvalues of the operator **A** is crucial to our technique of relating the solutions of the random walk and the master equations via their Green's function, Eqs. (2.14), (2.20), and (2.21). Since, however, eigenvalues with a large absolute value do not contribute appreciably to the long-time behavior of the solutions, it is possible to relate the long-time behavior of the solutions of a master equation with those of a random walk equation if the eigenvalues of **A** do not obey Eq. (2.28). In that case it is, however, impossible to relate **A** and **M** by Eq. (2.18); hence it is also impossible to give a simple relation between the various Green's functions. This will be explored further in Sec. 4.

3. POISSON PROCESSES

The following question immediately arises: Does there exist a probability density $\psi(t)$ such that the solutions of the master equation (2.2) and of the random walk equation in continuous time (2.7) are identical at all times for identical initial conditions? It follows from the Green's functions for both problems (2.13) and (2.17) that this will be the case if and only if **A** and **M** are related by Eq. (2.18) and if

$$\frac{s\tilde{\psi}(s)}{1-\tilde{\psi}(s)} = \frac{1}{\tau_1}. \tag{3.1}$$

Equation (3.1) has the solution

$$\tilde{\psi}(s) = (\tau_1 s + 1)^{-1}, \tag{3.2}$$

which then yields upon inversion

$$\psi(t) = \tau_1^{-1} e^{-t/\tau_1}. \tag{3.3}$$

This is the probability density for a Poisson process.[4] For such a process, the probability that the walker has made exactly n steps at time t is the Poisson distribution

$$\Phi_t(n) = \frac{1}{n!}\left(\frac{t}{\tau_1}\right)^n e^{-t/\tau_1}. \tag{3.4}$$

It is possible to show directly that the difference equation for a random walk with a Poisson density $\psi(t)$ is equivalent to a master equation at all times t. From the formal solution [Eq. (2.4)] of the random walk equation and the relation (2.7) one obtains

$$P(\alpha;t) = \sum_{n=0} \Phi_t(n)\, \mathbf{M}^n P(\alpha;0), \tag{3.5}$$

where $\sum_{n=0}^{\infty} \Phi_t(n)\mathbf{M}^n$ can be considered as an operator which translates the initial distribution to the distribution at time t.

Equation (3.5) yields:

$$\frac{P(\alpha;t+h)-P(\alpha;t)}{h} = h^{-1}\left[\exp\left(\frac{(t+h)}{\tau_1}(\mathbf{M}-1)\right) - \exp\left(\frac{t}{\tau_1}(\mathbf{M}-1)\right)\right] P(\alpha;0)$$

$$= h^{-1}\left[\exp\left(\frac{h}{\tau_1}(\mathbf{M}-1)\right) - 1\right] \exp\left(\frac{t}{\tau_1}(\mathbf{M}-1)\right) P(\alpha;0)$$

$$= h^{-1}\left[\exp\left(\frac{h}{\tau_1}(\mathbf{M}-1)\right) - 1\right] P(\alpha;t). \tag{3.6}$$

It is clear from what has been said above that Eq. (3.6) is valid for all $h > 0$. In the limit as $h \to 0$, Eq. (3.6) becomes

$$\frac{\partial P(\alpha;t)}{\partial t} = \frac{(\mathbf{M}-1)}{\tau_1} P(\alpha;t) = \mathbf{A}\, P(\alpha;t), \tag{3.7}$$

which is the master equation as given in Eq. (2.2).

The Poisson density $\psi(t)$ of Eq. (3.3) of time intervals between distinct events is characteristic of a large class of uncorrelated random processes developing in time. For such stochastic processes, where the random walk formulation with discrete steps is completely equivalent to the master equation formulation in continuous time for all times t, it is evidently only a matter of personal choice

which equation one wants to employ in the solution of the problem at hand.

4. PROCESSES WITH GENERAL DENSITIES $\psi(t)$

In this section we discuss the conditions under which the solutions of the random walk in continuous time and the random walk as a function of step number approach and are approximately equal to the solution of the master equation at long times. This analysis is subject to identical initial conditions for all processes, i.e.,

$$P(\alpha; n = 0) = P(\alpha; t = 0) = Q(\alpha; 0). \tag{4.1}$$

We will first analyze this problem for $\mathbf{A} = (\mathbf{M} - 1)\tau_1$ using the solutions of the random walk in continuous time and of the master equation, Eqs. (2.23) and (2.24):

$$P(\alpha; t) = \sum_a \theta_a(t) f_a(\alpha) P(a; 0), \tag{4.2}$$

$$Q(\alpha; t) = \sum_a e^{at} f_a(\alpha) P(a; 0), \tag{4.3}$$

where

$$\theta_a(t) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} ds e^{st} \left(s - \frac{s\tilde{\psi}(s)\tau_1 a}{1 - \tilde{\psi}(s)} \right)^{-1}. \tag{4.4}$$

Let τ_n by the nth moment of the probability density $\psi(t)$:

$$\tau_n \equiv \int_0^\infty t^n \psi(t) dt \tag{4.5}$$

and let us define

$$\gamma \equiv \sup(\tau_n / n!)^{1/n}. \tag{4.6}$$

The following theorem, which is the main result of this paper, is proved in the Appendix:

$$P(\alpha, t) - Q(\alpha, t) = Q(\alpha, t)\, O(\gamma/t) + O(e^{-t/\gamma}) \tag{4.7}$$

for $t \gg \gamma, \tau_1$ of order γ, and $\mathbf{A} = (\mathbf{M} - 1)/\tau_1$.

The order symbol O denotes that if $f(x) = O(x)$, then $f(x)/x$ remains bounded for all x. This theorem states that the solutions of the random walk in continuous time and the master equation are essentially equal for times much larger than γ. Our subsequent discussion in this section explores the implications of this theorem.

In order for theorem (4.7) to hold for all densities $\psi(t)$, it is necessary to note the following caveat. For a certain restricted class of sharply peaked densities $\psi(t)$, which are precisely defined in Eq. (A19) of the Appendix [an example would be $\psi(t) = \delta(t - T)$], it is necessary to exclude random walks with oscillatory solutions which persist at long times.[5] This corresponds to the exclusions of master equations with an operator $\mathbf{A}$ which has eigenvalues a in the range $0 \leq a + 2/\tau_1 \ll 1/\gamma$.

If the spectrum of $\mathbf{A}$ contains eigenvalues which satisfy the condition

$$0 < -a \ll 1/\gamma, \tag{4.8}$$

then the theorem implies that the two solutions are essentially equal *before* equilibrium is reached.

If any moment of $\psi(t)$ is infinite, it is clear from theorem (4.7) that the two solutions will become equal only *at* equilibrium. For probability densities $\psi(t)$ that decay at least exponentially at long times, all moments τ_n are finite. Typical examples of probability densities $\psi(t)$ with infinite γ are those that decay with negative powers of t for large times.

One quite frequently sees the statement that in the limit as the time interval between steps goes to zero, the random walk equation becomes equivalent to a master equation. We now give a mathematically precise formulation of this statement. We consider a sequence of random walks characterized by a sequence of transitions matrices $\{\mathbf{M}_i\}$ and a sequence of densities $\{\psi_i(t)\}$ with the property

$$\lim_{i\to\infty} \gamma_i = 0. \tag{4.9}$$

The meaning of condition (4.9) is that our sequence of random walks is so constructed that as the sequence index i increases, the moments of the probability density $\psi_i(t)$ all go to zero as specified. If we now define

$$P^{(\infty)}(\alpha; t) \equiv \lim_{i\to\infty} P^{(i)}(\alpha; t), \tag{4.10}$$

then $P^{(\infty)}(\alpha; t)$ is a solution of the master equation

$$\frac{\partial}{\partial t} P^{(\infty)}(\alpha; t) = \mathbf{A}^{(\infty)} P^{(\infty)}(\alpha; t) \tag{4.11}$$

if and only if the limit [see Eq. (2.18)]

$$\mathbf{A}^{(\infty)} \equiv \lim_{i\to\infty} (\mathbf{M}_i - 1)/\tau_{1,i} \tag{4.12}$$

exists, where as before, $\tau_{1,i}$ is of order γ_i. Therefore, the random walk equation indeed becomes equivalent to a master equation "in the limit as the time interval between steps goes to zero." Equation (4.12) is an analog of the well-known Kolmogoroff condition.[6]

It should be pointed out that if one considers the random walk in the continuous time and continuous space limits, which in general yields a diffusion-like equation, it is still necessary that Eq. (4.9) be satisfied. The conditions on the density $\psi(t)$ for passage to a diffusion equation are therefore the same as for passage to a master equation.

It is also possible to relate the solution of a random walk as a function of step number to the solution of the corresponding master equation. This is easily seen with the choice $\psi(t) = \delta(t - \tau_1)$, in which case $P(\alpha; n) = P(\alpha; t = n\tau_1)$. The use of

theorem (4.7) yields

$$P(\alpha;n) - Q(\alpha;n\tau_1) = Q(\alpha;n\tau_1)\, O(1/n) + O(e^{-n}) \tag{4.13}$$

for $n \gg 1$. Since the δ function is an example of a sharply peaked distribution as defined in Eq. (A19) of the Appendix, we must exclude transition operators **M** with eigenvalues m such that $0 \leq m + 1 \ll 1$.

We will now consider cases for which $\mathbf{A} \neq (\mathbf{M}-1)/\tau_1$ As will be proved in the Appendix, our main theorem, Eq. (4.7), and its consequences as discussed above, are still true, subject to a condition on the eigenvalues and eigenfunctions of **A** and **M**. If we define

$$\mathbf{B} \equiv (\mathbf{M} - 1)/\tau_1, \tag{4.14}$$

then this condition can be stated as follows: For eigenvalues a, b of the operators **A**, B which lie in the range

$$0 < -a \ll 1/\gamma, \quad 0 < -b \ll 1/\gamma, \tag{4.15}$$

there must be a one-to-one correspondence between the right eigenfunctions f_a, h_b of **A** and B such that

$$|a - b| \ll -a, \tag{4.16}$$

$$f_a(\alpha) - h_b(\alpha) = f_a(\alpha)\, O(\gamma a). \tag{4.17}$$

Theorem (4.7) thus still holds if the eigenvalues and eigenfunctions of A and $(\mathbf{M}-1)/\tau_1$ are approximately equal for eigenvalues close to zero. This extended version of theorem (4.7) specifies the class of random walk problems which have the same long-time behavior as that of a given master equation or vice versa.

The relation between **A** and **M** is determined by the physics of the problem. Consider, as an example, an open system in which the total probability is not conserved, due to irreversible trapping or evaporation. The transition rate A can now be written as the sum of two terms, $\mathbf{A} = \mathbf{A}_1 - \mathbf{A}_2$, where $\mathbf{A}_1$ conserves probability and $\mathbf{A}_2$ describes the irreversible loss process. Then the physically appropriate choice of **M** is[7]

$$\mathbf{M} = (1 + \tau_1\mathbf{A}_1)(1 + \tau_1\mathbf{A}_2)^{-1}. \tag{4.18}$$

The matrix **M** of Eq. (4.18) describes a situation in which the random walk within the system and the loss process therefrom are statistically independent and hence enter multiplicatively. That this choice of **M** is physically more reasonable for open systems than Eq. (2.18) is easily seen if one takes the case where $\mathbf{A}_2$ is a constant, $\mathbf{A}_2 = k$. Then $(1 + \tau_1\mathbf{A}_2)^{-1} = (1 + \tau_1 k)^{-1}$ is the probability per step that the random walker remains in the system. This probability ranges between 0 and 1 as the rate k in the master equation ranges between ∞ and 0. If $k \ll 1/\gamma$, then the conditions of Eqs. (4.15)–(4.17) are satisfied and the random walk solution and the master equation solution approach each other at long times.

A number of other examples could be discussed involving various physically plausible relations between **A** and **M** different from that in Eq. (2.18) for simple closed systems. In all such cases, conditions (4.14)–(4.17) determine whether and how rapidly the solution of the random walk and the master equation approach each other.

ACKNOWLEDGMENTS

One of us (K. L.-L.) would like to thank Dr. R. M. Pearlstein and Dr. R. P. Hemenger for many helpful and stimulating discussions.

APPENDIX:

We will now prove the theorem stated in Eq. (4.7).

We begin by considering densities $\psi(t)$ for which γ defined in Eq. (4.6) is finite. If this is the case, then $\tilde{\psi}(s)$ is analytic at $s = 0$ and can be expanded in a power series:

$$\tilde{\psi}(s) = 1 - \tau_1 s + \frac{\tau_2 s^2}{2!} - \frac{\tau_3 s^3}{3!} + \cdots, \tag{A1}$$

where the first term in the series is unity because $\psi(t)$ is normalized. The τ_n are defined in Eq. (4.5). Since $\psi(t)$ is positive-definite, all its moments τ_n are positive. Therefore the singularity of $\tilde{\psi}(s)$ closest to the origin will be on the negative real axis. The distance R of this pole from the origin is then the radius of convergence of the power series in Eq. (A1). This radius of convergence can be related to the moments of $\psi(t)$ by Hadamard's formula[8]

$$R = \left[\overline{\lim_{n\to\infty}}\, (\tau_n/n!)^{1/n}\right]^{-1} \geq \gamma - 1, \tag{A2}$$

where $\overline{\lim}$ indicates the limes superior, which is the greatest limit point of the sequence. This limit exists, since the sequence has the upper bound γ. Since $\tilde{\psi}(s)$ is analytic on the real axis to the right of $-R$, $\tilde{\psi}(s)$ is analytic for all s for which $\mathrm{Re}\, s > -R$, and for such s it is given by[9]

$$\tilde{\psi}(s) = \int_0^\infty e^{-st}\, \psi(t)dt, \quad \mathrm{Re}\, s > -R. \tag{A3}$$

The function $\theta_a(t)$ of Eq. (4.4) can be expressed in terms of the singularities of the integrand. Only the singularities to the right of $-R$ are important for the times we are interested in. For $\mathrm{Re}\, s > -R$, the only singularities are simple poles, which, for $a \neq 0$, are the zeros of the function $\tilde{\psi}(s) - (\tau_1 a + 1)^{-1}$ in that region:

$$\tilde{\psi}(s) - 1/(\tau_1 a + 1) = 0, \quad \mathrm{Re}\, s > -R. \tag{A4}$$

If a zero s_a is of nth order, then

$$\left.\frac{d^k\tilde{\psi}(s)}{ds^k}\right|_{s=s_a} = 0, \quad k = 1, 2, \cdots, n-1, \tag{A5}$$

and the pole in the integrand of Eq. (4.4) is also of nth order. For $a = 0$, it is easy to see that $\theta_0(t) = 1$.

Since $\tilde{\psi}(s)$ is real for s real and $s > -R$, it follows from the reflection principle[10] that

$$\tilde{\psi}(s^*) = \tilde{\psi}^*(s) \tag{A6}$$

for all s. This implies that the solutions of Eq. (A4) which do not lie on the real axis occur in complex conjugate pairs. From Eq. (A3) it follows that $\tilde{\psi}(s)$ is a monotonic decreasing function of s for s real and $s > -R$. Since $\tilde{\psi}(0) = 1$, Eq. (A4) has exactly one solution on the real axis for $-R < s \leq 0$ if

$$0 \geq a > \tau_1^{-1}\,[(\tilde{\psi}^+)^{-1} - 1] \geq -\tau_1^{-1}, \tag{A7}$$

where

$$\tilde{\psi}^+ \equiv \lim_{\epsilon \to 0^+} \tilde{\psi}(-R + \epsilon),$$

with ϵ positive and real. This solution is of first order. From Eq. (A3) it follows that

$$\mathrm{Re}\tilde{\psi}(r + i\Gamma) \leq \mathrm{Re}\tilde{\psi}(r) = \tilde{\psi}(r), \quad r > -R, \tag{A8}$$

where r and Γ are real. From this it follows that the complex solutions of Eq. (A4) do not lie to the right of the real solution discussed above. If Eq. (A7) is not satisfied, Eq. (A4) will have neither real nor complex solutions for $-1/\tau_1 \leq a \leq 0$. For $-2/\tau_1 < a < -1/\tau_1$, $(1 + \tau_1 a)^{-1}$ is negative. In this case Eq. (A4) can only have solutions on the real axis left of $-R$, but may very well have complex solutions to the right of $-R$. If we denote the solution on the real axis by s_a and the complex solutions in the upper half-plane by $s_{a,j} = \gamma_{a,j} + i\,\Gamma_{a,j}$, with $\gamma_{a,j}$ and $\Gamma_{a,j}$ real, then

$$-R < \gamma_{a,j} \leq s_a \quad \text{for} \quad -1/\tau_1 \leq a \leq 0. \tag{A9}$$

The complex solutions in the lower half-plane are than given by $s_{a,j} = \gamma_{a,j} - i\,\Gamma_{a,j}$, which follows from Eq. (A6).

For $t \gg \gamma$, the only contributions to $\theta_a(t)$ which are not of $O(e^{-t/\gamma})$ come from poles for which

$$\begin{aligned} &0 \leq -s_a \leq -\gamma_{a,j} \ll 1/\gamma \quad \text{for} \quad -1/\tau_1 \leq a \leq 0, \\ &0 \leq -\gamma_{a,j} \ll 1/\gamma \quad \text{for} \quad -2/\tau_1 < a < -1/\tau_1. \end{aligned} \tag{A10}$$

The function $\tilde{\psi}(s)$ can be expanded in a power series about any of the solutions $s_{a,j}$:

$$\tilde{\psi}(s) = \frac{1}{\tau_1 a + 1} - \tau_1^j (s - s_{a,j}) + \frac{\tau_2^j}{2!} \times (s - s_{a,j})^2 - \cdots, \tag{A11}$$

where

$$\tau_n^j \equiv \int_0^\infty t^n e^{-s_{a,j} t}\psi(t)dt. \tag{A12}$$

It is clear that

$$\begin{aligned} |\tau_n^j| &\leq \int_0^\infty t^n e^{-\gamma_{a,j} t}\psi(t)dt = \sum_{m=0}^{\infty} \frac{(-\gamma_{a,j})^m}{m!}\,\tau_{n+m} \\ &\leq \tau_n + \gamma^n \sum_{m=1}^{\infty} \frac{(n+m)!}{m!}\,(-\gamma_{a,j}\gamma)^m \\ &= \tau_n + \gamma^n \frac{d^n}{dx^n} \frac{x^{n+1}}{1-x}\Big|_{x=-\gamma_{a,j}\gamma}. \end{aligned} \tag{A13}$$

But

$$\frac{d^n}{dx^n}\,\frac{x^{n+1}}{(1-x)}\Big|_{x=-\gamma_{a,j}\gamma} \ll (n+1)! \tag{A14}$$

This yields

$$|\tau_n^j| \leq \tau_n + \Delta, \qquad \Delta \ll (n+1)!\gamma^n. \tag{A15}$$

From Eqs. (A15) and (A8) and the fact that $|(1 + \tau_1 a)^{-1}| \geq 1$, it follows that poles with real parts that fulfill Eq. (A10) can occur only for values of a in the ranges $0 \leq -a \ll 1/\gamma$ and $0 \leq a + 2/\tau_1 \ll 1/\gamma$. In addition, we will show that the complex poles with real parts $\gamma_{a,j}$ satisfying Eq. (A10) occur only in the extreme case where $\psi(t)$ is a superposition of very sharp peaks. For such $s_{a,j}$, Eqs. (A8), (A11), and (A15) show for the real and imaginary parts of $\tilde{\psi}(i\Gamma_{a,j})$ that

$$0 \leq 1 - \left|\int_0^\infty \cos\Gamma_{a,j}\, t\,\psi(t)dt\right| \ll 1 \tag{A16}$$

and

$$\left|\int_0^\infty \sin\Gamma_{a,j}\, t\,\psi(t)dt\right| \ll 1. \tag{A17}$$

Since $\psi(t)$ is normalized to unity, Eqs. (A16) and (A17) can hold for $0 \leq -a \ll 1/\gamma$ only if $\psi(t)$ is appreciably different from zero only for

$$\begin{aligned} &|t - 2\pi n/\Gamma_{a,j}| \ll 2\pi/\Gamma_{a,j}, \\ &n = 0 \text{ or } n = 1 \text{ or } n = 2 \text{ or } \cdots, \end{aligned} \tag{A18}$$

with $\Gamma_{a,j}$ of order $1/\gamma$ or greater. For $0 \leq a + 2/\tau_1 \ll 1$, Eqs. (A16) and (A17) can hold only if $\psi(t)$ is appreciably different from zero only for

$$\begin{aligned} &|t - 2\pi(n + 1/2)/\Gamma_{a,j}| \ll 2\pi/\Gamma_{a,j}, \\ &n = 0 \text{ or } n = 1 \text{ or } n = 2 \text{ or } \cdots, \end{aligned} \tag{A19}$$

with $\Gamma_{a,j}$ of order $1/\gamma$ or greater. It is possible that a density $\psi(t)$ belongs to both of the cases described above. An example of such a density is $\psi(t) = \delta(t - \Delta t)$. The values of $\psi(t)$ outside of these peaks must be so small that their contribution to $\int_0^\infty \psi(t)dt$ is $\ll 1$. This proves that complex poles satisfying Eq. (A10) can occur only if $\psi(t)$ is sharply peaked in the manner described above.

We now proceed to show that the poles corresponding to $0 \leq -a \ll 1/\gamma$, which lie close to the imaginary s axis, Eq. (A10), are of first order and that

the distances between them are of the order $2\pi/\gamma$. We first consider the real pole s_a. For $|s| \ll 1/\gamma$, Eq. (A1) becomes

$$\tilde{\psi}(s) = 1 - \tau_1 s + O(s^2\gamma^2). \tag{A20}$$

For τ_1 of order γ, the third term in Eq. (A20) is much smaller than the second term. Equation (A4) then yields

$$s_a = a[1 + O(\gamma a)], \tag{A21}$$

$$|s_a - a| \ll -a. \tag{A22}$$

From the way in which this solution was constructed, with τ_1 of order γ, it is clear that there are no other solutions of Eq. (A4) fulfilling Eq. (A22).

The poles due to a sharply peaked density $\psi(t)$ of the type described above which lie within a distance $\ll 1/\gamma$ of the imaginary s axis lie within a distance $\ll 1/\gamma$ of a finite subset of the poles of an appropriately chosen superposition of δ functions. To analyze the behavior of such sharply peaked densities, it is therefore sufficient to study distributions which are superpositions of δ functions. Consider, therefore,

$$\psi(t) = \sum_{n=0}^{\infty} c_n \delta(t - n\Delta t), \qquad c_n \geq 0, \tag{A23}$$

with Δt of order γ or less. Then

$$\tilde{\psi}(s) = \sum_{n=0}^{\infty} c_n e^{-ns\Delta t}. \tag{A24}$$

This is a periodic function:

$$\tilde{\psi}(s + 2\pi i/\Delta t) = \tilde{\psi}(s). \tag{A25}$$

Therefore, if s_a is the solution of Eq. (A4) on the real axis, then

$$s_{a,j} = s_a + 2\pi ij/\Delta t \tag{A26}$$

are also solutions of Eq. (A4). Since s_a is a first-order pole, so are the $s_{a,j}$. The distance between the poles is of the order $2\pi/\gamma$.

The results up to this point can be summarized as follows. It is possible to divide the probability densities $\psi(t)$ with finite γ and τ_1 of order γ into three classes:

(i) Densities which are superpositions of sharp peaks as specified in Eq. (A19). For such distributions we will consider only master operators **A** which have no eigenvalues in the range $0 \leq a + 2/\tau_1 \ll 1/\gamma$.

(ii) Densities which are superpositions of sharp peaks as specified in Eq. (A18) excluding those which are of the first class.

(iii) All other densities.

Densities in the third class produce only a single pole in the integrand of $\theta_a(t)$ which contributes for long times. This pole is real and of first order. Densities in the first and second classes produce, in addition, first order complex poles which lie sufficiently close to the imaginary s axis to contribute to $\theta_a(t)$ at long times for $0 \leq -a \ll 1/\gamma$. Densities in the first class produce additional poles in the range $0 \leq a + 2/\tau_1 \ll 1/\gamma$. Since in this class of densities we only consider master operators **A** which have no eignvalues in this range, these poles do not contribute to $P(\alpha, t)$. This corresponds to the exclusion of transition matrices **M** of the random walk with eigenvalues which cause oscillations in the random walk solution that persist at long times. These eigenvalues cause no persistent oscillations for distributions of the second and third classes and therefore need not be excluded for these cases.

We will first consider densities which are in the third class. In this case Eq. (4.4) can be written as

$$\theta_a(t) = -\tau_1 a(1 + \tau_1 a)^{-2} s_a^{-1} e^{s_a t} \times \left(\frac{d}{ds}\tilde{\psi}(s)\Big|_{s=s_a}\right)^{-1} + O(e^{-t/\gamma}) \tag{A27}$$

Via Eqs. (A20) and (A21), $\theta_a(t)$ becomes

$$\theta_a(t) = e^{at}[1 + O(\gamma a)] + O(e^{-t/\gamma}). \tag{A28}$$

The only values of a for which the first term is not of the same order as the second term are those for which at is of order 1 or less. Therefore,

$$\theta_a(t) = e^{at}[1 + O(\gamma/t)] + O(e^{-t/\gamma}). \tag{A29}$$

We will now show that this result is also valid for densities in the first and second classes. For these cases Eq. (4.4) can be written as

$$\theta_a(t) = -\tau_1 a(1 + \tau_1 a)^{-2} s_a^{-1} e^{s_a t} \left(\frac{d}{ds}\tilde{\psi}(s)\Big|_{s=s_a}\right)^{-1} - 2\tau_1 a(1 + \tau_1 a)^{-2} \sum_j e^{r_{a,j} t} \times \left| s_{a,j} \frac{d}{ds}\tilde{\psi}(s)\Big|_{s=s_{a,j}}\right|^{-1} \cos[\Gamma_{a,j} t + \eta(s_{a,j})] + O(e^{-t/\gamma}), \tag{A30}$$

where the phase shifts are

$$\exp[i\eta(s_{a,j})] \equiv \left| s_{a,j} \frac{d\tilde{\psi}(s)}{ds}\Big|_{s=s_{a,j}} \Big/ s_{a,j} \frac{d\tilde{\psi}(s)}{ds}\Big|_{s=s_{a,j}}\right| \tag{A31}$$

Equation (A30) differs from Eq. (A27) only in the second term of Eq. (A30):

$$\xi_a(t) = -2\tau_1 a(1 + \tau_1 a)^{-2} \sum_j e^{r_{a,j} t} \times \left| s_{a,j} \frac{d}{ds}\tilde{\psi}(s)\Big|_{s=s_{a,j}}\right|^{-1} \cos[\Gamma_{a,j} t + \eta(s_{a,j})]. \tag{A32}$$

For a density $\psi(t)$ which is a superposition of δ functions, we obtain

$$\xi_a(t) = 2\tau_1 a[1 + O(\gamma a)]e^{at} \sum_{j=1}^{\infty} \frac{\Delta t}{2\pi j \tau_1} \sin\frac{2\Delta j t}{\Delta t}$$

$$= a\Delta t[1 + O(\gamma a)]\, e^{at}\left(\tfrac{1}{2} - \frac{t}{\Delta t} + \left[\frac{t}{\Delta t}\right]\right), \qquad \text{(A33)}$$

where $[t/\Delta t]$ is the integer part of $t/\Delta t$. Since Δt is of order γ or less,

$$\xi_a(t) = O(\gamma a)e^{at}. \qquad \text{(A34)}$$

If $\psi(t)$ is a superposition of peaks of nonzero width as given in Eq. (A18) or Eq. (A19), then the poles $s_{a,j}$ which lie within a distance $\ll 1/\gamma$ of the imaginary s axis lie within a distance $\ll 1/\gamma$ of a finite subset of the poles of the corresponding superposition of δ functions. Then

$$\xi_a(t) = 2\tau_1 a[1 + O(\gamma a)]\, e^{at} \sum_j e^{(r_{a,j} - s_a)t}$$

$$\times \frac{\Delta t_v}{2\pi n_j \tau_1} \sin\frac{2\pi n_j t}{\Delta t}, \qquad \text{(A35)}$$

where the sum over j is over a finite number of poles and where n_j is the number of the pole of the corresponding superposition of δ functions. Since $r_{a,j} \leq s_a$ and the sum is a finite sum, this case also yields Eq. (A34). Substitution of Eq. (A34) into (A30) again yields (A28) for $\theta_a(t)$. Therefore, for both classes of distributions, Eq. (A29) holds.

Substitution of Eq. (A29) into (4.2) yields

$$P(\alpha, t) = Q(\alpha, t)\,[1 + O(\gamma/t)] + O(e^{-t/\gamma}) \qquad \text{(A36)}$$

for τ_1 of order γ, which proves theorem (4.7) for distributions $\psi(t)$ with finite γ.

For densities with infinite γ, the theorem is trivially true and is essentially empty. If γ is infinite, either $\tilde{\psi}(s)$ has an essential singularity at $s = 0$, or $s = 0$ is an accumulation point of singularities of $\tilde{\psi}(s)$. In both cases, the integrand of $\theta_a(t)$ will contribute in every open neighborhood of $s = 0$. Therefore, $\theta_a(t)$ can never approach an exponential at any time. For such distributions, $P(\alpha, t)$ and $Q(\alpha, t)$ will never approach each other even after long times, and will only be equal at equilibrium.

We will now prove the extended version of theorem (4.7). For this case the solution of the random walk equation can be written as

$$P(\alpha; t) = \sum_b \theta_b(t)\, h_b(\alpha)\, P(b; 0). \qquad \text{(A37)}$$

A consequence of the conditions (4.15)–(4.17) is that Eq. (A37) can be written as

$$P(\alpha; t) = \sum_a{}' \theta_b(t)\, h_b(\alpha)\, P(b; 0) + O(e^{-t/\gamma}), \qquad \text{(A38)}$$

where the prime on the sum indicates that the sum is taken only over those values of a that obey condition (4.15), and a and b are related by the one-to-one correspondence between the eigenfunctions f_a and h_b, Eqs. (4.16) and (4.17). A consequence of conditions (4.16) and (4.17) is that

$$\theta_b(t)\, h_b(\alpha)\, P(b; 0) = \theta_a(t) f_a(\alpha)\, P(a; 0)\,[1 + O(\gamma a)], \qquad \text{(A39)}$$

which immediately yields the desired result.

Note added in proof: A quantity that is often of physical interest is the mean time for a walker to reach state α for the first time. Let $R(\alpha, n|\alpha', 0)$ be the probability that a walker starting at state α' reaches state α for the first time on the nth step. Then

$$P(\alpha, n|\alpha', 0) = \sum_{n'=1}^{n} R(\alpha, n'|\alpha', 0) P(\alpha, n - n'|\alpha, 0).$$

The generating function

$$R_{\alpha\alpha'}(z) \equiv \sum_{n=1}^{\infty} z^n R(\alpha, n|\alpha', 0)$$

is then related to $G_{\alpha\alpha'}(z)$ by $R_{\alpha\alpha'}(z) = (G_{\alpha\alpha'}(z) - \delta_{\alpha\alpha'})/G_{\alpha\alpha}(z)$. The mean first passage time to state α is

$$\tau_{\alpha\alpha'} = \int_0^{\infty} dt \sum_{n=1}^{\infty} t\psi_n(t) R(\alpha, n|\alpha', 0)$$

$$= \tau_1 \frac{d}{dz} R_{\alpha\alpha'}(z)\Big|_{z=1} = \tau_1 \langle n\rangle_{\alpha\alpha'}$$

where $\langle n\rangle_{\alpha\alpha'} = (d/dz)R_{\alpha\alpha'}(z)|_{z=1}$ is the mean number of steps to reach state α for the first time. The mean first passage time is thus the same for all random walks for which τ_1 is the same, including the master equation. Higher moments of the first passage time can be similarly calculated. The nth moment will, in general, be a function of $\tau_1, \tau_2, \ldots, \tau_m$.

---

* Supported in part by the National Science Foundation, under Grant No. GP10536 and by the Advanced Research Projects Agency of the Department of Defense, monitored by the U.S. Office of Naval Research, under Contract No. N00014-69-A-0200-6018.

1 E.g., W. Feller, *An Introduction to Probability Theory and Its Applications* (Wiley, New York, 1966, 1968), 3rd ed., Vol. I, and 1st ed., Vol. II; F. Spitzer, in *Principles of Random Walks* (Van Nostrand, Princeton, N.J., 1964); *Selected Papers on Noise and Stochastic Processes*, edited by N. Wax (Dover, New York, 1954); I. Oppenheim, K. E. Shuler, and G. H. Weiss, Advan. Mol. Relaxation Process. **1**, 13 (1967).

2 E. W. Montroll and G. H. Weiss, J. Math. Phys. **6**, 167 (1965).

3 E. W. Montroll, Proc. Symp. Appl. Math. **16**, 193 (1964).

4 W. Feller, Ref. 1., Vol. I, p. 447, and Vol. II, p. 58.

5 For a discussion of such solutions and the associated eigenvalue spectrum see, e.g., J. G. Kemeny and J. L. Snell, *Finite Markov Chains* (Van Nostrand, Princeton, N.J., 1960) Chap. 5; E. W. Montroll, *Energetics in Metallurgical Phenomena* (Gordon and Breach, New York, 1967), Vol. 3, p. 139 et seq.

6 A. Kolmogoroff, Math. Ann. **104**, 451 (1931); **108**, 149 (1933).

7 K. Lakatos-Lindenberg, R. P. Hemenger, and R. M. Pearlstein (unpublished).

8 L. V. Ahlfors, *Complex Analysis* (McGraw-Hill New York, 1953), p. 141.

9 D. V. Widder, *The Laplace Transform* (Princeton U.P., Princeton, N.J., 1946), p. 36.

10 R. V. Churchill, *Complex Variables and Applications* (McGraw-Hill, New York, 1960), p. 265.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 10 OCTOBER 1971

The Korteweg–de Vries Equation: Transition Period

L. Y. Shih
National Research Council of Canada, Ottawa, Canada
(Received 13 November 1970)

A general form of the nonlinear wave equation with dispersive and dissipative terms involving small coefficients has been treated for the transition period when the time is around the breakdown time. In the first part of this paper, some relations have been established in order to see how the classical overtaking phenomenon is eventually prevented when the curve has steepened sufficiently. In the second part of this paper, the growth and damping of solitary waves due to the effects of dispersion and dissipation have been analyzed.

1. INTRODUCTION

Study of various physical problems (such as magnetosonic waves, ion-acoustic waves in plasmas, turbulence, shallow water waves, and long waves in anharmonic crystals) lead to a nonlinear partial differential equation of the general form

$$u_t + P(u)\, u_x = Q[u], \tag{1.1}$$

subject to the initial condition

$$u(x, 0) = f(x),$$

where x and t are normalized space and time coordinates, respectively. Here, $P(u)$ is a real function of u; $Q[u]$ is a linear operator which may consist of dissipative and dispersive terms with small coefficients, say

$$Q[u] = \mu L[u] - \delta^2 u_{xxx}, \tag{1.2}$$

where the first term stands for dissipation and the second term for dispersion.

In many cases, the function $P(u)$ is simply u itself. When $\mu = 0$, (1.1) reduces to Korteweg–de Vries equation,[1] where only the effect of dispersion is considered. On the other hand, if $\delta = 0$ and $L[u] = u_{xx}$, (1.1) reduces to Burgers' equation,[2] where only the effect of dissipation is considered.

In general, the dissipative operator has different forms for various physical systems. For example, Ott and Sudan[3] provided the expressions for four cases:

(a) magnetosonic waves damped by electron–ion collisions,
(b) ion-acoustic waves damped by ion–neutral collisions,
(c) ion-acoustic waves with electron Laudau damping, and
(d) shallow water waves damped by viscosity.

According to the numerical computation of Korteweg–de Vries equation,[4] it was observed that the time variation of the solution $u(x, t)$ can be divided into three stages, namely, initial, transition, and final. In the initial stage when the time is comparatively smaller than the breakdown time t_B, defined as the time when the solution of the equation $u_t + P(u)\, u_x = 0$ starts to yield an infinite slope, the effects of dispersion and dissipation are negligible and the classical overtaking phenomenon occurs; that is to say, $u(x, t)$ steepens in regions where $P'(u)\, u_x < 0$.

After $u(x, t)$ has steepened sufficiently (at $t \approx t_B$), the Q term becomes significant and serves to prevent the formation of discontinuities. In addition, due to the effect of dispersion, oscillations of small wavelength (of order δ) develop on the negative side of the front. This stage is called the transition period. The amplitudes of the oscillations grow and finally each oscillation achieves a quasisteady amplitude (which varies linearly with respect to x) and has a shape almost identical to that of an individual soliton solution. On the other hand, the dissipative term has an effect to damp these oscillations.

Existence and uniqueness of solution of Korteweg–de Vries equation have been proved by Sjöberg.[5] Equations of conservation form were treated by Miura, Gardner, Kruskal, and Su in a series of papers.[6–8]

In the present paper, a generalized nonlinear dispersive wave equation with dissipation is treated from the gometrical point of view. Efforts are concentrated on the transition period when solitary waves tend to emerge.

2. ABSENCE OF DISCONTINUITY

The curve representing the solution of (1.1) in the (u, x) diagram may be conceived as a vibrating string, any point of which possesses a longitudinal velocity $V_{\parallel} = P(u)$ and a transverse velocity $V_{\perp} = Q[u]$, where u is equivalent to the transverse displacement of the string. For the simplest equation

$$u_t + P(u)u_x = 0, \tag{2.1}$$

$V_{\perp} = 0$, and the points move longitudinally.

Let α denote the angle of velocity and θ the angle of the curve, both measured from the longitudinal axis. We have

$$\frac{\tan\alpha}{\tan\theta} = \frac{Q[u]}{P(u)u_x} = 1 + \frac{u_t}{P(u)u_x}. \tag{2.2}$$

If $P(u)Q[u]u_x < 0$, the velocity vectors $V_{\parallel}$ and $V_{\perp}$ lie on the same side of the curve $u(x, t)$. If $P(u)Q[u]u_x > 0$, the velocity vectors $V_{\parallel}$ and $V_{\perp}$ lie on different sides of the curve.

Differentiating (1.1) with respect to x, we get

$$\frac{D\theta}{Dt} \equiv \theta_t + P(u)\theta_x = -P'(u)\sin^2\theta + Q[u_x]\cos^2\theta, \tag{2.3}$$

where the prime designates differentiation. It may be written as

$$\frac{D}{Dt}(\tfrac{1}{2}\theta^2) = -P'(u)\theta \sin^2\theta + Q[u_x]\theta\cos^2\theta, \quad (2.4)$$

where $-\frac{1}{2}\pi < \theta < \frac{1}{2}\pi$.

For moderate values of u_x, the $Q[u_x]$ term as assumed is comparatively small. Equation (2.4) indicates that the curve steepens in regions where $P'(u)\theta$ is negative. As $|u_x|$ exceeds a sufficiently large value, the $Q[u_x]$ term becomes significant and counteracts the formation of discontinuities, provided that the following condition is satisfied.

Necessary condition 1: There exists a large but finite number K, such that for $|u_x| > K$,

$$Q[u_x]/P'(u)u_x^2 > 1. \quad (2.5)$$

Equation (1.1) may also be expressed in such a way that x is a function of u and t. Since $u_t = -x_t/x_u$, $u_x = 1/x_u$, and

$$\frac{\partial}{\partial x} = \frac{1}{x_u}\frac{\partial}{\partial u},$$

we have

$$x_t - P(u) = \frac{\delta^2}{2}\left(\frac{1}{x_u^2}\right)_{uu} - \mu\left(\frac{1}{x_u}\right)_u \quad (2.6)$$

for $L[u] = u_{xx}$. Differentiating (2.6) with respect to u, we get

$$\left(\frac{\partial\theta}{\partial t}\right)_{u=c} = -\sin^2\theta\left[P'(u) - \left(\frac{\delta^2 x_{uu}}{x_u^3} + \frac{\mu}{x_u}\right)_{uu}\right. \quad (2.7)$$

along constant u lines. A similar condition should be satisfied.

Necessary condition 2: There exists a small number $k > 0$, such that for $|x_u| < k$,

$$(\delta^2 x_{uu}/x_u^3 + \mu/x_u)_{uu}/P'(u) > 1. \quad (2.8)$$

It can easily be verified that for curves with shapes similar to those shown in Fig. 1, conditions (2.5) and (2.8) are true.

The velocity components which affect the shape of curve are those tangent and normal to the curve. They may be defined as

$$\begin{bmatrix} V_T \\ V_N \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} P(u) \\ Q[u] \end{bmatrix}. \quad (2.9)$$

Since

$$\sin\theta = u_x/(1+u_x^2)^{1/2}$$

and

$$\cos\theta = (1+u_x^2)^{-1/2},$$

we may write

$$V_T = \frac{P(u) + Q[u]u_x}{(1+u_x^2)^{1/2}} \quad (2.10)$$

and

$$V_N = u_t/(1+u_x^2)^{1/2}. \quad (2.11)$$

Differentiating (2.9) with respect to s, where s denotes the distance along the curve, we obtain

$$\frac{D\theta}{Dt} = \left(\frac{\partial V_N}{\partial s}\right)_{t=c} + KV_T, \quad (2.12)$$

$$\frac{1}{1+\epsilon}\frac{D\epsilon}{Dt} = \left(\frac{\partial V_T}{\partial s}\right)_{t=c} - KV_N, \quad (2.13)$$

where K denotes the curvature and ϵ the relative stretching of the curve. The rate of change of normal velocity along the curve represents the angular speed with which the curve turns locally. Equation (2.12) indicates that the curvature associated with the tangential velocity also plays an important role on the change of slope.

3. WAVE PROPAGATION

During the evolution of solitons, it may be assumed that disturbances propagating along the curve are small, thus the solution $u(x,t)$ can be approximately expressed as combination of a translation and a wave motion; i.e.,

$$u(x,t) = u_1(x,t) + A(x,t)\exp[i(kx - \omega t)], \quad (3.1)$$

where $u_1(x,t)$ satisfies

$$(u_1)_t + P(u)(u_1)_x = Q[u_1]$$

and

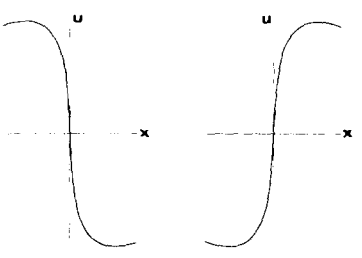
$$A(x,t) \ll u_1(x,t).$$

Substituting (3.1) into (1.1), we obtain

$$\omega/k = P(u) - \delta^2k^2 + 3\delta^2A_{xx}/A - 2\mu A_x/A, \quad (3.2)$$

$$A_t + (P(u) - 3\delta^2k^2)A_x + \mu(k^2A - A_{xx}) + \delta^2A_{xxx} = 0. \quad (3.3)$$

In order to shed light on the phenomena of waves propagating along the curve, we deal with two simple cases. For Korteweg–de Vries equation,



(a) $P'(u) > 0$ (b) $P'(u) < 0$

FIG. 1 Typical profiles

(3.2) and (3.3) reduce to

$$\omega/k = P(u) - \delta^2 k^2 + 3\delta^2 A_{xx}/A, \tag{3.4}$$

$$A_t + [P(u) - 3\delta^2 k^2]A_x + \delta^2 A_{xxx} = 0. \tag{3.5}$$

If $A(x, t)$ is a slowly varying function of x, we may assume that

$$A_{xx}/A,\ A_{xxx}/A_x \ll k^2.$$

Thus, from (3.4) and (3.5) we have

$$\omega/k \approx P(u) - 4\pi^2(\delta/\lambda)^2$$

and

$$A_t/A_x \approx 12\pi^2(\delta/\lambda)^2 - P(u).$$

It indicates that the wave with wavelengths of order δ propagates in the negative direction; the wave grows if the amplitude is an increasing function of x.[4]

Next, we consider Burgers' equation. For this case, (3.2) and (3.3) reduce to

$$\omega/k = P(u) - 2\mu A_x/A, \tag{3.6}$$

$$A_t + P(u)A_x + \mu(k^2 A - A_{xx}) = 0. \tag{3.7}$$

Assuming that $A(x, t)$ can be expressed in the form of $\exp(\alpha t + \beta x)$, from (3.6) and (3.7) we obtain

$$2\mu\beta = P(u) - [\{P(u)\}^2 + 4\mu(\mu k^2 + \alpha)]^{1/2}$$

and

$$\omega/k = [\{P(u)\}^2 + 4\mu(\mu k^2 + \alpha)]^{1/2}.$$

Consider the wavelength $\lambda \leq O(\mu)$, α and $P(u) \leq O(1)$, then $\beta = -O(1/\lambda)$. It indicates that the wave will be damped out within a distance of the order of a wavelength. This explains why Burgers' equation has no solitary wave as in the case of Korteweg-de Vries equation.

1 D. J. Korteweg and G. de Vries, Phil. Mag. 39, 422 (1895).
2 J. M. Burgers, Proc. Acad. Sci. Amsterdam 43, 2 (1940).
3 E. Ott and R. N. Sudan, Phys. Fluids 13, 1432 (1970).
4 N. J. Zabusky and M. D. Kruskal, Phys. Rev. Letters 15, 240 (1965).
5 A. Sjöberg, J. Math. Anal. Appl. 29, 569 (1970).
6 R. M. Miura, J. Math. Phys. 9, 1202 (1968).
7 R. M. Miura, C. S. Gardner, and M. D. Kruskal, J. Math. Phys. 9, 1204 (1968).
8 C. H. Su and C. S. Gardner, J. Math. Phys. 10, 536 (1969).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 10 OCTOBER 1971

Moments of Solutions of a Class of Stochastic Differential Equations

J. McKenna and J. A. Morrison
Bell Telephone Laboratories, Inc. Murray Hill, New Jersey 07974
(Received 19 February 1971)

This paper is concerned with the exact calculation of moments of solutions of the stochastic, ordinary differential equation

$$\frac{d^2u}{dz^2} + \beta_0^2[1 + \eta f(M(z))]u = 0,$$

where $M(z)$ is an arbitrary, finite state space Markov process and $f(\cdot)$ is an arbitrary, real, single-valued function of its argument. The process $f(M(z))$ is in general not a Markov process. The calculation of the moments is reduced to the solution of a system of ordinary, first-order differential equations. In the special case where the process $M(z)$ has a stationary transition mechanism, these systems of equations have constant coefficients, so that the moments must consist of sums of exponentials. A specific example of such an equation with $M(z)$ stationary is analyzed in some detail. The results obtained provide an important, nontrivial check of some useful approximate techniques.

1. INTRODUCTION

The subject of this paper is the stochastic, ordinary differential equation

$$\frac{d^2u}{dz^2} + \beta_0^2[1 + \eta N(z)]u = 0, \tag{1.1}$$

where β_0 and η are positive constants and $N(z)$ is a stochastic process. Each sample function $N(z)$ defines on $0 \leq z < \infty$ two new functions $u_m(z)$, $m = 1, 2$, which are the linearly independent solutions of (1.1) satisfying the nonstochastic initial conditions

$$u_1(0) = u_2'(0) = 1, \quad u_1'(0) = u_2(0) = 0. \tag{1.2}$$

Throughout this paper we use the notation

$$v_m(z) = u_m'(z) = \frac{du_m(z)}{dz}, \quad m = 1, 2. \tag{1.3}$$

Thus the ensemble of functions $N(z)$ defines, via Eq. (1.1) and initial conditions (1.2), four new "solution" stochastic processes or ensembles of functions, $\{u_m(z)\}$ and $\{v_m(z)\}$, $m = 1, 2$. Typically equations of this type can arise in the propagation of electromagnetic or acoustic waves through randomly stratified media and in the study of dielectric waveguides with randomly stratified dielectric constants.

In most such problems, $\eta N(z)$ represents a small

(3.2) and (3.3) reduce to

$$\omega/k = P(u) - \delta^2 k^2 + 3\delta^2 A_{xx}/A, \tag{3.4}$$

$$A_t + [P(u) - 3\delta^2 k^2]A_x + \delta^2 A_{xxx} = 0. \tag{3.5}$$

If $A(x, t)$ is a slowly varying function of x, we may assume that

$$A_{xx}/A,\ A_{xxx}/A_x \ll k^2.$$

Thus, from (3.4) and (3.5) we have

$$\omega/k \approx P(u) - 4\pi^2(\delta/\lambda)^2$$

and

$$A_t/A_x \approx 12\pi^2(\delta/\lambda)^2 - P(u).$$

It indicates that the wave with wavelengths of order δ propagates in the negative direction; the wave grows if the amplitude is an increasing function of x.[4]

Next, we consider Burgers' equation. For this case, (3.2) and (3.3) reduce to

$$\omega/k = P(u) - 2\mu A_x/A, \tag{3.6}$$

$$A_t + P(u)A_x + \mu(k^2 A - A_{xx}) = 0. \tag{3.7}$$

Assuming that $A(x, t)$ can be expressed in the form of $\exp(\alpha t + \beta x)$, from (3.6) and (3.7) we obtain

$$2\mu\beta = P(u) - [\{P(u)\}^2 + 4\mu(\mu k^2 + \alpha)]^{1/2}$$

and

$$\omega/k = [\{P(u)\}^2 + 4\mu(\mu k^2 + \alpha)]^{1/2}.$$

Consider the wavelength $\lambda \leq O(\mu)$, α and $P(u) \leq O(1)$, then $\beta = -O(1/\lambda)$. It indicates that the wave will be damped out within a distance of the order of a wavelength. This explains why Burgers' equation has no solitary wave as in the case of Korteweg-de Vries equation.

1 D. J. Korteweg and G. de Vries, Phil. Mag. 39, 422 (1895).
2 J. M. Burgers, Proc. Acad. Sci. Amsterdam 43, 2 (1940).
3 E. Ott and R. N. Sudan, Phys. Fluids 13, 1432 (1970).
4 N. J. Zabusky and M. D. Kruskal, Phys. Rev. Letters 15, 240 (1965).
5 A. Sjöberg, J. Math. Anal. Appl. 29, 569 (1970).
6 R. M. Miura, J. Math. Phys. 9, 1202 (1968).
7 R. M. Miura, C. S. Gardner, and M. D. Kruskal, J. Math. Phys. 9, 1204 (1968).
8 C. H. Su and C. S. Gardner, J. Math. Phys. 10, 536 (1969).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 10 OCTOBER 1971

Moments of Solutions of a Class of Stochastic Differential Equations

J. McKenna and J. A. Morrison
Bell Telephone Laboratories, Inc. Murray Hill, New Jersey 07974
(Received 19 February 1971)

This paper is concerned with the exact calculation of moments of solutions of the stochastic, ordinary differential equation

$$\frac{d^2u}{dz^2} + \beta_0^2[1 + \eta f(M(z))]u = 0,$$

where $M(z)$ is an arbitrary, finite state space Markov process and $f(\cdot)$ is an arbitrary, real, single-valued function of its argument. The process $f(M(z))$ is in general not a Markov process. The calculation of the moments is reduced to the solution of a system of ordinary, first-order differential equations. In the special case where the process $M(z)$ has a stationary transition mechanism, these systems of equations have constant coefficients, so that the moments must consist of sums of exponentials. A specific example of such an equation with $M(z)$ stationary is analyzed in some detail. The results obtained provide an important, nontrivial check of some useful approximate techniques.

1. INTRODUCTION

The subject of this paper is the stochastic, ordinary differential equation

$$\frac{d^2u}{dz^2} + \beta_0^2[1 + \eta N(z)]u = 0, \tag{1.1}$$

where β_0 and η are positive constants and $N(z)$ is a stochastic process. Each sample function $N(z)$ defines on $0 \leq z < \infty$ two new functions $u_m(z)$, $m = 1, 2$, which are the linearly independent solutions of (1.1) satisfying the nonstochastic initial conditions

$$u_1(0) = u_2'(0) = 1, \quad u_1'(0) = u_2(0) = 0. \tag{1.2}$$

Throughout this paper we use the notation

$$v_m(z) = u_m'(z) = \frac{du_m(z)}{dz}, \quad m = 1, 2. \tag{1.3}$$

Thus the ensemble of functions $N(z)$ defines, via Eq. (1.1) and initial conditions (1.2), four new "solution" stochastic processes or ensembles of functions, $\{u_m(z)\}$ and $\{v_m(z)\}$, $m = 1, 2$. Typically equations of this type can arise in the propagation of electromagnetic or acoustic waves through randomly stratified media and in the study of dielectric waveguides with randomly stratified dielectric constants.

In most such problems, $\eta N(z)$ represents a small

random perturbation, which suggests the use of perturbation techniques. However, the regions of validity of most of the perturbation techniques in current use are unknown, so that exactly soluble examples, in addition to their intrinsic physical interest, provide extremely valuable checks. This paper is devoted to showing that for an interesting class of processes $N(z)$, to be described below, the exact calculation of the various stochastic moments of $u_m(z)$ and $v_m(z)$ can be reduced to the solution of a system of nonstochastic, ordinary, linear differential equations.

In a recent paper[1] we showed that if $N(z)$ is the "random telegraph" process $T(z)$, then a phase space method can be used to calculate exactly the various moments of the solution processes of (1.1). This random telegraph process is a zero-mean, wide sense stationary Markov process. Specifically, it is an ensemble of square wavefunctions $\{T(z)\}$ such that each sample function of the ensemble, $T(z)$, can assume only the values ± 1. For fixed z, a sample function chosen at random will equal 1 with probability $\frac{1}{2}$ or -1 with probability $\frac{1}{2}$. The probability $p(n,z)$ of a given sample function changing sign n times in an interval of length z is given by the Poisson process

$$p(n,z) = [(bz)^n/n!]e^{-bz}, \quad n = 0, 1, 2, \cdots, \tag{1.4}$$

where b is the average number of changes per unit length. In addition,

$$\langle T(z) \rangle = 0, \quad \langle T(x)T(y) \rangle = \exp(-2b|x-y|), \tag{1.5}$$

where, here and in all that follows, $\langle\ \rangle$ denotes the stochastic average.

In this paper, we show that the phase space method can be generalized to calculate various moments of the solution processes for any $N(z)$ of the form

$$N(z) = f(M(z)), \tag{1.6}$$

where $M(z)$ is any continuous parameter, finite state space Markov chain[2] and f is any single valued function mapping the state space into the real numbers. The random telegraph process is perhaps the simplest example of a continuous parameter, finite state space Markov chain. The most general such process $M(z)$ is in general not stationary or wide-sense stationary, and, as is well known, for arbitrary $f(\cdot)$, $f(M(z))$ is in general not a Markov process.[3]

As will become clear later, the whole technique can be extended to more general systems of linear, stochastic differential equations, but we will not pursue these generalizations here. However, the damped harmonic oscillator equation of the form

$$\frac{d^2w}{dz^2} + 2a\frac{dw}{dz} + \gamma^2[1 + \epsilon N(z)]w = 0, \tag{1.7}$$

can be reduced by the transformation

$$w(z) = e^{-az}u(z) \tag{1.8}$$

to an equation of the form (1.1).

Ideally, one would like to determine all the possible distribution functions of the four solution processes. So far, however, this has been an unobtainable goal, and most research has centered on the problem of calculating the stochastic averages of various functions of the solution processes. There are very few exact solutions of even this limited problem, but a number of perturbation techniques for obtaining approximate solutions have been developed. In fact, while Eq. (1.1) is of considerable interest in itself, it is also an important example on which these perturbation techniques can be tested.

Most of the approximate techniques developed so far are either strictly formal or else the known extent of their validity is quite restricted. Thus Khas'minskii[4] has recently established a limit theorem, valid for $\eta \to 0$, which can be applied[5] to (1.1) to calculate averages of the form $\langle g(u_1, v_1, u_2, v_2) \rangle$ for a suitable class of functions g in the case where $N(z)$ is a bounded stochastic process.

Another important perturbation technique for calculating moments and correlation functions of solutions of stochastic equations, the so-called smoothing method,[6] has recently been developed by Bourret[7] and Keller[8] and applied to a number of problems by various authors.[9] No error estimates are available yet for this method.

We might mention a third perturbation scheme recently developed by Papanicolaou and Keller,[10] which is essentially an application of two variable perturbation procedures. At this time no error estimates are available for this scheme either. It has been shown,[5] however, that the smoothing technique, the Khas'minskii method, and the two variable procedure yield the same expressions to order $O(\eta^2)$ for the first- and second-order moments and correlation functions of the solutions of (1.1).

The preceeding remarks suggest the importance of having exactly soluble equations on which the various approximation techniques can be tested. Bourret[11] has shown that if $N(z)$ in (1.1) is a single random telegraph process, the smoothing method yields the exact expressions for the first-order moments. McKenna and Morrison[1] have calculated exactly the second-order moments and correlation functions in this case by the phase space method and have shown that the smoothing method, when properly applied, again yields the exact answer both for the moments[12] and correlation functions.[13]

In Sec. 2 of this paper we outline those properties of continuous parameter, finite state space Markov chains which we need.

In Sec. 3 we define the phase space density functions, and discuss their properties and the system

of partial differential equations they satisfy. In addition, the equations from which the first- and second-order moments can be derived are discussed.

In Sec. 4 we consider in some detail an application of the general theory to the case, where $N(z)$ is a linear combination of two independent random telegraph processes. The results of this example provide a nontrivial check of the validity of the smoothing method. The example where $N(z)$ is the sum of n identical, stochastically independent random telegraph processes is also discussed briefly.

There are also two appendices in which some of the details of the derivations and calculations are presented.

2. FINITE STATE MARKOV CHAINS

In this section we present a brief discussion of those parts of the theory of continuous parameter, finite state space Markov chains (finite state Markov chains or FSMC for short) which we will need in the sequel.[2] For more detailed information we refer the reader to the references listed. In all that follows subscripts can take on integer values only.

In such a Markov chain the sample functions $M(z)$ are defined on the half line $0 \le z < \infty$ and can take on only a finite number N of distinct values a_j, $1 \le j \le N$, where the a_j are points in some abstract space E. The collection of points a_j, denoted by E_N, is called the state space. An initial probability distribution vector is given,

$$\boldsymbol{\alpha} = (\alpha_1, \cdots, \alpha_n), \tag{2.1}$$

where

$$\alpha_j = \text{Prob}\{M(0) = a_j\}, \qquad 1 \le j \le N, \tag{2.2}$$

with $\alpha_j > 0$, and

$$\sum_{j=1}^{N} \alpha_j = 1. \tag{2.3}$$

The definition is completed by specifying the transition probabilities

$$P_{ij}(x,y) = \text{Prob}\{M(y) = a_j | M(x) = a_i\}, \quad y \ge x, \quad 1 \le i,j \le N. \tag{2.4}$$

We are going to consider only those processes which can be defined by means of continuous, bounded infinitesimal generators. Specifically, we assume, given an $N \times N$ matrix function,

$$\tau(z) = (\tau_{ij}(z)), \tag{2.5}$$

with the following properties holding for $0 \le z < \infty$ and $1 \le i,j \le N$:

(1) $\tau_{ij}(z) \ge 0, \ i \ne j, \ \tau_{ii}(z) \le 0.$ (2.6a)

(2) The $\tau_{ij}(z)$ are continuous, and there is a constant C such that $|\tau_{ij}(z)| \le C.$ (2.6b)

(3) $\sum_{j=1}^{N} \tau_{ij}(z) = 0.$ (2.6c)

Feller[14] has shown that given a matrix $\tau(z)$ with the properties (2.6a)–(2.6c), the $N \times N$ conditional probability matrix

$$\mathbf{P}(x,y) = (P_{ij}(x,y)) \tag{2.7}$$

satisfying the initial condition

$$\mathbf{P}(x,x) = \mathbf{I}_N, \tag{2.8}$$

where $\mathbf{I}_N$ is the $N \times N$ unit matrix, is the unique solution of the backward and forward Kolmogorov equations

$$\frac{\partial \mathbf{P}(x,y)}{\partial x} = -\tau(x)\mathbf{P}(x,y), \tag{2.9a}$$

$$\frac{\partial \mathbf{P}(x,y)}{\partial y} = \mathbf{P}(x,y)\tau(y). \tag{2.9b}$$

The random telegraph process described in the Introduction is perhaps the simplest example of an FSMC. In this case $N = 2$, $a_1 = 1$, $a_2 = -1$, and

$$\boldsymbol{\alpha} = (\tfrac{1}{2}, \tfrac{1}{2}), \qquad \tau = \begin{pmatrix} -b & b \\ b & -b \end{pmatrix},$$

$$\mathbf{P}(x, x+z) \equiv \mathbf{P}(z) = e^{-bz}\begin{pmatrix} \cosh bz & \sinh bz \\ \sinh bz & \cosh bz \end{pmatrix}. \tag{2.10}$$

Since τ is a constant and $\alpha\tau = 0$, the process is stationary. More complicated realizations of FSMC's will be considered in later sections.

3. THE PHASE SPACE AND MOMENT EQUATIONS

We return now to (1.1) and the solution stochastic processes $u_m(z)$ and $v_m(z)$, $m = 1, 2$, defined by (1.1)–(1.3), where we assume that $N(z)$ is an arbitrary function of an FSMC $M(z)$ as defined by (1.6). Clearly these four processes can be thought of as the components of the vector solution

$$\mathbf{w}(z) = (u_1(z), v_1(z), u_2(z), v_2(z))^t \tag{3.1}$$

(where t denotes transpose) of the first-order vector differential equation

$$\frac{d\mathbf{w}(z)}{dz} = \mathbf{B}(z)\mathbf{w}(z), \tag{3.2}$$

where

$$\mathbf{B}(z) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -\beta_0^2[1+\eta N(z)] & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -\beta_0^2[1+\eta N(z)] & 0 \end{bmatrix}, \tag{3.3}$$

and

$$\mathbf{w}(0) = (1, 0, 0, 1)^t. \tag{3.4}$$

We remark that the combined process $(\mathbf{w}(z), M(z))$ is a Markov process since a knowledge of $\mathbf{w}(z_0)$ and $M(z_0)$ determines the process for all $z \geq z_0$ via the differential equation and the fact that $M(z)$ is a Markov process. It is, however, also of interest to note that $N(z) = f(M(z))$ may or may not be a Markov process, depending on the function f. If f maps the N points of E_N onto N distinct points of the real line, $f(M(z))$ is also a Markov process. However, if $f(E_N)$ consists of less than N distinct points, $f(M(z))$ is in general not a Markov process. In general the problem of determining whether or not $f(M(z))$ is an FSMC is difficult. However, in the important special case where $M(z)$ has a stationary transition mechanism, a theorem due to Burke and Rosenblatt[15] provides a straightforward solution to the problem.

We now consider the phase space density functions. In all that follows we denote by $\mathbf{w}$ any vector in R_4 with components

$$\mathbf{w} = (u_1, v_1, u_2, v_2)^t. \tag{3.5}$$

The Euclidean norm is denoted by $\|\mathbf{w}\|$, $\mathbf{w} \geq \mathbf{v}$ means the inequality is satisfied component by component, $|\mathbf{w}| = (|u_1|, |v_1|, |u_2|, |v_2|)^t$,

$$\mathbf{e} = (1, 1, 1, 1)^t, \tag{3.6}$$

and $\mathbf{I}_N$ will denote the $N \times N$ unit matrix.

The phase space density functions $\sigma_j(\mathbf{w}, z)$ are defined, $1 \leq j \leq N$, by the relations

$$\sigma_j(\mathbf{w}, z)d^4\mathbf{w} = \text{Prob}\{\mathbf{w} \leq \mathbf{w}(z) \leq \mathbf{w} + d\mathbf{w},\ M(z) = a_j\}. \tag{3.7}$$

Since

$$\sum_{j=1}^{N} \sigma_j(\mathbf{w}, z)d^4\mathbf{w} = \text{Prob}\{\mathbf{w} \leq \mathbf{w}(z) < \mathbf{w} + d\mathbf{w}\}, \tag{3.8}$$

it follows that if we knew the $\sigma_j(\mathbf{w}, z)$, we could calculate the stochastic average of any appropriately smooth function of the solution process $g(\mathbf{w}(z), z)$ from the formula

$$\Big\langle g(\mathbf{w}(z), z)\Big\rangle = \sum_{j=1}^{N} \int_{R_4} g(\mathbf{w}, z)\sigma_j(\mathbf{w}, z)d^4\mathbf{w}. \tag{3.9}$$

In general, the $\sigma_j(\mathbf{w}, z)$ are generalized functions.

Since $N(z)$ can assume only a finite number of values, it follows that $\|\mathbf{B}(z)\|$ in (3.3) is bounded by some constant C for all z and all sample functions $M(z)$. Consequently, it is easy to show that every sample function satisfying (3.2) and initial conditions (3.4) has the bound

$$|\mathbf{w}(z)| \leq \sqrt{2}\exp(Cz)\mathbf{e}. \tag{3.10}$$

An easy consequence of this is that, for fixed z, the support of $\sigma_j(\mathbf{w}, z)$ is contained in $|\mathbf{w}| \leq \sqrt{2}\exp(Cz)\mathbf{e}$. For fixed z this implies in particular that all the moments of the solution process exist.

The $\sigma_j(\mathbf{w}, z)$ are determined as the weak solutions of the system of N, linear, first-order partial differential equations

$$\frac{\partial \sigma_j}{\partial z} + D_j\sigma_j - \sum_{i=1}^{N} \sigma_i\tau_{ij}(z) = 0, \qquad 1 \leq j \leq N \tag{3.11}$$

satisfying the initial conditions

$$\sigma_j(\mathbf{w}, 0) = \alpha_j\delta(u_1 - 1)\delta(v_1)\delta(u_2)\delta(v_2 - 1), \quad 1 \leq j \leq N, \tag{3.12}$$

where

$$D_j = \sum_{i=1}^{2}\left(v_i\frac{\partial}{\partial u_i} - \beta_j^2 u_i\frac{\partial}{\partial v_i}\right), \tag{3.13}$$

and

$$\beta_j^2 = \beta_0^2[1 + \eta f(a_j)], \qquad 1 \leq j \leq N, \tag{3.14}$$

and the a_j are the N possible values which $M(z)$ can assume.

The derivation of these "forward" equations is very similar to the derivation of the equations in the special case where $N(z) = T(z)$, given in (1), and so we do not give a derivation here. We should remark that Wonham[16] gives the infinitesimal generator of the joint process $\{\mathbf{w}(z), M(z)\}$, which was derived earlier by Krasovskii and Lidskii.[17,18] The corresponding "backward" equation may be derived by means of this generator.

In Ref. 1 we considered instead of $\sigma_j(\mathbf{w}, z)$ the conditional phase space distribution functions $q_j(\mathbf{w}, z)$ related to the $\sigma_j(\mathbf{w}, z)$ by

$$\sigma_j(\mathbf{w}, z) = q_j(\mathbf{w}, z)R_j(z), \tag{3.15}$$

where

$$R_j(z) = \text{Prob}\{M(z) = a_j\}. \tag{3.16}$$

It follows from (2.4) and (2.9b) that

$$\frac{dR_j(z)}{dz} = \sum_{k=1}^{N} R_k(z)\tau_{kj}(z). \tag{3.17}$$

Hence, from (3.11) and (3.17), it follows that the q_j are the weak solutions of

$$\frac{\partial q_j}{\partial z} + D_j q_j - \sum_{i=1}^{N} \frac{(q_i - q_j)R_i(z)\tau_{ij}(z)}{R_j(z)} = 0, \tag{3.18}$$

which satisfy the initial conditions, from (2.2) and (3.12),

$$q_j(\mathbf{w}, 0) = \delta(u_1 - 1)\delta(v_1)\delta(u_2)\delta(v_2 - 1), \qquad 1 \leq j \leq N. \tag{3.19}$$

As we will show by examples in Sec. 4 sometimes symmetries in the problem reduce the number of distinct $\sigma_j(\mathbf{w}, z)$. Unfortunately, in no case have we been able to solve the systems of equations (3.11) or (3.18). The use of Eqs. (3.11) seems preferable for the applications we have in mind, for they do not require a knowledge of $\mathbf{R}(z)$.

Although we have been unable to solve Eqs. (3.11), because they are homogeneous in u and v we derive systems of ordinary differential equations for moments of the solution processes. Define the N partial moments, $1 \le j \le N$,

$$\left\langle u_1(z)^p v_1(z)^q u_2(z)^r v_2(z)^s \right\rangle_j = \int_{R_4} u_1^p v_1^q u_2^r v_2^s \sigma_j(\mathbf{w}, z) d^4w, \tag{3.20}$$

where p, q, r, s are nonnegative integers. The expressions on the right of (3.20) are well defined since the $\sigma_j(\mathbf{w}, z)$ have compact support in R_4 for fixed z. Then from (3.19)

$$\left\langle u_1(z)^p v_1(z)^q u_2(z)^r v_2(z)^s \right\rangle = \sum_{j=1}^{N} \left\langle u_1(z)^p v_1(z)^q u_2(z)^r v_2(z)^s \right\rangle_j. \tag{3.21}$$

If we multiply each of Eqs. (3.11) by all $\binom{M+3}{3}$ products of the form $u_1^p v_1^q u_2^r v_2^s$ with $p + q + r + s = M$ and then integrate over R_4 with respect to $d^4\mathbf{w}$, on performing several integrations by parts, we obtain a system of $N\binom{M+3}{3}$ ordinary, linear differential equations involving only the $N\binom{M+3}{3}$ Mth-order conditional moments. Unless τ is constant, it is generally impossible to solve the moment equations analytically. However, even in the general case the moment equations are in a form suitable for numerical solution.

We conclude this section by writing down the equations for the first- and second-order moments. They can be written most compactly in matrix form. Define the four column vectors

$$\mathbf{U}_m(z) = (\langle u_m(z)\rangle_j), \quad \mathbf{V}_m(z) = (\langle v_m(z)\rangle_j), \quad m = 1, 2, \quad 1 \le j \le N, \tag{3.22}$$

where $\langle u_m(z)\rangle_j$ and $\langle v_m(z)\rangle_j$ are defined in (3.20). Let

$$\mathbf{B} = \operatorname{diag}(\beta_j^2) = \beta_0^2 \mathbf{I}_N + \eta\beta_0^2 \operatorname{diag}[f(a_j)], \tag{3.23}$$

from (3.14). Then the partial first-order moments are the solutions of the equations

$$\frac{d}{dz}\begin{bmatrix} \mathbf{U}_m(z) \\ \mathbf{V}_m(z) \end{bmatrix} + \begin{bmatrix} -\tau^t & -\mathbf{I}_N \\ \mathbf{B} & -\tau^t \end{bmatrix}\begin{bmatrix} \mathbf{U}_m(z) \\ \mathbf{V}_m(z) \end{bmatrix} = 0, \quad m = 1, 2, \tag{3.24}$$

which, from (3.12), satisfy the initial conditions

$$\mathbf{U}_1(0) = \mathbf{V}_2(0) = \boldsymbol{\alpha}^t, \quad \mathbf{U}_2(0) = \mathbf{V}_1(0) = 0, \tag{3.25}$$

where $\boldsymbol{\alpha}$ is the vector of initial probabilities defined in (2.11). Notice that the equations for $m = 1$ and 2 are uncoupled. From (3.21) and (3.22)

$$\langle u_m(z)\rangle = \mathbf{E}_N \mathbf{U}_m(z), \quad \langle v_m(z)\rangle = \mathbf{E}_N \mathbf{V}_m(z), \tag{3.26}$$

where $\mathbf{E}_N = (E_j)$, $1 \le j \le N$, is the row vector with components

$$E_j = 1. \tag{3.27}$$

Similarly, we define the nine column vectors

$$\mathbf{X}_m(z) = (X_m(z)_j), \quad \mathbf{Y}_m(z) = (Y_m(z)_j), \quad \mathbf{Z}_m(z) = (Z_m(z)_j), \quad m = 0, 1, 2, \quad 1 \le j \le N,$$

by

$$X_0(z)_j = \langle u_1(z)u_2(z)\rangle_j, \quad X_m(z)_j = \langle u_m^2(z)\rangle_j, \quad m = 1, 2, \tag{3.28}$$

$$Y_0(z)_j = \tfrac{1}{2}\{\langle u_1(z)v_2(z)\rangle_j + \langle u_2(z)v_1(z)\rangle_j\}, \quad Y_m(z)_j = \langle u_m(z)v_m(z)\rangle_j, \quad m = 1, 2, \tag{3.29}$$

$$Z_0(z)_j = \langle v_1(z)v_2(z)\rangle_j, \quad Z_m(z)_j = \langle v_m^2(z)\rangle_j, \quad m = 1, 2. \tag{3.30}$$

Then $\mathbf{X}_m$, $\mathbf{Y}_m$ and $\mathbf{Z}_m$ are the solutions of

$$\frac{d}{dz}\begin{bmatrix} \mathbf{X}_m(z) \\ \mathbf{Y}_m(z) \\ \mathbf{Z}_m(z) \end{bmatrix} + \begin{bmatrix} -\tau^t & -2\mathbf{I}_N & 0 \\ \mathbf{B} & -\tau^t & -\mathbf{I}_N \\ 0 & 2\mathbf{B} & -\tau^t \end{bmatrix}\begin{bmatrix} \mathbf{X}_m(z) \\ \mathbf{Y}_m(z) \\ \mathbf{Z}_m(z) \end{bmatrix} = 0, \tag{3.31}$$

which satisfy the initial conditions

$$\mathbf{X}_0(0) = \mathbf{X}_2(0) = \mathbf{Y}_1(0) = \mathbf{Y}_2(0) = \mathbf{Z}_0(0) = \mathbf{Z}_1(0) = 0, \quad \mathbf{X}_1(0) = 2\mathbf{Y}_0(0) = \mathbf{Z}_2(0) = \boldsymbol{\alpha}^t. \tag{3.32}$$

Notice again that the equations for $m = 0, 1$, and 2 are uncoupled. Since for each sample function $N(z)$ it is a direct consequence of (1.1) that $u_1(z)v_2(z) - u_2(z)v_1(z) \equiv 1$, it follows that

$$\langle u_1(z)v_1(z)\rangle - \langle u_2(z)v_1(z)\rangle \equiv 1. \tag{3.33}$$

Consequently, from (3.21), (3.27), and (3.33) we have

$$\langle u_m^2(z)\rangle = \mathbf{E}_N \mathbf{X}_m(z), \quad \langle u_m(z)v_m(z)\rangle = \mathbf{E}_N \mathbf{Y}_m(z), \quad \langle v_m^2(z)\rangle = \mathbf{E}_N \mathbf{Z}_m(z), \quad m = 1, 2, \tag{3.34}$$

and

$$\langle u_1(z)u_2(z)\rangle = \mathbf{E}_N \mathbf{X}_0, \quad \langle v_1(z)v_2(z)\rangle = \mathbf{E}_N \mathbf{Z}_0(z), \tag{3.35}$$

$$\langle u_1(z)v_2(z)\rangle = \mathbf{E}_N \mathbf{Y}_0(z) + \tfrac{1}{2},$$

$$\langle u_2(z)v_1(z)\rangle = \mathbf{E}_N \mathbf{Y}_0(z) - \tfrac{1}{2}.$$

4. APPLICATIONS OF THE GENERAL THEORY

In this section we illustrate the general theory by some relatively simple but interesting examples.

A. A Linear Combination of Two Random Telegraph Processes

Let $T_1(z)$ and $T_2(z)$ be two stochastically independent random telegraph processes with rates b_1 and b_2, respectively. If c_1 and c_2 are real numbers satisfying $c_1^2 + c_2^2 = 1$, define

$$N(z) = c_1 T_1(z) + c_2 T_2(z). \tag{4.1}$$

This process is clearly stationary, but, as we will show, whether or not it is Markov depends on the relationship between the parameters b_1, b_2, c_1, and c_2.

In order to fit $N(z)$ defined in (4.1) into the framework of Sec. 2, define the function $f(\mathbf{x})$, which maps the space R_2 of real two vectors onto the real line by

$$f((x_1, x_2)) = c_1 x_1 + c_2 x_2. \tag{4.2}$$

Then, if $\mathbf{M}(z)$ is the vector-valued process

$$\mathbf{M}(z) = (T_1(z), T_2(z)), \tag{4.3}$$

clearly $N(z) = f(\mathbf{M}(z))$.

The process $\mathbf{M}(z)$ has the state space E_4 consisting of the four vectors

$$\begin{aligned} &\mathbf{a}_1 = (1,1), \quad \mathbf{a}_2 = (1,-1), \\ &\mathbf{a}_3 = (-1,1), \quad \mathbf{a}_4 = (-1,-1), \end{aligned} \tag{4.4}$$

and has the initial probability vector [(2.1)]

$$\boldsymbol{\alpha} = (\tfrac{1}{4}, \tfrac{1}{4}, \tfrac{1}{4}, \tfrac{1}{4}). \tag{4.5}$$

Since $T_1(z)$ and $T_2(z)$ are stochastically independent and Markovian, it is easy to see that $\mathbf{M}(z)$ is an FSMC and that, for $y \geq x$,

$$\begin{aligned} &\text{Prob}\{\mathbf{M}(y) = ((-1)^{i-1}, (-1)^{j-1}) \,|\, \mathbf{M}(x) \\ &\quad = ((-1)^{k-1}, (-1)^{l-1})\} \\ &\quad = \text{Prob}\{T_1(y) = (-1)^{i-1} \,|\, T_1(x) = (-1)^{k-1}\} \\ &\quad \times \text{Prob}\{T_2(y) = (-k)^{j-1} \,|\, T_2(x) = (-1)^{l-1}\}. \end{aligned} \tag{4.6}$$

From (4.4) and (4.6) it follows that if $\mathbf{P}_j(x,y) = \mathbf{P}_j(y-x)$, $j = 1, 2$, are the 2×2 conditional probability matrices of $T_1(z)$ and $T_2(z)$, given by (2.10) with $b = b_1$ and b_2, respectively, then $\mathbf{P}(x,y)$, the 4×4 conditional probability matrix of $\mathbf{M}(z)$, is

$$\mathbf{P}(x,y) = \mathbf{P}_1(y-x) \times \mathbf{P}_2(y-x). \tag{4.7}$$

The product on the right-hand side of (4.7) is the Kronecker product of the two matrices.[19] From (2.7)–(2.9) it easily follows that then

$$\mathbf{P}(\delta z) = \mathbf{I}_4 + \tau \delta z + o(\delta z). \tag{4.8}$$

Combining (2.10), (4.7), and (4.8), we conclude that

$$\begin{aligned} \tau &= \tau_1 \times \mathbf{I}_2 + \mathbf{I}_2 \times \tau_2 \\ &= \begin{bmatrix} -(b_1+b_2) & b_2 & b_1 & 0 \\ b_2 & -(b_1+b_2) & 0 & b_1 \\ b_1 & 0 & -(b_1+b_2) & b_2 \\ 0 & b_1 & b_2 & -(b_1+b_2) \end{bmatrix}, \end{aligned} \tag{4.9}$$

where τ_1 and τ_2 are the infinitesimal generators of $T_1(z)$ and $T_2(z)$, respectively. Since from (4.7) $\mathbf{M}(z)$ clearly has a stationary transition mechanism and from (4.5) and (4.9) $\boldsymbol{\alpha}\tau = 0$, $\mathbf{M}(z)$ is stationary.

It is a straightforward matter to apply the theorem of Burke and Rosenblatt[15] to determine whether or not $N(z)$ is Markovian. We merely state the results. If $c_1 \neq \pm c_2$, then $N(z)$ is Markovian for any choice of b_1 and b_2. However, if $c_1 = \pm c_2$, then $N(z)$ is Markovian if and only if $b_1 = b_2$.

In general for this problem, there are four distinct phase space density functions $\sigma_j(\mathbf{w}, z)$, $1 \leq j \leq 4$. However, $\sigma_2(\mathbf{w}, z) \equiv \sigma_3(\mathbf{w}, z)$ in the special case $c_1 = c_2$ and $b_1 = b_2$, while $\sigma_1(\mathbf{w}, z) \equiv \sigma_4(\mathbf{w}, z)$ if $c_1 = -c_2$ and $b_1 = b_2$.

Excluding for the moment these two special cases, Eqs. (3.24) for the partial first-order moments consist of two sets of eight, first-order, constant-coefficient, ordinary differential equations. From the solutions of these we can calculate the first-order moments from (3.26). These calculations are outlined in Appendix A; we merely discuss the results here.

All the first-order moments can be obtained from $\langle u_2(z) \rangle$ by the relations

$$\langle u_1(z) \rangle = \langle v_2(z) \rangle = \frac{d}{dz} \langle u_2(z) \rangle, \qquad \langle v_1(z) \rangle = \frac{d}{dz} \langle u_1(z) \rangle. \tag{4.10}$$

Furthermore,

$$\langle u_2(z) \rangle = \sum_{j=1}^{8} l_j(\eta) \exp[z\lambda_j(\eta)], \tag{4.11}$$

where the $l_j(\eta)$ are constants and the $\lambda_j(\eta)$ are the roots of the eighth-order polynomial $D(s;\eta)$ given in (A13). If $\eta \ll 1$, $D(s,\eta)$ has four pairs of complex conjugate roots whose leading terms are shown in (A14), and all of them have negative real parts. In this case, from (A14) and (A15),

$$\langle u_2(z) \rangle = \beta_0^{-1} \exp(-\eta^2 \kappa z) \sin[\beta_0(1 - \eta^2\rho)z] + O(\eta^2), \tag{4.12}$$

where

$$\kappa = \frac{\beta_0^4}{8}\left[\frac{c_1^2}{b_1(b_1^2 + \beta_0^2)} + \frac{c_2^2}{b_2(b_2^2 + \beta_0^2)}\right], \tag{4.13}$$

$$\rho = \frac{\beta_0^2}{8}\left[\frac{c_1^2}{b_1^2+\beta_0^2} + \frac{c_2^2}{b_2^2+\beta_0^2}\right]. \tag{4.14}$$

In the special cases where $c_1^2 = c_2^2 = \frac{1}{2}$ and $b_1 = b_2 = b$, Eqs. (4.10) still hold and $\langle u_2(z)\rangle$ is still given by an expression of the form (4.11), but the sum now contains only six terms. The $\lambda_j(\eta)$, $1 \le j \le 6$, are the roots of a sixth-order polynomial which is a factor of $D(s;\eta)$ when $c_1^2 = c_2^2 = \frac{1}{2}, b_1 = b_2 = b$. Furthermore, (4.12) which expresses $\langle u_2(z)\rangle$ correctly to order η^2 still holds.

Again excluding the above special case, Eqs. (3.32) for second-order partial moments consist of three sets of 12, first-order, constant-coefficient, ordinary differential equations. Their solution is outlined in Appendix B.

The relations

$$\langle u_2 v_2\rangle = \tfrac{1}{2}\frac{d}{dz}\langle u_2^2\rangle, \tag{4.15}$$

$$\langle u_1^2\rangle = \langle v_2^2\rangle, \quad \langle u_1 v_1\rangle = \tfrac{1}{2}\frac{d}{dz}\langle v_2^2\rangle, \tag{4.16}$$

$$\langle u_1 u_2\rangle = \tfrac{1}{2}\frac{d}{dz}\langle u_2^2\rangle, \quad \langle v_1 v_2\rangle = \tfrac{1}{2}\frac{d}{dz}\langle v_2^2\rangle, \tag{4.17}$$

$$\begin{aligned}\langle u_1 v_2\rangle &= \tfrac{1}{2}\left(\frac{d}{dz}\langle u_2 v_2\rangle + 1\right),\\ \langle v_1 u_2\rangle &= \tfrac{1}{2}\left(\frac{d}{dz}\langle u_2 v_2\rangle - 1\right)\end{aligned} \tag{4.18}$$

show that all the second-order moments can be obtained from $\langle v_1^2\rangle$, $\langle u_2^2\rangle$ and $\langle v_2^2\rangle$. But,

$$\langle v_1^2\rangle = \sum_{k=1}^{12} m_k(\eta)\exp[z s_k(\eta)], \tag{4.19}$$

$$\langle u_2^2\rangle = \sum_{k=1}^{12} n_k(\eta)\exp[z s_k(\eta)], \tag{4.20}$$

$$\langle v_2^2\rangle = \sum_{k=1}^{12} p_k(\eta)\exp[z s_k(\eta)], \tag{4.21}$$

where the $s_k(\eta)$ are the roots of the twelfth-order polynomial $\Delta(s;\eta)$ given in (B19) and the $m_k(\eta)$, $n_k(\eta)$, and $p_k(\eta)$ are constants. If $\eta \ll 1$, $\Delta(s;\eta)$ has one root with a positive real part and eleven roots with negative real parts. In this case, from (B22)–(B25),

$$\begin{aligned}\langle v_1^2\rangle = {}& \tfrac{1}{2}\beta_0^2\{\exp(\eta^2\mu z) - \exp(-\eta^2\nu z)\\ &\times \cos[2\beta_0(1-\eta^2\rho)z]\} + O(\eta^2),\end{aligned} \tag{4.22}$$

$$\langle u_2^2\rangle = \langle v_1^2\rangle/\beta_0^4 + O(\eta^2) \tag{4.23}$$

$$\begin{aligned}\langle v_2^2\rangle = {}& \tfrac{1}{2}\{\exp(\eta^2\mu z) + \exp(-\eta^2\nu z)\cos[2\beta_0(1-\eta^2\rho)z]\}\\ &+ O(\eta^2),\end{aligned} \tag{4.24}$$

where

$$\mu = \frac{\beta_0^2}{2}\left[c_1^2\left(\frac{b_1}{b_1^2+\beta_0^2}\right) + c_2^2\left(\frac{b_2}{b_2^2+\beta_0^2}\right)\right], \tag{4.25}$$

$$\nu = \frac{\beta_0^2}{4}\left[\frac{c_1^2(b_1^2+2\beta_0^2)}{b_1(b_1^2+\beta_0^2)} + \frac{c_2^2(b_2^2+2\beta_0^2)}{b_2(b_2^2+\beta_0^2)}\right], \tag{4.26}$$

and ρ is given in (4.14).

In the case where $c_1^2 = c_2^2 = \frac{1}{2}$ and $b_1 = b_2 = b$, Eqs. (4.15)–(4.18) still hold and $\langle v_1^2\rangle$, $\langle u_2^2\rangle$, and $\langle v_2^2\rangle$ are still given by expressions of the form (4.19)–(4.21), but the sums now contain only nine terms. The $s_k(\eta)$, $1 \le k \le 9$, are the roots of a ninth-order polynomial which is a factor of $\Delta(s;\eta)$ when $c_1^2 = c_2^2 = \frac{1}{2}$ and $b_1 = b_2 = b$. Furthermore, (4.22)–(4.24) are still valid.

It is of some interest to compare our expressions for the first- and second-order moments of the solutions of Example 1 with the approximate expressions for the same quantities obtained by using the smoothing technique. Morrison[13] has calculated the first- and second-order moments of solutions of (1.1) by the smoothing method for $\eta \ll 1$ for quite general $N(z)$. His results are expressed in terms of the Laplace transform of the correlation function of $N(z)$. From (4.1) and (1.5)

$$\begin{aligned}\Gamma(z) = \langle N(z)N(0)\rangle = {}& c_1^2\exp(-2b_1|z|)\\ &+ c_2^2\exp(-2b_2|z|),\end{aligned} \tag{4.27}$$

$$\gamma(s) = \int_0^\infty e^{-sz}\Gamma(z)dz = \frac{c_1^2}{s+2b_1} + \frac{c_2^2}{s+2b_2}. \tag{4.28}$$

When $\gamma(s)$ given in (4.28) is substituted into Eqs. (4.8)–(4.17) of Ref. 13, it can be seen that the small η approximations to the moments given by the smoothing method agree with the expressions in this paper given in (4.12)–(4.14) and (4.22)–(4.26).

It has been pointed out previously[11,12] that the smoothing method yields the exact expressions for the moments of the solutions of (1.1) when $N(z)$ is a single random telegraph wave. However, the smoothing method does not yield the exact results for the moments in the example considered in this paper. It can be shown, for example, that in the special case $b_1 = b_2 = b$ the smoothing method yields the first-order moments as the sum of four exponentials and the second-order moments as the sum of seven exponentials. The exponents of the terms with coefficients which are $O(1)$ are given correctly through order η^2, but none of the remaining exponents are given correctly through order η^2. We can conclude that Example A of this paper provides an important, nontrivial check on the accuracy of the smoothing method.

B. The Sum of n Identical, Stochastically Independent, Random Telegraph Processes

Let $T_j(z)$, $1 \le j \le n$, be n identical, stochastically independent random telegraph processes with rate b, and define

$$N(z) = \sum_{j=1}^{n} T_j(z). \tag{4.29}$$

We will only set up the equations for the $\sigma_i(\mathbf{w}, z)$ to demonstrate the importance of making use of available symmetries. As in Example A, define the vector-valued process

$$\mathbf{M}(z) = (T_1(z), T_2(z), \cdots, T_n(z)). \tag{4.30}$$

This process has the state space E_{2n} consisting of the 2^n vectors of the form $(\pm 1, \pm 1, \cdots, \pm 1)$, and has the initial probability vector $\boldsymbol{\alpha} = (\alpha_i)$,

$$\alpha_i = 2^{-n}, \quad 1 \le i \le 2^n. \tag{4.31}$$

Since the $T_j(z)$ are stochastically independent, it is easy to see that $\mathbf{M}(z)$ is an FSMC. It is not hard to see that if the elements of the state space are labeled properly, then $\mathbf{P}_{2n}(x,y)$, the $2^n \times 2^n$ conditional probability matrix of $\mathbf{M}(z)$, is

$$\mathbf{P}_{2n}(x,y) = \mathbf{P}^{(n)}(x,y), \tag{4.32}$$

where $\mathbf{P}^{(n)}(x,y)$ is the n-fold Kronecker product with itself of the 2×2 conditional probability matrix of the random telegraph process $\mathbf{P}(x,y)$ given in (2.10). Then it follows just as in Example A that the infinitesimal generator of $\mathbf{M}(z)$ is

$$\tau_{2^n} = \tau \times \mathbf{I}_2^{(n-1)} + \sum_{j=1}^{n-2} \left(\mathbf{I}_2^{(j)} \times \tau \times \mathbf{I}_2^{(n-j-1)} \right) + \mathbf{I}_2^{(n-1)} \times \tau, \tag{4.33}$$

where τ is the infinitesimal generator of the random telegraph process given in (2.10) and $\mathbf{I}_2^{(k)}$ is the k-fold Kronecker product of the 2×2 unit matrix with itself. Thus, on substituting the matrix elements of τ_{2^n}, from (4.33) into (3.14), 2^n equations for the 2^n quantities $\sigma_j(\mathbf{w}, z)$ result.

However, with the aid of $\mathbf{P}_{2^n}(x,y)$ given in (4.32), the theorem of Burke and Rosenblatt[15] can be applied to show that $N(z) = f(\mathbf{M}(z))$ is also a Markov process. The state space of $N(z)$ consists of the $n+1$ points on the real line, $n - 2r + 2$, $r = 1, 2, \cdots, n+1$, and

$$\text{Prob}\{N(z) = n - 2r + 2\} = 2^{-n}\binom{n}{r-1}. \tag{4.34}$$

It can be shown that the infinitesimal generator of $N(z)$ is a tridiagonal matrix whose nonzero elements are

$$\tau_{jj} = -nb, \quad 1 \le j \le n+1, \tag{4.35a}$$

$$\tau_{j-1,j} = (n - j + 2)b, \quad 2 \le j \le n+1, \tag{4.35b}$$

$$\tau_{j+1,j} = jb, \quad 1 \le j \le n. \tag{4.35c}$$

Thus, of the original 2^n quantities $\sigma_j(\mathbf{w}, z)$, only $n+1$ are distinct from each other, and, substituting expression (4.35) for τ into (3.14), we obtain the system of $n+1$ equations from which they can be determined. Thus, in this case, making use of the available symmetries greatly reduces the difficulty of the problem for large n.

ACKNOWLEDGMENTS

We gratefully acknowledge many helpful discussions with our colleagues T. T. Kadota, S. P. Lloyd, and L. A. Shepp.

APPENDIX A:

In this appendix we outline the calculation of the first order moments for Example A of Sec. 4. To do this, we must solve Eqs. (3.24) with initial conditions (3.25), where $\boldsymbol{\alpha}$ is given in (4.5), $\tau = \tau^t$ is given in (4.9), and, from (3.23), (4.2), and (4.4), $\mathbf{B}$ is

$$\mathbf{B} = \beta_0^2 \mathbf{I}_4 + \eta\beta_0^2 \text{diag}(c_1 + c_2, c_1 - c_2, -c_1 + c_2, -c_1 - c_2). \tag{A1}$$

We assume initially that $c_1 \neq \pm c_2$.

If $\mathbf{w}(z)$ is a 4-vector function, its Laplace transform is

$$\hat{\mathbf{w}}(s) = \mathfrak{L}(\mathbf{w}(z)) = \int_0^\infty e^{-sz}\mathbf{w}(z)dz. \tag{A2}$$

Taking the Laplace transform of Eqs. (3.24) and making use of (3.25), we obtain

$$\mathbf{S}\hat{\mathbf{U}}_1 - \hat{\mathbf{V}}_1 = \boldsymbol{\alpha}^t, \quad \mathbf{B}\hat{\mathbf{U}}_1 + \mathbf{S}\hat{\mathbf{V}}_1 = 0, \tag{A3}$$

$$\mathbf{S}\hat{\mathbf{U}}_2 - \hat{\mathbf{V}}_2 = 0, \quad \mathbf{B}\hat{\mathbf{U}}_2 + \mathbf{S}\hat{\mathbf{V}}_2 = \boldsymbol{\alpha}^t, \tag{A4}$$

where

$$\mathbf{S} = s\mathbf{I}_4 - \boldsymbol{\tau}^t. \tag{A5}$$

These pairs of matrix equations are solved readily:

$$\hat{\mathbf{U}}_1(s) = (\mathbf{B} + \mathbf{S}^2)^{-1}\mathbf{S}\boldsymbol{\alpha}^t,$$
$$\hat{\mathbf{V}}_1(s) = \mathbf{S}(\mathbf{B} + \mathbf{S}^2)^{-1}\mathbf{S}\boldsymbol{\alpha}^t - \boldsymbol{\alpha}^t, \tag{A6}$$

$$\hat{\mathbf{U}}_2(s) = (\mathbf{B} + \mathbf{S}^2)^{-1}\boldsymbol{\alpha}^t, \quad \hat{\mathbf{V}}_2(s) = \mathbf{S}(\mathbf{B} + \mathbf{S}^2)^{-1}\boldsymbol{\alpha}^t. \tag{A7}$$

From (3.26) we see that

$$\mathfrak{L}(\langle u_m(z)\rangle) = \mathbf{E}_4\hat{\mathbf{U}}_m(s), \quad \mathfrak{L}(\langle v_m(z)\rangle) = \mathbf{E}_4\hat{\mathbf{V}}_m(s), \quad m = 1, 2. \tag{A8}$$

Since in this case $\mathbf{E}_4\boldsymbol{\tau} = \boldsymbol{\tau}\boldsymbol{\alpha}^t = 0$, it follows from (A5) that

$$\mathbf{E}_4\mathbf{S} = s\mathbf{E}_4, \quad \mathbf{S}\boldsymbol{\alpha}^t = s\boldsymbol{\alpha}^t, \tag{A9}$$

and consequently from (A6)–(A8) that

$$\mathfrak{L}(\langle u_1\rangle) = \mathfrak{L}(\langle v_2\rangle) = s\mathfrak{L}(\langle u_2\rangle),$$
$$\mathfrak{L}(\langle v_1\rangle) = s\mathfrak{L}(\langle u_1\rangle) - 1. \tag{A10}$$

Equations (A10) and the initial conditions imply Eqs. (4.10).

To determine $\langle u_2(z)\rangle$, from which the remaining first moments can be obtained by differentiation, the inverse Laplace transform of $\mathbf{E}_4\hat{\mathbf{U}}_2(s)$ must be found. However, $\mathbf{E}_4\hat{\mathbf{U}}_2(s)$ is a rational function of s with the denominator

$$D(s;\eta) = \det(\mathbf{B} + \mathbf{S}^2), \tag{A11}$$

which is an eighth-order polynomial in s. If $\lambda_j(\eta)$, $1 \le j \le 8$, are the eight roots of $D(s;\eta) = 0$, then

$$\langle u_2(z)\rangle = \sum_{j=1}^{8} l_j(\eta)\exp[\lambda_j(\eta)z], \tag{A12}$$

where $l_j(\eta)$ is the residue of $\mathbf{E}_4\hat{\mathbf{U}}_2(s)$ at $s = \lambda_j(\eta)$. After some algebra it can be shown that

$$\begin{aligned} D(s;\eta) &= (s^2 + \beta_0^2)[(s + 2b_1)^2 + \beta_0^2][(s + 2b_2)^2 + \beta_0^2] \\ &\times [(s + 2b_1 + 2b_2)^2 + \beta_0^2] - 2\eta^2\beta_0^4((c_1^2 + c_2^2) \\ &\times \{[(s + b_1 + b_2)^2 + \beta_0^2 + b_1^2 + b_2^2]^2 - 4b_1^2b_2^2\} \\ &- 4(c_1^2 - c_2^2)(b_1^2 - b_2^2)(s + b_1 + b_2)^2) \\ &+ \eta^4\beta_0^8(c_1^2 - c_2^2)^2. \end{aligned} \tag{A13}$$

If $\eta \ll 1$, the roots of $D(s;\eta)$ can be shown to consist of four complex conjugate pairs:

$$\begin{aligned} \lambda_1 = \lambda_2^* &= i\beta_0\left[1 - \frac{\beta_0^2\eta^2}{8}\left(\frac{c_1^2}{b_1(b_1 + i\beta_0)}\right.\right. \\ &\left.\left. + \frac{c_2^2}{b_2(b_2 + i\beta_0)}\right)\right] + O(\eta^4), \\ \lambda_3 = \lambda_4^* &= -2b_1 - 2b_2 + i\beta_0 + O(\eta^2), \\ \lambda_5 = \lambda_6^* &= -2b_1 + i\beta_0 + O(\eta^2), \\ \lambda_7 = \lambda_8^* &= -2b_2 + i\beta_0 + O(\eta^2). \end{aligned} \tag{A14}$$

The corresponding coefficients $l_j(\eta)$ in (A12) can be shown to have the form

$$\begin{aligned} l_1(\eta) &= l_2^*(\eta) = 1/(2i\beta_0) + O(\eta^2), \\ l_j(\eta) &= l_{j+1}^*(\eta) = O(\eta^2), \quad j = 3, 5, 7. \end{aligned} \tag{A15}$$

In the special cases $c_1 = \pm c_2$ and $b_1 = b_2 = b$, it is easily seen that $D(s;\eta)$ has the factor $(s + 2b)^2 + \beta_0^2$. It can be shown most simply in this case by making use from the start of the symmetry relation $\sigma_2(\mathbf{w}, z) \equiv \sigma_3(\mathbf{w}, z)$ or $\sigma_1(\mathbf{w}, z) = \sigma_4(\mathbf{w}, z)$ that the roots corresponding to the factor $(s + 2b)^2 + \beta_0^2$ do not appear. The expression for $\langle u_2(z)\rangle$ is then of the form (A12) but with only six terms appearing. The quantities $\lambda_j(\eta)$ and $l_j(\eta)$ are still given for $1 \le j \le 6$ by (A14) and (A15) with $c_1^2 = c_2^2 = \frac{1}{2}$ and $b_1 = b_2 = b$. Equations (A10) are independent of the specific form of $D(s;\eta)$ and are still valid in this special case.

APPENDIX B:

In this appendix we outline the calculation of the second-order moments for Example A of Sec. 4. We must solve (3.31) with initial conditions (3.32), and with $\boldsymbol{\alpha}, \tau$, and $\mathbf{B}$ as in Appendix A. Again assume initially that $c_1 \neq \pm c_2$.

Taking Laplace transforms of (3.31) and making use of (3.32), we obtain the three systems of equations, $m = 0, 1, 2$,

$$\mathbf{S}\hat{\mathbf{X}}_m - 2\hat{\mathbf{Y}}_m = \delta_{m,1}\boldsymbol{\alpha}^t, \tag{B1}$$

$$\mathbf{B}\hat{\mathbf{X}}_m + \mathbf{S}\hat{\mathbf{Y}}_m - \hat{\mathbf{Z}}_m = \tfrac{1}{2}\delta_{m,0}\boldsymbol{\alpha}^t, \tag{B2}$$

$$2\mathbf{B}\hat{\mathbf{Y}}_m + \mathbf{S}\hat{\mathbf{Z}}_m = \delta_{m,2}\boldsymbol{\alpha}^t, \tag{B3}$$

where δ_{jk} is the Kronecker delta, $\mathbf{S}$ is defined in (A5), $\hat{\mathbf{X}}_m = \mathcal{L}(\mathbf{X}_m)$, $\hat{\mathbf{Y}}_m = \mathcal{L}(\mathbf{Y}_m)$, and $\hat{\mathbf{Z}}_m = \mathcal{L}(\mathbf{Z}_m)$. These equations have the formal solutions

$$\begin{aligned} \hat{\mathbf{X}}_0 &= \tfrac{1}{2}\mathbf{J}\mathbf{S}\boldsymbol{\alpha}^t, \quad \hat{\mathbf{Y}}_0 = \tfrac{1}{2}\mathbf{S}\hat{\mathbf{X}}_0, \\ \hat{\mathbf{Z}}_0 &= (\mathbf{B} + \tfrac{1}{2}\mathbf{S}^2)\hat{\mathbf{X}}_0 - \tfrac{1}{2}\boldsymbol{\alpha}^t, \end{aligned} \tag{B4}$$

$$\begin{aligned} \hat{\mathbf{X}}_1 &= \mathbf{J}(\mathbf{B} + \tfrac{1}{2}\mathbf{S}^2)\boldsymbol{\alpha}^t, \quad \hat{\mathbf{Y}}_1 = \tfrac{1}{2}(\mathbf{S}\hat{\mathbf{X}}_1 - \boldsymbol{\alpha}^t), \\ \hat{\mathbf{Z}}_1 &= (\mathbf{B} + \tfrac{1}{2}\mathbf{S}^2)\hat{\mathbf{X}}_1 - \tfrac{1}{2}\mathbf{S}\boldsymbol{\alpha}^t, \end{aligned} \tag{B5}$$

$$\hat{\mathbf{X}}_2 = \mathbf{J}\boldsymbol{\alpha}^t, \quad \hat{\mathbf{Y}}_2 = \tfrac{1}{2}\mathbf{S}\hat{\mathbf{X}}_2, \quad \hat{\mathbf{Z}}_2 = (\mathbf{B} + \tfrac{1}{2}\mathbf{S}^2)\hat{\mathbf{X}}_2, \tag{B6}$$

where

$$\mathbf{J} = (\mathbf{B}\mathbf{S} + \mathbf{S}\mathbf{B} + \tfrac{1}{2}\mathbf{S}^3)^{-1}. \tag{B7}$$

A number of relations between the second-order moments can be deduced directly from (B4)–(B7) and (A9), which states that $\mathbf{E}_4$ and α^t are, respectively, left- and right-hand eigenvectors of $\mathbf{S}$ with eigenvalue s. From (B6) and (3.34)

$$\begin{aligned} \mathcal{L}(\langle u_2 v_2\rangle) &= \mathbf{E}_4\hat{\mathbf{Y}}_2 = \tfrac{1}{2}\mathbf{E}_4\mathbf{S}\hat{\mathbf{X}}_2 = \tfrac{1}{2}s\mathbf{E}_4\hat{\mathbf{X}}_2 \\ &= \tfrac{1}{2}s\mathcal{L}(\langle u_2^2\rangle), \end{aligned} \tag{B8}$$

which implies (4.15) since $\mathbf{U}_2(0) = 0$. In this example, $\boldsymbol{\alpha} = \frac{1}{4}\mathbf{E}_4$, and $\mathbf{S}, \mathbf{B}$, and hence $\mathbf{J}$ are symmetric matrices; hence, from (B5) and (B6), we can transpose the matrices in the scalar $\mathbf{E}_4\hat{\mathbf{X}}_1$ to get

$$\begin{aligned} \mathbf{E}_4\hat{\mathbf{X}}_1 &= \tfrac{1}{4}\mathbf{E}_4\mathbf{J}(\mathbf{B} + \tfrac{1}{2}\mathbf{S}^2)\mathbf{E}_4^t = \tfrac{1}{4}\mathbf{E}_4(\mathbf{B} + \tfrac{1}{2}\mathbf{S}^2)\mathbf{J}\mathbf{E}_4^t \\ &= \mathbf{E}_4\hat{\mathbf{Z}}_2. \end{aligned} \tag{B9}$$

From (3.34), this implies

$$\mathcal{L}(\langle u_1^2\rangle) = \mathcal{L}(\langle v_2^2\rangle), \tag{B10}$$

which in turn implies the first of relations (4.16). Similarly, we have from (3.34), (B5), (B9), and (B10)

$$\begin{aligned} \mathcal{L}(\langle u_1 v_1\rangle) &= E_4\hat{\mathbf{Y}}_1 = \tfrac{1}{2}(s\mathbf{E}_4\hat{\mathbf{X}}_1 - 1) = \tfrac{1}{2}(s\mathbf{E}_4\hat{\mathbf{Z}}_2 - 1) \\ &= \tfrac{1}{2}\mathcal{L}(\frac{d}{dz}\langle v_2^2\rangle), \end{aligned} \tag{B11}$$

which implies the second of relations (4.16). Next from (B4), (B6), (3.34), and (3.35),

$$\begin{aligned} \mathcal{L}(\langle u_1 u_2\rangle) &= \mathbf{E}_4\hat{\mathbf{X}}_0 = \tfrac{1}{2}s\mathbf{E}_4\mathbf{J}\boldsymbol{\alpha}^t = \tfrac{1}{2}s\mathbf{E}_4\hat{\mathbf{X}}_2 \\ &= \tfrac{1}{2}s\mathcal{L}(\langle u_2^2\rangle), \end{aligned} \tag{B12}$$

$$\mathbf{E}_4\hat{\mathbf{Y}}_0 = \tfrac{1}{2}s\mathbf{E}_4\hat{\mathbf{X}}_0 = \tfrac{1}{4}s^2\mathcal{L}(\langle u_2^2\rangle) = \tfrac{1}{2}s\mathcal{L}(\langle u_2 v_2\rangle), \quad \text{(B13)}$$

$$\mathcal{L}(\langle v_1 v_2\rangle) = \mathbf{E}_4\hat{\mathbf{Z}}_0 = \tfrac{1}{2}[\mathbf{E}_4(\mathbf{B} + \tfrac{1}{2}\mathbf{S}^2)\mathbf{J}\mathbf{S}\alpha^t - 1]$$
$$= \tfrac{1}{2}(s\mathbf{E}_4\hat{\mathbf{Z}}_2 - 1) = \tfrac{1}{2}[s\mathcal{L}(\langle v_2^2\rangle) - 1]. \quad \text{(B14)}$$

Relations (B12)–(B14), together with (3.35) and initial conditions, imply relations (4.17) and (4.18).

We have shown that all ten second-order moments can be obtained by at most simple differentiation from the three moments $\langle v_1^2\rangle$, $\langle u_2^2\rangle$, and $\langle v_2^2\rangle$. To obtain these moments, the inverse Laplace transform of $\mathbf{E}_4\hat{\mathbf{X}}_2(s)$, $\mathbf{E}_4\hat{\mathbf{Z}}_2(s)$, and $\mathbf{E}_4\hat{\mathbf{Z}}_1(s)$ must be found. However, each of these quantities is a rational function of s with the same denominator

$$\Delta(s;\eta) = \det(2\mathbf{S}\mathbf{B} + 2\mathbf{B}\mathbf{S} + \mathbf{S}^3), \quad \text{(B15)}$$

which is a twelfth-order polynomial in s. If $s_k(\eta)$, $1 \le k \le 12$, are the twelve roots of $\Delta(s;\eta) = 0$, then

$$\langle v_1^2(z)\rangle = \sum_{k=1}^{12} m_k(\eta)\exp[zs_k(\eta)], \quad \text{(B16)}$$

$$\langle u_2^2(z)\rangle = \sum_{k=1}^{12} n_k(\eta)\exp[zs_k(\eta)], \quad \text{(B17)}$$

$$\langle v_2^2(z)\rangle = \sum_{k=1}^{12} p_k(\eta)\exp[zs_k(\eta)], \quad \text{(B18)}$$

where $m_k(\eta)$, $n_k(\eta)$, and $p_k(\eta)$ are the residues, respectively, of $\mathbf{E}_4\hat{\mathbf{Z}}_1(s)$, $\mathbf{E}_4\hat{\mathbf{X}}_2(s)$, and $\mathbf{E}_4\hat{\mathbf{Z}}_2(s)$ at $s = s_k(\eta)$. After considerable algebra, it can be shown that

$$\begin{aligned}\Delta(s;\eta) &= \{s(s+2b_1)(s+2b_2)(s+2b_1+2b_2)(s^2+4\beta_0^2)\\ &\times[(s+2b_1)^2+4\beta_0^2][(s+2b_2)^2+4\beta_0^2]\\ &\times[(s+2b_1+2b_2)^2+4\beta_0^2]\} + 2^5\beta_0^4\eta^2\\ &\times[c_1^2F_1(s) + c_2^2F_2(s)] + 2^8\beta_0^8\eta^4[(c_1^2-c_2^2)\\ &\times(s+b_1+b_2)^2 + b_1^2c_2^2 - b_2^2c_1^2]^2, \end{aligned} \quad \text{(B19)}$$

where

$$\begin{aligned}F_1(s) &= [(s+b_1+b_2)f_+ - b_2g_+]\\ &\times[-(s+b_1+b_2)f_- + b_2g_-]\\ &+ [-(s+b_1+b_2)g_+ + b_2f_+]\\ &\times[(s+b_1+b_2)g_- - b_2f_-] \end{aligned} \quad \text{(B20)}$$

$$f_\pm = (s+b_1+b_2)[(s+b_1+b_2)^2 + 3(b_1 \pm b_2)^2 + 4\beta_0^2], \quad \text{(B21a)}$$

$$g_\pm = (\pm b_1 + b_2)[3(s+b_1+b_2)^2 + (b_1 \pm b_2)^2 + 4\beta_0^2], \quad \text{(B21b)}$$

and $F_2(s)$ is obtained from $F_1(s)$ by interchanging b_1 and b_2.

For $\eta \ll 1$ and sufficiently small compared to b_1 and b_2, $\Delta(s;\eta)$ can be shown to have one root with a positive real part, while the remaining roots have negative real parts:

$$s_1 = \frac{\eta^2\beta_0^2}{2}\left[c_1^2\left(\frac{b_1}{b_1^2+\beta_0^2}\right) + c_2^2\left(\frac{b_2}{b_2^2+\beta_0^2}\right)\right] + O(\eta^4), \quad \text{(B22a)}$$

$$s_2 = s_3^* = 2i\beta_0 - \frac{\eta^2\beta_0^2}{4}\left[\frac{c_1^2(b_1+2i\beta_0)}{b_1(b_1+i\beta_0)} + \frac{c_2^2(b_2+2i\beta_0)}{b_2(b_2+i\beta_0)}\right] + O(\eta^4), \quad \text{(B22b)}$$

$$s_4 = -2b_1 - 2b_2 + O(\eta^2), \quad \text{(B22c)}$$

$$s_5 = s_6^* = -2b_1 - 2b_2 + 2i\beta_0 + O(\eta^2), \quad \text{(B22d)}$$

$$s_7 = s_8^* = -2b_1 + 2i\beta_0 + O(\eta^2), \quad \text{(B22e)}$$

$$s_9 = -2b_1 + O(\eta^2), \quad \text{(B22f)}$$

$$s_{10} = s_{11}^* = -2b_2 + 2i\beta_0 + O(\eta^2), \quad \text{(B22g)}$$

$$s_{12} = -2b_2 + O(\eta^2). \quad \text{(B22h)}$$

The corresponding coefficients $m_k(\eta)$, $n_k(\eta)$, and $p_k(\eta)$ can be shown to have the form

$$m_1(\eta) = \tfrac{1}{2}\beta_0^2 + O(\eta^2), \quad \text{(B23a)}$$

$$m_2(\eta) = m_3(\eta) = -\tfrac{1}{4}\beta_0^2 + O(\eta^2), \quad \text{(B23b)}$$

$$m_k(\eta) = O(\eta^2), \quad 4 \le k \le 12, \quad \text{(B23c)}$$

$$n_1(\eta) = 1/(2\beta_0^2) + O(\eta^2), \quad \text{(B24a)}$$

$$n_2(\eta) = n_3(\eta) = -1/(4\beta_0^2) + O(\eta^2), \quad \text{(B24b)}$$

$$n_k(\eta) = O(\eta^2), \quad 4 \le k \le 12, \quad \text{(B24c)}$$

$$p_1(\eta) = \tfrac{1}{2} + O(\eta^2), \quad \text{(B25a)}$$

$$p_2(\eta) = p_3(\eta) = \tfrac{1}{4} + O(\eta^2), \quad \text{(B25b)}$$

$$p_k(\eta) = O(\eta^2), \quad 4 \le k \le 12. \quad \text{(B25c)}$$

In the special cases $c_1 = \pm c_2$ and $b_1 = b_2 = b$, it can again be easily seen that $\Delta(s;\eta)$ has the factors $(s+2b)^2$ and $[(s+2b)^2 + 4\beta_0^2]$. Again by making use from the start of the symmetry relations $\sigma_1(\mathbf{w}, z) \equiv \sigma_4(\mathbf{w}, z)$ or $\sigma_2(\mathbf{w}, z) \equiv \sigma_3(\mathbf{w}, z)$, it can be shown that one of the roots $s = -2b$ and both of the roots $s = -2b \pm 2i\beta_0$ do not appear in the expressions for $\langle v_1^2\rangle$, $\langle u_2^2\rangle$, or $\langle v_2^2\rangle$. The expressions for $\langle v_1^2\rangle$, $\langle u_2^2\rangle$, and $\langle v_2^2\rangle$ are still given by (B16)–(B18) although now only nine terms appear in each sum. The expressions (B22) are still valid for $s_k(\eta)$, $1 \le k \le 8$, when $c_1^2 = c_2^2 = \frac{1}{2}$ and $b_1 = b_2 = b$, but $s_9(\eta) = -2b$. The expressions (B23)–(B25) are also still valid for $1 \le k \le 9$ when $c_1^2 = c_2^2 = \frac{1}{2}$. Relations (B8)–(B14) do not depend on the specific form of $\Delta(s;\eta)$ and are still valid.

[1] J. McKenna and J. A. Morrison, J. Math. Phys. 11, 2348 (1970).
[2] J. L. Doob, *Stochastic Processes* (Wiley, New York, 1953), pp. 235-55.
[3] M. Rosenblatt, *Random Processes* (Oxford U.P., New York, 1962), p. 59.
[4] R. Z. Khas'minskii, Theory Prob. Appl. 11, 390 (1964). See also M. Lax, Rev. Mod. Phys. 38, 541 (1966)
[5] J. A. Morrison, J. Math. Anal. Appl., to appear.
[6] U. Frisch, W. Prop. Ran. M., in *Probabilistic Methods in Applied Mathematics*, edited by A. T. Bharucha-Reid (Academic, New York, 1968). p. 114.
[7] R. C. Bourret, Can. J. Phys. 43, 619 (1965), and earlier papers referred to in this reference.
[8] J. B. Keller, Proc. Symp. Appl. Math. 13, 227 (1962); 16, 145 (1964).
[9] See extensive bibliography in Ref. 6.
[10] G. Papanicolaou and J. B. Keller, SIAM J. Appl. Math., to appear.
[11] R. C. Bourret, Can. J. Phys. 44, 2519 (1966).
[12] J. A. Morrison and J. McKenna, J. Math. Phys. 11, 2361 (1970).
[13] J. A. Morrison, J. Math. Phys. 11, 3200 (1970).
[14] W. Feller, Trans. Am. Math. Soc. 48, 488 (1940).
[15] C. J. Burke and M. Rosenblatt, Ann. Math. Stat. 29, 1112 (1958).
[16] W. M. Wonham, in *Probabilistic Methods in Applied Mathematics*, edited by A. T. Bharucha-Reid (Academic, New York, 1970), Vol. II, p. 131.
[17] N. N. Krasovskii and E. A. Lidskii, Appl. Math. Mech. 25, 627 (1961).
[18] N. N. Krasovskii and E. A. Lidskii, Auto. Remote Control 22, 1021, 1141, 1289 (1961).
[19] R. Bellman, *Introduction to Matrix Analysis* (McGraw-Hill, New York, 1960), p. 227.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 10 OCTOBER 1971

Representations of $O(3)$

Kishor C. Tripathy*
Institute of Theoretical Physics, Fack, S-402 20, Göteborg 5, Sweden
(Received 8 March 1971)

We analyze the various classes of local irreducible representations (both finite and infinite-dimensional) of $O(3)$ realized on a linear complex vector space. The well-known finite-dimensional unitary representations are obtained when the scalar product is positive definite. The infinite-dimensional representations are realized on an indefinite metric space. The possible applications of these new classes of representations are briefly discussed.

1. INTRODUCTION

In recent years, various models have been developed to describe the evergrowing high-spin particles in a coherent fashion. The construction of a field theory over the homogeneous space of the Poincaré group P which contains the Minkowski space and which have continuous stabilizer groups is worth mentioning in this regard.[1] Another attempt was made to build up "Majorana-like" field equations based on the well-known notions of the hydrogen atom.[2] To us, it seems there are still a lot to be understood about this quantum mechanical system (namely, the hydrogen atom). It has long been well known that the symmetry of the bound-states of the two-dimensional hydrogen atom is $O(3)$ and $O(2,1)$ describes the symmetry of the scattering states.[3] The postulation of two different groups and two different representations for describing the different states of the *same* physical system is rather puzzling. One, in principle could invoke a bigger dynamical group as the symmetry group of this system describing both bound and scattering states.[4] However, this gives rise to additional difficulty of suitable interpretation of the various generators and the corresponding representation space.

Further, in conventional Regge theory, one is confined to a very restricted type of transformation associated with the continuous nonexceptional series representation of $O(2,1)$. It is possible to construct scattering amplitudes where the spin is continued simultaneously with the mass, so that the particles lie on the same Regge trajectory. So, one has to consider a more general class of transformations (generalized Sommerfeld–Watson transformation) to construct scattering amplitudes possessing momentum transfer as parameter as well. Some of these properties were investigated earlier in connection with the "local representation" of the rotation group $O(3)$. However, the analysis makes use of the continuous series representation of $O(3)$ [analogous to that of $O(2,1)$] and is far from complete.[5] It is the purpose of this paper to present a systematic analysis of the various classes of representations of $O(3)$ in addition to the well-known finite, unitary, irreducible representations and the continuous series representation which is infinite-dimensional. To furnish this, we follow the method outlined by Miller[6] and consider $O(3)$ as a group of transformations, the various representations of it being realized over a linear vector space. The linear vector space (in general) is a metric space, spanned by a set of eigenvectors of the Casimir operator Q and J^3. The dimensionality of the vector space furnishes the dimensionality of the underlying representations of $O(3)$. We have considered the following classes of representations.

Let $q = \phi(\phi + 1)$ and $E = E_0 + m$ (m, an integer) be the eigenvalues of Q and J^3, respectively. In general ϕ and E will be complex.

Class I: In this case, Q and J^3 are self-adjoint and Q takes integer values on the real axis and $E = -\phi, \cdots, +\phi$. Thus, we obtain the well-known finite dimensional unitary irreducible representations of $O(3)$, namely $D(2\phi)$.

Class II: The Class II representation is infinite dimensional and the Nelson operator Q is not self-adjoint, and the spectrum of J^3 is given by $E_0 = -\phi + n, n = 0, 1, 2, 3, \cdots$ and $E_0 - \phi \neq$ an integer.

[1] J. McKenna and J. A. Morrison, J. Math. Phys. 11, 2348 (1970).
[2] J. L. Doob, *Stochastic Processes* (Wiley, New York, 1953), pp. 235-55.
[3] M. Rosenblatt, *Random Processes* (Oxford U.P., New York, 1962), p. 59.
[4] R. Z. Khas'minskii, Theory Prob. Appl. 11, 390 (1964). See also M. Lax, Rev. Mod. Phys. 38, 541 (1966)
[5] J. A. Morrison, J. Math. Anal. Appl., to appear.
[6] U. Frisch, W. Prop. Ran. M., in *Probabilistic Methods in Applied Mathematics*, edited by A. T. Bharucha-Reid (Academic, New York, 1968). p. 114.
[7] R. C. Bourret, Can. J. Phys. 43, 619 (1965), and earlier papers referred to in this reference.
[8] J. B. Keller, Proc. Symp. Appl. Math. 13, 227 (1962); 16, 145 (1964).
[9] See extensive bibliography in Ref. 6.
[10] G. Papanicolaou and J. B. Keller, SIAM J. Appl. Math., to appear.
[11] R. C. Bourret, Can. J. Phys. 44, 2519 (1966).
[12] J. A. Morrison and J. McKenna, J. Math. Phys. 11, 2361 (1970).
[13] J. A. Morrison, J. Math. Phys. 11, 3200 (1970).
[14] W. Feller, Trans. Am. Math. Soc. 48, 488 (1940).
[15] C. J. Burke and M. Rosenblatt, Ann. Math. Stat. 29, 1112 (1958).
[16] W. M. Wonham, in *Probabilistic Methods in Applied Mathematics*, edited by A. T. Bharucha-Reid (Academic, New York, 1970), Vol. II, p. 131.
[17] N. N. Krasovskii and E. A. Lidskii, Appl. Math. Mech. 25, 627 (1961).
[18] N. N. Krasovskii and E. A. Lidskii, Auto. Remote Control 22, 1021, 1141, 1289 (1961).
[19] R. Bellman, *Introduction to Matrix Analysis* (McGraw-Hill, New York, 1960), p. 227.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 10 OCTOBER 1971

Representations of $O(3)$

Kishor C. Tripathy*
Institute of Theoretical Physics, Fack, S-402 20, Göteborg 5, Sweden
(Received 8 March 1971)

We analyze the various classes of local irreducible representations (both finite and infinite-dimensional) of $O(3)$ realized on a linear complex vector space. The well-known finite-dimensional unitary representations are obtained when the scalar product is positive definite. The infinite-dimensional representations are realized on an indefinite metric space. The possible applications of these new classes of representations are briefly discussed.

1. INTRODUCTION

In recent years, various models have been developed to describe the evergrowing high-spin particles in a coherent fashion. The construction of a field theory over the homogeneous space of the Poincaré group P which contains the Minkowski space and which have continuous stabilizer groups is worth mentioning in this regard.[1] Another attempt was made to build up "Majorana-like" field equations based on the well-known notions of the hydrogen atom.[2] To us, it seems there are still a lot to be understood about this quantum mechanical system (namely, the hydrogen atom). It has long been well known that the symmetry of the bound-states of the two-dimensional hydrogen atom is $O(3)$ and $O(2,1)$ describes the symmetry of the scattering states.[3] The postulation of two different groups and two different representations for describing the different states of the *same* physical system is rather puzzling. One, in principle could invoke a bigger dynamical group as the symmetry group of this system describing both bound and scattering states.[4] However, this gives rise to additional difficulty of suitable interpretation of the various generators and the corresponding representation space.

Further, in conventional Regge theory, one is confined to a very restricted type of transformation associated with the continuous nonexceptional series representation of $O(2,1)$. It is possible to construct scattering amplitudes where the spin is continued simultaneously with the mass, so that the particles lie on the same Regge trajectory. So, one has to consider a more general class of transformations (generalized Sommerfeld–Watson transformation) to construct scattering amplitudes possessing momentum transfer as parameter as well. Some of these properties were investigated earlier in connection with the "local representation" of the rotation group $O(3)$. However, the analysis makes use of the continuous series representation of $O(3)$ [analogous to that of $O(2,1)$] and is far from complete.[5] It is the purpose of this paper to present a systematic analysis of the various classes of representations of $O(3)$ in addition to the well-known finite, unitary, irreducible representations and the continuous series representation which is infinite-dimensional. To furnish this, we follow the method outlined by Miller[6] and consider $O(3)$ as a group of transformations, the various representations of it being realized over a linear vector space. The linear vector space (in general) is a metric space, spanned by a set of eigenvectors of the Casimir operator Q and J^3. The dimensionality of the vector space furnishes the dimensionality of the underlying representations of $O(3)$. We have considered the following classes of representations.

Let $q = \phi(\phi + 1)$ and $E = E_0 + m$ (m, an integer) be the eigenvalues of Q and J^3, respectively. In general ϕ and E will be complex.

Class I: In this case, Q and J^3 are self-adjoint and Q takes integer values on the real axis and $E = -\phi, \cdots, +\phi$. Thus, we obtain the well-known finite dimensional unitary irreducible representations of $O(3)$, namely $D(2\phi)$.

Class II: The Class II representation is infinite dimensional and the Nelson operator Q is not self-adjoint, and the spectrum of J^3 is given by $E_0 = -\phi + n, n = 0, 1, 2, 3, \cdots$ and $E_0 - \phi \neq$ an integer.

This furnishes the $D+(\phi, E_0)$ class of $O(3)$ (we are retaining the *same* nomenclature for these various classes of representations in the spirit of Barut and Fronsdal[7]).

Class III: $D^-(\phi, E_0)$: Like Class II, in this case, we obtain an infinite-dimensional representation of $O(3)$. The spectrum of J^3 is given by $E_0 = \phi - n$, n = a nonnegative integer, and $E_0 + \phi \neq$ an integer.

In Class II and Class III representations, the underlying space is equipped with an indefinite metric. The convergence properties in this space can be examined by introducing Hilbert or Frechet topologies.[8]

Class IV: $D(\phi, E_0)$: This class of representation is obtained for all complex values of ϕ and E:$(\phi + E_0)$ and $(\phi - E_0)$ being nonintegers.

It is our contention that the class I representation can be used as usual to construct the bound-state energy spectrum and Class IV for the scattering states of the two-dimensional hydrogen atom. We also suggest that the infinite-dimensional representations of the compact groups could as well be used to build up infinite-multiplet schemes for the hadrons. We have arranged our material as follows.

In Sec. 2, we describe the method and obtain the usual finite-dimensional unitary representations. Explicit expression for the infinitesimal generators, eigenfunctions, and matrix elements have been obtained. Following the method outlined in Sec. 2, we study the properties of the vector space when both ϕ and E take complex values. We obtain the infinite-dimensional representations for $D+(\phi, E_0)$, $D^-(\phi, E_0)$, and $D(\phi, E_0)$. This constitutes Sec. 3. Finally, we outline the application of our new class of representations for studying the symmetry of the two-dimensional hydrogen atom in Sec. 4. We conclude our discussion with a few remarks on the special features of the metric space.

2. THE METHOD AND CONSTRUCTION OF FINITE DIMENSIONAL REPRESENTATIONS

A. The Method

Let G be an n-dimensional local Lie-group and U an open set in C^m (complex fields). Let $U \times G \xrightarrow{F} C^m$, i.e., the mapping $F(\mathbf{z}, g) = \mathbf{z}g$ for $\mathbf{z} \in U$ and $g \in G$.

G is said to be a local Lie transformation group, if F satisfies the following conditions.

(1) $\mathbf{z}g$ is analytic in the coordinates of z and g;

(2) $\mathbf{z}e = \mathbf{z}$;

(3) $\mathbf{z}g \in U \Rightarrow (\mathbf{z}g)g' = \mathbf{z}(gg'), \forall g, g' \in G$.

Let G act on an open neighborhood of U of C^m, $O \in U$ and let $\mathfrak{C}$ be the set of all complex valued functions on U analytic in the neighborhood of O. The local multiplier representation T^ρ of G on H with multiplier ρ constitutes a mapping $T^\rho(g)$ of H onto H,

$$[T^\rho(g)f](\mathbf{z}) = \rho(\mathbf{z}, G)f(\mathbf{z}, g), \quad \mathbf{z} \in U, g \in G \text{ and } f \in H. \tag{2.1}$$

The multiplier $\rho(\mathbf{z}, g)$ satisfies the following properties, namely:

$$\text{(i)}\ \ \rho(\mathbf{z}, e) = 1, \quad \forall \mathbf{z} \in U, \tag{2.2}$$

$$\text{(ii)}\ \ \rho(\mathbf{z}, gg') = \rho(\mathbf{z}, g)\rho(\mathbf{z}, g'), \quad g, g' \in G. \tag{2.3}$$

Property (ii) implies the following. We define the generalized Lie derivatives ${}_\alpha Df(\mathbf{z})$ of the analytic function $f(\mathbf{z})$ under the one-parameter group $(\exp\alpha t)$ as

$$D_\alpha f(\mathbf{z}) = \frac{d}{dt}\,[T^\rho(\exp\alpha t)f](\mathbf{z})\Big|_{t=0}, \tag{2.4}$$

where $\alpha = (d/dt)g(t)|_{t\,0}$ is the tangent vector at $e = g(0)$ on the curve $g(t)$. For $\rho = 1$, we obtain the ordinary Lie derivative. From (2.4), we obtain

$$D_\alpha f(\mathbf{z}) = \sum_{j=1}^{n} \alpha_j P_{j\,i}(\mathbf{z})\,\frac{\partial f}{\partial z_i}(\mathbf{z}) + \sum \alpha_j P_j(\mathbf{z}) f(\mathbf{z}),$$

where

$$\sum_j \alpha_j P_j(\mathbf{z}) = \frac{d}{dt}\rho(\mathbf{z}, \exp\alpha t)\Big|_{t=0}$$

and

$$P_{j\,i}(\mathbf{z}) = \frac{\partial F_i}{\partial g_j}(\mathbf{z}, g)\Big|_{g=e}. \tag{2.5}$$

The D_αs form a Lie algebra under the operation of addition of derivatives and Lie product

$$[D_\alpha, D_\beta] = D_\alpha D_\beta - D_\beta D_\alpha$$

and is the homomorphic image of the Lie algebra of G, i.e., $\mathfrak{G}$;

$$D_{\alpha+\beta} = D_\alpha + D_\beta, \quad D_{[\alpha,\beta]} = [D_\alpha, D_\beta], D_{a\alpha} = aD_\alpha.$$

Let us define

$$D_j(\mathbf{z}) = \sum_{i=1}^{m} P_{j\,i}(\mathbf{z})\,\frac{\partial}{\partial z_i} + P_j(\mathbf{z}), \quad j = 1, 2, 3, \cdots, n \tag{2.6}$$

$\{D_j(\mathbf{z})\}$ are a set of linearly independent differential operators analytic in the open set $U \in C^m$ and satisfy

$$[D_j, D_k] = D_jD_k - D_kD_j = \sum_{l=1}^{n} C_{jk}^{l} D_l, \tag{2.7}$$

where C_{jk}^{i} are the structure constants. A complex linear combination of D_js spans the Lie algebra of the generalized Lie derivatives. The action of the group is obtained by integrating the following equations;

(i) $\frac{d}{dt} z_i(t) = \sum_{j=1}^{n} \alpha_j P_{ji}(\mathbf{z}(t)), \quad z_i(0) = z_i^0,$

$i = 1, 2, 3, \cdots$

(ii) $\frac{d}{dt} \rho(\mathbf{z}^0, e^{\alpha t}) = \rho(\mathbf{z}^0, e^{\alpha t}) \sum_j \alpha_j P_j(\mathbf{z}(t)),$

$\rho(\mathbf{z}^0, e) = 1,$

where

$$\mathbf{z}(t) = \mathbf{z}^0 e^{\alpha t},$$

(iii) $[T^\rho(e^{\alpha t})f](\mathbf{z}^0) = \rho(\mathbf{z}^0, e^{\alpha t}) f(\mathbf{z}^0, e^{\alpha t}).$ (2.8)

B. Construction of Class I representations of $O(3)$

Let us consider the differential operators[6]

$$J^3 = -\phi + z\frac{d}{dz}, \quad J^+ = -2\phi z + z^2\frac{d}{dz},$$

$$J^- = -\frac{d}{dz}. \tag{2.9}$$

z is a complex variable defined in the neighborhood of $0 \in C'$, and 2ϕ is a nonnegative integer. $\{J^\pm, J^3\}$ satisfy the following commutation relations [compare Eq. (2.7)]:

$$[J^3, J^\pm] = \pm J^\pm, \quad [J^+, J^-] = 2J^3,$$

and span the Lie algebra of generalized Lie derivatives of a local multiplier representation $T^\rho(g)$ of $O(3)$.

To obtain the action of the one-parameter group $e^{\gamma\sigma^3}, \gamma \in C'$, σ^3 being the Lie algebra element corresponding to J^3, we solve the following equations analogous to (2.8), namely

$$\frac{d}{d\gamma}(z) = z, \quad \frac{d}{d\gamma}\rho(z^0, e^{\gamma\sigma^3}) = -\phi\rho(z^0, e^{\gamma\sigma^3}), \tag{2.10}$$

with the initial conditions

$$z(0) = z^0, \quad \rho(z^0, e) = 1.$$

Thus, we obtain from (2.10)

$$z(\gamma) = z^0 e^\gamma, \quad \rho(z^0, e^{\gamma\sigma^3}) = e^{-\phi\gamma}.$$

Now,

$$[T^\rho(e^{\gamma\sigma^3})f](z^0) = e^{-\phi\gamma} f(z^0 e^\gamma), \tag{2.11}$$

for $|\gamma|$ sufficiently small. Similarly, we obtain

$$[T^\rho(e^{s\sigma^+})f](z^0) = (-sz^0 + 1)^{2\phi} f\left(\frac{z^0}{-sz^0 + 1}\right) \tag{2.12}$$

and

$$[T^\rho(e^{t\sigma^-})f](z^0) = f(z^0 - t), \tag{2.13}$$

$\sigma^3, \sigma^\pm$ are the usual Pauli matrices forming the two-dimensional representation of $O(3)$ and are given by

$$\sigma^+ = \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix}, \quad \sigma^- = \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix}, \quad \sigma^3 = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} \end{pmatrix}.$$

It follows then,

$$e^{\gamma\sigma^3} = \begin{pmatrix} e^{\gamma/2} & 0 \\ 0 & e^{-\gamma/2} \end{pmatrix}, \quad e^{s\sigma^+} = \begin{pmatrix} 1 & -s \\ 0 & 1 \end{pmatrix},$$

$$e^{t\sigma^-} = \begin{pmatrix} 1 & 0 \\ -t & 1 \end{pmatrix}, \quad \gamma, s, t \in C \tag{2.14}$$

Any arbitrary element $g \in O(3)$ can be written as

$$g = \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix}, \quad \det \|g\| = 1. \tag{2.15}$$

Further, in a small neighborhood of the identity we can write g as

$$g = \exp(s\sigma^+) \cdot \exp(t\sigma^-) \cdot \exp(\gamma\sigma^3). \tag{2.16}$$

Making use of (2.14) in (2.16), we obtain

$$g = \begin{pmatrix} e^{\gamma/2}(1 + st) & -se^{-\gamma/2} \\ -te^{\gamma/2} & e^{-\gamma/2} \end{pmatrix}. \tag{2.17}$$

Thus

$$e^{-\gamma/2} = \bar{a}, \quad s = -b/\bar{a}, \quad t = \bar{b} \cdot \bar{a}. \tag{2.18}$$

Then,

$$[T^\rho(g)f](z) = \{T^\rho[\exp(s\sigma^+) \cdot \exp(t\sigma^-) \cdot \exp(\gamma\sigma^3)]f\}(z). \tag{2.19}$$

Making use of (2.11)–(2.13) in (2.19) we finally obtain

$$[T^\rho(g)f](z) = e^{-\phi\gamma}(-sz + 1)^{2\phi} f\left(\frac{ze^\gamma(1 + st) - te^\gamma}{-sz + 1}\right)$$

$$= (bz + \bar{a})^{2\phi} f\left(\frac{az - \bar{b}}{bz + \bar{a}}\right). \tag{2.20}$$

Thus,

$$\rho(z, g) = (bz + \bar{a})^{2\phi}$$

and

$$z' = zg = \frac{az - \bar{b}}{bz + \bar{a}}. \tag{2.21}$$

It can now be easily checked that (2.20) and (2.21) indeed furnish the local multiplier representation of $O(3)$ and the differential operators $J^\pm, J^3$ are generalized Lie derivatives of T^ρ. Alternatively, given (2.20) and (2.21), we can obtain the generalized Lie derivatives of a local multiplier representation T^ρ of $O(3)$ from (2.4).

Writing $[f(z)] \equiv [f_E^{(\phi)}(z)]$, the action of the operators $J^\pm, J^3$ on this basis is given by

$$J^3 f_E^{(\phi)}(z) = E f_E^{(\phi)}(z),$$

$$J^\pm f_E^{(\phi)}(z) = [(\phi \mp E)(\phi \pm E + 1)]^{1/2} f_{E\pm1}^{(\phi)}(z). \tag{2.22}$$

Note that

$$J^+ f_\phi^{(\phi)}(z) \equiv 0, \text{ i.e., } f_{\phi+1}^{(\phi)}(z) = 0$$

and

$$J^- f^{(\phi)}_{-\phi}(z) \equiv 0, \text{ i.e., } f^{(\phi)}_{-\phi-1}(z) = 0.$$

Thus, an invariant linear-vector space is spanned by $(2\phi + 1)$ vectors $(f_{-\phi}, \cdots, f_{+\phi})$. We now compute the scalar product and show that the vector space is Hilbertian.

Using

$$1 + \left|\frac{az - \bar{b}}{\bar{a} + bz}\right|^2 = \frac{1 + z\bar{z}}{|\bar{a} + bz|^2}$$

[since, $|a|^2 + |b|^2 = 1$], we obtain from (2.21)

$$\frac{dz'}{dz} = |(\bar{a} + bz)|^{-2}(1 + |z|^2). \tag{2.23}$$

Thus, we define the inner product w.r.t. the invariant measure

$$d\mu(z) = \frac{2\phi + 1}{\pi}(1 + |z|^2)^{-2\phi-2}\, dz \cdot d\bar{z}$$

as

$$(f_1, f_2)_\phi = \int \overline{f_1(z)} f_2(z) d\mu(z)$$
$$= \frac{2\phi + 1}{\pi} \int \overline{f_1(z)} f_2(z)(1 + |z|^2)^{-2\phi-2}\, dz d\bar{z}, \tag{2.24}$$

where the integration is carried over the complex plane and $f_1, f_2 \in H$. Setting $f_1(z) = z^k, f_2(z) = z^l$, $k, l = 0, 1, 2, \cdots, 2\phi$, we obtain from (2.24)

$$(f_1, f_2)_\phi = \binom{-2\phi + k - 1}{k}^{-1}. \tag{2.25}$$

In (2.25) we have used $(z^k, z^l) = \delta_{kl}$. Thus, writing

$$f(z) = \sum_{k=0}^{2\phi} C_k z^k,$$

we obtain

$$\|f\|^2 = \sum_k \binom{-2\phi + k - 1}{k}^{-1} \|C_k\|^2 < \infty. \tag{2.26}$$

We now construct a complete set of orthonormal vectors $[h_k(z)]$ as

$$\|h\|^2 = 1, \text{ and } h(z) = i^k \left[\frac{\Gamma(k - 2\phi)}{\Gamma(k + 1)\Gamma(-2\phi)}\right]^{1/2} z^k. \tag{2.27}$$

In this basis, the matrix elements of $T^\rho(g)$ [which also furnishes the representations of $O(3)$] are given by

$$[T^\rho(g)h_k](z) = \sum_{l=0}^{2\phi} D_{lk}(g) h_l(z), \qquad 0 \le k \le 2\phi,$$

or

$$N_k(\bar{a} + bz)^{2\phi-k}(az - \bar{b})^k = \sum_l N_l D_{lk}(g) z^l, \tag{2.28}$$

where N_k and N_l are the normalization constants given in (2.27). Thus,

$$D_{lk}(g) = \frac{N_k}{N_l}(\bar{a})^{2\phi-k} \cdot a^l (-\bar{b})^{k-l} \frac{\Gamma(k + 1)}{\Gamma(l - 1)} \cdot \frac{1}{\Gamma(k + 1)}$$
$$\cdot F(-l, -2\phi + k, k - l + 1; -b\bar{b}/a\bar{a}),$$
$$\text{if } 2\phi \ge k \ge l \ge 0,$$
$$= \frac{N_l}{N_k} a^k (\bar{a})^{2\phi-l}(b)^{l-k}$$
$$\cdot \frac{\Gamma(-2\phi - k + 1)}{\Gamma(-2\phi - l + 1)} \cdot \frac{1}{\Gamma(l - k + 1)}$$
$$\cdot F(-k, -2\phi + l, l - k + 1; -b\bar{b}/a\bar{a}),$$
$$\text{if } 2\phi \ge l \ge k \ge 0. \tag{2.29}$$

Substituting for N_k, N_l and using

$$g = \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix}, \quad a = e^{-i(\alpha+\gamma)/2} \cdot \cos(\beta/2),$$
$$b = ie^{-i(\alpha-\gamma)/2} \sin(\beta/2)$$

(where α, β, γ are the Eulerian angles), and making use of the properties of the hypergeometric function $F(p, q, \gamma; z)$, namely

$$F(p, q, \gamma; z) = (1 - z)^{-p} F(p, \gamma - q, \gamma; z/(z - 1))$$
$$= (1 - z)^{\gamma-q-p} F(\gamma - p, \gamma - q, \gamma; z), \tag{2.30}$$

we obtain

$$D_{lk}(g) = \frac{1}{(k - l)!}\left[\frac{\Gamma(k - 2\phi)\Gamma(k + 1)}{\Gamma(l - 2\phi)\Gamma(l + 1)}\right]^{1/2} e^{i(\phi-k)\alpha}$$
$$\cdot e^{i(\phi-l)\gamma}(\cos\beta/2)^{2\phi-k-l} \cdot (\sin\beta/2)^{k-l}$$
$$\cdot F(-l, -l + 2\phi + 1, k - l + 1; \sin^2\beta/2). \tag{2.31}$$

(2.31) is exactly the expression given by Edmond for the matrix elements of the rotation group $O(3)$.[9]

3. CONSTRUCTION OF THE DISCRETE CLASSES II AND III AND THE CONTINUOUS SERIES REPRESENTATIONS

A. Class II Representations ($D^+(\phi, E_0)$)

Let the spectrum of J^3 be $E = (-\phi + m, m =$ a nonnegative integer). Using the differential form for the operators $J^3, J^\pm$ as given by (2.9) and substituting $E = -\phi + m$ in (2.22), we obtain

$$J^- f^{(\phi)}_E(z) = -i[m(m - 1 - 2\phi)]^{1/2} f^{(\phi)}_{E-1}(z),$$
$$m = 0, 1, 2, \cdots.$$

Then,

$$J^- f^{(\phi)}_{-\phi}(z) = 0 \Rightarrow f^{(\phi)}_{-\phi-1}(z) = 0.$$

However, $J^+ f^{(\phi)}_{-\phi+m}(z) = 0$ has no solution unless $m = 2\phi$ (compare finite-dimensional representations). Thus the infinite system of functions $[f^{(\phi)}_{-\phi+m}(z)] \equiv [f^{(\phi)}_{-\phi}(z), f^{(\phi)}_{-\phi+1}(z), \cdots]$ span the infinite-dimensional vector space for $O(3)$.

We further analyze the properties of this infinite-dimensional vector space by explicitly computing the scalar product. Then we obtain the multiplier representations [which also furnish the representations of $O(3)$] w.r.t. this basis. As usual, we define $[f_E^{(\phi)}(z)]$:

$$f_E^{(\phi)}(z) = N_m^{(\phi)} z^m.$$

$N_m^{(\phi)}$ is formally given by (2.27) except ϕ is now complex.

Now,

$$(f_1, f_2)_\phi = \int \overline{f_1(z)} f_2(z) d\mu(z)$$

or

$$(f_E^{(\phi)}(z), f_{E'}^{(\phi)}(z)) = \int \overline{f_E^{(\phi)}(z)} \cdot f_{E'}^{(\phi)}(z) d\mu(z)$$

$$= \frac{2\phi + 1}{\pi} \cdot \overline{N_E^{(\phi)}} \cdot N_E^{(\phi)} \int \bar{z}^m \cdot z^n \cdot (1 + |z|^2)^{-2\phi-2} \cdot dz \cdot d\bar{z}. \tag{3.1}$$

Substituting $z = \sqrt{\gamma} e^{i\theta}$, $z = x + iy$, $dz \cdot d\bar{z} = dx \cdot dy = \frac{1}{2} d\gamma \cdot d\theta$, we have

$$\int \bar{z}^m \cdot z^n (1 + |z|^2)^{-2\phi-2} dz \cdot d\bar{z}$$

$$= \tfrac{1}{2} \int_0^{2\pi} d\theta e^{-i(m-n)\theta} \int_0^\infty \gamma^{m+n/2} \cdot (1+\gamma)^{-2\phi-2} d\gamma$$

$$= \pi \delta_{mn} B(m+1, 2\phi - m + 1), \quad \mathrm{Re}\phi > m > -1, \tag{3.2}$$

where $B(p, q)$ is the beta function.

Thus, using $f_1 = \bar{f}$, $f_2 = f$, we have from (3.1)

$$(f, f)_\phi = (2\phi + 1) \cdot \overline{N_{\bar{E}}^{(\bar{\phi})}} \cdot N_E^{(\phi)} \cdot B(m+1, 2\phi - m + 1). \tag{3.3}$$

Substituting for $N_{\bar{E}}^{(\bar{\phi})}$ and $N_E^{(\phi)}$ and using

$$B(m+1, 2\phi - m + 1) = \Gamma(m+1)\Gamma(2\phi + 1 - m)/\Gamma(2\phi + 2)$$

and

$$\Gamma(-z) = \frac{(-1)^z}{\Gamma(z+1)},$$

we obtain from (3.3)

$$(f, f)_\phi = (-1)^m \left[\frac{\overline{\Gamma(m - 2\bar{\phi})}}{\Gamma(-2\bar{\phi})}\right]^{1/2} \cdot \left[\frac{\Gamma(m - 2\phi)}{\Gamma(-2\phi)}\right]^{-1/2}$$

$$\equiv (-1)^m \zeta. \tag{3.4}$$

Some properties of the metric ζ

(1) $m = 2\phi = 2\bar{\phi}$, (ϕ real), then $\zeta \equiv 1$.

This precisely produces the finite-dimensional unitary representations.

(2) $\zeta = +1$, if $2\,\mathrm{Re}\phi$ = even, $m > \mathrm{Re}\phi$

$= -1$, if $2\,\mathrm{Re}\phi$ = odd, $m > \mathrm{Re}\phi$.

(3) $\zeta = (-1)^m$, if $\mathrm{Re}\phi > 0$ and $m < 2\,\mathrm{Re}\phi$.

We now compute the multiplier representations $T^\rho(g)$ on the basis

$$f_n^{(\phi)}(z) = i^n \left[\frac{\Gamma(n - 2\phi)}{\Gamma(n+1)\Gamma(-2\phi)}\right]^{1/2} \cdot z^n, \quad n = 0, 1, 2, \cdots,$$

i.e.,

$$[T^\rho(g) f^{(\phi)}](z) = (\bar{a} + bz)^{2\phi} f\left(\frac{az - \bar{b}}{bz + \bar{a}}\right) \text{ for } z \in C \text{ and } g \in O(3)$$

$$= \sum_{l=0}^\infty D_{lk}^{(\phi, E_0)}(g) f_l^{(\phi, E_0)}, \quad k = 0, 1, 2, \cdots, \tag{3.5}$$

or

$$(az - \bar{b})^k (\bar{a})^{2\phi - k} \cdot (1 + \frac{bz}{\bar{a}})^{2\phi - k} \cdot N^k$$

$$= \sum_{l=0}^\infty N_l D_{lk}^{(\phi, E_0)}(g) z^l, \quad |\frac{bz}{\bar{a}}| < 1, \ |\frac{az}{\bar{b}}| < 1.$$

Comparing the coefficients of z^l on either side of the above equation, we get

$$D_{lk}^{(\phi, E_0)}(g) = \frac{N_k}{N_l} \cdot a^l \cdot (\bar{a})^{2\phi - k} (-\bar{b})^{k-l} \cdot \frac{\Gamma(k+1)}{\Gamma(l+1)\Gamma(k-l+1)}$$

$$\cdot F(-l, -2\phi + k, k - l + 1; -b\bar{b}/a\bar{a}),$$

if $k \geq l \geq 0$,

$$= \frac{N_l}{N_k} a^k (\bar{a})^{2\phi - l} (b)^{l-k} \frac{\Gamma(2\phi - k + 1)}{\Gamma(2\phi - l + 1)\Gamma(l - k + 1)}$$

$$\cdot F(-k, -2\phi + l, l - k + 1; -b\bar{b}/a\bar{a}),$$

if $l \geq k \geq 0$. (3.6)

We have to note here that the matrix elements (3.6) have the same form as (2.29); however k, $l = 0, 1, 2, 3, \cdots, \infty$.

B. Class III Representations $[D^-(\phi, E_0)]$

Let the spectrum of J^3 be $= [\phi - n; n =$ a non-negative integer$]$. We assume further the following form for the operators $J^\pm, J^3$:

$$J^3 = \phi - z \frac{d}{dz},$$

$$J^+ = -\frac{d}{dz},$$

$$J^- = -2\phi z + z^2 \frac{d}{dz}, \tag{3.7}$$

and $Q = \frac{1}{2}[J_+, J_-] + J_3^2 = \phi(\phi + 1)$.

With respect to the expressions (3.7), we obtain the multiplier representations $T^\rho(g)$ as

$$[T^\rho(g) f](z) = (a - \bar{b}z)^{2\phi} \cdot f\left(\frac{\bar{a}z + b}{a - \bar{b}z}\right). \tag{3.8}$$

(3.8) could be formally obtained from (2.18) by replacing $a \to \bar{a}$ and $b \to \bar{b}$. As usual, we consider the basis $[f_E^{(\phi)}(z)]$ and compute the matrix elements of $J^\pm$ and J^3. For $E = \phi - m$, we have

$$J^+ f_E^{(\phi)}(z) = i[m(m-2\phi-1)]^{1/2} f_{E+1}^{(\phi)}(z),$$
$$J^- f_E^{(\phi)}(z) = i[(m+1)(-2\phi+m)]^{1/2} \times f_{E-1}^{\phi}(z). \quad (3.9)$$

Now, $J^+ f_{\phi-m}^{(\phi)}(z) = 0 \Rightarrow f_{\phi+1}^{(\phi)}(z) = 0$, and $J^- f_{\phi-m}^{(\phi)}(z) = 0$ has no solution unless $m = 2\phi$ (compare finite-dimensional representation case). We thus obtain an infinite set of vectors

$$[f_{\phi-m}^{(\phi)}(z) = f_{\phi}^{(\phi)}(z), f_{\phi-1}^{(\phi)}(z), \cdots$$

spanning the linear complex vector space.

As usual,

$$f_E^{(\phi)}(z) \equiv f_m^{(\phi,E_0)}(z) \equiv N_m^{(\phi)} z^m$$
$$= i^m \left[\frac{\Gamma(m-2\phi)}{\Gamma(m+1)\Gamma(-2\phi)}\right]^{1/2} \cdot z^m, \quad m = 0, 1, 2, \cdots$$

On this basis, the multiplier representations (3.8) are given by

$$[T^{\rho}(g) f_k^{(\phi,E_0)}](z) = \sum_{l=0}^{\infty} D_{lk}(g) f_l^{(\phi,E_0)}(z),$$
$$k \geq 0, g \in O(3),$$

or

$$(-\bar{b}z + a)^{2\phi-k} \cdot (az+b)^k \cdot N_k = \sum_{l=0} D_{lk}(g) z^l \cdot N_l, \; |b/\bar{a} \cdot z| < 1,$$

or

$$D_{lk}^{(\phi,E_0)}(g) = \frac{N_k}{N_l} \cdot (\bar{a})^l \cdot (a)^{2\phi-k} \cdot (b)^{k-l} \cdot \frac{\Gamma(k+1)}{\Gamma(l+1)\Gamma(k-l+1)} \cdot F(-l, -2\phi+k, k-l+1; -|b|^2/|a|^2),$$
$$k \geq l \geq 0. \quad (3.10)$$

C. Continuous Series Representations $D(\phi, E_0)$

As usual, we construct this representation in the space of complex analytic functions $[f_E^{(\phi)}(z)]$ formally given by (2.27) for arbitrary complex values of ϕ and E. Thus,

$$f_m^{(\phi,E_0)}(z) = i^m \left[\frac{\Gamma(E-\phi)}{\Gamma(E+\phi+1)\Gamma(-2\phi)}\right]^{1/2} \cdot z^{E_0+m-\phi}$$
$$\equiv N_m^{(\phi,E_0)} \cdot z^{E_0+m-\phi}, \quad (3.11)$$

and

$$[T^{\rho} f_m^{(\phi,E_0)}](z) = N_m^{(\phi,E_0)} \cdot (\bar{a}+bz)^{\phi-E_0-m} \cdot (az-\bar{b})^{\phi+E_0+m}. \quad (3.12)$$

In (3.11) and (3.12), we have considered $E = E_0 + m$, and $m = 0, \pm 1, \pm 2, \pm 3, \ldots$. We now evaluate the inner product:

$$(f_p^{(\phi,E_0)}(z), f_q^{(\phi,E_0)}(z))_{\phi} = \int \overline{f_p^{(\phi,E_0)}(z)} \cdot f_q^{(\phi,E_0)}(z) d\mu(z)$$

and

$$d\mu(z) = \frac{2\phi+1}{\pi} \cdot (1+|z|^2)^{-2\phi-2} \cdot dz \cdot d\bar{z}. \quad (3.13)$$

Following an identical procedure as for discrete series representations, we obtain

$$(f_p^{(\phi,E_0)}(z), f_q^{(\phi,E_0)}(z))_{\phi} = \delta_{pq} \zeta_q^{(\phi,E_0)} \cdot N_q^{(\phi,E_0)} \cdot B(E_0+q+1, 2\phi+1-E_0-q), \quad (3.14)$$

where

$$\zeta_q^{(\phi,E0)} = \left[\frac{\Gamma(q+\bar{E}_0-\bar{\phi})}{\Gamma(q+\bar{E}_0+\bar{\phi}+1)\cdot\Gamma(-2\bar{\phi})}\right]^{1/2} \cdot \left[\frac{\Gamma(q+Eo-\phi)}{\Gamma(q+Eo+\phi+1)\cdot\Gamma(-2\phi)}\right]^{-1/2}. \quad (3.15)$$

We then evaluate the multiplier representations $T^{\rho}(g)$ in the basis (3, 11):

$$[T^{\rho}(g) f_q^{(\phi,Eo)}](z) = \sum_{p=-\infty}^{+\infty} D_{pq}^{(\phi,Eo)}(g) f_p^{(\phi,Eo)}(z)$$
$$\equiv N_q^{(\phi,Eo)} (\bar{a}+bz)^{\phi-Eo-q} (az-\bar{b})^{\phi+Eo+q}$$
$$q = 0, \pm 1, \pm 2, \ldots \quad (3.16)$$

After a little simplification we obtain

$$D_{pq}^{(\phi,Eo)}(g) = \frac{N_q}{N_p} \cdot a^{s+p} (\bar{a})^{t-q} (-\bar{b})^{q-p} \cdot \frac{\Gamma(s+q+1)}{\Gamma(s+p+1)\cdot\Gamma(q-p+1)} \cdot F(-s-p, -t+q, q-p+1; -b\bar{b}/a\bar{a}), \quad q \geq p$$
$$= \frac{N_p}{N_q} a^{s+q} (\bar{a})^{t-p} (b)^{p-q} \cdot \frac{\Gamma(t-q+1)}{\Gamma(t-p+1)\Gamma(p-q+1)} \cdot F(-s-q, -t+p, p-q+1; -|b|^2/|a|^2), \quad p \geq q. \quad (3.17)$$

Here, $s = \phi + E_0$ and $t = \phi - E_0$.

We have to note further that neither s nor t is an integer. The two expressions can be combined into one since

$$F(-s-p, -t+q, q-p+1; -|b|^2/|a|^2)/\Gamma(q-p+1)$$

is defined even when $(q-p+1)$ is a negative integer. Then using the properties of the hypergeometric functions, we obtain

$$D_{pq}^{(\phi,E0)}(g) = \frac{N_q}{N_p} a^{s+p} (\bar{a})^{t-q} (-\bar{b})^{q-p} \cdot \frac{\Gamma(s+q+1)}{\Gamma(s+p+1)\cdot\Gamma(q-p+1)} \cdot F(-s-p, -t+q, q-p+1; -|b|^2/|a|^2) \quad (3.18)$$

for all integers p, q.

Substituting for N_p and N_q from (3.11), we finally obtain

$$\begin{aligned} D^{(\phi,E0)}_{pq}(g) &= \Theta_{pq}(-1)^{q-p}(\bar{a})^{-(2E0+p+q)}(-\bar{b})^{q-p} \\ &\cdot F(-s-p,-t+q,q-p+1;b\bar{b}), q \geq p, \\ &\equiv \Theta_{pq}(\bar{a})^{-(2E0+p+q)}(-b)^{p-q} \\ &\cdot F(-q-s,-t+p,p-q+1;b\bar{b}), \\ &p \geq q, \end{aligned} \tag{3.19}$$

where

$$\Theta_{pq} = (i)^{q-p}\frac{1}{(q-p)!}\left[\frac{\Gamma(p+s+1)\Gamma(p-t}{\Gamma(q+s+1)\cdot\Gamma(q-t)}\right]. \tag{3.20}$$

We have to note here that the matrices $D^{(\phi,E0)}_{pq}(g)$ furnish an infinite-dimensional nonunitary representation for $O(3)$. In place of unitarity, we have

$$\zeta[D(g)]^{+}\zeta^{-1} = D(g^{-1}), \tag{3.21}$$

where ζ is the metric given by (3.15). Thus,

$$\sum_{\gamma} D_{p\gamma}(g)D_{\gamma q}(g^{-1}) \cong \delta_{pq}. \tag{3.22}$$

4. THE HYDROGEN ATOM

We will briefly outline the consequences of our new classes of representations in the context of the symmetry of the two dimensional hydrogen atom. A more realistic case of this problem will be published elsewhere.

The Hamiltonian H of the system is

$$H = \frac{1}{2\mu}(p_1^2 + p_2^2) - \frac{e^2}{\gamma}, \quad V(\gamma) = -\frac{e^2}{\gamma}. \tag{4.1}$$

The constants of the motion are

$$\begin{aligned} L_3 &= xp_2 - yp_1, \\ A_1 &= -1/2\mu e^2(L_3p_2 + p_2L_3) + x/r, \\ A_2 &= +(\tfrac{1}{2}\mu e^2)(L_3p_1 + p_1L_3) + y/r, \end{aligned} \tag{4.2}$$

satisfying the following commutation relations:

$$\begin{aligned} &[L_3,A_1] = iA_2, \quad [L_3,A_2] = -iA_1 \\ &[A_1,A_2] = \frac{i}{\mu e^4}(-2H)L_3. \end{aligned} \tag{4.3}$$

Since (4.2) do not close the algebra, we define a new set of operators as

$$\begin{aligned} J_1 &= \frac{\mu^{1/2}e^2}{(-2H)^{1/2}}\cdot A_1, \\ J_2 &= \frac{\mu^{1/2}e^2}{(-2H)^{1/2}}\cdot A_2, \\ J_3 &= L_3. \end{aligned} \tag{4.4}$$

It is easy to check that

$$[J_i,J_j] = i\,\epsilon_{ijk}J_k, \quad i,j,k = 1,2,3. \tag{4.5}$$

We now compute the Casimir operator Q:

$$Q = \tfrac{1}{2}(J_+,J_-) + J_3^2. \tag{4.6}$$

Substituting (4.2) and (4.4) in (4.6), we obtain, after a little algebraic manipulation,

$$\begin{aligned} Q &= \frac{\mu e^4}{(-2H)}\left[1 + \frac{2}{\mu e^4}HJ_3^2 + \frac{1}{2\mu e^4}\cdot H\right] + J_3^2 \\ &= -\frac{\mu e^4}{2H} - \tfrac{1}{4}. \end{aligned} \tag{4.7}$$

Solving for the eigenvalues, we obtain from (4.7)

$$\phi(\phi+1) = -\frac{\mu e^4}{2E} - \frac{1}{4}$$

or

$$E = \frac{-\mu e^4}{2[\phi(\phi+1)+\frac{1}{4}]} = \frac{-\mu e^4}{2\cdot[\phi+\frac{1}{2}]^2}. \tag{4.8}$$

Substituting $n = \phi + 1$, we get (in natural units, $\hbar \neq 1$),

$$E = \frac{-2\pi^2\mu e^4}{h^2(\phi+\frac{1}{2})^2} = \frac{-2\pi^2\mu e^4}{h^2(n-\frac{1}{2})^2}. \tag{4.9}$$

This is exactly the expression obtained by Jauch and Hill[3] for the bound-state energy spectrum of the two-dimensional hydrogen problem, when $n = 0,1,2,3,\ldots$

We have seen that for ϕ real, or, in other words for $n = 0,1,2,3,\ldots,E$ is aways negative and the symmetry of the system is $O(3)$. We make no secret of the fact that when ϕ is complex, namely for $\phi = -\frac{1}{2} + i\sigma, 0 < \sigma^2 \leq \infty$, ("continuous series representation"), we obtain,

$$E = 2\pi^2\mu e^4/h^2\sigma^2 > 0. \tag{4.10}$$

(4.10) gives the energy expression for the scattering state. We just want to emphasize here that, in case of (4.10), the underlying vector space is a metric space, whereas the bound-state spectrum (4.9) is obtained in Hilbert space.

5. TOPOLOGY OF H

We have demonstrated that for the discrete and continuous series representations of $O(3)$, the scalar product of the vectors $f(z)$ is highly indefinite. To obtain a positive scalar product on the space H of functions $f(z)$, let us define

$$(f,f')_{\phi} = (f_+,f'_+) - (f_-,f'_-), \tag{5.1}$$

where $f = f_+ + f_-$ and $f_{\pm}, f'_{\pm} \in H_{\pm}$.

We have split up the vector $f \in H$ into components $f_{\pm}$ belonging to the subspaces $H_{\pm}$ on which the metric ζ has a definite sign. The topology of this space satisfies the following properties (Hilbert topology):

(i) The norm $\|f\| = (f,f)^{1/2}$.

(ii) The geometry is given by the distance
$d(f,f') = \|f - f'\|$.

(iii) A sequence of vectors $f_\nu \to f$ if
$\lim \nu \to \infty \|f_\nu - f\| = \lim d(f_\nu, f) = 0$.

Alternatively, we can introduce Frechet's topology on the vector space H having the following properties.

For any pair of points f,f', the distance $d(f,f')$ satisfies

(i) $d(f,f') = d(f-f',0)$;

(ii) If $f_\nu \to 0$, then $d(f_\nu, 0) \to 0$. (5.2)

Indeed, we can set

$$d(f,f') = \sum_{p=1}^{\infty} \frac{\|f-f'\|_p}{2^p(1+\|f-f'\|_p)}, \tag{5.3}$$

where $\|\times\|_p$ is a denumerable set of seminorms. We can easily check that (5.3) obviously satisfies (5.2).

6. CONCLUSION

To summarize our discussion, we analyzed the possible "local representations" of $O(3)$ and demonstrated that in addition to the finite-dimensional unitary representations, there are infinite-dimensional representations realized on a more general linear vector space, rather than the Hilbert space. We briefly discussed the consequences of the continuous series representation in describing the scattering states of the two-dimensional hydrogen problem. This could be also used to construct multiparticle scattering amplitudes with a more general transformation (generalized Sommerfeld-Watson transformation) on the complex angular momentum plane. The formal correspondence with $0(2,1)$ could be established by replacing $\beta \to i\beta$ (the Eulerian angle). Under this case, in fact, we reproduce the Master analytic representations of $0(2,1)$ by Kuriyan et al.[10] We have also briefly outlined the possible geometry that could be associated with the vector space. Our method is general, and could as well be applied to study the infinite-dimensional representations of $0(n)$. The richness of the new series of representations could be useful to build up infinite-multiplet scheme for the ever-growing high spin particles.

ACKNOWLEDGMENTS

The author wishes to thank Professor J. Nilsson and Professor A. Kihlberg for many constructive criticisms and useful discussions. Thanks are due to Dr. J. Nagel and Dr. L. Brink for interesting conversations and to Mrs. S. Tripathy for a careful reading of the manuscript and for her constant encouragement. He also acknowledges Professor N. Svartholm and Professor J. Nilsson for their kind hospitality at the Institute of Theoretical Physics, Göteborg, Sweden.

---

* On leave of absence, Department of Physics, University of Delhi. Delhi-7, India.

[1] J. Nilsson and A. Kihlberg, "Group Theoretic Approach to Wavefunctions and Field," Göteborg (Report of work before publication) August, 1968; H. Bacry and A. Kihlberg, J. Math. Phys. **10**, 2132 (1969); G. Fuchs, CNRS, Paris (Report of work before publication), March, 1970.

[2a] C. Fronsdal, Phys. Rev. **156**, 1653 (1967); **171**, 1811 (1968); **182**, 1564 (1969); **185**, 1768 (1969); D. Stoyanov and I. T. Todorov, J. Math. Phys. **9**, 2146 (1968); K. C. Tripathy, *ibid.* **11**, 1901 (1970); and Phys. Rev. D **2**, 2955 (1970); E. C. G. Sudarshan and N. Mukunda, *ibid.* 571 (1970).

[2b] For the algebraic aspect of the infinite-multiplet scheme, see e.g., A. O. Barut, "Rapporteur's talk given at XV-th International Conference on High Energy Physics," Kiev (September 1970).

[3] J. M. Jauch and E. L. Hill, Phys. Rev. **57**, 641 (1970); S. P. Alliluev, Zh. Eksp. Teor. Fiz. **33**, 200 (1957) [Sov. Phys. JETP **6**, 157 (1958)]; A Cisneros and H. V. McIntosh, J. Math. Phys. **10**, 277 (1969).

[4] A. O. Barut (private communication).

[5] M. Andrews and J. Gunson, J. Math. Phys. **5**, 1391 (1964); J. Gunson, *ibid.* **6**, 852 (1965); S. Sannikov, Nucl. Phys. **87**, 834 (1967).

[6] W. Miller, *Lie Theory and Special Functions* (Academic, New York, 1968).

[7] A. O. Barut and C. Fronsdal, Proc. Roy. Soc. (London) **A287**, 532 (1965).

[8] I. M. Gel'fand and G. E. Shilov, *Generalized Functions* (Academic, New York, 1968), Vol. 2.

[9] A. R. Edmunds, *Angular momentum in Quantum Mechanics* (Princeton U.P., Princeton, N.J., 1957).

[10] J. G. Kuriyan, N. Mukunda, and E. C. G. Sudarshan, Commun. Math. Phys. **8**, 204 (1968).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 10 OCTOBER 1971

An Algebraic Realization of Spin-Wave Theory

C. J. Liu and Yutze Chow
Department of Physics, University of Wisconsin–Milwaukee, Milwaukee, Wisconsin 53201
(Received 15 March 1971)

The Heisenberg Hamiltonian of a ferromagnet with Zeeman term is treated by the method of algebraic realization of the spin algebra. The general features of this approach are first analyzed. A detailed calculation with a particular algebraic realization is then carried out by use of two boson operators; one of the operators is interpreted as the observable spin waves and the other as "spurions" which carry only spin quantum number without energy. The Bethe–Salpeter integral equation is solved in the long-wavelength region for the bound states of spin waves. A perturbation calculation yields a leading T^4 correction to the usual $T^{3/2}$ behavior of the magnetization.

1. INTRODUCTION

In the traditional way of handling many-body problems, the given Hamiltonians are usually the starting point of an *ab initio* calculation in the sense that the original Hamiltonians are *directly* responsible for the final results. However, it is our opinion that this is sometimes too restrictive since most Hamiltonians of physical models for many-body problems are not simple enough to yield solutions without some rather drastic approximations or mathematically doubtful operations. Therefore, it is physically interesting while mathematically not necessarily worse, to consider the initial Hamiltonians as *abstract* operator expressions capable of being *realized* by certain operators satisfying simple algebraic relations. These operators used for carrying out the algebraic realizations can be given a "particle" interpretation and, for convenience, can be called "semiparticles" here, when the operators satisfy simple commutation or anticommutation relations. More precisely, the criterion of establishing a "semiparticle" interpretation can be formulated in the following way. If the operators used for a particular realization enable us either to solve the eigenvalue problem of the given Hamiltonian or to derive an *effective* Hamiltonian of a simple form such that the problem becomes more manageable, then these operators are called "semiparticles" if they satisfy commutation or anticommutation relations, or even parastatistics. Such a criterion does not require a perfect diagonalization of the original Hamiltonian by the semiparticle operators, i.e., the semiparticles in our definition may have *interactions*. A perfect diagonalization which makes semiparticles *free* is therefore only a fortunate situation; in this case semiparticles are just quasiparticles in the usual language. The other feature of our approach is the commitment to identify the original operators as simple combinations of *finite* products of the new operators (i.e., the semiparticle operators), or possibly in closed forms. This differs from the usual methodology of expanding Heisenberg operators in terms of an irreducible ring of asymptotic operators. The latter expansion is, in general, an infinite series whose convergence may often be questionable. Besides, the most important problem lies in the truncation of such infinite series in order to make the calculation manageable. Any truncation, in fact, imposes conditions that must be satisfied by the original operators. The introduction of such new auxiliary conditions can, in general, be either inconsistent or bad enough to allow only trivial solutions. This point is traditionally ignored in the expansion of a Heisenberg operator into an infinite series. Therefore, the inconsistency may be hidden and thus completely overlooked. The method of algebraic realization usually has the advantage of having rather simple subsidiary conditions whose consistency can be readily verified.

Finally, one must ask the question of how to choose the *right* realization among many possible algebraic realizations. This question may be answered in the following tentative way. It is our opinion that an adequate realization is most probably a simple one since otherwise the advantage of an algebraic realization may be lost. Next, the auxiliary condition resulting from the particular realization should be both mathematically and physically consistent. In other words, a realization should be discarded if the auxiliary conditions resulting from the realization are inconsistent either on mathematical or on physical grounds. Furthermore, the realization should yield either a *simple* form of Hamiltonian or a Hamiltonian that has a *simple effective form*. The new operators introduced for algebraic realization should satisfy commutation or anticommutation relations or parastatistics. However, this last point is not absolutely necessary, if one does not insist on interpretations of semiparticles in terms of bosons, fermions, or paraparticles. Last but not least is the expectation that an adequate realization would lead to some physically *reasonable* results from the computation. In our study of the spin-wave theory with Heisenberg-type Hamiltonian we find the above set of criteria leaves very little freedom in choosing a realization. We succeeded in finding only *one*, the one that is discussed in this paper. This particular realization indeed gives a satisfactory result.

We first analyze the two most important papers in spin-wave theory, the work of Holstein–Primakoff[1] and that of Dyson,[2] from the viewpoint of algebraic realizations. By considering a different realization we rewrite the Hamiltonian in terms of two new operators. These new operators are boson fields. It is shown that one of them is, in fact, a "Goldstone particle," which carries no energy and can be interpreted as the existence of "condensation." The other boson operator can be interpreted as the existence of "spin waves" or "magnons." In other words, we have two kinds of semiparticles: One

takes care of the bookkeeping of the spin-projection quantum number and the other takes care of the spin-wave dynamics. The resulting Hamiltonian actually does not have a simple form. However, by means of the energy eigenvalues, we can write down an effective Hamiltonian which is of a much simpler form.

For long wavelengths, we show that there exist bound states of spin waves. Further, the Bethe-Salpeter integral equation leads to a relation between the spin-wave energy and the bound-state binding energy of two spin waves.

2. ON THE HOLSTEIN-PRIMAKOFF AND DYSON APPROACHES

In this section we consider the classic papers of Holstein-Primakoff and Dyson from the viewpoint of algebraic realizations. The Hamiltonian to be considered here is the Heisenberg model of ferromagnetism:

$$\mathcal{H} = -\sum_{l,m} J_{lm}\mathbf{S}_l \cdot \mathbf{S}_m - H\sum_l S_l^{(z)}, \tag{2.1}$$

where $\mathbf{S}_l$ is the spin operator of the atom at the lth site, J_{lm} is the exchange integral between atoms at the lth and mth sites, and H is the externally applied magnetic field. The spin operators satisfy the algebraic relation

$$[S_l^{(+)}, S_m^{(-)}] = 2\delta_{lm}S_l^{(z)}, \qquad [S_l^{(z)}, S_m^{(\pm)}] = \pm\,\delta_{lm}S_l^{(\pm)}, \tag{2.2}$$

where $S_l^{(+)} \equiv S_l^{(x)} + iS_l^{(y)}$; $S_l^{(-)} \equiv S_l^{(x)} - iS_l^{(y)}$; and $S_l^{(x)}, S_l^{(y)}, S_l^{(z)}$ are x, y, z components of the spin at the site l.

Holstein-Primakoff actually started from a more complicated Hamiltonian than Eq. (2.1); their Hamiltonian includes the dipolar and pseudodipolar terms. However, their approach is equivalent to the following choice of the realization of the algebra, Eq. (2.2):

$$\begin{aligned} S_l^{(-)} &\equiv (2s)^{1/2}a_l^\dagger(1-(1/2s)a_l^\dagger a_l)^{1/2},\\ S_l^{(+)} &\equiv (2s)^{1/2}(1-(1/2s)a_l^\dagger a_l)^{1/2}a_l,\\ S_l^{(z)} &\equiv s - a_l^\dagger a_l, \end{aligned} \tag{2.3}$$

where a_l and $a_l^\dagger$ form the commutator algebra:

$$[a_l, a_m^+] = \delta_{lm}, \qquad [a_l, a_m] = 0. \tag{2.4}$$

Under the realization, Eq. (2.3), the Heisenberg Hamiltonian of Eq. (2.1) splits into two parts: a *quadratic* part and a *nonquadratic* part containing an essentially *infinite formal series* in $a_l^\dagger a_l$ [due to the square roots appearing in Eq. (2.3)]. The quadratic part gives a result identical to Bloch's linear approximation. The other nonquadratic part is then interpreted by Holstein-Primakoff as scatterings of spin waves. Thus, if the nonquadratic part could be neglected, then the leading term of the Holstein-Primakoff result would agree with the experimental situation just as Bloch's[3] does. There was virtually no significant theoretical progress in this direction from then on until the appearance of Dyson's two papers in 1956. Dyson criticized the work of Holstein-Primakoff by pointing out that their nonquadratic term was actually too large to be neglected, so that the calculation was not really too useful. He attributed this to their improper choice of state vectors in the calculation and he proposed to use a different set of state vectors, though *nonorthogonal* among other things, to span the Hilbert space. His procedure consists of constructing, by means of his state vectors, an equivalent Hamiltonian from which his calculation proceeded. His result leads to a T^4 modification of the magnetization. Besides, he argued that there could not be any bound states in either two- or three-dimensional cases while Bethe[4] has shown the existence of bound states in the one-dimensional chain. This aspect of Dyson's paper was however shown to be false by Wortis.[5] Our results show that not only bound states do exist, but, in fact, two-particle scattering states can be ignored in the long-wavelength range. This is rather different from the conclusion reached by Wortis, who showed the presence of a threshold below which the bound states appear. Dyson's calculation can be considered as a result of the algebraic realization of the algebra:[6]

$$\begin{aligned} S_l^{(-)} &\equiv (2s)^{1/2}a_l^\dagger,\\ S_l^{(+)} &\equiv (2s)^{1/2}[1-(1/2s)a_l^\dagger a_l]a_l,\\ S_l^{(z)} &\equiv s - a_l^\dagger a_l. \end{aligned} \tag{2.5}$$

At this point, we note the interesting fact that the realizations corresponding to the Holstein-Primakoff and Dyson solutions both satisfy the relation

$$S_l^2 = s(s+1), \tag{2.6}$$

which is the usual eigenvalue equation if Eq. (2.6) is applied to an eigenstate of $\mathbf{S}^2$. Thus Eq. (2.6) should be considered as an "auxiliary condition" resulting from the particular realization. In other words, Eq. (2.6) is valid when applied to an eigenstate of $\mathbf{S}^2$. This is not at all a drawback; because the eigenvalue s is fixed for a given problem and because the Hamiltonian is expected to commute with $\mathbf{S}^2$ on physical grounds, it is natural to use simultaneous eigenstates of $\mathcal{H}$ and $\mathbf{S}^2$. Under these circumstances Eq. (2.6) is consistent if one keeps the preceding choice of eigenstates in mind. Since every realization will result in some sort of condition (or conditions) like Eq. (2.6), it therefore provides a useful guide in the selection of realizations (i.e., the realization should be discarded if the condition is inconsistent on physical or mathematical grounds) and it also tells what kind of eigenstate one must employ in the particular calculation. From the purely algebraic viewpoint, a drawback of the realization corresponding to the

Holstein–Primakoff solution is the appearance of square roots of operators which have to be interpreted mathematically as *formal* infinite series of operators. On the other hand, the realization corresponding to the Dyson's solution forces $S_l^{(+)}$ and $S_l^{(-)}$ is not to be related by Hermitian conjugation; this is also an uncomfortable situation since the resulting Hamiltonian is no longer Hermitian. Therefore, the question is whether one can find a simple enough realization, in terms of commutator or anticommutator algebras, if possible, of Eq. (2.2) without the drawbacks (either mathematically or physically) of the Holstein–Primakoff and Dyson approaches.

3. AN ALGEBRAIC REALIZATION AND SEMIPARTICLE INTERPRETATION

Our calculation in the rest of the paper is based upon the following realization in terms of two boson operators:

$$\begin{aligned} S_l^{(+)} &= \beta_l^\dagger b_l, \\ S_l^{(-)} &= b_l^\dagger \beta_l, \\ S_l^{(z)} &= \tfrac{1}{2}(\beta_l^\dagger \beta_l - b_l^\dagger b_l), \end{aligned} \tag{3.1}$$

with

$$\begin{aligned} &[\beta_l, \beta_m^\dagger] = \delta_{lm}, \quad [\beta_l, \beta_m] = 0, \\ &[b_l, b_m^\dagger] = \delta_{lm}, \quad [b_l, b_m] = 0, \\ &[\beta_l, b_m^\dagger] = 0, \quad [\beta_l, b_m] = 0. \end{aligned} \tag{3.2}$$

It is important to note that if the operators β and b are subjected to the transformation of a *c-number* addition,

$$\beta_l \to \beta_l + c, \quad b_l \to b_l + c', \tag{3.3}$$

the algebraic relations Eq. (2.2), are still satisfied. We shall make use of this property in what follows.

It is also easy to verify that the realization of Eq. (3.1) yields the following auxiliary condition:

$$\beta_l^\dagger \beta_l + b_l^\dagger b_l = 2s, \tag{3.4}$$

by which we can write $S_l^{(z)}$ into

$$S_l^{(z)} = s - b_l^\dagger b_l = -s + \beta_l^\dagger \beta_l. \tag{3.5}$$

In virtue of Eq. (3.3), we consider the following transformation:

$$\beta_l \to (2s)^{1/2} + \beta_l, \quad b_l \to b_l \tag{3.6}$$

and define the Fourier transform for β_l and b_l by

$$\beta_k = \frac{1}{\sqrt{N}} \sum_l e^{i\mathbf{k}\cdot\mathbf{l}} \beta_l, \tag{3.7}$$

$$b_k = \frac{1}{\sqrt{N}} \sum_l e^{i\mathbf{k}\cdot\mathbf{l}} b_l. \tag{3.8}$$

The Hamiltonian of Eq. (2.1) now becomes

$$\begin{aligned} \mathfrak{K} = E_0 &+ \sum_{\mathbf{k}} \{H + 2s[J(0) - J(\mathbf{k})]\} b_k^\dagger b_k \\ &- \left(\frac{2s}{N}\right)^{1/2} \sum_{\mathbf{k}_1 \mathbf{k}_2} J(k_1)(b_{k_1}^\dagger b_{k_1+k_2} \beta_{k_2}^\dagger + b_{k_1+k_2}^\dagger b_{k_1} \beta_{k_2}) \\ &- \frac{1}{N} \sum_{\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3} J(\mathbf{k}_1) b_{k_2+k_1}^\dagger \beta_{k_3-k_1}^\dagger b_{k_3} \beta_{k_2} \\ &- \frac{1}{N} \sum_{\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3} J(\mathbf{k}_1) b_{k_2+k_1}^\dagger b_{k_3-k_1}^\dagger b_{k_2} b_{k_3}, \end{aligned} \tag{3.9}$$

where

$$E_0 = -NSH - NS^2 J(0). \tag{3.10}$$

At this point, it is reasonable to expect particle interpretations for both operators β and b since they satisfy commutation relations. They are called *semiparticles*, whose physical properties are now to be analyzed. First, from the fact that there is no quadratic term for β particle in $\mathfrak{K}$ it is obvious that

$$\mathfrak{K} \beta_{k_1}^\dagger \cdots \beta_{k_n}^\dagger |0\rangle = E_0 \beta_{k_1}^\dagger \cdots \beta_{k_n}^\dagger |0\rangle \tag{3.11}$$

for arbitrary n. In other words, the semiparticle β actually carries no energy. Hence the β particle can simply be considered as a "spurion" which does nothing besides shuffling the spin-quantum numbers. It is clear that

$$b_{k_1}^\dagger \cdots b_{k_m}^\dagger |0\rangle, \quad m = 1, 2, \ldots \tag{3.12}$$

are not eigenvectors of $\mathfrak{K}$. Since the number operators of the b particles

$$\mathfrak{N}^{(b)} \equiv \sum_{\mathbf{p}} b_p^\dagger b_p \tag{3.13}$$

commute with $\mathfrak{K}$, the eigenvalues of $\mathfrak{K}$ and $\mathfrak{N}^{(b)}$ can be used as simultaneous labels for state vectors. Let us denote by $|\mathbf{p}_1, \cdots, \mathbf{p}_m\rangle$ a simultaneous eigenvector of $\mathfrak{K}$ and $\mathfrak{N}^{(b)}$ with b particles of momenta $\mathbf{p}_1, \ldots, \mathbf{p}_m$. Next, let us evaluate the expansion:

$$\begin{aligned} |\mathbf{p}_1, \ldots, \mathbf{p}_m\rangle \equiv \sum_{n=0}^{2SN} \sum_{\mathbf{q}_1 \cdots \mathbf{q}_m} C^{(n)}(\mathbf{p}_1 \cdots \mathbf{p}_m; \mathbf{q}_1 \cdots \mathbf{q}_m) \\ \times [1/(m!n!)^{1/2}](\beta_0^\dagger)^n b_{q_1}^\dagger \cdots b_{q_m}^\dagger |0\rangle, \end{aligned} \tag{3.14}$$

where the amplitude $C^{(n)}(\mathbf{p}_1 \cdots \mathbf{p}_m; \mathbf{q}_1 \cdots \mathbf{q}_m)$, to be determined, is the probability of finding m b particles of momenta $\mathbf{q}_1 \cdots \mathbf{q}_m$ together with n β particles of zero momentum. By orthonormality, Eq. (3.14) yields

$$\begin{aligned} C^{(n)}(\mathbf{p}_1 \cdots \mathbf{p}_m; \mathbf{q}_1 \cdots \mathbf{q}_m) \\ = [1/(m!n!)^{1/2}] \langle 0 | (\beta_0)^n b_{q_1} \cdots b_{q_m} |\mathbf{p}_1, \ldots, \mathbf{p}_m\rangle. \end{aligned} \tag{3.15}$$

Now, if $|\mathbf{p}_1,\ldots,\mathbf{p}_m\rangle$ is an eigenvector of $\mathcal{H}$ with eigenvalue $J_m(\mathbf{p}_1\cdots\mathbf{p}_m)$, i.e.,

$$\mathcal{H}|\mathbf{p}_1,\ldots,\mathbf{p}_m\rangle = J_m(\mathbf{p}_1\cdots\mathbf{p}_m)|\mathbf{p}_1,\ldots,\mathbf{p}_m\rangle, \quad (3.16)$$

then $C^{(n)}(\mathbf{p}_1\cdots\mathbf{p}_m;\mathbf{q}_1\cdots\mathbf{q}_m)$ has to satisfy the following *difference* equation:

$$\begin{aligned}J_m(\mathbf{p}_1\cdots\mathbf{p}_m)C^{(n)}(\mathbf{p}_1\cdots\mathbf{p}_m;\mathbf{q}_1\cdots\mathbf{q}_m) &= \left(E_0+\sum_{i=1}^{m}\{H+2s[J(0)-J(\mathbf{q}_i)]\}\right)C^{(n)}(\mathbf{p}_1\cdots\mathbf{p}_m;\mathbf{q}_1\cdots\mathbf{q}_m)\\ &-\left(\frac{2sn}{N}\right)^{1/2}\sum_{i=1}^{m}J(\mathbf{q}_i)C^{(n-1)}(\mathbf{p}_1\cdots\mathbf{p}_m;\mathbf{q}_1\cdots\mathbf{q}_m)-\left(\frac{2s(n+1)}{N}\right)^{1/2}\sum_{i=1}^{m}J(\mathbf{q}_i)C^{(n+1)}(\mathbf{p}_1\cdots\mathbf{p}_m;\mathbf{q}_1\cdots\mathbf{q}_m)\\ &-\frac{n}{N}\sum_{i=1}^{m}J(\mathbf{q}_i)C^{(n)}(\mathbf{p}_1\cdots\mathbf{p}_m;\mathbf{q}_1\cdots\mathbf{q}_m)-\frac{2}{N}\sum_{i=1}^{m}\sum_{j=1}^{i-1}\sum_{\mathbf{k}}J(\mathbf{k})\,C^{(n)}(\mathbf{p}_1\cdots\mathbf{p}_m;\mathbf{q}_1\cdots\mathbf{q}_{j-1},\mathbf{q}_j-\mathbf{k},\\ &\mathbf{q}_{j+1},\ldots,\mathbf{q}_i,\mathbf{q}_i+\mathbf{k},\mathbf{q}_{i+1},\ldots,\mathbf{q}_m). \end{aligned} \quad (3.17)$$

To solve for $C^{(n)}$ let us use the ansatz

$$C^{(n)}(\mathbf{p}_1\cdots\mathbf{p}_m;\mathbf{q}_1\cdots\mathbf{q}_m)=\left(-\frac{1}{(2SN)^{1/2}}\right)^n\frac{1}{\sqrt{n!}}F(n)W(\mathbf{p}_1\cdots\mathbf{p}_m;\mathbf{q}_1\cdots\mathbf{q}_m)\delta^{(3)}\left(\sum_{i=1}^{m}(\mathbf{p}_i-\mathbf{q}_i)\right). \quad (3.18)$$

The δ function in Eq. (3.18) corresponds to momentum conservation.

If one chooses $F(n)$ so that

$$F(n+2)-(n+1)F(n+1)+2SN(n+1)F(n)=0, \quad (3.19)$$

then Eq. (3.17) becomes

$$\begin{aligned}J_m(\mathbf{p}_1\cdots\mathbf{p}_m)W(\mathbf{p}_1\cdots\mathbf{p}_m;\mathbf{q}_1\cdots\mathbf{q}_m) &= \left(E_0+\sum_{i=1}^{m}\{H+2s[J(0)-J(\mathbf{q}_i)]\}\right)W(\mathbf{p}_1\ldots\mathbf{p}_m;\mathbf{q}_1\ldots\mathbf{q}_m)\\ &-\frac{2}{N}\sum_{i=1}^{m}\sum_{j=1}^{i-1}\sum_{\mathbf{k}}J(\mathbf{k})W(\mathbf{p}_1\cdots\mathbf{p}_m;\mathbf{q}_1\cdots\mathbf{q}_{j-1},\mathbf{q}_j-\mathbf{k},\mathbf{q}_{j+1},\ldots,\mathbf{q}_{i-1},\mathbf{q}_i+\mathbf{k},\mathbf{q}_{i+1}\cdots\mathbf{q}_m). \end{aligned} \quad (3.20)$$

Since Eq. (3.20) is entirely equivalent to the eigenvalue problem of the following Hamiltonian

$$\mathcal{H}_{eff}=E_0+\sum_{\mathbf{p}}\{H+2s[J(0)-J(\mathbf{p})]\}b_p^+b_p-\frac{1}{N}\sum_{\mathbf{p}_1\mathbf{p}_2\mathbf{p}_3}J(\mathbf{p}_1)b^+_{\mathbf{p}_2+\mathbf{p}_1}b^+_{\mathbf{p}_3-\mathbf{p}_1}b_{\mathbf{p}_2}b_{\mathbf{p}_3}, \quad (3.21)$$

we shall use the above $\mathcal{H}_{eff}$ as our "effective Hamiltonian" for carrying out the rest of the computation. We now define the eigenvector $|\mathbf{p}_1,\ldots,\mathbf{p}_m\rangle'$ by the following expansion

$$|\mathbf{p}_1,\ldots,\mathbf{p}_m\rangle' \equiv \sum_{\mathbf{q}_1\cdots\mathbf{q}_m}\frac{1}{\sqrt{m!}}W(\mathbf{p}_1\cdots\mathbf{p}_m;\mathbf{q}_1\cdots\mathbf{q}_m)\delta^{(3)}\left[\sum_{i=1}^{m}(\mathbf{p}_i-\mathbf{q}_i)\right]b^+_{\mathbf{q}_1}\cdots b^+_{\mathbf{q}_m}|0\rangle', \quad (3.22)$$

where

$$|O\rangle' \equiv \sqrt{R}\sum_{n=0}^{2SN}\left(-\frac{1}{(2SN)^{1/2}}\right)^n\frac{1}{n!}F(n)(\beta_0^+)^n|0\rangle. \quad (3.23)$$

By the normalization condition $\langle 0|0\rangle' = 1$, we have

$$R^{-1}=\sum_{n=0}^{2SN}\left(\frac{1}{2SN}\right)^n\frac{1}{n!}F^2(n). \quad (3.24)$$

It is an important feature of the effective Hamiltonian $\mathcal{H}_{eff}$ that the β particles are no longer present; they "dressed up" the b particles. Consequently, the dynamic properties of spin waves are entirely due to the b particles which are the only semiparticles in the Hamiltonian $\mathcal{H}_{eff}$. Thus, from now on we shall call b particles "spin waves" or "magnons."

Since the operators β and b both satisfy the same algebraic relation, namely the commutation relation, they can be considered as a *doublet* like the proton and neutron in nuclear physics.[7] In this case, one can interpret the operator β as the *condensed* particle while b is the actual spin wave observed physically.[8] The *doublet* feature can be made more elegant by introducing the notation

$$\Phi_l \equiv \begin{pmatrix}\beta_l\\ b_l\end{pmatrix}, \quad (3.25)$$

then Eq. (3.1) becomes

$$\begin{aligned}S_l^{(+)} &= \Phi_l^\dagger\sigma_+\Phi_l,\\ S_l^{(-)} &= \Phi_l^\dagger\sigma_-\Phi_l,\\ S_l^{(z)} &= \Phi_l^\dagger\sigma_z\Phi_l,\end{aligned} \quad (3.26)$$

with

$$\sigma_+ = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \sigma_- = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad \sigma_z = \tfrac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{3.27}$$

Thus, the doublet formalism yields a symmetric form which is very appealing.

4. THE BOUND STATES AND THE BETHE–SALPETER EQUATION

In the long-wavelength region, i.e., $ka \ll 1$ (a = the lattice constant), the effective Hamiltonian becomes

$$\mathcal{H} = E_0 + \sum_{\mathbf{k}} \left(H + \frac{1}{2m}\mathbf{k}^2 \right) b_k^\dagger b_k - \frac{J(0)}{N} \sum_{\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3} b_{k_2+k_1}^\dagger b_{k_3-k_1}^\dagger b_{k_2} b_{k_3}, \tag{4.1}$$

where

$$1/m \equiv 4SJa^2. \tag{4.2}$$

(For simplicity we consider only the simple cubic case.) Defining the Fourier transform:

$$\psi(\mathbf{x}, t) = \frac{1}{\sqrt{N}} \sum_{\mathbf{k}} b_k(t) e^{i\mathbf{k}\cdot\mathbf{x}} \tag{4.3}$$

therefore,

$$[\psi(\mathbf{x}, t), \psi^+(\mathbf{x}', t)] = \delta(\mathbf{x} - \mathbf{x}'),$$
$$\text{and} \quad [\psi(\mathbf{x}, t), \psi(\mathbf{x}', t)] = 0. \tag{4.4}$$

The equation of motion, in terms of $\psi(\mathbf{x}, t)$, is

$$\left(H - \frac{1}{2m}\nabla^2 + \frac{1}{i}\frac{\partial}{\partial t} \right) \psi(x) = 2J(0)\psi^+(x)\psi(x)\psi(x). \tag{4.5}$$

Define now the Bethe–Salpeter wavefunction:[9]

$$\chi_{\mathbf{k}E\alpha}(x_1, x_2) \equiv \langle 0 | T(\psi(x_1)\psi(x_2)) | \mathbf{k}, E, \alpha \rangle, \tag{4.6}$$

where $T(\psi(x_1)\psi(x_2))$ is the time-ordered product of $\psi(x_1)$ and $\psi(x_2)$. From the equation of motion and the commutation relations of ψ, it is straightforward to obtain the Bethe–Salpeter equation:

$$\left(H - \frac{1}{2m}\nabla_2^2 + \frac{1}{i}\frac{\partial}{\partial t_2} \right) \times \left(H - \frac{1}{2m}\nabla_1^2 + \frac{1}{i}\frac{\partial}{\partial t_1} \right) \chi_{\mathbf{k}E\alpha}(x_1, x_2) = i2J(0)\delta(t_1 - t_2)\delta^{(3)}(\mathbf{x}_1 - \mathbf{x}_2)\chi_{\mathbf{k}E\alpha}(x_1, x_2). \tag{4.7}$$

In order to write Eq. (4.7) as an integral equation, let us introduce the propagator function

$$S_F(x_1 - x_2) \equiv \langle 0 | T(\psi(x_1)\psi^+(x_2)) | 0 \rangle. \tag{4.8}$$

By Eq. (4.5) we obtain the equation for the Green's function:

$$\left(H - \frac{1}{2m}\nabla_1^2 + \frac{1}{i}\frac{\partial}{\partial t_1} \right) S_F(x_1 - x_2) = -i\delta(t_1 - t_2)\delta^{(3)}(\mathbf{x}_1 - \mathbf{x}_2). \tag{4.9}$$

Equations (4.7) and (4.9) lead to the Bethe–Salpeter integral equation:

$$\chi_{\mathbf{k}E\alpha}(x_1, x_2) = \chi^0_{\mathbf{k}E\alpha}(x_1, x_2) - i2J(0)\sum_{\mathbf{y}} \int dt_y S_F(x_1 - y) \times \chi_{\mathbf{k}E\alpha}(y, y) S_F(y - x_2), \tag{4.10}$$

where the inhomogeneous term satisfies

$$\left(H - \frac{1}{2m}\nabla_1^2 + \frac{1}{i}\frac{\partial}{\partial t_1} \right) \left(H - \frac{1}{2m}\nabla_2^2 + \frac{1}{i}\frac{\partial}{\partial t_2} \right) \times \chi^0_{\mathbf{k}E\alpha}(x_1, x_2) = 0. \tag{4.11}$$

For $x_1 = x_2$, Eq. (4.10) becomes

$$\chi_{\mathbf{k}E\alpha}(x, x) = \chi^0_{\mathbf{k}E\alpha}(x, x) - \sum_{\mathbf{Y}} \int dt_y Q(x - y) \times \chi_{\mathbf{k}E\alpha}(y, y), \tag{4.12}$$

where

$$Q(x) \equiv i2J(0)S_F(x)S_F(x). \tag{4.13}$$

It is useful to write the propagator function in the integral representation:

$$S_F(x) = \frac{i}{N}\sum_{\mathbf{k}} \int \frac{dE}{2\pi} \frac{1}{E - E_k + i\epsilon} e^{i(\mathbf{k}\cdot\mathbf{x} - Et)}, \tag{4.14}$$

where

$$E_k \equiv H + (1/2m)\mathbf{k}^2. \tag{4.15}$$

By use of Eq. (4.14), the kernel $Q(x)$ takes the form

$$Q(x) = \frac{1}{N}\sum_{\mathbf{k}} \int \frac{dE}{2\pi} Q(\mathbf{k}, E) e^{i(\mathbf{k}\cdot\mathbf{x} - Et)}, \tag{4.16}$$

where

$$Q(\mathbf{k}, E) = \frac{4J(0)}{N}\sum_{\mathbf{q}} \frac{1}{E - (E_q + E_{k-q})}. \tag{4.17}$$

Introduce now the two-particle state

$$\chi_{\mathbf{k}E\alpha}(x, x) \equiv \chi_\alpha(p, q) e^{i[(\mathbf{p}+\mathbf{q})\cdot\mathbf{x} - (E_p + E_p)t]}. \tag{4.18}$$

Substitution of Eq. (4.18) into the Bethe–Salpeter equation yields

$$\chi_\alpha(p, q) = \frac{1}{1 + Q(\mathbf{p} + \mathbf{q}, E_p + E_q)}. \tag{4.19}$$

Using Eq. (4.17) and replacing the summation by integration,

$$\frac{1}{N}\sum_{\mathbf{q}} \to \frac{1}{(2\pi)^3} \int_{-\pi}^{\pi} d^3q, \tag{4.20}$$

we have

$$Q(\mathbf{p}+\mathbf{q}, E_p+E_q) = -\frac{J(0)}{4\pi^2 SJ}\left(2\pi - a|\mathbf{p}-\mathbf{q}| \times \tanh^{-1}\frac{a|\mathbf{p}-\mathbf{q}|}{2\pi}\right). \quad (4.21)$$

Since the left-hand side of Eq. (4.21) depends on $(\mathbf{p}+\mathbf{q})$, while the right-hand side depends on $(\mathbf{p}-\mathbf{q})$, there are only two possibilities:

$$\text{(i)}\quad \mathbf{p}=\mathbf{q}, \qquad \text{(ii)}\quad \mathbf{p}\cdot\mathbf{q}=0. \quad (4.22)$$

Case (i) implies $\chi_\alpha(p,q) = \text{cons}$. Consequently, the particles are *free* in this case. On the other hand, case (ii) leads to an inconsistency. The argument is given in Appendix B. Consequently, two-particle scattering states may be ignored in the long-wavelength region.

The bound states may be written in the form

$$\chi_{B,\mathbf{k}E\alpha}(x,x) = \chi_{B\alpha}(k)e^{i(\mathbf{k}\cdot\mathbf{x}-\omega_k t)}. \quad (4.23)$$

Substitution of Eq. (4.23) into the homogeneous Bethe–Salpeter integral equation leads to

$$Q(\mathbf{k},\omega_k) = -1. \quad (4.24)$$

Hence Eq. (4.17) is reduced to

$$\frac{4J(0)}{N}\sum_{\mathbf{q}}\frac{1}{M+\frac{1}{m}\mathbf{q}^2} = -1, \quad (4.25)$$

with

$$\omega_k = 2H + (1/4m)\mathbf{k}^2 - M. \quad (4.26)$$

By changing the summation of Eq. (4.25) into an integration, according to Eq. (4.20), it yields a relation between the cutoff λ of the integral in terms of M:

$$\lambda - \sqrt{M}\tan^{-1}\frac{\lambda}{\sqrt{M}} = \frac{4\pi^2}{J(0)}(SJ)^{3/2}. \quad (4.27)$$

In the case of $S=\frac{1}{2}$, Eq. (4.26) leads to

$$\omega_k = 2H + \tfrac{1}{2}a^2 J\mathbf{k}^2 - M, \quad (4.28)$$

which is exactly the leading term found by Bethe.[4] The M defined by Eq. (4.26) is just the binding energy of the bound state.

By restricting ourselves only to two-particle states the Hamiltonian now takes the simple form

$$\mathfrak{H}C = E_0 + \sum_{\mathbf{k}} E_k b_k^+ b_k + \sum_{\mathbf{k}} \omega_k B_k^+ B_k, \quad (4.29)$$

where B_k and B_k^+ are the annihilation and creation operators of the bound states.

From the Hamiltonian of Eq. (4.29) we can write down the thermodynamic free energy of the system:

$$F = E_0 - N\beta^{-1}[Z_{5/2}(H\beta) + 2^{3/2}Z_{5/2}(2H\beta)]\theta^{3/2}, \quad (4.30)$$

with

$$Z_n(x) \equiv \sum_{r=1}^{\infty}\frac{1}{r^n}e^{-rx}, \quad (4.31)$$

$$\theta \equiv [\tfrac{4}{3}\pi\beta SJ(0)]^{-1}. \quad (4.32)$$

The spontaneous magnetization is therefore given by

$$M = N[S - (1+2^{5/2})\zeta(\tfrac{3}{2})\theta^{3/2}], \quad (4.33)$$

where $\zeta(\frac{3}{2}) = Z_{3/2}(0)$.

5. DISCUSSION

In the long-wavelength region, the bound states should appear from the preceding analysis. This conclusion is different from those of Dyson and Wortis. Dyson ruled out the existence of the bound states, while Wortis derived a bound-state condition (in three dimensions) which requires that the total momentum of the pair of particles be larger than a certain threshold value.

The existence of the bound state makes the Born approximation ineffective. Hence, we cannot simply take the interaction part of the Hamiltonian $\mathfrak{H}C$ of Eq. (3.21) as the perturbed part. However, this difficulty can be easily resolved by arranging the Hamiltonian of Eq. (3.21) in the following way:

$$\mathfrak{H}C = (E_0 + \mathfrak{H}C_0 + \mathfrak{H}C_I) + \mathfrak{H}C_{II}, \quad (5.1)$$

where

$$\mathfrak{H}C_0 = \sum_{\mathbf{k}}\{H + 2s[J(0) - J(\mathbf{k})]\}b_k^+ b_k, \quad (5.2)$$

$$\mathfrak{H}C_I = -\frac{J(0)}{N}\sum_{\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3} b_{k_2+k_1}^+ b_{k_3-k_1}^+ b_{k_2} b_{k_3}, \quad (5.3)$$

$$\mathfrak{H}C_{II} = \frac{1}{N}\sum_{\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3}[J(0) - J(\mathbf{k}_1)]b_{k_2+k_1}^+ b_{k_3-k_1}^+ b_{k_2} b_{k_3}. \quad (5.4)$$

By so doing we can take $(\mathfrak{H}C_0 + \mathfrak{H}C_I)$ as the unperturbed part and $\mathfrak{H}C_{II}$ as the perturbation. A perturbation calculation in this way gives the spontaneous magnetization (details of the perturbation calculation will be published in another paper by C. J. L.):

$$M = N(S - a_0\theta^{3/2} - a_1\theta^{5/2} - a_2\theta^{7/2} - a_3\theta^4 + \cdots), \quad (5.5)$$

where θ is defined by Eq. (4.32).

APPENDIX A: ON THE VALUE OF R

The R defined by Eq. (3.24) is given in terms of $F(n)$ which satisfies the equation

$$\frac{F(n+1)}{F(n)} = 2SN + \frac{1}{n+1}\frac{F(n+2)}{F(n)}. \tag{A1}$$

Defining

$$D \equiv \sum_{n=0}^{8SN}\left(\frac{1}{2SN}\right)^n \frac{1}{n!}F^2(n), \tag{A2}$$

we have

$$R^{-1} < D. \tag{A3}$$

To find whether D is convergent, it suffices to find the asymptotic solution of Eq. (A1) for $F(n)$ at $n = 8SN$. Using the ansatz

$$F(n) \sim (2SN)^{n/2-1}\sqrt{n!} \quad \text{for } n \gg 1 \tag{A4}$$

we have, therefore,

$$F(n+1)/F(n) = 4SN, \tag{A5}$$

$$2SN + \frac{1}{n+1}\frac{F(n+2)}{F(n)} = 4SN, \tag{A6}$$

when $n = 8SN$. Thus it verifies that the asymptotic solution of Eq. (A4) does satisfy Eq. (A1). By Eq. (A4) we obtain

$$\left(\frac{1}{2SN}\right)^n \frac{1}{n!}F^2(n) \sim \frac{1}{64S^2N^2}, \tag{A7}$$

which implies the convergence of D, i.e., $R \neq 0$.

APPENDIX B: THE INCONSISTENCY OF THE CASE $\mathbf{p}\cdot\mathbf{q} = 0$

Equations (4.19) and (4.21), under the condition $\mathbf{p}\cdot\mathbf{q} = 0$, lead to

$$\chi_\alpha(\mathbf{p}+\mathbf{q}) = \chi_\alpha(\mathbf{p}-\mathbf{q}) = \left[1 - \frac{J(0)}{4\pi^2 SJ}\left(2\pi - a|\mathbf{p}-\mathbf{q}|\tanh^{-1}\frac{a|\mathbf{p}-\mathbf{q}|}{2\pi}\right)\right]^{-1}. \tag{B1}$$

On the other hand, a two-particle state can be written as

$$\chi_\alpha(\mathbf{x}, t) = \sum_{\mathbf{p},\mathbf{q}} \chi_\alpha(\mathbf{p}+\mathbf{q}) e^{i[(\mathbf{p}+\mathbf{q})\cdot\mathbf{x} - E_{p+q}t]}, \tag{B2}$$

or

$$\chi_\alpha(\mathbf{x}, t) = \sum_{\mathbf{p},\mathbf{q}} \chi_\alpha(\mathbf{p}-\mathbf{q}) e^{i[(\mathbf{p}+\mathbf{q})\cdot\mathbf{x} - E_{p+q}t]}. \tag{B3}$$

Write $\mathbf{l} = \mathbf{p} + \mathbf{q}$; then we have

$$\chi_\alpha(\mathbf{x}, t) = \sum_{\mathbf{l},\mathbf{q}} \chi_\alpha(\mathbf{l}) e^{i[\mathbf{l}\cdot\mathbf{x} - E_l t]}, \tag{B4}$$

or

$$\chi_\alpha(\mathbf{x}, t) = \sum_{\mathbf{l},\mathbf{q}} \chi_\alpha(\mathbf{l}-2\mathbf{q}) e^{i[\mathbf{l}\cdot\mathbf{x} - E_l t]}, \tag{B5}$$

i.e.,

$$\chi_\alpha(\mathbf{l}) = \chi_\alpha(\mathbf{l}-2\mathbf{q}); \tag{B6}$$

hence $\chi_\alpha(\mathbf{l})$ is a periodic function. Comparison of Eq. (B6) with Eq. (B1) leads to the inconsistency.

1 T. Holstein and H. Primakoff, Phys. Rev. 58, 1098 (1940).
2 F. J. Dyson, Phys. Rev. 102, 1217, 1230 (1956).
3 F. Bloch, Z. Physik 61, 206 (1930); 74, 295 (1932).
4 H. Bethe, Z. Physik 71, 205 (1931).
5 M. Wortis, Phys. Rev. 132, 85 (1963).
6 S. Maleev, Zh, Eksp. Teor. Fiz. 33, 1010, (1957) [Sov. Phys. JETP 6, 776 (1957)].
7 See e.g., S. Gasiorowicz, *Elementary Particle Physics* (Wiley, New York, 1967).
8 The doublet feature of a Heisenberg ferromagnet Hamiltonian was discussed, from a different approach by L. Leplae, R. N. Sen, and H. Umezawa, Nuovo Cimento 49, 1 (1967).
9 See, e.g., D. Lurié, *Particle and Field* (Interscience, New York, 1968).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 10 OCTOBER 1971

Covariant Derivatives of Infinite-Dimensional Fields

Carl M. Bender*
Department of Mathematics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139
and
David J. Griffiths
Department of Physics, University of Utah, Salt Lake City, Utah 84112
(Received 13 May 1971)

We show that a field which transforms irreducibly under the homogeneous Lorentz group admits four kinds of covariant derivatives—gradient, divergence, and two species of curl. We construct these derivatives explicitly for both finite- and infinite-dimensional representations. Physical applications are discussed.

I. INTRODUCTION

Although the linear representations of the homogeneous Lorentz group have been exhaustively catalogued by mathematicians,[1] most physicists are reluctant to stray beyond the *finite*-dimensional (tensor and spinor) representations—except, perhaps, for an occasional look at an infinite-dimensional unitary representation. There is good reason for this wariness. A glance at the standard references[2] reveals that, while the theory is elegant, the notational complexity can be formidable. For *finite*-dimensional representations, Lorentz transformation behavior is absorbed into the tensor (or spinor) notation itself; and well-known tools and tricks (involving the use of $g_{\mu\nu}$, $\epsilon_{\mu\nu\lambda\sigma}$, γ_μ, ∂_μ, contraction of indices, symmetrizing, removing the trace, etc.) are available, which reduce most group theoretical problems (Clebsch–Gordan decompositions, formulation of covariant field equations, construction of scalar Lagrangians, etc.) to triviality. Unfortunately, there is no easy way to extend or generalize tensor notation to describe *infinite*-dimensional representations. Hence, these same group theoretical problems become extremely difficult calculations.

Our purpose in this article is to present a simple and general treatment of one such calculation one faces when attempting to formulate covariant field equations[3]: How does one covariantly differentiate an infinite-dimensional field?

II. COVARIANT DERIVATIVES OF FINITE-DIMENSIONAL FIELDS

We tentatively define a *covariant* derivative as follows: Given a field which transforms irreducibly under a group, a *covariant* derivative of this field is a new field which also transforms irreducibly. This new field is obtained by differentiating the original field in a way to be described later. We shall generalize this definition somewhat in Sec. IV.

A. The Rotation Group SO(3)

All irreducible representations of SO(3) are finite-dimensional, and can be labeled (l) where $l = 0, 1, 2, \ldots$. A field which transforms like (l) may be represented in tensor form as $Q^l_{a_1\ldots a_l}(x)$. (From here on, we suppress the space coordinate x.) $Q^l_{a_1\ldots a_l}$ is traceless and symmetric in its 3-space indices a_i, has $(2l+1)$ independent components, and is said to carry spin l.

There are precisely three covariant derivatives of $Q^l_{a_1\ldots a_l}$. These are the divergence

$$\mathrm{div}(Q^l) \equiv \nabla_k Q^l_{ka_1\ldots a_{l-1}}, \tag{1a}$$

which carries spin $l-1$, the curl

$$\mathrm{curl}(Q^l) \equiv \sum_{i=1}^{l} \epsilon_{a_i jk} \nabla_j Q^l_{a_1\ldots\bar{a}_i\ldots a_l k} \tag{1b}$$

which carries spin l, and the gradient

$$\mathrm{grad}(Q^l) \equiv \sum_{i=1}^{l+1} \nabla_{a_i} Q^l_{a_1\ldots\bar{a}_i\ldots a_{l+1}} - \frac{1}{l+2} \nabla_k \times \sum_{i,j=1, i\neq j}^{l+1} \delta_{a_i a_j} Q^l_{a_1\ldots\bar{a}_i\ldots\bar{a}_j\ldots a_{l+1}k}, \tag{1c}$$

which carries spin $l+1$. In Eq. (1) the overbar $\bar{a}_i$ indicates that this index is absent from the sequence. The derivatives in Eq. (1) are covariant (transform irreducibly) because they were constructed to be explicitly totally symmetric and traceless.

The derivative operator ∇_i is a vector (carries spin 1). We can interpret Eq. (1) group theoretically as the Clebsch–Gordan decomposition

$$(1) \otimes (l) = (l-1) \oplus (l) \oplus (l+1). \tag{2}$$

There is one exceptional case. A scalar, $l = 0$, allows only *one* kind of derivative, the gradient, because $(1) \otimes (0) = (1)$.

B. THE HOMOGENEOUS LORENTZ GROUP $O(3,1)$

Irreducible representations of the Lorentz group are characterized by a pair of numbers, (l_0, l_1).[2] l_0 is the lowest spin (that is, the smallest representation of the rotation subgroup) contained in (l_0, l_1). Thus, l_0 must be an integer or half-integer. Each higher spin, $(l_0+1), (l_0+2), (l_0+3), \ldots$, is contained exactly once. If either l_1 or $-l_1$ occurs in this sequence, then it and all higher spins are absent from the representation. The representation is then finite-dimensional, with dimension ${l_1}^2 - {l_0}^2$. For all other values of l_1, real or complex, the representation is infinite-dimensional.

Following subsection II A, we observe that the derivative operator ∂_μ is a vector [transforms

TABLE I. Examples of covariant derivatives of tensors showing the three exceptional cases and the simplest nonexceptional case. Note that (a) a tensor $A_{\kappa_1\kappa_2\ldots\kappa_{l-1}}$ transforming as $(0, l)$ is traceless and symmetric over the indices κ_i; (b) a tensor $B_{\mu\nu\kappa_1\kappa_2\ldots\kappa_{l-2}}$ transforming as $(1, l)$ is antisymmetric and self-dual over $\mu\nu$, traceless over μ and κ_i; (c) a tensor $C_{\mu\nu\lambda\sigma}$ transforming as $(2, 3)$ is antisymmetric and self-dual over $\mu\nu$, symmetric under $(\mu\nu) \longleftrightarrow (\lambda\sigma)$, and traceless over $\mu\lambda$. The dual of an antisymmetric tensor $D_{\mu\nu}$ is $\tilde{D}_{\mu\nu} \equiv \epsilon_{\mu\nu\lambda\sigma} D^{\lambda\sigma}$; $D_{\mu\nu}$ is self-dual if $\tilde{D}_{\mu\nu}$ is proportional to $D_{\mu\nu}$.

Tensor	Representation	Type	Divergence	Representation	grad	Representation	curl$^+$	Representation	curl$^-$	Representation
A	$(0,1)$	Case (i)	$\ldots$	$\ldots$	$\partial_\mu A$	$(0,2)$	$\ldots$	$\ldots$	$\ldots$	$\ldots$
$B_{\mu\nu}$	$(1,2)$	Case (ii)	$\ldots$	$\ldots$	$\partial_\lambda B_{\mu\nu} + (\epsilon_{\mu\nu\phi\phi}\epsilon_{\kappa\lambda\tau\varphi}) \times \frac{1}{2}\partial^\kappa B^{\phi\tau}$	$(1,3)$	$\ldots$	$\ldots$	$\partial_\mu B^{\mu\nu}$	$(0,2)$
A_μ	$(0,2)$	Case (iii)	$\partial_\mu A^\mu$	$(0,1)$	$\partial_\mu A_\nu + \partial_\nu A_\mu - \frac{1}{2} g_{\mu\nu}\partial_\lambda A^\lambda$	$(0,3)$	$\partial_\mu A_\nu - \partial_\nu A_\mu + i\epsilon_{\mu\nu\lambda\sigma}\partial^\lambda A^\sigma$	$(1,2)$	$\partial_\mu A_\nu - \partial_\nu A_\mu - i\epsilon_{\mu\nu\lambda\sigma}\partial^\lambda A^\sigma$	$(1,-2)$
$B_{\mu\nu\lambda}$	$(1,3)$	Simplest nonexceptional case	$\partial^\lambda B_{\mu\nu\lambda}$	$(1,2)$	$\partial^\sigma\{3(g_{\mu\sigma}B_{\mu\nu\lambda} + g_{\lambda\sigma}B_{\mu\nu\kappa}) - g_{\kappa\sigma}B_{\mu\nu\lambda} + g_{\mu\kappa}B_{\sigma\sigma\lambda} - g_{\nu\kappa}B_{\mu\sigma\lambda} + g_{\mu\lambda}B_{\nu\sigma\kappa} - g_{\nu\lambda}B_{\mu\sigma\kappa} + g_{\mu\sigma}B_{\kappa\nu\lambda} - g_{\nu\sigma}B_{\lambda\mu\kappa}\}$	$(1,4)$	$\partial_\mu B_{\kappa\lambda\nu} - \partial_\nu B_{\kappa\lambda\mu} + \partial_\kappa B_{\mu\nu\lambda} - \partial_\lambda B_{\mu\nu\kappa} + (\epsilon_{\mu\nu\tau\theta}\epsilon_{\kappa\lambda\phi\theta}) \times \frac{1}{2}\partial_\sigma(B^{\tau\sigma\varphi} + B^{\varphi\sigma\tau})$	$(2,3)$	$\partial^\mu(B_{\mu\nu\lambda} + B_{\mu\lambda\nu})$	$(0,3)$

as the representation $(0, 2)$]. Thus, to identify the covariant derivatives of the representation (l_0, l_1), we must examine the Clebsch–Gordan decomposition of $(0,2) \otimes (l_0, l_1)$. For finite-dimensional representations,[4] we obtain four covariant derivatives:

$$
\begin{aligned}
&(0,2)\otimes(l_0,l_1) = (l_0, l_1-1), \text{the "divergence"},\\
&\quad\oplus (l_0, l_1+1), \text{the "gradient"},\\
&\quad\oplus (l_0-1, l_1), \text{the "curl}^-\text{"},\\
&\quad\oplus (l_0+1, l_1), \text{the "curl}^+\text{"}.
\end{aligned}
\tag{3}
$$

There are three exceptional cases[5]: (i) If $l_0 = 0$ and $l_1 = 1$, then only the gradient $(0, 2)$ occurs. (ii) If $l_0 > 0$ and $l_1 = l_0 + 1$, then only the gradient and curl$^-$ occur. (iii) If $l_0 = 0$ and $l_1 \neq 1$, then curl$^-$ is $(1, -l_1)$; if $l_0 = \frac{1}{2}$, then curl$^-$ is $(\frac{1}{2}, -l_1)$.

Examples of covariant derivatives of finite-dimensional representations (including some exceptional cases) are given in Table I.

Observe that exceptional cases (i) and (ii) ensure that differentiating a finite-dimensional representation will never yield an infinite-dimensional one. The converse, as will be seen, is not merely *possible*, but *essential* to many physical applications. Exception (iii) explains how to avoid an (unacceptable) negative value for l_0.

III. COVARIANT DERIVATIVES OF INFINITE-DIMENSIONAL FIELDS

We describe infinite-dimensional representations by the following notation: Let

$$
Q = \{Q^{l_0}_{a_1\ldots a_{l_0}}, Q^{l_0+1}_{a_1\ldots a_{l_0+1}}, Q^{l_0+2}_{a_1\ldots a_{l_0+2}}, \ldots\}
$$

belong to the irreducible representation (l_0, l_1). The spin-N entry, $Q^N_{a_1\ldots a_N}$, is traceless and symmetric in the 3-space indices a_i. An infinitesimal Lorentz boost generated by J^{ok} gives

$$
\begin{aligned}
-i[Q^N_{a_1\ldots a_N}, J^{ok}] &= (x^k\partial^o - x^o\nabla_k)Q^N_{a_1\ldots a_N}\\
&\quad - (N+1-l_1)\frac{(N+1)^2 - l_0^2}{(N+1)^2}\, Q^{N+1}_{a_1\ldots a_N k}\\
&\quad + \frac{l_0 l_1}{iN(N+1)}\sum_{i=1}^{N}\epsilon_{a_i kq}Q^N_{a_1\ldots\bar{a}_i\ldots a_N q}\\
&\quad + \frac{N+l_1}{2N+1}\Bigg\{\sum_{i=1}^{N}\delta_{a_i k}Q^{N-1}_{a_1\ldots\bar{a}_i\ldots a_N} - \frac{1}{2N-1}\\
&\quad\times\sum_{i,j=1, i\neq j}^{N}\delta_{a_i a_j}Q^{N-1}_{a_1\ldots\bar{a}_i\ldots\bar{a}_j\ldots a_N k}\Bigg\}.
\end{aligned}
\tag{4}
$$

The choice of basis is, of course, arbitrary. We prefer a tensorial basis, because it treats all three directions symmetrically, and it permits easy connection with 4-tensors. For example, the normalization of Eq. (4) is such that if $T_{\mu_1\mu_2\ldots\mu_n}$ is a traceless symmetric tensor, then

$$
Q^o = T_{oo\ldots o}, \quad Q^1_i = T_{oo\ldots oi},
$$

$$
Q^2_{ij} = T_{oo\ldots oij} - \tfrac{1}{3}\delta_{ij}T_{oo\ldots okk},
$$

and so on. For another choice of basis, one could use the eigenstates of the z-component of the angular momentum, $\xi_{l,m}$.[6]

TABLE II. Explicit formulas for the curl$^+$, curl$^-$, div, and grad of an irreducible representation. The values of $\alpha, \beta, \gamma, \delta$ are to be inserted into Eq. (5).

Derivative	$(\bar{l}_0, \bar{l}_1)$	$\alpha(N)$	$\beta(N)$	$\gamma(N)$	$\delta(N)$
div	$(l_0, l_1 - 1)$	1	$(N+1)^2 - l_0{}^2$	$-l_0$	1
grad	$(l_0, l_1 + 1)$	$(N - l_1)(N + 1 + l_1)$	$(N - l_1)(N + 1 - l_1)[(N+1)^2 - l_0{}^2]$	$(N - l_1)(N + 1 + l_1) l_0$	$(N + l_1)(N + 1 + l_1)$
curl$^+$	$(l_0 + 1, l_1)$	$N - l_0$	$(N - l_0)(N + 1 - l_0)(N + 1 - l_1)$	$(N - l_0) l_1$	$N + l_1$
curl$^-$	$(l_0 - 1, l_1)$	$N + l_0$	$(N + l_0)(N + 1 + l_0)(N + 1 - l_1)$	$-(N + l_0) l_1$	$N + l_1$

We now proceed to construct the covariant derivatives of Q. Let $R = \{R^N\}$ be a covariant derivative of Q. The most general spin-N 3-tensor that one can construct by combining one derivative with the tensors $\{Q^N\}$ is

$$R^N_{a_1 \cdots a_N} = \alpha(N) \partial_o Q^N_{a_1 \cdots a_N} + \frac{\beta(N)}{(N+1)^2} \nabla_k Q^{N+1}_{a_1 \cdots a_N k} + \frac{\gamma(N)}{N(N+1)} \sum_{i=1}^{N} i \epsilon_{a_i pq} \nabla_p Q^N_{a_1 \cdots \bar{a}_i \cdots a_N q} + \frac{\delta(N)}{2N+1} \left\{ \sum_{i=1}^{N} \nabla_{a_i} Q^{N-1}_{a_1 \cdots \bar{a}_i \cdots a_N} - \frac{1}{2N-1} \times \sum_{i,j=1 \; i \neq j}^{N} \delta_{a_i a_j} \nabla_k Q^{N-1}_{a_1 \cdots \bar{a}_i \cdots \bar{a}_j \cdots a_N k} \right\}. \quad (5)$$

$\alpha(N), \beta(N), \gamma(N)$, and $\delta(N)$ are as yet unspecified coefficients. (In defining the last three, we have extracted various functions of N to simplify later results.)

That $R = \{R^N\}$ is a covariant derivative (transforms irreducibly) means that

$$R = \{R^{\bar{l}_0}_{a_1 \cdots a_{\bar{l}_0}}, R^{\bar{l}_0+1}_{a_1 \cdots a_{\bar{l}_0+1}}, R^{\bar{l}_0+2}_{a_1 \cdots a_{\bar{l}_0+2}}, \cdots\}$$

belongs to the representation $(\bar{l}_0, \bar{l}_1)$ for some new parameters $\bar{l}_0$ and $\bar{l}_1$ or equivalently that $-i[R^N_{a_1 \cdots a_N}, J^{ok}]$ is compatible with Eq. (4). This compatibility condition gives simultaneous difference equations satisfied by α, β, γ, and δ, which we have solved. We find that *for both infinite- and finite-dimensional representations, there are just four allowed values of* $(\bar{l}_0, \bar{l}_1)$ *and hence just four kinds of covariant derivatives.* The explicit derivatives, that is, formulas for $\alpha(N), \beta(N), \gamma(N)$, and $\delta(N)$, are given in Table II.

On Fig. 1, we exhibit these four covariant derivatives by graphically illustrating their effect upon the subclass of irreducible representations having integer l_1. We also illustrate the three exceptional cases for finite-dimensional representations.

On Fig. 1, we expect $l_0 = l_1 = l$ to be a special case because (l, l) stands at the threshold of the finite-dimensional region. At this threshold, we can hope to produce finite-dimensional representations by differentiating infinite-dimensional representations. However, a glance at Table II reveals that the four derivatives of (l, l) are *not linearly independent.* In fact, in a self-explanatory notation,

$$\text{curl}^-(l, l) = (N + l) \text{ div } (l, l),$$
$$\text{grad}(l, l) = (N + 1 + l) \text{curl}^+(l, l). \quad (6)$$

This is a startling result because div and curl$^-$ (and grad and curl$^+$) are supposed to transform as *different* representations of the Lorentz group. How can $(l - 1, l)$ be the same representation as $(l, l - 1)$?

To answer these questions we must introduce the concept of indecomposable representations.

Restated, this question is even more perplexing. We may think of R^N as a four-dimensional "vector," with "components" $\alpha(N), \beta(N), \gamma(N), \delta(N)$. Then Table II gives a convenient choice of basis in the four-dimensional space of derivatives. By resolving an arbitrary vector R along the four "directions," div, grad, curl$^+$, and curl$^-$, we have identified the representations of the Lorentz group it contains. However, when $l_0 = l_1 = l$, Table II no longer provides us with a complete basis—it gives only two linearly independent vectors—leaving two

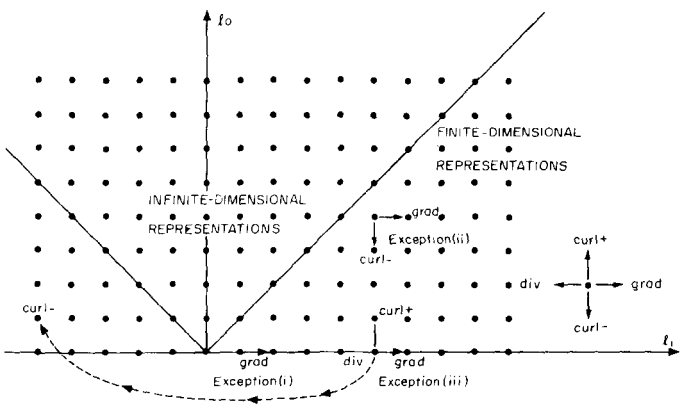


Fig. 1. Schematic diagram of irreducible representations (l_0, l_1) having integers l_0 and l_1. Each dot corresponds to a representation. The arrows show the results of taking a covariant derivative: curl$^+$, curl$^-$, grad, or div. Exceptional cases (i), (ii), and (iii) are shown. The line $|l_0| = |l_1|$ is the threshold between finite- and infinite-dimensional representations.

basis vectors unspecified. There must be two *other* kinds of derivatives belonging to representations of the Lorentz group which are not compatible with Eq. (4), and are thus not irreducible. How, then, do these two extraordinary derivatives of (l, l) transform?

IV. EXTENSION TO INDECOMPOSABLE REPRESENTATIONS

A. Description of Indecomposable Representations

In Sec. III we treated *infinite*-dimensional *irreducible* representations of the Lorentz group. However, the fundamental transformation law in Eq. (4), and hence also the derivatives in Table II, actually apply to a somewhat broader class of representations. Since the Lorentz group is not compact, it possesses representations which are indecomposable (cannot be decomposed into direct sums of irreducible representations). These indecomposable representations consist, in the colorful terminology of Gel'fand and Ponomarev,[1] of various compatible irreducible representations "glued" together. For example, when l_1 is an integer, one can "glue" a finite-dimensional representation (l_0, l_1) onto its infinite-dimensional "tail" (l_1, l_0).[7] Moreover, this gluing can be done in two ways: One way contains (l_0, l_1) as an invariant subrepresentation; the other way leaves the *tail* invariant.[8] Both species are described by Eq. (4).

If l_1 is a positive integer greater than l_0, then the first term on the right side of Eq. (4) exhibits a cutoff at $N = (l_1 - 1)$. This gives, as expected, a *finite*-dimensional representation, and no components with spin l_1 or higher can be generated from the lower-spin components. However, if we start with a component carrying spin $\geq l_1$, then we generate all spins including those below l_1. Thus Eq. (4) describes not only the finite-dimensional irreducible representation (l_0, l_1), but also the indecomposable representation consisting of (l_0, l_1) glued onto its tail and containing (l_0, l_1) as an invariant subrepresentation.

If l_1 is a negative integer and $-l_1 > l_0$, then the *last* term on the right of Eq. (4) has a cutoff at $N = -l_1$. If we start with some component carrying spin $> -l_1$ and work our way down, we never generate spins smaller than $-l_1$: This is the infinite-dimensional representation $(-l_1, -l_0)$. On the other hand, if we begin with a lower-spin component and work up, we recover all spins. Evidently, in this case, Eq. (4) describes the indecomposable representation consisting of (l_0, l_1) glued onto its tail $(-l_1, -l_0)$ and containing the *tail* as an invariant subrepresentation.

Observe that, as it stands, Eq. (4) cannot handle the *finite*-dimensional irreducible representation (l_0, l_1), when l_1 is a negative integer and $-l_1 > l_0$ (nor the indecomposable representation which invariantly contains it). But this defect is easily remedied by noting that $(l_0, -l_1)$ is the representation conjugate to (l_0, l_1). Thus, to get the transformation rule for $(1, -3)$, say, we merely write Eq. (4) for $(1, 3)$, and then change the sign of i in the middle term (formally, we switch the sign of l_0).

B. The Special Case $l_0 = l_1 =$ Integer

We can now answer the first question that we posed in Sec. III: How can $(l, l-1)$ be the same representation as $(l-1, l)$? In this one peculiar instance, the difference between these two is illusory. $\mathrm{div}(l, l) = (l, l-1)$ is the tail part of $\mathrm{curl}^-(l, l)$. The latter, as we have just found, normally consists of $(l-1, l)$ bound indecomposably to its tail $(l, l-1)$. Since, however, $\beta(l-1)$ happens to vanish in this case, only the tail remains. Apart from normalization, then, $\mathrm{div}(l, l)$ and $\mathrm{curl}^-(l, l)$ are *identical*, and both belong to the irreducible representation $(l, l-1)$. Similarly, $\mathrm{curl}^+(l, l)$ and $\mathrm{grad}(l, l)$ differ only in normalization, and both transform as $(l+1, l)$. Group theoretically, therefore,

$$\begin{aligned} &\mathrm{div}(l, l) \sim \mathrm{curl}^-(l, l) \sim (l, l-1), \\ &\mathrm{grad}(l, l) \sim \mathrm{curl}^+(l, l) \sim (l+1, l). \end{aligned} \tag{7}$$

C. Extraordinary Derivatives of (l, l)

The examples of indecomposable representations we have described thus far are rather simple. More complicated indecomposable representations are constructed by "gluing" *any number* of replicas of a representation to *any number* of replicas of its tail.[1] The particular structure of subrepresentations and subsubrepresentations is also somewhat arbitrary.

We can now answer the second question posed in Sec. III: What are the two extraordinary derivatives of (l, l)? Let

$$D = \{D^{l}_{a_1 \cdots a_l}, D^{l+1}_{a_1 \cdots a_{l+1}}, D^{l+2}_{a_1 \cdots a_{l+2}}, \cdots\}$$

and

$$G = \{G^{l+1}_{a_1 \cdots a_{l+1}}, G^{l+2}_{a_1 \cdots a_{l+2}}, G^{l+3}_{a_1 \cdots a_{l+3}}, \cdots\}$$

represent the divergence and curl^+ of (l, l). With each of these "ordinary" derivatives, we associate an "extraordinary" derivative

$$V = \{V^{l-1}_{a_1 \cdots a_{l-1}}, V^{l}_{a_1 \cdots a_l}, V^{l+1}_{a_1 \cdots a_{l+1}}, \cdots\}$$

and

$$W = \{W^{l}_{a_1 \cdots a_l}, W^{l+1}_{a_1 \cdots a_{l+1}}, W^{l+2}_{a_1 \cdots a_{l+2}}, \cdots\},$$

respectively. The coefficients which specify V are $\alpha(N) = 1, \beta(N) = (N+1)(N+1+l), \gamma(N) = 0$, and $\delta(N) = N(N+l)^{-1}$, and the coefficients which specify W are $\alpha(N) = 1, \beta(N) = (2N+1)(N+1-l)^2 (N+1+l)^{-1}, \gamma(N) = N$, and $\delta(N) = 0$.

We have found that D and V and G and W each combine to form indecomposable representations. The infinitesimal Lorentz transformation laws for V and W are

$$-i[V^N_{a_1\cdots a_N}, J^{ok}] = (x^k\partial^o - x^o\nabla_k)V^N_{a_1\cdots a_N} - \frac{N+1+l}{(N+1)^2}[(N+2-l)(N+1-l)V^{N+1}_{a_1\cdots a_N k} + lD^{N+1}_{a_1\cdots a_N k}]$$
$$- \frac{il}{N(N+1)}\sum_{i=1}^{N}\epsilon_{a_i kq}[(l-1)V^N_{a_1\cdots\bar{a}_i\cdots a_N q} + D^N_{a_1\cdots\bar{a}_i\cdots a_N q}] + \frac{1}{2N+1}\Big\{\sum_{i=1}^{N}\delta_{a_i k}\Big[(N+l-1)V^{N-1}_{a_1\cdots\bar{a}_i\cdots a_N}$$
$$+ \frac{l}{N+l}D^{N-1}_{a_1\cdots\bar{a}_i\cdots a_N}\Big] - \frac{1}{2N-1}\sum_{i,j=1,i\neq j}^{N}\delta_{a_i a_j}\Big[(N+l-1)V^{N-1}_{a_1\cdots\bar{a}_i\cdots\bar{a}_j\cdots a_N k}$$
$$+ \frac{l}{N+l}D^{N-1}_{a_1\cdots\bar{a}_i\cdots\bar{a}_j\cdots a_N k}\Big]\Big\}, \tag{8a}$$

$$-i[W^N_{a_1\cdots a_N}, J^{ok}] = (x^k\partial^o - x^o\nabla_k)W^N_{a_1\cdots a_N} - \frac{1}{(N+1)^2}\Big[(N+1-l)(N-l)(N+2+l)W^{N+1}_{a_1\cdots a_N k}$$
$$+ \frac{2[(N+1)^2 + l(l+1)]}{N+l+1}G^{N+1}_{a_1\cdots a_N k}\Big] - \frac{i}{N(N+1)}\sum_{i=1}^{N}\epsilon_{a_i kq}[l(l+1)W^N_{a_1\cdots\bar{a}_i\cdots a_N q} + G^N_{a_1\cdots\bar{a}_i\cdots a_N q}]$$
$$+ \frac{N+l}{2N+1}\Big\{\sum_{i=1}^{N}\delta_{a_i k}W^{N-1}_{a_1\cdots\bar{a}_i\cdots a_N} - \frac{1}{2N-1}\sum_{i,j=1,i\neq j}^{N}\delta_{a_i a_j}W^{N-1}_{a_1\cdots\bar{a}_i\cdots\bar{a}_j\cdots a_N k}\Big\}. \tag{8b}$$

Equation (8) is somewhat complicated, but it is easy to summarize schematically (see Fig. 2).[9] The arrows indicate a commutation with J^{ok}

From Eq. (8) (or Fig. 2) we observe the following: The transformation law for V^{l-1} gives D^l *only*, and no V^l. Hence, V^{l-1} is indecomposably "glued" to D. The other components of V are indecomposably joined to this combination. In other words, the divergence [the *irreducible* representation $(l, l-1)$] is imbedded in an *indecomposable* representation consisting of D joined to V^{l-1}; and this combination is, in turn, imbedded in a still *larger* indecomposable representation which includes $\{V^l, V^{l+1}, V^{l+2}, \ldots\}$. G and W are indecomposably united in an analogous way.

There are important instances in field theory (see Sec. V) when the divergence D vanishes. In such cases (only), V^{l-1} constitutes the *irreducible finite-dimensional* representation $(l-1, l)$. Similarly, if curl$^+$ vanishes, then W^l becomes the finite-dimensional representation $(l, l+1)$. It is precisely in these two ways that we can obtain finite-dimensional representations by differentiating infinite-dimensional representations. This mathematical curiosity is intimately connected with the construction of field-strength tensors in massless field theory.

D. Analogy to a Multisheeted Riemann Surface

There is a remarkable analogy between Fig. 1 and a Riemann surface. Let us think of the line $|l_0| = |l_1|$ (the threshold separating the finite- and infinite-dimensional representations) as a "branch cut" and the operations of curl$^+$, curl$^-$, div, and grad as "analytic continuation" from one point to another in the "complex plane." The "points" on Fig. 1 are all *irreducible* representations. We imagine an infinite stack of planes similar to Fig. 1, whose points are *indecomposable* representations of increasing complexity. When we analytically continue through the threshold, we jump to the next higher Riemann sheet. This jump always occurs except in the special instance described at the end of subsection IV C.

V. APPLICATIONS

As Bender first demonstrated,[10] free massless quantum fields in the radiation gauge belong to infinite-dimensional representations of the Lorentz group. In fact, the field $C(L)$ associated with a free massless particle of spin L transforms according to the representation $(L, 1)$. These fields obey complicated field equations.[11] In our present notation, these equations assume the delightfully simple form:

$$\begin{aligned} &\text{div } C = 0, \\ &\text{curl}^+ C = 0, \\ &\text{curl}^- C = 0, \\ &\text{grad } C = 2(L^2 - 1)(\partial_0 C). \end{aligned} \tag{9}$$

Eq. (9) is the infinite-dimensional analog of Maxwell's equations.

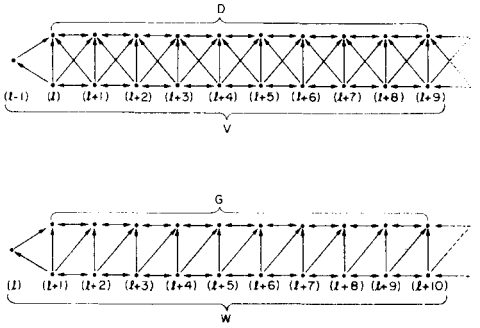


Fig. 2. Schematic view of the transformation law in Eq. (8). The invariant subspaces of the indecomposable representations, V and D, and W and G, are shown. Each arrow represents a commutation with J^{ok}, the generator of pure Lorentz transformations.

From the last equation, it is clear that $(\partial_0 C)$ belongs to the representation $(L, 2)$. Indeed, Bender found that the Pth time derivative of C [which we shall call (P) for short] belongs to the representation $(L, 1 + P)$, and obeys field equations which read, in our notation:

$$\begin{aligned} &\mathrm{div}(P) = 0, \\ &\mathrm{curl}^+(P) = 0, \\ &\mathrm{curl}^-(P) = 0, \\ &\mathrm{grad}(P) = 2[L^2 - (P+1)^2](P+1). \end{aligned} \tag{10}$$

When $P = (L - 1)$, the gradient vanishes. This is expected because we are now computing grad(L, L) which, as shown in Sec. IV, is proportional to curl$^+(L, L)$. As curl$^+(P)$ is *always* zero, it is not surprising that the gradient *also* vanishes at this stage. However, what about the two anomalous derivatives of (L, L)? Explicit calculation gives $V = 0$ and $W^N = 2(N + 1 - L)(N + 1 + L)^{-1}(L)^N$. In particular, since curl$^+$ (L, L) vanishes, $W^L = 2/(2L + 1)(L)^L$ belongs to the *finite*-dimensional irreducible representation $(L, L + 1)$. Thus, the Lth time derivative of $C(L)$ belongs to a finite-dimensional representation (this, also, was noted by Bender) called the "field-strength tensor." For spin 1 (electrodynamics) this is $F^{\mu\nu}$; for spin 2 (linearized gravity), it is the free Riemann tensor.

In a recent paper,[12] Bender and Griffiths demonstrated that the energy density (and, in fact, all the local densities of the Poincaré generators) for a massless field carrying spin >1 belongs to an infinite-dimensional representation of the Lorentz group. This result contrasts with the massive case (or massless spin 0, $\frac{1}{2}$, or 1) where, of course, there is a second-rank stress tensor $T^{\mu\nu}$. The ordinary finite-dimensional stress tensor transforms as

$$(0, 3) \oplus (0, 1),$$

whereas the massless infinite-rank stress "tensor" transforms as[12]

$$(0.3)' \oplus (0, 1)'.$$

The prime in $(0, l)'$ indicates that this is the infinite-dimensional representation consisting of $(0, l)$ glued indecomposably onto its tail $(l, 0)$ in such a way that the *tail* is the invariant subrepresentation.

$T^{\mu\nu}$ satisfies a local conservation law

$$\partial_\mu T^{\mu\nu} = 0,$$

or, in the terminology of this paper,

$$\mathrm{div}\ T_{(0,3)} \propto \mathrm{grad}\ T_{(0,1)}.$$

The divergence of the (0, 3) part and the gradient of the (0, 1) part both transform as (0, 2); the local conservation law says that they are in fact proportional.[13] The infinite-dimensional stress tensor was also found to obey a pair of local conservation laws which now read:

$$\begin{aligned} &\mathrm{div}\ T_{(0,3)'} \propto \mathrm{grad}\ T_{(0,1)'}, \\ &\mathrm{div}\ T_{(0,1)'} = 0. \end{aligned} \tag{11}$$

These equations as stated here are much simpler than their original form,[12] and reveal in an elegant way the striking parallel between finite- and infinite-dimensional stress tensors.[14]

---

* Work supported in part by the National Science Foundation.

1 I. M. Gel'fand and V. A. Ponomarev, Usp. Mat. Nauk. **23**, 3 (1968) [Russian Math. Surveys **23**, 1 (1968)].

2 I. M. Gel'fand, R. A. Minlos, and Z. Ya. Shapiro, *Representations of the Rotation and Lorentz Groups and Their Applications* (MacMillan, New York, 1963), hereafter, GMS; M.A. Naimark, *Linear Representations of the Lorentz Group* (MacMillan, New York, 1963).

3 Some of the results of this paper are to be found, in rather different form, in the last chapter of either work in Ref. 2.

4 See, for instance, GMS, Ref. 2, p. 264.

5 Equation (3) does not hold for these exceptional cases.

6 The explicit connection between our $Q^N_{a_1 \cdots a_N}$ and $\xi_{l,m}$ of GMS is

$$\xi_{l,m} = \beta(l, m) \sum_{n=0}^{m} (i)^n \binom{m}{n} Q^l_{\underbrace{11\cdots1}_{n}\ \underbrace{22\cdots2}_{m-n}\ \underbrace{33\cdots3}_{l-m}},$$

where β is given by

$$\frac{\beta(l, m)}{\beta(l-1, m)} = \frac{i}{l}\left[\frac{(l_1 - l)(l^2 - l_0^2)(4l^2 - 1)}{(l_1 + l)(l^2 - m^2)}\right]^{1/2}$$

With this translation, Eq. (3'), p. 189, of GMS reduces to our Eq. (2). The same applies to Eq. (54), p. 117, of Naimark (Ref. 2).

7 If l_1 is negative, then the tail is $(-l_1, -l_0)$.

8 These representations are sometimes called "integer point" representations. See I. M. Gel'fand, M. I. Graev, and N. Ya. Vilenkin, *Generalized Functions* (Academic, New York, 1966), Vol. V, Chap. III.

9 An arbitrary admixture of D in V (or of G in W) will neither affect the fundamental structure of the transformation law nor its group theoretical interpretation. In defining V and W we have chosen the admixtures which give the simplest coefficients and transformation laws.

10 C. M. Bender, Phys. Rev. **168**, 1809 (1968); see also Y. Frishman and C. Itzykson, *ibid.* **180**, 1556 (1969); **183**, 1520E (1969).

11 Bender (Ref. 10), Eqs. (11)–(14). The normalization convention in Bender's paper is not the same as our Eq. (2). The translation is

$$Q^N = \sigma(N) C(L)^N, \text{with } \sigma(N+1) = -[(N+1)/(2N+1)]\sigma(N).$$

12 C. M. Bender and D. J. Griffiths, Phys. Rev. D **1**, 2335 (1970).

13 We do not specify the proportionality constant because it depends on the normalization conventions for div and grad. We are concerned here with group theory only.

14 The absence of a second local conservation law for the "ordinary" stress tensor admits a simple interpretation: div(0, 1) does not exist [exceptional case (i)]!

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 10 OCTOBER 1971

Second Sound in a Low-Temperature Weakly Interacting Bose Gas*

Shang-keng Ma
Department of Physics and Institute for Pure and Applied Physical Sciences, University of California, San Diego, La Jolla, California 92037
(Received 6 May 1971)

Second sound in a weakly interacting Bose gas at low temperatures is described from a purely microscopic viewpoint. The simplicity of the model, which is specified by a single scattering length, allows a precise field theoretic analysis within the framework of equilibrium statistical mechanics in the weak coupling limit. The dispersion formula $\nu = c_2 q - i\tau q^2 + O(q^3)$ for the frequency of the second sound ν in terms of its wavenumber q is established. The leading terms in the low-temperature limit of the constants c_2 and τ are evaluated exactly.

1. INTRODUCTION

A weakly interacting Bose system at low temperatures may be viewed as a dilute gas of phonons. Analogous to the sound waves in ordinary gases, there is *second sound* in this gas of phonons. With this physical picture in mind, we present in this paper a purely microscopic description of second sound based on a simple model.

The purpose of such a microscopic description is to gain qualitative understanding in a clear-cut manner as simply as possible. We do not expect the simple results obtained here to be applicable to observations on He II. Second sound in He II has been very extensively investigated theoretically, especially by Khalantnikov.[1] In the absence of any tractable microscopic theory of He II, these investigations are necessarily phenomenological. They are complicated since they must include many effects to have a realistic description of He II. When the parameters are many, one may easily lose track of the basic physical features. An idealized microscopic model study can thus serve pedagogical purposes better and, being rigorous, it may offer clues for improving existing phenomenological theories, where ambiguities are still present.

A purely microscopic study of second sound is very important from a more theoretical viewpoint. Description of collective phenomena from first principles has always been a challenging problem. The weakly interacting Bose gas model has been one of the few which exhibit interesting physical features and allow some clear-cut calculations. It is responsible for much of our understanding of general characteristics of Bose systems. Previous field theoretic investigations of this model have been mostly concerned with zero-temperature properties.[2,3] Only recently has there been more extensive study on low-temperature features,[4,5] but not on second sound. It is of great interest to accumulate information about this model. A study of second sound is a step toward a more complete understanding of not only the model but also the phenomenon of second sound itself and related phenomena.

Second sound can also appear in a very different system, i.e., a crystal lattice, and has received much investigation. At low temperatures, a crystal lattice also behaves like a dilute gas of phonons. Several authors, notably Krumhansl and Guyer,[6] have studied the collective behavior, including the second sound, of these phonons via a kinetic equation. The first microscopic approach was due to Sham,[7] who derived the kinetic equation assumed by previous authors, and elucidated many basic features. The Krumhansl–Guyer solution to the kinetic equation was a very important step toward a general understanding.[6] It however requires phenomenological relaxation times. This is inevitable owing to the complex nature of any lattice model. The phonons needed to be described phenomenologically to start with, in any case.

A microscopic theory of second sound in a Bose gas must describe the phonons, which appear as a result of the interaction between bosons making up the system, from first principles and then describe the propagation of the second sound. The damping of the second sound must be also calculated in terms of microscopic parameters describing the bosons. In this respect, it is a more complicated problem than the one in solids mentioned above. On the other hand, the weakly interacting Bose gas model is probably the only one which is simple enough to allow a purely microscopic and tractable description of second sound.

We give an outline of the paper and summarize the results. In Sec. 2, we define the model, identify the small parameters, and give a qualitative discussion of the problem, emphasizing the distinctive features of the second sound in contrast to the zero sound. The small parameters in the theory of a weakly interacting Bose gas at low temperatures are

$$g = 4\pi a m s_0, \qquad t = T/ms_0^2, \tag{1.1}$$

where m is the mass of a boson, a is the boson–boson scattering length, s_0 is the speed of a phonon in the zeroth (Bogoliubov) approximation:

$$s_0 = (4\pi a n)^{1/2} m^{-1}, \tag{1.2}$$

and n is the average density.

Following the hints of the qualitative discussion, the small-g limit of the Bethe–Salpeter equation for certain vertex functions is obtained using the usual field theoretic finite-temperature perturbation theory. This small-g limit turns out to resemble a Boltzmann equation when we take into account emission and absorption of phonons. It is an integral equation with free phonon energy and on-shell absorption–emission amplitudes as inputs.

Some general features of the integral equation, as consequences of energy-momentum conservation,

space-time inversion, and rotation invariance, are discussed in Sec. 4.

The eigenvalues and eigenfunctions of the kernel of the integral equation are investigated in Sec. 5. Singular solutions with poles near the real frequency axis are identified as the second sound. We establish the dispersion formula for the second-sound frequency ν as a function of its wavenumber q:

$$\nu = c_2 q - i\tau q^2 + O(q^3) \tag{1.3}$$

for small q, and, in the low-temperature limit, we evaluate the leading terms of c_2 and τ exactly and find, in terms of the small parameters g and t defined in (1.1),

$$c_2/s_0 = 3^{-1/2}(1 + 10.57t^2) + O(t^4) + O(g), \tag{1.4}$$

$$\tau m = 0.003314 g^{-1} t^{-9} + g^{-1} O(t^{-7}) + O(1). \tag{1.5}$$

The quantity c_2 is the speed of second sound and τ/s_0^2 may be regarded as an effective relaxation time. Equation (1.3) is valid at low temperatures provided q is very small, i.e.,

$$q/ms_0 \ll gt^9. \tag{1.6}$$

The temperature dependence of the $O(g)$ term in (1.4) and $O(1)$ in (1.5) is not known.

Physical interpretation in terms of oscillations in phonon distribution is sketched in Sec. 6. Further discussions and remarks are made.

The results of this paper may be viewed as the leading terms in a systematic expansion in powers of g and t. There is neither an assumption nor any input other than g and t. As will be seen in the text, the picture of the second sound in this model is very simple indeed. A detailed analysis of the relationship between our description here and that of the two-fluid hydrodynamics has not been made. Such an analysis will undoubtedly be very fruitful.

We want to emphasize that the purpose of this investigation, like most of model calculations, is not to obtain exact numbers such as (1.4) and (1.5), but to extract qualitative features in a rigorous way and demonstrate new approaches.

The notation in this paper follows closely to that in Ref. 5, and the model is identical. For completeness, we shall redefine the model and notation here. The reader is assumed to be familiar with the elementary features of a Bose gas and the Bogoliubov approximation. Detailed discussions on the formulation of the perturbation theory and characteristics of phonons can be found in Refs. 2–5.

2. BASIC NOTION AND SIMPLE FEATURES

A. The Model

We are interested in a weakly interacting gas of spin-zero, mass-m bosons at temperature T and density n. The two-body force between a pair of bosons is assumed to be short-ranged and summarized by the s-wave scattering length a. These four quantities, m, T, n, and a are the input parameters of the theory. As will be seen shortly, there are effectively only two parameters in the theory after properly choosing the units.

We construct a grand canonical ensemble at temperature T, chemical potential μ, with the Hamiltonian

$$H - \mu N = \sum_{\mathbf{k}} (\epsilon_k - \mu) a_{\mathbf{k}}^\dagger a_{\mathbf{k}} + \sum_{\mathbf{k}} (\rho_{\mathbf{k}} \rho_{\mathbf{k}}^\dagger - N) v \tag{2.1}$$

and the grand potential

$$\Omega = -T \ln Tr \exp[-(H - \mu N)]/T, \tag{2.2}$$

where

$$\epsilon_k = k^2/2m, \tag{2.3}$$

$$\rho_{\mathbf{k}} = \sum_{\mathbf{p}} a_{\mathbf{p}}^\dagger a_{\mathbf{p}+\mathbf{k}}. \tag{2.4}$$

The volume of the system is taken to be unity. We have used a point potential as the model interaction so that v is a constant. This parameter v can be eliminated, to the orders of approximation of interest, by using the perturbation expansion for the scattering length a in powers of v.

To describe the condensate, we let

$$a_0 = a_0^\dagger = \sqrt{n_0}, \tag{2.5}$$

where n_0, a c-number, is the condensate density. Substituting (2.5) in (2.1), we have a model interaction with terms shown in Fig. 1. The condensate acts as a classical field. Bosons having nonzero momenta now have the Hamiltonian $H - \mu N'$ and the grand potential

$$\Omega' = -T \ln Tr \exp[-(H - \mu H')]/T = \Omega + \mu n_0, \tag{2.6}$$

where

$$N' = \sum_{\mathbf{k}=0} a_{\mathbf{k}}^\dagger a_{\mathbf{k}}. \tag{2.7}$$

The new parameters n_0 and μ can be eliminated by requiring $\partial\Omega/\partial n_0 = 0$ at the correct n_0, i.e.,

$$\frac{\partial \Omega'}{\partial n_0} = \mu, \tag{2.8}$$

and the condition

$$n = n_0 + n', \tag{2.9}$$

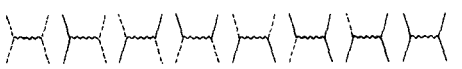


FIG. 1. Interaction terms. Dotted lines represent a factor $\sqrt{n_0}$ for the condensate particles. Solid lines represent particles of nonzero momenta. Wavy lines denote a factor v.

where

$$n = \langle N \rangle, \quad n' = \langle N' \rangle.$$

We have thus defined our model completely. The standard diagrammatic perturbation theory will be our method of analysis.

B. Bogoliubov Approximation and Small Parameters

The well-known Bogoliubov approximation is obtained by ignoring the last three of the interaction terms shown in Fig. 1 and setting $n_0 = n$. The excited states under this approximation are described by noninteracting particles, which we shall refer to as phonons, each having an energy

$$\omega_p = [s_0^2 p^2 + (p^2/2m)^2]^{1/2}, \tag{2.10}$$

where p is the phonon momentum, and

$$s_0^2 \equiv 4\pi a n/m^2. \tag{2.11}$$

For small momenta, we have $\omega_p = s_0 p$. When the last three terms in Fig. 1 are kept, one effectively switches on the interaction between phonons in addition to modifying the characteristics of individual phonons. At low temperatures, our model then describes a weakly interacting dilute gas of phonons.

Precisely, by "low temperature" and "weakly interacting" we mean that the dimensionless parameters t and g, defined by

$$t = \frac{T}{ms_0^2} = \frac{T}{4\pi an/m}, \qquad g = 4\pi a m s_0 = (4\pi a)^{3/2} n^{1/2}, \tag{2.12}$$

are small. Taking the Bogoliubov approximation as the zeroth approximation, we naturally define ms_0 and ms_0^2, as units of momentum and energy, respectively. In these units, we have

$$m = s_0 = 1, \quad 4\pi a = n^{-1} = g, \quad t = T. \tag{2.13}$$

We shall work in these units from now on. The usual perturbation theory can then be written in powers of g by systematically expanding the parameters v, n_0, and μ in powers of g. In the zero g approximation the system is described by the Hamiltonian of free phonons:

$$\sum_{\mathbf{p}} \alpha_{\mathbf{p}}^{\dagger} \alpha_{\mathbf{p}} \omega_p + \text{const}, \tag{2.14}$$

where α_p is related to the boson annihilation operator via the linear combination

$$\alpha_{\mathbf{p}} = A_p a_{\mathbf{p}} + B_p a_{-\mathbf{p}}^{\dagger}, \tag{2.15}$$

and

$$A_p = \tfrac{1}{2}\lambda_p^{-1/2}(1 + \lambda_p), \qquad B_p = \tfrac{1}{2}\lambda_p^{-1/2}(1 - \lambda_p),$$

$$\lambda_p = p^2/2\omega_p, \qquad \omega_p^2 = p^2 + \tfrac{1}{4}p^4. \tag{2.16}$$

C. Zero Sound and Second Sound

Before going into mathematical details, we briefly review some simple qualitative features.

A phonon is a particle, i.e., a discrete excited state of the system. (See Refs. 3–5.) The damping of a phonon at low temperatures is interpreted as due to absorption and scattering by other phonons. The very-low-frequency phonon is usually referred to as the "zero sound." It has little resemblance to the sound wave in ordinary gases. On the other hand, the second sound in a Bose system is analogous to the sound wave in ordinary gases. At low temperatures, we have a dilute gas of phonons. By virtue of the interaction between phonons, it is possible to propagate sound waves in this gas. This sound wave is the second sound, which is, as usually referred to, the "collision-dominated sound wave." The physical picture of a second sound is thus entirely different from that of a zero sound.

From a more mathematical viewpoint, the second sound like the zero sound appears as a nearly discrete singularity of certain response functions, which are functions of the complex frequency variable ν. The response functions are defined by the Fourier transform of retarded commutators:

$$F(X_{\mathbf{q}}, X'_{\mathbf{q}}) = -i\int dt e^{i\nu t}\langle [X_{\mathbf{q}}(t), X'_{-\mathbf{q}}]\rangle \theta(t), \tag{2.17}$$

where $X_{\mathbf{q}}, X'_{-\mathbf{q}}$ are dynamical variables such as density and current, and $\mathbf{q}$, the momentum, specifies the transformation property of the operators under translation. Writing F as a sum over matrix elements of X, X' between energy eigenstates, we have, in an obvious notation,

$$F = Z^{-1}\sum_{m,n} e^{-\beta E_m}\langle m|X_{\mathbf{q}}|n\rangle\langle n|X'_{-\mathbf{q}}|m\rangle(1 - e^{-\beta E_{nm}}) \times \frac{1}{\nu - E_{nm}}, \quad E_{nm} \equiv E_n - E_m, \tag{2.18}$$

and Z is the partition function. As a special case, at zero temperature we have

$$F_{T=0} = \sum_n \langle 0|X_{\mathbf{q}}|n\rangle\langle n|X'_{-\mathbf{q}}|0\rangle \times \left(\frac{1}{\nu - E_{n0}} - \frac{1}{\nu + E_{n0}}\right), \tag{2.19}$$

where $|0\rangle$ is the ground state, with zero energy. We see that $F_{T=0}$ has poles at the energies of excited states. The zero sound will appear as a pole in $F_{T=0}$ but the second sound will not. At a finite temperature F is singular at energy *differences* E_{nm} (including E_{n0}, of course) owing to transitions among states. The zero-sound pole will stay. The second sound now appears as a singularity resulting from specially favored transitions between certain sets of excited states. One must keep in mind these qualitative differences even though both the zero sound and the second sound appear as singularities of the response functions.

In determining the dispersion of zero sound, it is sufficient to use a straightforward perturbation expansion in powers of g.[5] This is because the phonon is a particle and is only slightly disturbed when a small g is switched on. It becomes unstable but the decay rate is small. At low temperatures, the phonon density is small and a phonon will feel only a small disturbance from other phonons.

In determining the dispersion of the second sound, the situation is reversed. As was mentioned above, the second sound owes its existence to the interaction among the phonons. There would be no second sound if $g = 0$. Also, there can be no second sound if $T = 0$ because there would be no phonon gas. In fact the temperature must be high enough to assure a high enough density of phonons for second-sound propagation. We thus expect the damping of the second sound to be a singular function of g and T. The mathematical approach will therefore be very different from a straightforward perturbation expansion in g and T.

Finally, we note that the range of frequency of the zero sound is large, i.e., anywhere from 0 up to 1 [in our units (2.13)]. On the contrary, the frequency range of the second sound is restricted to be very low. It must be much less than the "mean collision frequency" by analogy to the situation of sound waves in ordinary gases. This mean collision frequency may be estimated by the phonon damping rate, which is proportional to $\omega g T^4$, where ω is the phonon energy (see Ref. 5). Since $\omega \sim T$, we should have

$$0 < \nu \ll g T^5. \tag{2.20}$$

This estimate for the upper bound is wrong, however. We shall see later that, owing to the near proportionality of phonon energy and momentum, the condition turns out to be $\nu \ll g T^9$.

Thus, to describe the second sound in terms of interactions between phonons, it is necessary to describe multiple scattering events since the time interval of interest is very long. Mathematically, multiple scattering amplitudes as functions of energies and momenta of the particles involved have many singularities. The worst ones correspond to long-lasting intermediate states, or to the intermediate particles becoming on-energy shell. In the low-density limit, it is possible to sum the contribution of these leading singularities to various physical quantities. In the intermediate stages of calculation, an integral equation, which is essentially a Boltzmann equation, is often obtained and in it the on-shell amplitudes of simple events appear. What we shall accomplish in the next section is just the summation of leading singular terms using the techniques of finite-temperature perturbation theory. We emphasize this fact at this stage because it will not be very conspicuous when we go through the mathematics.

3. THE INTEGRAL EQUATION

Guided by the above qualitative results, we proceed to the mathematical details. The program is to look for singularities of response functions which can be identified as the second sound.

A. Definitions

Instead of using the operators $a_{\mathbf{p}}$ and $a_{\mathbf{p}}^{\dagger}$, we shall use $\alpha_{\mathbf{p}}$ and $\alpha_{\mathbf{p}}^{\dagger}$ [see (2.15)] to define the Green's functions. Let

$$\alpha_{\mathbf{p}}^{\mu} = \begin{cases} \alpha_{\mathbf{p}} & \text{if } \mu = +, \\ \alpha_{-\mathbf{p}}^{\dagger} & \text{if } \mu = -. \end{cases} \tag{3.1}$$

Greek indices will always denote + or −. Repeated indices are summed over + and −. We define the Green's function as the 2×2 matrix

$$\mathcal{G}_{\nu}^{\mu}(p) = -i \int dt e^{i\omega t} \langle [\alpha_{\mathbf{p}}^{\mu}(t), \alpha_{\mathbf{p}}^{\nu\dagger}] \rangle \theta(t), \tag{3.2}$$

where p stands for $(\omega, \mathbf{p})$ in the argument of $\mathcal{G}$. Let us define

$$\mathcal{G}_{\nu}^{\mu} = \mathfrak{N}_{\nu}^{\mu}/\mathfrak{D}, \quad \mathfrak{D} = -\det \mathcal{G}. \tag{3.3}$$

As expansions in powers of g, we have

$$\mathfrak{N} = \mathfrak{N}_0 + g\mathfrak{N}_1 + \cdots, \quad \mathfrak{D} = \mathfrak{D}_0 + g\mathfrak{D}_1 + \cdots. \tag{3.4}$$

The zeroth-order Green's function is diagonal:

$$\mathfrak{N}_0(p) = \begin{pmatrix} \omega + \omega_p & 0 \\ 0 & -\omega + \omega_p \end{pmatrix}, \tag{3.5}$$

$$\mathfrak{D}_0(p) = \omega^2 - \omega_p^2. \tag{3.6}$$

The Green's function fully describes the propagation of a phonon. Next, we define the vertex function $\Lambda_{\sigma}^{\mu}(p, q)$ by

$$\Lambda_{\sigma}^{\mu}(p, q) = \int_0^{\beta} d\tau d\tau' e^{\nu\tau + \omega\tau'} \langle T(X_{\mathbf{q}}(\tau)\alpha_{\mathbf{p}}^{\mu}(\tau') \times \alpha_{\mathbf{p}+\mathbf{q}}^{\sigma\dagger}) \rangle$$

$$p = (\omega, \mathbf{p}), \quad q = (\nu, \mathbf{q}), \quad X(\tau) = e^{\tau H} X e^{-\tau H}. \tag{3.7}$$

The discrete energy variables ω, ν are integral multiples of $2\pi i/\beta$ and $X_{\mathbf{q}}$ is a dynamical variable such as the momentum density, which is explicitly

$$P_{\mathbf{q}}^{i} = \sum_{\mathbf{p}} p^{i} \alpha_{\mathbf{p}}^{\dagger} \alpha_{\mathbf{p}+\mathbf{q}}, \quad i = 1, 2, 3, \tag{3.8}$$

for very small $\mathbf{q}$. From Λ_{σ}^{μ} one can obtain various response functions, for example,

$$F(P_{\mathbf{q}}^{i}, P_{-\mathbf{q}}^{j}) = -T \sum_{\mathbf{p}, \omega} p^{i} \Lambda_{+}^{+}(p, q). \tag{3.9}$$

Transforming the sum over ω into an integral along the real ω axis and analytically continuing Λ

to real ν above the real ν axis, we have, ignoring the ++ indices,

$$F(P^i_{\mathbf{q}}, P^j_{-\mathbf{q}}) = -i(2\pi)^{-4}\int d^3p \int_{-\infty}^{\infty} d\omega p^i \times \left(\frac{\Lambda_{12}(p,q)}{e^{\beta\omega}-1} + \frac{\Lambda_{23}(p,q)}{e^{\beta(\omega+\nu)}-1}\right),$$

$$\Lambda_{12} \equiv \Lambda_1 - \Lambda_2, \quad \Lambda_{23} \equiv \Lambda_2 - \Lambda_3. \tag{3.10}$$

The notation needs some explanation. The vertex function, defined by analytic continuation from discrete ω, ν as a function of complex ω and ν, becomes singular whenever one of ν, ω, $\nu+\omega$ becomes real. The cuts along $\mathrm{Im}\nu = 0$, $\mathrm{Im}\omega = 0$, $\mathrm{Im}(\omega+\nu) = 0$ divide the space of two complex variables into six regions, as shown in Fig. 2. Here we are only interested in $\mathrm{Im}\nu > 0$, i.e., regions 1–3. The subscripts in (3.10) specify the region in which Λ is defined. Evidently, if we find the singularities of the vertex functions, we then will know the singularities of the response functions.

B. The Integral Equation

The vertex function satisfies the Bethe–Salpeter-type integral equation depicted in Fig. 3, where I contains no isolated two particle lines. We have

$$\Lambda^{\alpha}_{\beta}(p,q) = \mathfrak{G}^{\alpha}_{\mu}(p)\mathfrak{G}^{\nu}_{\beta}(p+q)\left(X^{\mu}_{\mathbf{q}\nu}(p) + T \times \sum_{\mathbf{p}'\omega'} I^{\mu\sigma}_{\nu\lambda}(p,p',q)\Lambda^{\lambda}_{\sigma}(p',q)\right), \tag{3.11}$$

where the source term $X^{\mu}_{\mathbf{q}\nu}$ depends on the particular form of the operator $X_{\mathbf{q}}$. For example, for the operator $P^i_{\mathbf{q}}$ given in (3.8), we have

$$X^{+}_{+} = p^i, \quad X^{-}_{-} = -p^i, \quad X^{+}_{-} = X^{-}_{+} = 0. \tag{3.12}$$

The $\mathfrak{G}$'s in (3.11) and the solid lines in Fig. 3 denote exact Green's functions defined by (3.2). For simplicity of notation, we write

$$\mathfrak{D}' \equiv \mathfrak{D}(p+q), \quad \mathfrak{N}' \equiv \mathfrak{N}(p+q), \tag{3.13}$$

and substitute $\mathfrak{N}/\mathfrak{D}$ [see (3.3)] for the $\mathfrak{G}$'s in the integral equation (3.11). In a symbolic form, it reads

$$\Lambda(\mathfrak{N}\mathfrak{N}'/\mathfrak{D}\mathfrak{D}')(X + I\Lambda). \tag{3.14}$$

We now define the new variables $(\nu', \mathbf{q}')$ by

$$\mathbf{q} = g\mathbf{q}', \quad \nu = g\nu', \tag{3.15}$$

and thus regard $(\nu, \mathbf{q})$ as quantities of $O(g)$. This is appropriate in view of (2.20). It will facilitate our discussion because, when we drop terms of higher orders in g, we automatically drop higher powers of $(\nu, \mathbf{q})$.

Now let us consider the quantity $\mathfrak{D} - \mathfrak{D}'$ [see (3.4), (3.6), and (3.13)]. We have

$$\mathfrak{D}' - \mathfrak{D} = \mathfrak{D}'_0 - \mathfrak{D}_0 + g(\mathfrak{D}'_1 - \mathfrak{D}_1) + \cdots = 2g(\omega\nu' - \mathbf{q}'\cdot\mathbf{v}\omega_p) + g(\mathfrak{D}'_1 - \mathfrak{D}_1) + \cdots, \tag{3.16}$$

$$\mathbf{v} \equiv \partial\omega_p/\partial\mathbf{p}. \tag{3.17}$$

Now we observe that, by (3.14) Λ is proportional to $(\mathfrak{D}\mathfrak{D}')^{-1}$, which is

$$\frac{1}{\mathfrak{D}\mathfrak{D}'} = \left(\frac{1}{\mathfrak{D}} - \frac{1}{\mathfrak{D}'}\right)\frac{1}{\mathfrak{D}' - \mathfrak{D}} = \frac{1}{g}\left(\frac{1}{\mathfrak{D}_0} - \frac{1}{\mathfrak{D}'_0}\right)(2\omega\nu' - 2\mathbf{q}'\cdot\mathbf{v}\omega_p + \mathfrak{D}'_1 - \mathfrak{D}_1)^{-1} + O(1). \tag{3.18}$$

Assuming $\mathrm{Im}\nu > 0$, we notice that, in regions 1 and 3 shown in Fig. 2, $\mathrm{Im}\omega$ and $\mathrm{Im}(\omega+\nu)$ have the same sign, so that $\mathfrak{D}_0$ and $\mathfrak{D}'_0$ are simple poles on the same side of the real ω axis, and therefore

$$\frac{1}{\mathfrak{D}_0} - \frac{1}{\mathfrak{D}'_0} = O(g), \quad \mathfrak{D}'_1 - \mathfrak{D}_1 = O(g), \quad \text{in regions 1 and 3,} \tag{3.19}$$

since $(\nu, q) = O(g)$. On the other hand, $\mathrm{Im}\omega$ and $\mathrm{Im}(\omega+\nu)$ have opposite signs in region 2 so that

$$\frac{1}{\mathfrak{D}_0} - \frac{1}{\mathfrak{D}'_0} = \frac{2\pi i}{2\omega_p}[\delta(\omega-\omega_p) - \delta(\omega+\omega_p)] + O(g), \tag{3.20}$$

$$\mathfrak{D}'_1 - \mathfrak{D}_1 = 2i\,\mathrm{Im}\mathfrak{D}' + O(g) \equiv 2i\omega_p\Gamma(p) + O(g) \quad \text{in region 2.} \tag{3.21}$$

The function $g\Gamma(p)$ may be interpreted as the damping rate of a phonon. Therefore, by (3.14) and (3.18)–(3.20), we see that

$$\Lambda^{\alpha}_{\beta}(p,q) \propto \frac{1}{g}\left(\frac{2\pi i}{2\omega_p}\right)[\delta(\omega-\omega_p) - \delta(\omega+\omega_p)] \times \mathfrak{N}^{\alpha}_{0\mu}\mathfrak{N}^{\nu}_{0\beta}(X + I\Lambda)^{\mu}_{\nu} + O(1) \quad \text{in region 2} = O(1) \quad \text{in regions 1 and 3,} \tag{3.22}$$

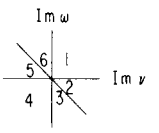


FIG. 2. Regions of analyticity of the vertex function.

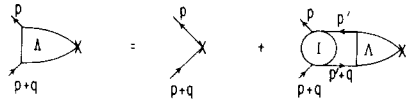


FIG. 3. Bethe-Salpeter equation satisfied by the vertex functions.

provided $I\Lambda$ is of $O(1)$, which is obviously the case as will be seen shortly. From now on, we shall only keep the leading term, i.e., the $O(1/g)$ term, in Λ. Since $\mathfrak{N}_0$ has the form (3.5), $\Lambda_+^\pm$ and Λ_+^- are of $O(1)$. Let us define $\phi^\pm(\mathbf{p},q)$ by writing

$$\Lambda_+^\pm(p,q) = (1/g)2\pi i\delta(\omega - \omega_p)\phi^+(\mathbf{p},q'),$$
$$\Lambda_-^-(p,q) = -(1/g)2\pi i\delta(\omega + \omega_p)\phi^-(\mathbf{p},q'), \tag{3.23}$$

in region 2 and zero in regions 1 and 3. The functions $\phi^\pm$ are clearly of $O(1)$. The leading term of the kernel I of the integral equation (3.14) is of $O(g)$ as given in Fig. 4. Each of the three line vertices is of $O(g^{1/2})$ and has a structure as shown in Fig. 5. Since $\Lambda = O(1/g)$, we see that $I\Lambda = O(1)$. To this order, the Green's function line in I can be replaced by the zeroth-order approximation. The term (a) in Fig. 4 generates a ladder series. The term (b) in Fig. 4 will be ignored because it, in fact, only gives an effect which is higher order in g. It generates diagrams like that shown in Fig. 6 and is responsible for the $O(g)$ correction to the zero sound-velocity and damping.

Now the integral equation (3.11) can be written as an integral equation for $\phi^\pm(\mathbf{p},q')$. We first transform the sum over ω' to integrals along the real ω' axis. After analytic continuation from discrete ω and ν to real variables in region 2, we substitute (3.23) for Λ. The energy δ functions in (3.23) allow the energy integrals to be performed easily. It turns out that only the imaginary part of I, which again is proportional to energy δ functions, contributes. All these mathematical steps are straightforward but uninteresting. Let us write down the results:

$$(\nu' - \mathbf{q}'\cdot\mathbf{v})\phi^+(\mathbf{p}) = X^+(\mathbf{p}) - i\Gamma(p)\phi^+(\mathbf{p}) + i(2\pi)^{-6}\int d^3p'd^3p''\{\sigma(pp'p'')(f'-f'')[\phi^+(\mathbf{p}'') + \phi^-(-\mathbf{p}')]$$
$$+ \sigma(p'p''p)\tfrac{1}{2}(1+f'+f'')[\phi^+(\mathbf{p}') + \phi^+(\mathbf{p}'')]\}, \tag{3.24}$$

$$(\nu' + \mathbf{q}'\cdot\mathbf{v})\phi^-(\mathbf{p}) = -X^-(\mathbf{p}) - i\Gamma(p)\phi^-(\mathbf{p}) + i(2\pi)^{-6}\int d^3p'd^3p''\{\sigma(pp'p'')(f'-f'')[\phi^-(\mathbf{p}'') + \phi^+(-\mathbf{p}')]$$
$$+ \sigma(p'p''p)\tfrac{1}{2}(1+f'+f'')[\phi^-(\mathbf{p}'') + \phi^-(\mathbf{p}')]\}, \tag{3.25}$$

where f', f'' are the Bose distribution functions

$$f' = \frac{1}{e^{\beta\omega'}-1}, \quad f'' = \frac{1}{e^{\beta\omega''}-1}, \quad \omega \equiv \omega_p, \quad \omega' \equiv \omega_{p'}, \quad \omega'' \equiv \omega_{p''}, \quad X^\pm \equiv X_\pm^\pm, \tag{3.26}$$

and

$$\sigma(pp'p'') = (2\pi)^4\delta(\omega + \omega' - \omega'')\delta(\mathbf{p} + \mathbf{p}' - \mathbf{p}'')|A(pp'p'')|^2 \tag{3.27}$$

measures the rate of the process $p + p' \to p''$, which has the amplitude $g^{1/2}A$ with

$$A(pp'p'') = \tfrac{1}{2}(\lambda\lambda'\lambda'')^{1/2}(\lambda + \lambda' - \lambda'' + 3\lambda\lambda'\lambda'') \tag{3.28}$$

and $\lambda \equiv \lambda_p = p^2/2\omega_p$ was given by (2.16). If p, p', p'' are small, we have, taking into account the energy conservation $\omega + \omega' = \omega''$,

$$A(pp'p'') = (\tfrac{9}{8}\omega\omega'\omega'')^{1/2}. \tag{3.29}$$

The dependence of $\phi^\pm(\mathbf{p})$ on $(\mathbf{q}', \nu') \equiv (g^{-1}\mathbf{q}, g^{-1}\nu)$ is understood.

Equations (3.24) and (3.25) represent the content of the general equation (3.11) in the small-g limit for small $(\mathbf{q}, \nu)$. The input to these equations are the free phonon energies, damping rate, and absorption-emission rates. These equations are exact for $g \to 0$,

$$(\mathbf{q}, \nu) \lesssim g, \quad T \lesssim 1. \tag{3.30}$$

Qualitatively, the appearance of energy δ functions reflects the fact that we are including in the integral equation only those scattering processes with very long-lasting intermediate states, so long-lasting that the particles in the intermediate states are essentially on-energy shell, so that the scattering process may be described as independent decay and absorption events put together. This point would be clearer if we formulated the

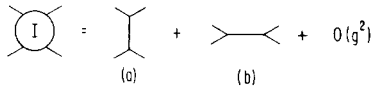


FIG. 4. Leading terms in I.

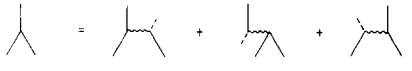


FIG. 5. Detail structure of a three-line vertex.

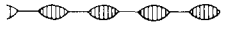


FIG. 6. Effect of the term (b) in Fig. 4 is to generate the diagrams of this kind. Each repetition of the shaded bubble turns out to give a power of g.

problem in terms of scattering amplitudes, but the mathematics would be much more complicated than what we have here.

Before solving these equations [(3.24) and (3.25)] for small T, we shall first elucidate some general features of these equations reflecting various conservation laws in the theory and transforming them into more convenient forms.

4. SOME CONSEQUENCES OF CONSERVATION LAWS

A. Energy-Momentum Conservation

Whenever one of the two operators in the response function (2.17) is a constant of motion, i.e.,

$$[X_{\mathbf{q}}, H] = 0, X_{\mathbf{q}}(t) = X_{\mathbf{q}}, \tag{4.1}$$

the response function is proportional to a simple pole at $\nu = 0$, i.e.,

$$F(X_{\mathbf{q}}, X'_{-\mathbf{q}}) = (1/\nu)\langle[X_{\mathbf{q}}, X'_{-\mathbf{q}}]\rangle \tag{4.2}$$

by simply performing the t integral. Analogous results can be derived for the vertex functions and thereby for $\phi^{\pm}$.

Let us replace the $X_{\mathbf{q}}$ in the definition of Λ [Eq. (3.7)], by $\mathbf{P}$:

$$\mathbf{P} = \sum_{\mathbf{p}} \alpha_{\mathbf{p}}^{\dagger}\alpha_{\mathbf{p}}\mathbf{P}, \tag{4.3}$$

with q set to zero. Since $\mathbf{P}$ is a constant of motion, we have

$$\begin{aligned} -\frac{\partial}{\partial\tau}\langle T(\mathbf{P}(\tau)\alpha_{\mathbf{p}}(\tau')\alpha_{\mathbf{p}}^{\dagger})\rangle \\ = -\langle T(\alpha_{\mathbf{p}}(\tau')[\mathbf{P}, \alpha_{\mathbf{p}}^{\dagger}])\rangle\delta(\tau) \\ -\langle T([\mathbf{P}(\tau), \alpha_{\mathbf{p}}(\tau)]\alpha_{\mathbf{p}}^{\dagger})\rangle\delta(\tau - \tau'). \end{aligned} \tag{4.4}$$

Since

$$[\mathbf{P}, \alpha_p^{\dagger}] = \mathbf{p}\alpha_{\mathbf{p}}^{\dagger}, \quad [\mathbf{P}, \alpha_{\mathbf{p}}] = -\alpha_{\mathbf{p}}, \tag{4.5}$$

$$\mathcal{G}_{\ddagger}^{\ddagger}(p) = -\int_0^{\beta}\langle T(\alpha_{\mathbf{p}}(\tau)\alpha_{\mathbf{p}}^{\dagger})\rangle e^{\tau\omega}d\tau, \tag{4.6}$$

we obtain from (4.4) and (3.7) that

$$\nu\Lambda_{\ddagger}^{\ddagger}(p, \nu) = \mathbf{p}[\mathcal{G}_{\ddagger}^{\ddagger}(p) - \mathcal{G}_{\ddagger}^{\ddagger}(p + \nu)]. \tag{4.7}$$

Using (3.23), it follows from (4.7) that to the leading order in g, we have

$$\nu'\phi^{+}(\mathrm{p}, \mathrm{q} = 0, \nu') = \mathbf{p}. \tag{4.8}$$

Similar arguments lead to

$$\nu'\phi^{-}(\mathbf{p}, \mathbf{q} = 0, \nu') = \mathbf{p}. \tag{4.9}$$

We can go through the above arguments using the total energy H in place of $\mathbf{P}$. Instead of (4.5), we now get

$$\begin{aligned} [H, \alpha_{\mathbf{p}}^{\dagger}] &= \omega_p\alpha_{\mathbf{p}}^{\dagger} + O(g), \\ [H, \alpha_{\mathbf{p}}] &= -\omega_p\alpha_{\mathbf{p}} + O(g). \end{aligned} \tag{4.10}$$

Keeping only the leading terms in g, we obtain identities similar to (4.8) and (4.9).

$$\begin{aligned} \nu'\phi^{+}(\mathbf{p}, \mathbf{q} = 0, \nu') &= \omega_p, \\ \nu'\phi^{-}(\mathbf{p}, \mathbf{q} = 0, \nu') &= -\omega_p. \end{aligned} \tag{4.11}$$

An immediate application of these identities is to get an expression for the damping rate $\Gamma(p)$ directly from the integral equations themselves. Let X_q be the total momentum $\mathbf{P}$. Then we have in (3.24) and (3.25)

$$X^{+}(\mathbf{p}) = \mathbf{p}, \quad X^{-}(\mathbf{p}) = -\mathbf{p}, \tag{4.12}$$

and $\mathbf{q} = 0$. We see that the identities (4.8) and (4.9) are already exhausted by the source terms X^{+} and X^{-}. Everything else on the right-hand side (rhs) must vanish. Both equations then imply, by virtue of the momentum δ function in σ, that

$$\begin{aligned} \Gamma(p) = (2\pi)^{-6}\int d^3p'd^3p''[\sigma(pp'p'')(f' - f'') \\ + \sigma(p'p''p)\tfrac{1}{2}(1 + f' + f'')]. \end{aligned} \tag{4.13}$$

The same conclusion is reached by setting $X_{\mathbf{q}} = H$, $\mathbf{q} = 0$.

B. Invariance Under Space and Time Inversions

Except for the source terms X^{+} and X^{-}, the rest of the terms on the rhs of the integral equations (3.24) and (3.25) may be regarded as an integral operator K operating on the pair (ϕ^{+}, ϕ^{-}). We see that K mixes ϕ^{+} and ϕ^{-}. The appearance of $\phi^{\pm}(-\mathbf{p}')$ instead of $\phi^{\pm}(\mathbf{p}')$ in the integrand is also unpleasant. Owing to the invariance of the theory under space and time inversions, these unpleasant features can be removed. We shall examine the explicit equations (3.24) and (3.25) directly instead of going back to Λ_{β}^{α}.

Let us define the inversion operators $\mathcal{P}$ and $\mathcal{P}'$ by

$$\mathcal{P}\phi^{\pm}(\mathbf{p}) = \phi^{\pm}(-\mathbf{p}), \quad \mathcal{P}'\phi^{\pm}(\mathbf{p}) = -\phi^{\mp}(\mathbf{p}). \tag{4.14}$$

Evidently $\mathcal{P}$ is the parity operator. $\mathcal{P}'$ is effectively $\mathcal{P} \times$ (time reversal) as far as the kernel of the integral equation is concerned. We can now construct eigenvectors of $\mathcal{P}$ and $\mathcal{P}'$ from $\phi^{\pm}(\mathbf{p})$ and $\phi^{\pm}(-\mathbf{p})$. Let

$$\begin{aligned} \varphi^{+}(\mathbf{p}) &= \tfrac{1}{4}[\phi^{+}(\mathbf{p}) - \phi^{-}(\mathbf{p}) + \phi^{+}(-\mathbf{p}) - \phi^{-}(-\mathbf{p})], \\ \varphi^{-}(\mathbf{p}) &= \tfrac{1}{4}[\phi^{+}(\mathbf{p}) + \phi^{-}(\mathbf{p}) - \phi^{+}(-\mathbf{p}) - \phi^{-}(-\mathbf{p})], \\ \bar{\varphi}^{+}(\mathbf{p}) &= \tfrac{1}{4}[\phi^{+}(\mathbf{p}) - \phi^{-}(\mathbf{p}) - \phi^{+}(-\mathbf{p}) + \phi^{-}(-\mathbf{p})], \\ \bar{\varphi}^{-}(\mathbf{p}) &= \tfrac{1}{4}[\phi^{+}(\mathbf{p}) + \phi^{-}(\mathbf{p}) + \phi^{+}(-\mathbf{p}) + \phi^{-}(-\mathbf{p})]. \end{aligned} \tag{4.15}$$

By construction, $\varphi^{\pm}, \bar{\varphi}^{\pm}$ satisfy

$$\mathcal{P}\varphi^{\pm} = \pm\varphi^{\pm}, \quad \mathcal{P}'\varphi^{\pm} = \pm\varphi^{\pm},$$
$$\mathcal{P}\varphi^{\pm} = \mp\bar{\varphi}^{\pm}, \quad \mathcal{P}'\bar{\varphi}^{\pm} = \pm\bar{\varphi}^{\pm}. \tag{4.16}$$

Since K commutes with $\mathcal{P}$ and $\mathcal{P}'$, it will not mix these eigenvectors of $\mathcal{P}$ and $\mathcal{P}'$. To obtain integral equations for $\varphi^{\pm}$ and $\bar{\varphi}^{\pm}$, we simply make linear combinations of Eqs. (3.24) and (3.25). Let us define

$$\varphi = \begin{pmatrix}\varphi^{+}\\ \varphi^{-}\end{pmatrix}, \quad \bar{\varphi} = \begin{pmatrix}\bar{\varphi}^{+}\\ \bar{\varphi}^{-}\end{pmatrix}, \quad \sigma_1 = \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix} \tag{4.17}$$

and construct similarly eigenvectors of $\mathcal{P}, \mathcal{P}'$ from the source term:

$$4x \equiv \begin{pmatrix}X^{+}(\mathbf{p}) + X^{-}(\mathbf{p}) + X^{+}(-\mathbf{p}) + X^{-}(-\mathbf{p})\\ X^{+}(\mathbf{p}) - X^{-}(\mathbf{p}) - X^{+}(-\mathbf{p}) + X^{-}(-\mathbf{p})\end{pmatrix} \equiv \begin{pmatrix}4x^{+}\\ 4x^{-}\end{pmatrix}, \tag{4.18}$$

$$4\bar{x} \equiv \begin{pmatrix}X^{+}(\mathbf{p}) + X^{-}(\mathbf{p}) - X^{+}(-\mathbf{p}) - X^{-}(-\mathbf{p})\\ X^{+}(\mathbf{p}) - X^{-}(\mathbf{p}) + X^{+}(-\mathbf{p}) - X^{-}(-\mathbf{p})\end{pmatrix} \equiv \begin{pmatrix}4\bar{x}^{+}\\ 4\bar{x}^{-}\end{pmatrix}. \tag{4.19}$$

Then we have, from (3.24) and (3.25),

$$(\nu' - \mathbf{q}'\cdot\boldsymbol{\nu}\sigma_1)\varphi/s = x/s - i\mathcal{K}(\varphi/s), \tag{4.20}$$

$$(\nu' - \mathbf{q}'\cdot\boldsymbol{\nu}\sigma_1)\bar{\varphi}/s = \bar{x}/s - i\bar{\mathcal{K}}(\bar{\varphi}/s), \tag{4.21}$$

where

$$s \equiv 2\sinh\tfrac{1}{2}\beta\omega_p \tag{4.22}$$

and $\mathcal{K}, \bar{\mathcal{K}}$ are now real and symmetric integral operators:

$$\mathcal{K}(\varphi/s) = (2\pi)^{-6}\int d^3p'd^3p''(s's'')^{-1} \times \{\sigma(pp'p'')[\varphi(\mathbf{p}) + \varphi(\mathbf{p}') - \varphi(\mathbf{p}'')] + \sigma(p'p''p)\tfrac{1}{2}[\varphi(p) - \varphi(\mathbf{p}') - \varphi(\mathbf{p}'')]\}, \tag{4.23}$$

$$\bar{\mathcal{K}}(\bar{\varphi}/s) = (2\pi)^{-6}\int d^3p'd^3p''(s's'')^{-1} \times [\sigma(pp'p'') + \tfrac{1}{2}\sigma(p'p''p)][\bar{\varphi}(p) - \bar{\varphi}(p') - \bar{\varphi}(p'')]. \tag{4.24}$$

Of course, s' and s'' are defined by (4.22) with p replaced by p' and p'', respectively. We have used (4.13) for $\Gamma(p)$ and taken advantage of the identities

$$f' - f'' = s(s's'')^{-1} \quad \text{if} \quad \omega'' = \omega + \omega',$$
$$1 + f' + f'' = s(s's'')^{-1} \quad \text{if} \quad \omega = \omega' + \omega''. \tag{4.25}$$

It will be seen later that only (4.20) gives rise to second sound but not (4.21). However, we shall carry (4.21) along until this becomes transparent.

C. Invariance Under Rotation

It is evident from (4.23) and (4.24) that $\mathcal{K}$ and $\bar{\mathcal{K}}$ are both rotationally invariant. As a consequence, the solutions of the integral equations without the $\mathbf{q}'\cdot\boldsymbol{\nu}\sigma_1$ term can be classified by (l, m), the angular momentum quantum numbers and $\mathcal{K}$ and $\bar{\mathcal{K}}$ are diagonal in the (l, m) representation. We shall leave the explicit construction to the next section.

5. SOLUTION OF THE INTEGRAL EQUATION

A. General Feature of the Kernel and the Second-Sound Pole

We are looking for singular solutions to the integral equations (4.20) and (4.21) for small q' which behave like simple poles near the real ν' axis. These poles may then be interpreted as reflecting the existence of the second sound. The precise location of the poles will tell us about the speed of propagation and the damping rate of the second sound. The location of poles is obtained by solving the homogeneous part of the integral equations (4.20) and (4.21), i.e., the eigenvalue equations

$$\lambda\Psi = (\mathcal{K} + \mathcal{K}')\Psi, \quad \bar{\lambda}\bar{\Psi} = (\bar{\mathcal{K}} + \mathcal{K}')\bar{\Psi}, \tag{5.1}$$

$$\mathcal{K}' \equiv i\mathbf{q}'\cdot\boldsymbol{v}\sigma_1 \equiv iq'v\cos\theta\sigma_1. \tag{5.2}$$

The poles on the ν' plane are then located at

$$\nu' = -i\lambda, \ -i\bar{\lambda}. \tag{5.3}$$

The z axis is taken along the direction of $\mathbf{q}'$. Since we are interested in the small-q' limit, the operator $\mathcal{K}'$ can be treated as a perturbation after we solve the equations without it. Since $\mathcal{K}, \bar{\mathcal{K}}$ are real and symmetric, their eigenvalues are all real. Also, the eigenvalues must be all positive because the integral equations were derived with ν in the upper half-plane and therefore there can be no singularity for $\mathrm{Im}\nu > 0$, i.e., all $\lambda, \bar{\lambda}$'s must be positive. To see this explicitly, let us write down the bilinear forms $(\varphi, \mathcal{K}\varphi)$, $(\bar{\varphi}, \bar{\mathcal{K}}\bar{\varphi})$. Here the scalar product is defined by

$$(\varphi, \varphi') = \int d^3p(2\pi)^{-3}\varphi(\mathbf{p})\varphi'(\mathbf{p})s^{-2}. \tag{5.4}$$

Writing $\Psi \equiv \varphi/s$, $\bar{\Psi} \equiv \bar{\varphi}/s$ and utilizing the symmetry properties of the integrands, we obtain from (4.23) and (4.24)

$$(\varphi, \mathcal{K}\varphi) = (2\pi)^{-9}\int d^3pd^3p'd^3p''(ss's'')^{-1} \times \tfrac{1}{2}\sigma(p'p''p)[\varphi(\mathbf{p}) - \varphi(\mathbf{p}') - \varphi(\mathbf{p}'')]^2, \tag{5.5}$$

$$(\bar{\varphi}, \bar{\mathcal{K}}\bar{\varphi}) = (2\pi)^{-9}\int d^3pd^3p'd^3p''(ss's'')^{-1} \times \tfrac{1}{4}\sigma(pp'p'')\{[\bar{\varphi}(\mathbf{p}) - \bar{\varphi}(\mathbf{p}')]^2 + [\bar{\varphi}(\mathbf{p}') - \bar{\varphi}(\mathbf{p}'')]^2 + [\bar{\varphi}(\mathbf{p}'') - \bar{\varphi}(\mathbf{p})]^2\}. \tag{5.6}$$

Since $\varphi, \bar{\varphi}$ are arbitrary, and $\sigma(pp'p'')$ is positive, it is then obvious that $\mathcal{K}$ and $\bar{\mathcal{K}}$ are positive definite.

It is also evident that the lowest eigenvalue of $\mathcal{K}$ and $\bar{\mathcal{K}}$ is zero. For $\mathcal{K}$, we see that (5.5) vanishes if $\varphi(\mathbf{p}) - \varphi(\mathbf{p}') - \varphi(\mathbf{p}'')$ vanishes identically. This happens if and only if

$$\varphi(\mathbf{p}) \propto \omega_p,$$

or

$$\varphi(\mathbf{p}) \propto \mathbf{p}, \tag{5.7}$$

in view of the energy-momentum δ function in $\sigma(pp'p'')$ [see (3.27)]. For $\bar{\mathcal{K}}$, we see that (5.6) vanishes if and only if

$$\bar{\varphi}(\mathbf{p}) \text{ is independent of } \mathbf{p}. \tag{5.8}$$

As we shall point out later, when $\sigma(pp'p'')$ takes certain special forms, there can be additional eigenfunctions with zero eigenvalue other than (5.7) and (5.8). In general, there is no more.

Of course, positive eigenvalues will not give rise to any propagating wave because $\nu' = -i$ (positive number) describes an exponentially damped amplitude. Therefore, we are only interested in the eigenvalue zero. The perturbation term $\mathcal{K}' = i\mathbf{q}' \cdot \boldsymbol{\nu}\sigma_1$, being imaginary, will supply an imaginary λ and hence a real ν' proportional to q' by a first-order perturbation calculation. This eigenvalue can be identified as describing a propagating wave. The second-order perturbation term will then give a real positive correction to λ, i.e., a damping term for the wave proportional to q'^2.

Since the eigenvalue $\bar{\lambda} = 0$ for $\bar{\mathcal{K}}$ is non degenerate and the wavefunction is a constant, the first-order eigenvalue is zero since the perturbing term $\mathcal{K}' \propto \cos\theta\sigma_1$ is purely nondiagonal and odd in parity. Thus, $\bar{\mathcal{K}}$ will not give us any sound wave and we shall concentrate on $\mathcal{K}$ alone from now on.

To sum up, our procedure of locating the second-sound frequency is to locate the nearly imaginary eigenvalues of $\mathcal{K} + \mathcal{K}'$: We start with the eigenvalue zero of $\mathcal{K}$ and calculate the correction due to $\mathcal{K}'$ by the perturbation theory. Since $\mathcal{K}'$ is purely imaginary and proportional to q', we have

$$-i(\text{eigenvalue of } \mathcal{K} + \mathcal{K}')$$

$$= c_2 q' - i\tau' q'^2 + O(q'^3), \tag{5.9}$$

where c_2 and τ' can be obtained using the elementary perturbation method in quantum mechanics.

B. First-Order Eigenvalue and the Speed of the Second Sound

No further information concerning $\mathcal{K}$ is needed to calculate the first-order correction to λ due to the perturbation $\mathcal{K}'$.

The eigenvalue zero of $\mathcal{K}$ is fourfold degenerate as (5.7) shows. The two relevant eigenfunctions are (not yet normalized),

$$\Psi_0(\mathbf{p}) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \frac{\omega}{s},$$

$$\Psi_1(\mathbf{p}) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \frac{p\cos\theta}{s}. \tag{5.10}$$

The other two eigenfunctions, being proportional to p_x and p_y, give rise to zero matrix elements. The matrix representation of $\mathcal{K}'$ in this two-dimensional space spanned by Ψ_0 and Ψ_1 is

$$\mathcal{K}' = iq' c_2 \sigma_1, \tag{5.11}$$

which clearly has the eigenvalues $\lambda = \pm iq' c_2$, where

$$c_2 = \frac{1}{\sqrt{3}} \left(\int_0^\infty dpp^3 \omega v s^{-2} \right) \left(\int_0^\infty dpp^2 \omega^2 s^{-2} \right)^{-1/2} \times \left(\int_0^\infty dpp^4 s^{-2} \right)^{-1/2}$$

$$= (1/\sqrt{3})[1 + \tfrac{15}{14}\pi^2 T^2 + O(T^4)]. \tag{5.12}$$

Recall that $s = 2\sinh\frac{1}{2}\beta\omega$ [see (4.23)], and p can be expanded in powers of ω:

$$p = \omega(1 - \tfrac{1}{8}\omega^2 + \tfrac{7}{128}\omega^4 - \cdots), \tag{5.13}$$

and $v = d\omega/dp$. The integrals in (5.12) are easily performed using this expansion. We have therefore found poles on ν' plane located at

$$\nu' = -i\lambda = \pm c_2 q'. \tag{5.14}$$

Since $\nu'/q' = \nu/q = c_2$, the speed of propagation is thus given by (5.12). The zeroth-order eigenfunctions are given by

$$\Psi_\pm \equiv (\Psi_0 \pm \Psi_1)2^{-1/2}, \tag{5.15}$$

where Ψ_0, Ψ_1 must be normalized to unity.

The first-order result is valid only if higher-order corrections are small. We must at least calculate the second-order term to see over what range of q' (5.14) makes sense. The second-order correction to λ is given by the formula

$$\sum_n{}' \frac{(\varphi_\pm, \mathcal{K}'\varphi_n)^2}{-\lambda_n} = -q'^2 \sum_n \frac{1}{-\lambda_n} (\varphi_n, \sigma_1 \cos\theta \varphi_\pm)^2 \equiv \tau' \tag{5.16}$$

in an obvious notation. Evidently, we must know more about the other eigenfunctions and eigenvalues of $\mathcal{K}$ in order to evaluate (5.16). We now proceed to study $\mathcal{K}$ in more detail.

C. Angular Momentum Decomposition

We have mentioned before that $\mathcal{K}$ is rotationally invariant and therefore its eigenfunctions can be

classified by (l,m), the angular momentum indices. Since $\mathcal{K}' \propto \cos\theta$, only $l=2, m=0$ enters in (5.16). It turns out that, for the leading term in the low-temperature limit, only the lowest $l=2$ level is needed to evaluate (5.16). Nevertheless, in order to exhibit some qualitative features of $\mathcal{K}$, we shall discuss an arbitrary l. We now classify the eigenfunctions by l and write

for l even,

$$\Psi_l(\mathbf{p}) = \begin{pmatrix}1\\0\end{pmatrix} P_l(\cos\theta)\left(\frac{2l+1}{2}\right)^{1/2} \frac{\chi_l(\omega)}{s},$$

for l odd,

$$\Psi_l(\mathbf{p}) = \begin{pmatrix}0\\1\end{pmatrix} P_l(\cos\theta)\left(\frac{2l+1}{2}\right)^{1/2} \frac{\chi_l(\omega)}{s}. \quad (5.17)$$

Recall that the two-component notation comes from the separation of even- and odd-parity functions [see (4.16) and (4.17)].

Substituting (5.17) in (4.23), we can perform the angular integral with the aid of the δ functions contained in $\sigma(pp'p'')$ [see (3.27)]. We obtain the radial part of $\mathcal{K}\Psi_l$ after a little algebra:

$$\begin{aligned}\mathcal{K}\chi_l = (2\pi p)^{-1} \int_0^\infty dp'dp''p'p''(s's'')^{-1} \\ \times \{\delta(\omega''-\omega'-\omega)|A(pp'p'')|^2[\chi_l(\omega) \\ + P_l(x')\chi_l(\omega') - P_l(x'')\chi_l(\omega'')] \\ + \delta(\omega-\omega'-\omega'')|A(p'p''p)|^2\tfrac{1}{2}[\chi_l(\omega) \\ - P_l(x')\chi_l(\omega') - P_l(x'')\chi_l(\omega'')]\},\end{aligned} \quad (5.18)$$

where

$$x' \equiv \hat{p}\cdot\hat{p}', \qquad x'' \equiv \hat{p}\cdot\hat{p}''. \quad (5.19)$$

are fixed by the triangles determined by the three energy variables ω, ω', and ω'' (see Fig. 7). Clearly, for $l=0, \chi_l = \omega$, and for $l=1, \chi_l = p$, the square brackets in (5.18) vanish as we have seen before.

This is as far as we can go without looking at the detail form of the amplitude A or the dependence of p on ω. Equation (5.18), as it is, appears to be very complicated, but it becomes tractable under the assumption of low temperature.

D. Low-Temperature Expansion of the Lowest Eigenvalues

Leaving higher eigenvalues to later discussion, let us determine the lowest eigenvalue of $\mathcal{K}$ for a definite l. We have the eigenvalue equation, leaving the subscript l understood:

$$\mathcal{K}\chi = \lambda\chi, \quad (5.20)$$

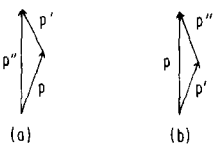


FIG. 7. Geometry of the three phonon momenta in the terms of the integral equation.

with the left-hand side (lhs) given by (5.18). Before doing systematic expansion in powers of the temperature, we note the following fact: At low temperatures, only phonons of very small momenta can appear. In the small-momentum limit, we have

$$p = \omega. \quad (5.21)$$

It then follows from energy-momentum conservation that

$$x' = x'' = 1, \qquad P_l(x') = P_l(x'') = 1, \quad (5.22)$$

i.e., $\mathbf{p}, \mathbf{p}'$, and $\mathbf{p}''$ are all parallel. In this case it is obvious that if

$$\chi(\omega) \propto \omega, \quad (5.23)$$

then the square brackets in (5.18) vanish, i.e., $\mathcal{K}\chi = 0$ for all l. In other words, the zero eigenvalue appears in more than $l=0,1$ noted before and the first-order calculation for the second-sound pole given in Sec. 5B is not meaningful in this limit. This is a general result. Whenever the phonon energy is strictly proportional to its momentum, emission and absorption processes cannot give rise to second-sound propagation. Physically this situation is clear. If gas particles always move along straight lines without ever changing directions, there cannot be any sound wave.

Here p is not strictly proportional to ω when higher-order terms [see (5.13)] are included. However, the deviation from $p=\omega$ is very small and delicate.

To expand in powers of T, it is convenient to change the variables $\omega, \omega'\omega''$ to y, y', y'' defined by

$$y = \beta\omega, \qquad y' = \beta\omega', \qquad y'' = \beta\omega''. \quad (5.24)$$

Then the functions s, s', s'' become

$$s = 2\sinh\tfrac{1}{2}y, \text{ etc.}, \quad (5.25)$$

and the expansion (5.13) for p in terms of ω becomes

$$p = Ty(1 - \tfrac{1}{8}T^2y^2 + \tfrac{7}{128}T^4y^4 + \cdots). \quad (5.26)$$

Expanding all the p variables in (5.18) by (5.26), we see that $\mathcal{K}$ is an expansion of the form

$$T^{-5}\mathcal{K} = \mathcal{K}^{(0)} + T^2\mathcal{K}^{(1)} + T^4\mathcal{K}^{(2)} + \cdots. \quad (5.27)$$

The fact that the leading term of $\mathcal{K}$ is of $O(T^5)$ can be seen easily from (5.18). One simply counts the power of p and takes into account the fact that $|A|^2 \propto \omega\omega'\omega''$, as shown in (3.29).

We can also expand the eigenfunction and the eigenvalue in powers of T^2:

$$\chi = \chi^{(0)} + \chi^{(1)}T^2 + \chi^{(2)}T^4 + \cdots, \quad (5.28)$$

$$T^{-5}\lambda = \lambda^{(0)} + \lambda^{(1)}T^2 + \lambda^{(2)}T^4 + \cdots. \quad (5.29)$$

The eigenvalue equation $\mathcal{K}\chi = \lambda\chi$ is then expanded.

We have

$$\mathcal{K}^{(0)}\chi^{(0)} = \lambda^{(0)}\chi^{(0)},$$
$$\mathcal{K}^{(0)}\chi^{(1)} + \mathcal{K}^{(1)}\chi^{(0)} = \lambda^{(1)}\chi^{(0)} + \lambda^{(0)}\chi^{(1)},$$
$$\mathcal{K}^{(0)}\chi^{(2)} + \mathcal{K}^{(1)}\chi^{(1)} + \mathcal{K}^{(2)}\chi^{(0)} = \lambda^{(1)}\chi^{(1)} + \lambda^{(2)}\chi^{(0)} + \lambda^{(0)}\chi^{(2)}, \quad \text{etc.} \tag{5.30}$$

Before proceeding, let us define the scalar product for the radial wavefunctions:

$$(\chi, \phi) \equiv \int_0^\infty dy\, y^2 s^{-2} \chi(y)\phi(y). \tag{5.31}$$

From what we just pointed out above, the lowest eigenvalue of $\mathcal{K}^{(0)}$ is zero since keeping the leading term in (5.26) implies $p = \omega$. Also $\chi^{(0)}$ is known via (5.23):

$$\lambda^{(0)} = 0, \quad \chi^{(0)}(y) = y, \tag{5.32}$$

apart from a normalization factor.

Let us now expand $\mathcal{K}$ in T^2. It is sufficient for the moment to concentrate on the square brackets in (5.18) which are the delicate parts. Everything outside the brackets are multiplicative factors and can be expanded separately later if necessary. Consider the first square bracket in (5.18). From Fig. 7a, we have

$$x' = \hat{p}\cdot\hat{p}' = (p''^2 - p^2 - p'^2)/2pp',$$
$$x'' = \hat{p}\cdot\hat{p}'' = (p^2 + p''^2 - p'^2)/2pp''. \tag{5.33}$$

Using the expansion (5.26) for the p variables, one finds

$$P_l(x') = 1 - \tfrac{3}{16}l(l+1)y''T^2 + \tfrac{1}{256}l(l+1)y''^2 \times \{3yy' - \tfrac{9}{2}y''^2[1 - \tfrac{1}{2}l(l+1)] - 35(y''^2 - yy')\}\, T^4 + O(T^6), \tag{5.34}$$

$$P_l(x'') = 1 - \tfrac{3}{16}l(l+1)y'^2T^2 - \tfrac{1}{256}l(l+1)y'^2 \times \{3yy'' + \tfrac{9}{2}y'^2[1 - \tfrac{1}{2}l(l+1)] + 35(y''^2 - yy')\}\, T^4 + O(T^6). \tag{5.35}$$

Let us write $\chi^{(1)} = a^{(1)}y^3$, where $a^{(1)}$ is a constant. Then the first square bracket of (5.18) becomes

$$[-\tfrac{3}{16}l(l+1) - 3a^{(1)}]yy'y''T^2 + O(T^4). \tag{5.36}$$

Therefore, if we set

$$\chi^{(1)}(y) = -\tfrac{1}{16}l(l+1)y^3, \tag{5.37}$$

then the first square bracket vanishes to $O(T^2)$. A similar calculation shows that the second square bracket of (5.18) also vanishes if we apply (5.37). Let us summarize what we have so far:

$$\lambda^{(0)} = 0, \quad \lambda^{(1)} = 0,$$
$$\chi^{(0)} = y, \quad \chi^{(1)} = -\tfrac{1}{16}l(l+1)y^3. \tag{5.38}$$

With this information we get a formula for $\lambda^{(2)}$ from (5.30):

$$\lambda^{(2)} = (\chi^{(0)}, \mathcal{K}^{(1)}\chi^{(1)} + \mathcal{K}^{(2)}\chi^{(0)})/(\chi^{(0)}, \chi^{(0)}). \tag{5.39}$$

This requires the $O(T^4)$ term of (5.36). A little algebra shows that this term is

$$[\tfrac{9}{4}y'y''(y''^3 - y'^3) - 3yy'^2y''^2] \times l(l+1)[\tfrac{1}{2}l(l+1) - 1](128)^{-1}T^4 + yy'y''(y''^2 - yy')l(l+1)\tfrac{35}{256}T^4. \tag{5.40}$$

The T^4 term of the second square bracket of (5.18) can be calculated in the same manner. We find

$$-[\tfrac{9}{4}y'y''(y'^3 + y''^3) + 3yy'^2y''^2] \times l(l+1)[\tfrac{1}{2}l(l+1) - 1](128)^{-1}T^4 - yy'y''(y^2 - y'y'')l(l+1)\tfrac{35}{256}T^4. \tag{5.41}$$

Substituting (5.40) and (5.41) for the square brackets in (5.18), the scalar product in (5.39) can now be calculated. We find

$$\lambda^{(2)} = l(l+1)[\tfrac{1}{2}l(l+1) - 1]C, \tag{5.42}$$

where the constant C is given by

$$C = 5\left(\frac{3}{8\pi}\right)^5 \int_0^\infty dydy'dy''(ss's'')^{-1}(yy'y'')^3 \times \delta(y - y' - y'') = \tfrac{5}{4}\left(\frac{3}{8\pi}\right)^5 \int_0^1 dx x^3(1-x)^3 \times \int_0^\infty dy y^{10} \coth\left(\frac{xy}{2}\right) \operatorname{csch}^2\left(\frac{y}{2}\right) = 3.353. \tag{5.43}$$

Thus, we conclude that the lowest eigenvalue of $\mathcal{K}$ in the low-temperature limit is, putting the subscript l in now for clarity,

$$\lambda_l = \lambda_l^{(2)}T^9 + O(T^{11}). \tag{5.44}$$

In view of (5.42), $\lambda_l^{(2)}$ vanishes for $l = 0, 1$ as it should.

E. Higher Eigenvalues

Higher eigenvalues can also be considered as expansions in T^2. We shall be satisfied here with an estimate of their leading temperature dependence. Consider, for an aribtrary $\chi(y)$, the bilinear form $(\chi, \mathcal{K}^{(0)}\chi)$. We have

$$(\chi, \mathcal{K}^{(0)}\chi) = \frac{9}{16\pi}\int_0^\infty dydy'dy''(ss's'')^{-1} \times (yy'y'')^2\delta(y - y' - y'')^{\frac{1}{2}}\,[\chi(y) - \chi(y') - \chi(y'')]^2. \tag{5.45}$$

This, of course, can be viewed as a special case of (5.5). As long as $\chi(y)$ is not $\chi^{(0)}(y) = y$, (5.45)

is of $O(1)$. Therefore, we conclude that all higher eigenvalues of $\mathcal{K}$ are proportional to T^5 and independent of l in the low-temperature limit.

F. Second-Order Perturbation and Damping of Second Sound

We are now in a position to evaluate the second-order term (5.16). The matrix element is nonzero only for $l = 2$ intermediate states. Since the lowest eigenvalue is of $O(T^9)$, while all the others are of $O(T^5)$, we need to keep only the lowest level in $l = 2$ levels to calculate the leading term in (5.16). The matrix element can be easily evaluated. We get for (5.16)

$$\tau' = \frac{2T^{-9}}{15\lambda\binom{2}{2}} + O(T^{-7}) = 0.003\ 314T^{-9} + O(T^{-7}), \tag{5.46}$$

which corresponds to a term in the imaginary part of ν':

$$\mathrm{Im}\nu' = -\tau' q'^2. \tag{5.47}$$

G. Summary of Results of Calculation

Together with (5.12), we have the frequency of the second-sound ν in terms of its wavenumber q. Recall that $(\nu', q') \equiv (\nu/g, q/g)$. We have

$$\nu = c_2 q - i\tau q^2 + O(q^3), \tag{5.48}$$

$$c_2 = 3^{-1/2}(1 + 10.57\ T^2) + O(T^4) + O(g),$$
$$\tau = 0.003\ 314 g^{-1}T^{-9} + g^{-1}O(T^{-7}) + O(1). \tag{5.49}$$

The damping must be much smaller than the real part of ν in order to have a propagating wave. We therefore must have

$$\nu \ll gT^9, \tag{5.50}$$

which is far more restrictive than the qualitative estimate $\nu \ll gT^5$ one would guess [see (2.20)]. The reason is that owing to the near proportionality of phonon energy to its momentum, the direction of propagation of a phonon is not easily altered by emission or absorption. This in a sense greatly hinders the 'relaxation' of phonons.

Clearly, it is a matter of further straightforward algebra to obtain high-order terms in T^2 for c_2 and τ. For the $O(T^{-7})$ term in τ, the lowest $l = 2$ level is still sufficient. For the $O(T^{-5})$ term, all higher $l = 2$ levels must be counted.

6. DISCUSSION AND SUMMARY

A. Oscillations in the Phonon Distribution Function

We have mentioned before that, in contrast to the zero sound, the second sound reflects oscillations between certain sets of excited states. In the zeroth-order approximation, the excited states are free phonons. Qualitatively, oscillations between excited states can be described as oscillations in the *distribution function*, which specifies the number of phonons per phase space volume.

To relate the concept of distribution function to the function $\varphi(\mathbf{p})$, we go back to (3.10), where a response function is expressed in terms of the vertex function. Substituting the expression (3.23) for Λ_+^+ in terms of ϕ^+ into (3.10), we have, for small ν

$$F(P_{\mathbf{q}}^i, P_{-\mathbf{q}}^j) = (2\pi)^{-3}\int d^3p p^i\, \delta f(\mathbf{p}, \mathbf{q}, \nu), \tag{6.1}$$

where we have defined

$$\delta f(\mathbf{p}, \mathbf{q}, \nu) = \phi^+(\mathbf{p}, \mathbf{q}', \nu')\nu' \frac{\partial f}{\partial \omega_p}, \tag{6.2}$$

and

$$f = \frac{1}{e^{\beta\omega_{p-1}}}, \nu' = \frac{\nu}{g} \tag{6.3}$$

as before. We have written out the $(\mathbf{q}', \nu')$ arguments of ϕ^+ explicitly.

The form (6.1) strongly suggests that δf be interpreted as the change in the distribution function due to an external perturbation. By the definition of response function,

$$F(P_{\mathbf{q}}^i, P_{-\mathbf{q}}^j)\, e^{i(\mathbf{q}\cdot\mathbf{x} - \nu t)} \tag{6.4}$$

is the average of momentum density at $(\mathbf{x}, t)$ if there is an external disturbance described by the perturbing Hamiltonian

$$H' = P_{-\mathbf{q}}^j e^{i\mathbf{q}\cdot\mathbf{x} - i\nu t}. \tag{6.5}$$

It is then clear that

$$\delta f(\mathbf{p}, \mathbf{q}, \nu) e^{i\mathbf{q}\cdot\mathbf{x} - i\nu t} \tag{6.6}$$

is the change in distribution function from the equilibrium one at $(\mathbf{x}, t)$ as a result of an external perturbation. This interpretation holds if $P_{-\mathbf{q}}^j$ is replaced by another operator. The poles of ϕ^+ near the real ν axis thus describes oscillations in the distribution function.

Now we return to the eigenfunctions in (5.7) and (5.8) for the eigenvalue zero of $\mathcal{K}$. For $\mathbf{q}' = 0$, it is clear that (4.20) and (4.21) have the solutions

$$\begin{aligned} \varphi_0 &= \delta\beta \binom{1}{0} \frac{\omega_p}{\nu'}, \\ \varphi_1 &= \delta u^i \binom{0}{1} \frac{p^i}{\nu'}, \\ \bar{\varphi}_0 &= \delta\epsilon \binom{1}{0} \frac{1}{\nu'}, \end{aligned} \tag{6.7}$$

where we have arbitrarily written $\delta\beta, \delta u^i$, and $\delta\epsilon$ for the source terms. Substituting (6.7) in (6.2), we obtain

$$\delta f \sim \delta\beta\omega_p \frac{\partial f}{\partial \omega_p} \sim \frac{1}{e^{(\beta+\delta\beta)\omega_p} - 1} - f, \tag{6.8}$$

$$\delta f \sim \delta \mathbf{u}\cdot\mathbf{p}\,\frac{\partial f}{\partial \omega_p} \sim \frac{1}{e^{\beta(\omega_p+\delta \mathbf{u}\cdot\mathbf{p})}-1} - f, \tag{6.9}$$

$$\delta f \sim \delta\epsilon\,\frac{\partial f}{\partial \omega_p} \sim \frac{1}{e^{\beta(\omega_p+\delta\epsilon)}-1} - f. \tag{6.10}$$

Thus (6.8) may be interpreted as a change in inverse temperature $\delta\beta$, (6.9) a uniform notion of the phonon gas with velocity $\delta\mathbf{u}$, and (6.10) a uniform shift of energy scale. When the term $\mathbf{q}'\cdot\mathbf{v}\sigma_1$ is turned on in (4.20) and (4.21), we showed that the second sound appears as a result of the coupling of φ_0 and φ_1, while $\bar{\varphi}_0$ gives no propagating wave. Therefore we can interpret the second sound, in view of (6.8) and (6.9) as a wave of drifting velocity of phonons coupled with temperature oscillations. For discussions along this line in problems of crystal lattices, the reader is referred to Refs. 6 and 7. Although the above physical picture is appealing, we wish to emphasize the qualitative nature of the interpretation in terms of oscillating distribution function, which is a concept valid only in the noninteracting limit.

B. Relaxation Time

For the second-sound problem in solids, there has been no concrete procedure from which the damping of second sound can be calculated. The authors in Refs. 6 and 7 simply introduced a relaxation time to count for the effect of a sum like (5.16). Their procedure is clearly consistent with our approach. In this paper, we have a well-defined model, which allows us to calculate the damping rate from first principles and display the physical mechanism responsible for it.

C. Higher-Order Terms

Keeping only the leading order in g seems to be essential in getting a tractable integral equation. This is rather unfortunate because we expect some of the higher-order processes such as phonon–phonon scattering to be important. They would give rise to large-angle scattering, which is more effective in supporting the second sound than the absorption–emission processes appearing in the leading equation. As was mentioned before, the absorption–emission processes are nearly all in the forward direction and are ineffective. Thus, higher-order terms may give qualitatively different temperature dependence of the damping rate.

The next-order calculation will also produce a correction to c_2 due to the lowest-order interaction between phonons. It is clear from our calculation in Sec. 5 that c_2 depends only on the free phonon dispersion curve and has nothing to do with the interaction between phonons even though the damping rate does depend on the interaction through $\sigma(pp'p'')$. This is not surprising in view of the following observation.

Our integral equations are equations at zero g limit, i.e., qualitatively a zero phonon density limit. Like the Boltzmann equation for dilute gases, it depends only on the *cross sections* of reactions, i.e., the square of the amplitudes A, but not the phase of the amplitudes. The quantity c_2 is related to the "compressibility" of the phonon gas which is obtained by differentiating the free energy of the interacting phonon gas. We know that the Boltzmann equation contains *no information* concerning the free energy except for the ideal-gas term, even though it can count for transport coefficients. The free energy depends on the *phase* of the scattering amplitude and there is no information of this kind in the Boltzmann equation.[8] Our integral equation contains no phase information in the phonon decay or absorption amplitude. Therefore only the compressibility for the *ideal* phonon gas can be obtained, even though the interaction plays a crucial role in the propagation and damping of the second sound.

D. Summary

We have described the phenomenon of second sound with a model containing only one dynamical parameter g. This is very likely the simplest possible model which allows a realization of second sound. In contrast to a general hydrodynamical approach, we have taken a specific perturbation theory approach. We started with the Bethe–Salpeter equation (3.11) after a qualitative investigation. We extracted the leading terms in the small-g limit. After taking symmetry principles into account, we arrived at the integral equations (4.20) and (4.21). The eigenvalues of the homogeneous equations were studied by repeated applications of perturbation theory treating the wavenumber q' and T^2 as small parameters. The eigenvalues giving rise to a nearly real frequency are found and interpreted as the second sound. The speed and damping rate are evaluated for small T, [(5.48) and (5.49)]. Finally, some qualitative interpretations are sketched.

The physical picture and mathematical apparatus are very elementary. Our analysis is strictly within the framework of equilibrium statistical mechanics. No concept of local thermal equilibrium is used. We have made no use of any macroscopic concepts such as viscosity or thermal conductivity. Indeed, the next step is to understand these concepts in a more precise manner in terms of a microscopic theory.

ACKNOWLEDGMENT

It is a pleasure to thank Professor L. J. Sham for suggesting this research and for helpful conversations.

* Research supported by the U.S. Atomic Energy Commission, under Contract No. AT(11-1)-GEN-10, P.A.11.

1 The literature on the second sound in He II is too extensive to be cited here. For complete references, the reader is referred to I. M. Khalatnikov, *Introduction to the Theory of Superfludity* (Benjamin, New York, 1965), and J. Wilks, The *Properties of Liquid and Solid Helium* (Clarendon, Oxford, England, 1967).

2 See collected papers in D. Pines, Many-Body Problem (Benjamin, New York, 1962).

3 J. Gavoret and P. Nozieres, Ann Phys. (N.Y.) **28**, 349 (1969).

4 P. Hohenberg and P. C. Martin, Ann. Phys. (N.Y.) **34**, 291 (1965), C. E. Carroll, Phys Rev. A **2**, 497 (1970).

5 S. Ma, H. Gould and V. K. Wong, Phys. Rev. A **3**, 1453 (1971).

6 For example, see J. A. Krumhansl, Proc. Phys. Soc. (London) **85**, 921 (1965), R. A. Guyer and J. A. Krumhansl, Phys. Rev. A 1411 (1964). See also P. C. Kwok and P. C. Martin, Phys. Rev. **142**, 495 (1966).

7 L. J. Sham, Phys. Rev **156**, 494 (1967); **163**, 401 (1967).

8 For a general discussion along this line, see R. Dashen and S. Ma. J. Math Phys. **12**, 1449 (1971).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 10 OCTOBER 1971

Radiation Scattering in Einstein–Maxwell Theory

W. E. Couch*

Department of Physics, University of Pittsburgh, Pittsburgh, Pennsylvania 15213

and

W. H. Hallidy

Department of Physics, California State College, California, Pennsylvania 15419

(Received 13 May 1971)

A linearized Einstein–Maxwell field composed of a mass monopole, a radiating mass quadrupole, a Coulomb field, and radiating electromagnetic dipole and quadrupole fields is examined to second order. If the first-order radiation is a pulse of finite duration, then the only second-order radiation present after the pulse has turned off is due to interactions involving the mass monopole. All other interactions contributing to the second order produce no such scattered radiation. At third order, the dipole–dipole-mass interaction is found to give rise to scattered radiation.

1. INTRODUCTION

Nonlinear effects that show some interesting aspects of the interaction of gravitational radiation fields in empty space–time with themselves and with mass monopoles have recently been discovered.[1-3] These previous investigations made use of a perturbative approximation scheme to calculate the nonlinear effects. In this paper we apply the same perturbative scheme to the case where an electromagnetic field is present in space-time and investigate certain interactions of electromagnetic fields with gravitational fields and with themselves.

We assume that all quantities occurring in our formulation of Einstein–Maxwell theory are expandable in a perturbation series in which the zeroth order represents flat space–time with no electromagnetic field present, the first-order fields are those of linearized Einstein–Maxwell theory, the second-order fields are those required by the field equations to first nonlinear order, and so forth.

Specifically, we take our combined linear gravitational and electromagnetic field Ψ to be composed of a mass monopole field m, a retarded gravitational quadrupole radiation field ψ, a Coulomb field e, and an electromagnetic radiation field ϕ consisting of a retarded dipole plus a retarded quadrupole. The radiating multipole fields are chosen to be pulses of radiation that turn on and off simultaneously. The field Ψ is taken to be axially symmetric.

We symbolize the field by writing

$$\Psi = m + e + \psi + \phi.$$

Formally, there are then ten possible second-order interactions:

$$e^2 + m^2 + \psi^2 + \phi^2 + me + m\psi + m\phi + e\psi + e\phi + \psi\phi.$$

Since the first-order gravitational and electromagnetic fields are not coupled (i.e., the first-order electromagnetic field does not contribute to the first-order metric and vice versa), a second-order electromagnetic field can arise only from the interaction of a first-order gravitational field and a first-order electromagnetic field. This is an obvious consequence of the fact that the interaction terms in Maxwell's equations are products of the electromagnetic field with quantities constructed only from the metric. A second-order gravitational field can arise only from the self-interaction of fields and not from the interactions of the gravitational field with the electromagnetic field. This follows obviously from the fact that the electromagnetic field occurs in the Bianchi identities only quadratically. Hence the electromagnetic parts of the interactions $e^2, m^2, \psi^2, \phi^2, m\psi$, and $e\phi$ do not occur, and the gravitational parts of the interactions $me, m\phi, e\psi$, and $\psi\phi$ do not occur. The gravitational part of m^2 and the electromagnetic part of me are not eliminated by the general nature of the nonlinearities of the Einstein–Maxwell equations, but direct calculation of these two interactions shows that they vanish. This is in agreement with the fact that they do not enter the Reissner–Nordström solution. The gravitational part of the e^2 interaction does not vanish. However, it is part of this well-known static solution.[4] The $m\psi$ and ψ^2 interactions have been analyzed completely in Ref. 1.

* Research supported by the U.S. Atomic Energy Commission, under Contract No. AT(11-1)-GEN-10, P.A.11.

1 The literature on the second sound in He II is too extensive to be cited here. For complete references, the reader is referred to I. M. Khalatnikov, *Introduction to the Theory of Superfludity* (Benjamin, New York, 1965), and J. Wilks, The *Properties of Liquid and Solid Helium* (Clarendon, Oxford, England, 1967).

2 See collected papers in D. Pines, Many-Body Problem (Benjamin, New York, 1962).

3 J. Gavoret and P. Nozieres, Ann Phys. (N.Y.) **28**, 349 (1969).

4 P. Hohenberg and P. C. Martin, Ann. Phys. (N.Y.) **34**, 291 (1965), C. E. Carroll, Phys Rev. A **2**, 497 (1970).

5 S. Ma, H. Gould and V. K. Wong, Phys. Rev. A **3**, 1453 (1971).

6 For example, see J. A. Krumhansl, Proc. Phys. Soc. (London) **85**, 921 (1965), R. A. Guyer and J. A. Krumhansl, Phys. Rev. A 1411 (1964). See also P. C. Kwok and P. C. Martin, Phys. Rev. **142**, 495 (1966).

7 L. J. Sham, Phys. Rev **156**, 494 (1967); **163**, 401 (1967).

8 For a general discussion along this line, see R. Dashen and S. Ma. J. Math Phys. **12**, 1449 (1971).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 10 OCTOBER 1971

Radiation Scattering in Einstein–Maxwell Theory

W. E. Couch*

Department of Physics, University of Pittsburgh, Pittsburgh, Pennsylvania 15213

and

W. H. Hallidy

Department of Physics, California State College, California, Pennsylvania 15419

(Received 13 May 1971)

A linearized Einstein–Maxwell field composed of a mass monopole, a radiating mass quadrupole, a Coulomb field, and radiating electromagnetic dipole and quadrupole fields is examined to second order. If the first-order radiation is a pulse of finite duration, then the only second-order radiation present after the pulse has turned off is due to interactions involving the mass monopole. All other interactions contributing to the second order produce no such scattered radiation. At third order, the dipole–dipole-mass interaction is found to give rise to scattered radiation.

1. INTRODUCTION

Nonlinear effects that show some interesting aspects of the interaction of gravitational radiation fields in empty space–time with themselves and with mass monopoles have recently been discovered.[1-3] These previous investigations made use of a perturbative approximation scheme to calculate the nonlinear effects. In this paper we apply the same perturbative scheme to the case where an electromagnetic field is present in space-time and investigate certain interactions of electromagnetic fields with gravitational fields and with themselves.

We assume that all quantities occurring in our formulation of Einstein–Maxwell theory are expandable in a perturbation series in which the zeroth order represents flat space–time with no electromagnetic field present, the first-order fields are those of linearized Einstein–Maxwell theory, the second-order fields are those required by the field equations to first nonlinear order, and so forth.

Specifically, we take our combined linear gravitational and electromagnetic field Ψ to be composed of a mass monopole field m, a retarded gravitational quadrupole radiation field ψ, a Coulomb field e, and an electromagnetic radiation field ϕ consisting of a retarded dipole plus a retarded quadrupole. The radiating multipole fields are chosen to be pulses of radiation that turn on and off simultaneously. The field Ψ is taken to be axially symmetric.

We symbolize the field by writing

$$\Psi = m + e + \psi + \phi.$$

Formally, there are then ten possible second-order interactions:

$$e^2 + m^2 + \psi^2 + \phi^2 + me + m\psi + m\phi + e\psi + e\phi + \psi\phi.$$

Since the first-order gravitational and electromagnetic fields are not coupled (i.e., the first-order electromagnetic field does not contribute to the first-order metric and vice versa), a second-order electromagnetic field can arise only from the interaction of a first-order gravitational field and a first-order electromagnetic field. This is an obvious consequence of the fact that the interaction terms in Maxwell's equations are products of the electromagnetic field with quantities constructed only from the metric. A second-order gravitational field can arise only from the self-interaction of fields and not from the interactions of the gravitational field with the electromagnetic field. This follows obviously from the fact that the electromagnetic field occurs in the Bianchi identities only quadratically. Hence the electromagnetic parts of the interactions $e^2, m^2, \psi^2, \phi^2, m\psi$, and $e\phi$ do not occur, and the gravitational parts of the interactions $me, m\phi, e\psi$, and $\psi\phi$ do not occur. The gravitational part of m^2 and the electromagnetic part of me are not eliminated by the general nature of the nonlinearities of the Einstein–Maxwell equations, but direct calculation of these two interactions shows that they vanish. This is in agreement with the fact that they do not enter the Reissner–Nordström solution. The gravitational part of the e^2 interaction does not vanish. However, it is part of this well-known static solution.[4] The $m\psi$ and ψ^2 interactions have been analyzed completely in Ref. 1.

We therefore restrict our attention to the second-order field that arises from the interactions ϕ^2, $m\phi, e\psi, e\phi$, and $\psi\phi$. The boundary condition that the second-order field have no gravitational or electromagnetic radiation that comes *in* from infinity is imposed. In fact we require that the second-order field vanish for all times prior to the emission of the first-order radiation, and further require that after emission, it has no radiating multipole singularities at the origin. With these conditions we are able to calculate the second-order gravitational and electromagnetic fields that arise, and we do this separately for each interaction. For the second-order field of each interaction we ascertain whether there is a gravitational or electromagnetic radiation field present in regions of space–time after the first-order radiation pulses have turned off. The boundary conditions ensure that if any radiation is present after the emission of the first-order radiation, it has been generated by the interaction. Whether such gravitational or electromagnetic radiation fields are present is the major invariant property of interest concerning the second-order field.

2. EINSTEIN–MAXWELL THEORY

We use a null tetrad of basis vectors[5,6] $(l_\mu, n_\mu, m_\mu, \bar{m}_\mu)$ satisfying

$$l_\mu n^\mu = -m_\mu \bar{m}^\mu = 1, \qquad l_\mu m^\mu = n_\mu m^\mu = 0. \tag{2.1}$$

The vector l_μ is chosen as the gradient of null hypersurfaces labeled by the coordinate $\chi^0 = u$ so $l_\mu = u_{,\mu}$. The vectors n_μ, m_μ, and $\bar{m}_\mu$ are parallel-propagated along the geodesics to which l^μ is tangent. The coordinate x^1 is taken to be the affine parameter r along l^μ; and the coordinates x^i label different null geodesics in each null hypersurface. Our tetrad then has the form:

$$l^\mu = \delta^\mu_1,$$
$$n^\mu = \delta^\mu_0 + U\delta^\mu_1 + X^i\delta^\mu_i, \tag{2.2}$$
$$m^\mu = \omega\delta^\mu_1 + \xi^i\delta^\mu_i.$$

The gravitational field is given by the tetrad components of the Weyl tensor:

$$\psi_0 \equiv -C_{\mu\nu\rho\sigma}l^\mu m^\nu l^\rho m^\sigma,$$
$$\psi_1 \equiv -C_{\mu\nu\rho\sigma}l^\mu n^\nu l^\rho m^\sigma, \tag{2.3}$$
$$\psi_2 \equiv -C_{\mu\nu\rho\sigma}\bar{m}^\mu n^\nu l^\rho m^\sigma,$$
$$\psi_3 \equiv -C_{\mu\nu\rho\sigma}\bar{m}^\mu n^\nu l^\rho n^\sigma,$$
$$\psi_4 \equiv -C_{\mu\nu\rho\sigma}\bar{m}^\mu n^\nu \bar{m}^\rho n^\sigma,$$

and the electromagnetic field is given by the tetrad components of the electromagnetic field tensor:

$$\phi_0 \equiv F_{\mu\nu}l^\mu m^\nu,$$
$$\phi_1 \equiv \tfrac{1}{2}F_{\mu\nu}(l^\mu n^\nu + \bar{m}^\mu m^\nu), \tag{2.4}$$
$$\phi_2 \equiv F_{\mu\nu}\bar{m}^\mu n^\nu.$$

The ψ_A satisfy the Bianchi identities:

$$\begin{aligned}\binom{D}{\delta}\psi_A - \binom{\bar{\delta}}{\Delta}\psi_{A-1} &= (1-A)\binom{\lambda}{\nu}\psi_{A-2} - (3-A)2\binom{\alpha}{\gamma}\psi_{A-1} + (5-A)\binom{\rho}{\tau}\psi_A + \binom{0}{1}[A\mu\psi_{A-1} + 2(2-A)\beta\psi_A \\ &- (4-A)\sigma\psi_{A+1}] - \binom{\bar{\phi}_0}{\bar{\phi}_1}((\delta_{A1}+\delta_{A3})\{[\delta - 2(2-A)\beta]\phi_{A-1} + [(3-A)\sigma + (1-A)\mu]\phi_1\} + (\delta_{A2}+\delta_{A4}) \\ &\times\{[\Delta - 2(3-A)\gamma]\phi_{A-2} + [(4-A)\tau + (2-A)\nu]\phi_1\}) + \binom{\bar{\phi}_1}{\bar{\phi}_2}(\delta_{A1}+\delta_{A3})D\phi_{A-1} + (\delta_{A2}+\delta_{A4}) \\ &\times\{[\bar{\delta} - 2(3-A)\alpha]\phi_{A-2} + [(4-A)\rho + (2-A)\lambda]\phi_1\}),\end{aligned} \tag{2.5}$$

with $A = 1,2,3,4$, where δ_{AB} is the Kronecker symbol,

$$\rho \equiv l_{\mu;\nu}m^\mu\bar{m}^\nu,$$
$$\alpha \equiv \tfrac{1}{2}(l_{\mu;\nu}n^\mu\bar{m}^\nu - m_{\mu;\nu}\bar{m}^\mu\bar{m}^\nu),$$
$$\beta \equiv \tfrac{1}{2}(l_{\mu;\nu}n^\mu m^\nu - m_{\mu;\nu}\bar{m}^\mu m^\nu),$$
$$\tau \equiv \bar{\alpha} + \beta, \tag{2.6}$$
$$\nu \equiv -n_{\mu;\nu}\bar{m}^\mu n^\nu,$$
$$\gamma \equiv \tfrac{1}{2}(l_{\mu;\nu}n^\mu n^\nu - m_{\mu;\nu}\bar{m}^\mu n^\nu),$$
$$\mu \equiv -n_{\mu;\nu}\bar{m}^\mu m^\nu,$$
$$\sigma \equiv l_{\mu;\nu}m^\mu m^\nu,$$
$$\lambda \equiv -n_{\mu;\nu}\bar{m}^\mu\bar{m}^\nu,$$

and the differential operators $(D, \Delta, \delta, \bar{\delta})$ are the intrinsic derivatives with respect to the tetrad. Maxwell's equations in terms of the ϕ_A are

$$\begin{aligned}\binom{D}{\delta}\phi_A - \binom{\bar{\delta}}{\Delta}\phi_{A-1} &= (1-A)\binom{\lambda}{\nu}\phi_{A-2} - 2(2-A) \\ &\times\binom{\alpha}{\gamma}\phi_{A-1} + (3-A)\binom{\rho}{\tau}\phi_A + \binom{0}{1} \\ &\times[A\mu\phi_{A-1} + 2(1-A)\beta\phi_A - (2-A)\sigma\phi_{A+1}],\end{aligned} \tag{2.7}$$

with $A = 1, 2$.

The perturbation scheme applied to these equations gives the differential equations that govern the fields at each perturbative order.

The zeroth-order field is given by $\psi_A = \phi_A = 0$ and we have

$$\rho = -1/r, \quad U = r\mu = -\tfrac{1}{2},$$

$$\alpha = -\beta = -(\sqrt{2}/4r)\cot\theta,$$

$$\xi^j = (\sqrt{2}/2r)(1, i/\sin\theta),$$

(u, r, θ, ϕ) are null spherical polar coordinates. All other spin coefficients and tetrad vector components vanish in flat space.

The first-order field of our problem is given by

$$\psi_A = \frac{m}{r^3}\,{}_0Y_{00}\,\delta^2_A + (-\sqrt{2})^{A-2}\left(\frac{(4-A)!}{A!}\right)^{1/2} \times\sqrt{6}\;{}_{2-A}Y_{20}\left(\frac{\partial}{\partial u} - \frac{1}{2}\frac{\partial}{\partial r}\right)^A \frac{Q}{r^{5-A}}, \tag{2.8a}$$

with $A = 0, 1, 2, 3, 4$;

$$\phi_A = \frac{e}{r^2}\,{}_0Y_{00}\delta^1_A + (-1)^A\,{}_{1-A}Y_{10}\left(\frac{\partial}{\partial u} - \frac{1}{2}\frac{\partial}{\partial r}\right)^A \times \frac{p}{r^{3-A}} + (-1)^A(\sqrt{2})_{A-1}\left(\frac{(3-A)!}{(1+A)!}\right)^{1/2} \times\sqrt{3}\,{}_{1-A}Y_{20}\,r\left(\frac{\partial}{\partial u} - \frac{1}{2}\frac{\partial}{\partial r}\right)^{A+1}\frac{q}{r^{4-A}}, \tag{2.8b}$$

with $A = 0, 1, 2$ where m is the mass, e the electric charge, Q the mass quadrupole moment, p the electric dipole moment, and q the electric quadrupole moment; the quantities Q, p, and q vanish outside the interval $u_2 > u > u_1$. The functions ${}_sY_{lo}$ are spin-weighted spherical harmonics of spin weight s.[7]

The first-order spin coefficients and metric variables are

$$\rho = 0,$$

$$\sigma = \left(\frac{\ddot{\bar{Q}}}{r^2} + \frac{3Q}{r^4}\right){}_2Y_{20},$$

$$\alpha = \tfrac{1}{4}\sqrt{2}\left(\frac{\ddot{\bar{Q}}}{r^2} + \frac{\bar{Q}}{r^4}\right)\cot\theta\,{}_2Y_{20},$$

$$\tau = -\frac{3}{2}\sqrt{2}\left(\frac{\dot{Q}}{r^3} + \frac{4}{3}\frac{Q}{r^4}\right){}_1Y_{20},$$

$$\beta = \tau - \bar{\alpha},$$

$$\mu = \frac{-m}{r^2}\,{}_0Y_{00} + \sqrt{6}\left(\frac{\ddot{Q}}{r^2} + \frac{3\dot{Q}}{2r^3} + \frac{Q}{r^4}\right){}_0Y_{20},$$

$$\gamma = -\frac{m}{2r^2}\,{}_0Y_{00} + \sqrt{6}\left(\frac{\ddot{Q}}{2r^2} + \frac{\dot{Q}}{r^3} + \frac{3Q}{4r^4}\right){}_0Y_{20} - \frac{1}{4}\left(\frac{\dot{\bar{Q}}}{r^3} + \frac{\bar{Q}}{r^4}\right)\cot\theta\,{}_{-1}Y_{20} - \frac{1}{4}\left(\frac{\dot{Q}}{r^3} + \frac{Q}{r^4}\right)\cot\theta\,{}_1Y_{20}, \tag{2.9}$$

$$\lambda = \frac{1}{2}\left(\frac{2\dddot{\bar{Q}}}{r} + \frac{\ddot{\bar{Q}}}{r^2} + \frac{\bar{Q}}{r^4}\right){}_{-2}Y_{20},$$

$$\nu = -\sqrt{2}\left(\frac{\dddot{\bar{Q}}}{r} + \frac{3\ddot{\bar{Q}}}{2r^2} + \frac{6\dot{Q} - \dot{\bar{Q}}}{4r^3} + \frac{3Q - \bar{Q}}{4r^4}\right){}_{-1}Y_{20},$$

$$U = \frac{-m}{r}\,{}_0Y_{00} + \frac{1}{2}\sqrt{6}\left(\frac{\ddot{Q} + \ddot{\bar{Q}}}{r} + \frac{\dot{Q} + \dot{\bar{Q}}}{r^2} + \frac{Q + \bar{Q}}{2r^3}\right){}_0Y_{20},$$

$$\omega = \sqrt{2}\left(\frac{\ddot{\bar{Q}}}{r} - \frac{3\dot{Q}}{2r^2} - \frac{Q}{r^3}\right){}_1Y_{20},$$

$$X^j = \tfrac{1}{2}\left(1, \frac{-i}{\sin\theta}\right)\left(\frac{\dot{Q}}{r^3} + \frac{Q}{r^4}\right){}_1Y_{20} - \tfrac{1}{2}\left(1, \frac{i}{\sin\theta}\right)\left(\frac{\dot{\bar{Q}}}{r^3} + \frac{\bar{Q}}{r^4}\right){}_{-1}Y_{20},$$

$$\xi^j = -\frac{\sqrt{2}}{2}\left(1, \frac{-i}{\sin\theta}\right)\left(\frac{\ddot{\bar{Q}}}{r^2} + \frac{Q}{r^4}\right){}_2Y_{20},$$

where the dot denotes $\partial/\partial u$.

These quantities can be obtained from the metric and the field equations, $G_{\mu\nu} = -8\pi k\,T_{\mu\nu} = -8\pi k(F^{\rho}_{\mu}F_{\nu\rho} - \frac{1}{4}g_{\mu\nu}F_{\rho\sigma}F^{\rho\sigma})$, when both the metric and the field equations are written in their tetrad form.[1,5]

The first-order field satisfies the equations

$$\begin{pmatrix} \frac{\partial}{\partial r} + \frac{5-A}{r} & \frac{\sqrt{2}}{2r}\,\bar{\eth} \\ -\frac{\sqrt{2}}{2r}\,\eth & -\frac{\partial}{\partial u} + \frac{1}{2}\frac{\partial}{\partial r} + \frac{A}{2r} \end{pmatrix}\begin{pmatrix}\psi_A \\ \psi_{A-1}\end{pmatrix} = 0, \tag{2.10}$$

with $A = 1, 2, 3, 4$;

$$\begin{pmatrix} \frac{\partial}{\partial r} + \frac{3-A}{r} & \frac{\sqrt{2}}{2r}\,\bar{\eth} \\ -\frac{\sqrt{2}}{2r}\,\eth & -\frac{\partial}{\partial u} + \frac{1}{2}\frac{\partial}{\partial r} + \frac{A}{2r} \end{pmatrix}\begin{pmatrix}\phi_A \\ \phi_{A-1}\end{pmatrix} = 0, \tag{2.11}$$

with $A = 1, 2$. The operator $\eth$ is a raising operator for spin-weighted quantities and is defined by[7]

$$\eth\eta = -(\sin\theta)^s\left(\frac{\partial}{\partial\theta} + \frac{i}{\sin\theta}\frac{\partial}{\partial\phi}\right)[(\sin\theta)^{-s}\eta], \tag{2.12}$$

and $\bar{\eth}$ is a lowering operator defined by

$$\bar{\eth}\eta = -(\sin\theta)^{-s}\left(\frac{\partial}{\partial\theta} - \frac{i}{\sin\theta}\frac{\partial}{\partial\phi}\right)[(\sin\theta)^{s}\eta], \tag{2.13}$$

where η is a quantity of spin weight s. At each perturbative order the ψ_A have spin weight $2 - A$ and the ϕ_A have spin weight $1 - A$.

The second-order field satisfies a modified form of Eqs. (2.10) and (2.11) that has instead of zero for the right-hand sides driving terms composed of products of known first-order quantities. These driving terms for Eqs. (2.10) and (2.11) are, respectively, given by

$$
\begin{aligned}
&(1-A)\begin{pmatrix}\lambda\\ \nu\end{pmatrix}\psi_{A-2}+\left[\begin{pmatrix}\bar{\delta}\\ \Delta\end{pmatrix}-2(3-A)\begin{pmatrix}\alpha\\ \gamma\end{pmatrix}\right]\psi_{A-1}\\
&\quad+\begin{pmatrix}0\\ 1\end{pmatrix}\{A\mu\psi_{A-1}-[\delta-2(2-A)\beta-(5-A)\tau]\psi_A\\
&\quad-(4-A)\sigma\psi_{A+1}\}\\
&\quad+\begin{pmatrix}\bar{\phi}_0\\ \bar{\phi}_1\end{pmatrix}\left[(\delta_{A1}+\delta_{A3})\left(\frac{\sqrt{2}}{2r}\,\delta\phi_{A-1}+\frac{1-A}{2r}\,\phi_1\right)\right.\\
&\quad\left.+(\delta_{A2}+\delta_{A4})\left(\frac{\partial}{\partial u}-\frac{1}{2}\,\frac{\partial}{\partial r}\right)\phi_{A-2}\right]\\
&\quad+\begin{pmatrix}\bar{\phi}_1\\ \bar{\phi}_2\end{pmatrix}\left[(\delta_{A1}+\delta_{A3})\,\frac{\partial}{\partial r}\,\phi_{A-1}-(\delta_{A2}+\delta_{A4})\right.\\
&\quad\left.+\left(\frac{\sqrt{2}}{2r}\,\bar{\delta}\phi_{A-2}+\frac{4-A}{r}\,\phi_1\right)\right]
\end{aligned}
$$

and

$$
\begin{aligned}
&(1-A)\begin{pmatrix}\lambda\\ \nu\end{pmatrix}\phi_{A-2}+\left[\begin{pmatrix}\bar{\delta}\\ \Delta\end{pmatrix}-2(2-A)\begin{pmatrix}\alpha\\ \gamma\end{pmatrix}\right]\phi_{A-1}\\
&\quad+\begin{pmatrix}0\\ 1\end{pmatrix}\{A\mu\phi_{A-1}-[\delta-2(1-A)\beta-(3-A)\tau]\phi_A\\
&\quad-(2-A)\sigma\phi_{A+1}\}.
\end{aligned}
$$

The second-order field is obtained by solving the driven form of Eqs. (2.10) and (2.11) in conjunction with the boundary conditions.

3. THE SECOND-ORDER FIELD

We find the second-order Einstein–Maxwell field caused by each interaction by solving for ψ_0 and ϕ_0 in the driven form of Eqs. (2.10) and (2.11) with $A=1$. The other ϕ_A and ψ_A can then be obtained from the remaining equations by straightforward calculation. We will only display ϕ_0 and ψ_0.

A. The ϕ^2 Interaction

There is no electromagnetic field arising in this case. The solution for the gravitational field is given by

$$
\begin{aligned}
\psi_0 &= r^{-5}[(W+V)_2Y_{20}+(\dot{X}+\dot{O})_2Y_{30}\\
&\quad+(\ddot{Y}+\dot{S})_2Y_{40}]+r^{-6}[3(X+O)_2Y_{30}\\
&\quad+7(\dot{Y}+\dot{S})_2Y_{40}]+r^{-7}14[Y+S]_2Y_{40}+Z,
\end{aligned}
\tag{3.1}
$$

where V, O, and S are, respectively, the second-order quadrupole, octupole, and sixteen-pole moments due to second-order outgoing gravitational news. This news is present only in the pulse region[8] and is unrestricted there except for the following relations which must be satisfied at $u=u_2$:

$$
\frac{d^nV}{du^n}=-\frac{d^nW}{du^n},\qquad n=0,1,2,
\tag{3.2a}
$$

$$
\frac{d^nO}{du^n}=-\frac{d^nX}{du^n},\qquad n=0,1,2,3,
\tag{3.2b}
$$

$$
\frac{d^nS}{du^n}=-\frac{d^nY}{du^n},\qquad n=0,1,2,3,4.
\tag{3.2c}
$$

The terms W, X, Y and Z are defined in the Appendix.

We see that $\psi_0\equiv 0$ outside of the pulse region, so that the radiative part of this field vanishes in space-time regions outside of the first-order pulse. If the conditions at $u=u_2$ expressed by Eqs. (3.2) were not required, the second-order field would still be nonradiative but ψ_0 would not vanish for $u>u_2$.

B. The $m\phi$ Interaction

There is no gravitational field arising in this case. The solution for the electromagnetic field is given by

$$
\phi_0=\frac{m}{\sqrt{4\pi}}\left\{\left[-D^2\left(\frac{I}{r}\right)+\frac{K}{r^3}\right]{}_1Y_{10}+\left[rD^3\left(\frac{J}{r^2}\right)+\frac{L}{r^4}+\frac{\dot{L}}{2r^3}\right]{}_1Y_{20}\right\},
\tag{3.3}
$$

where

$$
\begin{aligned}
I &\equiv \int_{u_1}^{u}\frac{p(u')du'}{(u+2r-u')^2},\qquad u\geq u_1,\\
K(u) &\equiv 2\int_{u_1}^{u}\frac{p(u')du'}{(u-u')^2},\qquad u\geq u_1,\\
J &\equiv \frac{3q}{20r}+\tfrac{1}{2}\int_{u_1}^{u}\frac{q(u')du'}{(u+2r-u')^2},\qquad u\geq u_1,\\
L(u) &\equiv 12\int_{u_1}^{u}\frac{q(u')du'}{(u-u')^2},\qquad u\geq u_1.
\end{aligned}
$$

The outgoing electromagnetic news for this field is given by

$$
\phi_2^0=\frac{m}{\sqrt{4\pi}}\left(\ddot{K}_{-1}Y_{10}+\tfrac{1}{6}\dddot{L}_{-1}Y_{20}\right)
\tag{3.4}
$$

in the region after the pulse. This shows the scattering of the electromagnetic radiation off the Schwarzschild mass.

C. The $e\psi$ Interaction

There is no gravitational field arising in this case. The solution for the electromagnetic field is given by

$$
\phi_0=\left(\frac{2}{\pi}\right)^{1/2}e\left(-\frac{\dot{\bar{Q}}}{r^4}+\frac{3\bar{Q}}{r^5}\right){}_1Y_{20},
\tag{3.5}
$$

which vanishes outside of the pulse region. This field is therefore nonradiative in the region after the pulse.

D. The $e\phi$ Interaction

There is no electromagnetic field arising in this

case. The solution for the gravitational field is given by

$$\psi_0 = \left(\frac{2}{\pi}\right)^{1/2} e\left(\frac{\dot{\bar{q}}}{r^5} - \frac{6q}{r^6}\right) {}_2Y_{20}, \tag{3.6}$$

which vanishes outside of the pulse region. This field is therefore nonradiative in the region after the pulse.

E. The $\phi\psi$ Interaction

There is no gravitational field arising in this case. The solution for the electromagnetic field has been calculated for the electromagnetic dipole-gravitational quadrupole interaction. The field is found to vanish outside of the pulse, and is therefore nonradiative in that region. This is also found to be true when the first-order field is not restricted to be axially symmetric.

It has previously been found[1] that there is scattered radiation arising from the $m\psi$ interaction and no scattered radiation in the case of the ψ^2 interaction. These results and the interactions calculated here show that for the complete second-order field arising from ψ, the only scattered radiation present after the pulse is due to the scattering of the first-order gravitational and electromagnetic radiation off the mass monopole; no radiation is scattered if the mass monopole vanishes. In particular, the interaction of these first-order electromagnetic radiation fields with themselves or with a radiating gravitational quadrupole produces no scattered radiation after the pulse.

4. THE THIRD-ORDER FIELD

We will discuss only the mass-dipole–dipole interaction. There is no third-order electromagnetic field arising in this case. The solution for the gravitational field is given by

$$\psi_0 = \left[\frac{M+N}{r^5} + D^4\left(\frac{R}{r}\right) - T\right] {}_2Y_{20}, \tag{4.1}$$

where

$$N(u) \equiv \frac{m}{2\sqrt{4\pi}}\,\dddot{\bar{V}}(u), \qquad u \geq u_1,$$

$$R \equiv \int_{u_1}^{u} \frac{F(u')du'}{(u + 2r - u')^2}, \qquad u \geq u_1,$$

$$M(u) + N(u) \equiv -24 \int_{u_1}^{u} \frac{F(u')du'}{(u-u')^2}, \qquad u \geq u_1,$$

$$F(u) \equiv -\frac{m}{2\sqrt{4\pi}}\left[W + V - \frac{2}{7}\left(\frac{3}{10\pi}\right)^{1/2}(\bar{p}\dot{p} - 6\dot{\bar{p}}\dot{p})\right],$$

$$T \equiv \frac{\dot{\bar{p}}}{r^3} D\left(rD^2I - 4DI + \frac{6I}{r}\right) + \frac{\bar{p}}{2r^2} D^2\left(\frac{DI}{r}\right) + 15\frac{p\bar{p}}{7r^7}$$

and now the terms involving q are not present in W and V.

The third-order outgoing gravitational news for this field is given by

$$\dot{\sigma}^0 = -\tfrac{1}{6}\dddot{\bar{M}}_2Y_{20} \tag{4.2}$$

in the region after the pulse. We see that we will have scattering in the region after the pulse unless the second-order outgoing gravitational news $-\frac{1}{6}\dddot{\bar{V}}_2Y_{20}$ is chosen such that $F(u) = 0$ everywhere. In that case, in the region after the pulse we have $\psi_0 = \dddot{M} = 0$ and no radiation is present. We interpret this to mean that in general, third-order gravitational radiation is produced by the interaction of first-order electromagnetic radiation with the mass. In the special case that $F \equiv 0$ the produced radiation is canceled by radiative scattering of second-order gravitational radiation off the mass.

ACKNOWLEDGMENTS

The authors wish to thank Professor A. I. Janis and Professor E. T. Newman for helpful discussions.

APPENDIX

The quantities W, X, and Y are defined as follows:

$$W \equiv -\left(\frac{3}{10\pi}\right)^{1/2}\left[\int^u\int\int \ddot{p}\ddot{\bar{p}} + 2\int^u\int \dot{p}\dddot{\bar{p}} + 3\int^u p\ddot{\bar{p}}\right] + \tfrac{1}{14}\left(\frac{30}{\pi}\right)^{1/2}\left[\int^u\int\int \ddot{q}\dddot{\bar{q}} - 2\int^u\int \ddot{q}\ddot{\bar{q}} - 9\int^u \dot{q}\dddot{\bar{q}}\right],$$

$$X \equiv -\left(\frac{6}{7\pi}\right)^{1/2}\left[\tfrac{3}{2}\int^u\int\int\int(\ddot{\bar{p}}\dddot{q} + \ddot{p}\ddot{\bar{q}}) + \int^u\int\int(3\dddot{\bar{p}}\ddot{q} + 2\dot{p}\dddot{\bar{q}}) + \tfrac{3}{2}\int^u\int(3\ddot{\bar{p}}\dot{q} + p\dddot{\bar{q}}) + 2\int^u(2\ddot{\bar{p}}q + \dot{p}\dot{q})\right],$$

$$Y \equiv -\tfrac{3}{7}\left(\frac{10}{\pi}\right)^{1/2}\left[2\int^u\int\int\int\int \ddot{q}\dddot{\bar{q}} + 3\int^u\int\int\int \dddot{q}\dddot{\bar{q}} + 3\int^u\int\int \dddot{q}\dddot{\bar{q}} - \tfrac{2}{7}\int^u\int(3\ddot{q}\dot{q} - 4\dddot{q}q) - \tfrac{3}{14}\int^u(3\dot{q}\dddot{\bar{q}} - \ddot{\bar{q}}\dddot{q})\right],$$

where each integral in these multiple integrals has u_1 for its lower limit and a dummy variable of integration for its upper limit except when it is the final integral. The quantity Z is defined by

$$Z = -\tfrac{6}{7}\left(\frac{30}{\pi}\right)^{1/2}\left(\frac{2\ddot{\bar{q}}q}{r^6} + \frac{3\dot{\bar{q}}q}{r^7} + \frac{2q\bar{q}}{r^8}\right) {}_2Y_{20} + \left(\frac{6}{7\pi}\right)^{1/2}\left(\frac{4\ddot{\bar{p}}q + 2\dot{\bar{p}}\dot{q}}{r^5} + \frac{3\bar{p}q}{r^7}\right) {}_2Y_{30}$$
$$-\tfrac{3}{98}\left(\frac{10}{\pi}\right)^{1/2}\left(\frac{9\dot{q}\ddot{q} + 18\ddot{\bar{q}}\dot{q} - 19\dddot{\bar{q}}q}{r^5} + \frac{21(3\dot{\bar{q}}\dot{q} - \ddot{\bar{q}}q)}{r^6} - \frac{84q\bar{q}}{r^8}\right) {}_2Y_{40}.$$

* Research supported in part by the National Science Foundation.

1 W. E. Couch, R. J. Torrence, A. I. Janis, and E. T. Newman, J. Math. Phys. 9, 484 (1968).
2 W. E. Couch, W. M. Kinnersley, and R. J. Torrence, Phys. Letters 31A, 576 (1970).
3 W. E. Couch and R. J. Torrence, J. Math. Phys. 11, 2096 (1970).
4 E. T. Newman *et al.*, J. Math. Phys. 6, 918 (1965).
5 E. T. Newman and R. Penrose, J. Math. Phys. 3, 566 (1962).
6 Greek indices range and sum from 0 to 3, lower case Latin indices range and sum from 2 to 3; covariant differentiation is denoted by a semicolon.
7 J. N. Goldberg *et al.*, J. Math. Phys. 8, 2155 (1967).
8 This radiation cannot be interpreted as scattered radiation because in the space-time region where the pulse is present we have not required the appropriate boundary conditions that make the scattering problem well defined. In the case of the ψ^2 interaction (see Ref. 3) it has been found that there is no scattered radiation at ∞ in the pulse region; this result is to be expected for the ϕ^2 interaction also.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 10 OCTOBER 1971

The Statistical Mechanical Derivation of the Stress Tensor and Heat Flux for a System of Spherical Molecules

Bruce N. Miller*
Physics Department, Rice University, Houston, Texas 77001
(Received 8 March 1971)

The equations of hydrodynamics are derived from the principles of classical statistical mechanics for a single component system of spherical molecules. Exact expressions for the stress tensor and heat flux are obtained without resorting to any methods of approximation. The derivations are carried out for a system of molecules which interact through a continuous pair potential, and for a system consisting solely of rigid spheres. The "long-wavelength" expansion employed by Irving and Kirkwood [J. Chem Phys. 18, 817 (1950)] in their expressions for the stress tensor and heat flux is examined. It is demonstrated that this expansion converges only if the intermolecular potential goes to zero faster than any positive power of $(1/r)$ in the limit of large r (r is the internuclear distance) and hence diverges for any realistic intermolecular potential. An elementary example is considered to demonstrate the effect of finite wavelength on the stress tensor.

1. INTRODUCTION

The derivation of the hydrodynamic equations, the general conservation laws of macroscopic physics, from the principles of molecular mechanics has always been a central problem in statistical mechanics. One of the earliest approaches to this problem starts with the Boltzman transport equation[1] which arises in the study of dilute gases. Although this approach yielded the correct kinetic contributions to the equations of hydrodynamics, the role played by intermolecular forces was not generally understood until about twenty years ago. At that time Irving and Kirkwood developed their general distribution function formalism[2] to obtain explicit expressions for the stress tensor and heat flux in terms of microscopic quantities. Since then their results have been widely used as the starting point of more detailed investigations.[3]

In order to obtain expressions for the stress tensor and heat flux, Irving and Kirkwood arbitrarily expanded a pair distribution function occurring in the general conservation expressions in a Taylor series. The underlying assumption for performing this expansion is that the pair distribution function is a slowly varying function of the center of mass of the two molecules over distances comparable to the "range" of the intermolecular forces. Similar expansions, appearing frequently in the literature,[4-7] may be characterized as "long-wavelength" expansions. Irving and Kirkwood, as well as Green, Mori, Zwanzig, and Frölich, assume that for most situations the convergence of the expansion is sufficiently rapid to justify disregarding all but the first nonvanishing term.

Unfortunately, as shown in Sec. 14 of this paper, the long-wavelength expansions for the stress tensor and heat flux do not converge unless the intermolecular pair potential vanishes faster than any positive power of $1/r$ in the limit of large r(r is the internuclear distance). This is a requirement no realistic potential can meet and poses the problem of obtaining a satisfactory derivation of the hydrodynamic equations. It is our intention to supply solutions for these problems under the limitations of classical statistical mechanics.

We rigorously derive the equations of hydrodynamics for a single-component system of spherical molecules and obtain exact expressions for the stress tensor and heat flux. We consider not only molecules exhibiting continuous, differentiable potentials, but also a system of rigid spheres. The effects of finite wavelength on the intermolecular contribution to the stress tensor, as opposed to the kinetic contribution, are then considered in an elementary example. Any external forces on the systems considered are assumed to be conservative, and functions of position only.

Although the results for a system with a continuous intermolecular potential were obtained by Choh[8] from the BBGKY hierarchy,[9] he immediately employed the long-wavelength expansion of Irving and Kirkwood. It should be noted that the derivation presented here has two distinct advantages: It gives the physical origin of all terms appearing in the transport equations and is independent of the equations governing the evolution in time of the required distribution functions.

* Research supported in part by the National Science Foundation.

1 W. E. Couch, R. J. Torrence, A. I. Janis, and E. T. Newman, J. Math. Phys. 9, 484 (1968).
2 W. E. Couch, W. M. Kinnersley, and R. J. Torrence, Phys. Letters 31A, 576 (1970).
3 W. E. Couch and R. J. Torrence, J. Math. Phys. 11, 2096 (1970).
4 E. T. Newman *et al.*, J. Math. Phys. 6, 918 (1965).
5 E. T. Newman and R. Penrose, J. Math. Phys. 3, 566 (1962).
6 Greek indices range and sum from 0 to 3, lower case Latin indices range and sum from 2 to 3; covariant differentiation is denoted by a semicolon.
7 J. N. Goldberg *et al.*, J. Math. Phys. 8, 2155 (1967).
8 This radiation cannot be interpreted as scattered radiation because in the space-time region where the pulse is present we have not required the appropriate boundary conditions that make the scattering problem well defined. In the case of the ψ^2 interaction (see Ref. 3) it has been found that there is no scattered radiation at ∞ in the pulse region; this result is to be expected for the ϕ^2 interaction also.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 10 OCTOBER 1971

The Statistical Mechanical Derivation of the Stress Tensor and Heat Flux for a System of Spherical Molecules

Bruce N. Miller*
Physics Department, Rice University, Houston, Texas 77001
(Received 8 March 1971)

The equations of hydrodynamics are derived from the principles of classical statistical mechanics for a single component system of spherical molecules. Exact expressions for the stress tensor and heat flux are obtained without resorting to any methods of approximation. The derivations are carried out for a system of molecules which interact through a continuous pair potential, and for a system consisting solely of rigid spheres. The "long-wavelength" expansion employed by Irving and Kirkwood [J. Chem Phys. 18, 817 (1950)] in their expressions for the stress tensor and heat flux is examined. It is demonstrated that this expansion converges only if the intermolecular potential goes to zero faster than any positive power of $(1/r)$ in the limit of large r (r is the internuclear distance) and hence diverges for any realistic intermolecular potential. An elementary example is considered to demonstrate the effect of finite wavelength on the stress tensor.

1. INTRODUCTION

The derivation of the hydrodynamic equations, the general conservation laws of macroscopic physics, from the principles of molecular mechanics has always been a central problem in statistical mechanics. One of the earliest approaches to this problem starts with the Boltzman transport equation[1] which arises in the study of dilute gases. Although this approach yielded the correct kinetic contributions to the equations of hydrodynamics, the role played by intermolecular forces was not generally understood until about twenty years ago. At that time Irving and Kirkwood developed their general distribution function formalism[2] to obtain explicit expressions for the stress tensor and heat flux in terms of microscopic quantities. Since then their results have been widely used as the starting point of more detailed investigations.[3]

In order to obtain expressions for the stress tensor and heat flux, Irving and Kirkwood arbitrarily expanded a pair distribution function occurring in the general conservation expressions in a Taylor series. The underlying assumption for performing this expansion is that the pair distribution function is a slowly varying function of the center of mass of the two molecules over distances comparable to the "range" of the intermolecular forces. Similar expansions, appearing frequently in the literature,[4-7] may be characterized as "long-wavelength" expansions. Irving and Kirkwood, as well as Green, Mori, Zwanzig, and Frölich, assume that for most situations the convergence of the expansion is sufficiently rapid to justify disregarding all but the first nonvanishing term.

Unfortunately, as shown in Sec. 14 of this paper, the long-wavelength expansions for the stress tensor and heat flux do not converge unless the intermolecular pair potential vanishes faster than any positive power of $1/r$ in the limit of large r(r is the internuclear distance). This is a requirement no realistic potential can meet and poses the problem of obtaining a satisfactory derivation of the hydrodynamic equations. It is our intention to supply solutions for these problems under the limitations of classical statistical mechanics.

We rigorously derive the equations of hydrodynamics for a single-component system of spherical molecules and obtain exact expressions for the stress tensor and heat flux. We consider not only molecules exhibiting continuous, differentiable potentials, but also a system of rigid spheres. The effects of finite wavelength on the intermolecular contribution to the stress tensor, as opposed to the kinetic contribution, are then considered in an elementary example. Any external forces on the systems considered are assumed to be conservative, and functions of position only.

Although the results for a system with a continuous intermolecular potential were obtained by Choh[8] from the BBGKY hierarchy,[9] he immediately employed the long-wavelength expansion of Irving and Kirkwood. It should be noted that the derivation presented here has two distinct advantages: It gives the physical origin of all terms appearing in the transport equations and is independent of the equations governing the evolution in time of the required distribution functions.

Transport equations for a system of perfectly rigid spheres were obtained by Enskog[10] for moderately dense, as well as dilute systems. All of his results depend on approximation methods of limited applicability and are incorrect for highly dense systems.

2. DISTRIBUTION FUNCTIONS

Let $g^{(N)}(\mathbf{x}_1, \mathbf{p}_1; \cdots \mathbf{x}_N, \mathbf{p}_N; t)$ be the canonical distribution function, normalized to unity, of the entire system, where $\mathbf{x}_i$ is the position and $\mathbf{p}_i$ the momentum of the ith molecule, and N is the total number of molecules in the system. The two-particle distribution function in configuration and momentum space is defined by

$$g^{(2)}(\mathbf{x}_1, \mathbf{p}_1; \mathbf{x}_2, \mathbf{p}_2) = N(N-1) \times \int_{6(N-2)dim} d\mathbf{x}_3 \cdots d\mathbf{x}_N d\mathbf{p}_3 \cdots d\mathbf{p}_N g^{(N)} \quad (2.1)$$

and the one-particle distribution function in configuration and momentum space is given by

$$g^{(1)}(\mathbf{x}_1, \mathbf{p}_1) = N \int_{6(N-1)dim.} d\mathbf{x}_2 \cdots d\mathbf{x}_N d\mathbf{p}_2 \times \cdots d\mathbf{p}_N g^{(N)}. \quad (2.2)$$

The two- and one-particle distribution functions in configuration space are, respectively,

$$f^{(2)}(\mathbf{x}_1, \mathbf{x}_2) = \int d\mathbf{p}_1 \int d\mathbf{p}_2 g^{(2)}(\mathbf{x}_1, \mathbf{p}_1; \mathbf{x}_2, \mathbf{p}_2), \quad (2.3)$$

$$f^{(1)}(\mathbf{x}_1) = \int d\mathbf{p}_1 g^{(1)}(\mathbf{x}_1, \mathbf{p}_1). \quad (2.4)$$

For convenience we have suppressed the dependence of these functions on the time t, but it is assumed that there may be explicit dependence on t in any or all of them.

It will be convenient to introduce the vectors $\mathbf{w}, \mathbf{p}, \mathbf{r}$, and $\mathbf{x}$, where

$$\mathbf{w} = \mathbf{p}_2 - \mathbf{p}_1, \quad \mathbf{p} = \tfrac{1}{2}(\mathbf{p}_2 + \mathbf{p}_1), \quad \mathbf{r} = \mathbf{x}_2 - \mathbf{x}_1, \quad \mathbf{x} = \tfrac{1}{2}(\mathbf{x}_2 + \mathbf{x}_1); \quad (2.5)$$

and the functions Γ and γ defined by

$$\Gamma(\mathbf{w}, \mathbf{r}; \mathbf{p}, \mathbf{x}) = g^{(2)}(\mathbf{x}_1, \mathbf{p}_1; \mathbf{x}_2, \mathbf{p}_2), \quad \gamma(\mathbf{r}, \mathbf{x}) = f^{(2)}(\mathbf{x}_1, \mathbf{x}_2). \quad (2.6)$$

Because the molecules are identical, $g^{(2)}$ and $f^{(2)}$ are such that

$$g^{(2)}(\mathbf{x}_1, \mathbf{p}_1; \mathbf{x}_2, \mathbf{p}_2) = g^{(2)}(\mathbf{x}_2, \mathbf{p}_2; \mathbf{x}_1, \mathbf{p}_1), \quad f^{(2)}(\mathbf{x}_1, \mathbf{x}_2) = f^{(2)}(\mathbf{x}_2, \mathbf{x}_1); \quad (2.7)$$

so that

$$\Gamma(\mathbf{w}, \mathbf{r}; \mathbf{p}, \mathbf{x}) = \Gamma(-\mathbf{w}, -\mathbf{r}; \mathbf{p}, \mathbf{x}), \quad \gamma(\mathbf{r}, \mathbf{x}) = \gamma(-\mathbf{r}, \mathbf{x}). \quad (2.8)$$

We mention that in hydrodynamics the mass density ρ is frequently employed in place of $f^{(1)}$. If m is the mass of one molecule, then ρ is defined by

$$\rho(\mathbf{x}_1) = m f^{(1)}(\mathbf{x}_1).$$

Having defined all of the distribution functions we shall require, we may now proceed to derive the statistical transport equations.

3. TRANSPORT OF MASS

Throughout the following work we shall be concerned with the rate of change of some property (mass, momentum, etc.) of a simply connected region R in configuration space. The velocity of a boundary point of R is defined to be the average velocity $\mathbf{v}_0$ of a molecule at that point:

$$\mathbf{v}_0(\mathbf{x}_1) = \frac{1}{m} \int d\mathbf{p}_1 \frac{g^{(1)}(\mathbf{x}_1, \mathbf{p}_1)}{f^{(1)}(\mathbf{x}_1)} \mathbf{p}_1 = \frac{1}{m} \mathbf{p}_0(\mathbf{x}_1), \quad (3.1)$$

where $\mathbf{p}_0$ is the average momentum of a molecule at $\mathbf{x}_1$.

The mass M_R contained in the region R is given by

$$M_R = \int_R d\mathbf{x}_1 \rho(\mathbf{x}_1) \quad (3.2)$$

so that, from the definition of R, the rate of change of M_R is

$$\frac{dM_R}{dt} = \int_R d\mathbf{x}_1 \frac{\partial \rho}{\partial t} + \int_R da \hat{\mathbf{n}} \cdot \mathbf{v}_0 \rho, \quad (3.3)$$

where da is an element of the boundary surface of R and $\hat{\mathbf{n}}$ is the unit vector normal to the boundary surface and directed outward. Applying Gauss's theorem to the second term on the right-hand side (rhs) of Eq. (3.3)

$$\frac{dM_R}{dt} = \int_R d\mathbf{x}_1 \left(\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial \mathbf{x}_1} \cdot \rho \mathbf{v}_0 \right). \quad (3.4)$$

Now, the only way in which M_R may change is by the flow of molecules across the boundary surface of R. Consider a molecule with position $\mathbf{x}_1$ and momentum $\mathbf{p}_1$ such that $\mathbf{x}_1$ is in the region $(1/m)(\mathbf{p}_1 - \mathbf{p}_0) \cdot \hat{\mathbf{n}}\, da\, \delta t$ just inside the surface and $(\mathbf{p}_1 - \mathbf{p}_0) \cdot \hat{\mathbf{n}} > 0$ for some surface element da: Clearly, if δt is positive but very small, the molecule will leave R in the time δt. Thus, the number of molecules with $\mathbf{p}_1 \in d\mathbf{p}_1$ and $\mathbf{x}_1 \in (1/m)(\mathbf{p}_1 - \mathbf{p}_0) \cdot \hat{\mathbf{n}}\, da\, \delta t$ leaving R in the time δt is

$$(1/m)(\mathbf{p}_1 - \mathbf{p}_0) \cdot \hat{\mathbf{n}}\, g^{(1)}(\mathbf{x}_1, \mathbf{p}_1) H[(\mathbf{p}_1 - \mathbf{p}_0) \cdot \mathbf{n}] d\mathbf{p}_1 da \delta t,$$

where $H(\chi) = 1$ if $\chi > 0$ and $H(\chi) = 0$ if $\chi < 0$. Similarly, the number entering in δt from just outside da and with $\mathbf{p}_1 \in d\mathbf{p}_1$ is

$$-\frac{1}{m} (\mathbf{p}_1 - \mathbf{p}_0) \cdot \hat{\mathbf{n}} g^{(1)}(\mathbf{x}_1, \mathbf{p}_1) H[(\mathbf{p}_0 - \mathbf{p}_1) \cdot \hat{\mathbf{n}}] d\mathbf{p}_1 da \delta t$$

so that, as each molecule carries with it the mass m,

$$\frac{dM_R}{dt} = -\oint_R da \int d\mathbf{p}_1 \hat{\mathbf{n}} \cdot (\mathbf{p}_1 - \mathbf{p}_0) g^{(1)}(\mathbf{x}_1, \mathbf{p}_1) = 0, \tag{3.5}$$

by Eqs. (3.1) and (2.4). As this is true for any R in the system, we find from Eq. (3.4) that

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial \mathbf{x}_1} \cdot \rho \mathbf{v}_0 = 0, \tag{3.6}$$

which is the familiar equation of continuity in hydrodynamics.

From the foregoing analysis it should be clear that if $\chi(\mathbf{x}_1, \mathbf{p}_1)$ is some property associated with a molecule located at $\mathbf{x}_1$ with momentum $\mathbf{p}_1$, then the rate of change of X_R, the quantity of χ contained in R, due to the flow of molecules across the boundary of R is given by

$$\left(\frac{dX_R}{dt}\right)_{\text{flow}} = -\frac{1}{m} \oint_R da \int d\mathbf{p}_1 (\mathbf{p}_1 - \mathbf{p}_0) \cdot \hat{\mathbf{n}} \chi g^{(1)} \tag{3.7}$$
$$= -\frac{1}{m} \int_R d\mathbf{x}_1 \frac{\partial}{\partial \mathbf{x}_1} \cdot \int d\mathbf{p}_1 (\mathbf{p}_1 - \mathbf{p}_0) \chi g^{(1)},$$

where we have used Gauss's theorem. We shall make frequent use of Eq. (3.7) in what follows. Moreover, it should be noted that Eq. (3.7) applies regardless of whether the property χ is a scalar, vector, or tensor quantity.

4. TRANSPORT OF MOMENTUM

The momentum $\mathbf{p}_R$ contained in the region R at the time t is given by

$$\mathbf{p}_R = \int_R d\mathbf{x}_1 \int d\mathbf{p}_1 \mathbf{p}_1 g^{(1)}(\mathbf{x}_1, \mathbf{p}_1) = \int_R d\mathbf{x}_1 \rho \mathbf{v}_0, \tag{4.1}$$

by Eq. (3.1), so that the rate of change of $\mathbf{p}_R$ is

$$\frac{d\mathbf{P}_R}{dt} = \int_R d\mathbf{x}_1 \frac{\partial}{\partial t} \rho \mathbf{v}_0 + \oint_R da \hat{\mathbf{n}} \cdot \mathbf{v}_0 \rho \mathbf{v}_0$$
$$= \int_R d\mathbf{x}_1 \left(\frac{\partial}{\partial t} \rho \mathbf{v}_0 + \frac{\partial}{\partial \mathbf{x}_1} \cdot \mathbf{v}_0 \rho \mathbf{v}_0 \right)$$
$$= \int_R d\mathbf{x}_1 \rho \left(\frac{\partial}{\partial t} + \mathbf{v}_0 \cdot \frac{\partial}{\partial \mathbf{x}_1} \right) \mathbf{v}_0, \tag{4.2}$$

where, in the above, we first applied Gauss's theorem and then the continuity equation, Eq. (3.6), to obtain the result.

The contribution to the rate of change of $\mathbf{P}_R$ due to the flow of molecules across the boundary of R is, from Eq. (3.7),

$$\left(\frac{d\mathbf{P}_R}{dt}\right)_{\text{flow}} = -\frac{1}{m} \int_R d\mathbf{x}_1 \frac{\partial}{\partial \mathbf{x}_1} \cdot \int d\mathbf{p}_1 (\mathbf{p}_1 - \mathbf{p}_0) \times (\mathbf{p}_1 - \mathbf{p}_0) g^{(1)} \tag{4.3}$$

as $\int d\mathbf{p}_1 (\mathbf{p}_1 - \mathbf{p}_0) g^{(1)} = 0$.

If $\mathbf{F}(\mathbf{x}_1)$ is the external force per unit mass acting on the molecules in R, then $\mathbf{F}$ contributes

$$\left(\frac{d\mathbf{P}_R}{dt}\right)_{\text{ext}} = \int_R d\mathbf{x}_1 \int d\mathbf{p}_1 m \mathbf{F} g^{(1)} = \int_R d\mathbf{x}_1 \rho \mathbf{F} \tag{4.4}$$

to the rate of change of $\mathbf{P}_R$.

The force exerted on a molecule at $\mathbf{x}_1$ due to the presence of another molecule at $\mathbf{x}_2$ is

$$-\frac{\partial \phi(|\mathbf{x}_2 - \mathbf{x}_1|)}{\partial \mathbf{x}_1} = -\frac{\partial \phi}{\partial \mathbf{x}_1}(|\mathbf{r}|) = \frac{\mathbf{r}}{r} \frac{d\phi(r)}{dr}$$
$$= \hat{\mathbf{r}} \phi'(r)$$

where $\phi(r)$ is the intermolecular potential. Now, given a molecule at $\mathbf{x}_1$, the probability that there is another molecule located in $d\mathbf{x}_2$ is

$$\frac{f^{(2)}(\mathbf{x}_1, \mathbf{x}_2)}{f^{(1)}(\mathbf{x}_1)} d\mathbf{x}_2.$$

Thus, the rate of change of $\mathbf{P}_R$ due to intermolecular forces is

$$\left(\frac{d\mathbf{P}_R}{dt}\right)_{\text{int}} = \int_R d\mathbf{x}_1 \int d\mathbf{r} \hat{\mathbf{r}} \phi'(r) f^{(2)}(\mathbf{x}_1, \mathbf{x}_1 + \mathbf{r})$$
$$= \int_R d\mathbf{x}_1 \int d\mathbf{r} \hat{\mathbf{r}} \phi'(r) \gamma(\mathbf{r}, \mathbf{x}_1 + \tfrac{1}{2}\mathbf{r}). \tag{4.5}$$

We shall now express the integrand as the divergence of a tensor. We note that

$$\left(\frac{d\mathbf{P}_R}{dt}\right)_{\text{int}} = -\int_R d\mathbf{x}_1 \int d\mathbf{r} \hat{\mathbf{r}} \phi'(r) \gamma(-\mathbf{r}, \mathbf{x}_1 - \tfrac{1}{2}\mathbf{r})$$
$$= \tfrac{1}{2} \int_R d\mathbf{x}_1 \int d\mathbf{r} \hat{\mathbf{r}} \phi'(r) [\gamma(\mathbf{r}, \mathbf{x}_1 + \tfrac{1}{2}\mathbf{r})$$
$$- \gamma(\mathbf{r}, \mathbf{x}_1 - \tfrac{1}{2}\mathbf{r})], \tag{4.6}$$

where, we first let $\mathbf{r} \to -\mathbf{r}$ (without altering the limits) and then used Eq. (2.8). Now, since γ is a continuous function of the components of $\mathbf{x}$, we find

$$\gamma(\mathbf{r}, \mathbf{x}_1 + \tfrac{1}{2}\mathbf{r}) - \gamma(\mathbf{r}, \mathbf{x}_1 - \tfrac{1}{2}\mathbf{r})$$
$$= \int_{-1/2}^{1/2} d\eta \frac{\partial}{\partial \eta} \gamma(\mathbf{r}, \mathbf{x}_1 + \eta \mathbf{r})$$
$$= \int_{-1/2}^{1/2} d\eta \frac{\partial}{\partial \eta} \gamma(\mathbf{r}, \mathbf{y})$$
$$= \int_{-1/2}^{1/2} d\eta \frac{\partial \gamma}{\partial y_i} \frac{\partial y_i}{\partial \eta}$$
$$= \int_{-1/2}^{1/2} d\eta r_i \frac{\partial \gamma}{\partial y_i}, \tag{4.7}$$

where $\mathbf{y} = \mathbf{x}_1 + \eta \mathbf{r}$, the y_i and r_i are the Cartesian components of $\mathbf{y}$ and $\mathbf{r}$, and repeated indices are summed over. But,

$$\frac{\partial \gamma}{\partial \mathbf{x}_i} = \frac{\partial \gamma}{\partial y_j} \frac{\partial y_j}{\partial x_i} = \frac{\partial \gamma}{\partial y_i}.$$

Hence

$$\gamma(\mathbf{r}, \mathbf{x}_1 + \tfrac{1}{2}\mathbf{r}) - \gamma(\mathbf{r}, \mathbf{x}_1 - \tfrac{1}{2}\mathbf{r})$$
$$= \int_{-1/2}^{1/2} d\eta \, r_i \frac{\partial \gamma}{\partial x_i}$$

$$= \mathbf{r} \cdot \frac{\partial}{\partial \mathbf{x}_1} \int_{-1/2}^{1/2} d\eta \gamma(\mathbf{r}, \mathbf{x}_1 + \eta\mathbf{r})$$
$$= \mathbf{r} \cdot \frac{\partial}{\partial \mathbf{x}_1} \bar{\gamma}(\mathbf{r}, \mathbf{x}_1), \tag{4.8}$$

where we have defined $\bar{\gamma}$ by

$$\bar{\gamma}(\mathbf{r}, \mathbf{x}_1) = \int_{-1/2}^{1/2} d\eta \gamma(\mathbf{r}, \mathbf{x}_1 + \eta\mathbf{r}). \tag{4.9}$$

The intermolecular contribution to the rate of change of $\mathbf{P}_R$ may now be written as

$$\left(\frac{d\mathbf{P}_R}{dt}\right)_{\text{int}} = \int_R d\mathbf{x}_1 \frac{\partial}{\partial \mathbf{x}_1} \cdot \tfrac{1}{2} \int d\mathbf{r} \hat{\mathbf{r}} \mathbf{r} \phi'(\mathbf{r}) \bar{\gamma}(\mathbf{r}, \mathbf{x}_1). \tag{4.10}$$

Adding up all of the contributions and using the fact that R is arbitrary, we obtain the balance equation for the transport of momentum:

$$\rho\left(\frac{\partial}{\partial t} + \mathbf{v}_o \cdot \frac{\partial}{\partial \mathbf{x}_1}\right)\mathbf{v}_o$$
$$= \rho\mathbf{F} + \frac{\partial}{\partial \mathbf{x}_1} \cdot \left(-\frac{l}{m} \int d\mathbf{p}_1 (\mathbf{p}_1 - \mathbf{p}_0)(\mathbf{p}_1 - \mathbf{p}_0) g^{(1)}\right.$$
$$\left. + \tfrac{1}{2} \int d\mathbf{r}\, \hat{\mathbf{r}} \mathbf{r} \phi' \bar{\gamma}\right). \tag{4.11}$$

5. TRANSPORT OF ANGULAR MOMENTUM

Because we have assumed that the molecules of our system have no internal angular momentum, a molecule with position $\mathbf{x}_1$ and momentum $\mathbf{p}_1$ possesses solely its orbital angular momentum, $\mathbf{x}_1 \times \mathbf{p}_1$. We shall see later that such a system must have a symmetric stress tensor. Here, however, we shall simply find the balance equation for the transport of orbital angular momentum.

The angular momentum contained in R is

$$\mathbf{L}_R = \int_R d\mathbf{x}_1 \int d\mathbf{p}_1 (\mathbf{x}_1 \times \mathbf{p}_1) g^{(1)}$$
$$= \int_R d\mathbf{x}_1 \mathbf{x}_1 \times \int d\mathbf{p}_1 \mathbf{p}_1 g^{(1)}$$
$$= \int_R d\mathbf{x}_1 \mathbf{x}_1 \times \rho\mathbf{v}_o, \tag{5.1}$$

so that the rate of change of $\mathbf{L}_R$ is found in the usual way to be

$$\frac{d\mathbf{L}_R}{dt} = \int_R d\mathbf{x}_1 \mathbf{x}_1 \times \frac{\partial}{\partial t} \rho\mathbf{v}_o + \oint_R da\, \hat{\mathbf{n}} \cdot \mathbf{v}_o (\mathbf{x}_1 \times \rho\mathbf{v}_o)$$
$$= \int_R d\mathbf{x}_1 \rho\left(\frac{\partial}{\partial t} + \mathbf{v}_o \cdot \frac{\partial}{\partial \mathbf{x}_1}\right)(\mathbf{x}_1 \times \mathbf{v}_o), \tag{5.2}$$

where we have used the arguments leading to Eq. (4.2) to obtain Eq. (5.2).

The rate of flow of angular momentum into R is, from Eq. (3.7),

$$\left(\frac{d\mathbf{L}_R}{dt}\right)_{\text{flow}}$$
$$= -\frac{1}{m} \int_R d\mathbf{x}_1 \frac{\partial}{\partial \mathbf{x}_1} \cdot \int d\mathbf{p}_1 (\mathbf{p}_1 - \mathbf{p}_o)(\mathbf{x}_1 \times \mathbf{p}_1) g^{(1)}$$
$$= \frac{1}{m} \int_R d\mathbf{x}_1 \frac{\partial}{\partial \mathbf{x}_1}$$
$$\cdot \left(\int d\mathbf{p}_1 (\mathbf{p}_1 - \mathbf{p}_o)(\mathbf{p}_1 - \mathbf{p}_o) g^{(1)}\right) \times \mathbf{x}_1 . \tag{5.3}$$

We define the symmetric tensor $\mathbf{A}$ by

$$\mathbf{A} = \frac{1}{m} \int d\mathbf{p}_1 (\mathbf{p}_1 - \mathbf{p}_o)(\mathbf{p}_1 - \mathbf{p}_o) g^{(1)}, \tag{5.4}$$

so that

$$\left(\frac{d\mathbf{L}_R}{dt}\right)_{\text{flow}} = \int_R d\mathbf{x}_1 \frac{\partial}{\partial \mathbf{x}_1} \cdot (\mathbf{A} \times \mathbf{x}_1)$$
$$= \int_R d\mathbf{x}_1 \frac{\partial}{\partial \chi_i} (A_{ij} \chi_k \hat{\mathbf{e}}_j \times \hat{\mathbf{e}}_k)$$
$$= -\int_R d\mathbf{x}_1 \mathbf{x}_1 \times \frac{\partial}{\partial \mathbf{x}_1} \cdot \mathbf{A} + \int_R d\mathbf{x}_1 A_{ij} \hat{\mathbf{e}}_j \times \hat{\mathbf{e}}_i, \tag{5.5}$$

where the $\hat{\mathbf{e}}_i$ are the Cartesian unit vectors, the χ_i are the components of $\mathbf{x}_1$, the A_{ij} are the elements of $\mathbf{A}$, and we have employed the convention of summing over repeated indices. Now, since $\mathbf{A}$ is symmetric,

$$A_{ij} \hat{\mathbf{e}}_j \times \hat{\mathbf{e}}_i = 0, \tag{5.6}$$

and the contribution due to flow may be written as

$$\left(\frac{d\mathbf{L}_R}{dt}\right)_{\text{flow}} = \frac{1}{m} \int_R d\mathbf{x}_1 \frac{\partial}{\partial \mathbf{x}_1} \cdot \left[\left(\int d\mathbf{p}_1 (\mathbf{p}_1 - \mathbf{p}_o)\right.\right.$$
$$\left.\left. \times (\mathbf{p}_1 - \mathbf{p}_o) g^{(1)}\right) \times \mathbf{x}_1\right]$$
$$= -\frac{1}{m} \int_R d\mathbf{x}_1 \mathbf{x}_1 \times \frac{\partial}{\partial \mathbf{x}_1} \cdot \int d\mathbf{p}_1 (\mathbf{p}_1 - \mathbf{p}_0)$$
$$\times (\mathbf{p}_1 - \mathbf{p}_0) g^{(1)}. \tag{5.7}$$

Since the external forces are only functions of position, the contribution they make is simply

$$\left(\frac{d\mathbf{L}_R}{dt}\right)_{\text{ext}} = \int_R d\mathbf{x}_1 \mathbf{x}_1 \times \rho\mathbf{F}. \tag{5.8}$$

In the previous section we demonstrated that the average force on a molecule at $\mathbf{x}_1$ due to the presence of all other molecules is

$$\int d\mathbf{r} \hat{\mathbf{r}} \phi'(r) \frac{f^{(2)}(\mathbf{x}_1, \mathbf{x}_1 + \mathbf{r})}{f^{(1)}(\mathbf{x}_1)} .$$

Hence, the intermolecular forces contribute

$$\int_R d\mathbf{x}_1 \mathbf{x}_1 \times \int d\mathbf{r} \hat{\mathbf{r}} \phi' f^{(2)}(\mathbf{x}_1, \mathbf{x}_1 + \mathbf{r})$$

to the torque acting on the molecules in R. Applying Eqs. (4.5)–(4.9) to the above, we obtain

$$\left(\frac{d\mathbf{L}_R}{dt}\right)_{\text{int}} = \int_R d\mathbf{x}_1 \mathbf{x}_1 \times \left(\frac{\partial}{\partial \mathbf{x}_1} \cdot \tfrac{1}{2} \int d\mathbf{r}\, \hat{\mathbf{r}} \mathbf{r} \phi' \bar{\gamma}\right). \tag{5.9}$$

Now, the tensor

$$\int d\mathbf{r}\, \hat{\mathbf{r}}\mathbf{r}\phi'\bar{\gamma}$$

is obviously symmetric, so that the arguments used on **A** in Eqs. (5.5)–(5.7) also apply here, and Eq. (5.9) may be alternatively expressed by

$$\left(\frac{d\mathbf{L}_R}{dt}\right)_{\text{int}} = -\int_R d\mathbf{x}_1 \frac{\partial}{\partial \mathbf{x}_1} \cdot \left(\tfrac{1}{2}\int d\mathbf{r}\hat{\mathbf{r}}\mathbf{r}\phi'\bar{\gamma} \times \mathbf{x}_1\right). \tag{5.10}$$

Adding the contributions due to flow, external forces, and intermolecular forces, equating the sum to the rhs of Eq. (5.2), and using the fact that R is arbitrary, we obtain the balance equation for the transport of angular momentum:

$$\begin{aligned}\rho\left(\frac{\partial}{\partial t} + \mathbf{v}_0 \cdot \frac{\partial}{\partial \mathbf{x}_1}\right)(\mathbf{x}_1 \times \mathbf{v}_0) &= \mathbf{x}_1 \times \rho\mathbf{F} - \frac{\partial}{\partial \mathbf{x}_1} \cdot \left[\left(-\frac{1}{m}\int d\mathbf{p}_1(\mathbf{p}_1 - \mathbf{p}_0)\right.\right. \\ &\left.\left. \times(\mathbf{p}_1 - \mathbf{p}_0)g^{(1)} + \tfrac{1}{2}\int d\mathbf{r}\hat{\mathbf{r}}\mathbf{r}\phi'\bar{\gamma}\right) \times \mathbf{x}_1\right] \\ &= \mathbf{x}_1 \times \left[\rho\mathbf{F} + \frac{\partial}{\partial \mathbf{x}_1} \cdot \left(-\frac{1}{m}\int d\mathbf{p}_1(\mathbf{p}_1 - \mathbf{p}_0)\right.\right. \\ &\left.\left. \times (\mathbf{p}_1 - \mathbf{p}_0)g^{(1)} + \tfrac{1}{2}\int d\mathbf{r}\hat{\mathbf{r}}\mathbf{r}\phi'\bar{\gamma}\right)\right].\end{aligned} \tag{5.11}$$

The motivation for expressing this balance equation in two alternative forms will be made apparent when we discuss the stress tensor.

6. TRANSPORT OF ENERGY

We ascribe to each molecule in a given pair the potential energy $\frac{1}{2}\phi$, and we write $m\Phi(\mathbf{x}_1)$ as the scalar potential due to the external forces of a molecule at $\mathbf{x}_1$, so that

$$\mathbf{F} = -\frac{\partial}{\partial \mathbf{x}_1}\Phi. \tag{6.1}$$

With these definitions, the average energy of a molecule at $\mathbf{x}_1$ with momentum $\mathbf{p}_1$ is

$$e(\mathbf{x}_1, \mathbf{p}_1) = \frac{1}{2m}|\mathbf{p}_1|^2 + m\Phi + \tfrac{1}{2}\int d\mathbf{r}\int d\mathbf{p}_2\phi\frac{g^{(2)}(\mathbf{x}_1, \mathbf{p}_1; \mathbf{x}_1 + \mathbf{r}, \mathbf{p}_2)}{g^{(1)}(\mathbf{x}_1, \mathbf{p}_1)}, \tag{6.2}$$

while the average energy of a molecule located at $\mathbf{x}_1$ is

$$\bar{e}(\mathbf{x}_1) = m\Phi + \frac{1}{f^{(1)}} \times \left(\frac{1}{2m}\int d\mathbf{p}_1|\mathbf{p}_1|^2 g^{(1)} + \tfrac{1}{2}\int d\mathbf{r} f^{(2)}(\mathbf{x}_1, \mathbf{x}_1 + \mathbf{r})\phi\right). \tag{6.3}$$

The total energy E_R contained in the region R is then

$$E_R = \int_R d\mathbf{x}_1 f^{(1)}\bar{e} = \int_R d\mathbf{x}_1 \rho\epsilon, \tag{6.4}$$

where ϵ is the energy per unit mass:

$$\epsilon = \frac{1}{m}\bar{e}. \tag{6.5}$$

In the usual manner we find that the rate of change of E_R is given by

$$\frac{dE_R}{dt} = \int_R d\mathbf{x}_1 \rho\left(\frac{\partial}{\partial t} + \mathbf{v}_0 \cdot \frac{\partial}{\partial \mathbf{x}_1}\right)\epsilon. \tag{6.6}$$

The contribution due to flow is simply

$$\left(\frac{d\mathbf{E}_R}{dt}\right)_{\text{flow}} = -\frac{1}{m}\int_R d\mathbf{x}_1 \frac{\partial}{\partial \mathbf{x}_1} \cdot \int d\mathbf{p}_1(\mathbf{p}_1 - \mathbf{p}_0)g^{(1)}e \tag{6.7}$$

and it should be noted that, since Φ depends solely on $\mathbf{x}_1$, the external forces make no contribution to the flow term.

The average increase in the kinetic energy of a molecule at $\mathbf{x}_1$ with momentum $\mathbf{p}_1$ in the time δt is

$$\left(\mathbf{F}\cdot\mathbf{p}_1 + \frac{1}{mg^{(1)}}\mathbf{p}_1 \cdot \int d\mathbf{r}\int d\mathbf{p}_2\hat{\mathbf{r}}\phi' g^{(2)}(\mathbf{x}_1, \mathbf{p}_1; \mathbf{x}_1 + \mathbf{r}, \mathbf{p}_2)\right)\delta t$$

and the average increase of the potential energy of such a molecule in δt is

$$\left(\mathbf{p}_1 \cdot \frac{\partial\Phi}{\partial \mathbf{x}_1} + \frac{1}{2mg^{(1)}}\int d\mathbf{r}\int d\mathbf{p}_2(\mathbf{p}_2 - \mathbf{p}_1) \cdot \hat{\mathbf{r}}\phi' g^{(2)}(\mathbf{x}_1, \mathbf{p}_1; \mathbf{x}_1 + \mathbf{r}, \mathbf{p}_2)\right)\delta t,$$

where we have expanded ϕ and Φ through first order in δt to obtain the result. Multiplying the last two expressions by $g^{(1)}$, integrating over all $\mathbf{p}_1$ and over $\mathbf{x}_1 \epsilon R$, dividing by δt, and adding on the flow contribution, we find

$$\begin{aligned}\frac{d\mathbf{E}_R}{dt} &= -\int_R d\mathbf{x}_1 \frac{\partial}{\partial \mathbf{x}_1} \cdot \frac{1}{m}\int d\mathbf{p}_1(\mathbf{p}_1 - \mathbf{p}_0)eg^{(1)} \\ &\quad + \int_R d\mathbf{x}_1 \int d\mathbf{r}\int d\mathbf{p}_2 \frac{1}{2m}(\mathbf{p}_1 + \mathbf{p}_2) \\ &\quad \cdot \hat{\mathbf{r}}\phi' g^{(2)}(\mathbf{x}_1, \mathbf{p}_1; \mathbf{x}_1 + \mathbf{r}, \mathbf{p}_2),\end{aligned} \tag{6.8}$$

where we have used Eq. (6.1) in the above.

We shall now proceed in a manner similar to that of Sec. 4 to express the second term on the rhs of Eq. (6.8), the intermolecular term, as the integral over $\mathbf{x}_1 \epsilon R$ of the divergence of a vector quantity. Transforming to **w**, **r**, and **p** coordinates, the intermolecular term may be written as

$$\int_R d\mathbf{x}_1 \int d\mathbf{r}\phi'\hat{\mathbf{r}} \cdot \int d\mathbf{w}\int d\mathbf{p}\frac{1}{m}\mathbf{p}\Gamma(\mathbf{w}, \mathbf{r}; \mathbf{p}, \mathbf{x}_1 + \tfrac{1}{2}\mathbf{r}).$$

Defining $\mathbf{b}(\mathbf{r},\mathbf{x})$ by

$$\mathbf{b}(\mathbf{r},\mathbf{x}) = \int d\mathbf{w}\int d\mathbf{p}\mathbf{p}\Gamma(\mathbf{w},\mathbf{r};\mathbf{p},\mathbf{x}), \tag{6.9}$$

we have

$$\begin{aligned}\mathbf{b}(-\mathbf{r},\mathbf{x}) &= \int d\mathbf{w}\int d\mathbf{p}\mathbf{p}\Gamma(\mathbf{w},-\mathbf{r};\mathbf{p},\mathbf{x})\\ &= \int d\mathbf{w}\int d\mathbf{p}\mathbf{p}\Gamma(-\mathbf{w},\mathbf{r};\mathbf{p},\mathbf{x})\\ &= \int d\mathbf{w}\int d\mathbf{p}\mathbf{p}\Gamma(\mathbf{w},\mathbf{r};\mathbf{p},\mathbf{x})\\ &= \mathbf{b}(\mathbf{r},\mathbf{x}),\end{aligned} \tag{6.10}$$

where we first used Eq. (2.8) and then let $\mathbf{w}\to-\mathbf{w}$ without altering the limits. We may now write the intermolecular term as

$$\begin{aligned}&\frac{1}{m}\int_R d\mathbf{x}_1\int d\mathbf{r}\,\phi'\hat{\mathbf{r}}\cdot\mathbf{b}(\mathbf{r},\mathbf{x}_1+\tfrac{1}{2}\mathbf{r})\\ &\quad=-\frac{1}{m}\int_R d\mathbf{x}_1\int d\mathbf{r}\,\phi'\hat{\mathbf{r}}\cdot\mathbf{b}(-\mathbf{r},\mathbf{x}_1-\tfrac{1}{2}\mathbf{r})\\ &\quad=\frac{1}{2m}\int_R d\mathbf{x}_1\int d\mathbf{r}\,\phi'\hat{\mathbf{r}}\cdot[\mathbf{b}(\mathbf{r},\mathbf{x}_1+\tfrac{1}{2}\mathbf{r})\\ &\quad\quad-\mathbf{b}(\mathbf{r},\mathbf{x}_1-\tfrac{1}{2}\mathbf{r})]\\ &\quad=\frac{1}{2m}\int_R d\mathbf{x}_1\int d\mathbf{r}\,\phi'\hat{\mathbf{r}}\cdot\int d\mathbf{w}\int d\mathbf{p}\\ &\quad\quad\times\mathbf{p}[\Gamma(\mathbf{w},\mathbf{r};\mathbf{p},\mathbf{x}_1+\tfrac{1}{2}\mathbf{r})-\Gamma(\mathbf{w},\mathbf{r};\mathbf{p},\mathbf{x}_1-\tfrac{1}{2}\mathbf{r})].\end{aligned}$$

But, applying the arguments used in Eqs. (4.7)–(4.9), we find that

$$\Gamma(\mathbf{w},\mathbf{r};\mathbf{p},\mathbf{x}_1+\tfrac{1}{2}\mathbf{r})-\Gamma(\mathbf{w},\mathbf{r};\mathbf{p},\mathbf{x}_1-\tfrac{1}{2}\mathbf{r}) = \mathbf{r}\cdot\frac{\partial}{\partial\mathbf{x}_1}\overline{\Gamma}, \tag{6.11}$$

where $\overline{\Gamma}$ is given by

$$\overline{\Gamma}(\mathbf{w},\mathbf{p};\mathbf{r},\mathbf{x}_1) = \int_{-1/2}^{1/2} d\eta\;\Gamma(\mathbf{w},\mathbf{r};\mathbf{p},\mathbf{x}_1+\eta\mathbf{r}). \tag{6.12}$$

Finally, the intermolecular contribution may be expressed as

$$\frac{1}{2m}\int_R d\mathbf{x}_1\frac{\partial}{\partial\mathbf{x}_1}\cdot\int d\mathbf{r}\,\hat{\mathbf{r}}\phi'\mathbf{r}\cdot\int d\mathbf{w}\int d\mathbf{p}\,\mathbf{p}\overline{\Gamma}.$$

Adding up the contributions and using the fact that R is arbitrary, we find that the balance equation for the transport of energy may be written as

$$\begin{aligned}&\rho\left(\frac{\partial}{\partial t}+\mathbf{v}_0\cdot\frac{\partial}{\partial\mathbf{x}_1}\right)\epsilon\\ &\quad=\frac{\partial}{\partial\mathbf{x}_1}\cdot\left(-\frac{1}{m}\int d\mathbf{p}_1(\mathbf{p}_1-\mathbf{p}_0)g^{(1)}e\right.\\ &\quad\quad\left.+\frac{1}{2m}\int d\mathbf{r}\,\hat{\mathbf{r}}\phi'\mathbf{r}\cdot\int d\mathbf{w}\int d\mathbf{p}\,\mathbf{p}\overline{\Gamma}\right).\end{aligned} \tag{6.13}$$

While the term in large parentheses may be interpreted as the total energy flux, in order to find the heat flux we shall first need to determine the stress tensor.

7. THE STRESS TENSOR

In hydrodynamics and the theory of continuous media it is assumed that the rate of change of $\mathbf{P}_R$ is due to forces acting on the surface of R (stresses) and external forces (or body forces). If da is an element of the boundary surface of R, and if $\mathbf{T}$ is the stress tensor, then the force acting on da is $\hat{\mathbf{n}}\cdot\mathbf{T}da$, where $\hat{\mathbf{n}}$ is the outward directed unit vector normal to da. Thus, if $\mathbf{T}$ exists,

$$\begin{aligned}\frac{d\mathbf{P}_R}{dt} &= \oint_R da\,\hat{\mathbf{n}}\cdot\mathbf{T}+\int_R d\mathbf{x}_1\rho\mathbf{F}\\ &= \int_R d\mathbf{x}_1\left(\rho\mathbf{F}+\frac{\partial}{\partial\mathbf{x}_1}\cdot\mathbf{T}\right),\end{aligned} \tag{7.1}$$

so that the momentum balance equation of hydrodynamics is, from Eq. (4.2),

$$\rho\left(\frac{\partial}{\partial t}+\mathbf{v}_0\cdot\frac{\partial}{\partial\mathbf{x}_1}\right)\mathbf{v}_0 = \frac{\partial}{\partial\mathbf{x}_1}\cdot\mathbf{T}+\rho\mathbf{F}. \tag{7.2}$$

The rate of change of $\mathbf{L}_R$ in hydrodynamics is

$$\begin{aligned}\frac{d\mathbf{L}_R}{dt} &= \oint_R da\,\mathbf{x}_1\times\hat{\mathbf{n}}\cdot\mathbf{T}+\int_R d\mathbf{x}_1\,\mathbf{x}_1\times\rho\mathbf{F}\\ &= \int_R d\mathbf{x}_1\left[\mathbf{x}_1\times\rho\mathbf{F}-\frac{\partial}{\partial\mathbf{x}_1}\cdot(\mathbf{T}\times\mathbf{x}_1)\right]\end{aligned} \tag{7.3}$$

giving us, with Eq. (5.2), the balance equation

$$\begin{aligned}&\rho\left(\frac{\partial}{\partial t}+\mathbf{v}_0\cdot\frac{\partial}{\partial\mathbf{x}_1}\right)(\mathbf{x}_1\times\mathbf{v}_0)\\ &\quad=\mathbf{x}_1\times\rho\mathbf{F}-\frac{\partial}{\partial\mathbf{x}_1}\cdot(\mathbf{T}\times\mathbf{x}_1).\end{aligned} \tag{7.4}$$

Now, it is easy to verify that

$$\begin{aligned}&\left(\frac{\partial}{\partial t}+\mathbf{v}_0\cdot\frac{\partial}{\partial\mathbf{x}_1}\right)(\mathbf{x}_1\times\mathbf{v}_0)\\ &\quad=\mathbf{x}_1\times\left(\frac{\partial}{\partial t}+\mathbf{v}_0\cdot\frac{\partial}{\partial\mathbf{x}_1}\right)\mathbf{v}_0,\end{aligned} \tag{7.5}$$

which may be combined with Eqs. (7.2)–(7.4) to give us the further condition on $\mathbf{T}$:

$$\mathbf{x}_1\times\left(\frac{\partial}{\partial\mathbf{x}_1}\cdot\mathbf{T}\right)+\frac{\partial}{\partial\mathbf{x}_1}\cdot(\mathbf{T}\times\mathbf{x}_1)=0. \tag{7.6}$$

This condition is satisfied if and only if $\mathbf{T}$ is symmetric.

Hence, the conditions that any candidate for the stress tensor must satisfy are, first of all, that it satisfy the momentum and angular momentum balance equations, secondly, that it be symmetric, and thirdly, that it reduce to the familiar result obtained from the virial theorem for a system in thermal and mechanical equilibrium. Thus, from Eqs. (4.10) and (5.11), and from the fact that, in

equilibrium, γ depends only on $\mathbf{r}$, the stress tensor is given by

$$\mathbf{T} = -\frac{1}{m}\int d\mathbf{p}_1(\mathbf{p}_1 - \mathbf{p}_0)(\mathbf{p}_1 - \mathbf{p}_0)g^{(1)} + \tfrac{1}{2}\int d\mathbf{r}\,\hat{\mathbf{r}}\mathbf{r}\phi'\bar{\gamma}. \tag{7.7}$$

8. THE HEAT FLUX

In hydrodynamics the rate of change of E_R is ascribed to the work done by the surface forces and the flow of heat across the boundary surface of R. Thus,

$$\frac{dE_R}{dt} = \oint_R da\,\hat{\mathbf{n}}\cdot\mathbf{T}\cdot\mathbf{v}_0 - \oint_R da\,\hat{\mathbf{n}}\cdot\mathbf{q} = \int_R d\mathbf{x}_1 \frac{\partial}{\partial \mathbf{x}_1}\cdot(\mathbf{T}\cdot\mathbf{v}_0 - \mathbf{q}), \tag{8.1}$$

where $\mathbf{q}$ is defined as the heat flux, or conductive heat current. The balance equation for the transport of energy is, from Eq. (6.6),

$$\rho\left(\frac{\partial}{\partial t} + \mathbf{v}_0\cdot\frac{\partial}{\partial \mathbf{x}_1}\right)\epsilon = \frac{\partial}{\partial \mathbf{x}_1}\cdot(\mathbf{T}\cdot\mathbf{v}_0 - \mathbf{q}) \tag{8.2}$$

so that, from Eq. (6.13),

$$\mathbf{q} = \mathbf{T}\cdot\mathbf{v}_0 + \frac{1}{m}\int d\mathbf{p}_1(\mathbf{p}_1 - \mathbf{p}_0)g^{(1)}e - \frac{1}{2m}\int d\mathbf{r}\,\hat{\mathbf{r}}\phi'\mathbf{r}\cdot\int d\mathbf{w}\int d\mathbf{p}\,\mathbf{p}\bar{\Gamma}. \tag{8.3}$$

Inserting the rhs of Eq. (7.7) in the above, and noting that

$$\bar{\gamma}(\mathbf{r}, \mathbf{x}_1) = \int d\mathbf{w}\int d\mathbf{p}\bar{\Gamma}, \tag{8.4}$$

we find, with a little rearrangement,

$$\mathbf{q} = \frac{1}{2m^2}\int d\mathbf{p}_1(\mathbf{p}_1 - \mathbf{p}_0)|\mathbf{p}_1 - \mathbf{p}_0|^2 g^{(1)} + \frac{1}{2m}\int d\mathbf{p}_1\int d\mathbf{p}_2\int d\mathbf{r}(\mathbf{p}_1 - \mathbf{p}_0)\phi \times g^{(2)}(\mathbf{x}_1, \mathbf{p}_1; \mathbf{x}_1 + \mathbf{r}, \mathbf{p}_2) - \frac{1}{2m}\int d\mathbf{r}\,\hat{\mathbf{r}}\phi'\mathbf{r}\cdot\int d\mathbf{p}\int d\mathbf{w}(\mathbf{p} - \mathbf{p}_0)\bar{\Gamma}. \tag{8.5}$$

Any candidate for the heat flux not only must satisfy Eq. (8.2), but must also vanish for a system in thermal and mechanical equilibrium. By inserting the Maxwellian form of the momentum distribution functions on the rhs of Eq. (8.5), it is easily verified that our expression for $\mathbf{q}$ vanishes in equilibrium, and hence is the correct choice.

9. TRANSPORT OF INTERNAL ENERGY

The internal energy of a molecule located at $\mathbf{x}_1$ is defined as the sum of its average intermolecular potential energy and the difference between its average kinetic energy and $(1/2m)|\mathbf{p}_0|^2$. The appropriate transport equation is obtained easily from the balance equations for momentum and energy transport.

We observe that

$$\frac{\partial}{\partial \mathbf{x}_1}\cdot(\mathbf{T}\cdot\mathbf{v}_0) = \mathbf{v}_0\cdot\left(\frac{\partial}{\partial \mathbf{x}_1}\cdot\mathbf{T}\right) + \mathbf{T}:\frac{\partial \mathbf{v}_0}{\partial \mathbf{x}_1} = \rho\left(\frac{\partial}{\partial t} + \mathbf{v}_0\cdot\frac{\partial}{\partial \mathbf{x}_1}\right)\tfrac{1}{2}|\mathbf{v}_0|^2 + \mathbf{T}:\frac{\partial}{\partial \mathbf{x}_1} - \rho\mathbf{F}, \tag{9.1}$$

where we first asserted the symmetry of $\mathbf{T}$ and then used Eq. (8.2). Inserting this result into Eq. (8.2) we obtain

$$\rho\left(\frac{\partial}{\partial t} + \mathbf{v}_0\cdot\frac{\partial}{\partial \mathbf{x}_1}\right)\left[\frac{1}{2m^2}\int d\mathbf{p}_1(|\mathbf{p}_1|^2 - |\mathbf{p}_0|^2)\frac{g^{(1)}}{f^{(1)}} + \frac{1}{2m}\int d\mathbf{r}\,\phi\frac{f^{(2)}(\mathbf{x}_1, \mathbf{x}_1 + \mathbf{r})}{f^{(1)}}\right] = \mathbf{T}:\frac{\partial \mathbf{v}_0}{\partial \mathbf{x}_1} - \frac{\partial}{\partial \mathbf{x}_1}\cdot\mathbf{q}, \tag{9.2}$$

where the term in square brackets is the internal energy per unit mass. Now, the kinetic temperature θ is defined by

$$\theta(\mathbf{x}_1) = \frac{1}{3m}\int d\mathbf{p}_1|\mathbf{p}_1 - \mathbf{p}_0|^2\frac{g^{(1)}}{f^{(1)}} = \frac{1}{3m}\int d\mathbf{p}_1(|\mathbf{p}_1|^2 - |\mathbf{p}_0|^2)\frac{g^{(1)}}{f^{(1)}}, \tag{9.3}$$

and we define the internal potential energy u by

$$u = \tfrac{1}{2}\int d\mathbf{r}\phi\frac{f^{(2)}(\mathbf{x}_1, \mathbf{x}_1 + \mathbf{r})}{f^{(1)}} \tag{9.4}$$

so that u is the average potential energy of a molecule at $\mathbf{x}_1$. Inserting these definitions into Eq. (9.2) we obtain

$$\frac{1}{m}\rho\left(\frac{\partial}{\partial t} + \mathbf{v}_0\cdot\frac{\partial}{\partial \mathbf{x}_1}\right)(\tfrac{3}{2}\theta + u) = \mathbf{T}:\frac{\partial \mathbf{v}_0}{\partial \mathbf{x}_1} - \frac{\partial}{\partial \mathbf{x}_1}\cdot\mathbf{q}, \tag{9.5}$$

the balance equation for the transport of internal energy.

10. PROPERTIES OF A SYSTEM OF RIGID SPHERES

In the previous sections we have shown that well-defined representations for the stress tensor and heat flux may be found for systems with a continuous intermolecular potential ϕ. We shall now demonstrate that this is also true for a system consisting of perfectly rigid spheres of mass m and radius σ.

Before we start, we need to consider two more characteristics of the distribution functions Γ and γ for a system consisting specifically of rigid spheres (the definitions and some of the general properties of these functions are in Sec. 2). First of all, since two rigid spheres cannot overlap, we note that Γ and γ must vanish if $r < 2\sigma$; clearly these is a discontinuity at $r = 2\sigma$. In this work we shall define $\Gamma(\mathbf{w}, 2\sigma\hat{\mathbf{r}}; \mathbf{p}, \mathbf{x})$ and $\gamma(2\sigma\hat{\mathbf{r}}, \mathbf{x})$ by

$$\Gamma(\mathbf{w}, 2\sigma\hat{\mathbf{r}}; \mathbf{p}, \mathbf{x}) = \lim_{\epsilon \to +0} \Gamma(\mathbf{w}, (1+\epsilon)2\sigma\hat{\mathbf{r}}; \mathbf{p}, \mathbf{x}),$$
$$\gamma(2\sigma\hat{\mathbf{r}}, \mathbf{x}) = \lim_{\epsilon \to +0} \gamma((1+\epsilon)2\sigma\hat{\mathbf{r}}, \mathbf{x}), \tag{10.1}$$

that is, as the limit $r \to 2\sigma$ taken from above. These quantities are perfectly well defined and have a definite physical interpretation.

In order to develop the second property, we consider the collision of two rigid spheres at the time t. Suppose that, at the time $t - 0$, their momenta are

$$\mathbf{p}_1 = \mathbf{p}_{1\perp} + \mathbf{p}_1 \cdot \hat{\mathbf{r}}\hat{\mathbf{r}} \quad \text{and} \quad \mathbf{p}_2 = \mathbf{p}_{2\perp} + \mathbf{p}_2 \cdot \hat{\mathbf{r}}\hat{\mathbf{r}},$$

where $\mathbf{p}_{1\perp}$ and $\mathbf{p}_{2\perp}$ are the projections of $\mathbf{p}_1$ and $\mathbf{p}_2$ in a plane perpendicular to $\hat{\mathbf{r}}$, so that after the collision, at the time $t + 0$, their momenta are

$$\mathbf{p}_1' = \mathbf{p}_{1\perp} + \mathbf{p}_2 \cdot \hat{\mathbf{r}}\hat{\mathbf{r}} \quad \text{and} \quad \mathbf{p}_2' = \mathbf{p}_{2\perp} + \mathbf{p}_1 \cdot \hat{\mathbf{r}}\hat{\mathbf{r}}.$$

We observe that the collision has duration zero, so that no simultaneous triple, or higher multiplicity, collisions occur. Thus, if two molecules are approaching each other at the time $t - 0$ with separation $2\sigma + 0$, the outcome of the collision is completely determined. The probability that two molecules are approaching each other at $t - 0$ with momenta $\mathbf{p}_1$ and $\mathbf{p}_2$ must be equal to the probability that they are moving apart at $t + 0$ with momenta $\mathbf{p}_1'$ and $\mathbf{p}_2'$, so that Γ must have the property

$$\Gamma(\mathbf{w}_\perp + \mathbf{w} \cdot \hat{\mathbf{r}}\hat{\mathbf{r}}, 2\sigma\hat{\mathbf{r}}; \mathbf{p}, \mathbf{x}) = \Gamma(\mathbf{w}_\perp - \mathbf{w} \cdot \hat{\mathbf{r}}\hat{\mathbf{r}}, 2\sigma\hat{\mathbf{r}}; \mathbf{p}, \mathbf{x}) \tag{10.2}$$

for a system consisting specifically of rigid spheres.

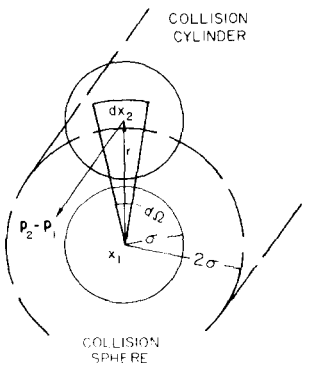


FIG. 1. Collision of two rigid spheres

11. COLLISIONAL CONTRIBUTIONS TO THE EQUATIONS OF TRANSPORT

The differences in the equations of transport between a system of rigid spheres and a system with a continuous intermolecular potential arise only in the intermolecular contributions. Rather than repeat all of the formalism of the first nine sections, our prescription will be first to calculate the collisional, or intermolecular, contributions of a system of rigid spheres to the various transport processes, and then to modify the previous results accordingly. We shall be concerned with the average increase of some property of a hard-sphere molecule located at $\mathbf{x}_1$, with momentum $\mathbf{p}_1$ due to collisions with other molecules in the time δt. Hence, we need to know the probability that such a collision occurs in the time δt.

Suppose we are given a molecule at $\mathbf{x}_1$ with momentum $\mathbf{p}_1$: the probability that it has a collision with another molecule having $\mathbf{p}_2 \in d\mathbf{p}_2$ and $\mathbf{x}_1 \in d\mathbf{x}_1 = (2\sigma)^2(1/m)|(\mathbf{p}_2 - \mathbf{p}_1)\cdot\hat{\mathbf{r}}|d\Omega\delta t$ in the time δt is

$$\frac{g^{(2)}(\mathbf{x}_1, \mathbf{p}_1; \mathbf{x}_1 + 2\sigma\hat{\mathbf{r}}, \mathbf{p}_2)}{g^{(1)}(\mathbf{x}_1, \mathbf{p}_1)} (2\sigma)^2 \frac{1}{m} |\mathbf{w}\cdot\hat{\mathbf{r}}| H(-\mathbf{w}\cdot\hat{\mathbf{r}}) \times d\Omega d\mathbf{p}_2 \delta t, \tag{11.1}$$

where $d\Omega$ is the solid angle subtended by $d\mathbf{x}_2$ at $\mathbf{x}_1$ (see Fig. 1), 2σ is the radius of the collision sphere, and the requirement $\mathbf{w}\cdot\hat{\mathbf{r}} < 0$ ($H(\chi)$ is defined in Sec. 3) ensures that the molecules will collide in the limit $\delta t \to 0$. Thus, if $\Delta\chi$ is the increase in the property χ of the molecule at $\mathbf{x}_1$ with momentum $\mathbf{p}_1$ due to a collision with a molecule at $\mathbf{x}_1 + 2\sigma\hat{\mathbf{r}}$ with momentum $\mathbf{p}_2$, then the rate of change of X_R due to collisions is

$$\left(\frac{dX_R}{dt}\right)_{\text{col}} = \int_R d\mathbf{x}_1 \int d\mathbf{p}_1 \int d\mathbf{p}_2 \int d\Omega \frac{(2\sigma)^2}{m} \times \Delta\chi |\mathbf{w}\cdot\hat{\mathbf{r}}| H(-\mathbf{w}\cdot\hat{\mathbf{r}}) g^{(2)}(\mathbf{x}_1, \mathbf{p}_1; \mathbf{x}_1 + 2\sigma\hat{\mathbf{r}}, \mathbf{p}_2)$$
$$= -\int_R d\mathbf{x}_1 \int d\mathbf{p} \int d\mathbf{w} \int d\Omega \frac{(2\sigma)^2}{m} \Delta\chi\, \mathbf{w}\cdot\hat{\mathbf{r}}\, H(-\mathbf{w}\cdot\hat{\mathbf{r}}) \times \Gamma(\mathbf{w}, 2\sigma\hat{\mathbf{r}}; \mathbf{p}, \mathbf{x}_1 + \sigma\hat{\mathbf{r}}), \tag{11.2}$$

where X_R is the amount of χ contained in the region R.

The momentum increase of the molecule at $\mathbf{x}_1$ due to the collision is $\mathbf{w}\cdot\hat{\mathbf{r}}\hat{\mathbf{r}}$ so that

$$\left(\frac{d\mathbf{P}_R}{dt}\right)_{\text{col}} = -\int_R d\mathbf{x}_1 \int d\mathbf{p} \int d\mathbf{w} \int d\Omega \frac{(2\sigma)^2}{m} \times (\mathbf{w}\cdot\hat{\mathbf{r}})^2 \hat{\mathbf{r}}\, H(-\mathbf{w}\cdot\hat{\mathbf{r}}) \Gamma; \tag{11.3}$$

the increase in angular momentum is $\mathbf{x}_1 \times (\Delta\mathbf{p}_1)$ giving us

$$\left(\frac{d\mathbf{L}_R}{dt}\right)_{\text{col}} = -\int_R d\mathbf{x}_1 \mathbf{x}_1 \times \int d\mathbf{p} \int d\mathbf{w} \int d\Omega \frac{(2\sigma)^2}{m}$$

$$\times (\mathbf{w}\cdot\hat{\mathbf{r}})^2\hat{\mathbf{r}}\,H(-\mathbf{w}\cdot\hat{\mathbf{r}})\Gamma; \tag{11.4}$$

and, as the increase in energy of the molecule at $\mathbf{x}_1$ due to the collision is $(1/m)(\mathbf{w}\cdot\hat{\mathbf{r}})(\mathbf{p}\cdot\hat{\mathbf{r}})$,

$$\left(\frac{dE_R}{dt}\right)_{\text{col}} = -\int_R d\mathbf{x}_1 \int d\mathbf{p}\int d\mathbf{w}\int d\Omega\left(\frac{2\sigma}{m}\right)^2 \times (\mathbf{w}\cdot\hat{\mathbf{r}})^2\mathbf{p}\cdot\hat{\mathbf{r}}\,H(-\mathbf{w}\cdot\hat{\mathbf{r}})\Gamma. \tag{11.5}$$

The integrands in Eqs. (11.3)–(11.5) depend on $\mathbf{w}\cdot\hat{\mathbf{r}}$ precisely in the same manner; namely, through the function $(\mathbf{w}\cdot\hat{\mathbf{r}})^2 H(-\mathbf{w}\cdot\hat{\mathbf{r}})\Gamma$. But, from Eq. (10.2), we see that Γ is an even function of $\mathbf{w}\cdot\hat{\mathbf{r}}$. Hence, if we integrate over $\mathbf{w}$ before we integrate over Ω, and if we break up the integration over $\mathbf{w}$ into an integration over $\mathbf{w}\cdot\hat{\mathbf{r}}\hat{\mathbf{r}}$ (where $\hat{\mathbf{r}}$ is fixed) and then an integration over $\mathbf{w}_\perp$, we may replace $H(-\mathbf{w}\cdot\hat{\mathbf{r}})$ by $\frac{1}{2}$ in the integrand without altering the result. As the order of integration is arbitrary, we shall assume in what follows that this substitution has been executed.

12. THE BALANCE EQUATIONS, STRESS TENSOR, AND HEAT FLUX FOR A SYSTEM OF RIGID SPHERES

Equations (11.3)–(11.5) are the collisional contributions to the transport of momentum, angular momentum, and energy for a system of rigid spheres. Each shall replace the appropriate intermolecular contribution (with subscript int) in the earlier work for a system with a continuous intermolecular potential. As in the earlier work, we shall express each contribution as an integral over $\mathbf{x}_1 \in R$ of the divergence of a tensor, or vector, quantity. This is quickly accomplished by observing that each collisional contribution contains

$$\begin{aligned}
&\int d\Omega\int d\mathbf{w}\,\hat{\mathbf{r}}(\mathbf{w}\cdot\hat{\mathbf{r}})^2\Gamma(\mathbf{w}, 2\sigma\hat{\mathbf{r}}; \mathbf{p}, \mathbf{x}_1 + \sigma\hat{\mathbf{r}})\\
&\quad = -\int d\Omega\int d\mathbf{w}\,\hat{\mathbf{r}}(\mathbf{w}\cdot\hat{\mathbf{r}})^2\Gamma(\mathbf{w}, -2\sigma\hat{\mathbf{r}}; \mathbf{p}, \mathbf{x}_1 - \sigma\hat{\mathbf{r}})\\
&\quad = -\int d\Omega\int d\mathbf{w}\,\hat{\mathbf{r}}(\mathbf{w}\cdot\hat{\mathbf{r}})^2\Gamma(-\mathbf{w}, 2\sigma\hat{\mathbf{r}}; \mathbf{p}, \mathbf{x}_1 - \sigma\hat{\mathbf{r}})\\
&\quad = -\int d\Omega\int d\mathbf{w}\,\hat{\mathbf{r}}(\mathbf{w}\cdot\hat{\mathbf{r}})^2\Gamma(\mathbf{w}, 2\sigma\hat{\mathbf{r}}; \mathbf{p}, \mathbf{x}_1 - \sigma\hat{\mathbf{r}})\\
&\quad = \tfrac{1}{2}\int d\Omega\int d\mathbf{w}\hat{\mathbf{r}}(\mathbf{w}\cdot\hat{\mathbf{r}})^2[\Gamma(\mathbf{w}, 2\sigma\hat{\mathbf{r}}; \mathbf{p}, \mathbf{x}_1 + \sigma\hat{\mathbf{r}})\\
&\quad\quad - \Gamma(\mathbf{w}, 2\sigma\hat{\mathbf{r}}; \mathbf{p}, \mathbf{x}_1 - \sigma\hat{\mathbf{r}}]\\
&\quad = \sigma\int d\Omega\int d\mathbf{w}\,\hat{\mathbf{r}}\hat{\mathbf{r}}(\mathbf{w}\cdot\hat{\mathbf{r}})^2\cdot\frac{\partial}{\partial\mathbf{x}_1}\overline{\Gamma}(\mathbf{w}, \mathbf{p}; 2\sigma\hat{\mathbf{r}}, \mathbf{x}_1),
\end{aligned} \tag{12.1}$$

where, in the first step we let $\mathbf{r}\to-\mathbf{r}$ without changing the limits on Ω, in the second step we used Eq. (2.8), in the third step we let $\mathbf{w}\to-\mathbf{w}$ without changing the limits, and in the last step we used Eq. (6.11). Inserting this result into Eqs. (11.3)–(11.5), we obtain

$$\left(\frac{d\mathbf{P}_R}{dt}\right)_{\text{col}} = -\int_R d\mathbf{x}_1\,\frac{\partial}{\partial\mathbf{x}_1}\cdot\frac{2\sigma^3}{m}\int d\mathbf{p}\int d\mathbf{w}\int d\Omega\,\hat{\mathbf{r}}\hat{\mathbf{r}}(\mathbf{w}\cdot\hat{\mathbf{r}})^2\overline{\Gamma}, \tag{12.2}$$

$$\left(\frac{d\mathbf{L}_R}{dt}\right)_{\text{col}} = -\int_R d\mathbf{x}_1\mathbf{x}_1\times\frac{\partial}{\partial\mathbf{x}_1}\cdot\frac{2\sigma^3}{m}\int d\mathbf{p}\int d\mathbf{w}\int d\Omega \times\hat{\mathbf{r}}\hat{\mathbf{r}}(\mathbf{w}\cdot\hat{\mathbf{r}})^2\overline{\Gamma}, \tag{12.3}$$

$$\left(\frac{dE_R}{dt}\right)_{\text{col}} = -\int_R d\mathbf{x}_1\,\frac{\partial}{\partial\mathbf{x}_1}\cdot\frac{2\sigma^3}{m^2}\int d\mathbf{p}\int d\mathbf{w}\int d\Omega \times\hat{\mathbf{r}}(\mathbf{p}\cdot\hat{\mathbf{r}})(\mathbf{w}\cdot\hat{\mathbf{r}})^2\overline{\Gamma}. \tag{12.4}$$

It is now a straightforward matter to obtain the balance equations of transport for a system of rigid spheres; we simply perform the aforementioned substitution of each collisional contribution for the appropriate intermolecular contribution in the balance equations obtained for a system with a continuous intermolecular potential [Eqs. (4.11), (5.11), and (6.13)]. We must also delete the terms containing ϕ (as opposed to ϕ') in the energy per unit mass and in the flow contribution in the earlier energy balance equation, Eq. (6.13), since there is no long-range intermolecular potential in our hard-sphere system. Performing these operations, we find that, for a system of rigid spheres, the momentum balance equation is

$$\rho\left(\frac{\partial}{\partial t} + \mathbf{v}_0\cdot\frac{\partial}{\partial\mathbf{x}_1}\right)\mathbf{v}_0 = \rho\mathbf{F} - \frac{\partial}{\partial\mathbf{x}_1}\cdot\left(\frac{1}{m}\int d\mathbf{p}_1(\mathbf{p}_1-\mathbf{p}_0)(\mathbf{p}_1-\mathbf{p}_0)g^{(1)} + \frac{2\sigma^3}{m}\int d\mathbf{p}\int d\mathbf{w}\int d\Omega\,\hat{\mathbf{r}}\hat{\mathbf{r}}(\mathbf{w}\cdot\hat{\mathbf{r}})^2\overline{\Gamma}\right), \tag{12.5}$$

the angular momentum equation is

$$\rho\left(\frac{\partial}{\partial t} + \mathbf{v}_0\cdot\frac{\partial}{\partial\mathbf{x}_1}\right)(\mathbf{x}_1\times\mathbf{v}_0) = \mathbf{x}_1\times\rho\mathbf{F} - \mathbf{x}_1\times\frac{\partial}{\partial\mathbf{x}_1}\cdot\left(\frac{1}{m}\int d\mathbf{p}_1(\mathbf{p}_1-\mathbf{p}_0)(\mathbf{p}_1-\mathbf{p}_0)g^{(1)} + \frac{2\sigma^3}{m}\int d\mathbf{p}\int d\mathbf{w}\int d\Omega\,\hat{\mathbf{r}}\hat{\mathbf{r}}(\mathbf{w}\cdot\hat{\mathbf{r}})^2\overline{\Gamma}\right), \tag{12.6}$$

and the energy balance equation is

$$\rho\left(\frac{\partial}{\partial t} + \mathbf{v}_0\cdot\frac{\partial}{\partial\mathbf{x}_1}\right)\epsilon = -\frac{\partial}{\partial\mathbf{x}_1}\cdot\left(\frac{1}{2m^2}\int d\mathbf{p}_1(\mathbf{p}_1-\mathbf{p}_0)|\mathbf{p}_1|^2 g^{(1)} + \frac{2\sigma^3}{m^2}\int d\mathbf{p}\int d\mathbf{w}\int d\Omega\,\hat{\mathbf{r}}(\mathbf{p}\cdot\hat{\mathbf{r}})(\mathbf{w}\cdot\hat{\mathbf{r}})^2\overline{\Gamma}\right). \tag{12.7}$$

From the above, and the discussion of hydrodynamics in Secs. 7–9, it is clear that the stress tensor $\mathbf{T}$ is given by

$$\mathbf{T} = -\frac{1}{m}\int d\mathbf{p}_1(\mathbf{p}_1-\mathbf{p}_0)(\mathbf{p}_1-\mathbf{p}_0)g^{(1)} - \frac{2\sigma^3}{m}\int d\mathbf{p}\int d\mathbf{w}\int d\Omega\ \hat{\mathbf{r}}\hat{\mathbf{r}}(\mathbf{w}\cdot\hat{\mathbf{r}})^2\overline{\Gamma} \tag{12.8}$$

for a system of rigid spheres. Inserting the equilibrium form of the distribution functions, one quickly obtains agreement with the virial theorem for a system in thermal and mechanical equilibrium.

The heat flux $\mathbf{q}$ is found as before by setting the term in large parentheses in Eq. (12.7) equal to $\mathbf{q} - \mathbf{T}\cdot\mathbf{v}_0$ to obtain

$$\mathbf{q} = \frac{1}{2m^2}\int d\mathbf{p}_1(\mathbf{p}_1-\mathbf{p}_0)(\mathbf{p}_1-\mathbf{p}_0)g^{(1)} + \frac{2\sigma^3}{m^2}\int d\mathbf{p}\int d\mathbf{w}\int d\Omega\ \hat{\mathbf{r}}(\mathbf{p}_1-\mathbf{p}_0)\cdot\hat{\mathbf{r}}(\mathbf{w}\cdot\hat{\mathbf{r}})^2\overline{\Gamma}. \tag{12.9}$$

Since a hard-sphere molecule possesses no intermolecular potential energy, the transport equation for the internal energy is simply

$$\rho\left(\frac{\partial}{\partial t}+\mathbf{v}_0\cdot\frac{\partial}{\partial\mathbf{x}_1}\right)\frac{3}{2m}\theta = \mathbf{T}:\frac{\partial\mathbf{v}_0}{\partial\mathbf{x}_1} - \frac{\partial}{\partial\mathbf{x}_1}\cdot\mathbf{q}. \tag{12.10}$$

13. SUMMARY OF RESULTS

We have shown that a single-component system of spherical molecules possessing either a continuous intermolecular potential $\phi(r)$, or a hardcore interaction of radius σ, obeys the general hydrodynamic equations of motion:

$$\frac{\partial\rho}{\partial t}+\frac{\partial}{\partial\mathbf{x}_1}\cdot(\rho\mathbf{v}_0)=0, \quad \text{mass transport,} \tag{13.1}$$

$$\rho\left(\frac{\partial}{\partial t}+\mathbf{v}_0\cdot\frac{\partial}{\partial\mathbf{x}_1}\right)\mathbf{v}_0 = \rho\mathbf{F}+\frac{\partial}{\partial\mathbf{x}_1}\cdot\mathbf{T}, \quad \text{momentum transport,} \tag{13.2}$$

$$\rho\left(\frac{\partial}{\partial t}+\mathbf{v}_0\cdot\frac{\partial}{\partial\mathbf{x}_1}\right)(\mathbf{x}_1\times\mathbf{v}_0) = \mathbf{x}_1\times\rho\mathbf{F}-\frac{\partial}{\partial\mathbf{x}_1}\cdot(\mathbf{T}\times\mathbf{x}_1), \quad \text{angular momentum transport,} \tag{13.3}$$

$$\rho\left(\frac{\partial}{\partial t}+\mathbf{v}_0\cdot\frac{\partial}{\partial\mathbf{x}_1}\right)\epsilon = \frac{\partial}{\partial\mathbf{x}_1}\cdot(\mathbf{T}\cdot\mathbf{v}_0-\mathbf{q}), \quad \text{energy transport,} \tag{13.4}$$

$$\rho\left(\frac{\partial}{\partial t}+\mathbf{v}_0\cdot\frac{\partial}{\partial\mathbf{x}_1}\right)\left(\frac{3}{2m}\theta+\frac{1}{m}u\right) = \mathbf{T}:\frac{\partial\mathbf{v}_0}{\partial\mathbf{x}_1}-\frac{\partial}{\partial\mathbf{x}_1}\cdot\mathbf{q}, \quad \text{internal energy transport,} \tag{13.5}$$

where m = mass of one molecule, ρ = mass density, $\mathbf{v}_0$ = macroscopic velocity $[=(1/m)\mathbf{p}_0]$, $\mathbf{F}$ = external force per unit mass, $\mathbf{T}$ = stress tensor, ϵ = energy per unit mass, $\mathbf{q}$ = heat flux, θ = kinetic temperature, u = average intermolecular potential energy of a molecule at $\mathbf{x}_1$.

For a system having a continuous intermolecular potential ϕ, we have found that the stress tensor and heat flux are given by

$$\mathbf{T} = -\frac{1}{m}\int d\mathbf{p}_1(\mathbf{p}_1-\mathbf{p}_0)(\mathbf{p}_1-\mathbf{p}_0)g^{(1)}+\tfrac{1}{2}\int d\mathbf{r}\,\hat{\mathbf{r}}\mathbf{r}\phi'\overline{\gamma} \tag{13.6}$$

and

$$\mathbf{q} = \frac{1}{2m^2}\int d\mathbf{p}_1(\mathbf{p}_1-\mathbf{p}_0)|\mathbf{p}_1-\mathbf{p}_0|^2g^{(1)} + \frac{1}{2m}\int d\mathbf{p}_1\int d\mathbf{p}_2\int d\mathbf{r}\,(\mathbf{p}_1-\mathbf{p}_0)\phi \times g^{(2)}(\mathbf{x}_1,\mathbf{p}_1;\mathbf{x}_1+\mathbf{r},\mathbf{p}_2) - \frac{1}{2m}\int d\mathbf{r}\,\hat{\mathbf{r}}\phi'\mathbf{r}\cdot\int d\mathbf{w}\int d\mathbf{p}\,(\mathbf{p}-\mathbf{p}_0)\overline{\Gamma}. \tag{13.7}$$

For a system consisting of rigid spheres of radius σ we have found that

$$\mathbf{T} = -\frac{1}{m}\int d\mathbf{p}_1\,(\mathbf{p}_1-\mathbf{p}_0)(\mathbf{p}_1-\mathbf{p}_0)g^{(1)} - \frac{2\sigma^3}{m}\int d\mathbf{p}\int d\mathbf{w}\int d\Omega\ \hat{\mathbf{r}}\hat{\mathbf{r}}(\mathbf{w}\cdot\hat{\mathbf{r}})^2\overline{\Gamma}(\mathbf{w},\mathbf{p};2\sigma\hat{\mathbf{r}},\mathbf{x}_1), \tag{13.8}$$

$$\mathbf{q} = \frac{1}{2m^2}\int d\mathbf{p}_1\,(\mathbf{p}_1-\mathbf{p}_0)|\mathbf{p}_1-\mathbf{p}_0|^2g^{(1)} + \frac{2\sigma^3}{m^2}\int d\mathbf{p}\int d\mathbf{w}\int d\Omega\ \hat{\mathbf{r}}(\mathbf{p}-\mathbf{p}_0)\cdot\hat{\mathbf{r}}(\mathbf{w}\cdot\hat{\mathbf{r}})^2 \times\overline{\Gamma}\ (\mathbf{w},\mathbf{p};2\sigma\hat{\mathbf{r}},\mathbf{x}_1). \tag{13.9}$$

In the last four equations we have used the notation

$$\overline{\gamma}(\mathbf{r},\mathbf{x}_1)=\int_{-1/2}^{1/2}d\eta\ \gamma(\mathbf{r},\mathbf{x}_1+\eta\mathbf{r}),$$
$$\overline{\Gamma}(\mathbf{w},\mathbf{p};\mathbf{r},\mathbf{x}_1)=\int_{-1/2}^{1/2}d\eta\ \Gamma(\mathbf{w},\mathbf{r};\mathbf{p},\mathbf{x}_1+\eta\mathbf{r}). \tag{13.10}$$

In obtaining these results we have assumed that the interactions between molecules may be described by classical statistical mechanics, that the external forces are conservative and depend only on position, and that the interactions of molecules far from the system boundary with molecules in the boundary material are negligible.

14. DISCUSSION

The stress tensor and the heat flux may each be considered as the sum of two parts which we label a local and a nonlocal contribution. The local part depends only on the distribution function $g^{(1)}$, whereas the nonlocal contribution depends on $f^{(2)}$ $(=\gamma)$, or $g^{(2)}$ $(=\Gamma)$. The local contribution is a familiar result of the kinetic theory of gases and is discussed in great detail elsewhere.[1] It is the nonlocal contribution that interests us here. We call any contribution depending on $f^{(2)}$ or $g^{(2)}$ nonlocal because it depends on what is happening to

the system not simply at $\mathbf{x}_1$, but also at points near $\mathbf{x}_1$. Thus, we expect the nonlocal contribution to be highly sensitive to the distribution of molecules in the substance making up the boundary wall (we have ignored this contribution here) when $\mathbf{x}_1$ is near the system boundary.

The functions $\bar{\gamma}$ and $\bar{\Gamma}$ play an important role in our theory. They appear in the stress tensor and heat flux in the terms due to the force exerted and work done, by molecules outside of the region R on the molecules inside R. They are obtained by averaging γ and Γ in their center-of-mass dependence over the straight line connecting $\mathbf{x}_1 - \frac{1}{2}\mathbf{r}$ to $\mathbf{x}_1 + \frac{1}{2}\mathbf{r}$. To understand the importance of this averaging process, consider the case when the system is perturbed by a plane wave of such short wavelength that any significant variation of ϕ or ϕ' occurs over many wavelengths. In the limit of such short-wavelength phenomena, it can be expected that the dependence of $\bar{\gamma}$ and $\bar{\Gamma}$ on $\mathbf{r}$ will become insensitive to the disturbance. We shall see that this is so for the elementary example considered in the next section.

In order to understand the difficulties encountered with the Taylor expansion obtained by Irving and Kirkwood, we shall now examine the trace of the nonlocal part of the stress tensor, $\triangle p\mathbf{T}_{\text{int}}$, for a system with a continuous intermolecular potential. From Eq. (13.6) we find

$$\triangle p\mathbf{T}_{\text{int}} = \tfrac{1}{2}\int d\mathbf{r}\, r\phi'\bar{\gamma} = \tfrac{1}{2}\int d\Omega \int_0^\infty dr\, r^3\phi'\bar{\gamma}. \quad (14.1)$$

We quickly obtain their result by expanding $\bar{\gamma}$ about $\gamma(\mathbf{r}, \mathbf{x}_1)$:

$$\begin{aligned}\bar{\gamma}(\mathbf{r}, \mathbf{x}_1) &= \int_{-1/2}^{1/2} d\eta\, \gamma(\mathbf{r}, \mathbf{x}_1 + \eta\mathbf{r}) \\ &= \int_{-1/2}^{1/2} d\eta \sum_0^\infty \frac{1}{n!}\left(\eta\mathbf{r}\cdot\frac{\partial}{\partial\mathbf{x}_1}\right)^n \gamma(\mathbf{r}, \mathbf{x}_1) \\ &= \sum_0^\infty \frac{1}{n!} r^n \left(\int_{-1/2}^{1/2} d\eta\, \eta^n\right)\left(\hat{\mathbf{r}}\cdot\frac{\partial}{\partial\mathbf{x}_1}\right)^n \gamma(\mathbf{r}, \mathbf{x}_1) \\ &= \sum_0^\infty \frac{1}{(2n+1)!} r^{2n}\left(\tfrac{1}{2}\hat{\mathbf{r}}\cdot\frac{\partial}{\partial\mathbf{x}_1}\right)^{2n}\gamma(\mathbf{r}, \mathbf{x}_1), \end{aligned} \quad (14.2)$$

giving us

$$\triangle p\mathbf{T}_{\text{int}} = \tfrac{1}{2}\sum_0^\infty \frac{1}{(2n+1)!}\int d\Omega\left(\tfrac{1}{2}\hat{\mathbf{r}}\cdot\frac{\partial}{\partial\mathbf{x}_1}\right)^{2n}\int_0^\infty dr \times r^{2n+3}\phi'(r)\gamma(\mathbf{r}, \mathbf{x}_1).$$

Let us examine the integrand of

$$\int_0^\infty dr\, r^{2n+3}\phi'(r)\gamma(\mathbf{r}, \mathbf{x}_1)$$

in the limit of large r. As γ becomes the product of the number densities,

$$\lim_{r\to\infty}\gamma(\mathbf{r}, \mathbf{x}_1) = f^{(1)}(\mathbf{x}_1 - \tfrac{1}{2}\mathbf{r})f^{(1)}(\mathbf{x}_1 + \tfrac{1}{2}\mathbf{r}), \quad (14.3)$$

it does not vanish for large r. Thus, either ϕ' goes to zero faster than $(1/r)^{2n+3}$ or the integral blows up and the nth term in the expansion is not defined. Hence, for a square well potential or a Yukawa potential the Irving and Kirkwood expansion produces meaningful results, but for a Leonard–Jones potential the expansion is ambiguous. The reader may easily verify that the same difficulties are encountered with the heat flux. It is a straightforward matter, however, to avoid this pitfall. If one wants to determine T, one must know, or assume some form for, $\gamma(\mathbf{r}, \mathbf{x})$. But, if one knows γ, it is just as easy to integrate over η to obtain $\bar{\gamma}$ and then perform the integration over $\mathbf{r}$ as to perform an expansion under the integral. Any long-wavelength expansion can then be made on the expression for T obtained in closed form.

We gain further insight into the Irving and Kirkwood expansion by contructing a Fourier decomposition of $\gamma(\mathbf{r}, \mathbf{x})$ in $\mathbf{x}$:

$$\gamma(\mathbf{r}, \mathbf{x}) = \int d\mathbf{k}\, e^{i\mathbf{k}\cdot\mathbf{x}}\mathcal{G}(\mathbf{r}, \mathbf{k}). \quad (14.4)$$

By inserting Eq. (14.4) into Eqs. (13.10) and (14.4), we quickly find

$$\bar{\gamma}(\mathbf{r}, \mathbf{x}_1) = \int d\mathbf{k}\, e^{i\mathbf{k}\cdot\mathbf{x}_1}\mathcal{G}(\mathbf{r}, \mathbf{k})\frac{\sin\frac{1}{2}\mathbf{k}\cdot\mathbf{r}}{\frac{1}{2}\mathbf{k}\cdot\mathbf{r}} \quad (14.5)$$

and

$$\triangle p\mathbf{T}_{\text{int}} = \tfrac{1}{2}\int d\mathbf{k}\, e^{i\mathbf{k}\cdot\mathbf{x}_1}\int d\mathbf{r}\, r\phi'\mathcal{G}(\mathbf{r},\mathbf{k})\frac{\sin\frac{1}{2}\mathbf{k}\cdot\mathbf{r}}{\frac{1}{2}\mathbf{k}\cdot\mathbf{r}}. \quad (14.6)$$

The reader may easily verify, by applying the power series of Eq. (14.2) to

$$\gamma(\mathbf{r}, \mathbf{x}_1) = \int d\mathbf{k}\, e^{i\mathbf{k}\cdot\mathbf{x}_1}\mathcal{G}(\mathbf{r}, \mathbf{k}), \quad (14.7)$$

that the Irving and Kirkwood expansion consists of expanding $\sin\frac{1}{2}\mathbf{k}\cdot\mathbf{r}/\frac{1}{2}\mathbf{k}\cdot\mathbf{r}$ in powers of $(\frac{1}{2}\mathbf{k}\cdot\mathbf{r})$ in the integrand of Eq. (14.6):

$$\triangle p\mathbf{T}_{\text{int}} = \tfrac{1}{2}\sum_0^\infty \frac{(-1)^n}{(2n+1)!}\int d\mathbf{k}\, e^{i\mathbf{k}\cdot\mathbf{x}_1} \times \int d\mathbf{r}\, r\phi'\mathcal{G}(\mathbf{r}, \mathbf{k})(\tfrac{1}{2}\mathbf{k}\cdot\mathbf{r})^{2n}.$$

While the lack of convergence of the series is immediately apparent for a realistic potential, we shall find yet another shortcoming of the expansion, even for well-behaved potential functions, in the next section. Although the first term of the series provides an excellent approximation for long-wavelength phenomena, the remaining terms may not provide useful corrections to this limit.

15. AN ELEMENTARY EXAMPLE

Suppose our system is a fluid which is slightly perturbed by a density wave of angular frequency ω and wave vector $\mathbf{k}$. We shall assume that the amplitude of the wave is small and that it may be considered a constant over distances of many

wavelengths. For such a disturbance the density is given by

$$f^{(1)}(\mathbf{x}_1) = n[1 + \epsilon \cos(\mathbf{k}_0 \cdot \mathbf{x}_1 - \omega t], \tag{15.1}$$

where n is the number density in equilibrium. We shall also assume that in the perturbed state the two-particle distribution function is of the form

$$f^{(2)}(\mathbf{x}_1, \mathbf{x}_2) = f^{(1)}(\mathbf{x}_1) f^{(1)}(\mathbf{x}_2) h(r), \tag{15.2}$$

where we require

$$\lim_{r \to \infty} h(r) = 1 \tag{15.3}$$

in order that $f^{(2)}$ have the proper asymptotic behavior.

With these conditions we find

$$\gamma(\mathbf{r}, \mathbf{x}) = n^2 h(r)[1 + 2\epsilon \cos(\mathbf{k}_0 \cdot \mathbf{x} - \omega t) \cos \tfrac{1}{2}\mathbf{k}_0 \cdot \mathbf{r}] + \text{order } \epsilon^2, \tag{15.4}$$

where we shall ignore all terms in ϵ^2, so that

$$\bar{\gamma}(\mathbf{r}, \mathbf{x}_1) = \int_{-1/2}^{1/2} d\eta \gamma(\mathbf{r}, \mathbf{x}_1 + \eta \mathbf{r}) = n^2 h(r)\,[1 + 2\epsilon \cos(\mathbf{k}_0 \cdot \mathbf{x}_1 - \omega t) \sin k_0 \cdot \mathbf{r}/k_0 \cdot \mathbf{r}]. \tag{15.5}$$

The intermolecular contribution to the stress tensor for a system with a continuous potential is, from Eq. (13.6),

$$\mathbf{T}_{\text{int}} = \tfrac{1}{2} \int d\mathbf{r}\, \hat{\mathbf{r}}\mathbf{r} \phi'(r) \bar{\gamma}(\mathbf{r}, \mathbf{x}_1) \tag{15.6}$$

so that, for our example,

$$\mathbf{T}_{\text{int}} = \tfrac{1}{2} n^2 \int d\mathbf{r}\, \hat{\mathbf{r}}\mathbf{r} \phi'(r) h(r) + \epsilon \cos(\mathbf{k}_0 \cdot \mathbf{x}_1 - \omega t) n^2 \int d\mathbf{r}\, \hat{\mathbf{r}}\mathbf{r} \phi'(r) h(r) \times \sin k_0 \cdot \mathbf{r}/k_0 \cdot \mathbf{r}. \tag{15.7}$$

In order to determine the Irving and Kirkwood expansion for this example, we construct the Fourier transform of $\gamma(\mathbf{r}, \mathbf{x})$ [see Eqs. (14.4)–(14.8)]:

$$\mathcal{G}(\mathbf{r}, \mathbf{k}) = n^2 h\{\delta(\mathbf{k}) + \epsilon \cos \tfrac{1}{2}\mathbf{k}_0 \cdot \mathbf{r}[e^{-i\omega t} \delta(\mathbf{k} - \mathbf{k}_0) + e^{i\omega t} \delta(\mathbf{k} + \mathbf{k}_0)]\},$$

where the δ's are three-dimensional Dirac δ functions in $\mathbf{k}$ space. Inserting this into Eq. (14.5) for $\bar{\gamma}$, and the resulting expression into Eq. (15.6) we find that the Irving and Kirkwood expansion does not result in an expansion of the coefficient of $\cos(\mathbf{k}_0 \cdot \mathbf{x}_1 - \omega t)$ in Eq. (15.7)! Writing

$$\frac{\sin k_0 \cdot \mathbf{r}}{k_0 \cdot \mathbf{r}} = \left(\frac{\sin \tfrac{1}{2}\mathbf{k}_0 \cdot \mathbf{r}}{\tfrac{1}{2}\mathbf{k}_0 \cdot \mathbf{r}}\right) \cos \tfrac{1}{2}\mathbf{k}_0 \cdot \mathbf{r},$$

we find that the expansion consists of expressing $\sin \tfrac{1}{2}\mathbf{k}_0 \cdot \mathbf{r} / \tfrac{1}{2}\mathbf{k}_0 \cdot \mathbf{r}$ as a power series in $(\tfrac{1}{2}\mathbf{k}_0 \cdot \mathbf{r})$ without expressing $\cos \tfrac{1}{2}\mathbf{k}_0 \cdot \mathbf{r}$ in similar fashion. This is clearly a shortcoming of the theory as every term in the expansion of the coefficient of $\cos(\mathbf{k}_0 \cdot \mathbf{x}_1 - \omega t)$ contains all the even powers of $\tfrac{1}{2}\mathbf{k}_0 \cdot \mathbf{r}$. We still encounter the previous difficulty in that the contributions of most of the terms are not defined unless ϕ' goes to zero in the limit of large r faster than any positive power of $(1/r)$. The *correct* long-wavelength limit is found by setting $\sin k_0 \cdot \mathbf{r}/k_0 \cdot \mathbf{r} = 1$ in Eq. (15.7), giving us

$$(\mathbf{T}_{\text{int}})_{lw} = [(1 + 2\epsilon \cos(\mathbf{k}_0 \cdot \mathbf{x}_1 - \omega t)) \tfrac{2\pi}{3} n^2 \times \int_0^\infty dr\, r^3 \phi' h]\mathbf{I}, \tag{15.8}$$

where $\mathbf{I}$ is the unit tensor. Hence, in the long-wavelength limit $\mathbf{T}_{\text{int}}$ is spherical and there are no shears present unless the kinetic contribution is nonspherical.

In order to determine $\mathbf{T}_{\text{int}}$ in the general case we shall assume that $\mathbf{k}_0$ is in the $\hat{\mathbf{e}}_3$ direction. Performing the integration over angles in Eq. (15.7) we find

$$(\mathbf{T}_{\text{int}})_{3,3} = \tfrac{2\pi}{3} n^2 \int_0^\infty dr\, r^3 \phi' h + 4\pi\epsilon \cos(\mathbf{k}_0 \cdot \mathbf{x}_1 - \omega t) n^2 \int_0^\infty dr\, r^3 \phi' h \left(\frac{1}{k_0 r}\right)^3 \times (\sin k_0 r - k_0 r \cos k_0 r) \tag{15.9}$$

and

$$(\mathbf{T}_{\text{int}})_{1,1} = (\mathbf{T}_{\text{int}})_{2,2} = \tfrac{2\pi}{3} n^2 \int_0^\infty dr\, r^3 \phi' h + 2\pi\epsilon \cos(\mathbf{k}_0 \cdot \mathbf{x}_1 - \omega t) n^2 \int_0^\infty dr\, r^3 \phi' h \times \left[\frac{1}{k_0 r} \operatorname{Si}(k_0 r) - \left(\frac{1}{k_0 r}\right)^3 (\sin k_0 r - k_0 r \cos k_0 r\right], \tag{15.10}$$

where,[11]

$$\operatorname{Si}(\chi) = \int_0^\chi dy\, \frac{\sin y}{y}. \tag{15.11}$$

All of the off-diagonal elements vanish, but the stress quadric is now an ellipsoid of revolution, so that the fluid is undergoing shear.

It is interesting to note that, in the short-wavelength limit $(k_0 \to \infty)$, $\operatorname{Si}(k_0 r) \to \tfrac{1}{2}\pi$, so that

$$(\mathbf{T}_{\text{int}})_{3,3} \to \tfrac{2\pi}{3} n^2 \int_0^\infty dr\, r^3 \phi' h + \text{order} \left(\frac{1}{k_0}\right)^2 \tag{15.12}$$

and

$$(\mathbf{T}_{\text{int}})_{1,1} = (\mathbf{T}_{\text{int}})_{2,2} \to \tfrac{2\pi}{3} n^2 \int_0^\infty dr\, r^3 \phi' h + \epsilon \frac{\pi^2}{k_0} \cos(\mathbf{k}_0 \cdot \mathbf{x}_1 - \omega t) n^2 \int_0^\infty dr\, r^2 \phi' h + \text{order} \left(\frac{1}{k_0}\right)^2. \tag{15.13}$$

Hence, in the limit of short wavelength the stress tensor returns to its spherical form. Moreover, if in this limit $h(r)$ approaches the equilibrium radial distribution function, then $\mathbf{T}_{\text{int}}$ is unaffected by the disturbance and the perturbed part of $\mathbf{T}_{\text{int}}$

in the $\mathbf{k}_0$ "direction" vanishes more quickly than the perturbed part of $\mathbf{T}_{int}$ in the $\hat{e}_1$ and $\hat{e}_2$ "directions."

ACKNOWLEDGMENTS

We are indebted to Dr. Z. Salsburg for his encouragement during the course of this investigation and for his valuable comments on the preparation of the manuscript for publication. We are also grateful for some stimulating discussions with G. Bettoney. This work was supported by the National Science Foundation, under Grant No. GP-6447 and the Robert A. Welch Foundation.

* Present address: Physics Dept., Texas Christian University, Fort Worth, Texas 76129.
1 S. Chapman and T. Cowling, *The Mathematical Theory of Non-uniform Gases* (Cambridge U. P., Cambridge, 1939).
2 J. H. Irving and J. G. Kirkwood, J. Chem. Phys. **18**, 817 (1950).
3 F. M. Snell, R. Aranow, and R. A. Spangler, J. Chem. Phys. **47**, 4959 (1967).
4 M. S. Green, J. Chem. Phys. **22**, 398 (1954).
5 H. Mori, Phys. Rev. **112**, 1829 (1958).
6 H. Frölich, Physica **37**, 215 (1967).
7 J. H. Irving and R. W. Zwanzig, J. Chem. Phys. **19**, 1173 (1951).
8 S. T. Choh, "The Kinetic Theory of Phenomena in Dense Gases" (Thesis), University of Michigan, 1958 (unpublished).
9 See N. N. Bogolubov in *Studies in Statistical Mechanics*, edited by J. de Boer and G. E. Uhlenbeck (North-Holland, Amsterdam, 1962), Vol. I.
10 D. Enskog, Svensk. Akad., Arkiv. f. Math., Ast. Och. Fys. **21A**, no.13 (1928).
11 M. Abramowitz and I. Stegan, *Handbook of Mathematical Functions* (Nat. Bur. Stds. Applied Mathematics Series, U.S. GPO, Washington, D.C., 1964), p. 231.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 10 OCTOBER 1971

Multiple Scattering Theory of Radiative Transfer in the Light of Singular Modes*

Madhoo Kanal
Department of Physics and Astronomy, University of Massachusetts, Amherst, Massachusetts 01002
(Received 27 April 1971)

The multiple scattering theory of radiative transfer, in the light of Case's normal modes, is investigated. In particular, it is shown that in the Neumann expansion of the total intensity the expansion coefficients are related to the normal modes in the form of certain integrals over the continuum Case spectrum. In the general formulation of the theory, the phase function is kept arbitrary. Two types of boundary-value problems are considered which involve semi-infinite and bounded media. To illustrate the structure of the general formulation, the isotropic case is treated in detail.

INTRODUCTION

In boundary-value problems of radiative transfer one often seeks solutions of the transfer equation expressed as a powers series in the single scattering albedo ω.[1-3] The coefficients in such a power series, the so-called "Neumann coefficients," are associated with the physical process of multiple scattering of the radiation traversing some given medium (c.f. Van de Hulst[4]). The radiation at any given point in the medium is decomposed into components according to whether it has suffered no scattering, or has been scattered once, twice, or n times. At every stage of this process it is required that the equation of transfer be satisfied such that the scattering of the intensity in the nth process will give rise to a source function of the $(n+1)$ times scattered radiation. The nth coefficient in the Neumann series then represents the nth order of scattering.

In this paper we seek the relation of such an elementary process to Case's singular normal modes.[5] In other words, the relation of Neumann coefficients to the elementary solutions of the transfer equation is investigated. The general formulation is presented for an arbitrary phase function. The two types of boundary-value problems we consider involve semi-infinite and bounded media. To illustrate the detailed structure of solutions, we treat the isotropic case in detail. However, the whole procedure is kept sufficiently general so that, for a gray atmosphere, the extension to other geometries remains straightforward. The basic tool we use is Case's Green's function technique[6] and its further developments are discussed in Refs. 7 and 8

1. GENERAL FORMULATION FOR MULTIPLE SCATTERING

The equation of radiative transfer we consider is

$$\mu \frac{\partial}{\partial \tau} I(\tau, \Omega) + I(\tau, \Omega) = \frac{\omega}{4\pi} \int d\Omega' P(\Omega \cdot \Omega') I(\tau, \Omega'), \tag{1.1}$$

where we have employed the usual notation. Thus, τ is the optical depth, Ω is the unit vector pointing into the direction of propagation of radiation. If $\hat{n}$ represents the unit normal at $\tau = 0$ pointing into the volume of interest, then $\mu = \hat{n} \cdot \Omega$. Finally ω is the albedo for single scattering, i.e., $\omega = \sigma(\sigma + K)^{-1}$, where σ and K are the coefficients of scattering and extinction, respectively, $P(\Omega \cdot \Omega')$ is a rotationally invariant phase function, and $I(\tau, \Omega)$ is the total intensity. Since the frequency occurs as a parameter in Eq. (1.1), this dependence is not explicitly indicated. In general, of course, I and ω will be functions of the frequency.

Let $I_n(\tau, \Omega)$ represent the intensity of the n-times

in the $\mathbf{k}_0$ "direction" vanishes more quickly than the perturbed part of $\mathbf{T}_{int}$ in the $\hat{e}_1$ and $\hat{e}_2$ "directions."

ACKNOWLEDGMENTS

We are indebted to Dr. Z. Salsburg for his encouragement during the course of this investigation and for his valuable comments on the preparation of the manuscript for publication. We are also grateful for some stimulating discussions with G. Bettoney. This work was supported by the National Science Foundation, under Grant No. GP-6447 and the Robert A. Welch Foundation.

* Present address: Physics Dept., Texas Christian University, Fort Worth, Texas 76129.
1 S. Chapman and T. Cowling, *The Mathematical Theory of Non-uniform Gases* (Cambridge U. P., Cambridge, 1939).
2 J. H. Irving and J. G. Kirkwood, J. Chem. Phys. **18**, 817 (1950).
3 F. M. Snell, R. Aranow, and R. A. Spangler, J. Chem. Phys. **47**, 4959 (1967).
4 M. S. Green, J. Chem. Phys. **22**, 398 (1954).
5 H. Mori, Phys. Rev. **112**, 1829 (1958).
6 H. Frölich, Physica **37**, 215 (1967).
7 J. H. Irving and R. W. Zwanzig, J. Chem. Phys. **19**, 1173 (1951).
8 S. T. Choh, "The Kinetic Theory of Phenomena in Dense Gases" (Thesis), University of Michigan, 1958 (unpublished).
9 See N. N. Bogolubov in *Studies in Statistical Mechanics*, edited by J. de Boer and G. E. Uhlenbeck (North-Holland, Amsterdam, 1962), Vol. I.
10 D. Enskog, Svensk. Akad., Arkiv. f. Math., Ast. Och. Fys. **21A**, no.13 (1928).
11 M. Abramowitz and I. Stegan, *Handbook of Mathematical Functions* (Nat. Bur. Stds. Applied Mathematics Series, U.S. GPO, Washington, D.C., 1964), p. 231.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 10 OCTOBER 1971

Multiple Scattering Theory of Radiative Transfer in the Light of Singular Modes*

Madhoo Kanal
Department of Physics and Astronomy, University of Massachusetts, Amherst, Massachusetts 01002
(Received 27 April 1971)

The multiple scattering theory of radiative transfer, in the light of Case's normal modes, is investigated. In particular, it is shown that in the Neumann expansion of the total intensity the expansion coefficients are related to the normal modes in the form of certain integrals over the continuum Case spectrum. In the general formulation of the theory, the phase function is kept arbitrary. Two types of boundary-value problems are considered which involve semi-infinite and bounded media. To illustrate the structure of the general formulation, the isotropic case is treated in detail.

INTRODUCTION

In boundary-value problems of radiative transfer one often seeks solutions of the transfer equation expressed as a powers series in the single scattering albedo ω.[1-3] The coefficients in such a power series, the so-called "Neumann coefficients," are associated with the physical process of multiple scattering of the radiation traversing some given medium (c.f. Van de Hulst[4]). The radiation at any given point in the medium is decomposed into components according to whether it has suffered no scattering, or has been scattered once, twice, or n times. At every stage of this process it is required that the equation of transfer be satisfied such that the scattering of the intensity in the nth process will give rise to a source function of the $(n+1)$ times scattered radiation. The nth coefficient in the Neumann series then represents the nth order of scattering.

In this paper we seek the relation of such an elementary process to Case's singular normal modes.[5] In other words, the relation of Neumann coefficients to the elementary solutions of the transfer equation is investigated. The general formulation is presented for an arbitrary phase function. The two types of boundary-value problems we consider involve semi-infinite and bounded media. To illustrate the detailed structure of solutions, we treat the isotropic case in detail. However, the whole procedure is kept sufficiently general so that, for a gray atmosphere, the extension to other geometries remains straightforward. The basic tool we use is Case's Green's function technique[6] and its further developments are discussed in Refs. 7 and 8

1. GENERAL FORMULATION FOR MULTIPLE SCATTERING

The equation of radiative transfer we consider is

$$\mu \frac{\partial}{\partial \tau} I(\tau, \Omega) + I(\tau, \Omega) = \frac{\omega}{4\pi} \int d\Omega' P(\Omega \cdot \Omega') I(\tau, \Omega'), \tag{1.1}$$

where we have employed the usual notation. Thus, τ is the optical depth, Ω is the unit vector pointing into the direction of propagation of radiation. If $\hat{n}$ represents the unit normal at $\tau = 0$ pointing into the volume of interest, then $\mu = \hat{n} \cdot \Omega$. Finally ω is the albedo for single scattering, i.e., $\omega = \sigma(\sigma + K)^{-1}$, where σ and K are the coefficients of scattering and extinction, respectively, $P(\Omega \cdot \Omega')$ is a rotationally invariant phase function, and $I(\tau, \Omega)$ is the total intensity. Since the frequency occurs as a parameter in Eq. (1.1), this dependence is not explicitly indicated. In general, of course, I and ω will be functions of the frequency.

Let $I_n(\tau, \Omega)$ represent the intensity of the n-times

scattered radiation. Then for the nth-order scattering the equation of transfer to be satisfied is

$$\mu \frac{\partial}{\partial \tau} I_n(\tau, \Omega) + I_n(\tau, \Omega) = \frac{\omega}{4\pi} \int d\Omega' P(\Omega \cdot \Omega') I_{n-1}(\tau, \Omega'). \tag{1.2}$$

One may readily check upon summing both sides of Eq. (1.2) with respect to n from 0 to ∞, that the total intensity

$$I(\tau, \Omega) = \sum_{n=0}^{\infty} I_n(\tau, \Omega) \tag{1.3}$$

is a solution of Eq. (2.1), as it should be.

It is convenient to separate out the azimuth angle dependence in Eq. (1.2). This we do as follows. Writing

$$P(\Omega \cdot \Omega') = \sum_{m=-\infty}^{\infty} e^{im(\phi' - \phi)} S_m(\mu, \mu'), \tag{1.4}$$

and

$$I_n^{(m)}(\tau, \mu) = \frac{1}{2\pi} \int_0^{2\pi} d\phi e^{im\phi} I_n(\tau, \Omega), \tag{1.5}$$

we see that Eq. (1.2) then becomes

$$\mu \frac{\partial}{\partial \tau} I_n^{(m)}(\tau, \mu) + I_n^{(m)}(\tau, \mu) = \frac{\omega}{2} \int_{-1}^{1} d\mu' S_m(\mu, \mu') I_{n-1}^{(m)}(\tau, \mu'). \tag{1.6}$$

Now, since $e^{im\phi}$ are complete, we have

$$I_n(\tau, \Omega) = \sum_{m=-\infty}^{\infty} e^{im\phi} I_n^{(m)}(\tau, \mu). \tag{1.7}$$

Integral Representation of $I_n^{(m)}$ (Half-Space)

Following the procedure discussed in Ref. 8, we first consider the time-reversed Green's function for a purely absorbing medium, i.e.,

$$-\mu \frac{\partial}{\partial \tau} G(\tau, -\mu; \tau_0, -\mu_0) + G(\tau, -\mu; \tau_0, -\mu_0) = \delta(\tau - \tau_0)\delta(\mu - \mu_0). \tag{1.8}$$

Noting that G satisfies the reciprocity relation

$$G(\tau, -\mu; \tau_0, -\mu_0) = G(\tau_0, \mu_0; \tau, \mu), \tag{1.9}$$

we combine Eqs. (1.6) and (1.8) in the conventional way and obtain

$$I_n^{(m)}(\tau, \mu)\Theta(\tau) = \mu I_n^{(m)}(0, \mu) \frac{1}{2\pi} \int_{-\infty}^{\infty} dk \frac{e^{ik\tau}}{1 + ik\mu} + \frac{\omega}{4\pi} \int_{-\infty}^{\infty} dk \frac{e^{ik\tau}}{1 + ik\mu} R_{n-1}^{(m)}(k; \mu), \tag{1.10}$$

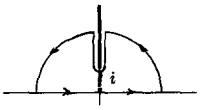


Fig. 1. Contour in the complex K plane.

where the Fourier coefficients $R_n^{(m)}$ are defined as [see Ref. 9]

$$R_n^{(m)}(k; \mu) = \int_0^{\infty} d\tau e^{-ik\tau} \int_{-1}^{1} d\mu' S_m(\mu, \mu') I_n^{(m)}(\tau, \mu'), \tag{1.11}$$

and $\Theta(\tau)$ is the Heaviside step function. This symbol (Θ) will be reserved throughout for the step function.

We note that once the coefficients $R_{n-1}^{(m)}$ are determined, then $I_n^{(m)}$ are known everywhere. In the usual fashion, an equation which determines these coefficients is obtained by letting $\tau \to 0$ in Eq. (1.10). Before such a limit is considered we shall first examine certain singular properties of $R_{n-1}^{(m)}(k; \mu)$ in the complex k plane. Then re-expressing the right-hand side of Eq. (1.10) over the Case spectrum,[7] we will utilize those singular properties and obtain certain modified coefficients which are relatively simple to evaluate. We will find that the modified coefficients are directly related to Case's normal modes. Let us remark here that Neumann coefficients N_n are defined simply as

$$I(\tau, \Omega) = \sum_{n=0}^{\infty} \omega^n N_n(\tau, \Omega), \tag{1.12}$$

where N_n are independent of ω and correspond to the multiple scattering process in a purely scattering medium ($\omega = 1$).

Singular Properties of $R_n^{(m)}$

Pursuing the procedure of Ref. 8, we take the Fourier transform of Eq. (1.10) with respect to τ, multiply both sides by $S_m(\mu, \mu')$ and integrate over μ, to obtain

$$R_n^{(m)}(z; \mu) = J_n^{(m)}(z; \mu) + \frac{\omega z}{2} \int_{-1}^{1} d\mu' \frac{S_m(\mu, \mu')}{z - \mu'} \times R_{n-1}^{(m)}(z; \mu'), \tag{1.13}$$

where for convenience we have set $k = i/z$ and defined

$$J_n^{(m)}(z; \mu) = z \int_{-1}^{1} d\mu' \mu' I_n^{(m)}(0, \mu') \frac{S_m(\mu', \mu)}{z - \mu'} \tag{1.14}$$

and used the symmetry of $S_m(\mu, \mu')$, i.e.,

$$S_m(\mu, \mu') = S_m(\mu', \mu). \tag{1.15}$$

From Eq. (1.13) it is clear that $R_n^{(m)}(z; \mu)$ as functions of z are sectionally holomorphic in the complex z plane cut from -1 to 1. If $R_n^{(m)} \pm (\nu; \mu)$ represent the boundary values of $R_n^{(m)}(z; \mu)$ as z approaches the cut (from $\begin{matrix}\text{top}\\ \text{bottom}\end{matrix}$), then the sums and differences of these boundary values are re-

lated in the manner as shown below,

$$R_{n-1}^{(m)+}(\mu;\mu) + R_{n-1}^{(m)-}(\mu;\mu) = -\frac{4\mu}{\omega} I_n^{(m)}(0,\mu) - \frac{2}{i\pi\omega\mu}\frac{R_n^{(m)+}(\mu;\nu) - R_n^{(m)-}(\mu;\nu)}{S_m(\mu,\nu)} + \frac{1}{i\pi}\mathcal{P}\int_{-1}^{1}\frac{d\mu'}{\mu-\mu'}\frac{S_m(\mu',\nu)}{S_m(\mu,\nu)}(R_{n-1}^{(m)+}(\mu;\mu') - R_{n-1}^{(m)-}(\mu;\mu')), \quad (1.16)$$

where $\mathcal{P}$ represents the principle value and we have used Plemelj's formula,

$$\frac{1}{(\nu-\mu)_\pm} = \mathcal{P}\frac{1}{\nu-\mu} \mp i\pi\delta(\nu-\mu). \quad (1.17)$$

It is interesting to note that the left-hand side of Eq. (1.16) is independent of ν. The reason is simply that by virtue of Plemelj's formula (Eq. 1.17) the term associated with the sum of the boundary values of $R_{n-1}^{(m)}(z;\mu)$, on the right-hand side of Eq. (1.13), is a delta function. This "factorization" property of Eq. (1.13) turns out to be very useful in obtaining the modified coefficients which are directly related to continuum normal modes, as we shall see presently.

Keeping the recurrence relation (1.13) in mind, we note that other than the branch cut in the complex z-plane extending from -1 to 1, $R_n^{(m)}(z;\mu)$ have no other singularities. This implies that in the spectral representation of $I_n^{(m)}(\tau,\mu)$, the discrete modes, which correspond to the discrete spectrum of the transport operator of Eq. (1.1) [the integro-differential operator operating on the total intensity $I(\tau,\Omega)$], will not occur explicitly. This is really not surprising, for the integro-differential operator of Eq. (1.6) has only the continuous spectrum for any fixed value of n. However, in the infinite sum over $I_n^{(m)}(\tau,\mu)$, the discrete modes will be generated from the dynamics of the recurrence relation (1.13). To illustrate this point we shall later examine the isotropic case in detail.

Spectral Representation of $I_n^{(m)}$

Now consider Eq. (1.10) and a contour around the branch cut of $R_{n-1}^{(m)}(k;\mu)$ in the upper half complex k plane as shown in the figure. Since $\tau > 0$ and $R_{n-1}^{(m)}(k;\mu)$ in the half-space are functions of k (and μ) only, the integrals over the semicircle do not contribute. Consequently,

$$I_n^{(m)}(\tau,\mu)\Theta(\tau) = I_n^{(m)}(0,\mu)e^{-\tau/\mu}\Theta(\mu) + \frac{\omega i}{4\pi}\mathcal{P}\int_0^1\frac{d\nu}{\nu^2}e^{-\tau/\nu}\frac{\nu}{\nu-\mu} \times (R_{n-1}^{(m)+}(\nu;\mu) - R_{n-1}^{(m)-}(\nu;\mu)) + \frac{\omega}{4\mu}e^{-\tau/\mu}\Theta(\mu) \times (R_{n-1}^{(m)+}(\mu;\mu) + R_{n-1}^{(m)-}(\mu;\mu), \quad (1.18)$$

where $+(-)$ on $R_{n-1}^{(m)}(z;\mu)$, as before, represent the boundary values as z approaches the cut in the z-plane ($z = i/k$) from the top (bottom). The last term on the right-hand side, involving the sum of boundary values, may be eliminated by means of Eq. (1.9). The result, after change of variables and rearrangement of terms, is

$$I_n^{(m)}(\tau,\mu)\Theta(\tau)S_m(\mu,\nu) = \Gamma_n^{(m)}(\mu;\nu)e^{-\tau/\mu}\Theta(\mu) + \omega S_m(\mu,\nu) \times \int_0^1 d\nu'\phi^{\nu'}(\mu)e^{-\tau/\nu'}\Gamma_{n-1}^{(m)}(\nu';\mu) - \omega e^{-\tau/\mu}\Theta(\mu) \times \int_{-1}^1 d\nu'\phi^{\mu}(\nu')S_m(\nu',\nu)\Gamma_{n-1}^{(m)}(\mu;\nu'), \quad (1.19)$$

where we have introduced the modified coefficients $\Gamma_n^{(m)}$ and distributions $\phi^\nu(\mu)$ which are defined by

$$\Gamma_n^{(m)}(\mu;\nu) = -1/2\pi i\mu^2[R_n^{(m)+}(\mu;\nu) - R_n^{(m)-}(\mu;\nu)] \quad (1.20)$$

and

$$\phi^\nu(\mu) = \frac{\nu}{2}\mathcal{P}\frac{1}{\nu-\mu} + \frac{\Lambda^+(\nu)+\Lambda^-(\nu)}{2}\delta(\nu-\mu), \quad (1.21)$$

with

$$\Lambda(z) = 1 - \frac{z}{2}\int_{-1}^1\frac{d\mu}{z-\mu}. \quad (1.22)$$

One may readily check that $\phi^\nu(\mu)$ are Case's continuum normal modes[5] of the equation of transfer for a conservative ($\omega = 1$) isotropically scattering medium. For the sake of convenience, we introduce another function defined by

$$\Psi_m^{\nu'}(\mu;\nu) = \phi^{\nu'}(\mu)S_m(\mu,\nu), \quad (1.23)$$

and rewrite Eq. (1.19) as

$$I_n^{(m)}(\tau,\mu)\Theta(\tau)S_m(\mu,\nu) = \Gamma_n^{(m)}(\mu;\nu)e^{-\tau/\mu}\Theta(\mu) + \omega\int_0^1 d\nu' e^{-\tau/\nu'}\Psi_m^{\nu'}(\mu;\nu)\Gamma_{n-1}^{(m)}(\nu';\mu) - \omega e^{-\tau/\mu}\Theta(\mu)\int_{-1}^1 d\nu'\Psi_m^{\mu}(\nu';\nu)\Gamma_{n-1}^{(m)}(\mu;\nu'). \quad (1.24)$$

We remark here that functions Ψ_m^ν are independent of n. In the spectral representation of n-times scattered radiation $I_n^{(m)}$ [see Eq. (1.24)], the order of scattering occurs only via coefficients $\Gamma_n^{(m)}$. Consequently, the functions Ψ_m^ν are fundamental in the sense that they involve the scattering properties of the medium and not the order of scattering. Finally, the coefficients $\Gamma_n^{(m)}$ are to be determined by the boundary condition at $\tau = 0$. In other words, at $\tau = 0$, we assume that

$$I_n^{(m)}(0,\mu) = \delta_{n0}I_0^{(m)}(0,\mu), \quad \mu > 0, \quad (1.25)$$

where $I_0^{(m)}(0,\mu)$ for $\mu > 0$ is the known incident radiation appropriately integrated over the azimuth angle Φ [see Eq. (1.5)]. Then, the set of equations which determine $\Gamma_n^{(m)}$ is

$$\Gamma_n^{(m)}(\mu;\nu) = \omega\int_{-1}^1 d\nu'\Psi_m^{\mu}(\nu';\nu)\Gamma_{n-1}^{(m)}(\mu;\nu') - \omega\int_0^1 d\nu'\Psi_m^{\nu'}(\mu;\nu)\Gamma_{n-1}^{(m)}(\nu';\mu) + \delta_{n0}I_0^{(m)}(0,\mu)S_m(\mu,\nu), \quad \mu > 0. \quad (1.26)$$

In particular

$$\Gamma_0^{(m)}(\mu;\nu) = I_0^{(m)}(0,\mu)S_m(\mu,\nu), \quad \mu > 0, \tag{1.27}$$

and

$$\Gamma_n^{(m)}(\mu;\nu) = \omega\int_{-1}^{1} d\nu' \Psi_m^{\mu}(\nu';\nu)\Gamma_{n-1}^{(m)}(\mu;\nu') - \omega\int_0^1 d\nu' \Psi_m^{\nu'}(\mu;\nu)\Gamma_{n-1}^{(m)}(\nu';\mu), \quad \mu > 0, \quad n \geq 1. \tag{1.28}$$

This is the recurrence relation which determines all the coefficients $\Gamma_n^{(m)}$, in terms of normal modes, and the scattering kernel $S_m(\mu,\nu)$. Clearly $\Gamma_n^{(m)}$ are proportional to ω^n. Therefore if we set

$$\Gamma_n^{(m)}(\mu;\nu) = \omega^n A_n^{(m)}(\mu;\nu), \tag{1.29}$$

then, A_n^m satisfy

$$A_n^{(m)}(\mu;\nu) = \int_{-1}^{1} d\nu' \Psi_m^{\mu}(\nu';\nu)A_{n-1}^{(m)}(\mu;\nu') - \int_0^1 d\nu' \Psi_m^{\nu'}(\mu;\nu)A_{n-1}^{(m)}(\nu';\mu), \quad \mu > 0, n \geq 1, \tag{1.30}$$

with

$$A_0^{(m)}(\mu;\nu) = I_0^{(m)}(0,\mu)S_m(\mu;\nu). \tag{1.31}$$

Now in Eq. (19) if we divide both sides by $S_m(\mu,\nu)$, then in terms of the coefficients A_n^m defined by Eq. (1.29), the component of the radiation n times scattered is given by

$$I_n^{(m)}(\tau,\mu)\Theta(\tau) = \omega^n\left[e^{-\tau/\mu}\Theta(\mu)\left(\frac{A_n^{(m)}(\mu;\nu)}{S_m(\mu,\nu)} - \int_{-1}^{1} d\nu' \phi^{\mu}(\nu')\frac{S_m(\nu',\nu)}{S_m(\mu,\nu)}A_{n-1}^{(m)}(\mu;\nu')\right) + \int_0^1 d\nu' \phi^{\nu'}(\mu)e^{-\tau/\nu'}A_{n-1}^{(m)}(\nu';\mu)\right]. \tag{1.32}$$

In particular, at any $\tau > 0$ the nth order back-scattered ($\mu < 0$) radiation is

$$I_n^{(m)}(\tau,\mu) = \omega^n\int_0^1 d\nu' \phi^{\nu'}(\mu)e^{-\tau/\nu'}A_{n-1}^{(m)}(\nu';\mu), \quad \mu < 0. \tag{1.33}$$

The emergent intensity ($\tau = 0$) is

$$I_n^{(m)}(0,\mu) = \omega^n\int_0^1 d\nu' \phi^{\nu'}(\mu)A_{n-1}^{(m)}(\nu';\mu), \quad \mu < 0. \tag{1.34}$$

From Eqs. (1.33) and (1.34) we conclude that for $n = 0$, corresponding to the process in which the radiation has suffered no encounters, the back-scattered and emergent intensities are zero, as they should be.

Let us now introduce the Neumann coefficients by defining

$$N_n(\tau,\Omega) = \sum_{m=-\infty}^{\infty} e^{im\phi}\left[e^{-\tau/\mu}\Theta(\mu)\left\{\frac{A_{n-1}^{(m)}(\mu;\nu)}{S_m(\mu,\nu)} - \int_{-1}^{1} d\nu' \phi^{\mu}(\nu')\frac{S_m(\nu',\nu)}{S_m(\mu,\nu)}A_{n-1}^{(m)}(\mu;\nu')\right\} + \int_0^1 d\nu' \phi^{\nu'}(\mu)e^{-\tau/\nu'}A_{n-1}^{(m)}(\nu';\mu)\right]. \tag{1.35}$$

In terms of these coefficients the total intensity $I(\tau,\Omega)$ at any optical depth τ is given by

$$I(\tau,\Omega) = \sum_{n=0}^{\infty} \omega^n N_n(\tau,\Omega). \tag{1.36}$$

We remark here that in practical calculations this so-called "Neumann expansion" [Eq. (1.36)] of $I(\tau,\Omega)$ is very convenient for such purposes as computation of absorption line strengths, since it eliminates the problem of repeatedly solving the transfer equation numerically for various albedos ω (the albedo being a function of frequency). In other words, the fact that the Neumann coefficients N_n correspond to the multiple scattering process in a conservative medium ($\omega = 1$), and the same coefficients appear in the case of an arbitrary nonconservative ($\omega < 1$) medium, permits us to reduce the problem of solving the transfer equation for various frequency dependent albedos to that of obtaining the coefficients N_n once and for all. The solution for a nonconservative medium then results by trivial summation as given by Eq. (1.36). Unfortunately, the Neumann series [Eq. (1.36)] for a near conservative medium ($\omega \simeq 1$) converges very slowly. This, in a way, limits the usefulness of such an expansion (c.f. however, Uesugi and Irvine[10]). However, by a simple modification of the multiple scattering process described earlier, an expansion in powers of $(1 - \omega)$ can be obtained. We shall treat that case in a forthcoming paper.

In the next section we examine the structure of coefficients $A_n^{(m)}$ for the simple isotropically scattering medium. Since the Neumann coefficients N_n are related to $A_n^{(m)}$ by means of Eq. (1.35), the convergence properties of the series in Eq. (1.36) then become transparent.

2. APPLICATIONS

Isotropic Case (Semi-Infinite Atmosphere)

For the isotropically scattering medium, the formulas derived in Sec. 1 are considerably simplified. Let us first examine the recurrence relation for $A_n^{(m)}$ [Eq. (1.30)]. Since for the case under consideration $m = 0$, we have

$$A_n^{(0)}(\mu;\nu) = \int_{-1}^{1} d\nu' \Psi_0^{\mu}(\nu';\nu)A_{n-1}^{(0)}(\mu;\nu') - \int_0^1 d\nu' \Psi_0^{\nu'}(\mu;\nu)A_{n-1}^{(0)}(\nu';\mu). \tag{2.1}$$

Now, by definition [see Eq. (1.23)]

$$\Psi_0^{\nu'}(\mu;\nu) = \phi^{\nu'}(\mu)S_0(\mu,\nu)$$

and for the isotropic case

$$S_0(\mu,\nu) = 1.$$

Then $\Psi_0^{\nu'}(\mu;\nu)$ reduce to normal modes, i.e.,

$$\Psi_0^{\mu}(\nu';\nu) = \phi^{\mu}(\nu'). \tag{2.2}$$

Also, from definitions of $A_n^{(0)}(\mu;\nu)$, $\Gamma_n^{(0)}(\mu;\nu)$ [see Eqs. (1.29), (1.20), and (1.11)], we conclude that

$$A_n^{(0)}(\mu;\nu) = A_n(\mu), \tag{2.3}$$

i.e., $A_n^{(0)}(\mu;\nu)$ are independent of the second argument ν.

With these simplifications, Eq. (2.1) becomes

$$A_n(\mu) = A_{n-1}(\mu) - \int_0^1 d\nu\phi^{\nu}(\mu)A_{n-1}(\nu), \tag{2.4}$$

where we have used the normalization

$$\int_{-1}^1 d\mu\phi^{\nu}(\mu) = 1 \tag{2.5}$$

of the modes.

Before we solve the recurrence relation (2.4) for $A_n(\mu)$, let us introduce the functions

$$\begin{aligned} H_0(\mu;\nu) &= \delta(\mu-\nu), \\ H_1(\mu;\nu) &= \phi^{\nu}(\mu), \\ H_2(\mu;\nu) &= \int_0^1 d\nu_1\phi^{\nu_1}(\mu)H_1(\nu_1;\nu), \end{aligned} \tag{2.6a}$$

and in general

$$H_l(\mu;\nu) = \int_0^1 d\nu_1\phi^{\nu_1}(\mu)H_{l-1}(\nu_1;\nu), \quad l \geq 1. \tag{2.6b}$$

Now one may readily show that Eq. (4) has the solution

$$A_n(\mu) = \sum_{l=0}^{n} (-1)^l \frac{n!}{(n-l)!\,l!} \int_0^1 d\nu A_0(\nu)H_l(\mu;\nu), \tag{2.7}$$

with

$$A_0(\mu) = I_0(0,\mu), \quad \mu > 0, \tag{2.8}$$

being the incident radiation.

Finally for the intensity $I_n^{(0)}(\tau,\mu)$, Eq. (1.32) gives us

$$\begin{aligned} &I_n(\tau,\mu)\Theta(\tau) \\ &\quad = \omega^n[A_n(\mu)e^{-\tau/\mu}\Theta(\mu) + \int_0^1 d\nu\phi^{\nu}(\mu)e^{-\tau/\nu}A_{n-1}(\nu) \\ &\quad - e^{-\tau/\mu}\Theta(\mu)A_{n-1}(\mu),], \end{aligned} \tag{2.9}$$

where we have deleted the superscript (0) on $I_n(\tau,\mu)$. In particular, for $\mu < 0$ (the backscattered radiation) we have

$$\begin{aligned} &I_n(\tau,\mu) = \omega^n \int_0^1 d\nu\phi^{\nu}(\mu)e^{-\tau/\nu}A_{n-1}(\nu), \\ &\quad \mu < 0, \tau > 0, \end{aligned} \tag{2.10}$$

whence the reflected intensity at $\tau = 0$ is given by

$$I_n(0,\mu) = \omega^n \int_0^1 d\nu\phi^{\nu}(\mu)A_{n-1}(\nu), \quad \mu < 0. \tag{2.11}$$

For $n = 0$,

$$I_0(0,\mu) = I_0(\tau,\mu) = 0, \quad \mu < 0. \tag{2.12}$$

Inserting the explicit form of $A_n(\nu)$ [given by Eq. (2.7)] in Eq. (2.11) yields

$$\begin{aligned} I_{n+1}(0,\mu) &= \omega^{n+1} \sum_{l=0}^{n} (-1)^l \frac{n!}{(n-l)!\,l!} \\ &\quad \times \int_0^1 d\nu A_0(\nu)H_{l+1}(\mu;\nu), \quad \mu < 0. \end{aligned} \tag{2.13}$$

The total emergent intensity is then

$$I(0,\mu) = \sum_{n=0}^{\infty} I_{n+1}(0,\mu), \quad \mu < 0,$$

i.e.,

$$\begin{aligned} I(0,\mu) &= \omega \sum_{n=0}^{\infty} \sum_{l=0}^{n} (-1)^l \omega^n \frac{n!}{(n-l)!\,l!} \\ &\quad \times \int_0^1 d\nu A_0(\nu)H_{l+1}(\mu;\nu), \quad \mu < 0. \end{aligned} \tag{2.14}$$

From this we conclude that the Neumann coefficients associated with the emergent intensity are

$$\begin{aligned} N_n(0,\mu) &= \sum_{l=0}^{n} (-1)^n \frac{n!}{(n-l)!\,l!} \\ &\quad \int_0^1 d\nu A_0(\nu)H_{l+1}(\mu;\nu), \quad \mu < 0, \end{aligned} \tag{2.15}$$

where, as a reminder, $H_{l+1}(\mu;\nu)$ defined by Eq. (6) are the multiple integrals of normal modes.

In Eq. (2.14) the right-hand side may be reduced to a single sum by an application of the Cauchy sum formula[11]

$$\sum_{n=0}^{\infty} \sum_{k=0}^{n} B(n;k) = \sum_{n=0}^{\infty} \sum_{k=0}^{\infty} B(n+k;k). \tag{2.16}$$

The result is

$$\begin{aligned} I(0,\mu) &= \frac{\omega}{1-\omega} \sum_{n=0}^{\infty} (-1)^n \left(\frac{\omega}{1-\omega}\right)^n \\ &\quad \times \int_0^1 d\nu A_0(\nu)H_{n+1}(\mu;\nu), \quad \mu < 0. \end{aligned} \tag{2.17}$$

The slow convergence of the Neumann expansion for $\omega \simeq 1$ should now be obvious from Eq. (2.17).

3. GENERAL FORMULATION FOR FINITE REGIONS

To see how the general formulation presented in Sec. 1 for the semi-infinite region is modified in the case of a bounded region, let us consider an atmosphere bounded by two planes at $\tau = 0$ and $\tau = d$. We assume that the incident intensity at $\tau = 0, \mu > 0$ and the intensity at $\tau = d$ for $\mu < 0$ are known. Following the procedure of Sec. 1, one may readily show that the integral representation of the n-times scattered radiation $I_n^{(n)}(\tau,\mu)$ is

$$I_n^{(m)}(\tau,\mu)\Theta(\tau)\Theta(d-\tau)=\frac{1}{2\pi}\int_{-\infty}^{\infty}dk\,\frac{e^{ik\tau}}{1+ik\mu}\times[\mu I_n^{(m)}(0,\mu)-e^{-ikd}\mu I_n^{(m)}(d,\mu)]+\frac{\omega}{4\pi}\int_{-\infty}^{\infty}dk\,\frac{e^{ik\tau}}{1+ik\mu}R_{n-1}^{(m)}(k;\mu;d),\tag{3.1}$$

where, in contrast to the half-space problems [see Sec. 1, Eq. (11)], $R_n^{(m)}(k;\mu;d)$ are now the finite Fourier transforms of the Source function, i.e.,

$$R_n^{(m)}(k;\mu;d)=\int_0^d d\tau e^{-ik\tau}\int_{-1}^{1}d\mu' S_m(\mu,\mu')I_n^{(m)}(\tau,\mu').\tag{3.2}$$

By the same token we have the recurrence relation for $R_n^{(m)}$ of the form

$$R_n^{(m)}(k;\mu;d)=J_n^{(m)}(k;\mu;0)-e^{-ikd}J_n^{(m)}(k;\mu;d)+\frac{\omega}{2}\int_{-1}^{1}d\mu'\,\frac{S_m(\mu,\mu')}{1+ik\mu'}R_{n-1}^{(m)}(k;\mu';d),\tag{3.3}$$

where for $x=0$ or d

$$J_n^{(m)}(k;\mu;x)=\int_{-1}^{1}d\mu'\mu' I_n^{(m)}(x,\mu')\,S_m(\mu,\mu')/(1+ik\mu')\tag{3.4}$$

From Eq. (3.3) it is clear that $R_n^{(m)}(k;\mu;d)$ are explicitly dependent on the total optical thickness d. Aside from that, these coefficients share the same analytical properties as those discussed in Sec. 1 for similar coefficients for half-space problems. Thus, to remind us that $R_n^{(m)}(k;\mu;d)$ are sectionally holomorphic in the complex k plane cut from $-i\infty$ to $-i$ and i to $i\infty$. However, due to the exponential dependence on d they diverge as $|k|\to\infty$ in the upper half k plane. Consequently, in obtaining the spectral representation of $I_n^{(m)}(\tau,\mu)$ by considering the contour (see figure in Sec. 1) in the upper or the lower half k plane, it is necessary to decompose $R_n^{(m)}$ into two parts which converge in the respective parts of the k plane. Such a decomposition is readily obtained by noting that upon solving Eq. (3.3) one has

$$R_n^{(m)}(k;\mu;d)=U_n^{(m)}(k;\mu;0)-e^{-ikd}U_n^{(m)}(k;\mu;d),\tag{3.5}$$

where $U_n^{(m)}$, for $x=0$, or d, satisfy the following recurrence relation:

$$U_n^{(m)}(k;\mu;x)=J_n^{(m)}(k;\mu;x)+\frac{\omega}{2}\int_{-1}^{1}d\mu'\times\frac{S_m(\mu,\mu')}{1+ik\mu'}U_{n-1}^{(m)}(k;\mu';x).\tag{3.6}$$

By inserting the decomposition, given by Eq (3.5), into Eq. (3.1), and solving the first integral on the right-hand side, we obtain

$$I_n^{(m)}(\tau,\mu)\Theta(\tau)\Theta(d-\tau)=I_n^{(m)}(0,\mu)e^{-\tau/\mu}\Theta(\mu)+\frac{\omega}{4\pi}\int_{-\infty}^{\infty}dk\,\frac{e^{ik\tau}}{1+ik\mu}U_{n-1}^{(m)}(k;\mu;0)-I_n^{(m)}(d;\mu)e^{(d-\tau)/\mu}\Theta(-\mu)-\frac{\omega}{4\pi}\int_{-\infty}^{\infty}dk\,\frac{e^{ik(d-\tau)}}{1+ik\mu}U_{n-1}^{(m)}(k;\mu;d).\tag{3.7}$$

Now since $U_n^{(m)}(k;\mu;0)$ are independent of d and $U_n^{(m)}(k;\mu;d)$ involve d only through surface distribution at d, we distort the path of integration in the upper half K plane in the first integral and in the lower half in the second integral given on the right-hand side of Eq. (3.7), respectively. In the manner of the procedure followed in Sec. 1 we obtain the following spectral representation of $I_n^{(m)}(\tau,\mu)$:

$$\begin{aligned}I_n^{(m)}(\tau,\mu)\Theta(\tau)\Theta(d-\tau)S_m(\mu;\nu)&=\omega\int_0^1 d\nu' e^{-\tau/\nu'}\Psi_m^{\nu'}(\mu;\nu)\Gamma_{n-1}^{(m)}(\nu';\mu;0)\\&-e^{-\tau/\mu}\Theta(\mu)\omega\int_{-1}^{1}d\nu'\Psi_m^{\mu}(\nu';\nu)\\&\times\Gamma_{n-1}^{(m)}(\mu;\nu';0)+\Gamma_n^{(m)}(\mu;\nu,0)e^{-\tau/\mu}\Theta(\mu)\\&-\omega\int_{-1}^{0}d\nu' e^{(d-\tau)/\nu'}\Psi_m^{\nu'}(\mu;\nu)\Gamma_{n-1}^{(m)}(\nu';\mu;d)\\&+e^{(d-\tau)/\mu}\Theta(-\mu)\omega\int_{-1}^{1}d\nu'\Psi_m^{\mu}(\nu';\nu)\\&\times\Gamma_{n-1}^{(m)}(\mu;\nu';d)-\Gamma_n^{(m)}(\mu;\nu;d)e^{(d-\tau)/\mu}\Theta(-\mu),\end{aligned}\tag{3.8}$$

where, for $x=0$ or d,

$$\Gamma_n^{(m)}(\nu;\mu;x)=-(1/2\pi i\nu^2)(U_n^{(m)+}(\nu;\mu;x)-U_n^{(m)-}(\nu;\mu;x)),\tag{3.9}$$

and, as before

$$\Psi_m^{\nu'}(\mu;\nu)=\phi^{\nu'}(\mu)S_m(\mu;\nu),\tag{3.10}$$

with $\phi^{\nu'}(\mu)$ being the continuum normal modes defined by Eq. (1.21). Equations which determine coefficients $\Gamma_n^{(m)}$ are obtained by letting $\tau\to 0$ and $\tau\to d$ in Eq. (3.9), so that

$$\begin{aligned}I_n^{(m)}(0,\mu)S_m(\mu,\nu)&=\omega\int_0^1 d\nu'\Psi_m^{\nu'}(\mu;\nu)\Gamma_{n-1}^{(m)}(\nu';\mu;0)\\&+\Gamma_n^{(m)}(\mu;\nu;0)-\omega\int_{-1}^{1}d\nu'\Psi_m^{\mu}(\nu';\nu)\\&\times\Gamma_{n-1}^{(m)}(\mu;\nu';0)-\omega\int_{-1}^{0}d\nu' e^{d/\nu'}\Psi_m^{\nu'}(\mu;\nu)\\&\times\Gamma_{n-1}^{(m)}(\nu';\mu;d),\qquad\mu>0,\end{aligned}\tag{3.11}$$

and

$$\begin{aligned}I_n^{(m)}(d,\mu)S_m(\mu,\nu)&=\omega\int_0^1 d\nu' e^{-d/\nu'}\Psi_m^{\nu'}(\mu;\nu)\Gamma_{n-1}^{(m)}(\nu';\mu;0)\\&-\omega\int_{-1}^{0}d\nu'\Psi_m^{\nu'}(\mu;\nu)\Gamma_{n-1}^{(m)}(\nu';\mu;d)\\&-\Gamma_n^{(m)}(\mu;\nu;d)+\omega\int_{-1}^{1}d\nu'\Psi_m^{\mu}(\nu';\nu)\\&\times\Gamma_{n-1}^{(m)}(\mu;\nu';d),\qquad\mu<0,\end{aligned}\tag{3.12}$$

where

$$I_n^{(m)}(0,\mu) = I_0^{(m)}(0,\mu)\delta_{n0}, \quad \mu > 0,$$
$$I_n^{(m)}(d,\mu) = I_0^{(m)}(d,\mu)\delta_{n0}, \quad \mu < 0, \tag{3.13}$$

are known from the boundary conditions.

Finally, the reflected and the transmitted components of the intensity are

$$I_n^{(m)}(0,\mu)S_m(\mu,\nu) = \omega\int_0^1 d\nu' \Psi_m^{\nu'}(\mu;\nu)\Gamma_{n-1}^{(m)}(\nu';\mu;0)$$
$$- \omega\int_{-1}^0 d\nu' e^{d/\nu'}\Psi_m^{\nu'}(\mu;\nu)\Gamma_{n-1}^{(m)}(\nu';\mu;d)$$
$$- \Gamma_n^{(m)}(\mu;\nu;d)e^{d/\mu} + e^{d/\mu}\omega\int_{-1}^1 d\nu' \Psi_m^{\mu}(\nu';\nu)$$
$$\times \Gamma_{n-1}^{(m)}(\mu;\nu';d), \quad \mu < 0, \tag{3.14}$$

and

$$I_n^{(m)}(d,\mu)S_m(\mu,\nu) = \omega\int_0^1 d\nu' e^{-d/\nu'}\Psi_m^{\nu'}(\mu;\nu)$$
$$\times \Gamma_{n-1}^{(m)}(\nu';\mu;0) - e^{-d/\mu}\omega\int_{-1}^1 d\nu' \Psi_m^{\mu}(\nu';\nu)$$
$$\times \Gamma_{n-1}^{(m)}(\mu;\nu';0) + \Gamma_n^{(m)}(\mu;\nu;0)e^{-d/\mu}$$
$$- \omega\int_{-1}^0 d\nu' \Psi_m^{\nu'}(\mu;\nu)\Gamma_{n-1}^{(m)}(\nu';\mu;d), \quad \mu > 0, \tag{3.15}$$

respectively.

4. APPLICATIONS

Isotropic Case (Bounded Media)

For the isotropically scattering medium, various equations derived in Sec. 3 are again considerably simplified. Thus, in the notation used in Sec. 2 the n-times scattered radiation $I_n(\tau,\mu)$ at any point [see Eq. (3.8)] is given by

$$I_n(\tau,\mu)\Theta(\tau)\Theta(d-\tau) = \omega\int_0^1 d\nu e^{-\tau/\nu}\phi^{\nu}(\mu)\Gamma_{n-1}(\nu;0)$$
$$- e^{-\tau/\mu}\Theta(\mu)\omega\,\Gamma_{n-1}(\mu;0) + \Gamma_n(\mu;0)e^{-\tau/\mu}\Theta(\mu)$$
$$- \omega\int_{-1}^0 d\nu e^{(d-\tau)/\nu}\phi^{\nu}(\mu)\Gamma_{n-1}(\nu;d)$$
$$+ e^{(d-\tau)/\mu}\Theta(-\mu)\omega\,\Gamma_{n-1}(\mu;d)$$
$$- \Gamma_n(\mu;d)e^{(d-\tau)/\mu}\Theta(-\mu). \tag{4.1}$$

Equations which determine the unknown coefficients $\Gamma_n(\mu;0)$ and $\Gamma_n(\mu;d)$ from the boundary conditions at $\tau = 0$ ($\mu > 0$) and $\tau = d$ ($\mu < 0$) are

$$I_n(0,\mu) = \omega\int_0^1 d\nu\phi^{\nu}(\mu)\Gamma_{n-1}(\nu;0) + \Gamma_n(\mu;0)$$
$$- \omega\Gamma_{n-1}(\mu;0) - \omega\int_{-1}^0 d\nu e^{d/\nu}\phi^{\nu}(\mu)\Gamma_{n-1}(\nu;d),$$
$$\mu > 0, \tag{4.2}$$

and

$$I_n(d,\mu) = \omega\int_0^1 d\nu e^{-d/\nu}\phi^{\nu}(\mu)\Gamma_{n-1}(\nu;0) - \Gamma_n(\mu;d)$$
$$- \omega\int_{-1}^0 d\nu\phi^{\nu}(\mu)\Gamma_{n-1}(\nu;d) + \omega\Gamma_{n-1}(\mu;d),$$
$$\mu < 0, \tag{4.3}$$

where

$$I_n(0,\mu) = \delta_{n0}I_0(0,\mu), \quad \mu > 0, \tag{4.4a}$$

and

$$I_n(d,\mu) = \delta_{n0}I_0(d,\mu), \quad \mu < 0, \tag{4.4b}$$

are known. These are coupled recurrence relations which have the following formal solutions expressed in terms of the H_j functions defined by Eq. (2.6b):

$$\Gamma_n(\mu;0) = \omega^n\int_0^1 d\nu E_n(\mu;\nu)I_0(0,\nu) - \sum_{l=0}^n \omega^{l+1}$$
$$\times \int_0^1 d\nu\int_{-1}^0 d\nu' e^{d/\nu'}E_l(\mu;\nu)\phi^{\nu'}(\nu)$$
$$\times \Gamma_{n-l-1}(\nu';d), \quad \mu > 0, \tag{4.5}$$

and

$$\Gamma_n(-\mu;d) = -\omega^n\int_0^1 d\nu E_n(\mu;\nu)I_0(d,-\nu) + \sum_{l=0}^n \omega^{l+1}$$
$$\times \int_0^1 d\nu\int_{-1}^0 d\nu' e^{d/\nu'}E_l(\mu;\nu)\phi^{\nu'}(\nu)$$
$$\times \Gamma_{n-l-1}(-\nu';0), \quad \mu > 0, \tag{4.6}$$

where for convenience we have set

$$E_n(\mu;\nu) = \sum_{j=0}^n \frac{(-1)^j n!}{(n-j)!\,j!}H_j(\mu;\nu). \tag{4.7}$$

Equations (4.5) and (4.6) are particularly suited for asymptotic calculations for large d. For example, the zeroth approximation is obtained by neglecting the terms involving the exponentials. The results is then

$$\Gamma_n(\mu;0) = \omega^n\int_0^1 d\nu E_n(\mu;\nu)I_0(0,\nu) \quad \mu > 0, \tag{4.8}$$

$$\Gamma_n(-\mu;d) = -\omega^n\int_0^1 d\nu E_n(\mu;\nu)I_0(d,-\nu) \quad \mu > 0, \tag{4.9}$$

which are merely the coefficients associated with the half-space problems. By using these values of $\Gamma_n(\mu;0)$ and $\Gamma_n(-\mu;d)$, correction terms may then be obtained by calculating the terms previously neglected and so on. For arbitrary values of d, exact solutions of Eqs. (5) and (6) pose no real difficulties.

Finally the reflected and the transmitted components of the intensities are given by

$$I_n(0,\mu) = \omega\int_0^1 d\nu\phi^{\nu}(\mu)\Gamma_{n-1}(\nu;0) - \omega\int_{-1}^0 d\nu e^{d/\nu}$$
$$\times \phi^{\nu}(\mu)\Gamma_{n-1}(\nu;d) + e^{d/\mu}\omega\Gamma_{n-1}(\mu;d)$$
$$- \Gamma_n(\mu;d)e^{d/\mu}, \quad \mu < 0, \tag{4.10}$$

and

$$I_n(d;\mu) = \omega\int_0^1 d\nu e^{-d/\nu}\phi^{\nu}(\mu)\Gamma_{n-1}(\nu;0) - e^{-d/\mu} \times \omega\Gamma_{n-1}(\mu;0) + \Gamma_n(\mu;0)e^{-d/\mu} - \omega\int_{-1}^0 d\nu\phi^{\nu}(\mu) \times \Gamma_{n-1}(\nu;d), \quad \mu > 0, \tag{4.11}$$

respectively.

ACKNOWLEDGMENTS

The author is grateful to Professor William M. Irvine for suggesting the problem and for many valuable discussions. This work was supported by NASA NGL 22-010-023.

* Contribution No. 104 of the Five College Observatories.

1 W. M. Irvine, Astrophys. J., **152**, 823 (1968).
2 A. Uesugi and W. M. Irvine, Astrophys. J., **159**, 127 (1970).
3 A. Uesugi and W. M. Irvine, Astrophys. J., **161**, 243 (1970).
4 H. C. van de Hulst, Astrophys. J., **107**, No. 220 (1948).
5 K. M. Case and P. F. Zweifel, *Linear Transport Theory* (Addison-Wesley, Reading, Mass. 1967).
6 K. M. Case, "On the Boundary-Value Problems of Linear Transport Theory" (The University of Michigan, 1967).
7 M. Kanal, J. Math. Phys., **11**, 3042 (1970).
8 M. Kanal, "Boundary-Value Problems of Radiative Transfer-Further Applications of Case's Green's Function Approach" (to be published).
9 The Fourier coefficients $R_n^{(m)}(k;\mu)$, for the half-space problems, are one-sided Fourier Transforms of the Source function. In the case of finite regions, say bounded between $\tau = 0$ and $\tau = d$, these coefficients as we shall see in Sec. 3 become the finite Fourier transforms of the source function. One important fact we should keep in mind is that $R_n^{(m)}(k;\mu)$ are not symmetric under the exchange of the arguments k, μ. For that reason the two arguments are separated by a semicolon.
10 A. Uesugi and W. M. Irvine, J. Atmos. Sci. **26**, 973 (1969).
11 E. D. Rainville, *Special Functions*, (Macmillan, New York, 1960).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 10 OCTOBER 1971

Remarks on the Existence of Solutions of the Two-Particle Lippmann–Schwinger Equation. II

C. S. Shastry* and A. K. Rajagopal
Department of Physics and Astronomy, Louisiana State University, Baton Rouge, Louisiana 70803
(Received 22 April 1971)

We prove the following: The kernel $K_s^{(l)}$ corresponding to the lth partial wave of the modified partial wave Lippmann–Schwinger equation having symmetric kernel $K_s = V^{1/2}G_0^+V^{1/2}$ belongs to the Hilbert–Schmidt (L^2) class for spherically symmetric potential V such that $V(r) \to r^{-\delta}$, $\delta < 2$ as $r \to 0$; $V(r) \to r^{-\eta}$, $\eta > 1$ as $r \to \infty$. For $V(r) = g/r^{\alpha}$, $1 < \alpha < 2$, the kernel K_s of the modified Lippmann–Schwinger equation satisfies $\mathrm{Tr}\{(K_s^{\dagger})^m K_s^m\} < \infty$, if $\alpha > 1 + 1/m$, $m = 2, 3, \cdots$. The above results are valid even for positive *real energy*. It is also shown that for potentials for which both K_s and the kernel $K = G_0^+V$ of the usual Lippmann–Schwinger equation belong to L^2 class, the corresponding off-shell two-particle T matrix $\langle \mathbf{k}|T(E + i\epsilon)|\mathbf{k}'\rangle$ has a unique limit as $\epsilon \to 0$, $E > 0$. Some important consequences of these results are discussed.

I. INTRODUCTION

In a recent paper[1] (hereafter referred to as I) we examined the conditions on the two-particle potential V for which the associated kernel $K(= G_0^+V)$ (or some power of it) of the Lippmann–Schwinger (LS) equation belongs to the Hilbert–Schmidt (L^2) class, for complex energies only. It is found that when the energy is made purely real and positive by letting the imaginary part of it vanish, the trace of (K^+K) is unbounded, showing that K does not belong to the L^2 class. This happens even for short-ranged potentials of the Yukawa or cut-off types. This points to the fact that if the two-particle T matrix associated with K is obtained as a solution of the LS equation for complex energies, its limit as the energy becomes real and positive may not tend to a unique value. To circumvent this difficulty, it is often suggested[2-7] that one should examine if the symmetrized form of the LS kernel defined by

$$K_s = V^{1/2}KV^{-1/2} \tag{1.1}$$

belongs to the L^2 class for complex as well as real energies. The associated T matrix, denoted by T_s, is related to that defined for the usual LS equation denoted by T, through the expression

$$T_s = V^{-1/2}T. \tag{1.2}$$

These conditions hold for a class of potentials which obey the conditions:

$$\mathrm{Tr}\{K^{\dagger}K\} = (m^2/2\pi\,|\mathrm{Im}k|)\int d^3r\,|\,V^2(\mathbf{r})\,|^2 < \infty, \tag{1.3}$$

$$\mathrm{Tr}\{K_s^{\dagger}Ks\} = \frac{m^2}{4\pi^2}\int d^3r\int d^3r'\,|V(\mathbf{r})||\,V(\mathbf{r}')\,| \times \frac{e^{-2\,\mathrm{Im}k\,|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|^2} < \infty. \tag{1.4}$$

Here m is the reduced mass, Imk is related to the imaginary part of the energy via the relation $k^2/2m = E + i\epsilon$. Consider for instance, the Yukawa potential $V(r) = V_0e^{-\frac{1}{2}r}/r$ for which both (1.3) and (1.4) hold for Im$k > 0$, while for Im$k = 0$ (1.3) is divergent, whereas (1.4) is not. This indicates that if it can be shown that the T matrix obtained from K and that obtained from K_s are identical for complex energies, then even though (1.3) does not hold for real positive energies, a unique limit of it is assured because of (1.4). We prove this here for a wide class of potentials, including those covered by our earlier work.

In this paper we establish the statements similar to those obtained in I for K, K_l, the partial wave

$$I_n(d;\mu) = \omega\int_0^1 d\nu e^{-d/\nu}\phi^{\nu}(\mu)\Gamma_{n-1}(\nu;0) - e^{-d/\mu} \times \omega\Gamma_{n-1}(\mu;0) + \Gamma_n(\mu;0)e^{-d/\mu} - \omega\int_{-1}^0 d\nu\phi^{\nu}(\mu) \times \Gamma_{n-1}(\nu;d), \quad \mu > 0, \tag{4.11}$$

respectively.

ACKNOWLEDGMENTS

The author is grateful to Professor William M. Irvine for suggesting the problem and for many valuable discussions. This work was supported by NASA NGL 22-010-023.

* Contribution No. 104 of the Five College Observatories.

1 W. M. Irvine, Astrophys. J., **152**, 823 (1968).
2 A. Uesugi and W. M. Irvine, Astrophys. J., **159**, 127 (1970).
3 A. Uesugi and W. M. Irvine, Astrophys. J., **161**, 243 (1970).
4 H. C. van de Hulst, Astrophys. J., **107**, No. 220 (1948).
5 K. M. Case and P. F. Zweifel, *Linear Transport Theory* (Addison-Wesley, Reading, Mass. 1967).
6 K. M. Case, "On the Boundary-Value Problems of Linear Transport Theory" (The University of Michigan, 1967).
7 M. Kanal, J. Math. Phys., **11**, 3042 (1970).
8 M. Kanal, "Boundary-Value Problems of Radiative Transfer-Further Applications of Case's Green's Function Approach" (to be published).
9 The Fourier coefficients $R_n^{(m)}(k;\mu)$, for the half-space problems, are one-sided Fourier Transforms of the Source function. In the case of finite regions, say bounded between $\tau = 0$ and $\tau = d$, these coefficients as we shall see in Sec. 3 become the finite Fourier transforms of the source function. One important fact we should keep in mind is that $R_n^{(m)}(k;\mu)$ are not symmetric under the exchange of the arguments k, μ. For that reason the two arguments are separated by a semicolon.
10 A. Uesugi and W. M. Irvine, J. Atmos. Sci. **26**, 973 (1969).
11 E. D. Rainville, *Special Functions*, (Macmillan, New York, 1960).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 10 OCTOBER 1971

Remarks on the Existence of Solutions of the Two-Particle Lippmann–Schwinger Equation. II

C. S. Shastry* and A. K. Rajagopal
Department of Physics and Astronomy, Louisiana State University, Baton Rouge, Louisiana 70803
(Received 22 April 1971)

We prove the following: The kernel $K_s^{(l)}$ corresponding to the lth partial wave of the modified partial wave Lippmann–Schwinger equation having symmetric kernel $K_s = V^{1/2}G_0^+V^{1/2}$ belongs to the Hilbert–Schmidt (L^2) class for spherically symmetric potential V such that $V(r) \to r^{-\delta}$, $\delta < 2$ as $r \to 0$; $V(r) \to r^{-\eta}$, $\eta > 1$ as $r \to \infty$. For $V(r) = g/r^{\alpha}$, $1 < \alpha < 2$, the kernel K_s of the modified Lippmann–Schwinger equation satisfies $\mathrm{Tr}\{(K_s^{\dagger})^m K_s^m\} < \infty$, if $\alpha > 1 + 1/m$, $m = 2, 3, \cdots$. The above results are valid even for positive *real energy*. It is also shown that for potentials for which both K_s and the kernel $K = G_0^+V$ of the usual Lippmann–Schwinger equation belong to L^2 class, the corresponding off-shell two-particle T matrix $\langle \mathbf{k}|T(E + i\epsilon)|\mathbf{k}'\rangle$ has a unique limit as $\epsilon \to 0$, $E > 0$. Some important consequences of these results are discussed.

I. INTRODUCTION

In a recent paper[1] (hereafter referred to as I) we examined the conditions on the two-particle potential V for which the associated kernel $K(= G_0^+V)$ (or some power of it) of the Lippmann–Schwinger (LS) equation belongs to the Hilbert–Schmidt (L^2) class, for complex energies only. It is found that when the energy is made purely real and positive by letting the imaginary part of it vanish, the trace of (K^+K) is unbounded, showing that K does not belong to the L^2 class. This happens even for short-ranged potentials of the Yukawa or cut-off types. This points to the fact that if the two-particle T matrix associated with K is obtained as a solution of the LS equation for complex energies, its limit as the energy becomes real and positive may not tend to a unique value. To circumvent this difficulty, it is often suggested[2-7] that one should examine if the symmetrized form of the LS kernel defined by

$$K_s = V^{1/2}KV^{-1/2} \tag{1.1}$$

belongs to the L^2 class for complex as well as real energies. The associated T matrix, denoted by T_s, is related to that defined for the usual LS equation denoted by T, through the expression

$$T_s = V^{-1/2}T. \tag{1.2}$$

These conditions hold for a class of potentials which obey the conditions:

$$\mathrm{Tr}\{K^{\dagger}K\} = (m^2/2\pi\,|\mathrm{Im}k|)\int d^3r\,|\,V^2(\mathbf{r})\,|^2 < \infty, \tag{1.3}$$

$$\mathrm{Tr}\{K_s^{\dagger}Ks\} = \frac{m^2}{4\pi^2}\int d^3r\int d^3r'\,|V(\mathbf{r})||\,V(\mathbf{r}')\,| \times \frac{e^{-2\,\mathrm{Im}k\,|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|^2} < \infty. \tag{1.4}$$

Here m is the reduced mass, Imk is related to the imaginary part of the energy via the relation $k^2/2m = E + i\epsilon$. Consider for instance, the Yukawa potential $V(r) = V_0e^{-\frac{1}{2}r}/r$ for which both (1.3) and (1.4) hold for Im$k > 0$, while for Im$k = 0$ (1.3) is divergent, whereas (1.4) is not. This indicates that if it can be shown that the T matrix obtained from K and that obtained from K_s are identical for complex energies, then even though (1.3) does not hold for real positive energies, a unique limit of it is assured because of (1.4). We prove this here for a wide class of potentials, including those covered by our earlier work.

In this paper we establish the statements similar to those obtained in I for K, K_l, the partial wave

decomposition of K, and its iterated counterparts, for the corresponding symmetrized kernel K_s. This is given in Sec. 2. Section 3 proves the statements that the T matrixes obtained from K and K_s are identical for those potentials for which both (1.3) and (1.4) hold, and that a unique on-shell limit of T exists for such potentials. The results are discussed in Sec. 4. The notations and definitions employed here are the same as those in I.

2. STATEMENT OF THE THEOREM AND ITS PROOF

Theorem: (a) The kernel K_s of the modified LS equation belongs to the L^2 class provided

$$\int d^3r \int d^3r' |V(\mathbf{r})||V(\mathbf{r}')| e^{-2\,\mathrm{Im}k|\mathbf{r}-\mathbf{r}'|}/|\mathbf{r}-\mathbf{r}'|^2 < \infty.$$

(b) The kernel $K_s^{(l)}$ corresponding to the lth partial wave decomposition of K_s belongs to the L^2 class for spherically symmetric $V(\mathbf{r})$ such that

$$V(r) \xrightarrow[r\to 0]{} r^{-\delta}, \quad \delta < 2, \tag{2.1a}$$

$$V(r) \xrightarrow[r\to\infty]{} r^{-\eta}, \quad \eta > 1. \tag{2.1b}$$

(c) For the potential $V(r) = g/r^\alpha$, $1 < \alpha < 2$,

$$\mathrm{Tr}\{(K_s^\dagger)^m K_s^m\} < \infty$$

provided

$$\alpha > 1 + 1/m, \quad m = 2, 3, \cdots, \tag{2.2}$$

and all statements (a), (b), and (c) are true even when $\mathrm{Im}k = 0$.

Proof: Theorem (a) is well known[2] and therefore will not be proved here. To establish (b) and (c) we proceed as follows:

We have

$$K_s = V^{1/2} G_0^+(E) V^{1/2}, \tag{2.3}$$

$$K_s^\dagger = V^{1/2} G_0^-(E) V^{1/2}, \tag{2.4}$$

$$\mathrm{Tr}\{K_s^\dagger K_s\} = \int d^3r \langle \mathbf{r}|K_s^+ K_s|\mathbf{r}\rangle \tag{2.5}$$

$$= \int d^3r \int d^3r' V(\mathbf{r}) G_0^-(\mathbf{r},\mathbf{r}') G_0^+(\mathbf{r}',\mathbf{r}) V(\mathbf{r}') \tag{2.6}$$

$$= (2\pi)^{-6} \int d^3r \int d^3r' \int \frac{d^3k}{\Delta_k^-} \int \frac{d^3k'}{\Delta_{k'}^+} e^{i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')} \times e^{i\mathbf{k}'\cdot(\mathbf{r}'-\mathbf{r})} V(r) V(r') \tag{2.7}$$

$$= \left(\frac{2}{\pi}\right) \sum_{l=0}^{2\infty} (2l+1) \int_0^\infty r^2 dr\, V(r) \int_0^\infty r'^2 dr'\, V(r') \times \int_0^\infty \frac{k^2 dk}{\Delta_k^-} j_l(kr) j_l(kr') \int_0^\infty \frac{k'^2 dk'}{\Delta_{k'}^+} j_l(k'r) j_l(k'r') \tag{2.8}$$

$$= \left(\frac{2}{\pi}\right)^2 \sum_{l=0}^{\infty} (2l+1) \int_0^\infty \frac{k^2 dk}{\Delta_k^-} \int_0^\infty \frac{k'^2 dk'}{\Delta_{k'}^+} [\langle lk|V|lk'\rangle]^2 \tag{2.9a}$$

$$= \sum_{l=0}^{\infty} (2l+1) \{\mathrm{Tr}(K_s^\dagger K_s)\}_l. \tag{2.9b}$$

It may be shown by the same procedures as in I, that

$$\mathrm{Tr}\{K_s^\dagger K_s\} = 2^3 m^2 |k_+|^2 \sum_{l=0}^{\infty} (2l+1) \int_0^\infty dr r^2 V(r) \times \int_0^\infty dr' r'^2 V(r') |j_l(k_+ r_<) h_l^{(1)}(k_+ r_>)|^2, \tag{2.10}$$

with $k_+ = (2mE + 2mi\epsilon)^{1/2}$. Since $\lim \epsilon \to 0$ exists, we note that k_+ can also be made real. Thus $\mathrm{Im}k_+ \geq 0$. By considering in detail the integral in (2.10), Theorem (b) can be verified.

We now use (2.9) to study the Coulomb-like potentials $V(r) = g/r^\alpha$. Via the procedure developed in I, it is easy to show that

$$\mathrm{Tr}\{(K_s^\dagger)^m K_s^m\} = \left(\frac{2}{\pi}\right)^{2m} \sum_{l=0}^{\infty} (2l+1) \int_0^\infty \frac{k_1^2 dk_1}{\Delta_{k_1}^-} \cdots \int_0^\infty \frac{k_m^2 dk_m}{\Delta_{k_m}^-} \int_0^\infty \frac{k_{m+1}^2}{\Delta_{k_{m+1}}^+} dk_{m+1} \cdots \int_0^\infty \frac{k_{2m}^2}{\Delta_{k_{2m}}^+} dk_{2m} \times \langle lk_1|V|lk_2\rangle \langle lk_2|V|lk_3\rangle \cdots \langle lk_{m-1}|V|lk_m\rangle \cdots \langle lk_{2m}|V|lk_1\rangle. \tag{2.11}$$

We now use Eq. (16) of I to obtain a bound on $\langle lk|V|lk'\rangle$ for $V(r) = g/r^\alpha$ and get after some algebra similar to the one described in I,

$$\{\mathrm{Tr}(K_s^\dagger K_s)\}_l \leq \left[\frac{g\Gamma[\mu + \tfrac{1}{2}(1-\lambda)]\Gamma(2\mu+1)C_1(\lambda)2^{-\lambda/2}}{\Gamma(\mu+1)\Gamma(1+\lambda)\Gamma(\mu+\tfrac{1}{2})\Gamma(\mu+\tfrac{1}{2})}\right]^2 \int_0^\infty \frac{k^2 dk}{|\Delta_k^-|} \int_0^\infty \frac{k'^2 dk'}{|\Delta_{k'}^+|} \left(\frac{kk'}{(k+k')^2}\right)^{2\mu-1} (k+k')^{2\lambda-4} \tag{2.12}$$

$$\leq D^2(\lambda,\mu) 2^{\lambda-2} \int_0^\infty \frac{k^\lambda}{|\Delta_k^-|} dk \int_0^\infty \frac{k'^\lambda dk'}{|\Delta_{k'}^+|} < \infty, \quad \epsilon > 0, \tag{2.13}$$

$D(\lambda, \mu)$ the term inside large brackets in (2.12). All other symbols are defined in I. $\lambda = \alpha - 1$ and so for $1 < \alpha < 2$, we have $0 < \lambda < 1$. The integral

$$\int_0^\infty \frac{k^{\alpha-1}}{\Delta_k^{\pm}}$$

is the straightforward Mellin transform which can be seen to exist even when $\epsilon = 0$ in the sense of the principal value. Therefore the integrals (2.9a) exist even for $\epsilon = 0$. The convergence is uniform in l. In a similar fashion following I, one can show the convergence of the multiple integrals (2.11) for $1 < \alpha < 2$. This can be seen as follows. As $\epsilon \to 0$, the most divergent part of the multiple integral in (2.11) can be isolated and expressed in the form

$$g^{(l)} = \int_{k_0-\xi_0}^{k_0+\xi_0} \frac{k_1^2 dk_1}{\Delta_{k_1}^-} \cdots \int_{k_0-\xi_0}^{k_0+\xi_0} \frac{k_m^2 dk_m}{\Delta_{k_m}^-} \times \int_{k_0-\xi_0}^{k_0+\xi_0} \frac{k_{m+1}^2 dk_{m+1}}{\Delta_{k_{m+1}}^+} \cdots \int_{k_0-\xi_0}^{k_0+\xi_0} \frac{k_{2m}^2 dk_{2m}}{\Delta_{k_{2m}}^+} \times \langle lk_1 | V | lk_2 \rangle \langle lk_2 | V | lk_3 \rangle \cdots \langle lk_{2m} | V | lk_1 \rangle, \tag{2.14}$$

where $k_0 = [2m(E + i\epsilon)]^{1/2}$ and ξ_0 is an arbitrarily small positive definite number, which is set equal to zero after the integration. Now let us define a real function F_l of $2m$ variables:

$$F_l(k_1, k_2, \cdots, k_{2m}) = \langle lk_1 | V | lk_2 \rangle \langle lk_2 | V | lk_3 \rangle \times \cdots \langle lk_{2m} | V | lk_1 \rangle, \tag{2.15}$$

which is regular for $0 < k_n < \infty, n = 1, 2, \cdots, 2m$, for $V(r) = g/r^\alpha$.

Now we write

$$F_l(k_1, k_2, \cdots, k_{2m}) = [F_l(k_1, k_2, \cdots, k_{2m}) - F_l(k_0, k_0, \cdots, k_0)] + F_l(k_0, k_0, \cdots, k_0). \tag{2.16}$$

This leads to the separation of the most singular part g_s in the multiple integral (2.14):

$$g_s = F_l(k_0, k_0, \cdots, k_0) \int_{k_0-\xi_0}^{k_0+\xi_0} k_1^2 \frac{dk_1}{\Delta_{k_1}^-} \cdots \int_{k_0-\xi_0}^{k_0+\xi_0} \frac{k_m^2 dk_m}{\Delta_{k_m}^-} \cdots \int_{k_0-\xi_0}^{k_0+\xi_0} k_{m+1}^2 \frac{dk_{m+1}}{\Delta_{k_{m+1}}^+} \cdots \int_{k_0-\xi_0}^{k_0+\xi_0} \frac{k_{2m}^2 dk_{2m}}{\Delta_{k_{2m}}^+}.$$

But the integral

$$\int_{k_0-\xi_0}^{k_0+\xi_0} \frac{k^2 dk}{\Delta_k^{\pm}}$$

exists with the sense of principal value even for $\epsilon = 0$ and hence $|g_s| < \infty$ for $\epsilon = 0$. The integrations remaining after separating the most singular part given by (2.14) from (2.11) can similarly seen to be convergent for $\epsilon = 0$. This proves the existence of the multiple integrals in (2.11) for $\epsilon = 0$ for the potential $V(r) = g/r^\alpha$ with $1 < \alpha < 2$. One also notices that the matrix element $\langle lk | V | lk' \rangle$ falls off exponentially in l as $l \to \infty$ if $k \neq k'$. When $k = k'$ we can use the asymptotic behavior established in Eq. (19) of I. As in I, we estimate the convergence of the l sum in (2.11) by considering the most divergent part of the multiple k integrals as a function of l. This most divergent part of the l series occurs when $k_1 = k_2 = \cdots k_{2m}$; corresponding to this, the l series is convergent if

$$\alpha > 1 + m^{-1}, \quad m = 2, 3, \cdots.$$

This proves Theorem (c).

3. EVALUATION OF T FOR REAL POSITIVE ENERGIES

From the Theorem it is clear that the class of potentials for which $\mathrm{Tr}\{K_s^\dagger K_s\} < \infty$ is not identical with the class of potentials for which $\mathrm{Tr}\{K^\dagger K\} < \infty$. In addition, even for the potential for which both of the above two traces exist, it is not yet clear that $\lim \langle \mathbf{k} | T(E + i\epsilon) | \mathbf{k}' \rangle$ as $\epsilon \to 0$ is unique and is the same as the one that one would get by solving $\langle \mathbf{k} | V^{1/2} T_s(E + i\epsilon) | \mathbf{k}' \rangle$ for $\epsilon = 0$. Let ϵ be nonzero and $V(r)$ be such that both the kernels K and K_s belong to L^2. Let $T' = V^{1/2} T_s$. Then it is easy to see that $\mathbf{T} = T - T'$ obeys the homogeneous LS. equation:

$$\mathbf{T} = V G_0^+ \mathbf{T} \equiv K\mathbf{T}. \tag{3.1}$$

Since the kernel K belongs to the L^2 class for complex E, the usual Fredholm theory can be applied to the integral equation (3.1). Then we find that (3.1) has nontrivial solutions for some discrete values and $E_n = E < 0$. These correspond to one of the discrete eigensolutions of $H = H_0 + V$. Since in the scattering problem we are interested in

$$\lim_{\epsilon \to 0} \langle \mathbf{k} | T(E + i\epsilon) | \mathbf{k}' \rangle \text{ for } E > 0,$$

we deduce

$$\mathbf{T} \equiv 0, \quad E > 0, \quad \epsilon \neq 0 \tag{3.2}$$

as the solution to (3.1). In fact $\mathbf{T} \equiv 0$ except for negative real E_n where $K = 1$. Thus

$$\langle \mathbf{k} | V^{1/2} T_s(E + i\epsilon) | \mathbf{k}' \rangle \equiv \langle \mathbf{k} | T(E + i\epsilon) | \mathbf{k}' \rangle, \quad \epsilon > 0. \tag{3.3}$$

Since lhs of (3.3) has a well-defined limit for $\epsilon \to 0$ in view of the results of Sec. 2, we can conclude that $\langle \mathbf{k}|V^{1/2}T_s(E+i\epsilon)|\mathbf{k}'\rangle$ gives unique analytic continuation of $\langle \mathbf{k}|T(E+i\epsilon)|\mathbf{k}'\rangle$ on the positive real axis.

4. DISCUSSION

We wish to stress the fact that the proof given above for the existence of a unique limit for $\langle \mathbf{k}|T(E+i\epsilon)|\mathbf{k}'\rangle$ as $\epsilon \to 0$ is valid only for that $V(r)$ for which both K_s and K belong to L^2 class. In I we showed that for $V(r) = g/r^\alpha$, the partial wave LS equation has L^2 kernel for $\frac{1}{2} < \alpha < \frac{3}{2}$. On the other hand, the symmetrized partial wave LS equation has L^2 kernel if $1 < \alpha < 2$. Therefore the common region for which both have L^2 kernel is for $1 < \alpha < \frac{3}{2}$. Thus, for $1 < \alpha < \frac{3}{2}$ with $V(r) = g/r^\alpha$, a unique on-shell limit of the partial wave T matrix exists. It can be noted that the Coulomb potential $V_c(r) = g/r$ is not covered by our analysis. This is not surprising in view of the fact that plane wave representation of the off-shell Coulomb T matrix has no unique limit as one approaches $\epsilon \to 0$[8].

The full symmetrized LS equation requires twice as many iterations as the usual LS equation in order to have L^2 solutions.

Another aspect to be emphasized is that $G_0^+V \Rightarrow V^{1/2}G_0^+V^{1/2}$ is not the only way to obtain a modified LS equation. In fact Coester[4] shows that one can in general express $V = V^+V^-$ and obtain the correspondingly modified LS equations. As an example let $V = V^\beta V^{1-\beta}$; then the modified LS equation:

$$[V^{1-\beta}G^+] = V^{1-\beta}G_0^+ + \{V^{1-\beta}G_0^+V^\beta\}[V^{1-\beta}G^+].$$

The evaluation of various traces corresponding to $K_s^\beta = V^{1-\beta}G_0^+V^\beta$ will lead to the same conclusions that we obtain from K_s for which $\beta = \frac{1}{2}$. In the differential equation method employed in potential scattering, the constraint on the potential is $\int_0^\infty r|V(r)|dr < \infty$. It may be pointed out that potentials obeying this condition are incorporated in Theorem (b), whereas the converse is not true. For example, $V(r) = g/r^\alpha$, $1 < \alpha < 2$ satisfies the requirements of Theorem (b), whereas the integral $\int_0^\infty r|V(r)|dr$ is not convergent for this potential.

There is another method to construct the modified LS equation, by employing the kernel[5]

$$K_m = G_0^{+1/2}VG_0^{+1/2}. \tag{4.1}$$

For $\epsilon \neq 0$, one can obtain for spherically symmetric potentials the result

$$\{\mathrm{Tr}(K_m^\dagger K_m)\} = \left(\frac{2}{\pi}\right)^2 \sum_{l=0}^{\infty}(2l+1)\int \frac{k^2dk}{|\Delta_k^+|}\int \frac{k'^2dk'}{|\Delta_{k'}^+|} \times [\langle lk|V|lk'\rangle]^2. \tag{4.2}$$

We notice that (4.2) differs from (2.9) in the presence of $|\Delta_k^+|$ and $|\Delta_{k'}^+|$ instead of Δ_k^- and $\Delta_{k'}^+$. Due to this difference (4.2) diverges when $\epsilon = 0$ (the principal value integral does not exist), whereas for $\epsilon = 0$ (2.9) exists in the sense of principal value. However for $\epsilon > 0$, the existence of $Tr\{K_m^\dagger K_m\}$ has the same restrictions on the potential $V(r)$ as stated in (2.1a) and (2.1b).

Note added in proof: In I and the present paper the specific case of $E = 0$ was not investigated. This case was studied subsequently and we summarize the results below

The existence of the limit $E \to 0$ (from the negative side) for the trace

$$\mathrm{Tr}(K^\dagger K)_l = \left(\frac{2}{\pi}\right)^2 \int_0^\infty r^2dr \int_0^\infty \frac{k^2dk}{|\Delta_k^+|^2} j_l^2(kr)V^2(r)$$

can be obtained by explicit evaluation. Thus (with $\delta = 0$)

$$\mathrm{Tr}(K^\dagger K)_l = -\left(\frac{2}{\pi}\right)^2 4m^2 \int_0^\infty r^2V^2(r)dr \frac{d}{d\beta^2}\left[\int_0^\infty \frac{k^2dkj_l^2(kr)}{(k^2+\beta^2)}\right]$$

$$= -\left(\frac{2}{\pi}\right)^2 4m^2 \int_0^\infty r^2V^2(r)dr \times\left[\frac{1}{2r}\frac{d}{d\beta^2}\{I_{l+1/2}(r\beta)K_{l+1/2}(r\beta)\}\right]$$

Here $\beta^2 = -2mE > 0$, and $I_{l+1/2}$ and $K_{l+1/2}$ are the modified Bessel functions of first and third kind, respectively. Via the properties of these functions, it is found that, for $l = 0$, $\mathrm{Tr}(K^\dagger K)_l$ diverges like β^{-1} as $\beta \to 0$ (i.e., $E \to 0$) and, for $l = 1, 2, 3, \cdots$,

$$\lim_{E\to 0}\{\mathrm{Tr}(K^\dagger K)_l\}$$

exists provided that

$$V(r) \underset{r\to 0}{\longrightarrow} r^{-\delta}, \quad \delta < 2$$

and

$$V(r) \underset{r\to 0}{\longrightarrow} r^{-\eta}, \quad \eta > \tfrac{1}{2}.$$

Similarly, from the study of (2.16) of the present paper, $\mathrm{Tr}\{K_s^\dagger K_s\}_l$ can be shown to exist as $k_+ \to 0$ (i.e., as $E \to 0$), provided that $V(r)$ satisfies the conditions (2.1a) and (2.1b).

---

* Address after September 1971: Department of Physics, Birla Institute of Technology and Science, Pilani, Rajasthan, India.

1 A. K. Rajagopal and C. S. Shastry, J. Math. Phys. (to be published).
2 R. G. Newton, *Scattering Theory of Waves and Particles* (McGraw-Hill, New York, 1966).
3 S. Weinberg, Phys. Rev. **131**, 440 (1963).
4 F. Coester, Phys. Rev. **133**, B1516 (1964).

[5] M. Scadron, S. Weinberg, and J. Wright, Phys. Rev. 135, B202 (1964).
[6] A. Grossman and T. Wu, J. Math. Phys. 2, 710 (1961).
[7] K. Meetz, J. Math. Phys. 3, 690 (1962).
[8] A. K. Rajagopal and C. S. Shastry (unpublished).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 10 OCTOBER 1971

Asymptotic Form of the Wavefunction for Neutral Particle–Hydrogen Atom Scattering*

G. Doolen
Texas A&M University, College Station, Texas 77843
and
J. Nuttall
University of California, Los Alamos Scientific Laboratory, Los Alamos, New Mexico 87544
and
Texas A&M University, College Station, Texas 77843†

(Received 11 March 1971)

The asymptotic form of the wavefunction for the scattering of a neutral particle on a hydrogen atom bound state is discussed, with particular reference to the case where the two charged particles may be close together in the asymptotic region.

I. INTRODUCTION

In spite of the great progress that has been made in the past decade in understanding the mathematics of multiparticle scattering problems, there are still problems that remain when long-range Coulomb forces are involved.

If we look at e-H scattering (for example) in coordinate space, we need to impose suitable boundary conditions before the solution of Schrödinger's equation is uniquely specified. Let us use a six-dimensional vector $\hat{\rho} = (\mathbf{r}_1, \mathbf{r}_2)$ (with $\mathbf{r}_1, \mathbf{r}_2$ the positions of the electrons). Then the only boundary in the problem is a five-dimensional hypersphere at infinity, i.e., $\rho \equiv (r_1^2 + r_2^2)^{1/2} \to \infty$. The boundary conditions involve specifying the form of the wavefunction $\Phi(\hat{\rho})$ as $\rho \to \infty$.

The asymptotic form of $\Phi(\hat{\rho})$ has already been given by Peterkop[1] and others[2] for almost all the hypersphere, and is, for a wavefunction describing scattering from the ground state of hydrogen above the ionization threshold,

$$\Phi(\hat{\rho}) \underset{\rho\to\infty}{\sim} \frac{e^{i\sqrt{E}\rho}}{\rho^{5/2}} \rho^{i\eta(\hat{\rho}_u)} f_0(\hat{\rho}_u), \tag{1}$$

where $f_0(\hat{\rho}_u)$ is proportional to the amplitude for ionization with the electrons emitted in directions specified by the unit vector $\hat{\rho}_u = \hat{\rho}/\rho$. The function $\eta(\hat{\rho}_u)$ is related to the total energy E and the complete potential V by

$$\eta(\hat{\rho}_u) = \lim_{\rho\to\infty} -\rho V/2\sqrt{E}. \tag{2}$$

Form (1) is valid on all but three small portions of the hypersphere corresponding to the places where a pair of particles is close together.

It is the purpose of this paper to describe the asymptotic behavior of $\Phi(\hat{\rho})$ in those regions where two charged particles are relatively near each other. An understanding of the region where particles 1 and 3 approach each other, with a Coulomb interaction, is an essential step in understanding the behavior of three charged particles. For simplicity, instead of e-H scattering, we here study a model in which V_3, V_2 (the potentials between 12, 23) are short range, e.g., a superposition of Yukawa potentials, but V_1 is a pure Coulomb potential $-2c/r_1$, taken to be attractive for the following discussion ($c > 0$). Particle 3 is infinitely heavy and $m_1 = m_2 = 1/2$.

A principal result of our work is that (1) becomes invalid if the inequality

$$r_1^3 E >> 2cr_2^2 \tag{3}$$

is not satisfied. We call the region where (3) holds the outer region and its complement the inner region.[3] In the inner region, Φ is described by the well-known sum over the bound states of the hydrogen atom, with coefficients proportional to the amplitudes for elastic scattering and excitation. In addition there is a continuum contribution which Peterkop[1] has written as an integral over two-body Coulomb scattering wavefunctions $\psi_q(\mathbf{r}_2)$. We evaluate this continuum contribution in the inner region, and in particular find that, for fixed r_1, it falls off as r_2^{-2}. We also find that, in the outer region, the bound-state sum falls off no faster than $r_2^{-1} r_1^{-9/4}$, but that the continuum contribution cancels this term in such a way that (1) and the corrections to it given by Peterkop[1] are the only parts of Φ that do not decay exponentially.

II. BASIC EQUATIONS

Let us suppose that, with the model we have described, particle 2 is incident on a bound state of particles 1 and 3, so that the total energy is E. This initial state will have a wavefunction $\chi(\mathbf{r}_1, \mathbf{r}_2)$ which is a product of a plane wave for particle 2 and the bound-state wavefunction. The three-body

[5] M. Scadron, S. Weinberg, and J. Wright, Phys. Rev. 135, B202 (1964).
[6] A. Grossman and T. Wu, J. Math. Phys. 2, 710 (1961).
[7] K. Meetz, J. Math. Phys. 3, 690 (1962).
[8] A. K. Rajagopal and C. S. Shastry (unpublished).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 10 OCTOBER 1971

Asymptotic Form of the Wavefunction for Neutral Particle–Hydrogen Atom Scattering*

G. Doolen
Texas A&M University, College Station, Texas 77843
and
J. Nuttall
University of California, Los Alamos Scientific Laboratory, Los Alamos, New Mexico 87544
and
Texas A&M University, College Station, Texas 77843†

(Received 11 March 1971)

The asymptotic form of the wavefunction for the scattering of a neutral particle on a hydrogen atom bound state is discussed, with particular reference to the case where the two charged particles may be close together in the asymptotic region.

I. INTRODUCTION

In spite of the great progress that has been made in the past decade in understanding the mathematics of multiparticle scattering problems, there are still problems that remain when long-range Coulomb forces are involved.

If we look at e-H scattering (for example) in coordinate space, we need to impose suitable boundary conditions before the solution of Schrödinger's equation is uniquely specified. Let us use a six-dimensional vector $\hat{\rho} = (\mathbf{r}_1, \mathbf{r}_2)$ (with $\mathbf{r}_1, \mathbf{r}_2$ the positions of the electrons). Then the only boundary in the problem is a five-dimensional hypersphere at infinity, i.e., $\rho \equiv (r_1^2 + r_2^2)^{1/2} \to \infty$. The boundary conditions involve specifying the form of the wavefunction $\Phi(\hat{\rho})$ as $\rho \to \infty$.

The asymptotic form of $\Phi(\hat{\rho})$ has already been given by Peterkop[1] and others[2] for almost all the hypersphere, and is, for a wavefunction describing scattering from the ground state of hydrogen above the ionization threshold,

$$\Phi(\hat{\rho}) \underset{\rho\to\infty}{\sim} \frac{e^{i\sqrt{E}\rho}}{\rho^{5/2}} \rho^{i\eta(\hat{\rho}_u)} f_0(\hat{\rho}_u), \tag{1}$$

where $f_0(\hat{\rho}_u)$ is proportional to the amplitude for ionization with the electrons emitted in directions specified by the unit vector $\hat{\rho}_u = \hat{\rho}/\rho$. The function $\eta(\hat{\rho}_u)$ is related to the total energy E and the complete potential V by

$$\eta(\hat{\rho}_u) = \lim_{\rho\to\infty} -\rho V/2\sqrt{E}. \tag{2}$$

Form (1) is valid on all but three small portions of the hypersphere corresponding to the places where a pair of particles is close together.

It is the purpose of this paper to describe the asymptotic behavior of $\Phi(\hat{\rho})$ in those regions where two charged particles are relatively near each other. An understanding of the region where particles 1 and 3 approach each other, with a Coulomb interaction, is an essential step in understanding the behavior of three charged particles. For simplicity, instead of e-H scattering, we here study a model in which V_3, V_2 (the potentials between 12, 23) are short range, e.g., a superposition of Yukawa potentials, but V_1 is a pure Coulomb potential $-2c/r_1$, taken to be attractive for the following discussion ($c > 0$). Particle 3 is infinitely heavy and $m_1 = m_2 = 1/2$.

A principal result of our work is that (1) becomes invalid if the inequality

$$r_1^3 E >> 2cr_2^2 \tag{3}$$

is not satisfied. We call the region where (3) holds the outer region and its complement the inner region.[3] In the inner region, Φ is described by the well-known sum over the bound states of the hydrogen atom, with coefficients proportional to the amplitudes for elastic scattering and excitation. In addition there is a continuum contribution which Peterkop[1] has written as an integral over two-body Coulomb scattering wavefunctions $\psi_q(\mathbf{r}_2)$. We evaluate this continuum contribution in the inner region, and in particular find that, for fixed r_1, it falls off as r_2^{-2}. We also find that, in the outer region, the bound-state sum falls off no faster than $r_2^{-1} r_1^{-9/4}$, but that the continuum contribution cancels this term in such a way that (1) and the corrections to it given by Peterkop[1] are the only parts of Φ that do not decay exponentially.

II. BASIC EQUATIONS

Let us suppose that, with the model we have described, particle 2 is incident on a bound state of particles 1 and 3, so that the total energy is E. This initial state will have a wavefunction $\chi(\mathbf{r}_1, \mathbf{r}_2)$ which is a product of a plane wave for particle 2 and the bound-state wavefunction. The three-body

wavefunction $\Phi(\mathbf{r}_1, \mathbf{r}_2)$ may be obtained from

$$|\Phi\rangle = |\chi\rangle + \lim_{\mathrm{Im}W \to 0} (W - H)^{-1}(V_2 + V_3)|\chi\rangle. \quad (4)$$

Here W is the total energy with real part E and positive imaginary part, and $H = T + V_1 + V_2 + V_3$ is the Hamiltonian for the problem. We work with this equation before the limit is taken.

Equation (4) can be rearranged to read

$$|\Phi\rangle = |\chi'\rangle + (W - T - V_1)^{-1}(V_2 + V_3)|\Phi\rangle. \quad (5)$$

In the coordinate representation, (5) may be written as

$$\begin{aligned}\Phi(\mathbf{r}_1,\mathbf{r}_2) &= \chi'(\mathbf{r}_1,\mathbf{r}_2) + (1/8\pi^3)\int d\mathbf{q}d\mathbf{p}\psi_{\mathbf{q}}^{-}(\mathbf{r}_1)\exp(i\mathbf{p}\cdot\mathbf{r}_2)\\ &\times (W - p^2 - q^2)^{-1}T(\mathbf{q},\mathbf{p})\\ &+ (1/8\pi^3)\sum_{LMN}\int d\mathbf{p}\psi_{NLM}(\mathbf{r}_1)\exp(i\mathbf{p}\cdot\mathbf{r}_2)\\ &\times (W - p^2 + B_N)^{-1}T_{NLM}(\mathbf{p}), \quad (6)\end{aligned}$$

with

$$\begin{aligned}T(\mathbf{q},\mathbf{p}) &= \int d\mathbf{r}_1 d\mathbf{r}_2 \psi_{\mathbf{q}}^{-*}(\mathbf{r}_1)\exp(-i\mathbf{p}\cdot\mathbf{r}_2)\\ &\times (V_2 + V_3)\Phi(\mathbf{r}_1,\mathbf{r}_2) \quad (7)\end{aligned}$$

and

$$\begin{aligned}T_{NLM}(\mathbf{p}) &= \int d\mathbf{r}_1 d\mathbf{r}_2 \psi_{NLM}^{*}(\mathbf{r}_1)\exp(-i\mathbf{p}\cdot\mathbf{r}_2)\\ &\times (V_2 + V_3)\Phi(\mathbf{r}_1,\mathbf{r}_2). \quad (8)\end{aligned}$$

Here $\psi_{NLM}(\mathbf{r}_1)$ are the normalized Coulomb bound-state wavefunctions and $\psi_{\mathbf{q}}^{-}(\mathbf{r}_1)$ are the corresponding continuum functions for in-going scattered waves. The inhomogeneous term χ' approaches χ as $W \to 0$.

In our model problem in which V_2 and V_3 are of short range, we shall see that, for fixed $\mathbf{q}$, $T(q,p)$ and $T_{NLM}(\mathbf{p})$ are analytic functions of $\mathbf{p}$ for real $\mathbf{p}$ near $p(q) = (E - q^2)^{1/2}$ or near $p_N = (E + B_N)^{1/2}$. Consequently, for large r_2, we may approximate (6) by an extension of the method of steepest descents, described in more detail for this case elsewhere,[4] with the result that

$$\begin{aligned}\Phi(\mathbf{r}_1,\mathbf{r}_2) &\sim \chi'(\mathbf{r}_1,\mathbf{r}_2) - \frac{1}{4\pi}\int d\mathbf{q}\psi_{\mathbf{q}}^{-}(\mathbf{r}_1)\\ &\times \frac{\exp[ip(q)r_2]}{r_2} T(\mathbf{q}, p(q)\hat{\mathbf{r}}_2)\\ &- \frac{1}{4\pi}\sum_{NLM}\psi_{NLM}(\mathbf{r}_1)\frac{\exp(ip_N r_2)}{r_2} T_{NLM}(p_N\hat{\mathbf{r}}_2). \quad (9)\end{aligned}$$

It should be noted that this formula is not valid in regions $r_2/\rho \to 0$ or $|\mathbf{r}_1 - \mathbf{r}_2|/\rho \to 0$.

Equation (9) may be further simplified by the method of steepest descents, but care must be taken of the singularity that the integrand contains at $q = 0$, both in $\psi_{\mathbf{q}}^{-}(\mathbf{r}_1)$ and in $T(\mathbf{q}, p(q)\hat{\mathbf{r}}_2)$. The behavior of $T(q,p)$ near $q = 0$ may be studied from its definition [Eq. (7)]. An important point, which is the basis of our whole analysis, is that apart from two exceptional regions, not near $\mathbf{q} = 0$, we may assume that the integral in (7) is over a bounded region when studying the analyticity of $T(\mathbf{q},\mathbf{p})$. This is due to the fact that we know enough about the phase of Φ for large ρ (dominant term $\sqrt{W}\rho$) to see that at no point where $V_2 + V_3$ is not exponentially damped is the phase of the integrand stationary for large ρ.

At this point it is convenient to perform a partial wave resolution in the $\hat{\mathbf{q}}$ and $\hat{\mathbf{r}}_1$ variables to obtain a partial wave analysis of $T(\mathbf{q},\mathbf{p})$.

We set $\psi_{\mathbf{q}}^{-}(\mathbf{r}_1) = \sum_{LM} Y_L^M(\hat{\mathbf{r}}_1)Y_L^{M*}(\hat{\mathbf{q}})q^L e^{-n\pi/2}\Gamma(L + 1 - in)\mathfrak{F}_L^*(q, r_1)$,

where $n = -c/q$, with c positive in the attractive case under discussion.

The function $\mathfrak{F}_L(q, r_1)$ is defined in terms of Kummer's function[5] M by the relation

$$\begin{aligned}\mathfrak{F}_L(q, r_1) &= i^L\left(\frac{2}{\pi}\right)^{1/2}\frac{(2r_1)^L}{\Gamma(2L + 2)}\, e^{-iqr_1}\\ &\times M(L + 1 - in, 2L + 2, 2iqr_1)\end{aligned}$$

From (7), we have

$$\begin{aligned}T(\mathbf{q},\mathbf{p}) &= \sum_{LM} Y_L^M(\hat{q})(-q)^L e^{-n\pi/2}\\ &\times \Gamma(L + 1 + in)t_{LM}(q,\mathbf{p}), \quad (10)\end{aligned}$$

with

$$\begin{aligned}t_{LM}(q,\mathbf{p}) &= \int d\mathbf{r}_1 d\mathbf{r}_2 Y_L^{M*}(\hat{\mathbf{r}}_1)\mathfrak{F}_L(q, r_1)\\ &\times e^{-\mathbf{p}\cdot\mathbf{r}_2}(V_2 + V_3)\Phi(\mathbf{r}_1,\mathbf{r}_2).\end{aligned}$$

The function $\mathfrak{F}_L(q, r)$ [and thus $t_{LM}(q,\mathbf{p})$ from the argument above] is an even function of q and is analytic in q near $q = 0$.

The behavior of $T(\mathbf{q},\mathbf{p})$ near $q = 0$ is governed by the factors given explicitly in (10).

We can use (9) and (10) to write the asymptotic form in a more compact notation. The continuum contribution to (9) is

$$\begin{aligned}&-\frac{1}{4\pi}\sum_{LM}(-1)^L\int_0^\infty dq q^{2L+2}e^{-n\pi}\Gamma(L + 1 - in)\\ &\times \Gamma(L + 1 + in)\mathfrak{F}_L(q, r_1)Y_L^M(\hat{\mathbf{r}}_1)\\ &\times \frac{e^{ip(q)r_2}}{r_2}t_{LM}(q, p(q)\hat{\mathbf{r}}_2). \quad (11)\end{aligned}$$

The integrand has poles at $L + 1 - in = L + 1 - N$, $N \geq L + 1$, or $q = -ic/N$.

We note that if $t_{LM}(q,\mathbf{p})$ is evaluated at $q = -ic/N$ where N is a positive integer and $N \geq L + 1$, it will be found to be related to the excitation amplitude T_{NLM} by

$$\begin{aligned}T_{NLM}(\mathbf{p}) &= \left(\frac{2\pi\Gamma(N + L + 1)}{\Gamma(N - L)}\right)^{1/2} i^{-L} c^{L+3/2} N^{-L-2}\\ &\times t_{LM}\left(\frac{-ic}{N},\mathbf{p}\right), \quad (12)\end{aligned}$$

$$\psi_{NLM}(\mathbf{r}) = \left(\frac{2\pi\Gamma(N+L+1)}{\Gamma(N-L)}\right)^{1/2} i^{-L} c^{L+3/2} N^{-L-2} \times \mathfrak{F}_L\left(\frac{-ic}{N}, r\right) Y_L^M(\hat{\mathbf{r}}). \quad (13)$$

Thus the behavior of $T_{NLM}(\mathbf{p})$ for large N is related to the behavior of the continuum scattering amplitude $t_{LM}(q,\mathbf{p})$ for small q.

Using (12) and (13), we may show that if the integral in (11) is taken on a contour consisting of a small circle about one of these poles, then the result is exactly the corresponding bound-state contribution to (9),

$$-\frac{1}{4\pi}\sum_{LM}\psi_{NLM}(\mathbf{r}_1)\frac{e^{ip_N r_2}}{r_2}T_{NLM}(p_N\hat{\mathbf{r}}_2).$$

Thus we may replace the sum over N by an integral over q and we write (9) as

$$\Phi(\mathbf{r}_1,\mathbf{r}_2) \sim \chi'(\mathbf{r}_1,\mathbf{r}_2) - \frac{1}{4\pi}\sum_{LM}\int_\alpha dq q^{2L+2} e^{-n\pi} \times \Gamma(L+1-in)\Gamma(L+1+in) \times \mathfrak{F}_L(q,r_1)Y_L^M(\hat{\mathbf{r}}_1)\frac{e^{ip(q)r_2}}{r_2}t_{LM}(q,p(q)\hat{\mathbf{r}}_2),$$

where the contour α is shown in Fig. 1.

It is now useful to introduce the irregular Whittaker function[5] and write

$$\mathfrak{F}_L(q,r) = \left(\frac{2}{\pi}\right)^{1/2}\frac{q^{-L-1}}{2ir}e^{n\pi} \times \left(\frac{W_{-in,L+1/2}(-2iqr)}{\Gamma(L+1-in)} + (-1)^{L+1}\frac{W_{in,L+1/2}(2iqr)}{\Gamma(L+1+in)}\right). \quad (14)$$

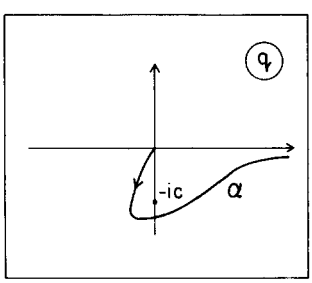


FIG. 1. The contour α used to include the bound-state sum.

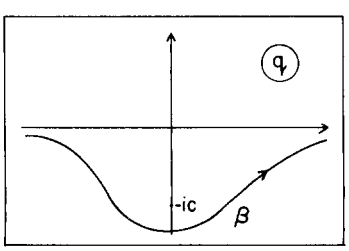


FIG. 2. The contour β used with the Whittaker function to include the bound-state sum.

Although this does not appear to be an even function of q, it must be remembered that $W_{k,m}(z)$ has a cut from $z = 0$ to $-\infty$. As we take q along a path from real positive to real negative values either above or below the origin, the argument of one of the W functions crosses its cut.

In view of (14) we can rewrite (11), the continuum contribution to (9), as

$$-2^{-5/2}\pi^{-3/2}i\sum_{LM}\int_0^\infty dq q^{L+1}[\Gamma(L+1-in) \times W_{in,L+1/2}(2iqr_1) + (-1)^{L+1} \times \Gamma(L+1+in)W_{-in,L+1/2}(-2iqr_1)] \times Y_L^M(\hat{\mathbf{r}}_1)\frac{e^{ip(q)r_2}}{r_2 r_1}t_{LM}(q,p(q)\hat{\mathbf{r}}_2) \quad (15)$$

or

$$-2^{-5/2}\pi^{-3/2}i\sum_{LM}\int_{-\infty}^\infty dq q^{L+1}\Gamma(L+1-in) \times W_{in,L+1/2}(2iqr_1)Y_L^M(\hat{\mathbf{r}}_1)\frac{e^{ip(q)r_2}}{r_1 r_2}t_{LM}(q,p(q)\hat{\mathbf{r}}_2). \quad (16)$$

Since the first term of (14) is zero at the bound-state poles $q = -ic/N$, we can include the contributions of the bound-state poles to Φ by integrating (16) along the contour β shown in Fig. 2:

$$\Phi(\mathbf{r}_1,\mathbf{r}_2) \sim \chi'(\mathbf{r}_1,\mathbf{r}_2) - 2^{-5/2}\pi^{-3/2}i\sum_{LM}\int_\beta dq q^{L+1} \times \Gamma(L+1-in)W_{in,L+1/2}(2iqr_1) \times Y_L^M(\hat{\mathbf{r}}_1)\frac{e^{ip(q)r_2}}{r_1 r_2}t_{LM}(q,p(q)\hat{\mathbf{r}}_2). \quad (17)$$

III. ASYMPTOTIC FORM IN OUTER REGION

It is obvious that the asymptotic form (1) cannot be valid for all values of r_1/r_2, however small, since $\eta \to \frac{1}{2}E(r_2/r_1)$ as $r_1 \to 0$. This form does not even satisfy Schrödinger's equation in this region.[6] It is of interest to determine how small r_1/r_2 may be taken before (1) must be modified. Also, it might be asked whether the bound-state sum in (9) gives a contribution to the asymptotic form in the region where (1) is valid, since as $N \to \infty$, $\psi_{NLM}(\mathbf{r}_1)$ is nonnegligible for increasing values of r_1.

We show below that, provided $r_1^3 E \gg 2cr_2^2$, the leading term of the asymptotic expansion is indeed (1), and corrections are given by the expansion of Peterkop.[1] This is done from a consideration of representation (17) of the asymptotic form which includes both bound-state and continuum contributions. In Sec. IV it will be shown that the

continuum contribution alone leads to a correction to (1) in the outer region which does not fall off exponentially. We are therefore led to the conclusion that the bound-state sum does not fall off exponentially in the outer region, but must in fact cancel the anomalous part from the continuum term.

Let us choose the contour β used in (17) to follow the path shown in Fig. 3. If we assume that r_1 is such that, for all points on β, $|q^2(r_1/2c)| \gg 1$, then we may use the asymptotic form of the Whittaker function,[5]

$$W_{in,L+\frac{1}{2}}(2iqr_1) \sim \exp[-iqr_1 - (ic/q)\ln 2qr_1 + c\pi/2q].$$

The dominant part of the phase of the integrand in (17) is $ir_2(E-q^2)^{1/2} - iqr_1$, which leads to the saddlepoint at $-q_0 = -r_1\sqrt{E}(r_1^2+r_2^2)^{-1/2}$ through which β passes. It may be shown that, as we move along β away from $-q_0$ in either direction, the integrand falls off exponentially, even though the contour passes close to the poles of the Γ function. For large ρ the width of the saddle at $-q_0$ is much less than the distance from $-q_0$ to the point B in Fig. 3. Thus the entire contribution to the asymptotic form comes from the saddlepoint, which leads to

$$\Phi(\mathbf{r}_1,\mathbf{r}_2) \underset{\rho\to\infty}{\sim} \chi'(\mathbf{r}_1,\mathbf{r}_2) + \frac{i^{1/2}E^{3/4}}{4\pi\rho^{5/2}} \times \exp\left(iq_0r_1 + \frac{ic}{q_0}\ln 2q_0r_1 + i\sqrt{E}\frac{r_2^2}{\rho}\right) \times T\left(\mathbf{r}_1\frac{\sqrt{E}}{\rho}, \mathbf{r}_2\frac{\sqrt{E}}{\rho}\right), \quad (18)$$

which agrees with (1) if we set

$$f_0(\hat{p}_u) = \frac{i^{1/2}E^{3/4}}{4\pi}\exp\left[\frac{ic}{q_0}\ln(2q_0^2/\sqrt{E})\right] \times T\left(\mathbf{r}_1\frac{\sqrt{E}}{\rho}, \mathbf{r}_2\frac{\sqrt{E}}{\rho}\right) \quad (19)$$

The above argument is valid only when $(q_0^2r_1/2c) \gg 1$, which is equivalent to $r_1^3E \gg 2cr_2^2$. This relation describes the boundary between the outer region where all particles are well separated and (18) holds, and the inner region where particles 1 and 3 are close enough together to have an interaction that significantly changes the nature of the asymptotic form. Note that, as $\rho\to\infty$, the inner region subtends a vanishingly small solid angle at the origin of the six-dimensional space.

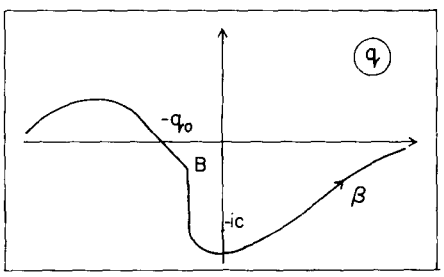


FIG. 3. The version of contour β used in the discussion of asymptotic form in the outer region. The point B is $-\frac{1}{2}[q_0(1+i)]$.

It would be interesting if the location of the boundary we have found between inner and outer regions could be understood in classical terms. In this connection, we point out that the inequality $r_2/\sqrt{E} \ll r_1^{3/2}(2c)^{-1/2}$ states that the time taken for particle 2 travelling at velocity $\sqrt{E}$ to reach the origin from r_2 is much less than the time taken by particle 1 to complete an orbit with average radius r_1. We do not understand the significance of this.

IV. CONTINUUM CONTRIBUTION TO THE ASYMPTOTIC FORM

We now discuss the form of the continuum contribution. We shall see that the integral in (15) may be simplified, rather as is the case with short-range forces. Unfortunately, we do not know of any comparable approximation to the bound-state sum in the inner region.

In the region r_1 fixed, $r_2\to\infty$, it is best to use the representation (11) of the continuum contribution. Provided that $|\arg q| < \frac{1}{2}\pi$ we have, as $|q|\to 0$,

$$e^{-n\pi}\Gamma(L+1-in)\Gamma(L+1+in)\mathfrak{F}_L(q,r_1) \sim -(2\sqrt{\pi}/\sqrt{r_1})i^Lq^{-(2L+1)}c^{L+1/2}J_{2L+1}(-(8r_1c)^{1/2}), \quad (20)$$

where J_{2L+1} is a Bessel function. The rapidly varying part of the phase of the integrand in (11) in this region is $p(q)r_2 = r_2(E-q^2)^{1/2}$, which gives rise to a saddlepoint at $q=0$. To achieve maximum fall-off from $q=0$, we must rotate the contour near $q=0$ through $\frac{1}{4}\pi$ in a clockwise direction. Equation (19) is therefore valid. Thus we find that the continuum contribution Φ_c has the form

$$\Phi_c(r_1,r_2) \sim \tfrac{1}{2}\pi^{1/2}\frac{\sqrt{E}}{r_2^2r_1^{1/2}}\exp(i\sqrt{E}r_2) \times \sum_{LM}(-c)^{L+1/2}i^LY_L^M(\hat{r}_1)J_{2L+1}[-(8r_1c)^{1/2}] \times t_{LM}(0,\sqrt{E}\hat{\mathbf{r}}_2), \text{ for } r_2\to\infty \text{ and } r_1 \text{ fixed.} \quad (21)$$

This result disagrees with a formula given by Peterkop[7], which is based on an oversimplified analogy with the case of short-range forces.

To study the continuum contribution for larger values of r_1, it is convenient to use the representation (15). We use WKB-type approximation to the

Whittaker function[8], valid in the region $|\arg q| < \frac{1}{2}\pi$, r_1 large:

$$W_{\pm in,L+1/2}(\pm 2iqr_1) \sim (1+\alpha^{-1})^{-1/4}\exp[\pi c/2q \pm i\theta],$$

$$\theta = \frac{-c}{q}\ln\frac{c}{q} + \frac{c}{q} - \frac{2c}{q}\{[\alpha(1+\alpha)]^{1/2} - \ln[(1+\alpha)^{1/2} - \alpha^{1/2}]\}, \quad (22)$$

with

$$\alpha = q^2 r_1/2c.$$

The phase in (22) is too complicated to allow us to obtain in closed form the location of the saddlepoints of the two terms in the integrand of (15). However, it may be seen that the first term in (15) has its only saddlepoint at $q = 0$. Using the expansion of θ about $q = 0$, which reads

$$\theta \approx -4c\left(\frac{r_1}{2c}\right)^{1/2} - \frac{2c}{3}q^2\left(\frac{r_1}{2c}\right)^{3/2} + \frac{c}{10}q^4\left(\frac{r_1}{2c}\right)^{5/2} \quad (23)$$

to evaluate the saddlepoint integral, we obtain for the contribution Φ_{1c} of the first term in (15)

$$\Phi_{1c} \sim \frac{\exp[i\sqrt{E}r_2 - i(8rc)^{1/2}]\sum_{LM}(ic)^{L+1/2}}{8\pi r_2 r_1^{3/4}(2c)^{1/4}[r_2/2\sqrt{E} + \frac{2}{3}c(r_1/2c)^{3/2}]} \times Y_L^M(\hat{\mathbf{r}}_1)t_{LM}(0,\sqrt{E}\hat{\mathbf{r}}_2), \text{ for } r_2 \to \infty \text{ and } r_1 \text{ large.}$$

The second term in (15) also has a saddlepoint fixed at $q = 0$, but in addition there is a moving saddlepoint that lies at q_0 for $r_1^3 E \gg 2cr_2^2$. This saddlepoint gives rise to the usual term (1), while in this region the $q = 0$ point gives a contribution Φ_{2c},

$$\Phi_{2c} \sim -\frac{3i(2c)^{1/4}\exp[i\sqrt{E}r_2 - i(8r_1c)^{1/2}]}{8\pi r_2 r_1^{9/4}} \times \sum_{LM}(ic)^{L+1/2}Y_L^M(\hat{\mathbf{r}}_1)t_{LM}(0,\sqrt{E}\hat{\mathbf{r}}_2), \text{ for } r_2 - \infty \text{ and } r_1^3 E \gg 2cr_2^2.$$

As the ratio $r_1^3 E/(2cr_2^2)$ decreases from large values, the moving saddlepoint approaches $q = 0$, which it reaches when $r_1^3 E/(2cr_2^2) = \frac{9}{4}$. In this region, we need to include the fourth-order term in (23) when evaluating the second term of (15), with the result that

$$\Phi_{2c} \sim \frac{1}{16\pi}\left(\frac{\pi}{i\beta}\right)^{1/2} \times \frac{\exp[i\sqrt{E}r_2 + i(8r_1c)^{1/2} + i\alpha^2/(4\beta)]}{r_1^{3/4}r_2(2c)^{1/4}} \times \operatorname{erfc}\left(\frac{i\alpha}{2(i\beta)^{1/2}}\right)\sum_{LM}(ic)^{L+1/2}Y_L^M(\hat{\mathbf{r}}_1)t_{LM}(0,\sqrt{E}\hat{\mathbf{r}}_2), \text{ for } r_2 \to \infty, r_1^3 E \approx 2cr_2^2,$$

where

$$\alpha = \frac{r_2}{2\sqrt{E}} - \frac{2c}{3}\left(\frac{r_1}{2c}\right)^{3/2}$$

and

$$\beta = \frac{r_2}{8E^{3/2}} + \frac{c}{10}\left(\frac{r_1}{2c}\right)^{5/2}.$$

As the quantity $r_1^3 E/(2cr_2^2)$ becomes much less than unity, we may use an expansion of erfc to deduce, with the help of (24), that in the case r_1, r_2 large $r_1^3 E \ll 2cr_2^2$, the entire continuum contribution is given by

$$\Phi_c \sim \frac{\sqrt{E}\exp(i\sqrt{E}r_2)}{4\pi r_2^2 r_1^{3/4}(2c)^{1/4}}\{\exp[-i(8r_1c)^{1/2}] + i\exp[i(8r_1c)^{1/2}]\}\sum_{LM}(ic)^{L+1/2} \times Y_L^M(\hat{\mathbf{r}}_1)t_{LM}(0,\sqrt{E}\hat{\mathbf{r}}_2).$$

It will be seen that this result merges into (21) if an expansion of J_{2L+1} for large argument is substituted into (21).

V. REPULSIVE CASE AND GENERALIZATIONS

The case of a repulsive Coulomb interaction between particles 1 and 3 is much easier to discuss. The amplitudes $t_{LM}(q,\mathbf{p})$ still approach constants as $q \to 0$, but the quantity

$$e^{-n\pi}\Gamma(L+1-in)\Gamma(L+1+in) \to 2e^{-2n\pi}\pi n^3 = 2\pi c\exp(-2c/q)\pi/q^3 \quad \text{as} \quad q \to 0$$

in (11) now falls to zero exponentially. Hence, the contributions to the stationary phases at and near $q = 0$ are suppressed by this factor. The width of the saddle in the complex q plane goes as r_1/r_2. Thus for large r_2, $\Phi(\mathbf{r}_1,\mathbf{r}_2)$ goes to zero exponentially as r_1/r_2 decreases to zero.

It is also possible to modify our results[9] to include the possibility of an additional short-range interaction between particles 1 and 3.

We believe that the form of the results changes little if particle 2 is charged. Probably it is merely necessary to add a factor to our results for the form of Φ in the region. We hope to return to this question elsewhere.

---

* Work supported in part by the Air Force Office of Scientific Research, Office of Aerospace Research, U.S. Air Force, under Grant No. 71-1979 and, in part, performed under the auspices of the A.E.C. at Los Alamos Scientific Laboratory.

† Permanent address.

1 R. K. Peterkop, Bull. Acad. Sci. USSR, Phys. Ser., **27**, 987 (1963).

2 R. W. Hart, E. P. Gray, and W. H. Guier, Phys. Rev. **108**, 1512 (1957); E. Gerjuoy, *ibid.* **109**, 1806 (1958).

3 Note that there are really three inner regions, but we shall discuss only the one where the two charged particles of our model are relatively close together.

4 J. Nuttall, Phys. Rev. Letters **19**, 473 (1967); for a general

discussion see A. Erdelyi, *Asymptotic Expansions* (Dover, New York, 1956).
5 L. J. Slater, *Confluent Hypergeometric Functions* (Cambridge U. P., Cambridge, 1960).
6 A. Temkin, Phys. Rev. Letters **16**, 835 (1966).
7 R. K. Peterkop, "Proceedings of the Sommerfeld Centennial Memorial Conference and Symposium on the Physics of One- and Two-Electron Atoms" (unpublished).
8 See Ref. 5, Eq. (4.6.82), in which $te^{-2i\theta}$ should be written equal to $x/4k$.
9 J. Nuttall, in *Three Body Problem*, edited by J. S. C. McKee and P. M. Rolph (North-Holland, Amsterdam, 1970).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 10 OCTOBER 1971

The Riemannian Structure of Space–Time as a Consequence of a Measurement Method

Mario A. Castagnino
Departamento de Matemáticas, Facultad de Ciencias Exactas, Universidad de Buenos Aires, Argentina
(Received 6 December 1968; Revised Manuscript Received 18 February 1971)

Following the graphical method of Marzke and Wheeler it is shown analytically that if we know the space–time paths of all particles and light pulses we can deduce the connection and the metric of the space–time manifold. Specifically if such space–time paths fulfill some reasonable physical hypotheses, it is proved that space–time is a Riemannian manifold. To reach this conclusion a new definition of parallelism is introduced, based only on ideal experiments. This parallelism is entirely different from the ordinary parallel transfers, if the manifold is non-Riemannian; therefore it opens new ways of modifying gravitational theory.

1. THE INVERSE PROBLEM OF GENERAL RELATIVITY

Usually general relativity is based on the postulate that space–time is a Riemannian manifold. Once the metric of such structures is known (by the integration of corresponding equations), we can deduce the paths of particles and light rays as geodesic lines of the metric. But the Marzke–Wheeler method of measurement[1,2] allows us to invert the procedure. If the space–time paths of particles and light rays are experimentally known, Marzke and Wheeler show graphically how one can draw parallel lines, and construct an ideal geodesic clock that defines the metric over the whole manifold.

The finest achievement of Marzke–Wheeler method is the definition of the metric without clocks and measuring rods of atomic constitution. The metric only depends on the paths of particles and light rays.

In this paper we shall adopt the same procedure, but from an anlytical point of view. First, we suppose that we know the space–time paths of particles and light pulses, and that they satisfy some natural physical hypotheses which lead us to their differential equations. Second, we show how to define a natural, both geometrical and physical, parallelism (the *Desargues* transfer). Third, we find the parameter measured by a geodesic clock, i.e., the proper time. Fourth, we analyze the way in which this time defines a metric on the manifold. Finally, we conclude that if the proper time and the metric are defined unambiguously, the natural Desargues parallelism coincides with the ordinary parallel transfer of the Riemannian manifold which has such a metric.

In this way the Riemannian structure of space–time appears as a consequence of a set of natural and physical hypotheses.[3] We think that this work might allow us to begin a critical study of these hypotheses and to classify them according to their physical soundness. To acquire alternative theories to the classical one, we should keep those which are well founded and replace the weaker ones.

But, perhaps, the most interesting novelty is the introduction, based on physical considerations, of a new kind of parallel transfer which can also be interpreted within the manifold itself; without embedding the V_4 in a Euclidean space of more than four dimensions. This parallel *Desargues* transfer is entirely different from the usual *Levi-Civita* transfer, if the manifold is non-Riemannian and thereby opens new ways of modifying classical gravitational theory.

2. DIFFERENTIAL EQUATION OF THE SPACE–TIME PATH OF A FREE PARTICLE

To begin with, we must adopt a hypothesis that fixes the geometric background in which we shall work:

Hypothesis H_0: Space–time is an analytical (differentiable) manifold V_4. Every physical field is defined by an analytical function over V_4.

Specifically, each point $x \in V_4$ belongs to a neighborhood U which is mapped on an open sphere $S(U)$ which belongs to a four-dimensional Euclidean space R^4 by a coordinate system S, so that to each point $x \in U$ correspond for real numbers $S(x) = (x^i)$.

The changes of coordinates and the functions that define physical fields in V_4 are analytical functions.

To assume that these functions are analytical, i.e., that they lack any type of singularity or any other mathematical problem, is a usual hypothesis in physics, one that is probably not completely essential but is useful in this first approach.

We must define a *fundamental congruence* of curves, within V_4, which is formed by every free-particle space–time path and by the space–time

discussion see A. Erdelyi, *Asymptotic Expansions* (Dover, New York, 1956).
5 L. J. Slater, *Confluent Hypergeometric Functions* (Cambridge U. P., Cambridge, 1960).
6 A. Temkin, Phys. Rev. Letters **16**, 835 (1966).
7 R. K. Peterkop, "Proceedings of the Sommerfeld Centennial Memorial Conference and Symposium on the Physics of One- and Two-Electron Atoms" (unpublished).
8 See Ref. 5, Eq. (4.6.82), in which $te^{-2i\theta}$ should be written equal to $x/4k$.
9 J. Nuttall, in *Three Body Problem*, edited by J. S. C. McKee and P. M. Rolph (North-Holland, Amsterdam, 1970).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 10 OCTOBER 1971

The Riemannian Structure of Space–Time as a Consequence of a Measurement Method

Mario A. Castagnino
Departamento de Matemáticas, Facultad de Ciencias Exactas, Universidad de Buenos Aires, Argentina
(Received 6 December 1968; Revised Manuscript Received 18 February 1971)

Following the graphical method of Marzke and Wheeler it is shown analytically that if we know the space–time paths of all particles and light pulses we can deduce the connection and the metric of the space–time manifold. Specifically if such space–time paths fulfill some reasonable physical hypotheses, it is proved that space–time is a Riemannian manifold. To reach this conclusion a new definition of parallelism is introduced, based only on ideal experiments. This parallelism is entirely different from the ordinary parallel transfers, if the manifold is non-Riemannian; therefore it opens new ways of modifying gravitational theory.

1. THE INVERSE PROBLEM OF GENERAL RELATIVITY

Usually general relativity is based on the postulate that space–time is a Riemannian manifold. Once the metric of such structures is known (by the integration of corresponding equations), we can deduce the paths of particles and light rays as geodesic lines of the metric. But the Marzke–Wheeler method of measurement[1,2] allows us to invert the procedure. If the space–time paths of particles and light rays are experimentally known, Marzke and Wheeler show graphically how one can draw parallel lines, and construct an ideal geodesic clock that defines the metric over the whole manifold.

The finest achievement of Marzke–Wheeler method is the definition of the metric without clocks and measuring rods of atomic constitution. The metric only depends on the paths of particles and light rays.

In this paper we shall adopt the same procedure, but from an anlytical point of view. First, we suppose that we know the space–time paths of particles and light pulses, and that they satisfy some natural physical hypotheses which lead us to their differential equations. Second, we show how to define a natural, both geometrical and physical, parallelism (the *Desargues* transfer). Third, we find the parameter measured by a geodesic clock, i.e., the proper time. Fourth, we analyze the way in which this time defines a metric on the manifold. Finally, we conclude that if the proper time and the metric are defined unambiguously, the natural Desargues parallelism coincides with the ordinary parallel transfer of the Riemannian manifold which has such a metric.

In this way the Riemannian structure of space–time appears as a consequence of a set of natural and physical hypotheses.[3] We think that this work might allow us to begin a critical study of these hypotheses and to classify them according to their physical soundness. To acquire alternative theories to the classical one, we should keep those which are well founded and replace the weaker ones.

But, perhaps, the most interesting novelty is the introduction, based on physical considerations, of a new kind of parallel transfer which can also be interpreted within the manifold itself; without embedding the V_4 in a Euclidean space of more than four dimensions. This parallel *Desargues* transfer is entirely different from the usual *Levi-Civita* transfer, if the manifold is non-Riemannian and thereby opens new ways of modifying classical gravitational theory.

2. DIFFERENTIAL EQUATION OF THE SPACE–TIME PATH OF A FREE PARTICLE

To begin with, we must adopt a hypothesis that fixes the geometric background in which we shall work:

Hypothesis H_0: Space–time is an analytical (differentiable) manifold V_4. Every physical field is defined by an analytical function over V_4.

Specifically, each point $x \in V_4$ belongs to a neighborhood U which is mapped on an open sphere $S(U)$ which belongs to a four-dimensional Euclidean space R^4 by a coordinate system S, so that to each point $x \in U$ correspond for real numbers $S(x) = (x^i)$.

The changes of coordinates and the functions that define physical fields in V_4 are analytical functions.

To assume that these functions are analytical, i.e., that they lack any type of singularity or any other mathematical problem, is a usual hypothesis in physics, one that is probably not completely essential but is useful in this first approach.

We must define a *fundamental congruence* of curves, within V_4, which is formed by every free-particle space–time path and by the space–time

paths of pulses of light. We can assume that the fundamental congruence is experimentally known *a priori*. Of course we know that our experimental knowledge of the fundamental congruence is not yet complete. This is why we are forced to make various hypotheses about the behavior of these curves. The origin and physical meaning of these hypotheses are well known.

To begin with we must require that the congruence satisfies the *weak equivalence principle*; that is to say that paths of free particles are independent of the mass and of all other properties of the particle. At each point $x \in V_4$ the tangent vectors to the curves of the fundamental congruence form a subset I_x of the tangent space at $x : \tau x$. We can then postulate:

Hypothesis H_1: For each point $x \in V_4$ and for each vector $(A^i) \in I_x$ there exists one and only one curve of the fundamental congruence that passes through x and is tangent to (A^i).

In other words: if u is a parameter such that $A^i = dx^i/du$ and if we know x and dx^i/du, then we know the corresponding curve and also the second derivative at a point x:

$$\frac{d^2x^i}{du^2} = f^i\left(x; \frac{dx^j}{du}\right), \tag{1}$$

so that the functions f^i are experimentally known.

A subset of the fundamental congruence is formed by the paths of light pulses. The tangent vectors to these curves define in each tangent space τ_x, a set of vectors $C_x \subseteq \tau_x$ that satisfy the following postulate:

Hypothesis H_2: C_x is the boundary of I_x and is also a hypercone of signature $+,+,+,-$. I_x is the "interior" of this hypercone.

We shall call this hypercone the *null cone*. Its equation is

$$\lambda_{ij}A^iA^j = 0, \tag{2}$$

where $(A^i) \in \tau_x$. Naturally we shall call a vector (A^i) "interior" to the null cone if $\lambda_{ij}X^iX^j = \lambda_{ij}A^iA^j$ is the equation, in variables X^i, of a hyperboloid of two sheets in the space τ_x. In this way, at each point of V_4 a tensor λ_{ij} is clearly defined [λ_{ij} can be considered to be symmetric because its antisymmetrical part is irrelevant, as can be seen from Eq. (2)], λ_{ij} is of signature $+,+,+,-$ and can be multiplied by an arbitrary factor without changing the null cone; i.e., we have a *conformal metric*. We can also say that λ_{ij} is experimentally known. We shall find an adequate factor λ such that $\lambda\lambda_{ij}$ could be considered as the true metric.

If we develop the analytic function of the right-hand side (rhs) of Eq. (1), we have

$$-\frac{d^2x^i}{du^2} = \overset{*}{\Gamma}{}^i + \overset{*}{\Gamma}{}^i_j\frac{dx^j}{du} + \overset{*}{\Gamma}{}^i_{jk}\frac{dx^j}{du}\frac{dx^k}{du} + \cdots + \overset{*}{\Gamma}{}^i_{j_1\cdots j_n}\frac{dx^{j_1}}{du}\cdots\frac{dx^{j_n}}{du} + \cdots. \tag{3}$$

We know that "gravitational forces" do not exist in a system in free fall, so that the paths of free particles must be, locally, straight lines in such a system; thus we may postulate:

Hypothesis H_3: There is a system S' in which the second derivative of the space–time path of all particles is proportional to the first derivative.

This is of course a well-known feature of the equivalence principle. We observe that this statement of the principle (although it is stronger than the weak principle) is weaker than the strong equivalence principle.

By calculating the new coefficients $\overset{*}{\Gamma}{}'$ in S' as a function of the $\overset{*}{\Gamma}$ in S and imposing the condition H_3 it is easy to prove that Eq. (3) becomes

$$-\frac{d^2x^i}{du^2} = \overset{*}{\Gamma}{}^i_{jk}\frac{dx^j}{du}\frac{dx^k}{du} + C\frac{dx^i}{du}, \tag{4}$$

and if we use an adequate parameter,

$$\frac{d^2x^i}{du^2} + \overset{*}{\Gamma}{}^i_{jk}\frac{dx^j}{du}\frac{dx^k}{du} = 0. \tag{5}$$

We have ∞^2 parameters of this type, which can be obtained from each other by linear transformations. We shall call them *affine parameters* and we shall use them from now on.

In this way the coefficients $\overset{*}{\Gamma}{}^i_{jk}$ (which are symmetric in jk for the same reason as λ_{ij} is) are defined at every point of V_4. Thus $\overset{*}{\Gamma}{}^i_{jk}$ can be considered as the coefficients of an *affine connection*, which allows us to introduce a parallel transfer according to *Levi-Civita* and a covariant derivation ($\overset{*}{\nabla}_i$ and $\overset{*}{D}$). But we shall see that this notion of parallelism is not physically correct, in general. We shall use the covariant derivation as an auxiliary notation, so Eq. (5) can be written as

$$\frac{\overset{*}{D}}{du}\frac{dx^i}{du} = \frac{dx^j}{du}\overset{*}{\nabla}_j\frac{dx^i}{du} = 0. \tag{6}$$

It is known that a light pulse never becomes a particle of velocity less than c, nor is the inverse case possible. (At least this is true in all V_4 except for a set of isolated points where phenomena such as photoproduction can take place.) Then we can postulate that, except possibly for a set of isolated points, the following hypothesis is valid in all V_4:

Hypothesis H_4: If the vector dx^i/du is the tangent to a curve of the fundamental congruence and it belongs to a null cone at one point of the curve, the tangent vectors at all the other points of the

curve belong to their corresponding null cones. (We shall call this particular curve a null curve; it is the space–time path of a light pulse.)

So if dx^i/du has the property

$$\lambda_{ij}\frac{dx^i}{du}\frac{dx^j}{du}=0 \tag{7}$$

at a point of a path of the fundamental congruence, Eq. (7) holds for all other points of the path. Taking the derivative with respect to u and remembering formula (6), which is satisfied by dx^i/du, we have

$$\overset{*}{\nabla}_k\lambda_{ij}\frac{dx^i}{du}\frac{dx^j}{du}\frac{dx^k}{du}=0. \tag{8}$$

Then if dx^i/du is a root of polynomial (7), it is also a root of polynomial (8), so the latter is divisible by the former, i.e.,

$$\overset{*}{\nabla}_k\lambda_{ij}\frac{dx^i}{du}\frac{dx^j}{du}\frac{dx^k}{du}=\lambda_{ij}\frac{dx^i}{du}\frac{dx^j}{du}\lambda_k\frac{dx^k}{du}, \tag{9}$$

where λ_k is a vector to be determined. Then[4]

$$\underset{ijk}{S}\overset{*}{\nabla}_i\lambda_{jk}=\underset{ijk}{S}\lambda_i\lambda_{jk}. \tag{10}$$

Since this equation holds true except for a set of isolated points and we have assumed that λ_{jk} and $\overset{*}{\Gamma}{}^i_{jk}$ are analytic functions of x, we can say that (10) is valid at every point of V_4.

3. DEFINITION OF PARALLEL TRANSFER ON THE MANIFOLD V_4

We can give a physical definition of parallel transfer on space–time based on *affine geometry*. We know by H_3 that for every $x \in V_4$ there exists at least one privileged system S' that maps all curves of the fundamental congruence that belong to $U(x \in U \subseteq V_4)$ on $S'(U)[(x^i) \in S'(U) \subseteq R_4]$; in such a way that they will have a vanishing second derivative in (x^i) (if we use an affine parameter); i.e., in $S'(U)$ all the curves of the fundamental congruence are mapped on to approximately straight lines; we can make the approximation as good as we wish by reducing U. From now on we shall speak of paths of particles and light pulses as if they were straight lines, because we shall refer to their images under the S' mapping. For this set of straight lines of Euclidean space R_4 all theorems of *projective geometry* are obviously valid.

On the other hand it is evident that we can consider the paths of two light pulses emitted by a particle and lying on a two-plane of R_4 as parallel. Indeed, in the system S', where gravitational fields do not exist, all events take place locally as in the *special relativity*, and in the flat space these paths of light pulses cannot intersect. (If such a thing could happen, two light wavefronts emitted by one particle would intersect, which would only be possible if the particles were travelling faster than light.) So we have physically defined the notion of parallelism, because in each (space–time-like) two-plane there are two pairs of parallel lines (i.e., two *improper points*). These are the intersections of the two null cones (i.e., the two wavefronts) with the two-plane. According to the *Desargues theorem* (and some others such as the *Pappus theorem*), given two pairs of parallel lines, we can obtain other pairs of parallel lines.

In its affine version the Desargues theorem says: Given two triangles ABC and $A'B'C'$ (Fig. 1) such that the three straight lines that join the corresponding vertices intersect at point 0, if AC is parallel to $A'C'$ and if AB is parallel to $A'B'$, then BC is parallel to $B'C'$.

If, on a two-surface, we repeat the drawing of Fig. 1, OA, OB, and OC are paths of particles, AC and $A'C'$ are paths of light pulses that could be considered parallel, and the same happens with AB and $A'B'$; then BC is parallel to $B'C'$. This method is essentially the one used by Marzke and Wheeler to define parallelism, and, as the author has proved, it gives rise to an ideal experiment that allows parallel paths to be constructed.[5] The reader interested in a more detailed explanation of the physical meaning of this parallelism is referred to Refs. 1, 2, and 5.

These considerations are only approximate, but they become exact if they are repeated on V_4 with paths of particles, instead of straight lines of R_4 and if we make Fig. 1 become infinitely small.

To prove this, let us take a point $0 \in V_4$ of coordinates (0, 0, 0, 0) in an arbitrary system S, and two vectors A^i (interior to the null cone) and B^i (arbitrary). A^i and B^i define a two-plane (Fig. 2). Let us consider a two-surface formed by all the particles paths which pass through 0 and are tangent to the two-plane, and let us define

$$\lambda_{ij}A^iA^j=A,\quad \lambda_{ij}B^iB^j=B,\quad \lambda_{ij}A^iB^j=C. \tag{11}$$

Let us define an affine parameter u on each path

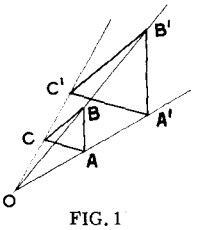


FIG. 1

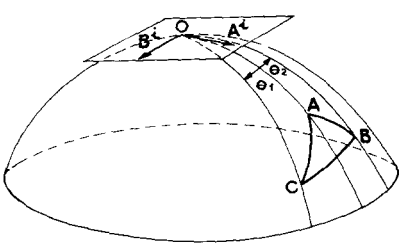


FIG. 2

of the two-surface that passes through 0 in such a way that $u = 0$ for the point 0. All the tangent vectors inside the null cone can be written as

$$\left(\frac{\partial x^i}{\partial u}\right)_0 = \cosh\theta A^i + \sinh\theta B^i. \tag{12}$$

In this way each point of the two-surface inside the null conoid at 0 is defined by two corrdinates u and $\theta: x^i = x^i(u, \theta)$.

Let OA be the path tangent to the vector A^i (Fig. 2), and let us take two other paths OB and OC and two paths of light pulses AB and AC. Let $u = u_1(\theta)$ and $u = u_2(\theta)$ be the equations of these last paths. Then the coordinates of the points A, B, and C are, respectively,

$$\begin{aligned} &\text{A: } x^i = x^i(u; \theta), \\ &\text{B: } x^i = x^i(u_2(\theta_2); \theta_2), \\ &\text{C: } x^i = x^i(u_1(\theta_1); \theta_1). \end{aligned} \tag{13}$$

Let us take $\theta_1 \to 0$ and $\theta_2 \to 0$ and calculate the tangent vector

$$\delta' x^i = \lim_{\theta_1 \to 0,\ \theta_2 \to 0} \frac{x^i(u_2(\theta_2); \theta_2) - x^i(u_1(\theta_1); \theta_1)}{\theta_2 - \theta_1}. \tag{14}$$

Calling the arbitrary limits

$$\lim_{\theta_1 \to 0,\ \theta_2 \to 0} \frac{\theta_2}{\theta_2 - \theta_1} = \alpha', \quad \lim_{\theta_1 \to 0,\ \theta_2 \to 0} \frac{\theta_1}{\theta_2 - \theta_1} = \beta', \tag{15}$$

we have

$$\begin{aligned} \delta' x^i &= \alpha' \frac{\partial}{\partial\theta} x^i(u_2(\theta); \theta) + \beta' \frac{\partial}{\partial\theta} x^i(u_1(\theta); \theta) \\ &= \alpha'\left(\frac{\partial x^i}{\partial u}\frac{du_2}{d\theta} + \frac{\partial x^i}{\partial\theta}\right) + \beta'\left(\frac{\partial x^i}{\partial u}\frac{du_1}{d\theta} + \frac{\partial x^i}{\partial\theta}\right). \end{aligned} \tag{16}$$

If we take α' and β' to be constant along the curve $\theta = 0$ (i.e., $0A$) and we repeat the construction of $\delta' x^i$ at each point of the curve $0A$, we can consider $\delta' x^i$ to undergo parallel transfer along this curve according to the Desargues theorem. In fact, we have only drawn triangles ABC that satisfy the hypotheses of this theorem in such a way that the sides AB and AC can be considered parallel. After we have made angles θ_1 and θ_2 tend to zero, the vector δx^i tangent to CB may be considered parallel to itself along the curve. $du_2/d\theta$ and $du_1/d\theta$ must be taken in such a way that AB and AC are light pulses, i.e.,

$$\frac{\partial x^i}{\partial u}\frac{du_2}{\partial\theta} + \frac{\partial x^i}{\partial\theta} \quad \text{and} \quad \frac{\partial x^i}{\partial u}\frac{du_1}{d\theta} + \frac{\partial x^i}{\partial\theta}$$

must belong to the null cone.

Then

$$\left(\lambda_{ij}\frac{\partial x^i}{\partial u}\frac{\partial x^j}{\partial u}\right)\left(\frac{du_{1,2}}{\partial\theta}\right)^2 + 2\left(\lambda_{ij}\frac{\partial x^i}{\partial u}\frac{\partial x^j}{\partial u}\right)\frac{du_{1,2}}{\partial\theta} + \lambda_{ij}\frac{\partial x^i}{\partial\theta}\frac{\partial x^j}{\partial\theta} = 0. \tag{17}$$

Calling

$$\begin{aligned} a &= \lambda_{ij}\frac{\partial x^i}{\partial u}\frac{\partial x^j}{\partial u}, \quad b = \lambda_{ij}\frac{\partial x^i}{\partial u}\frac{\partial x^j}{\partial\theta}, \\ c &= \lambda_{ij}\frac{\partial x^i}{\partial\theta}\frac{\partial x^j}{\partial\theta}, \quad d = +\,(b^2 - ac)^{1/2}, \end{aligned} \tag{18}$$

we have

$$\frac{du_1}{d\theta} = -\frac{b}{a} + \frac{d}{a}, \quad \frac{du_2}{d\theta} = -\frac{b}{a} - \frac{d}{a}; \tag{19}$$

by replacing (19) in (16), this last equation becomes

$$\delta' x^i = \left(\frac{b}{a}\alpha + \frac{d}{a}\beta\right)\frac{\partial x^i}{\partial u} - \alpha\frac{\partial x^i}{\partial\theta}, \tag{20}$$

where we have introduced new constants

$$\alpha = -(\alpha' + \beta'), \quad \beta = \alpha' - \beta'. \tag{21}$$

Let us develop the analytic function $x^i = x^i(u, \theta)$ in a power series:

$$x^i = \left(\frac{\partial x^i}{\partial u}\right)_0 u + \tfrac{1}{2}\left(\frac{\partial^2 u^i}{\partial u^2}\right)_0 u^2 + o(u^2). \tag{22}$$

But since each curve is a path, it satisfies (5). Using (11) we have

$$\begin{aligned} x^i = (\cosh\theta\, A^i + \sinh\theta\, A^i)u - \tfrac{1}{2}(\overset{*}{\Gamma}{}^i_{jk})_0(\cosh\theta\, A^j \\ + \sinh\theta\, A^j)(\cosh\theta\, A^k + \sinh\theta\, A^k)u^2 + o(u^2). \end{aligned} \tag{23}$$

We can define a new vector:

$$\delta x^i = (1/u)\delta' x^i \tag{24}$$

parallel to $\delta' x^i$. Then Eq. (20), with our computation of a, b, c, d, of (18) with (23) yields

$$\delta x^i = (\alpha C + \beta D)A^i - \alpha A B^i + X^i u + o(u^2), \tag{25}$$

where

$$\begin{aligned} X^i = &-(\overset{*}{\Gamma}{}^i_{jk})_0 A^k[(\alpha C + \beta D)A^j - \alpha A B^j] \\ &+ \alpha A^k A^m (A^i B^n - B^i A^n)\overset{*}{\nabla}_k \lambda_{mn} \\ &+ \beta A^i A^k (\overset{*}{\nabla}_k \lambda_{mn})(1/2D)(-AB^m B^n + 2CA^m B^n \\ &- BA^m A^n) \end{aligned} \tag{26}$$

and

$$D^2 = C^2 - AB. \tag{27}$$

Differentiating, we have

$$\frac{d\delta x^i}{du} = X^i + o(u). \tag{28}$$

Let us take $u \to 0$. In this way we consider only the points in an infinitely small neighborhood of 0, and Eqs. (25) and (28) become

$$\delta x^i = (\alpha C + \beta D)A^i - \alpha A B^i, \tag{29}$$

$$\frac{d\delta x^i}{du} = X^i, \tag{30}$$

i.e., we have

$$\frac{\overset{*}{D}}{du}\delta x^i = \alpha(A^i B^n A^m - B^i A^n A^m)(\overset{*}{\nabla}_k \lambda_{mn})A^k + \beta A^i A^k (\overset{*}{\nabla}_k \lambda_{mn})(1/2D)(-AB^m B^n + 2CA^m B^n - BA^m A^n). \tag{31}$$

This is the differential equation of the parallel transfer of the vector δx^i along the curve.

From Eq. (12) we obtain that A^i is the tangent vector to the curve, and B^i is an arbitrary vector of the two-plane defined by A^i and δx^i, as Eq. (29) shows. We shall call this transfer a *Desargues transfer* and we must observe that it is not the same as the ordinary *Levi-Civita transfer*. But the Desargues transfer can claim to be the real physical parallel transfer, as we have seen.

Nevertheless, the tangent vector A^i can be obtained by making $\alpha = 0$ in (29). Thus, it satisfies a differential equation of the form

$$\frac{\overset{*}{D}\delta x^i}{du} = f(u)\delta x^i. \tag{32}$$

In this case, then, both parallel transfers coincide.

For further developments let us calculate the *Grassmann tensor* of the two-plane defined by A^i and δx^i:

$$\overset{\tau}{W}{}^{ij} = \begin{vmatrix} A^i & \delta x^i \\ A^j & \delta x^j \end{vmatrix}. \tag{33}$$

Bearing in mind that $\overset{*}{D}A^i/du = 0$ and that δx^i satisfies (31), we have

$$\frac{\overset{*}{D}\overset{\tau}{W}{}^{ij}}{du} = g(u)\overset{\tau}{W}{}^{ij}, \tag{34}$$

i.e., the two-plane is subject to a parallel transfer according to *Levi-Civita*. Then there exist two vectors that lie on the two-plane and that are subject to parallel transfer.[6] One of them is A^i, the other B^i. We have

$$\frac{\overset{*}{D}A^i}{du} = 0, \qquad \frac{\overset{*}{D}B^i}{du} = 0. \tag{35}$$

Let us now consider a new Grassmann tensor of the two-plane, proportional to $\overset{\tau}{W}{}^{ij}$:

$$W^{ij} = \begin{vmatrix} A^i & B^i \\ A^j & B^j \end{vmatrix}, \tag{36}$$

which satisfies

$$\frac{\overset{*}{D}}{du}W^{ij} = 0. \tag{37}$$

Equation (29) can be written:

$$\delta x^i = \alpha\lambda_{jk}W^{ij}A^k + \beta DA^i. \tag{38}$$

If we take α and β to be constant and we differentiate Eq. (38), using Eqs. (35) and (37), we obtain Eq. (31), so that Eqs. (38) and (35) also define the Desargues transfer.

Let us finally observe that δx^i can be decomposed into two components of peculiar characteristics. In fact, in Eq. (38) the second term of the rhs is parallel to A^i and the first one is orthogonal in the sense of the conformal metric, because

$$\lambda_{mn}A^n\lambda_{jk}W^{mj}A^k = W^{nj}A_nA_j = 0. \tag{39}$$

Then if we put $\beta = 0$ in Eq. (38), we obtain a vector that undergoes parallel transfer and which is orthogonal to the curve everywhere. So orthogonality in the sense of the metric λ_{ij} is conserved under the Desargues transfer.

4. THE PARALLEL RIBBON

The vector δx^i given by Eqs. (35) and (38) can be regarded as parallel to itself but not as equipollent. In fact it can be multiplied by an arbitrary scalar (variable along the curve) and it still satisfies (35) and (38), as can easily be shown. Now we want to determine this scalar in such a way that the origin and the end point of δx^i define two parallel curves, i.e., two curves whose tangent vectors are parallel, as we shall define precisely later.

We shall limit ourselves, for the moment, to the case of a vector orthogonal to the curve C (Fig. 3). Therefore Eq. (35) reduces to

$$\delta x^i = \alpha(u)\lambda_{mn}W^{im}A^n. \tag{40}$$

We want to determine $\alpha(u)$ so that δx^i may be considered equipollent to itself along the curve. At each point of the curve C we can construct a curve tangent to the corresponding δx^i and geodesic, in the sense of the connection $\overset{*}{\Gamma}{}^i_{jk}$, i.e.; δx^i will be tangent to these curves and will undergo parallel transfer along these curves according to Levi-Civita. Then if v is an affine parameter to these curves, we find

$$\delta x^i = \frac{\partial x^i}{\partial v}, \qquad \frac{\overset{*}{D}}{dv}\delta x^i = 0. \tag{41}$$

The set of points of these new curves generates a two-surface endowed with a system of coordinates $x^i = x^i(u, v)$. We know that along the curve C

$$\lambda_{ij}A^i\delta x^j = 0. \tag{42}$$

If we want the curves $v = \text{const}$ to be parallel to the curve C (let C be the curve $v = 0$), their tangents must be obtained by the parallel Desargues transfer of vectors A^i along the curves $u = \text{const}$. But

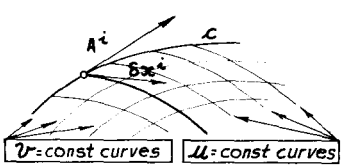


FIG. 3

the Desargues transfer preserves orthogonality; therefore the former condition is equivalent to demanding that the net of coordinate curves be orthogonal. Now if the curves $v = \text{const}$ are parallel, the vectors δx^i, obtained by (41) as limits when $v \to 0$, may be considered equipollent. Therefore

$$\frac{\overset{*}{D}}{dv}(\lambda_{ij}A^i\delta x^j) = 0. \tag{43}$$

Bearing in mind that $A^i = \partial x^i/\partial u$, that $\delta x^i = \partial x^i/\partial v$, and that the coefficients $\overset{*}{\Gamma}{}^i_{jk}$ are symmetric, we have

$$\frac{\overset{*}{D}}{dv}A^i = \frac{\overset{*}{D}}{du}\delta x^i. \tag{44}$$

Consequently Eq. (43) becomes

$$\delta x^k(\overset{*}{\nabla}_k\lambda_{ij})A^i\delta x^j + \lambda_{ij}\delta x^j\frac{D\delta x^i}{du} = 0. \tag{45}$$

This is the differential equation that must be satisfied if we want the coordinate net to be orthogonal.

Equation (45) defines the coefficient $\alpha(u)$ of (40) and it can be integrated if we use Eq. (10). This leads us to

$$-\left(\frac{\overset{*}{D}}{du}\lambda_{jk}\right)\delta x^j\delta x^k + 2\lambda_{ij}\delta x^i\frac{\overset{*}{D}}{du}\delta x^j + A^i\lambda_i\delta x^j\delta x^k\lambda_{jk} = 0. \tag{46}$$

If we call

$$\delta = \lambda_{ij}\delta x^i\delta x^j \tag{47}$$

and

$$C^i = \lambda_{mn}A^mW^{in}, \tag{48}$$

and if we calculate

$$\lambda_{ij}C^iC^j = -AD^2, \tag{49}$$

then

$$\delta = -\alpha^2AD^2. \tag{50}$$

If δ is introduced in (46), we obtain

$$AD^2\frac{d}{du}\log\alpha + \tfrac{1}{2}AD^2A^i\lambda_i + \tfrac{1}{2}\frac{d}{du}(AD^2) + C^iC^j\frac{\overset{*}{D}}{du}\lambda_{ij} = 0. \tag{51}$$

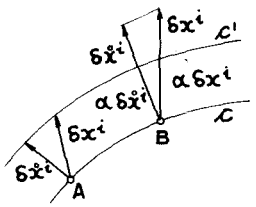


FIG. 4

If we now calculate dD^2/du and C^iC^j, we obtain

$$C^iC^j\frac{\overset{*}{D}}{du}\lambda_{ij} = -A\frac{dD^2}{du} + D^2\frac{dA}{du}. \tag{52}$$

On the other hand, using Eq. (10), we have

$$\frac{dA}{du} = A^kA^iA^j\overset{*}{\nabla}_k\lambda_{ij} = A^k\lambda_k\lambda_{ij}A^iA^j = A\lambda_kA^k. \tag{53}$$

Introducing (50) and (53) into (51), we have

$$\frac{d}{du}\log\alpha + 2\frac{d}{du}\log A - \tfrac{1}{2}\log D^2 = 0, \tag{54}$$

so finally,

$$\alpha = KD/A^2, \tag{55}$$

and the equipollent δx^i is given by

$$\delta x^i = (KD/A^2)\lambda mnW^{im}A^n, \tag{56}$$

where, in addition to the conditions (35), K is a constant.

We shall call the curve C and a curve $v = \text{const}$ a parallel ribbon that can be approximately interpreted as the curves generated by the origin and the end point of δx^i, when δx^i undergoes parallel transfer.

Now it is easy to understand that the coefficient α can be used to make the parallel transfer equipollent in the general case (38), when δx^i is not orthogonal to A^i. Specifically, let δ^0x^i, the normal component of δx^i at a point A of the curve C (Fig. 4), generate a parallel ribbon CC'. With formula (38) we transfer both δx^i and δ^0x^i in the Desargues parallel way, to the point B. To adjust δ^0x^i so that it would belong to the ribbon, we must multiply it by α, and so, we must do the same with δx^i. Therefore equipollent transfer in the general case is given by

$$\delta x^i = K\frac{D}{A^2}\lambda_{mn}W^{im}A^n + H\frac{D^2}{A^2}A^i, \tag{57}$$

where K and H are constants, together with Eqs. (35).

5. THE GEODESIC CLOCK

Now we can define a geodesic clock as a pulse of light that bounces back and forth between two boundary curves of a parallel ribbon. If δx^i is the vector that defines the parallel ribbon we want to find a parameter v (in general a nonaffine one) in such a way that if the diagonal curves of Fig. 5 are the paths of the bouncing light pulses, the increment Δv of the parameters is the same for each oscillation of the light pulse.

$\delta x^i + \frac{1}{2}\Delta v dx^i/dv$ is the vector that can be considered as the path of the light pulse to a first approximation, so

$$\lambda_{ij}\left(\delta x^i + \frac{dx^i}{dv}\frac{\Delta v}{2}\right)\left(\delta x^j + \frac{dx^j}{dv}\frac{\Delta v}{2}\right) = 0, \tag{58}$$

and if δx^i is normal to the boundary curve of the ribbon, as in Sec. 4, we have

$$\delta = \lambda_{ij}\delta x^i \delta x^j = -\lambda_{ij}\left(\frac{dx^i}{dv}\right)\left(\frac{dx^j}{dv}\right)\left(\frac{\Delta v}{2}\right)^2. \tag{59}$$

By Eqs. (50) and (55) we have

$$K^2\frac{D^4}{A^4} = \left(\frac{du}{dv}\right)^2\left(\frac{\Delta v}{2}\right)^2, \tag{60}$$

thus if we want Δv to be constant we must have

$$v = K'\int_0^u \frac{A^2}{D^2}\,du, \tag{61}$$

where K' is a constant and v is the time parameter measured by the geodesic clock.

In this way the notion of parallel transfer and the existence of null cones allow us to measure times (and distances too) and will also allow us to construct a metric on V_4 if we add some hypotheses to those already given. Fundamentally the method of measurement given up to now must lead us to a unique result to make the definition of a metric possible. For example, v depends upon the parallel ribbon we are using, i.e.; two parallel ribbons with a common boundary curve C lead, in general, to different times v. Let us introduce the assumption that such a thing would not take place:

Hypothesis H_5 : Formula (61) gives us a unique time, independent of the parallel ribbon we are using.

If δx^i and $\delta' x^i$ define two parallel ribbons with the same boundary curve C we must have

$$\int_0^u \frac{A^2}{D^2}\,du = K''\int_0^u \frac{A^2}{D'^2}\,du, \tag{62}$$

where D^2 and D'^2 are the correspondent determinants and K'' is a constant. Then

$$\frac{dD'^2/du}{D'^2} = \frac{dD^2/du}{D^2} = k(A^i), \tag{63}$$

where $k(A^i)$ is function of λ_{ij} and A^i and is not a function of B^i or δx^i. Then

$$A^iA^jA^k\overset{*}{\nabla}_k\begin{vmatrix}\lambda_{ij} & \lambda_{ia}\\ \lambda_{jb} & \lambda_{ab}\end{vmatrix} = k(A^k)A^iA^j\begin{vmatrix}\lambda_{ij} & \lambda_{ia}\\ \lambda_{jb} & \lambda_{ab}\end{vmatrix}. \tag{64}$$

When contracted with λ^{ab}, the inverse matrix of λ_{ab}, this equation leads us to

$$k(A^k) = \tfrac{1}{3}\left(A^k d_k + \frac{2}{A}\frac{dA}{du}\right), \tag{65}$$

where

$$d_k = \lambda^{ij}\overset{*}{\nabla}_k\lambda_{ij}. \tag{66}$$

Then, using Eq. (53), we have

$$k(A^k) = A^k\mu_k, \tag{67}$$

where

$$\mu_k = \tfrac{1}{3}(d_k + 2\lambda_k). \tag{68}$$

μ_k is a vector independent of A^i. Equations (64) and (67) must be satisfied by all A^i. Then hypothesis H^5 implies that

$$\underset{ijk}{S}\overset{*}{\nabla}_k\begin{vmatrix}\lambda_{ai} & \lambda_{ij}\\ \lambda_{ab} & \lambda_{bj}\end{vmatrix} = \underset{ijk}{S}\mu_k\begin{vmatrix}\lambda_{ai} & \lambda_{ij}\\ \lambda_{ab} & \lambda_{bj}\end{vmatrix}. \tag{69}$$

Contracting with λ_{ij} and using (10), we have

$$\overset{*}{\nabla}_k\lambda_{ab} = a_k\lambda_{ab} + b_a\lambda_{bk} + b_b\lambda_{ak}, \tag{70}$$

where

$$a_k = \tfrac{1}{10}(2\lambda_k + 10\mu_k - 2d_k - 4c_k),$$
$$b_k = \tfrac{1}{10}(2\lambda_k - 2\mu_k + 2c_k), \tag{71}$$

and

$$c_k = \lambda^{ij}\overset{*}{\nabla}_i\lambda_{jk}. \tag{72}$$

Contracting Eq. (70) with λ^{ab}, we obtain

$$c_k = 4\lambda_k - \tfrac{3}{2}\mu_k, \qquad d_k = -2\lambda_k - 3\mu_k. \tag{73}$$

Then

$$a_k = \mu_k - \lambda_k, \qquad b_k = \lambda_k - \tfrac{1}{2}\mu_k, \tag{74}$$

and Eq. (70) becomes

$$\overset{*}{\nabla}_k\lambda_{ab} = (\mu_k - \lambda_k)\lambda_{ab} + (\lambda_a - \tfrac{1}{2}\mu_a)\lambda_{bk} + (\lambda_b - \tfrac{1}{2}\mu_b)\lambda_{ak}. \tag{75}$$

In this way hypothesis H_5 allows us to give an explicit expression of the covariant derivative of λ_{ab} as a function of two unknown vectors λ_k and μ_k. From now on we shall suppose that H_5 is valid, so that time v is uniquely defined along each path.

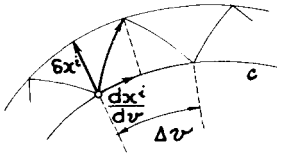


FIG. 5

6 THE EXTENSION OF THE METRIC

If it is possible to define a real metric, it must be one of the conformal metrics, so

$$g_{ij} = \lambda\lambda_{ij}. \tag{76}$$

Besides, on any curve of the fundamental congruence, the proper time v must be equal to the length

of the corresponding piece of path measured by the metric g_{ij}, i.e.,

$$dv^2 = -g_{ij}dx^i dx^j. \tag{77}$$

From (76) and (77) we can find λ:

$$\lambda = -\left[\lambda_{ij}\frac{dx^i}{du}\frac{dx^j}{du}\left(\frac{du}{dv}\right)^2\right]^{-1} = \frac{1}{A}\left(\frac{dv}{du}\right)^2 = -K'^2\frac{A^3}{D^4}, \tag{78}$$

Let C be a curve of the fundamental congruence that joins point ${}^0x^i$ to point x^i. By definition we take $\lambda = 1$ at ${}^0x^i$, which means that we take λ_{ij} as the real g_{ij} in ${}^0x^i$. In other words, in ${}^0x^i$ we define the standards of time or length. Then if A_0 and D_0 are the values of A and D in ${}^0x^i$, then we have

$$K'^2 = -\frac{D_0^4}{A_0^3}, \tag{79}$$

and

$$\lambda = \frac{D_0^4}{A_0^3}\frac{A^3}{D^4}. \tag{80}$$

But, of course, λ depends on the curve C that joins ${}^0x^i$ to x^i. This new ambiguity must be eliminated by a new hypothesis.

Finally, we observe that if in (61) we substitute for K' its value given by (79), we shall obtain the proper time τ according to the standard taken in ${}^0x^i$.

7. THE RIEMANNIAN MANIFOLD

Let us then make a final hypothesis:

Hypothesis H_6: The preceding method leads us to a unique metric, independent of the curve we use.

Let us express this hypothesis as a mathematical equation. From Eq. (80) we have

$$\frac{d}{du}\log\lambda = \frac{3}{A}\frac{dA}{du} - \frac{2}{D^2}\frac{dD^2}{du}. \tag{81}$$

Taking (53), (63), and (67) into account, we have

$$\frac{d}{du}\log\lambda = A^k(3\lambda_k - 2\mu_k). \tag{82}$$

H_6 requires that $d\log\lambda$ be a total exact differential; i.e., λ_k and μ_k are related by the following equation:

$$3\lambda_k - 2\mu_k = \partial_k\log\lambda, \tag{83}$$

where now λ is defined all over V_4. We have now unequivocally defined the metric g_{ij} and the Desargues parallel transfer. We shall see that if we accept all the hypotheses H_0–H_6, the Desargues parallel transfer coincides with that of Levi-Civita, and is simply the one that corresponds to a Riemannian manifold; i.e., the connection is given by the Christoffel symbols, so we would arrive at the classical theory of *Einstein*. Then we shall prove the following:

Theorem: If the hypotheses H_0–H_6 are valid, the Desargues transfer coincides with the ordinary parallel transfer of a Riemannian manifold with metric g_{ij}.

Specifically, as $g_{ij} = \lambda\lambda_{ij}$, using (75) and (83), we have

$$\overset{*}{\nabla}_k g_{ij} = 2(\lambda_k - \tfrac{1}{2}\mu_k)g_{ij} + (\lambda_i - \tfrac{1}{2}\mu_i)g_{jk} + (\lambda_j - \tfrac{1}{2}\mu_j)g_{ik}. \tag{84}$$

We can restudy the formulas that define parallel transfer, but now since g_{ij} is defined in a unique way, we can take it as the original λ_{ij}. Then we have $g_{ij} = \lambda_{ij}$, $\lambda = 1$, and from (83), $\mu_k = \frac{3}{2}\lambda_k$. Thus

$$\overset{*}{\nabla}_k g_{ij} = \tfrac{1}{4}[2\lambda_k g_{ij} + \lambda_i g_{jk} + \lambda_j g_{ik}]. \tag{85}$$

From (63) and (65) we have

$$\frac{dD^2}{du} = \tfrac{3}{2}A^i\lambda_i D^2, \tag{86}$$

and from (53), (57), (85), and (86) we obtain the differential equation of Desargues transfer:

$$\frac{\overset{*}{D}\delta x^i}{du} = -\tfrac{1}{2}A^j\lambda_j\delta x^i + \frac{K}{4}\frac{D}{A}W^{ij}\lambda_j, \tag{87}$$

and from Eq. (57)

$$\frac{A^2}{D}\delta x^i = (KC + HD)A^i - KAB^i. \tag{88}$$

Then

$$B^i = -\frac{A}{KD}\delta x^i + \left(\frac{KC + HD}{KA}\right)A^i \tag{89}$$

and

$$W^{ij}\lambda_j = (A/KD)(-A^i\delta x^j + \delta x^i A^j)\lambda_j. \tag{90}$$

Replacing (90) in (87) we have

$$\frac{\overset{*}{D}\delta x^i}{du} = -\tfrac{1}{4}(\delta^i_k\lambda_j + \delta^i_j\lambda_k)A^j\delta x^k, \tag{91}$$

i.e.,

$$\frac{d\delta x^i}{du} + (\overset{*}{\Gamma}{}^i_{jk} + \tfrac{1}{4}\delta^i{}_{(j}\lambda_{k)})\frac{dx^j}{du}\delta x^k = 0. \tag{92}$$

Thus in this particular case the Desargues transfer is simply the Levi-Civita transfer given by the new connection:

$$\Gamma^i_{jk} = \overset{*}{\Gamma}{}^i_{jk} + \tfrac{1}{4}\delta^i{}_{(j}\lambda_{k)}. \tag{93}$$

Γ^i_{jk} defines a new covariant derivative, with ∇ and D such that the Desargues transfer equation is

$$\frac{d\delta x^i}{du} + \Gamma^i{}_{jk}\delta x^j\frac{dx^k}{du} = \frac{D}{du}\delta x^i = 0. \tag{94}$$

Finally Eq. (85) can be written

$$\partial_i g_{jk} - \overset{*}{\Gamma}{}^h_{ji} g_{hk} - \tfrac{1}{4}\lambda_{(j}\delta^h_{i)} g_{hk} - \overset{*}{\Gamma}{}^h_{ki} g_{jh} - \tfrac{1}{4}\lambda_{(k}\delta^h_{i)} g_{jh} = 0, \tag{95}$$

and, using the new connection:

$$\nabla_i g_{ik} = 0. \tag{96}$$

From this equation one can immediately deduce that

$$\Gamma^i_{jk} = \left\{{}^{i}_{jk}\right\},$$

which proves the theorem.

Let us now see that the paths of the fundamental congruence are geodesics of the Riemannian V_4, by (61) and (80):

$$\frac{d^2x^i}{d\tau^2} = \frac{1}{(K')^2}\frac{D^2}{A^2}\left[\frac{d^2x^i}{du^2}\frac{D^2}{A^2} + \frac{dx^i}{du}\left(\frac{D^2}{A^2}\right)\right] . \tag{97}$$

Thus, bearing in mind (53), (86), and (87), the differential equations of the paths of the fundamental congruence are

$$\frac{d^2x^i}{d\tau^2} = -\overset{*}{\Gamma}{}^i_{jk}\frac{dx^j}{d\tau}\frac{dx^k}{d\tau} - \tfrac{1}{2}\frac{dx^j}{d\tau}\lambda_j\frac{dx^i}{d\tau}$$

$$= -(\overset{*}{\Gamma}{}^i_{jk} + \tfrac{1}{4}\delta^i_{(k}\lambda_{j)})\frac{dx^j}{du}\frac{dx^k}{du}, \tag{98}$$

i.e.,

$$\frac{D}{d\tau}\left(\frac{dx^i}{d\tau}\right) = 0. \tag{99}$$

Then follows the law of motion of particles: The second covariant derivative, in the new connection Γ is zero. The paths are then geodesics and the proper time τ is an affine parameter.

8. CONCLUSIONS

If we accept hypotheses H_0–H_6, the preceding measurement method leads us to conclude that the structure of space–time is a Riemannian manifold and to formulate the classical theory of Einstein. H_0 must be accepted as a reasonable hypothesis in the macroscopic world, although we know it is not true microscopically. H_1 is solidly based on the experiment of *Eötvös*. H_2 is a consequence of special relativity. H_3 is a reasonable formulation of the principle of equivalence. H_4 is a physical fact. H_5 and H_6 are the less solid hypotheses, any possible modifications of the Riemannian theory must be sought by rejecting one of them. In particular, if we eliminate H_6 and keep H_5 it can be shown that we reach the *Weyl geometry*.

Finally we must observe that if the manifold is non-Riemannian, the Desargues transfer does not coincide, in general, with the Levi-Civita transfer. This fact perhaps clarifies the failure of non-Riemannian theories that use the nonphysical Levi-Civita transfer.

ACKNOWLEDGMENTS

I wish to thank Professor G. Beck, who pointed out to me the interest of this research, and Professor C. Bollini, Professor C. Cattaneo, Professor V. Cantoni, Professor A. Lichnerowicz, Professor G. Tallini, and Professor L. A. Santaló for general valuable discussions on the subject of this article and the previous ones [cf. Refs. (3) and (5)].

---

[1] R. F. Marzke, "The Theory of Measurement in General Relativity," A. B. Senior Thesis (Princeton, 1959) (unpublished).

[2] R. F. Marzke and J. A. Wheeler, *Gravitation as a Geometry. The Geometry of Space-Time and Geometro-Dynamics Standard Meters* Enclosed in *Gravitation and Relativity* (Benjamin, New York, 1964)

[3] M. Castagnino, Rend. Accad. Naz. Lincei 44, 533 (1968).

[4] We indicate the complete symmetrization by

$$\underset{ijk}{S} T_{ijk} = \tfrac{1}{6}(T_{ijk} + T_{ikj} + T_{jik} + T_{jki} + T_{kij} + T_{kji}).$$

[5] M. Castagnino, Nuovo Cimento 54, 149 (1968).

[6] M. Castagnino, Math. Notae 3, 1, 117 (1965).

The Algebras of Representations and Classes of Finite Groups

G. de B. Robinson
University of Toronto, Toronto 5, Ontario, Canada
(Received 25 March 1971)

In a previous paper the author proposed a representation r for the algebra of representations of a finite group G. Here a corresponding representation c for the algebra of classes of G is established, and the relations between the two representations are investigated. In particular, the idempotents of r and c are studied in some detail.

1. INTRODUCTION

Having become interested in the *geometry of group representations*[1] as applied to theoretical physics, the author has attempted elsewhere[2] to show the significance of these ideas for finite groups. Subsequently, he discovered the papers by Gamba[3] and Killingbeck[4] in which some of the same results were described. The present paper develops the duality between the irreducible representations λ of G over the complex field and the classes C_i of G in the foregoing context.

2. REPRESENTATIONS OF THE DUAL ALGEBRAS

We begin then by assuming G to be of finite order g with classes C_i, each containing g_i elements with corresponding character $\mathbf{X}_i^{\lambda}$. We assume the table of characters $\mathbf{X}$ has rows λ and columns C_i, and write the reduction of the Kronecker or tensor product[5]

$$\lambda \times \mu = \lambda\mu = \mu\lambda = \sum_{\nu} g_{\lambda\mu}^{\nu}\, \nu \tag{2.1}$$

and the corresponding product of classes

$$C_i C_j = C_j C_i = \sum_k c_{ij}^k C_k, \tag{2.2}$$

where $C_1 = \mathbf{I}$.

If we denote the sum of the elements of the class C_i by the same symbol, it follows from Schur's lemma that the f^{λ} eigenvalues of the representation of C_i are all equal to γ_i^{λ}, called the *class multiplier* of C_i, where[6]

$$\gamma_i^{\lambda} = g_i \mathbf{X}_i^{\lambda}/f^{\lambda}.$$

It follows from (2.2) that

$$\gamma_i^{\lambda}\gamma_j^{\lambda} = \sum_k c_{ij}^k \gamma_k^{\lambda}. \tag{2.3}$$

As in the former paper,[2] the clue to what follows is the writing of the characters $\mathbf{X}_i^{\lambda}$ along the principal diagonal of a *diagonal matrix* D^{λ}, for fixed λ, and similarly the class multiplier γ_i^{λ} as a *diagonal matrix* D_i, for fixed i. We shall distinguish those formulas relating to the representations λ of G by the suffix r and those relating to the classes C_i of G by the suffix c. Thus from (2.1)–(2.3) we have

$$D^{\lambda'}D^{\mu'} = \sum_{\nu} g_{\lambda\mu}^{\nu} D^{\nu'} \tag{2.4r}$$

and

$$D_{i'}D_{j'} = \sum_k c_{ij}^k D_{k'}, \tag{2.4c}$$

where we denote the representation with conjugate complex character to that of λ by λ' and the class inverse to i by i'.

In what follows we think of the character table of G as a matrix $\mathbf{X}$, and we transform (2.4r) and (2.4c) to yield

$$\mathbf{X}D^{\lambda'}\mathbf{X}^{-1}\cdot\mathbf{X}D^{\mu'}\mathbf{X}^{-1} = \sum_{\nu} g_{\lambda\mu}^{\nu}\mathbf{X}D^{\nu'}\mathbf{X}^{-1} \tag{2.5r}$$

and

$$\mathbf{X}^{-1}D_{i'}\mathbf{X}\cdot\mathbf{X}^{-1}D_{j'}\mathbf{X} = \sum_k c_{ij}^k \mathbf{X}^{-1}D_{k'}\mathbf{X}. \tag{2.5c}$$

The desired representations of the dual algebras are obtained by setting

$$\{\lambda\} = \mathbf{X}D^{\lambda'}\mathbf{X}^{-1} = (g_{\beta\gamma}^{\lambda}) \tag{2.6r}$$

and

$$\{C_i\} = \mathbf{X}^{-1}D_{i'}\mathbf{X} = (c_{rs}^{i}). \tag{2.6c}$$

While relations analogous to (2.5r) and (2.5c) remain valid for transformation of (2.4r) and (2.4c) by any nonsingular matrix M, it is the special choice of $M = \mathbf{X}$ which yields the matrices $\{\lambda\}$ and $\{C_i\}$ with the $g_{\beta\gamma}^{\lambda}$ and c_{rs}^{i} as elements.

3. THEOREM AND COROLLARY

In order to prove the statements just made, it is necessary to recall the expressions in terms of the characters for the multiplicities $g_{\lambda\mu}^{\nu}$ and c_{ij}^{k}. We quote the result only[7]:

$$\begin{aligned} g_{\beta\gamma}^{\lambda} &= (1/g)\sum_i g_i \mathbf{X}_i^{\beta}\mathbf{X}_i^{\gamma}\mathbf{X}_i^{\lambda'} \\ &= \sum_i \mathbf{X}_i^{\beta}\mathbf{X}_i^{\lambda'}(g_i\mathbf{X}_i^{\gamma}/g), \end{aligned} \tag{3.1r}$$

which is symmetrical in β, γ, and λ' and corresponds to the matrix multiplication in (2.6r). Similarly, the value of c_{rs}^{i} is given by[8]

$$\begin{aligned} c_{rs}^{i} &= \frac{g_r g_i}{g}\sum_{\rho}\frac{1}{f^{\rho}}\mathbf{X}_r^{\rho}\mathbf{X}_{i'}^{\rho}\mathbf{X}_s^{\rho} \\ &= \sum_{\rho}\left(\frac{g_r\mathbf{X}_r^{\rho}}{g}\right)\gamma_{i'}^{\rho}\mathbf{X}_s^{\rho}, \end{aligned} \tag{3.1c}$$

which is symmetrical in r and i', but *not* in s, and corresponds to the matrix multiplication in (2.6c).

It follows from the above that we may obtain the desired matrix representations of the two algebras directly from the *rows* or *columns* of the tables of products, as will be illustrated in Sec. 5. In addition, the extra symmetry in (3.1r) leads to the

derivation already noted.[2] Thus we have proved the following.

Theorem 1: (i) $\{\lambda\}\{\mu\} = \sum_\nu g^\nu_{\lambda\mu}\{\nu\}$, (ii) $\{C_i\}\{C_j\} = \sum_k c^k_{ij}\{C_k\}$ with the following corollary.

Corollary: If for some γ (s) we have

$$\text{(i)}\quad g^\lambda_{\beta\gamma'} g^\mu_{\gamma\alpha'} \neq 0, \qquad \text{(ii)}\quad c^i_{rs'} c^j_{st'} \neq 0,$$

then for some $\nu(k)$ we have

$$\text{(i)}\quad g^\nu_{\lambda\mu} g^\nu_{\beta\alpha'} \neq 0, \qquad \text{(ii)}\quad c^k_{ij} c^k_{rt'} \neq 0,$$

which is illustrated in Sec. 6.

Perhaps it is worth writing out in detail the matrix multiplication involved in Theorem 1 (ii) to illustrate the complications which are involved. If we set

$$\{C_i\}\{C_j\} = (c^i_{rs'})(c^j_{st'}) = \sum_k c^k_{ij}\{C_k\},$$

then

$$\begin{aligned}
&\sum_s c^i_{rs'} c^j_{st'} \\
&\quad = \sum_s \left[\sum_\rho \left(\frac{g_r \mathbf{X}^\rho_r}{g}\right) \gamma^\rho_{i'} \mathbf{X}^\rho_{s'}\right]\left[\sum_\sigma \left(\frac{g_s \mathbf{X}^\sigma_s}{g}\right) \mathbf{X}^\sigma_{j'} \mathbf{X}^\sigma_{t'}\right] \\
&\quad = \sum_\rho \left(\frac{g_r \mathbf{X}^\rho_r}{g}\right)(\gamma^\rho_{i'} \gamma^\rho_{j'}) \mathbf{X}^\rho_{t'} \\
&\quad = \sum_k c^k_{ij} \left[\sum_\rho \left(\frac{g_r \mathbf{X}^\rho_r}{g}\right) \gamma^\rho_{k'} \mathbf{X}^\rho_{t'}\right] \\
&\quad = \sum_k c^k_{ij} c^k_{rt'}, \qquad (3.2c)
\end{aligned}$$

using the orthogonality relations of the characters. A similar reduction applies in the case of Theorem 1(i).

It is evident from (2.6r) that the eigenvalues of $\{\lambda\}$ are the elements of the λ' *row* of X. Since the eigenvectors of $D^{\lambda'}$ are the $E_i(0, \ldots, 0, 1, 0, \ldots, 0)$, it follows that the eigenvectors of $\{\lambda\}$ are the $\mathbf{X}E_i$, i.e., the corresponding *columns* of X. Similarly, the eigenvalues of $\{C_i\}$ are the elements of the C_i *column* of γ and the eigenvectors of $\{C_i\}$ are the corresponding *rows* of γ. This last statement follows from the fact that an eigenvector of $\{C_i\}$ must have the form $\mathbf{X}^{-1}E^\rho = (f^\rho/g)\mathbf{X}^\rho_i$, where f^ρ/g is constant for varying i.

4. IDEMPOTENTS OF THE ALGEBRAS

Crucial in the development of the representation theory of any finite group G is the notion of the *idempotent* of the group algebra[9]

$$T^\rho = \frac{f^\rho}{g} \sum_i \mathbf{X}^\rho_i C_{i'}, \qquad T^\rho T^\sigma = \begin{cases} T^\rho & \text{for } \sigma = \rho \\ 0 & \text{for } \sigma \neq \rho \end{cases},$$

$$\sum_\rho T^\rho = I. \qquad (4.1c)$$

Substituting $\{C_i\}$ for $C_{i'}$, we obtain

$$\begin{aligned}
\mathbf{X}\{T^\rho\}\mathbf{X}^{-1} &= \frac{f^\rho}{g} \sum_i \mathbf{X}^\rho_i (\mathbf{X}\{C_i\}\mathbf{X}^{-1}) \\
&= \frac{f^\rho}{g} \sum_i \mathbf{X}^\rho_i D_{i'}, \qquad (4.2c)
\end{aligned}$$

which is an idempotent I^ρ of the matrix algebra with 1 in the ρ place of the diagonal and zeros elsewhere.

If we seek to construct the analog of (4.1c) involving representations ρ instead of classes C_i, we run into difficulties. However, their *representations* $\{\rho\}$ are all of the same dimension, and so we may set

$$\{S_i\} = \frac{g_i}{g} \sum_\rho \mathbf{X}^\rho_i \{\rho\}$$

and, transforming again by X, we have

$$\begin{aligned}
\mathbf{X}^{-1}\{S_i\}\mathbf{X} &= \frac{g_i}{g} \sum_\rho \mathbf{X}^\rho_i \mathbf{X}^{-1}\{\rho\}\mathbf{X} \\
&= \frac{g_i}{g} \sum_\rho \mathbf{X}^\rho_i D^{\rho'}, \qquad (4.2r)
\end{aligned}$$

which is also an idempotent I_i with 1 in the i place of the diagonal and zeros elsewhere.

One continues to hope for an explicit mechanism, analogous to the Young diagram for S_n, to relate the classes to the representations of an arbitrary group G, even if it does not extend to products!

5. EXAMPLE

We illustrate these ideas in the case of the group of order 10 defined by[10]

$$S^5 = T^2 = I, \qquad TST = S^{-1},$$

where

$$C_1 = I, \quad C_2 = S, S^{-1}, \quad C_3 = S^2, S^{-2}, \quad C_4 = T, TS, TS^2, TS^3, TS^4.$$

Setting $\rho = \frac{1}{2}(-1 + \sqrt{5})$ and $\sigma = \frac{1}{2}(-1 - \sqrt{5})$, we have the following tables for X and γ:

$$\begin{array}{c} \\ a \\ b \\ c \\ d \end{array} \begin{array}{c} \begin{array}{cccc} C_1 & C_2 & C_3 & C_4 \end{array} \\ \begin{pmatrix} 1 & 1 & 1 & 1 \\ 2 & \rho & \sigma & 0 \\ 2 & \sigma & \rho & 0 \\ 1 & 1 & 1 & -1 \end{pmatrix} \end{array} : \mathbf{X} \qquad \begin{pmatrix} 1 & 2 & 2 & 5 \\ 1 & \rho & \sigma & 0 \\ 1 & \sigma & \rho & 0 \\ 1 & 2 & 2 & -5 \end{pmatrix} : \gamma,$$

whence we derive the multiplication tables

$$\begin{array}{c} \\ a \\ b \\ c \\ d \end{array}\!\!\begin{array}{c} \begin{array}{cccc} a & b & c & d \end{array} \\ \begin{pmatrix} a & b & c & d \\ b & a+c+d & b+c & b \\ c & b+c & a+b+d & c \\ d & b & c & a \end{pmatrix} \end{array}\!\!: \pi_r \qquad \begin{array}{c} \\ C_1 \\ C_2 \\ C_3 \\ C_4 \end{array}\!\!\begin{array}{c} \begin{array}{cccc} C_1 & C_2 & C_3 & C_4 \end{array} \\ \begin{pmatrix} C_1 & C_2 & C_3 & C_4 \\ C_2 & 2C_1+C_3 & C_2+C_3 & 2C_4 \\ C_3 & C_2+C_3 & 2C_1+C_2 & 2C_4 \\ C_4 & 2C_4 & 2C_4 & 5C_1+5C_2+5C_3 \end{pmatrix} \end{array}\!\!: \pi_c .$$

From π_r we obtain

$$\{a\} = \begin{pmatrix} 1&0&0&0 \\ 0&1&0&0 \\ 0&0&1&0 \\ 0&0&0&1 \end{pmatrix}, \quad \{b\} = \begin{pmatrix} 0&1&0&0 \\ 1&0&1&1 \\ 0&1&1&0 \\ 0&1&0&0 \end{pmatrix}, \quad \{c\} = \begin{pmatrix} 0&0&1&0 \\ 0&1&1&0 \\ 1&1&0&1 \\ 0&0&1&0 \end{pmatrix}, \quad \{d\} = \begin{pmatrix} 0&0&0&1 \\ 0&1&0&0 \\ 0&0&1&0 \\ 1&0&0&0 \end{pmatrix},$$

so that

$$\{S_1\} = \tfrac{1}{10}\begin{pmatrix} 1&2&2&1 \\ 2&4&4&2 \\ 2&4&4&2 \\ 1&2&2&1 \end{pmatrix}, \quad \{S_2\} = \tfrac{2}{10}\begin{pmatrix} 1&\rho&\sigma&1 \\ \rho&\rho^2&\rho\sigma&\rho \\ \sigma&\rho\sigma&\sigma^2&\sigma \\ 1&\rho&\sigma&1 \end{pmatrix},$$

$$\{S_3\} = \tfrac{2}{10}\begin{pmatrix} 1&\sigma&\rho&1 \\ \sigma&\sigma^2&\sigma\rho&\sigma \\ \rho&\sigma\rho&\rho^2&\rho \\ 1&\sigma&\rho&1 \end{pmatrix}, \quad \{S_4\} = \tfrac{5}{10}\begin{pmatrix} 1&0&0&-1 \\ 0&0&0&0 \\ 0&0&0&0 \\ -1&0&0&1 \end{pmatrix},$$

where, e.g.,

$$\{b\}\{b\} = \begin{pmatrix} 1&0&1&1 \\ 0&3&1&0 \\ 1&1&2&1 \\ 1&0&1&1 \end{pmatrix} = \{a\} + \{c\} + \{d\}$$

and

$$\{S_i\}\{S_j\} = \begin{cases} \{S_i\} & \text{for} \quad i = j \\ 0 & \text{for} \quad i \neq j \end{cases}$$

with

$$\{S_1\} + \{S_2\} + \{S_3\} + \{S_4\} = I.$$

Similarly, from π_c,

$$\{C_i\} = \begin{pmatrix} 1&0&0&0 \\ 0&1&0&0 \\ 0&0&1&0 \\ 0&0&0&1 \end{pmatrix}, \quad \{C_2\} = \begin{pmatrix} 0&1&0&0 \\ 2&0&1&0 \\ 0&1&1&0 \\ 0&0&0&2 \end{pmatrix}, \quad \{C_3\} = \begin{pmatrix} 0&0&1&0 \\ 0&1&1&0 \\ 2&1&0&0 \\ 0&0&0&2 \end{pmatrix}, \quad \{C_4\} = \begin{pmatrix} 0&0&0&1 \\ 0&0&0&2 \\ 0&0&0&2 \\ 5&5&5&0 \end{pmatrix},$$

$$\{T^a\} = \tfrac{1}{10}\begin{pmatrix} 1&1&1&1 \\ 2&2&2&2 \\ 2&2&2&2 \\ 5&5&5&5 \end{pmatrix}, \quad \{T^b\} = \tfrac{2}{10}\begin{pmatrix} 2&\rho&\sigma&0 \\ 2\rho&\rho^2&\sigma\rho&0 \\ 2\sigma&\sigma\rho&\sigma^2&0 \\ 0&0&0&0 \end{pmatrix}, \quad \{T^c\} = \tfrac{2}{10}\begin{pmatrix} 2&\sigma&\rho&0 \\ 2\sigma&\sigma^2&\rho\sigma&0 \\ 2\rho&\rho\sigma&\rho^2&0 \\ 0&0&0&0 \end{pmatrix},$$

$$\{T^d\} = \tfrac{1}{10}\begin{pmatrix} 1 & 1 & 1 & -1 \\ 2 & 2 & 2 & -2 \\ 2 & 2 & 2 & -2 \\ -5 & -5 & -5 & 5 \end{pmatrix},$$

where, e.g.,

$$\{C_2\}\{C_2\} = \begin{pmatrix} 2 & 0 & 1 & 0 \\ 0 & 3 & 1 & 0 \\ 2 & 1 & 2 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix} = 2\{C_1\} + \{C_3\}$$

and

$$\{T^i\}\{T^j\} = \begin{cases} \{T^i\} & \text{for} \quad i = j \\ 0 & \text{for} \quad i \neq j \end{cases}$$

with $\{T^a\} + \{T^b\} + \{T^c\} + \{T^d\} = I$. Here

$$X^{-1}\{S_1\}X = X\{T^a\}X^{-1}, \quad X^{-1}\{S_2\}X = X\{T^b\}X^{-1}, \quad X^{-1}\{S_3\}X = X\{T^c\}X^{-1}, \quad X^{-1}\{S_4\}X = X\{T^d\}X^{-1}.$$

The form of the idempotent matrices suggests

Theorem 2: The idempotents of the representations r and c can be written as Kronecker products of suitably chosen row and column vectors.

The proof follows immediately by rewriting (4.2r) and (4.2c) in the forms

$$\{S_i\} = XI_iX^{-1}, \tag{5.1r}$$

$$\{T^\rho\} = X^{-1}I^\rho X, \tag{5.1c}$$

where I_i and I^ρ are defined in § 4.

6. COROLLARY (ii)

Having illustrated Corollary (i) in the previous paper,[2] we confine our attention here to (ii), and associate the six classes of S_5, omitting C_1, with the accompanying Pasch figure (see Fig. 1). Since $S_5 = A_5 + (12)A_5$, it is clear that the even classes $(3, 1^2)$, (2^21), and (5) must be collinear while the odd ones $(2, 1^3)$, $(4, 1)$, and $(3, 2)$ may be associated with the three remaining points of the figure in all possible ways. One could develop the analog of the $6j$-symbol here also.

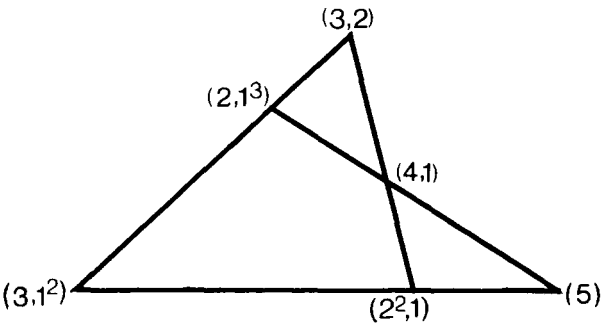


FIG. 1

[1] G. de B. Robinson, J. Math. Phys. 11, 3428 (1970).
[2] G. de B. Robinson, "Tensor Product Representations," J. Algebra (to be published).
[3] A. Gamba, J. Math. Phys. 9, 186 (1968).
[4] J. Killingbeck, J. Math. Phys. 11, 2268 (1970).
[5] W. Burnside, *The Theory of Groups* (Cambridge U.P., Cambridge, 1910), Chap. 15.
[6] M. Hamermesh, *Group Theory* (Addison-Wesley, Reading, Mass., 1962), p. 109.
[7] Reference 5, p. 291.
[8] Reference 5, p. 285.
[9] Reference 5, p. 305.
[10] Reference 5, p. 296.

Perturbation Formulas*

Masataka Mizushima
Department of Physics and Astrophysics, University of Colorado, Boulder, Colorado
(Received 12 May 1971)

A simple form of perturbation formula is shown to be correct to the fifth-order of the perturbation. It also has the variational character, namely, it gives an upper limit for a correct eigenvalue when applied to a bound state. The formula is in the form of a Padé approximant. The next approximation is correct to the seventeenth order and has the same variational character. Some examples are given and compared to conventional theories.

1. INTRODUCTION

The perturbation theory of Schrödinger[1] is so powerful and useful that it is discussed in almost all text books on quantum mechanics. The explicit expression of the perturbation series, however, is not commonly given beyond the second-order term. An expression for the higher-order terms is given in Condon and Shortley's book[2] but is quite complicated. Brillouin and Wigner[3] gave a simple expression for the perturbation series, but in actual calculations one has to reduce it to the complicated formula of Condon and Shortley. A modification of the Brillouin–Wigner method is given in the form of a continued fraction,[4] which makes its application easier. All existing formulations are applicable only to the nondegenerate case. In this paper we present perturbation formulas which are simple enough for practical applications and yet accurate. This scheme is also applicable to the degenerate case.

We assume that

$$\hat{H} = \hat{H}_0 + \hat{H}_1 \equiv \hat{H}_0 + \alpha\hat{h}_1, \tag{1}$$

where α is a parameter which indicates the strength of the perturbation $\hat{H}_1$. We also assume that all eigenvalues and eigenfunctions of $\hat{H}_0$ are known:

$$\hat{H}_0|n^0\rangle = \epsilon_n^0|n^0\rangle, \quad \hat{H}_0|m^0\rangle = \epsilon_m^0|m^0\rangle, \ldots. \tag{2}$$

The problem is to find approximate expressions for the eigenvalues ϵ_n of $\hat{H}$,

$$\hat{H}|n\rangle = \epsilon_n|n\rangle \tag{3}$$

by means of $\epsilon_n^0, \epsilon_m^0, \ldots$, and $|n^0\rangle, |m^0\rangle, \ldots$.

2. FIRST-ORDER PERTURBATION

If we express the Hamiltonian matrix by taking the eigenvectors of $\hat{H}_0$ as the basis, we see that the diagonal matrix elements are of the form $\epsilon_n^0 + \langle n^0|\hat{H}_1|n^0\rangle$, while the nondiagonal matrix elements are of the form $\langle n^0|\hat{H}_1|m^0\rangle$. The secular equation is then

$$(\epsilon_n^0 + \langle n^0|\hat{H}_1|n^0\rangle - \epsilon)(\epsilon_m^0 + \langle m^0|\hat{H}_1|m^0\rangle - \epsilon)(\cdots - \epsilon) + O(\alpha^2) = 0, \tag{4}$$

where ϵ is the unknown and the last term $O(\alpha^2)$ represents all terms in which nondiagonal matrix elements are involved. Since nondiagonal matrix elements always appear in products such as $\langle n^0|\alpha\hat{h}_1|m^0\rangle\langle m^0|\alpha\hat{h}_1|n^0\rangle$ in the expansion (4), $O(\alpha^2)$ is of the order of α^2 at least. If there is no degeneracy,

$$\epsilon_n^{(1)} = \epsilon_n^0 + \langle n^0|\hat{H}_1|n^0\rangle \tag{5}$$

is correct to the first order in α, because $\epsilon_m^0 + \langle m^0|\hat{H}_1|m^0\rangle - \epsilon_n^{(1)}$ is of the order of α^0 for all m's except for n.

3. SPACE CONTRACTION OPERATOR

Let us define an operator $\hat{S}_n^0$, called a space contraction operator

$$\hat{S}_n^0 = \hat{1} + (\epsilon_n^0 - \hat{H}_0 - \hat{Q}_n\hat{H}_1)^{-1}\hat{Q}_n\hat{H}_1 = \hat{1} + (\epsilon_n^0 - \hat{H}_0)^{-1}\hat{Q}_n\hat{H}_1\hat{S}_n^0, \tag{6}$$

where $\hat{1}$ is the identity operator and $\hat{Q}_n$ is a projection operator defined by

$$\hat{Q}_n = \hat{1} - |n^0\rangle\langle n^0|. \tag{7}$$

The well-known property of a resolvent

$$(\epsilon_n^0 - \hat{H}_0 - \hat{Q}_n\hat{H}_1)^{-1} = (\epsilon_n^0 - \hat{H}_0)^{-1} + (\epsilon_n^0 - \hat{H}_0)^{-1} \times \hat{Q}_n\hat{H}_1(\epsilon_n^0 - \hat{H}_0 - \hat{Q}_n\hat{H}_1)^{-1} \tag{8}$$

is used in obtaining the last expression of (6).

It is easy to see that

$$(\hat{H} - \epsilon_n^0)\hat{S}_n^0 = \hat{H}_0 - \epsilon_n^0 + (\hat{1} - \hat{Q}_n)\hat{H}_1\hat{S}_n^0 \tag{9}$$

holds.

4. THIRD-ORDER PERTURBATION

Let us transform the original vectors $|n^0\rangle, |m^0\rangle, \cdots$ into $\hat{S}_n^0|n^0\rangle, \hat{S}_m^0|m^0\rangle, \ldots$. We will show that the nondiagonal elements $\langle n^0|\hat{S}_n^{0\dagger}\hat{S}_m^0|m^0\rangle$ and $\langle n^0|\hat{S}_n^{0\dagger}\hat{H}\hat{S}_m^0|m^0\rangle$ are all of the order α^2 or higher.

By a straightforward calculation we obtain

$$\hat{S}_n^{0\dagger}\hat{S}_m^0 = \hat{1} + (\epsilon_m^0 - \hat{H}_0)^{-1}\hat{Q}_m\hat{H}_1 + \hat{H}_1\hat{Q}_n(\epsilon_n^0 - \hat{H}_0)^{-1} + [(\epsilon_m^0 - \hat{H}_0)^{-1}\hat{Q}_m\hat{H}_1 - (\epsilon_m^0 - \epsilon_n^0)^{-1}\hat{H}_1\hat{Q}_n] \times (\epsilon_m^0 - \hat{H}_0)^{-1}\hat{Q}_m\hat{H}_1 + \hat{H}_1\hat{Q}_n(\epsilon_n^0 - \hat{H}_0)^{-1} \times [\hat{H}_1\hat{Q}_n(\epsilon_n^0 - \hat{H}_0)^{-1} - \hat{Q}_m\hat{H}_1(\epsilon_n^0 - \epsilon_m^0)^{-1}] + O(\alpha^3), \tag{10}$$

where the last term $O(\alpha^3)$ represents all terms which are of the order α^3 or higher. Therefore

$$\langle n^0|\hat{S}_n^{0\dagger}\hat{S}_m^0|m^0\rangle = (\epsilon_m^0-\epsilon_n^0)^{-2}\,(\langle n^0|\hat{H}_1|n^0\rangle + \langle m^0|\hat{H}_1|m^0\rangle)\langle n^0|\hat{H}_1|m^0\rangle + O(\alpha^3), \quad \text{if } m\neq n. \tag{11}$$

The transformed vectors $\hat{S}_n^0|n^0\rangle, \hat{S}_m^0|m^0\rangle, \cdots$ are not orthogonal to each other, but the nonorthogonality is of the order α^2 or higher.

In a similar way we obtain

$$\langle n^0|\hat{S}_n^{0\dagger}\hat{H}\hat{S}_m^0|m^0\rangle = (\epsilon_m^0\langle n^0|\hat{H}_1|n^0\rangle + \epsilon_n^0\langle m^0|\hat{H}_1|m^0\rangle)\times\langle n^0|\hat{H}_1|m^0\rangle\,(\epsilon_m^0-\epsilon_n^0)^{-2} + O(\alpha^3), \text{ if } m\neq n, \tag{12}$$

which is of the order α^2 or higher.

In the matrix of $\hat{H}-\epsilon\hat{1}$ we therefore see that all nondiagonal matrix elements are of the order α^2 or higher, if we use $\hat{S}_n^0|n^0\rangle, \hat{S}_m^0|m^0\rangle, \cdots$ as the basis. Applying the same argument as we gave in relation to (4), we see that

$$\epsilon_n^{(3)} = \langle n^0|\hat{S}_n^{0\dagger}\hat{H}\hat{S}_n^0|n^0\rangle/\langle n^0|\hat{S}_n^{0\dagger}\hat{S}_n^0|n^0\rangle \tag{13}$$

is correct to the order α^3.

5. FIFTH-ORDER PERTURBATION

When we look at (11) and (12) we see that if we can get rid of the diagonal elements of $\hat{H}_1$ in the original representation, we can make all nondiagonal matrix elements of $\hat{H}-\epsilon\hat{1}$ of the order α^3 or higher, and, therefore, obtain an expression for the approximate eigenvalue which is correct to the order of α^5. It is easy to see that this can be actually done if we reformulate the problem by replacing (1) by

$$\hat{H} = \hat{H}_0' + \hat{H}_1', \tag{14a}$$

where

$$\hat{H}_0'|n^0\rangle = \epsilon_n^{(1)}|n^0\rangle \tag{14b}$$

that is, making the perturbation $\hat{H}_1'$ purely nondiagonal, and take

$$\hat{S}_n^{(1)} = \hat{1} + (\epsilon_n^{(1)}-\hat{H}_0')^{-1}\hat{Q}_n\hat{H}_1'\hat{S}_n^{(1)} \tag{15}$$

for the space contraction operator.

Following the same argument as before, we see that

$$\epsilon_n^{(5)} = \langle n^0|\hat{S}_n^{(1)\dagger}\hat{H}\hat{S}_n^{(1)}|n^0\rangle/\langle n^0|\hat{S}_n^{(1)\dagger}\hat{S}_n^{(1)}|n^0\rangle \tag{16}$$

is correct to the order of α^5.

6. SEVENTEENTH-ORDER PERTURBATION

Although the fifth-order formula is accurate enough for almost all practical applications of perturbation theory, it is worthwhile to indicate how we can obtain higher-order perturbation formulas.

We have seen that when $\hat{S}_n^{(1)}|n^0\rangle, \hat{S}_m^{(1)}|m^0\rangle, \cdots$ are taken as the basis, all nondiagonal matrix elements of $\hat{H}-\epsilon\hat{1}$ are of the order α^3 or higher. If we reformulate the problem by replacing (14a) by

$$\hat{H} = \hat{H}_0^{(5)} + \hat{H}_1^{(5)}, \tag{17}$$

where

$$\hat{H}_0^{(5)}|n^0\rangle = \epsilon_n^{(5)}|n^0\rangle \tag{18}$$

and

$$\langle m^0|\hat{H}_1^{(5)}|n^0\rangle = \langle m^0|\hat{S}_m^{(1)\dagger}\hat{H}\hat{S}_n^{(1)}|n^0\rangle/\times[\langle m^0|\hat{S}_m^{(1)\dagger}\hat{S}_m^{(1)}|m^0\rangle\,\langle n^0|\hat{S}_n^{(1)\dagger}\hat{S}_n^{(1)}|n^0\rangle]^{1/2} \tag{19}$$

and by taking

$$\hat{S}_n^{(5)} = \hat{1} + (\epsilon_n^{(5)}-\hat{H}_0^{(5)})^{-1}\hat{Q}_n\hat{H}_1^{(5)}\hat{S}_n^{(5)} \tag{20}$$

instead of (15), then all nondiagonal matrix elements of $\hat{H}-\epsilon\hat{1}$ in the $\hat{S}_n^{(5)}|n^0\rangle$ representation must be of the order $(\alpha^3)^3=\alpha^9$ or higher. Therefore, we can neglect all nondiagonal matrix elements in the resultant secular equation up to the order of α^{17} to obtain

$$\epsilon_n^{(17)} = \langle n^0|\hat{S}_n^{(5)\dagger}\hat{H}\hat{S}_n^{(5)}|n^0\rangle/\langle n^0|\hat{S}_n^{(5)\dagger}\hat{S}_n^{(5)}|n^0\rangle \tag{21}$$

from the diagonal element only.

7. PERTURBATION SERIES

From (9) we obtain

$$\langle n^0|\hat{S}_n^{0\dagger}\hat{H}\hat{S}_n^0|n^0\rangle = \epsilon_n^0\langle n^0|\hat{S}_n^{0\dagger}\hat{S}_n^0|n^0\rangle + \langle n^0|\hat{H}_1\hat{S}_n^0|n^0\rangle, \tag{22}$$

therefore,

$$\epsilon_n^{(3)} = \epsilon_n^0 + \frac{\langle n^0|\hat{H}_1\hat{S}_n^0|n^0\rangle}{\langle n^0|\hat{S}_n^{0\dagger}\hat{S}_n^0|n^0\rangle}. \tag{23}$$

Corresponding to (9) we easily obtain

$$(\hat{H}-\epsilon_n^{(1)})\hat{S}_n^{(1)} = \hat{H}_0-\epsilon_n^{(1)} + (\hat{1}-\hat{Q}_n)\hat{H}_1\hat{S}_n^{(1)}, \tag{24}$$

so that

$$\epsilon_n^{(5)} = \epsilon_n^{(1)} + \frac{\langle n^0|\hat{H}_1'\hat{S}_n^{(1)}|n^0\rangle}{\langle n^0|\hat{S}_n^{(1)\dagger}\hat{S}_n^{(1)}|n^0\rangle}. \tag{25}$$

In the same way we obtain

$$\epsilon_n^{(16)} = \epsilon_n^{(5)} + \frac{\langle n^0|\hat{H}_1^{(5)}\hat{S}_n^{(5)}|n^0\rangle}{\langle n^0|\hat{S}_n^{(5)\dagger}\hat{S}_n^{(5)}|n^0\rangle}. \tag{26}$$

Explicit expressions are obtained by using (6), (15), and (20). For example, (6) gives

$$\hat{S}_n^0|n^0\rangle = |n^0\rangle + \sum_{m\neq n}\frac{|m^0\rangle(mn)}{m} + \sum_{m\neq n,l\neq n}\frac{|l^0\rangle(lm)(mn)}{lm} + \sum_{m\neq n,l\neq n,k\neq n}\frac{|k^0\rangle(kl)\,(lm)(mn)}{klm} + \cdots, \tag{27}$$

where

$(mn) = \langle m^0|\hat{H}_1|n^0\rangle$ in the numerators, and
$m = \epsilon_n^0 - \epsilon_m^0$ in the denominators. (28)

Therefore

$$\epsilon_n^{(3)} = \epsilon_n^0 + \frac{(nn) + \sum_{m\neq n}[(nm)(mn)/m] + \sum_{m\neq n,l\neq n}[(nl)(lm)(mn)/lm] + \cdots}{1 + \sum_{m\neq n}[(nm)(mn)/m^2] + \cdots}. \tag{29}$$

In the same way we obtain

$$\epsilon_n^{(5)} = \epsilon_n^{(1)} + \Big\{\sum_{m\neq n}{}'[(nm)\,(mn)/m] + \sum_{m\neq n,l\neq n}{}'[(nl)(lm)(mn)/lm] + \sum_{m\neq n,l\neq n,k\neq n}[(nk)\,(kl)(lm)(mn)/klm] + \sum_{m\neq n,l\neq n,k\neq n,h\neq n}{}'(nh)(hk)(kl)(lm)(mn)/hklm] + \cdots\Big\}\Big\{1 + \sum_{m\neq n}{}'[(nm)(mn)/m^2] + \sum_{m\neq n,l\neq n}{}'(nm)(ml)\,(ln)(1/m^2l + 1/ml^2) + \cdots\Big\}^{-1}, \tag{30}$$

where diagonal elements of $\hat{H}_1$ are excluded from the summations, and

$m = \epsilon_n^{(1)} - \epsilon_m^{(1)}$ in the denominators. (31)

A considerably more complicated expression can be obtained for $\epsilon_n^{(1\,6)}$ when in (30) we replace $\epsilon_n^{(1)}$ by $\epsilon_n^{(5)}$ in front of the fraction and take

$m = \epsilon_n^{(5)} - \epsilon_m^{(5)}$ in the denominators (32)

and substitute (kh) in the numerators by

$$\Big\{\sum{}''(kl)\,(lj)\,(jh)\Big(\frac{k'}{(h-k)\,(h-l)\,(h-j)} + \frac{l'}{(k-l)\,(h-l)\,(h-j)} + \frac{j'}{(k-l)\,(k-j)\,(j-j)} + \frac{h'}{(k-l)\,(k-j)\,(k-h)} + \frac{1}{(h-l)\,(h-j)} + \frac{1}{(k-l)\,(h-j)} + \frac{1}{(k-l)\,(k-j)}\Big) + \sum{}''(kl)\,(lj)\,(ji)\,(ih)(\cdots) + \cdots\Big] \times\Big(1 + \sum_{l\neq k}{}'(kl)\,(lk)\,\frac{1}{(k-l)^2} + \cdots\Big)^{-1/2} \times\Big(1 + \sum_{l\neq h}{}'(hl)\,(lh)\,\frac{1}{(h-l)^2} + \cdots\Big)^{-1/2}, \tag{33}$$

where (kl) and $\sum'$ have the same meanings as before, but

$(h-k) = \epsilon_h^{(1)} - \epsilon_k^{(1)}$ in the denominators, (34)
$k' = \epsilon_k^{(1)}$ in the numerators, (35)

and $\sum''$ is a summation over all intermediate states excluding diagonal matrix elements of $\hat{H}_1$ and those intermediate states for which the corresponding denominators are zero.

8. DEGENERATE CASE

Let us assume that $|n_1^0\rangle, |n_2^0\rangle, \ldots, |n_b^0\rangle$ have a common eigenvalue ϵ_n^0 of $\hat{H}_0$. Since the argument in relation to (4) still holds even if such degeneracy exists, formula (5) is correct to the first-order of $\hat{H}_1$ without any modification. In order to obtain a third-order perturbation formula we replace $\hat{Q}_n$ of (7) by

$$\hat{Q}_N = \hat{1} - |n_1^0\rangle\langle n_1^0| - |n_2^0\rangle\langle n_2^0| - \cdots - |n_b^0\rangle\langle n_b^0| \tag{36}$$

and replace $\hat{S}_n^0$ of (6) by

$$\hat{S}_N^0 = \hat{1} + (\epsilon_n^0 - \hat{H}_0)^{-1}\hat{Q}_N\hat{H}_1\hat{S}_N^0. \tag{37}$$

We obtain

$$\langle m^0|\hat{S}_m^{0\dagger}\hat{S}_N^0|n_i^0\rangle = (\epsilon_n^0 - \epsilon_m^0)^{-2}(\langle m^0|\hat{H}_1|m^0\rangle \times\langle m^0|\hat{H}_1|n_i^0\rangle + \sum_{j=1}^{b}\langle m^0|\hat{H}_1|n_j^0\rangle\langle n_j^0|\hat{H}_1|n_i^0\rangle) + O(\alpha^3), \quad \text{if } m \text{ is not among } n_1,\ldots,n_b \tag{38}$$

and

$$\langle m^0|\hat{S}_m^{0\dagger}\hat{H}\hat{S}_N^0|n_i^0\rangle = (\epsilon_n^0 - \epsilon_m^0)^{-2}(\epsilon_n^0\langle m^0|\hat{H}_1|m^0\rangle \times\langle m^0|\hat{H}_1|n_i^0\rangle) + \epsilon_m^0\sum_{j=1}^{b}\langle m^0|\hat{H}_1|n_j^0\rangle\langle n_j^0|\hat{H}_1|n_i^0\rangle + O(\alpha^3), \quad \text{if } m \text{ is not among } n_1,\ldots,n_b. \tag{39}$$

If we solve the secular equation

$$\begin{vmatrix} \langle n_1^0|\hat{S}_N^{0\dagger}\hat{H}\hat{S}_N^0|n_1^0\rangle - \epsilon\langle n_1^0|\hat{S}_N^{0\dagger}\hat{S}_N^0|n_1^0\rangle & \cdots & \langle n_b^0|\hat{S}_N^{0\dagger}\hat{H}\hat{S}_N^0|n_1^0\rangle - \epsilon\langle n_b^0|\hat{S}_N^{\dagger}\hat{S}_N|n_1^0\rangle \\ \langle n_1^0|\hat{S}_N^{0\dagger}\hat{H}\hat{S}_N^0|n_2^0\rangle - \epsilon\langle n_1^0|\hat{S}_N^{0\dagger}\hat{S}_N^0|n_2^0\rangle & \cdots & \\ \cdots & & \\ \langle n_1^0|\hat{S}_N^{0\dagger}\hat{H}\hat{S}_N^0|n_b^0\rangle - \epsilon\langle n_1^0|\hat{S}_N^{0\dagger}\hat{S}_N^0|n_b^0\rangle & \cdots & \langle n_b^0|\hat{S}_N^{0\dagger}\hat{H}\hat{S}_N^0|n_b^0\rangle - \epsilon\langle n_b^0|\hat{S}_N^{0\dagger}\hat{S}_N^0|n_b^0\rangle \end{vmatrix} = 0, \tag{40}$$

then these solutions are correct to the order of α^3, because all nondiagonal matrix elements, except for those in the $b \times b$ submatrix, remain of the order α^2 or higher during the process of diagonalizing the $b \times b$ submatrix.

Formulas (38) and (39) show that if $\hat{H}_1$ has no matrix element at all inside the $b \times b$ submatrix, then the above procedure gives solutions which are correct to the order of α^5. If $\hat{H}_1$ has any nonzero matrix element inside the $b \times b$ submatrix, then we should diagonalize this submatrix first to obtain a fifth-order perturbation formula. The diagonalization will remove the degeneracy, and the problem in this case will be reduced to that of the nondegenerate case.

9. EXAMPLES

In the one-dimensional simple harmonic oscillator, where

$$\hat{H}_0 = (\hat{p}_x^2/2m) + \tfrac{1}{2}kx^2, \tag{41}$$

we may take

$$\hat{H}_1 = \alpha x \tag{42}$$

as a perturbation. In this case we know the exact solution as

$$\epsilon_n = \hbar\omega(n + \tfrac{1}{2}) - \tfrac{1}{2}(\alpha^2/k). \tag{43}$$

Our formula (30) gives

$$\epsilon_n^{(5)} = \hbar\omega(n + \tfrac{1}{2}) - \frac{\tfrac{1}{2}(\alpha^2/k) + \tfrac{1}{4}(2n + 1)(\alpha^2/k)^2 + \tfrac{1}{16}(5n^2 + 5n + 3)(\alpha^2/k)^3 + \cdots}{1 + \tfrac{1}{2}(2n + 1)(\alpha^2/k) + \tfrac{5}{8}(n^2 + n + 1)(\alpha^2/k)^2 + \cdots}$$

$$= \hbar\omega(n + \tfrac{1}{2}) - \tfrac{1}{2}(\alpha^2/k)[1 - \tfrac{1}{4}(\alpha^2/k)^2 + \cdots], \tag{44}$$

which deviate from (43) only in α^6 and higher terms.

It is interesting to compare this result to that of the Brillouin–Wigner perturbation method.[3] In their method, an approximate eigenvalue which is correct to the fifth order of α is given from

$$\epsilon_n' = \frac{\tfrac{1}{2}(n + 1)\alpha^2}{\epsilon_n' - \hbar\omega}\left(\frac{\hbar}{m\omega}\right) + \frac{\tfrac{1}{2}n\alpha^2}{\epsilon_n' + \hbar\omega}\left(\frac{h}{m\omega}\right) + \frac{\tfrac{1}{4}(n + 1)(n + 2)(\hbar/m\omega)^2\alpha^4}{(\epsilon_n' - \hbar\omega)^2(\epsilon_n' - 2\hbar\omega)} + \frac{\tfrac{1}{4}n(n - 1)(\hbar/m\omega)^2\alpha^4}{(\epsilon_n' + \hbar\omega)^2(\epsilon_n' + 2\hbar\omega)}, \tag{45}$$

where

$$\epsilon_n' = \epsilon_n - \hbar\omega(n + \tfrac{1}{2}).$$

Equation (45) cannot be solved analytically but the result can be obtained as the expansion

$$\epsilon_n' = -\tfrac{1}{2}(\alpha^2/k)[1 - \tfrac{1}{8}(5n^2 + 5n + 3)(\alpha^2/k)^2 + \cdots], \tag{46}$$

which is to be compared with our result (44).

When we take

$$\hat{H}_1 = \tfrac{1}{2}\alpha x^2 \tag{47}$$

as a perturbation to (41), the exact solution is

$$\epsilon_n = \hbar\omega(n + \tfrac{1}{2})[(k + \alpha)/k]^{1/2} = \hbar\omega(n + \tfrac{1}{2})[1 + \tfrac{1}{2}(\alpha/k) - \tfrac{1}{8}(\alpha/k)^2 + \tfrac{1}{16}(\alpha/k)^3 - \tfrac{5}{128}(\alpha/k)^4 + \tfrac{7}{256}(\alpha/k)^5 - \cdots], \tag{48}$$

while the perturbation formula (30) gives

$$\epsilon_n^{(5)} = \hbar\omega(n + \tfrac{1}{2}) + (2n + 1)\beta - \frac{[(2n + 1)/(\hbar\omega + 2\beta)]\beta^2 + [(2n + 1)(n^2 + n + 3)/2(\hbar\omega + 2\beta)^3]\beta^4 + \cdots}{1 + [(n^2 + n + 1)/2(\hbar\omega + 2\beta)^2]\beta^2 + \cdots}, \tag{49}$$

TABLE I. Magnitudes of exact and approximate eigenvalues of Eq. (51). ($a = 1$.)

ξ^2	0.01	0.05	0.1	0.5	1.0
exact	1.019 803 90	1.095 445	1.1832	1.732	2.24
$\epsilon^{(17)}$	1.019 803 87	1.095 418	1.1829	1.693	2.06
$\epsilon^{(5)}$	1.019 802 0	1.095 24	1.1818	1.667	2.00
$\epsilon^{(2)}$	1.020 00	1.100	1.200	2.00	3.0

where

$$\beta = \alpha\hbar\omega/4k. \tag{50}$$

We indeed see that when (49) is expanded into a power series of β or α, the result agrees with (48) to the fifth order of α.

The exact eigenvalues of the 2×2 matrix

$$\begin{bmatrix} -a & b \\ b & a \end{bmatrix} \tag{51}$$

are well known as

$$\epsilon_{\pm} = \pm (a^2 + b^2)^{1/2} = \pm a(1 + 4\xi^2)^{1/2} = \pm a(1 + 2\xi^2 - 2\xi^4 + 4\xi^6 - \cdots), \tag{52}$$

where

$$\xi = b/2a. \tag{53}$$

The perturbation formula (30) gives

$$\epsilon_n^{(5)} = \pm a \pm \frac{2a\xi^2}{1 + \xi^2} = \pm a(1 + 2\xi^2 - 2\xi^4 + 2\xi^6 - \cdots), \tag{54}$$

which deviates from (52) at ξ^6. In this simple case we can calculate the next approximation without much trouble. The result is

$$\epsilon_n^{(17)} = \epsilon_n^{(5)} \pm 2a \left(\frac{\xi^3}{1 + \xi^2}\right)^2 \left(1 + \frac{2\xi^2}{1 + \xi^2}\right)^{-1} \Bigg/ \left[1 + \left(\frac{\xi^3}{1 + \xi^2}\right)^2 \left(1 + \frac{2\xi^2}{1 + \xi^2}\right)^{-2}\right]$$

$$= \epsilon_n^{(5)} \pm \frac{2a\xi^6}{(1 + \xi^2)(1 + 5\xi^2 + 9\xi^4 + \xi^6)}$$

$$= \pm a \left[1 + \frac{2\xi^2(1 + 6\xi^2 + 10\xi^4 + \xi^6)}{(1 + \xi^2)(1 + 6\xi^2 + 9\xi^4 + \xi^6)}\right]. \tag{55}$$

In Table I numbers obtained by our perturbation formulas are compared with the exact ones and those given by the conventional second-order perturbation formula, which is

$$\epsilon_n^{(2)} = \pm a(1 + 2\xi^2) \tag{56}$$

in this case.

10. VARIATIONAL CHARACTER

Since our perturbation formulas are obtained from the diagonal elements of unitarily transformed matrices, they obviously have variational character: When they are applied to a bound state we obtain its upper limit, and so forth.[5] This property is clearly seen in Table I.

Our formula for $\epsilon_n^{(5)}$ is of the form of the Padé approximant,[6] and that for $\epsilon_n^{(17)}$ may be regarded as a generalized Padé approximant.

* Supported by NSF Grant No. GP-27444.

1 E. Schrödinger, Ann. Physik **80**, 437 (1926).
2 E. U. Condon and G. H. Shortley, *The Theory of Atomic Spectra* (Cambridge U. P., London, 1935).
3 P. Wigner, Math. Naturw. Anz. Ungar. Akad. Wiss. **53**, 475 (1935); L. Brillouin, J. Phys. **4**, 1 (1933).
4 P. Goldhammer and E. Feenberg, Phys. Rev. **101**, 1233 (1956); R. C. Young, L. C. Biedenharn, and E. Feenberg, Phys. Rev. **106**, 1151 (1957).
5 R. Courant and D. Hilbert, *Methods of Mathematical Physics* (Interscience, New York, 1953), Chap. 1.
6 G. A. Baker and J. L. Gammel, Eds., *The Padé Approximations in Theoretical Physics* (Academic, New York, 1970).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 10 OCTOBER 1971

Explicit Decomposition of a General Two-Body Hamiltonian into Irreducible Symplectic Tensors*

Shou-yung Li
Department of Physics and Astronomy, University of Rochester, Rochester, New York 14627
(Received 1 April 1971)

The seniority classification of a general two-body Hamiltonian according to the group chain $U(N) \supset Sp(N) \supset R(3)$ is considered. It is found that by using the standard contraction and symmetrization process, one can explicitly decompose such a Hamiltonian into irreducible tensors with respect to the symplectic group. An example is given where the explicit angular momentum coupled form of these tensors are also worked out.

I. INTRODUCTION

The method of classifying nuclear or atomic many-particle states according to irreducible representations of various linear groups is well known. Among the commonly used groups is, for example, the group chain $U(N) \supset Sp(N) \supset R(3)$ which specifies the particle number, seniority, and total angular momentum of many-particle states formed from a set of N single-particle states.

The corresponding decomposition of a general two-body Hamiltonian into irreducible tensors of this group chain is, however, more difficult than the classification of states, and not much work has been done in this direction. This is because a Hamiltonian is generally a Kronecker product of irreducible tensors. In order to reduce it, one needs the vector-coupling coefficients of the group considered, but these coefficients are usually not known for the general representations.

In the case of the unitary decomposition, Chang, French, and Thio[1] have obtained an explicit form by a contraction procedure. For the symplectic group, Judd[2] has considered the decomposition of a one-body Hamiltonian, and Wybourne[3] has listed the irreducible symplectic tensors contained in a general two-body Hamiltonian, but did not give the explicit form of these tensors.

In this paper we shall show that by writing a general two-body Hamiltonian in the second quantized form, one can obtain explicit expressions of the irreducible symplectic parts by the standard contraction and symmetrization process.[4]

In Sec. II, we give the general form of a two-body Hamiltonian in second-quantized form. We also write down the unitary tensors using the results of Ref. 1. In Sec. III, we consider the general method of constructing symplectic tensors. Finally in the last section, we investigate a special case which may have physical interest, which is when the single-particle space consists of a single j shell in a j–j coupling scheme.

II. THE HAMILTONIAN AND ITS UNITARY DECOMPOSITION

In the second-quantized picture, an orthonomal set of single-particle states can be expressed as

$$|i\rangle \equiv A_i^+|0\rangle,$$

where $|0\rangle$ is the vacuum state, and $i = 1, 2, \cdots$.

The fermion creation operators a_i^+ together with their Hermitian conjugates a_i satisfy the anti-commutation relations

$$\{a_i, a_j^+\} = \delta_{ij}$$

and

$$\{a_i, a_j\} = \{a_i^+ a_j^+\} = 0.$$

A general two-body Hamiltonian can now be written as

$$H_2 = \tfrac{1}{4} \sum_{ijkl} (ij|H_2|kl) a_i^+ a_j^+ a_l a_k, \tag{1}$$

in which the two-body matrix elements have been antisymmetrized so that

$$\begin{aligned}(ij|H_2|kl) &= -(ij|H_2|lk) \\ &= (ji|H_2|lk).\end{aligned} \tag{2}$$

Similarly, a one-body Hamiltonian has the form

$$H_1 = \sum_{ij} (i|H_1|j) a_i^+ a_j. \tag{3}$$

In this paper we consider a Hamiltonian

$$\begin{aligned}H &\equiv h + V \\ &\equiv \sum_{ij} (i|h|j) a_i^+ a_j - \tfrac{1}{4} \sum_{ijkl} (ij|V|kl) a_i^+ a_j^+ a_k a_l,\end{aligned} \tag{4}$$

which is assumed to be Hermitian so that

$$(i|h|j) = (j|h|i)^* \tag{5a}$$

and

$$(ij|V|kl) = (kl|V|ij)^*. \tag{5b}$$

Generally the single-particle space is infinite dimensional, but in practical calculations one has to restrict himself in a subspace with a finite number N of states. Hereafter we shall consider only such a subspace, which is then subjected to a unitary transformation of dimension N and forms an irreducible representation $\{1\}$ of the group $U(N)$,[5] and the many-particle states formed from them can be decomposed into irreducible representation of $U(N)$. However, the Hamiltonian H defined on this many-particle space is not irreducible with respect to $U(N)$. The two-body part V, for example, contains the representations $\{2^2 1^{N-4}\}$, $\{21^{N-2}\}$, and $\{0\}$. Likewise h contains $\{21^{N-2}\}$ and $\{0\}$.

The decomposition of H into irreducible unitary tensors is done in Ref. 1. Here we shall only quote

the results. For the two-body Hamiltonian V, we shall write the irreducible parts as $H_2(\nu)$, where $\nu = 0$ corresponds to $\{0\}$, $\nu = 1$ to $\{21^{N-2}\}$, and $\nu = 2$ to $\{2^2 1^{N-4}\}$. We have

$$H_2(0) = W\binom{n}{2}, \tag{6a}$$

$$H_2(1) = (N-2)^{-1}(n-1)\left[\sum_{ij}\left(\sum_k (ik|V|jk)\right) \times a_i^\dagger a_j - \frac{2W}{N}\binom{N}{2}n\right], \tag{6b}$$

and

$$H_2(2) = -\tfrac{1}{4}\sum_{ijkl}(ij|V|kl)a_i^\dagger a_j^\dagger a_l a_k - H_2(1) - H_2(0) \equiv -\tfrac{1}{4}\sum_{ijkl}(ij|\bar{V}|kl)a_i^\dagger a_j^\dagger a_l a_k. \tag{6c}$$

For the decomposition of the one-body part h, one has the similar result

$$H_1(0) = \epsilon n \tag{7a}$$

and

$$H_1(1) = \sum_{ij}(i|h|j)a_i^\dagger a_j - \epsilon n. \tag{7b}$$

Here n is the number operator

$$n \equiv \sum_{i=1}^{N} a_i^\dagger a_i \tag{8}$$

and $\binom{n}{2} = n(n-1)/2$ is the binomial coefficient. The quantities ϵ, W, and $(ij|\bar{V}|kl)$ are defined as

$$\epsilon \equiv (1/N)\sum_i (i|h|i), \tag{9a}$$

$$W \equiv \binom{N}{2}^{-1}\tfrac{1}{2}\sum_{ij}(ij|V|ij), \tag{9b}$$

and

$$(ij|\bar{V}|kl) \equiv (ij|V|kl) - (N-2)^{-1}\{\delta_{ik}\bar{V}_{jl} + \delta_{jl}\bar{V}_{ik} - \delta_{il}\bar{V}_{jk} - \delta_{jk}\bar{V}_{il}\} - (\delta_{ik}\delta_{jl} - \delta_{il}\delta_{jk})W, \tag{9c}$$

with

$$\bar{V}_{ik} \equiv \sum_j (ij|V|kj) - \delta_{ik}\frac{2W}{N}\binom{N}{2}. \tag{9d}$$

It is not difficult to verify that $H_2(2)$ has the following very important particle–hole symmetry:

$$H_2(2) = -\tfrac{1}{4}\sum_{ijkl}(ij|\bar{V}|kl)a_i^\dagger a_j^\dagger a_k a_l = \tfrac{1}{4}\sum_{ijkl}(ij|\bar{V}|kl)a_i^\dagger a_k a_j^\dagger a_l = -\tfrac{1}{4}\sum_{ijkl}(ij|\bar{V}|kl)a_k a_l a_i^\dagger a_j^\dagger, \tag{10}$$

which shows that for the irreducible unitary parts, the various creation and annihilation operators essentially commute.

We shall define the quantity

$$j \equiv \tfrac{1}{2}(N-1),$$

and rewrite the single-particle operators as

$$A^j_m \equiv a^\dagger_{j+1+m}, \tag{11a}$$

and

$$B^j_m \equiv (-)^{j+m}a_{j+1-m}, \tag{11b}$$

with $m = -j, -j+1, \cdots, j$. For the group Sp(N), N must be an even integer, therefore j is half-integral. It must be noted however that our change of notation is only for later convenience, and j does not necessarily have any physical meaning. Next we define the quantities

$$U^J_M \equiv (A^j \times B^j)^J_M \equiv \sum_{m(m')}\begin{bmatrix} j & j & J \\ m & m' & M\end{bmatrix} A^j_m B^j_{m'}, \tag{12a}$$

$$Z^I_M \equiv -(1/\sqrt{2})(A^j \times A^j)^I_M, \tag{12b}$$

and

$$\bar{Z}^I_M \equiv (1/\sqrt{2}), (B^j \times B^j)^I_M, \tag{12c}$$

where square brakets surround the ordinary Clebsch–Gordan coefficient, and x denotes ordinary vector coupling. Hereafter we will follow the notations of Ref. 6. Equations (5) and (6) can now be written as

$$H(0) = H_1(0) + H_2(0) = \epsilon n + W\binom{n}{2}, \tag{13a}$$

$$H(1) = H_1(1) + H_2(1) = \sum_{J=0,M}[h_{JM} + (N-2)^{-1}(n-1)V_{JM}][\hat{j}]^{1/2}U^J_M, \tag{13b}$$

and

$$H(2) = H_2(2) = \tfrac{1}{2}\sum_{I_1 I_2 JM}\bar{V}^J_M(I_1 I_2)(Z^{I_1} \times Z^{I_2})^J_M, \tag{13c}$$

with the following definitions

$$\hat{j} \equiv 2j+1, \tag{14}$$

$$n = \sum_{m=-j}^{j} A_m A^+_m = [\hat{j}]^{1/2}U^0, \tag{15}$$

$$h_{JM} \equiv \frac{1}{[\hat{j}]}\sum_{m(m')}[\hat{J}]^{1/2}\begin{bmatrix} j & J & j \\ m & M & m\end{bmatrix}(m|h|m'), \tag{16a}$$

$$V_{JM} \equiv \frac{1}{[\hat{j}]}\sum_{m(m')}[\hat{J}]^{1/2}\begin{bmatrix} j & J & j \\ m' & M & m\end{bmatrix} \times \left(\sum_k (mk|V|m'k)\right) \tag{16b}$$

and

$$\overline{V}_M^J(I_1I_2) = \sum_{m_1m_2(m_1'm_2')M_1(M_2)} \begin{bmatrix} j & j & I_1 \\ m_1 & m_1' & M_1 \end{bmatrix} \times \begin{bmatrix} j & j & I_2 \\ m_2 & m_2' & M_2 \end{bmatrix} (-)^{M_2} \begin{bmatrix} I_1 & I_2 & J \\ M_1 & M_2 & M \end{bmatrix} \times (m_1m_1'|\overline{V}|m_2m_2'), \tag{16c}$$

where the subscripts in the parentheses are not summed over. We note again that since j does not correspond to any real physical angular momentum, the Wigner–Eckart theorem cannot be applied to the matrix elements $(m|h|m')$ or $(m_1m_1'|V|m_2m_2')$. We shall also notice that

$$h_{00} \equiv \frac{1}{[\hat{j}]} \sum_m (m|h|m) = \epsilon \tag{17a}$$

and

$$V_{00} \equiv \frac{1}{[\hat{j}]} \sum_{mk} (mk|V|mk) = \frac{2W}{N}\binom{N}{2}. \tag{17b}$$

III. THE $U(N) \supset \mathrm{Sp}(N)$ DECOMPOSITION OF THE HAMILTONIAN

In this section we shall further reduce Eq. (13) into irreducible symplectic tensors. First, the branching rule of $U(N) \to \mathrm{Sp}(N)$ gives

$$\{0\} \to \langle 0\rangle, \tag{18a}$$

$$\{21^{N-2}\} \to \langle 1^2\rangle + \langle 2\rangle, \tag{18b}$$

and

$$\{2^21^{N-4}\} \to \langle 0\rangle + 2\langle 1^2\rangle + \langle 2^2\rangle + \langle 21^2\rangle + \langle 1^4\rangle. \tag{18c}$$

This is true for $N \geq 8$. For $N < 8$ some of the symplectic representations do not appear. Also $\langle 21^2\rangle$ vanishes for a Hermitian operator.[3]

The reason of rewriting the Hamiltonian in the form of Eq. (13) is that in this way the Hamiltonian is already partially decomposed into irreducible symplectic tensors. It is well known[2] that the unit tensors U^J with odd J are the generators of the symplectic group and belong to the representation $\langle 2\rangle$, while those with even $J \neq 0$ belong to $\langle 1^2\rangle$. Therefore we immediately have

$$H(0, \langle 0\rangle) = H(0). \tag{19a}$$

$$H(1, \langle 2\rangle) = \sum_{J\,\mathrm{odd},\,M} [h_{JM} + (N-2)^{-1}(n-1)V_{JM}] \times [\hat{j}]^{1/2} U_M^J \tag{19b}$$

and

$$H(1, \langle 1^2\rangle) = \sum_{J\,\mathrm{even}\neq 0,\,M} [h_{JM} + (N-2)^{-1}(n-1)V_{JM}] \times [\hat{j}]^{1/2} U_M^J. \tag{19c}$$

We are now left only with the tensor $H_2(2)$ given in Eq. (10), which according to the branching rule (18c) contains symplectic tensors with rank 0, 2, and 4. It is however well known[4] that symplectic transformations are invariant under contractions with an antisymmetic metric. Consequently we shall define such a metric tensor

$$\eta_{mn} \equiv [\hat{j}]^{1/2} \begin{bmatrix} j & j & 0 \\ m & n & 0 \end{bmatrix} = (-)^{j-m}\delta_{m,-n} = -\eta_{nm} \tag{20}$$

and proceed to separate the symplectic tensors of different ranks.

We write $H_2(2)$ as

$$H_2(2) = -\tfrac{1}{4} \sum_{m_1m_2m_3m_4} (m_1m_2|\overline{V}|-m_3-m_4) \times (-)^{2j+m_3+m_4} A_{m_1}A_{m_2}B_{m_3}B_{m_4} \equiv -\tfrac{1}{4} \sum_{1234} (12|U|34)A_1A_2B_3B_4, \tag{21a}$$

such that

$$(12|U|34) = -(12|U|43) = (21|U|43) \tag{21b}$$

and

$$(12|U|34) = (34|U|12), \tag{22}$$

where the last equality comes from the assumption that $\overline{V}$ is invariant under time reversal. Note that the particle–hole contractions of U vanish as a consequence of being irreducible with respect to the unitary group. For example, we have

$$\sum_{13} \eta_{13}(12|U|34) = \sum_{m_1} (m_1m_2|\overline{V}|m_1 - m_4)(-)^{j+m_4} = 0. \tag{23}$$

Thus we can decompose U as follows

$$(12|U|34) = \sum_{i=0}^{2} (12|\overline{U}^{(i)}|34), \tag{24}$$

with

$$(12|\overline{U}^{(0)}|34) \equiv [\eta_{12}\eta_{34} + (N-1)^{-1}(\eta_{14}\eta_{23} - \eta_{13}\eta_{24})]\overline{E}^{(0)}, \tag{25a}$$

$$\overline{E}^{(0)} \equiv \frac{N-1}{N(N+1)(N-2)} \sum_{1234} \eta_{12}\eta_{34}(12|U|34), \tag{25b}$$

$$(12|\overline{U}^{(1)}|34) = \eta_{12}\overline{E}_{34}^{(1)} + \eta_{34}\overline{E}_{12}^{(1)} + 2(N-2)^{-1}[\eta_{14}\overline{E}_{23}^{(1)} + \eta_{23}\overline{E}_{14}^{(1)} - \eta_{13}\overline{E}_{24}^{(1)} - \eta_{24}\overline{E}_{13}^{(1)}], \tag{26a}$$

$$\overline{E}_{12}^{(1)} \equiv \frac{N-2}{(N-4)(N+2)} [\sum_{34} \eta_{34}(12|U|34) - \frac{\eta_{12}}{N} \sum_{1234} \eta_{12}\eta_{34}(12|U|34)], \tag{26b}$$

and

$$(12|\bar{U}^{(2)}|34) \equiv (12|U|34) - (12|\bar{U}^{(1)}|34) - (12|\bar{U}^{(0)}|34). \quad (27)$$

The above formulas are derived in Appendix A.

In the branching rule (18c), there is only one representation,[7] $\langle 1^2\rangle$, which has symplectic rank 2; hence we can identify it as $\bar{U}^{(1)}$. Also the zero-rank tensor $\bar{U}^{(0)}$ must belong to $\langle 0\rangle$; then we have

$$H(2,\langle 0\rangle) = -\tfrac{1}{4}\sum(12|\bar{U}^{(0)}|34)A_1A_2B_3B_4 \quad (28)$$

and

$$H(2,\langle 1^2\rangle) = -\tfrac{1}{4}\sum(12|\bar{U}^{(1)}|34)A_1A_2B_3B_4. \quad (29)$$

Note that the above tensors still possess the particle–hole symmetry of Eq. (10).

Finally, the different irreducible symplectic tensors contained in $\bar{U}^{(2)}$ can be separated by using the Young symmetrizers.[4]

$$(12|\bar{U}^{(2)}_{\langle 1^4\rangle}|34) = \tfrac{1}{3}[(12|\bar{U}^{(2)}|34) + (14|\bar{U}^{(2)}|23) - (13|\bar{U}^{(2)}|24)], \quad (30)$$

$$(12|\bar{U}^{(2)}_{\langle 2^2\rangle}|34) = \tfrac{1}{3}[2(12|\bar{U}^{(2)}|34) - (14|\bar{U}^{(2)}|23) + (13|\bar{U}^{(2)}|24)], \quad (31)$$

$$(12|\bar{U}^{(2)}_{\langle 21^2\rangle}|34) = 0, \quad (32)$$

where the last equation is a consequence of hermiticity and time-reversal symmetry. We now have

$$H(2,\langle 1^4\rangle) = -\tfrac{1}{4}\sum\tfrac{1}{3}[(12|\bar{U}^{(2)}|34) + 2(14|\bar{U}^{(2)}|23)] \times A_1A_2B_3B_4 \quad (33)$$

and

$$H(2,\langle 2^2\rangle) = -\tfrac{1}{4}\sum\tfrac{2}{3}[(12|\bar{U}^{(2)}|34) - (14|\bar{U}^{(2)}|23)] \times A_1A_2B_3B_4. \quad (34)$$

Equations (28), (29), (33), and (34) can easily be written in the coupled form. In particular, if we define

$$\bar{U}^{(i)}(I_1I_2JM) \equiv \sum_{m_1m_2(m_1'm_2'),M_1(M_2)} \begin{bmatrix} j & j & I_1 \\ m_1 & m_1' & M_1 \end{bmatrix} \times \begin{bmatrix} j & j & I_2 \\ m_2 & m_2' & M_2 \end{bmatrix} \begin{bmatrix} I_1 & I_2 & J \\ M_1 & M_2 & M \end{bmatrix} \times (m_1m_1'|\bar{U}^{(i)}|m_2m_2'), \quad (35)$$

then Eq. (33) can be written as

$$H(2,\langle 1^4\rangle) = (-\tfrac{1}{4})(\tfrac{1}{3})\sum(12|\bar{U}^{(2)}|34) \times (\tfrac{1}{2}A_1A_2B_3B_4 + \tfrac{1}{2}B_1B_2A_3A_4 + 2A_1B_2A_3B_4) = -\tfrac{1}{12}\sum_{I_1\neq 0,JM,I_2\neq 0}\bar{U}^{(2)}(I_1I_2JM) \times [-(Z^{I_1}\times\bar{Z}^{I_2})^J_M - (\bar{Z}^{I_1}\times Z^{I_2})^J_M + 2(U^{I_1}\times U^{I_2})^J_M], \quad (36)$$

where the first equality is a consequence of particle–hole symmetry. From Eq. (B7c) of Appendix B we can write $H(2,\langle 1^4\rangle)$ as

$$H(2,\langle 1^4\rangle) = -\sqrt{\tfrac{3}{2}}\sum_{I_1,I_2\neq 0,JM}\bar{U}^{(2)}(I_1I_2JM) \times [T^{I_1}1 \times T^{I_2}1]^{J\,2}_{M0}, \quad (37)$$

which verifies the fact that $H(2,\langle 1^4\rangle)$ is a pure quasispin-two tensor.[3] Similarly, one can show that $H(2,\langle 2^2\rangle)$ is a quasispin-zero tensor.

IV. THE SCALAR HAMILTONIAN IN A SINGLE j SHELL

We now consider the special case when the single-particle space consists of a single j shell in a j–j coupling shell model with a Hamiltonian

$$H = h[\hat{j}]^{1/2}U^0 + \sum_I [\hat{I}]^{1/2}V^I(Z^I\times\bar{Z}^I)^0, \quad (38)$$

which is a scalar with respect to $R(3)$. The operators U^J and Z^I were defined in Eq. (12), with j now corresponding to real angular momentum. We get from Eq. (19)

$$H(0) = hm + W\binom{n}{2} \quad (39)$$

and

$$H(1) = 0, \quad (40)$$

with

$$W = \binom{N}{2}^{-1}\sum_I[\hat{I}]V^I. \quad (41)$$

The remaining part is

$$H_2(2) = \sum_I[\hat{I}]^{1/2}V^I(Z^I\times\bar{Z}^I)^0 - (W/2)n(n-1) = \sum_I[\hat{I}]^{1/2}V^I(Z^I\times\bar{Z}^I)^0 - \tfrac{1}{4}\sum_{ijkl}(2W)\delta_{jk}\delta_{il} \times a_i^+a_j^+a_ka_l \equiv \tfrac{1}{2}\sum_I\bar{V}^I(Z^I\times\bar{Z}^I)^0, \quad (42a)$$

with

$$\bar{V}^I \equiv 2[\hat{I}]^{1/2}(V^I - W). \quad (42b)$$

From Eqs. (25) and (26), we can write

$$\bar{E}^{(0)} = \frac{N-1}{(N+1)(N-2)}\bar{V}^{(0)}, \quad (43)$$

$$H(2,\langle 0\rangle) = \tfrac{1}{2}\sum[\bar{U}^{(0)}]^I(Z^I\times\bar{Z}^I)^0, \quad (44a)$$

$$[\bar{U}^{(0)}]^I = \bar{E}^{(0)}\{N\delta_{I0} - 2(N-1)^{-1}[\hat{I}]^{1/2}\} = \frac{2\bar{V}^{(0)}}{(N+1)(N-2)}\left\{\binom{N}{2}\delta_{I0} - [\hat{I}]^{1/2}\right\}, \quad (44b)$$

and

$$H(2,\langle 1^2\rangle) = 0. \quad (45)$$

Finally, we give the explicit forms of the second-rank symplectic tensors:

$$H(2,\langle 1^4\rangle) = \tfrac{1}{6}\sum_{I\neq 0}[\overline{U}^{(2)}]^I[(Z^I\times\overline{Z}^I)^0 + (U^I\times U^I)^0]$$
$$= \tfrac{1}{6}\sum_{IK}[\overline{U}^{(2)}]^K\left(\delta_{IK} - [\hat{I}\hat{K}]^{1/2}\begin{Bmatrix} j & j & I\\ j & j & K\end{Bmatrix}\right)(Z^I\times\overline{Z}^I)^0, \quad (46)$$

$$H(2,\langle 2^2\rangle) = \tfrac{1}{6}\sum_{IK}[\overline{U}^{(2)}]^K\left(\delta_{IK} + [\hat{I}\hat{K}]^{1/2}\begin{Bmatrix} j & j & I\\ j & j & K\end{Bmatrix}\right)$$
$$\times (Z^I\times\overline{Z}^I)^0, \quad (47)$$

where the curly brakets indicate a $6-j$ symbol, and

$$[\overline{U}^{(2)}]^I \equiv \overline{V}^I - [\overline{U}^{(0)}]^I$$
$$= (1-\delta_{I0})\left(\overline{V}^I - \frac{2[\hat{I}]^{1/2}\overline{V}^0}{(N+1)(N-2)}\right). \quad (48)$$

V. CONCLUSIONS

We have shown that, starting from the irreducible unitary tensors of a general two-body Hamiltonian, the subsequent irreducible symplectic tensors can be obtained by first separating those of different ranks with the standard contraction procedure with an antisymmetric metric, and then constructing tensors with different symmetry by using the Young operators. Finally we considered a special case when the single-particle space consists of a single j shell in the j–j coupling scheme. In that case one can easily obtain the explicit angular momentum coupled form of these irreducible tensors.

ACKNOWLEDGMENTS

The author is indebted to Professor J. B. French and Dr. F. S. Chang for many helpful discussions.

APPENDIX A: THE DECOMPOSITION OF SYMPLECTIC TENSORS

We consider the decomposition of U defined in Eq. (20). In general, such a tensor can be written as[4]

$$(12|U|34) \equiv (12|U^{(2)}|34)$$
$$+ \eta_{12}U^{(1)}_{34} + \eta_{13}V^{(1)}_{24} - \eta_{23}V^{(1)}_{14}$$
$$- \eta_{14}V^{(1)}_{23} + \eta_{24}V^{(1)}_{13} + \eta_{34}U^{(1)}_{12}$$
$$+ \eta_{12}\eta_{34}U^{(0)} + (\eta_{13}\eta_{24} - \eta_{14}\eta_{23})V^{(0)}, \quad (A1)$$

where we have used the symmetry properties Eq. (21). Here $U^{(i)}$ is an ith rank tensor. Contracting, we get

$$\sum_{1234}\eta_{12}\eta_{34}(12|U|34) = N^2U^{(0)} + 2NV^{(0)} \quad (A2a)$$

and

$$\sum_{1234}\eta_{13}\eta_{24}(12|U|34) = NU^{(0)} + (N^2-N)V^{(0)} = 0. \quad (A2b)$$

Solving these two equations gives

$$V^{(0)} = -(N-1)^{-1}U^{(0)} \quad (A3)$$

and

$$U^{(0)} = \frac{N-1}{N(N+1)(N-2)}\sum_{1234}\eta_{12}\eta_{34}(12|U|34). \quad (A4)$$

Next, we have

$$\sum_{13}\eta_{13}(12|U|34) = 2U^{(1)}_{24} + (N-2)V^{(1)}_{24} = 0, \quad (A5)$$

which gives

$$V^{(1)}_{24} = -2(N-2)^{-1}U^{(1)}_{24}. \quad (A6)$$

Finally, from the equation

$$\sum_{12}\eta_{12}(12|U|34)$$
$$= NU^{(1)}_{34} + 4V^{(1)}_{34} + \eta_{34}[N - 2(N-1)^{-1}]U^{(0)}, \quad (A7)$$

we get

$$U^{(1)} = \frac{N-2}{(N-4)(N+2)}E^{(1)} \quad (A8)$$

and

$$V^{(1)} = \frac{-2}{(N-4)(N+2)}E^{(1)}, \quad (A9)$$

with

$$E^{(1)}_{12} \equiv \sum_{34}\eta_{34}(12|U|34) - \frac{\eta_{12}}{N}\sum_{1234}\eta_{12}\eta_{34}(12|U|34). \quad (A10)$$

APPENDIX B: THE QUASISPIN TENSORS

It is well known[8] that if we define

$$S_+ \equiv \sqrt{\Omega}Z^0 = -(\Omega/2)^{1/2}(A^i\times A^i)^0, \quad (B1a)$$

$$S_- \equiv S_+^+ = \sqrt{\Omega}\overline{Z}^0, \quad (B1b)$$

and

$$S_0 \equiv \tfrac{1}{2}(N-\Omega), \quad (B1c)$$

where $\Omega = \frac{1}{2}[\hat{j}]$, then these three operators form the quasispin algebra,

$$[S_+, S_-] = 2S_0 \quad (B2a)$$

and

$$[S_0, S_\pm] = \pm S_\pm. \quad (B2b)$$

From Eq. (B1), we easily derive the relations

$$[S_+, A^j_m] = 0, \quad (B3a)$$

$$[S_-, A^j_m] = B^j_m, \quad (B3b)$$

$$[S_0, A^j_m] = \tfrac{1}{2}A^j_m, \quad (B3c)$$

which together with similar ones for B^j_m demonstrate that the operators A^j_m and B^j_m form an irreducible quasispin tensor with rank $\frac{1}{2}$. We shall write

$$T^{j\ 1/2}_{m\ 1/2} \equiv A^j_m \tag{B4a}$$

and

$$T^{j\ 1/2}_{m\ 1/2} \equiv B^j_m . \tag{B4b}$$

From them we now construct a quasispin tensor of rank 1

$$T^{J\ 1}_{M\ \mu} \equiv -(1/\sqrt{2})[T^{j1/2} \times T^{j1/2}]^{j\ 1}_{M\ \mu}$$

$$\equiv (1/\sqrt{2}) \sum_{m s (m s')} \begin{bmatrix} j & j & J \\ m & m' & M \end{bmatrix} \begin{bmatrix} 1/2 & 1/2 & 1 \\ s & s' & \mu \end{bmatrix} \times T^{j\ 1/2}_{m\ s} T^{j\ 1/2}_{m\ s'}, \tag{B5}$$

which gives the following identifications

$$T^{I\ 1}_{M\ 1} = -(1/\sqrt{2})(A^j \times A^j)^I_M = Z^I, \tag{B6a}$$

$$T^{I\ 1}_{M\ 0} = -(1/\sqrt{2})\{(1/\sqrt{2})(A^j \times B^j)^i + (1/\sqrt{2})(B^j \times A^j)^I\} = -U^I + (\Omega/2)^{1/2}\delta_{I0}, \quad (I \text{ even}), \tag{B6b}$$

and

$$T^{I\ 1}_{M\ -1} = -(1/\sqrt{2})(B^j \times B^j)^I_M = -\bar{Z}^I. \tag{B6c}$$

Finally one can form the quasispin tensors $(I_1, I_2 \neq 0)$,

$$-[T^{I_1 1} \times T^{I_2 1}]^{J\ 0}_{M\ 0} = (1/\sqrt{3})(Z^{I_1} \times \bar{Z}^{I_2})^J_M + (1/\sqrt{3})(\bar{Z}^{I_1} \times Z^{I_2})^J_M + (1/\sqrt{3})(U^{I_1} \times U^{I_2})^J_M \tag{B7a}$$

$$-[T^{I_1 1} \times T^{I_2 1}]^{J\ 1}_{M\ 0} = (1/\sqrt{2})\{(Z^{I_1} \times \bar{Z}^{I_2})^J_M - (\bar{Z}^{I_1} \times Z^{I_2})^J_M\} \tag{B7b}$$

and

$$-[T^{I_1 1} \times T^{I_2 1}]^{J\ 2}_{M\ 0} = (1/\sqrt{6})(Z^{I_1} \times \bar{Z}^{I_2})^J_M + (1/\sqrt{6})(\bar{Z}^{I_1} \times Z^{I_2})^J_M - \sqrt{\tfrac{2}{3}}(U^{I_1} \times U^{I_2})^J_M. \tag{B7c}$$

---

* Supported in part by the U.S. Atomic Energy Commission.

1 F. S. Chang, J. B. French, and T. H. Thio, Ann. Phys. (N.Y.) (to be published).

2 B. R. Judd, in *Group Theory and Its Applications*, edited by E. M. Loebl (Academic, New York, 1968).

3 B. G. Wybourne, J. de Phys. **30**, 39 (1969).

4 M. Hamermesh, *Group Theory* (Addison-Wesley, Reading, Mass., 1962).

5 The irreducible representation of $U(N)$ is denoted as $\{\lambda_1 \lambda_i \cdots \lambda_N\}$ which corresponds to a Young diagram of N rows, the ith row being of length λ_i.

6 J. B. French, in *Many-Body Description of Nuclear Structure and Reactions*, edited by C. Bloch (Academic, New York, 1966).

7 Under time reversal the two $\langle 1^2 \rangle$ representations are mixed.

8 R. D. Lawson and M. H. MacFarlane, Nucl. Phys. **66**, 80 (1965).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 10 OCTOBER 1971

The Canonization of Nice Variables

David P. Stern

Laboratory for Space Physics, Goddard Space Flight Center, Greenbelt, Maryland 20771

(Received 8 September 1970)

In a perturbed periodic classical motion, the angle variable may be eliminated either by Kruskal's transformation to "nice variables" or, if the system is canonical, by the Poincaré–Von Zeipel method. For systems that possess a Hamiltonian, the present work (i) shows that Kruskal's transformation may be made canonical order by order, (ii) derives a practical formula for achieving this result, and (iii) shows that the two methods are equivalent and may be matched order by order.

I. INTRODUCTION

A basic perturbation problem in celestial mechanics and in guiding center motion involves a set of n first-order differential equations which can be represented vectorially as

$$d\mathbf{x}/dt = \mathbf{F}(\mathbf{x}) \tag{1}$$

and which has the following properties:

(i) $\mathbf{F}(\mathbf{x})$ depends on a small parameter $\epsilon \ll 1$ and may be expanded in it:

$$\mathbf{F}(\mathbf{x}) = \mathbf{F}^{(0)}(\mathbf{x}) + \epsilon \mathbf{F}^{(1)}(\mathbf{x}) + \cdots . \tag{2}$$

(ii) In the limit $\epsilon \to 0$ (unperturbed case) the system may be solved and its solution is then periodic in time.

The problem is then to find an approximate solution valid for small but finite ϵ and useful for time intervals of the order of ϵ^{-1} periods.

Kruskal[1] devised a method for achieving this, which in many ways resembles the method of Bogoliubov and Zubarev and of Krylov and Bogoliubov.[2] The calculation in this case proceeds in two steps. First, a transformation to "intermediate variables" $\mathbf{y}(\mathbf{x})$ is performed, such that in the limit $\epsilon \to 0$, y_n is an angle variable linear in time while the remaining $n-1$ components of $\mathbf{y}$ (which we shall collectively denote by $\bar{\mathbf{y}}$) are constant. The equations according to which $\mathbf{y}$ evolves, derived from Eq. (1), then have the form

$$\frac{d\mathbf{y}}{dt} = \sum_{k=0} \epsilon^k \mathbf{g}^{(k)}(\mathbf{y}). \tag{3}$$

which together with similar ones for B^j_m demonstrate that the operators A^j_m and B^j_m form an irreducible quasispin tensor with rank $\frac{1}{2}$. We shall write

$$T^{j\ 1/2}_{m\ 1/2} \equiv A^j_m \tag{B4a}$$

and

$$T^{j\ 1/2}_{m\ 1/2} \equiv B^j_m . \tag{B4b}$$

From them we now construct a quasispin tensor of rank 1

$$T^{J\ 1}_{M\ \mu} \equiv -(1/\sqrt{2})[T^{j1/2} \times T^{j1/2}]^{j\ 1}_{M\ \mu}$$

$$\equiv (1/\sqrt{2}) \sum_{m s (m s')} \begin{bmatrix} j & j & J \\ m & m' & M \end{bmatrix} \begin{bmatrix} 1/2 & 1/2 & 1 \\ s & s' & \mu \end{bmatrix} \times T^{j\ 1/2}_{m\ s} T^{j\ 1/2}_{m\ s'}, \tag{B5}$$

which gives the following identifications

$$T^{I\ 1}_{M\ 1} = -(1/\sqrt{2})(A^j \times A^j)^I_M = Z^I, \tag{B6a}$$

$$T^{I\ 1}_{M\ 0} = -(1/\sqrt{2})\{(1/\sqrt{2})(A^j \times B^j)^i + (1/\sqrt{2})(B^j \times A^j)^I\} = -U^I + (\Omega/2)^{1/2}\delta_{I0}, \quad (I \text{ even}), \tag{B6b}$$

and

$$T^{I\ 1}_{M\ -1} = -(1/\sqrt{2})(B^j \times B^j)^I_M = -\bar{Z}^I. \tag{B6c}$$

Finally one can form the quasispin tensors $(I_1, I_2 \neq 0)$,

$$-[T^{I_1 1} \times T^{I_2 1}]^{J\ 0}_{M\ 0} = (1/\sqrt{3})(Z^{I_1} \times \bar{Z}^{I_2})^J_M + (1/\sqrt{3})(\bar{Z}^{I_1} \times Z^{I_2})^J_M + (1/\sqrt{3})(U^{I_1} \times U^{I_2})^J_M \tag{B7a}$$

$$-[T^{I_1 1} \times T^{I_2 1}]^{J\ 1}_{M\ 0} = (1/\sqrt{2})\{(Z^{I_1} \times \bar{Z}^{I_2})^J_M - (\bar{Z}^{I_1} \times Z^{I_2})^J_M\} \tag{B7b}$$

and

$$-[T^{I_1 1} \times T^{I_2 1}]^{J\ 2}_{M\ 0} = (1/\sqrt{6})(Z^{I_1} \times \bar{Z}^{I_2})^J_M + (1/\sqrt{6})(\bar{Z}^{I_1} \times Z^{I_2})^J_M - \sqrt{\tfrac{2}{3}}(U^{I_1} \times U^{I_2})^J_M. \tag{B7c}$$

---

* Supported in part by the U.S. Atomic Energy Commission.

1 F. S. Chang, J. B. French, and T. H. Thio, Ann. Phys. (N.Y.) (to be published).

2 B. R. Judd, in *Group Theory and Its Applications*, edited by E. M. Loebl (Academic, New York, 1968).

3 B. G. Wybourne, J. de Phys. **30**, 39 (1969).

4 M. Hamermesh, *Group Theory* (Addison-Wesley, Reading, Mass., 1962).

5 The irreducible representation of $U(N)$ is denoted as $\{\lambda_1 \lambda_i \cdots \lambda_N\}$ which corresponds to a Young diagram of N rows, the ith row being of length λ_i.

6 J. B. French, in *Many-Body Description of Nuclear Structure and Reactions*, edited by C. Bloch (Academic, New York, 1966).

7 Under time reversal the two $\langle 1^2 \rangle$ representations are mixed.

8 R. D. Lawson and M. H. MacFarlane, Nucl. Phys. **66**, 80 (1965).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 10 OCTOBER 1971

The Canonization of Nice Variables

David P. Stern

Laboratory for Space Physics, Goddard Space Flight Center, Greenbelt, Maryland 20771

(Received 8 September 1970)

In a perturbed periodic classical motion, the angle variable may be eliminated either by Kruskal's transformation to "nice variables" or, if the system is canonical, by the Poincaré–Von Zeipel method. For systems that possess a Hamiltonian, the present work (i) shows that Kruskal's transformation may be made canonical order by order, (ii) derives a practical formula for achieving this result, and (iii) shows that the two methods are equivalent and may be matched order by order.

I. INTRODUCTION

A basic perturbation problem in celestial mechanics and in guiding center motion involves a set of n first-order differential equations which can be represented vectorially as

$$d\mathbf{x}/dt = \mathbf{F}(\mathbf{x}) \tag{1}$$

and which has the following properties:

(i) $\mathbf{F}(\mathbf{x})$ depends on a small parameter $\epsilon \ll 1$ and may be expanded in it:

$$\mathbf{F}(\mathbf{x}) = \mathbf{F}^{(0)}(\mathbf{x}) + \epsilon \mathbf{F}^{(1)}(\mathbf{x}) + \cdots . \tag{2}$$

(ii) In the limit $\epsilon \to 0$ (unperturbed case) the system may be solved and its solution is then periodic in time.

The problem is then to find an approximate solution valid for small but finite ϵ and useful for time intervals of the order of ϵ^{-1} periods.

Kruskal[1] devised a method for achieving this, which in many ways resembles the method of Bogoliubov and Zubarev and of Krylov and Bogoliubov.[2] The calculation in this case proceeds in two steps. First, a transformation to "intermediate variables" $\mathbf{y}(\mathbf{x})$ is performed, such that in the limit $\epsilon \to 0$, y_n is an angle variable linear in time while the remaining $n-1$ components of $\mathbf{y}$ (which we shall collectively denote by $\bar{\mathbf{y}}$) are constant. The equations according to which $\mathbf{y}$ evolves, derived from Eq. (1), then have the form

$$\frac{d\mathbf{y}}{dt} = \sum_{k=0} \epsilon^k \mathbf{g}^{(k)}(\mathbf{y}). \tag{3}$$

In these equations, y_n appears in the $\mathbf{g}^{(k)}$ only as the angle argument of periodic functions and $\mathbf{g}^{(0)}$ has only one nonzero component, the last one[3]

$$\mathbf{g}^{(0)} = (0, 0, \ldots, 0, 1). \tag{4}$$

The second step involves a near-identity transformation to "nice variables" z

$$\mathbf{z} = \mathbf{y} + \sum_{k=1} \epsilon^k \boldsymbol{\zeta}^{(k)}(\mathbf{y}), \tag{5}$$

such that in the new equations of evolution, the transformed angle variable z_n no longer appears on the right-hand side:

$$\frac{d\mathbf{z}}{dt} = \sum_{k=0} \epsilon^k \mathbf{h}^{(k)}(\tilde{\mathbf{z}}), \tag{6}$$

where as before

$$\mathbf{z} = (\tilde{\mathbf{z}}, z_n), \quad \text{and} \quad \mathbf{h}^{(0)} = \mathbf{g}^{(0)}. \tag{7}$$

A general recursion scheme for deriving z order by order has been described in an article by Stern[4] (henceforth referred to as I). We shall follow the notation introduced there, which differs slightly from Kruskal's (a similar scheme for the related Krylov–Bogoliubov expansion has been given by Musen[5]). After Eqs. (6) have been derived in this fashion, their first $n-1$ components constitute an autonomous system for deriving the components of $\tilde{\mathbf{z}}$, which can then be independently solved. The problem is thus reduced to one with $n-1$ variables.

If the system furthermore possesses a Hamiltonian, an additional variable may be eliminated by deriving a constant J of the motion. It is defined as

$$J = \oint \mathbf{p} \cdot d\mathbf{q} = \int_0^1 \mathbf{p} \cdot \frac{\partial \mathbf{q}}{\partial z_n} dz_n, \tag{8}$$

with the integration performed using an arbitrary canonical set $(\mathbf{p}, \mathbf{q})$, over a group of points ("ring"), all of which evolve according to (6) and possess the same $\tilde{\mathbf{z}}$ but with values of z_n that cover the full range for that variable (this property, if established, is maintained throughout the ring's evolution in time).

If z itself forms a canonical set, with z_1 the momentum conjugate to z_n, then one may use this set in (8), leading to

$$J = \int_0^1 z_1 dz_n = z_1, \tag{9}$$

which is a great simplification.

If this is not the case, one is forced to retrace the transformations $\mathbf{x} \to \mathbf{y} \to \mathbf{z}$ until some canonical set, with which (8) can be evaluated, is reached.

Kruskal did not derive nice canonical variables, but he showed them to be possible. Specifically, he proved that in any "nice" set, the following Poisson bracket relations are always satisfied:

$$\frac{\partial}{\partial z_n}[z_i, z_j] = 0, \tag{10}$$

$$[z_i, J] = \delta_{in}. \tag{11}$$

Following Nordheim and Fues,[6] Kruskal then shows that J and z_n may be augmented by $n-2$ functions of the z_i to form a complete canonical set.

In what follows, we shall assume that the intermediate variables y_i form a canonical set, in which case a near-identity transformation like (5) can, in principle, lead to nice canonical variables. This assumption is not unreasonable since, if a Hamiltonian for the system is known, such y_i can in general be derived by solving the unperturbed motion via the Hamilton–Jacobi method.

We then

(i) Show that the freedom allowed by Kruskal's method in the derivation of each order of Eq. (5) is sufficient to assure the canonical character of the "nice variables" z_i.
(ii) Derive a method for obtaining such z_i.
(iii) Show that the result is equivalent to what is obtained by conventional perturbation methods based on the Hamilton–Jacobi equation.

II. THE POSSIBILITY OF STEP-BY-STEP DERIVATION

If $\mathbf{y}$ is canonical

$$\mathbf{y} = (\mathbf{p}, \mathbf{q}),$$

then from (10)

$$\frac{\partial}{\partial z_n}\{[z_i, z_j] - [y_i, y_j]\} = 0. \tag{12}$$

If (5) is substituted here, the zero-order part cancels identically and the expression remaining inside the curly brackets separates into orders of ϵ and gives

$$\frac{\partial}{\partial z_n} \sum_{m=1} \epsilon^m \{[\zeta_i^{(m)}, y_j] - [\zeta_j^{(m)}, y_i] + \sum_{s=1}^{m-1} [\zeta_i^{(s)}, \zeta_j^{(m-s)}]\} = 0. \tag{13}$$

This condition is satisfied for *any* nice set z. If z is not merely nice but also canonical, then the expressions in the curly brackets of both (12) and (13) are not only independent of z_n but actually vanish.

Let us now assume that the expansion (5) has already been derived and brought to canonical form, up to and including order $k-1$. Then the first $k-1$ orders of (13) do, in fact, vanish, leaving (after division by ϵ^k)

$$\frac{\partial}{\partial z_n}\{[\zeta_i^{(k)}, y_j] - [\zeta_j^{(k)}, y_i] + \sum_{s=1}^{k-1}[\zeta_i^{(s)}, \zeta_j^{(k-s)}]\} + O(\epsilon) = 0. \quad (14)$$

As the next step, one may derive $\zeta^{(k)}$ and thus extend the calculation of z one more order. The equation satisfied by $\zeta^{(k)}$ is [I, Eq. (15)]

$$\partial\zeta^{(k)}/\partial y_n - \mathbf{h}^{(k)}(\bar{\mathbf{y}}) = \lambda^{(k)}, \quad (15)$$

where $\lambda^{(k)}$ depends only on lower orders of $\zeta^{(m)}$ and $\mathbf{h}^{(m)}$, assumed to be known at this stage. If $\langle\lambda^{(k)}\rangle$ denotes the result of averaging over the angle variable y_n

$$\langle\lambda^{(k)}\rangle = \int_0^1 \lambda^{(k)}(\mathbf{y})dy_n, \quad (16)$$

one gets [see I, Eq. (19)]

$$\zeta^{(k)} = \int_0^{y_n}(\lambda^{(k)} - \langle\lambda^{(k)}\rangle)dy'_n + \mu^{(k)}(\bar{\mathbf{y}}) = \hat{\zeta}^{(k)} + \mu^{(k)}(\bar{\mathbf{y}}). \quad (17)$$

Here $\mu^{(k)}$ is an arbitrary additive vector independent of y_n, allowed by the fact that the derivative of $\zeta^{(k)}$ which enters (15) is unaffected by such an addition. The question now arises whether $\mu^{(k)}$ may be selected so as to make the expansion (5) canonical to order k.

If this occurs, then the $O(1)$ part of (14) must vanish. It helps here to introduce the concept of the conjugate vectors $\bar{\mathbf{y}}$ [I, Eq. (4)]

$$\bar{\mathbf{y}} = (\mathbf{q}, -\mathbf{p}). \quad (18)$$

Then

$$[\zeta_i^{(m)}, y_j] - [\zeta_j^{(m)}, y_i] = \partial\zeta_i^{(m)}/\partial\bar{y}_j - \partial\zeta_j^{(m)}/\partial\bar{y}_i. \quad (19)$$

This has the form of a component of a curl dyadic in $\bar{\mathbf{y}}$ space, which one may call the conjugate curl, with components denoted by

$$(\bar{\nabla}\times\zeta^{(m)})_{ij}. \quad (20)$$

Now it has been shown in an article by Stern[7] (henceforth referred to as II) that the general condition for an expansion (5) to be canonical is

$$\zeta^{(m)} = \mathbf{f}^{(m)}(\zeta) + \bar{\nabla}\chi^{(m)}, \quad m = 1, 2, \ldots, \quad (21)$$

where $\bar{\nabla}$ is a gradient operator in $\bar{\mathbf{y}}$ space, $\chi^{(m)}$ is an arbitrary scalar, and $\mathbf{f}^{(m)}(\zeta)$ are vectors of a certain form, depending on orders of ζ *lower* than the mth and on their derivatives. Various choices of $\mathbf{f}^{(m)}$ are derived in II, all of which satisfy the identity (for canonical $\zeta^{(s)}$)

$$(\bar{\nabla}\times\mathbf{f}^{(m)})_{ij} = -\sum_{s=1}^{m-1}[\zeta_i^{(s)}, \zeta^{(m-s)}]. \quad (22)$$

Because the $\zeta^{(s)}$ in this case are known to be canonical up to and including order $k-1$, Eq. (22) will hold for the lowest $k-1$ values of m; however, it will also hold for $m = k$, for even then the orders of ζ appearing in the equation are all lower than the kth. Substitution in (14) then yields

$$(\partial/\partial z_n)\{\bar{\nabla}\times[\zeta^{(k)} - \mathbf{f}^{(k)}(\zeta)]\} + O(\epsilon) = 0. \quad (23)$$

Now let the transformation inverse to (5) be

$$y = \mathbf{z} + \sum_{s=1}\epsilon^s\eta^{(s)}(\mathbf{z}). \quad (24)$$

Given the expansion (5), the $\eta^{(k)}$ may easily be derived (see the Appendix). Alternatively, they may be directly obtained from expanding the relation between (3) and (6), in a manner similar to what has been done for the expansion (5) in I. This is essentially the method of Krylov and Bogoliubov,[2] as expanded by Musen.[5] Indeed, the elimination of the angle variable by Kruskal's method so resembles the Krylov–Bogoliubov approach that the two ought perhaps to be regarded as one method (which could be called the Krylov-Bogoliubov-Kruskal method; however, Kruskal's work proceeds past the elimination to the derivation of invariants).

Since the inner part of (23) is a function of $\mathbf{y}$, one must transform

$$\frac{\partial}{\partial z_n} = \sum\frac{\partial y_i}{\partial z_n}\frac{\partial}{\partial y_i} = \sum\left[\delta_{in} + \sum\epsilon^s\frac{\partial\eta_i^{(s)}}{\partial z_n}\right]\frac{\partial}{\partial y_i}, \quad (25)$$

or

$$\frac{\partial}{\partial z_n} = \frac{\partial}{\partial y_n} + O(\epsilon). \quad (26)$$

Equation (23) thus becomes

$$\frac{\partial}{\partial y_n}\{\bar{\nabla}\times[\zeta^{(k)} - \mathbf{f}^{(k)}(\zeta)]\} + O(\epsilon) = 0. \quad (27)$$

One may now assume that all variables in (27) are expressed in terms of $\mathbf{y}$; in that case, each order in ϵ vanishes separately, including the zeroth. Because differential operators commute, this means

$$\bar{\nabla}\times\left\{\frac{\partial}{\partial y_n}[\zeta^{(k)} - \mathbf{f}^{(k)}(\zeta)]\right\} = 0. \quad (28)$$

This integrates to

$$\frac{\partial}{\partial y_n}[\zeta^{(k)} - \mathbf{f}^{(k)}(\zeta)] = \bar{\nabla}\psi, \quad (29)$$

$$\zeta^{(k)} - \mathbf{f}^{(k)}(\zeta) = \bar{\nabla}\tau + \mathbf{u}(\bar{\mathbf{y}}), \quad (30)$$

where τ is the indefinite integral of ψ and $\mathbf{u}$ is an additive function independent of z_n, allowed by the integration.

The above condition is satisfied by *any* $\zeta^{(k)}$ belonging to a nice set of variables which is canonical to order $k-1$. For instance, $\hat{\zeta}^{(k)}$ defined in (17) will correspond to a certain τ and to a certain additive function, which may be denoted by $\hat{\mathbf{u}}$:

$$\hat{\zeta}^{(k)} = \mathbf{f}^{(k)}(\zeta) + \bar{\nabla}\tau + \hat{\mathbf{u}}(\bar{\mathbf{y}}). \quad (31)$$

If one now chooses in (17)

$$\mu^{(k)} = -\hat{\mathbf{u}},$$

then

$$\begin{aligned}\zeta^{(k)} &= \hat{\zeta}^{(k)} - \hat{\mathbf{u}} \\ &= \mathbf{f}^{(k)}(\zeta) + \bar{\nabla}\tau,\end{aligned} \tag{32}$$

and by (21), $\zeta^{(k)}$ satisfies the condition for canonical variables, making z canonical to order k. Thus the requirement can be met.

III. PRACTICAL CANONIZATION

In order to actually derive the "canonizing" $\mu^{(k)}$, one must first investigate the amount of arbitrariness inherent in that vector. Let

$$\zeta^{(k)} = \hat{\zeta}^{(k)} + \mu^{(k)}(\tilde{\mathbf{y}}) \tag{33}$$

extend the canonical properties to $O(\epsilon^k)$, i.e. let it satisfy

$$\hat{\zeta}^{(k)} + \mu^{(k)}(\tilde{\mathbf{y}}) = \mathbf{f}^{(k)}(\zeta) + \bar{\nabla}\chi^{(k)}. \tag{34}$$

Then if one replaces $\mu^{(k)}$ with

$$\mu'^{(k)} = \mu^{(k)}(\tilde{\mathbf{y}}) + \bar{\nabla}\Phi(\tilde{\mathbf{y}}), \tag{35}$$

with Φ an arbitrary function independent of y_n, one finds

$$\hat{\zeta}^{(k)} + \mu'^{(k)}(\tilde{\mathbf{y}}) = \mathbf{f}^{(k)}(\zeta) + \nabla(\chi^{(k)} + \Phi), \tag{36}$$

which still has the required form (21). The canonizing choice of $\mu^{(k)}$ is thus arbitrary within the addition of a conjugate gradient of some scalar Φ which does not contain y_n.

In order to isolate $\mu^{(k)}$ one operates on (34) with the averaging operator of (16); since $\mu^{(k)}$ is independent of the angle variable, it equals its own average, giving

$$\begin{aligned}\mu^{(k)}(\tilde{\mathbf{y}}) &= \langle \mathbf{f}^{(k)}(\zeta) - \hat{\zeta}^{(k)} \rangle + \langle \bar{\nabla}\chi^{(k)} \rangle \\ &= \langle \mathbf{f}^{(k)}(\zeta) - \hat{\zeta}^{(k)} \rangle + \bar{\nabla}\langle \chi^{(k)} \rangle.\end{aligned} \tag{37}$$

The last term on the right is a conjugate gradient of some function of $\tilde{\mathbf{y}}$, and it has already been established that such vectors, when added to $\mu^{(k)}$ do not affect canonization. One may thus drop this part and obtain[8]

$$\mu^{(k)} = \langle \mathbf{f}^{(k)}(\zeta) - \hat{\zeta}^{(k)} \rangle. \tag{38}$$

To evaluate this, $\hat{\zeta}^{(k)}$ must be derived from (17), while $\mathbf{f}^{(k)}$ may be obtained by methods given in II. The *most general* canonizing additive function is then

$$\mu^{(k)} = \langle \mathbf{f}^{(k)}(\zeta) - \hat{\zeta}^{(k)} \rangle + \bar{\nabla}\Phi(\tilde{\mathbf{y}}), \tag{39}$$

with Φ arbitrary.

IV. EQUIVALENCE TO CANONICAL PERTURBATION THEORY

A widely used method for solving perturbed periodic canonical systems is due to Poincaré and Von Zeipel[5,9-12] and operates in the following manner.

First, one expresses the Hamiltonian in terms of the solution

$$\mathbf{y} = (\mathbf{p}, \mathbf{q})$$

of the unperturbed Hamilton–Jacobi equation, with y_n an angle variable and y_1 the conjugate action variable. In the absence of "slowly varying" quantities, the Hamiltonian then assumes the form

$$H = y_1 + \sum_{k=1} \epsilon^k H^{(k)}(\mathbf{y}). \tag{40}$$

Next a near-identity transformation to new variables

$$\mathbf{z} = (\mathbf{P}, \mathbf{Q})$$

is sought, produced by the generating function

$$\sigma(\mathbf{P}, \mathbf{q}) = \sum P_i q_i + \sum_{k=1} \epsilon^k \sigma^{(k)}(\mathbf{P}, \mathbf{q}), \tag{41}$$

and having the property that the new Hamiltonian $H^*(\mathbf{z})$ is independent of the transformed angle variable z_n, making the conjugate z_1 a constant of the motion. Since the transformation is a near-identity one, the lowest order of H^* has the same form as that of H, giving

$$H^*(\tilde{\mathbf{z}}) = z_1 + \sum_{k=1} \epsilon^k H^{*(k)}(\tilde{\mathbf{z}}). \tag{42}$$

Methods then exist for deriving $\sigma^{(k)}$ and $H^{*(k)}$ order by order.

Suppose now that two near-identity canonical transformations are given:

$$\begin{aligned}\mathbf{z} &= \mathbf{y} + \sum_{s=1} \epsilon^s \zeta^{(s)}(\mathbf{y}), \\ \mathbf{w} &= \mathbf{y} + \sum_{s=1} \epsilon^s \psi^{(s)}(\mathbf{y}),\end{aligned} \tag{43}$$

either of which eliminates the angle variable from the Hamiltonian. Furthermore, let

$$\zeta^{(s)}(\mathbf{y}) = \psi^{(s)}(\mathbf{y}) \quad \text{for} \quad s = 1, 2, \ldots, (k-1). \tag{44}$$

Then it will be shown that $\zeta^{(k)}$ differs from $\psi^{(k)}$ at most by a conjugate gradient of a function $\Phi(\tilde{\mathbf{y}})$ independent of y_n.

Clearly such a property would allow the two methods discussed in this work to be matched order by order. If w, for instance, represents a solution of a given problem by the Poincaré–Von Zeipel method and z represents a solution by the Kurskal method, then z can be made equal to w to any desired order. To achieve this, one must first choose $\mu^{(k)}$ so as to make z canonical and then add to each order the appropriate $\bar{\nabla}\Phi(\tilde{\mathbf{y}})$ which makes the corresponding $\zeta^{(k)}$ and $\psi^{(k)}$ equal to each other.

Actually, one could also work in the opposite direction, since the Poincaré–Von Zeipel method also contains a certain amount of arbitrariness in each order, but we shall not consider this possibility here.

Let

$$H^*(\tilde{\mathbf{z}}) = \sum_{s=0} \epsilon^s H^{*(s)}(\tilde{\mathbf{z}})$$

and

$$H^{**}(\tilde{\mathbf{w}}) = \sum_{s=0} \epsilon^s H^{**(s)}(\tilde{\mathbf{w}})$$

be the two alternative forms of the Hamiltonian. Since the transformation is time independent,

$$H^*(\tilde{\mathbf{z}}) = H^{**}(\tilde{\mathbf{w}}) = H(\mathbf{y}). \tag{45}$$

Also, since it is a near-identity transformation, the $O(1)$ parts of the above equations must be equal in form, which leads to

$$\begin{aligned} H^{*(0)}(\tilde{\mathbf{z}}) &= z_1, \\ H^{**(0)}(\mathbf{w}) &= w_1. \end{aligned} \tag{46}$$

We shall furthermore assume (and later justify)

$$H^{*(s)} = H^{**(s)}, \qquad s = 1, 2, \ldots, (k-1). \tag{47}$$

Substituting the expansions in (45)

$$\sum \epsilon^s H^{*(s)}(\tilde{\mathbf{y}} + \sum \epsilon^m \zeta^{(m)}) = \sum \epsilon^s H^{**(s)}(\tilde{\mathbf{y}} + \sum \epsilon^m \psi^{(m)}). \tag{48}$$

By means of expansion operators $T^{(m)}$ and $S^{(m)}$ (see I) this may be broken up into a series of equations, one for each order in ϵ. The equation for $O(\epsilon^k)$ then is

$$\sum_{m=0}^{k} T^{(m)} * H^{*(k-m)}(\tilde{\mathbf{y}}) = \sum_{m=0}^{k} S^{(m)} * H^{**(k-m)}(\tilde{\mathbf{y}}), \tag{49}$$

where $*$ marks "operates on" and where,

$$\begin{aligned} T^{(0)} &= S^{(0)} = 1, \\ T^{(1)} &= \boldsymbol{\zeta}^{(1)} \cdot \nabla, \\ S^{(1)} &= \boldsymbol{\psi}^{(1)} \cdot \nabla, \end{aligned}$$

and in general

$$\begin{aligned} T^{(m)} &= \boldsymbol{\zeta}^{(m)} \cdot \nabla + N^{(m)}(\zeta), \\ S^{(m)} &= \boldsymbol{\psi}^{(m)} \cdot \nabla + N^{(m)}(\psi), \end{aligned} \tag{50}$$

with $N^{(m)}$ an operator depending only on orders of its argument that are lower than the mth. Substitution and use of (46) give

$$\begin{aligned} &H^{*(k)} + \boldsymbol{\zeta}^{(k)} \cdot \nabla y_1 + \{N^{(k)}(\zeta) * y_1 + \sum_{m=1}^{k-1} T^{(m)} * H^{*(k-m)}\} \\ &\quad = H^{**(k)} + \boldsymbol{\psi}^{(k)} \cdot \nabla y_1 + \{N^{(k)}(\psi) * y_1 \\ &\quad + \sum_{m=1}^{k-1} S^{(m)} * H^{**(k-m)}\}. \end{aligned} \tag{51}$$

The two expressions in curly brackets depend only on lower orders; they are therefore equal to each other and may be dropped, leaving

$$H^{*(k)}(\tilde{\mathbf{y}}) - H^{**(k)}(\tilde{\mathbf{y}}) = (\boldsymbol{\psi}^{(k)} - \boldsymbol{\zeta}^{(k)}) \cdot \nabla y_1. \tag{52}$$

Now by (24), scalars $\chi^{(k)}$ and $\tau^{(k)}$ must exist such that

$$\begin{aligned} \boldsymbol{\zeta}^{(k)} &= \mathbf{f}^{(k)}(\zeta) + \bar{\nabla}\chi^{(k)}, \\ \boldsymbol{\psi}^{(k)} &= \mathbf{f}^{(k)}(\psi) + \bar{\nabla}\tau^{(k)}. \end{aligned} \tag{53}$$

Since lower orders of the two expansions are equal, the two $\mathbf{f}^{(k)}$ vectors are equal too, leaving

$$\boldsymbol{\zeta}^{(k)} - \boldsymbol{\psi}^{(k)} = \bar{\nabla}(\chi^{(k)} - \tau^{(k)}) = \bar{\nabla}\Phi. \tag{54}$$

To prove our assertion we must show that Φ is independent of y_n. Substituting in (52), we obtain

$$H^{*(k)}(\tilde{\mathbf{y}}) - H^{**(k)}(\tilde{\mathbf{y}}) = \frac{\partial \Phi}{\partial \tilde{y}_1} = \frac{\partial \Phi}{\partial y_n} \tag{55}$$

Now Φ is allowed to depend on y_m only in a periodic manner, from which it follows that $\partial\Phi/\partial y_n$ also depends periodically on y_n. However, the left-hand side is independent of that variable, so that Φ must be independent of y_n. This proves the main assertion. Incidentally, one also finds

$$H^{*(k)} = H^{**(k)}, \tag{56}$$

which justifies Eq. (47).

APPENDIX: THE INVERSE TRANSFORMATION

Adding (5) and (24) and cancelling zeroth-order terms gives

$$\begin{aligned} \sum \epsilon^k \eta^{(k)}(\mathbf{z}) &= -\sum \epsilon^k \boldsymbol{\zeta}^{(k)}(\mathbf{y}) \\ &= -\sum \epsilon^k \boldsymbol{\zeta}^{(k)}(\mathbf{z} + \sum \epsilon^m \boldsymbol{\eta}^{(m)}(\mathbf{z})), \end{aligned} \tag{A1}$$

if ∇ is in z space and $*$ denotes the operation

$$\begin{aligned} &\boldsymbol{\zeta}^{(k)}(\mathbf{z} + \sum \epsilon^m \boldsymbol{\eta}^{(m)}(\mathbf{z})) \\ &\quad = \exp\left[\sum_{m=1} \epsilon^m (\boldsymbol{\eta}^{(m)}(\mathbf{z}) \cdot \nabla)\right] * \boldsymbol{\zeta}^{(k)}(\mathbf{z}) \\ &\quad = \sum_{m=0} \epsilon^m S^{(m)} * \boldsymbol{\zeta}^{(k)}(\mathbf{z}), \end{aligned} \tag{A2}$$

where

$$\begin{aligned} S^{(0)} &= 1, \qquad S^{(1)} = \boldsymbol{\eta}^{(1)} \cdot \nabla, \\ S^{(2)} &= \boldsymbol{\eta}^{(2)} \cdot \nabla + \tfrac{1}{2} \boldsymbol{\eta}^{(1)} \boldsymbol{\eta}^{(1)} : \nabla\nabla, \end{aligned} \tag{A3}$$

and so forth. If this is substituted in (A1), orders of ϵ may be individually equated since everything is now expressed in z. Separating the $m = 0$ term from the rest then brings the $O(\epsilon^k)$ relation to the form

$$\boldsymbol{\eta}^{(k)} = -\boldsymbol{\zeta}^{(k)} - \sum_{m=1}^{k-1} S^{(m)} * \boldsymbol{\zeta}^{(k-m)}. \tag{A4}$$

If all lower orders of $\boldsymbol{\zeta}^{(m)}$ are known, those of $\boldsymbol{\eta}^{(m)}$ may be derived and used for constructing the $S^{(m)}$.

[1] M. Kruskal, J. Math. Phys. 3, 806 (1962).
[2] N. N. Bogoliubov and D. N. Zubarev, Ukr. Math. J. 7, 5 (1955) English transl: Space Technology Laboratories, 60. The method of Krylov and Bogoliubov is described in N. M. Krylov and N. N. Bogoliubov, *Introduction to Non-Linear Mechanics* (Kiev, 1936); N. N. Bogoliubov and Y. A. Mitropolsky, *Asymptotic Methods in the Theory of Non-linear Oscillations* (Gordon and Breach, New York, 1961), Sec. 25.
[3] Kruskal's original method is slightly more general than this in allowing the nonzero component of (4) to be a function independent of the angle variable. The recursion can then still be performed, but in each order the derivation of the nth component must be postponed until all other components have been found. If y is canonical, this more general behavior only occurs if some of the variables are "slowly varying."
[4] D. Stern, J. Math. Phys. 11, 2771 (1971).
[5] P. Musen, J. Astron. Sci. 12, 129 (1965).
[6] L. Nordheim and E. Fues, *Handbuch der Physik*, Ed. S. Flügge (Springer, Berlin, 1927).
[7] D. Stern, J. Math. Phys. 11, 2776 (1971).
[8] Results equivalent to the lower orders of Eq. (38), have been derived by E. L. Burshtein and L. S. Solovev, Dokl. Akad. Nauk SSSR 139, 855 (1961) [Sov. Phys. Dokl. 6, 731 (1962)].
[9] H. Poincaré, *Les Methodes Nouvelles de la Mecanique Celeste*, (Gauthier-Villars, Paris, 1893), Vol. II, reprint (Dover, New York, 1957); NASA translation TT F-451, 1967.
[10] H. Von Zeipel, Ark. Astr. Mat. Fys. 11, No. 1 (1916).
[11] H. C. Corben and P. Stehle, *Classical Mechanics* 2nd ed. (Wiley, New York, 1960), Sec. 75.
[12] D. Ter Haar, *Elements of Hamiltonian Mechanics* (North-Holland, Amsterdam, 1961), Chap. 7.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 10 OCTOBER 1971

Classical Adiabatic Perturbation Theory

David P. Stern
Laboratory for Space Physics, Goddard Space Flight Center, Greenbelt, Maryland 20771
(Received 21 December 1970)

Methods of classical perturbation theory developed for small perturbations are extended to slowly (or adiabatically) perturbed systems, with slow dependence either on time or on dynamical variables. Specifically, the extension is performed for the canonical perturbation theory of Poincaré and Von Zeipel, for the Krylov–Bogoliubov–Kruskal method of eliminating angle variables, for the general form of direct near-identity canonical transformations and for two of its realizations, based on the "conventional" generating function and on the Lie transform. In addition, the concepts of slow (or adiabatic) perturbations and of an implicit "small parameter" ϵ are clarified, as is the distinction between two alternative definitions of adiabatic invariance, and as an example the solution of the slow perturbed harmonic oscillator up to and including $O(\epsilon^3)$ is derived.

INTRODUCTION

Perhaps the most widely studied perturbation problem in classical mechanics is that of *perturbed periodic motion.* If a motion is given that is soluble and periodic, the problem may be concisely defined as the derivation of an approximate solution for a motion that is slightly different.

This "slight change" applied to the motion is termed the *perturbation* and it usually belongs to one of two types: "*small*" perturbations and "*slow*" (or "adiabatic") ones. The difference between the two is best explained by assuming that the motion can be described by a Hamiltonian, although this condition is not essential. In a *slightly* perturbed motion the Hamiltonian may then be written

$$H = H^{(0)} + \epsilon H^{(1)} + \epsilon^2 H^{(2)} + \cdots, \tag{1}$$

where $\epsilon \ll 1$ is a small numerical parameter characterizing the magnitude of the perturbation and where the limit $\epsilon \to 0$ corresponds to the unperturbed motion. A typical example would be the motion of a planet around the sun as perturbed by the planet Jupiter. In that case $H^{(0)}$ describes the planet's Keplerian motion in the sun's gravity field while $H^{(1)}$ describes the lowest order of the perturbation induced by Jupiter. The zero-order Hamiltonian is then proportional to the solar mass m_s while $\epsilon H^{(1)}$ is proportional to the mass m_j of Jupiter: the ratio of the two terms (m_j/m_s) will be of the order 10^{-3} and this dimensionless quantity provides a natural choice for ϵ.

To illustrate a slow perturbation, consider a Hamiltonian that is slowly dependent on the time t (slow dependence may also involve canonical variables):

$$H = H(\mathbf{p}, \mathbf{q}, t). \tag{2}$$

Then the dependence is said to be *slow* if the terms produced by the operation $\partial/\partial t$ are by an order in ϵ smaller than the terms from which they are derived, e.g.,

$$\frac{\partial H}{\partial t} = O(\epsilon H). \tag{3}$$

Equation (3) is not quite precise, since it implies that ϵ has the dimension of t^{-1}. In fact, one always requires some natural time period T against which the rapidity of the time variation may be gauged, this usually being the period of the unperturbed system. With this taken into account, (3) becomes

$$\frac{\partial H}{\partial t} = O\left(\frac{\epsilon H}{T}\right), \tag{4}$$

and ϵ is clearly dimensionless.

We restrict ourselves here to problems having a single zero-order periodicity. If several distinct zero-order periodicities exist, the methods presented here must be extended and furthermore, complications due to resonance effects may arise.

In either type of problem there generally exists a steadily increasing "angle variable" appearing in the argument of sines and cosines, describing the nearly periodic part of the motion. One way of solving the problem then involves finding a transformation to new variables, such that the angle

[1] M. Kruskal, J. Math. Phys. 3, 806 (1962).
[2] N. N. Bogoliubov and D. N. Zubarev, Ukr. Math. J. 7, 5 (1955) English transl: Space Technology Laboratories, 60. The method of Krylov and Bogoliubov is described in N. M. Krylov and N. N. Bogoliubov, *Introduction to Non-Linear Mechanics* (Kiev, 1936); N. N. Bogoliubov and Y. A. Mitropolsky, *Asymptotic Methods in the Theory of Non-linear Oscillations* (Gordon and Breach, New York, 1961), Sec. 25.
[3] Kruskal's original method is slightly more general than this in allowing the nonzero component of (4) to be a function independent of the angle variable. The recursion can then still be performed, but in each order the derivation of the nth component must be postponed until all other components have been found. If y is canonical, this more general behavior only occurs if some of the variables are "slowly varying."
[4] D. Stern, J. Math. Phys. 11, 2771 (1971).
[5] P. Musen, J. Astron. Sci. 12, 129 (1965).
[6] L. Nordheim and E. Fues, *Handbuch der Physik*, Ed. S. Flügge (Springer, Berlin, 1927).
[7] D. Stern, J. Math. Phys. 11, 2776 (1971).
[8] Results equivalent to the lower orders of Eq. (38), have been derived by E. L. Burshtein and L. S. Solovev, Dokl. Akad. Nauk SSSR 139, 855 (1961) [Sov. Phys. Dokl. 6, 731 (1962)].
[9] H. Poincaré, *Les Methodes Nouvelles de la Mecanique Celeste*, (Gauthier-Villars, Paris, 1893), Vol. II, reprint (Dover, New York, 1957); NASA translation TT F-451, 1967.
[10] H. Von Zeipel, Ark. Astr. Mat. Fys. 11, No. 1 (1916).
[11] H. C. Corben and P. Stehle, *Classical Mechanics* 2nd ed. (Wiley, New York, 1960), Sec. 75.
[12] D. Ter Haar, *Elements of Hamiltonian Mechanics* (North-Holland, Amsterdam, 1961), Chap. 7.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 10 OCTOBER 1971

Classical Adiabatic Perturbation Theory

David P. Stern
Laboratory for Space Physics, Goddard Space Flight Center, Greenbelt, Maryland 20771
(Received 21 December 1970)

Methods of classical perturbation theory developed for small perturbations are extended to slowly (or adiabatically) perturbed systems, with slow dependence either on time or on dynamical variables. Specifically, the extension is performed for the canonical perturbation theory of Poincaré and Von Zeipel, for the Krylov–Bogoliubov–Kruskal method of eliminating angle variables, for the general form of direct near-identity canonical transformations and for two of its realizations, based on the "conventional" generating function and on the Lie transform. In addition, the concepts of slow (or adiabatic) perturbations and of an implicit "small parameter" ϵ are clarified, as is the distinction between two alternative definitions of adiabatic invariance, and as an example the solution of the slow perturbed harmonic oscillator up to and including $O(\epsilon^3)$ is derived.

INTRODUCTION

Perhaps the most widely studied perturbation problem in classical mechanics is that of *perturbed periodic motion.* If a motion is given that is soluble and periodic, the problem may be concisely defined as the derivation of an approximate solution for a motion that is slightly different.

This "slight change" applied to the motion is termed the *perturbation* and it usually belongs to one of two types: "*small*" perturbations and "*slow*" (or "adiabatic") ones. The difference between the two is best explained by assuming that the motion can be described by a Hamiltonian, although this condition is not essential. In a *slightly* perturbed motion the Hamiltonian may then be written

$$H = H^{(0)} + \epsilon H^{(1)} + \epsilon^2 H^{(2)} + \cdots, \tag{1}$$

where $\epsilon \ll 1$ is a small numerical parameter characterizing the magnitude of the perturbation and where the limit $\epsilon \to 0$ corresponds to the unperturbed motion. A typical example would be the motion of a planet around the sun as perturbed by the planet Jupiter. In that case $H^{(0)}$ describes the planet's Keplerian motion in the sun's gravity field while $H^{(1)}$ describes the lowest order of the perturbation induced by Jupiter. The zero-order Hamiltonian is then proportional to the solar mass m_s while $\epsilon H^{(1)}$ is proportional to the mass m_j of Jupiter: the ratio of the two terms (m_j/m_s) will be of the order 10^{-3} and this dimensionless quantity provides a natural choice for ϵ.

To illustrate a slow perturbation, consider a Hamiltonian that is slowly dependent on the time t (slow dependence may also involve canonical variables):

$$H = H(\mathbf{p}, \mathbf{q}, t). \tag{2}$$

Then the dependence is said to be *slow* if the terms produced by the operation $\partial/\partial t$ are by an order in ϵ smaller than the terms from which they are derived, e.g.,

$$\frac{\partial H}{\partial t} = O(\epsilon H). \tag{3}$$

Equation (3) is not quite precise, since it implies that ϵ has the dimension of t^{-1}. In fact, one always requires some natural time period T against which the rapidity of the time variation may be gauged, this usually being the period of the unperturbed system. With this taken into account, (3) becomes

$$\frac{\partial H}{\partial t} = O\left(\frac{\epsilon H}{T}\right), \tag{4}$$

and ϵ is clearly dimensionless.

We restrict ourselves here to problems having a single zero-order periodicity. If several distinct zero-order periodicities exist, the methods presented here must be extended and furthermore, complications due to resonance effects may arise.

In either type of problem there generally exists a steadily increasing "angle variable" appearing in the argument of sines and cosines, describing the nearly periodic part of the motion. One way of solving the problem then involves finding a transformation to new variables, such that the angle

variable is eliminated from the equations of motion. If the system also possesses a Hamiltonian H, the absence of the angle variable from H implies that its conjugate "action variable" is a constant of the motion, and this eliminates an additional variable from consideration. In *slowly* perturbed systems, such constants are call *adiabatic invariants*.[1] In *slightly perturbed* systems, no generally accepted name exists (Contopoulos, who investigated the relation between the two types of constants,[2] has termed them "third integrals"), but they are well known in celestial mechanics and may be derived in a variety of ways.

The purpose of this work is to show how two standard methods of celestial mechanics, designed to handle small perturbations, may be modified to deal with slow perturbation and lead to the derivation of adiabatic invariants. The two methods considered here are the Poincaré–Von Zeipel method[3-7] for solving the Hamilton–Jacobi equation and the Krylov–Bogoliubov procedure[7-12] (or the related method of Kruskal). In addition, it will be shown that the direct form of near-identity canonical transformations can also be adapted to cases in which some variables are slow.

EXPLICIT AND IMPLICIT ϵ

In the example of perturbed planetary motion the small parameter ϵ can be given an explicit numerical value. In problems of slowly perturbed motion this is often difficult to do and one may then speak of an *implicit* ϵ.

As the archtype of a slowly perturbed system, consider the "pulled-up pendulum"[13,14]: a simple pendulum is suspended from a hole in the ceiling and its suspension string is pulled up (or released) at a slow, though not necessarily constant rate. Obviously, the angular frequency ω of the pendulum will vary and, since work is being done against the centrifugal force of the oscillation, so will its energy E. However, as long as the rate at which the string is withdrawn is sufficiently slow (and does not resonate with the oscillation of the pendulum) an adiabatic invariant may be found, equaling E/ω in the lowest order.

Two points should be noted here. First, the perturbation need not be small: by the time the withdrawal is complete, the length of the pendulum may well have changed by a large factor. Secondly, while one can devise an explicit ϵ for the problem —e.g., $\epsilon = \omega\tau$, where τ is the time in which the length of the pendulum is reduced to $1/e$ of its value, at the given (instantaneous) rate—its value nowhere enters the calculation.

A more complicated example is provided by the motion of a charged particle in a slightly inhomogeneous magnetic field B. Here "slightly" means that the derivatives $\partial B_i/\partial x_j$ are all of order ϵ smaller than the components of the field intensity and its magnitude B. Thus the slowness is in the dependence on spatial coordinates and a scale length for gauging it is provided by the gyration radius ρ, giving, in analogy to Eq. (4),

$$\frac{\partial B_i}{\partial x_j} = O\left(\frac{\epsilon B}{\rho}\right). \tag{5}$$

Again, the value of ϵ does not explicitly enter, except through the requirement that for the perturbation approach (known as the guiding center theory) to be valid the problem must satisfy "Alfvén's criterion"

$$\frac{\rho}{B}\ \frac{\partial B_i}{\partial x_j} \ll 1.$$

An implicit ϵ may be "made visible" by the following device. Consider a Hamiltonian with slow time dependence: One may artificially introduce ϵ into its time derivative by writing

$$\frac{\partial H}{\partial t} = \epsilon \frac{\partial H}{\partial(\epsilon t)}. \tag{6}$$

Since

$$\frac{\partial H}{\partial \epsilon t} = 0(1),$$

this notation clearly displays the fact that the term is of order ϵ, and for this reason the Hamiltonian (2) is often written

$$H = H(\mathrm{p}, \mathrm{q}, \epsilon t).$$

A similar device may be used when there exists a slow dependence on dynamical variables; this can be quite useful in arranging the terms according to their orders in ϵ, but two things must be remembered. First, because of the way in which ϵ is introduced, expressions of the kth order which have a factor ϵ^k standing in front of them will also have "hidden inside" a factor ϵ^{-k}. Secondly, because a definite value of ϵ is never stated, such factors must be canceled before the final result is obtained.

An example may be useful here. Suppose a one-dimensional motion is given with a Hamiltonian that has a slow dependence on t, and it is also given that if this dependence is "frozen" (limit $\epsilon = 0$), the motion is periodic. The solution of such a motion usually begins with a canonical transformation to new variables $(\mathbf{P}, \mathbf{Q})$ which are the action-angle variables of the unperturbed motion. If S is the generating function of this transformation, which in general is also slowly dependent on t, then the new Hamiltonian H' is

$$\begin{aligned} H'(\mathbf{P}, \mathbf{Q}) &= H + \frac{\partial S}{\partial t} \\ &= H + \epsilon \frac{\partial S}{\partial(\epsilon t)} \\ &= H'^{(0)} + \epsilon H'^{(1)}. \end{aligned} \tag{7}$$

In the transformed Hamiltonian, the first-order correction $H'^{(1)}$ has a factor ϵ preceding it, but this factor is artificial and is balanced by a factor

ϵ^{-1} that is "hidden inside" the term, as is evident from the derivation. In practice, these factors must be canceled before, say, the canonical equations of motion are used.

THE POINCARE–VON ZEIPEL METHOD FOR SLOW TIME DEPENDENCE

Consider a canonical system with $2N$ variables which has a slow dependence on time. We assume that the Hamiltonian H may be expanded in powers of ϵ:

$$H = \sum \epsilon^k H^{(k)}(\mathbf{p}, \mathbf{q}, t). \tag{8}$$

We further assume that the Hamilton–Jacobi equation for $\epsilon \to 0$ has been solved and that the transformation derived by it has already been applied, deriving as action-angle variables for the unperturbed motion

$$(J, \Omega) = (p_1, q_1)$$

and giving

$$H^{(0)} = J\omega/2\pi, \tag{9}$$

with $\omega = \omega(t)$ the slowly varying angular frequency. In the limit $\epsilon \to 0$, evidently, ω is a constant and so are all the canonical variables, except for Ω which is then linear in time.

To 'solve' the motion we now seek a near-identity canonical transformation to new variables $(\mathbf{P}, \mathbf{Q})$, with

$$(J^*, \Omega^*) = (P_1, Q_1)$$

generated by

$$\sigma(\mathbf{P}, \mathbf{q}, t) = \sum P_i q_i + \sum \epsilon^k \sigma^{(k)}(\mathbf{P}, \mathbf{q}, t) \tag{10}$$

such that the new Hamiltonian H^* does not depend on Ω^*. This is somewhat similar to, but simpler than, an approach advocated by Gardner[15] and investigated by Contopoulos,[2] in which the same result is obtained by a succession of canonical transformations, each of which pushes the elimination of Ω^* from H^* one order higher.

If H^* is expanded in a manner similar to (8) and the time derivative is expressed as in (7), one obtains

$$\sum \epsilon^k H^{*(k)}(\mathbf{P}, \mathbf{Q}, t) = \sum \epsilon^k H^{(k)}(\mathbf{p}, \mathbf{q}, t) + \sum \epsilon^k \frac{\partial \sigma^{(k-1)}}{\partial \epsilon t}. \tag{11}$$

This equation contains $4N$ canonical variables, but half of them can be eliminated by means of the transformation equations

$$p_i = P_i + \sum \epsilon^m \frac{\partial \sigma^{(m)}}{\partial q_i}, \tag{12}$$

$$Q_i = q_i + \sum \epsilon^m \frac{\partial \sigma^{(m)}}{\partial P_i}. \tag{13}$$

To facilitate the elimination it is best to follow a method introduced by Musen[7] and use expansion operators[12] (* denotes operation, $\partial/\partial\mathbf{P}$, etc. are gradient-type operators):

$$\begin{aligned} H^{*(k)}(\mathbf{P}, \mathbf{Q}, t) &= H^{*(k)}\left(\mathbf{p}, \mathbf{q} + \sum \epsilon^m \frac{\partial \sigma^{(m)}}{\partial \mathbf{P}, t}\right) \\ &= \exp \sum_{m=1} \left(\epsilon^m \frac{\partial \sigma^{(m)}}{\partial \mathbf{P}} \cdot \frac{\partial}{\partial \mathbf{q}}\right) * H^{*(k)}(\mathbf{P}, \mathbf{q}, t) \\ &= \sum_{m=0} \epsilon^m T^{(m)} * H^{*(k)}(\mathbf{P}, \mathbf{q}, t), \end{aligned} \tag{14}$$

where

$$\begin{aligned} T^{(0)} &= 1, \\ T^{(1)} &= \sum_i \frac{\partial \sigma^{(1)}}{\partial P_i} \frac{\partial}{\partial q_i}, \\ T^{(2)} &= \sum_i \frac{\partial \sigma^{(2)}}{\partial P_i} \frac{\partial}{\partial q_i} + \\ &\quad + \tfrac{1}{2} \sum_{i,j} \frac{\partial \sigma^{(1)}}{\partial P_i} \frac{\partial \sigma^{(1)}}{\partial P_j} \frac{\partial^2}{\partial q_i \partial q_j}, \end{aligned} \tag{15}$$

etc. Similarly

$$H^{(k)}(\mathbf{p}, \mathbf{q}, t) = \sum_{m=0} \epsilon^m S^{(m)} * H^{(k)}(\mathbf{P}, \mathbf{q}, t), \tag{16}$$

$$\begin{aligned} S^{(0)} &= 1, \\ S^{(1)} &= \sum \frac{\partial \sigma^{(1)}}{\partial q_i} \frac{\partial}{\partial P_i}, \\ S^{(2)} &= \sum_i \frac{\partial \sigma^{(2)}}{\partial q_i} \frac{\partial}{\partial P_i} \\ &\quad + \tfrac{1}{2} \sum_{i,j} \frac{\partial \sigma^{(1)}}{\partial q_i} \frac{\partial \sigma^{(1)}}{\partial q_j} \frac{\partial^2}{\partial P_i \partial P_j}, \end{aligned} \tag{17}$$

and so forth. Substituting all this in (11) and collecting terms associated with ϵ^k gives

$$\sum_{m=0}^{k} T^{(m)} * H^{*(k-m)} = \sum_{m=0}^{k} S^{(m)} * H^{(k-m)} + \frac{\partial \sigma^{(k-1)}}{\partial(\epsilon t)}. \tag{18}$$

The terms with $m = 0$ simply equal $H^{*(k)}$ and $H^{(k)}$ and will be taken outside the summation. The terms with $m = k$ also have simple form, for in general

$$S^{(k)} = \sum \frac{\partial \sigma^{(k)}}{\partial q_i} \frac{\partial}{\partial P_i} + N^{(k)}, \tag{19}$$

where $N^{(k)}$ contains only terms with at least two differentiations. Substituting (9) then gives

$$S^{(k)} * H^{(0)} = \frac{\omega}{2\pi} \frac{\partial \sigma^{(k)}}{\partial \Omega}. \tag{20}$$

Because the transformation reduces to the identity

transformation in the limit of vanishing ϵ, $H^{*(0)}$ equals $H^{(0)}$ and due to (9) it satisfies

$$T^{(k)}*H^{*(0)} = 0$$

since $T^{(k)}$ operators involve only differentiation by the q_i, which $H^{(0)}$ does not contain. One then obtains the basic recursion relation

$$\frac{\omega}{2\pi}\frac{\partial\sigma^{(k)}}{\partial\Omega} - H^{*(k)}(\mathbf{P},\mathbf{q},t) = \Lambda^{(k)}(\mathbf{P},\mathbf{q},t), \tag{21}$$

with

$$\Lambda^{(k)} = \sum_{m=1}^{k-1}\left(T^{(m)}*H^{*(k-m)} - S^{(m)}*H^{(k-m)}\right) - H^{(k)} - \frac{\partial\sigma^{(k-1)}}{\partial\epsilon t} \tag{22}$$

depending only on orders lower than the kth. If Ω enters only as an angle variable with period unity, any function $F(\mathbf{P},\mathbf{q},t)$ may be resolved into an "average" part

$$\langle F\rangle = \int_0^1 F d\Omega \tag{23}$$

and a "purely periodic" part with zero average

$$(F)_{\mathrm{per}} = F - \langle F\rangle.$$

The derivative of a purely periodic function is also purely periodic and therefore, so is

$$\frac{\partial\sigma^{(k)}}{\partial\Omega} = \frac{\partial}{\partial\Omega}\left((\sigma^{(k)})_{\mathrm{per}} + \langle\sigma^{(k)}\rangle\right) \tag{24}$$

since the contribution of $\langle\sigma^{(k)}\rangle$ vanishes. On the other hand, $H^{*(k)}$ does not depend on Ω, so one gets the recursive relations

$$H^{*(k)} = -\langle\Lambda^{(k)}\rangle, \tag{25}$$

$$\frac{\partial\sigma^{(k)}}{\partial\Omega} = \frac{2\pi}{\omega}(\Lambda^{(k)})_{\mathrm{per}}. \tag{26}$$

Once these are solved, the calculation may be advanced to the next order.

EXAMPLE: THE HARMONIC OSCILLATOR[16]

The Hamiltonian H'' of a harmonic oscillator with a slowly time-dependent angular frequency $\omega(t)$ is

$$H'' = \frac{1}{2m}(\mathrm{P}^2 + \omega^2 m^2 Q^2). \tag{27}$$

If one 'freezes' the time dependence, one can solve the Hamilton–Jacobi equation and derive a canonical transformation to action-angle variables (J,Ω), generated by

$$W = \int\left(\frac{J\omega m}{\pi} - m^2\omega^2 Q^2\right)^{1/2} dQ. \tag{28}$$

Following this transformation, the new Hamiltonian H becomes

$$H = \frac{J\omega}{2\pi} + \epsilon J\left(\frac{\omega'}{4\pi\omega}\right)\sin(4\pi\Omega), \tag{29}$$

where the prime henceforth signifies the operation $\partial/\partial(\epsilon t)$. Let σ of (10) generate a transformation to (J^*,Ω^*) such that all orders $H^{*(k)}$ of the new Hamiltonian are independent of Ω^*. This, combined with the fact that in the present case the only differentiation performed by $T^{(m)}$ of (15) is $\partial/\partial\Omega$, allows all such operators to be ignored *except* for $T^{(0)}$.

A further simplification is obtained by noting that H contains only two orders, both linear in J: Using the argument of (19) this gives, for the terms of (18) depending on H,

$$\sum S^{(m)}*H^{(k-m)} = S^{(k)}*H^{(0)} + S^{(k-1)}*H^{(1)} = \frac{\omega}{2\pi}\frac{\partial\sigma^{(k)}}{\partial\Omega} + \frac{\partial\sigma^{(k-1)}}{\partial\Omega}\frac{\omega'}{4\pi\omega}\sin(4\pi\Omega). \tag{30}$$

In what follows, we will for conciseness *write* J instead of J^*, restoring the superscript—if necessary—only at the end. In analogy with (21) we then obtain as the basic recursion relation, for $k > 1$:

$$\frac{\omega}{2\pi}\frac{\partial\sigma^{(k)}}{\partial\Omega} - H^{*(k)} = -\frac{\partial\sigma^{(k-1)}}{\partial\Omega}\frac{\omega'}{4\pi\omega}\sin(4\pi\Omega) - (\sigma^{(k-1)})'. \tag{31}$$

Using (18) directly for $k = 0$, one simply gets the equality of $H^{(0)}$ and $H^{*(0)}$, while for $k = 1$ this yields

$$\frac{\omega}{2\pi}\frac{\partial\sigma^{(1)}}{\partial\Omega} - H^{*(1)} = -J\frac{\omega'}{4\pi\omega}\sin(4\pi\Omega), \tag{32}$$

from which we get

$$H^{*(1)} = 0, \tag{33}$$

$$\sigma^{(1)} = J(\omega'/8\pi\omega^2)\cos(4\pi\Omega). \tag{34}$$

Higher orders, derived by the use of (31), are

$$H^{*(2)} = -(J/16\pi)(\omega')^2/\omega^3,$$
$$\sigma^{(2)} = -(J/64\pi)(\omega'/\omega^2)^2\sin(8\pi\Omega) - (J/16\pi\omega)(\omega'/\omega^2)'\sin(4\pi\Omega), \tag{35}$$

$$H^{*(3)} = 0,$$
$$\sigma^{(3)} = -\frac{J}{384\pi}\left(\frac{\omega'}{\omega^2}\right)^3\cos(12\pi\Omega) - \frac{J}{128\pi\omega}\left[\left(\frac{\omega'}{\omega^2}\right)^2\right]'\cos(8\pi\Omega) + \frac{J}{128\pi}\left(\frac{\omega'}{\omega^2}\right)^3\cos(4\pi\Omega) - \frac{J}{32\pi\omega}\left[\left(\frac{\omega'}{\omega^2}\right)'/\omega\right]'\cos(4\pi\Omega). \tag{36}$$

Note that any term in an $O(\epsilon^k)$ expression contains the prime operator exactly k times, corresponding to the factor ϵ^{-k} "hidden inside."

At this stage Eqs. (12) and (13) could be used to express (J, Ω^*) in terms of (J^*, Ω), up to $O(\epsilon^3)$. In fact, expressing the result in this manner, in terms of mixed variables, is not too useful, and it pays to "invert" the result and express (J^*, Ω^*) in terms of (J, Ω), or vice versa. The shortest way to achieve this is by means of the *direct transformation* technique.[17] If

$$\mathbf{y} \equiv (\mathbf{p}, \mathbf{q})$$

are the "old" variables and

$$\mathbf{z} \equiv (\mathbf{P}, \mathbf{Q})$$

are the "new" ones, and if the relation between the two sets has the "direct" form

$$\mathbf{z} = \mathbf{y} + \sum_{k=1} \epsilon^k \zeta^{(k)}(\mathbf{y}), \tag{37}$$

then for this to be a *canonical* transformation, $\zeta^{(k)}$ must have the form

$$\zeta^{(k)} = \bar{\nabla} \chi^{(k)} + \mathbf{f}^{(k)}, \tag{38}$$

where $\bar{\nabla}$ is a gradient operator in "conjugate phase space"

$$\bar{\mathbf{y}} \equiv (\mathbf{q}, -\mathbf{p}),$$

the $\chi^{(k)}$ are arbitrary functions, and $\mathbf{f}^{(k)}$ are prescribed expressions involving lower orders. In particular, if (37) is the "direct" form of the transformation generated by (10), one may choose

$$\chi^{(k)}(\mathbf{y}) = -\sigma^{(k)}(\mathbf{p}, \mathbf{q}) = -\sigma^{(k)}(\mathbf{y}) \tag{39}$$

(i.e., $\mathbf{P}$ is everywhere replaced by $\mathbf{p}$). The corresponding $\mathbf{f}^{(k)}$ is

$$\mathbf{f}^{(k)} = -\sum_{m=1}^{k-1} U^{(m)} * \bar{\nabla} \sigma^{(k-m)}(\mathbf{y}), \tag{40}$$

with $U^{(m)}$ expansion operators depending only on the momentumlike components $\pi^{(m)}$ of $\zeta^{(m)}$:

$$\pi^{(m)} = (\zeta_1^{(m)}, \ldots, \zeta_N^{(m)}, 0, \ldots, 0),$$

with

$$U^{(1)} = \pi^{(1)} \cdot \nabla, \tag{41}$$

$$U^{(2)} = \pi^{(2)} \cdot \nabla + \tfrac{1}{2} \pi^{(1)} \pi^{(1)} : \nabla\nabla,$$

and so on.

Of particular interest is the derivation of the adiabatic invariant

$$z_1 = J^* = \sum \epsilon^k J^{*(k)}, \tag{42}$$

which will now be outlined.

To obtain $\sigma^{(k)}(\mathbf{y})$ one simply uses the expressions (34)–(36) *without* restoring the asterisk superscript (as was originally planned). To derive (41), note that only one component of canonical momentum enters the calculation, so that

$$U^{(1)} = J^{*(1)} \frac{\partial}{\partial J}, \qquad U^{(2)} = J^{*(2)} \frac{\partial}{\partial J} + \tfrac{1}{2} (J^{*(1)})^2 \frac{\partial^2}{\partial J^2}. \tag{43}$$

The second-derivative terms may be safely ignored, since all orders of $\sigma^{(k)}$ used here are found to be linear in J. Finally, the components of the conjugate gradient $\bar{\nabla}$ contributing to z_1 are simply

$$\frac{\partial \sigma^{(k)}}{\partial \bar{y}_1} = \frac{\partial \sigma^{(k)}}{\partial \Omega}.$$

For the first order, $\mathbf{f}^{(1)}$ vanishes and one obtains

$$J^{*(1)} = -\frac{\partial \sigma^{(1)}}{\partial \Omega} = J \frac{\omega'}{2\omega^2} \sin(4\pi\Omega). \tag{44}$$

The next terms are

$$J^{*(2)} = \frac{J}{8}\left(\frac{\omega'}{\omega^2}\right)^2 + \frac{J}{4\omega}\left(\frac{\omega'}{\omega^2}\right)' \cos(4\pi\Omega),$$

$$J^{*(3)} = \frac{J}{16}\left(\frac{\omega'}{\omega^2}\right)^3 \sin(4\pi\Omega), \qquad -\frac{J}{8\omega}\left[\left(\frac{\omega'}{\omega^2}\right)'/\omega\right] \sin(4\pi\Omega). \tag{45}$$

THE "OLD" NOTION OF ADIABATIC INVARIANCE

In some texts of mechanics[5] and in the older literature, the definition of adiabatic invariance differs somewhat from the one given here. The alternative definition is usually applied to one-dimensional systems (though generalizations for several dimensions exist) and is as follows:

Given a slowly perturbed periodic motion, consider the action integral

$$J = \oint p \, dq \tag{46}$$

evaluated over one period of the *unperturbed* system. As the system is perturbed, an "instantaneous" J may be evaluated at any time by "freezing" slowly varying quantities. Then J has the property of adiabatic invariance: If the system undergoes a finite perturbation—e.g., a finite change of the Hamiltonian from H_1 to H_2—the corresponding change in J may be made arbitrarily small by stretching out the perturbation over a sufficiently long time.

The action variable J of (46) is the same as the zero-order action variable with which the previously developed perturbation scheme begins, but

its "adiabatic invariance" differs in two respects from what was earlier defined as adiabatic invariance. First, there exists here no hierarchy of invariants each of which is conserved to some specified order and secondly, the definition concerns itself with the cumulative change in J over a long period in time. In fact, this property does not follow automatically from the definition of adiabatic invariance used earlier. It is nevertheless an extremely useful property, since it allows one to derive, using only the unperturbed variables, a quantity with long-term invariance properties, *without even specifying the perturbation.*

Since J is the zero-order part of J^*, we may use (12) to obtain [compare also Eq. (44)]

$$J^* = J - \epsilon \frac{\partial \sigma^{(1)}}{\partial \Omega} + O(\epsilon^2) = J + \epsilon J^{*(1)} + O(\epsilon^2). \tag{47}$$

As in (12), $\sigma^{(1)}$ means $\sigma^{(1)}(J^*, \Omega, t)$; since J^* is a constant of the motion, only Ω and the slow direct dependence on t contribute to the variation of the first-order correction $J^{*(1)}$. The basic reason for the long-term adiabatic invariance of J, stated earlier, is that by the arguments of Eq. (24) $J^{*(1)}$ is purely periodic in Ω, and therefore "nearly" purely periodic in t. Over long time intervals, its variation is therefore bounded, causing the long term conservation of J to be better than might otherwise be expected.

To demonstrate this, expand (47) to

$$J = J^* - \epsilon J^{*(1)}(J^*, \Omega, t) + O(\epsilon^2) = J^* - \epsilon J^{*(1)}(J^*, \Omega, 0) - \epsilon^2 t \frac{\partial J^{*(1)}}{\partial(\epsilon t)} + \cdots + O(\epsilon^2). \tag{48}$$

Let a time $T = O(\epsilon^{-1})$ pass. The first term on the right is conserved, while the second one will vary only through the variation of Ω. Since the dependence of this term on Ω is periodic, the resulting contribution is bounded and due to the factor preceding it, of order ϵ. The next term is also $O(\epsilon)$ and the same holds for higher terms in the expansion of the slow direct time dependence of $J^{*(1)}$. The $O(\epsilon^2)$ terms may contribute to dJ/dt a term of form $\epsilon^2\psi$, but its contribution to the total change of J will again be of order ϵ:

$$\epsilon^2 \int_0^T \psi dt = O(\epsilon^2 T) = O(\epsilon).$$

Hence, the long term variation of J is $O(\epsilon)$.

The variation of other dynamical quantities, on the other hand, will be finite. For instance, for H

$$\Delta H = \int_0^T \frac{dH}{dt} dt = \int_0^T \frac{\partial H}{\partial t} dt = T\left(\frac{\partial H}{\partial t}\right)_{\text{av}} = (\epsilon T)\left(\frac{\partial H}{\partial(\epsilon t)}\right)_{\text{av}} \tag{49}$$

and each factor here is $O(1)$. Thus by making ϵ arbitrarily small, but keeping $T = O(\epsilon^{-1})$, the variation of J may be made as small as is desired while that of H remains finite.

THE POINCARE–VON ZEIPEL METHOD FOR SLOW DEPENDENCE ON CANONICAL VARIABLES

Let a perturbed periodic motion be given, represented by a Hamiltonian

$$H = \sum \epsilon^k H^{(k)}(\mathbf{p}, \mathbf{q}), \tag{50}$$

with (p_1, q_1) the action-angle variables (J, Ω) of the unperturbed motion; since we have already derived methods dealing with slow time dependence, we will simplify matters by not including such a dependence here. The motion represented by $H^{(0)}$ alone is assumed to be periodic and soluble: we shall *not* require at this stage that $H^{(0)}$ has the form (9), but we note that it must be independent of Ω, since J is a constant of the unperturbed motion.

Instead, we shall assume that the canonical variables y_i fall into two groups: "normal" variables for which $\partial/\partial y_i$ maintains the same order in ϵ and "slow" ones for which it raises the order by one level. It is useful to define parameters that distinguish between the two groups: let γ_i equal 0 or 1 depending on whether q_i is normal or slow, and let δ_i play the same role for p_i. One can then define

$$\mathbb{Q}_i = \epsilon^{\gamma_i} q_i \tag{51}$$

$$\mathbb{P}_i = \epsilon^{\delta_i} p_i, \tag{52}$$

so that (for example) $\partial H/\partial \mathbb{Q}_i$ and $\partial H/\partial \mathbb{P}_i$ are always of the same order as H itself.

As before, let a generating function

$$\sigma(\mathbf{p}, \mathbf{q}) = \sum_i P_i q_i + \sum_{k=1} \epsilon^k \sigma^{(k)}(\mathbf{P}, \mathbf{q}) \tag{53}$$

define a near-identity transformation to a new canonical set $(\mathbf{P}, \mathbf{Q})$, with the new Hamiltonian H^* independent of the transformed angle variable Ω^* (a term with $k = 0$ could be included, but since it may not depend on Ω^*, it is not useful here). Again, the basic equation is

$$H^*(\mathbf{P}, \mathbf{Q}) = H(\mathbf{p}, \mathbf{q}) \tag{54}$$

and this again is expressed in powers of ϵ and expressed solely in terms of $(\mathbf{P}, \mathbf{q})$. Since $\mathbf{p}$ no longer appears, it helps to *redefine* $\mathbb{P}_i$ as

$$\mathbb{P}_i = \epsilon^{\delta_i} P_i; \tag{55}$$

this will be the definition used in the remainder of this section. In analogy with (14) one finds

$$H^{*(k)}(\mathbf{P},\mathbf{Q}) = \exp\left[\sum_{m=1}\epsilon^m \sum_i \frac{\partial\sigma^{(m)}}{\partial P_i}\frac{\partial}{\partial q_i}\right] * H^{*(k)}(\mathbf{P},\mathbf{q}) = \exp\left[\sum_{i,m}\epsilon^{m+\delta_i+\gamma_i}\frac{\partial\sigma^{(m)}}{\partial \mathbb{P}_i}\frac{\partial}{\partial \mathbb{Q}_i}\right] * H^{*(k)}(\mathbf{P},\mathbf{q})$$

$$= \exp\sum_{\boldsymbol{m}}\epsilon^m \sum_i \frac{\partial\sigma^{(m-\delta_i-\gamma_i)}}{\partial \mathbb{P}_i}\frac{\partial}{\partial \mathbb{Q}_i} * H^{*(k)}(\mathbf{P},\mathbf{q}) = \sum_{m=0}\epsilon^m V^{(m)} * H^{*(k)}, \tag{56}$$

with $V^{(m)}$ suitable operators and $\sigma^{(m)}$ vanishing for all nonpositive values of m. Expanding the exponential gives

$$V^{(0)} = 1,$$

$$V^{(1)} = \sum_i \frac{\partial\sigma^{(1-\delta_i-\gamma_i)}}{\partial \mathbb{P}_i}\frac{\partial}{\partial \mathbb{Q}_i}, \tag{57}$$

and so forth; because these operators are expressed solely in terms of $\mathbb{P}_i$ and $\mathbb{Q}_i$, their action on any function maintains the ordering in powers of ϵ.

Similarly,

$$H^{(k)}(\mathbf{p},\mathbf{q}) = \sum_{m=0}\epsilon^m R^{(m)} * H^{(k)}(\mathbf{P},\mathbf{q}), \tag{58}$$

with

$$R^{(0)} = 1,$$

$$R^{(1)} = \sum_i \frac{\partial\sigma^{(1-\delta_i-\gamma_i)}}{\partial \mathbb{Q}_i}\frac{\partial}{\partial \mathbb{P}_i}, \tag{59}$$

and so forth. Substituting these operators and collecting terms associated with ϵ^k then gives, in analogy with (18)

$$\sum_{m=0}^{k}\left(V^{(m)} * H^{*(k-m)} - R^{(m)} * H^{(k-m)}\right) = 0. \tag{60}$$

Again, the terms with $m = 0$ and $m = k$ are separated. For the latter terms one gets, in analogy with (19),

$$V^{(k)} = \sum_i \frac{\partial\sigma^{(k-\delta_i-\gamma_i)}}{\partial \mathbb{P}_i}\frac{\partial}{\partial \mathbb{Q}_i} + M^{(k)},$$

$$R^{(k)} = \sum_i \frac{\partial\sigma^{(k-\delta_i-\gamma_i)}}{\partial \mathbb{Q}_i}\frac{\partial}{\partial \mathbb{P}_i} + N^{(k)}, \tag{61}$$

with $M^{(k)}$ and $N^{(k)}$ involving only lower orders. By taking $k = 0$ in (60), one again finds

$$H^{*(0)} = H^{(0)} \tag{62}$$

so that (60) becomes, for the general case

$$H^{*(k)} + \sum_i\left(\frac{\partial\sigma^{(k-\delta_i-\gamma_i)}}{\partial \mathbb{P}_i}\frac{\partial H^{(0)}}{\partial \mathbb{Q}_i} - \frac{\partial\sigma^{(k-\delta_i-\gamma_i)}}{\partial \mathbb{Q}_i}\frac{\partial H^{(0)}}{\partial \mathbb{P}_i}\right) = G^{(k)}, \tag{63}$$

where

$$G^{(k)} = H^{(k)} + \sum_{m=1}^{k-1}\left(R^{(m)} * H^{(k-m)} - V^{(m)} * H^{*(k-m)}\right) + \left(M^{(k)} - N^{(k)}\right) * H^{(0)} \tag{64}$$

involves only given functions and lower orders. For every k a relation of this type is obtained, constituting a kth-order recursion formula for the derivation of $H^{*(k)}$ and $\sigma^{(k)}$.

Now the action-angle variables associated with the zero-order periodicity (and used here in mixed form)

$$(J^*, \Omega) \equiv (P_1, q_1)$$

are assumed to be "normal," so that the left-hand side of (63) will include a term

$$\frac{\partial\sigma^{(k)}}{\partial\Omega}\frac{\partial H^{(0)}}{\partial J^*}.$$

If $\sigma^{(k)}$ appears *nowhere else*, the equation assumes the form of (21) and is solved in the same manner.

On the other hand, if $\sigma^{(k)}$ appears anywhere in (63), it may not be possible to derive it, for the equation then becomes a partial differential equation for $\sigma^{(k)}$. To prevent this from happening, it is required that for all (p_i, q_i) appearing in $H^{(0)}$ other than the action-angle pair, we have the relation

$$\delta_i + \gamma_i \geq 1. \tag{65}$$

Hence the recursion can be carried out *if*:

(i) $H^{(0)}$ has normal dependence on p_1 but does not depend on q_1;
(ii) $H^{(0)}$ may depend on any "slow" variable;
(iii) $H^{(0)}$ may depend on any "normal" variable, provided its canonical conjugate is slow.

Furthermore, it may be shown by extending the present calculation that

(iv) H may include a term $H^{(-1)}$ of order ϵ^{-1}, provided it depends only on slow variables having slow conjugates. Such terms are then transformed intact to the new Hamiltonian.

As an example, the Hamiltonian of a charged particle in a time-dependent electromagnetic field, in the regime of guiding-center motion, may be brought to the form[15,18]

$$H = {p_2}^2/2m + p_1\omega/2\pi + \epsilon^{-1}e\phi^{(0)} + O(\epsilon). \tag{66}$$

Here (p_1, q_1) are canonical variables associated with the rapid gyration, (p_2, q_2) represent the motion along field lines, and (p_3, q_3) describe the identity of the guiding field line, which changes slowly with time (in the references, subscripts (1) and (3) have reversed meanings); the variables (p_1, q_1, p_2) are normal, whereas the remaining ones are slow. Furthermore, the gyration frequency ω and the lowest order $\phi^{(0)}$ of the electric

potential are both functions of slow variables only. The last term is of order ϵ^{-1}, since its derivatives are proportional to the components of the lowest order of the electric field **E**, which are of order 1.

Evidently, H meets all the previously stated conditions except for one: If $\phi^{(0)}$ contains q_2, condition (iv) is violated, since p_2 is not slow. One therefore must impose an additional requirement that $\partial\phi^{(0)}/\partial q_2$ vanishes: This reduces to the well-known restriction in guiding center theory that the electric field may have no zero-order component parallel to the magnetic field.

DIRECT CANONICAL TRANSFORMATIONS WITH SLOW VARIABLES

The generating function $\sigma^{(k)}$ gives the transformation equations as in (12) and (13), in mixed form. To bring them to the "direct" form (37) it is useful to generalize (38) for cases in which slow variables are present, and this will now be done.

Let

$$\mathbf{y} = (\mathbf{p}, \mathbf{q})$$

be a canonical set and

$$\bar{\mathbf{y}} = (\mathbf{q}, -\mathbf{p})$$

be its conjugate.[17] One may now define an index Γ_i equaling 0 or 1 depending on whether y_i is normal or slow, and an index Δ_i which has a similar relation to $\bar{y}_i$. With this notation, it is possible to define vectors $\mathbf{Y}$ and $\bar{\mathbf{Y}}$ satisfying relations similar to (51) and (52):

$$Y_i = \epsilon^{\Gamma_i} y_i, \tag{67}$$

$$\bar{Y}_i = \epsilon^{\Delta_i} \bar{y}_i. \tag{68}$$

As with quantities defined in (51) and (52), $\partial/\partial Y_i$ and $\partial/\partial \bar{Y}_i$ are always $O(1)$.

We now seek the condition for a near-identity transformation (37) to be canonical. Actually, in what follows the recursion may still be carried out even if the transformation is not a near-identity one and $\zeta_i^{(0)}$ terms are included (satisfying appropriate conditions), but we shall not develop this possibility here. One then finds, as a condition for canonical behavior

$$\begin{aligned}[y_i, y_j] &= [z_i, z_j] \\ &= [y_i + \textstyle\sum \epsilon^k \zeta_i^{(k)}, y_j + \sum \epsilon^m \zeta_j^{(m)}] \\ &= [y_i, y_j] + \textstyle\sum \epsilon^k \{[\zeta_i^{(k)}, y_j] - [\zeta_j^{(k)}, y_i] \\ &\quad + \sum_{m=1}^{k-1} [\zeta_i^{(m)}, \zeta_j^{(k-m)}]\}.\end{aligned} \tag{69}$$

Expressing derivatives in terms of $\mathbf{Y}$ and $\bar{\mathbf{Y}}$ gives

$$\begin{aligned}[a, b] &= \sum_s \frac{\partial a}{\partial \bar{y}_s} \frac{\partial b}{\partial y_s} \\ &= \sum_s \epsilon^{\Delta_s + \Gamma_s} \frac{\partial a}{\partial \bar{Y}_s} \frac{\partial b}{\partial Y_s}.\end{aligned} \tag{70}$$

In particular

$$[a, y_i] = \epsilon^{\Delta_i} \frac{\partial a}{\partial \bar{Y}_i}. \tag{71}$$

Thus

$$0 = \sum_{k=1} \left\{ \epsilon^{k+\Delta_j} \frac{\partial \zeta_i^{(k)}}{\partial \bar{Y}_j} - \epsilon^{k+\Delta_i} \frac{\partial \zeta_j^{(k)}}{\partial \bar{Y}_i} + \sum_s \epsilon^{k+\Delta_s+\Gamma_s} \sum_{m=1}^{k-1} \frac{\partial \zeta^{(m)}}{\partial \bar{Y}_s} \frac{\partial \zeta_j^{(k-m)}}{\partial Y_s} \right\}. \tag{72}$$

Dividing by $\epsilon^{\Delta_i + \Delta_j}$,

$$0 = \sum_{k=\Delta_i+\Delta_j+1} \left(\epsilon^{k-\Delta_i} \frac{\partial \zeta_i^{(k)}}{\partial \bar{Y}_j} - \epsilon^{k-\Delta_j} \frac{\partial \zeta_j^{(k)}}{\partial \bar{Y}_i} + \sum_s \epsilon^{k+\Gamma_s+\Delta_s-\Delta_i-\Delta_j} \sum_{m=1}^{k-1} \frac{\partial \zeta_i^{(m)}}{\partial \bar{Y}_s} \frac{\partial \zeta_j^{(k-m)}}{\partial Y_s} \right). \tag{73}$$

It is useful at this point to redefine k for each term so that all powers of ϵ become ϵ^k and also to replace m by

$$M = m - \Delta_i. \tag{74}$$

Because the exponent of ϵ differs for each term in (73), the new summation over k will begin at a different value for each term; this summation limit may, however, be uniformly set equal to 1 if it is assumed that $\zeta_i^{(u)}$ vanishes for nonpositive values of u. With these changes, (73) gives

$$0 = \sum_{k=1} \epsilon^k \left(\frac{\partial \zeta_i^{(k+\Delta_i)}}{\partial \bar{Y}_j} - \frac{\partial \zeta_j^{(k+\Delta_j)}}{\partial \bar{Y}_i} + \sum_{M=1-\Delta_i}^{k-\Delta_i-1} \sum_s \frac{\partial \zeta_i^{(M+\Delta_i)}}{\partial \bar{Y}_s} \frac{\partial \zeta_j^{(k-M-\Gamma_s-\Delta_s+\Delta_j)}}{\partial Y_s} \right). \tag{75}$$

This suggests the introduction of new "staggered" vectors

$$\boldsymbol{\vartheta}^{(k)} = \boldsymbol{\zeta}^{(k+\Delta_i)}, \tag{76}$$

i.e.,

$$\begin{aligned} \vartheta_i^{(k)} &= \zeta_i^{(k)}, \quad \text{if } \Delta_i = 0, \\ \vartheta_i^{(k)} &= \zeta_i^{(k+1)}, \quad \text{if } \Delta_i = 1. \end{aligned} \tag{77}$$

It is also useful (in analogy to what was done in Ref. 17) to introduce a curl operator in $\bar{\mathbf{Y}}$ space. With this notation (75) may be rewritten as

$$(\bar{\nabla}_Y \times \boldsymbol{\vartheta}^{(k)})_{ij} = -\sum_s \sum_{m=1}^{k-1-\Gamma_s-\Delta_s+\Delta_j} \frac{\partial \vartheta_i^{(M)}}{\partial \bar{Y}_s} \frac{\partial \vartheta_j^{(k-M-\Gamma_s-\Delta_s)}}{\partial Y_s}. \tag{78}$$

This equation may in principle be used to derive $\mathfrak{z}^{(k)}$ recursively, but this turns out to be a rather inconvenient approach. It is more useful in determining the degree of arbitrariness associated with a near-identity canonical transformation of the form (37). Let two such transformations be given, characterized by staggered vectors $\mathfrak{z}^{(m)}$ and $\mathbf{x}^{(m)}$ which are identical for orders up to and including the $(k-1)$. For the kth order one finds that the right-hand side of (78), which depends only on lower orders, is identical for both expansions, giving

$$\bar{\nabla}_Y \times (\mathfrak{z}^{(k)} - \mathbf{x}^{(k)}) = 0, \tag{79}$$

from which we get

$$\mathfrak{z}^{(k)} = \mathbf{x}^{(k)} + \bar{\nabla}_Y \chi^{(k)}. \tag{80}$$

Thus the arbitrariness in specifying the canonical transformation at each level of $\mathfrak{z}^{(k)}$ is contained in the gradient in $\bar{Y}$ space of an arbitrary scalar $\chi^{(k)}$. The general form of $\mathfrak{z}^{(k)}$ for a canonical transformation may be written, in analogy with (38),

$$\mathfrak{z}^{(k)} = \bar{\nabla}_Y \chi^{(k)} + \mathbf{F}^{(k)}, \tag{81}$$

where $\mathbf{F}^{(k)}$ is a vector involving orders of $\mathfrak{z}^{(m)}$ lower than the kth and constitutes one particular solution of (78). In the following sections two such particular solutions will be derived, analogous to those found in Ref. 17 for *small* perturbations.

DERIVATION BASED ON $\sigma(\mathbf{P}, \mathbf{q})$

Let a near-identity transformation of $n = 2N$ variables

$$\mathbf{y} = (\mathbf{p}, \mathbf{q}) \rightarrow \mathbf{z} = (\mathbf{P}, \mathbf{Q}) \tag{82}$$

be given by (37), and let a generating function (53) be assumed to produce the same transformations via Eqs. (12) and (13). In what follows, the relation between (37) and (53) will be established in a way resembling what was done in Ref. 17 for the case when no slow variables are present. As before, the calculation may be broadened somewhat beyond what is done here, since the method only requires that canonical *momenta* transform in near-identical fashion.

With the notation of (51) and (55) Eqs. (12) and (13) give

$$P_i = p_i - \sum \epsilon^k \frac{\partial \sigma^{(k)}}{\partial q_i} = p_i - \sum \epsilon^{k+\delta_i} \frac{\partial \sigma^{(k)}}{\partial \mathbb{Q}_i}, \tag{83}$$

$$Q_i = q_i + \sum \epsilon^k \frac{\partial \sigma^{(k)}}{\partial P_i} = q_i + \sum \epsilon^{k+\delta_i} \frac{\partial \sigma^{(k)}}{\partial \mathbb{P}_i}. \tag{84}$$

All functions on the right depend on mixed variables $(\mathbf{P}, \mathbf{q})$; to introduce a dependence on $\mathbf{y}$, it is useful to define "partial vectors" adding up to $\zeta^{(k)}$

$$\begin{aligned} \pi^{(k)} &= (\zeta_1^{(k)}, \cdots, \zeta_{N-1}^{(k)}, 0, \cdots, 0), \\ \theta^{(k)} &= (0, \cdots, 0, \zeta_N^{(k)}, \cdots, \zeta_n^{(k)}). \end{aligned} \tag{85}$$

From this

$$P_i = p_i + \sum \epsilon^k \pi_i^{(k)}(\mathbf{y}), \tag{86}$$

$$Q_i = q_i + \sum \epsilon^k \theta_{N+i}^{(k)}(\mathbf{y}). \tag{87}$$

If the vectors $\pi^{(k)}$ are known, they may also be used to expand any function of mixed variables $(\mathbf{P}, \mathbf{q})$ in terms of $\mathbf{y}$, e.g.,

$$\begin{aligned} F(\mathbf{P}, \mathbf{q}) &= F(\mathbf{y} + \textstyle\sum \epsilon^k \boldsymbol{\pi}^{(k)}) \\ &= \exp\left[\sum \epsilon^k \left(\boldsymbol{\pi}^{(k)} \cdot \frac{\partial}{\partial \mathbf{y}}\right)\right] * F(\mathbf{y}) \\ &= \exp\left[\sum \epsilon^k \sum_s \left(\pi_s^{(k-\Gamma_s)} \frac{\partial}{\partial Y_s}\right)\right] * F(\mathbf{y}) \\ &= \sum_{k=0} \epsilon^k L^{(k)} * F(\mathbf{y}), \end{aligned} \tag{88}$$

where, if we collectively denote all "slow" components of $\mathbf{y}$ by $\mathbf{R}$ and all "normal" ones by $\mathbf{r}$ and if $\boldsymbol{\pi}_R^{(k)}$ and $\boldsymbol{\pi}_r^{(k)}$ denote vectors composed of the corresponding components of $\boldsymbol{\pi}^{(k)}$, then

$$\begin{aligned} L^{(0)} &= 1, \\ L^{(1)} &= \boldsymbol{\pi}_r^{(1)} \cdot \frac{\partial}{\partial \mathbf{r}}, \\ L^{(2)} &= \boldsymbol{\pi}_r^{(2)} \cdot \frac{\partial}{\partial \mathbf{r}} + \boldsymbol{\pi}_R^{(1)} \cdot \frac{\partial}{\partial \epsilon \mathbf{R}} + \tfrac{1}{2} \boldsymbol{\pi}_r^{(1)} \boldsymbol{\pi}_r^{(1)} : \frac{\partial}{\partial \mathbf{r}} \frac{\partial}{\partial \mathbf{r}}, \end{aligned} \tag{89}$$

and so forth. Note that since ϵ is implicit, $\pi^{(1)}$ and $\pi^{(2)}$ should have factors ϵ^{-1} and ϵ^{-2} "hidden inside," since they are teamed up with the corresponding positive powers in (86). For the same reason $L^{(2)}$ should contain a factor ϵ^{-2} and, indeed, inspection of the last equality in (89) shows that all terms have such a factor.

Substitution in (83) yields

$$\begin{aligned} P_i &= p_i - \sum \epsilon^{k+\gamma_i} \sum_m \epsilon^m L^{(m)} * \frac{\partial \sigma^{(k)}}{\partial \mathbb{Q}_i} \\ &= p_i - \sum \epsilon^k \left(\frac{\partial \sigma^{(k-\gamma_i)}}{\partial \mathbb{Q}_i} + \sum_{m=1}^{k-\gamma_i} L^{(m)} * \frac{\partial \sigma^{(k-m-\gamma_i)}}{\partial \mathbb{Q}_i} \right), \end{aligned} \tag{90}$$

where all terms of σ are viewed as functions of $\mathbf{y}$, i.e., with $\mathbf{p}$ replacing $\mathbf{P}$, wherever the latter originally appeared. This should be identical to (83), and therefore

$$\pi_i^{(k)} = -\frac{\partial \sigma^{(k-\gamma_i)}}{\partial \mathbb{Q}_i} - \sum_{m=1}^{k-1} L^{(m)} * \frac{\partial \sigma^{(k-m-\gamma_i)}}{\partial \mathbb{Q}_i}. \tag{91}$$

The highest order of $L^{(m)}$ appearing on the right is $k-1$ [$\sigma^{(0)}$ only appears if $\gamma_i = 1$, for if it depends on "normal" variables, the transformation is no longer one of near-identity: this is the reason for the change in summation limit] and this is therefore also the highest order of $\pi^{(s)}$ appearing on the right. Thus, (91) is a usable recursion relation for deriving $\pi^{(k)}$.

Expanding (84) in a similar manner gives

$$\theta_i^{(k)} = \frac{\partial \sigma^{(k-\gamma_i)}}{\partial \mathbb{P}_i} + \sum_{m=1}^{k-1} L^{(m)} * \frac{\partial \sigma^{(k-m-\gamma_i)}}{\partial \mathbb{P}_i}, \tag{92}$$

where the definition of $\mathbb{P}_i$ reverts to (52).

Now if

$$y_i = q_j,$$

then

$$\bar{Y}_i = -\mathbb{P}_j, \qquad \Delta_i = \delta_j;$$

and if

$$y_i = p_j,$$

then

$$\bar{Y}_i = \mathbb{Q}_j, \qquad \Delta_i = \gamma_j.$$

Inspection then shows that (91) and (92) may be combined to one equation:

$$z_i = y_i - \sum \epsilon^k \left(\frac{\partial \sigma^{(k-\Delta_i)}}{\partial \bar{Y}_i} + \sum_{m=1}^{k-1} L^{(m)} * \frac{\partial \sigma^{(k-m-\Delta_i)}}{\partial \bar{Y}_i} \right). \tag{93}$$

The dependence on Δ_i may be removed by introducing $\mathfrak{z}^{(k)}$ defined in (76). Then, using the gradient operator in $\bar{\mathbf{Y}}$ space, (93) becomes

$$\mathfrak{z}^{(k)} = -\bar{\nabla}_Y \sigma^{(k)} - \sum_{m=1}^{k-1} L^{(m)} * \bar{\nabla}_Y \sigma^{(k-m)}. \tag{94}$$

Since it has already been established in (81) that $\mathfrak{z}^{(k)}$ is arbitrary within some gradient in $\bar{\mathbf{Y}}$ space, the summation term represents a particular solution of (78).

LIE TRANSFORMS WITH SLOW VARIABLES[16]

If L_W is the operator denoting Poisson bracketing with a function W of the canonical variables

$$L_W(f) = [f, W], \tag{95}$$

then it may be shown[17, 19-21] that the transformation from $y = (\mathbf{p}, \mathbf{q})$ to

$$\mathbf{z} = \exp(\epsilon L_W) * \mathbf{y} \tag{96}$$

(with the exponential operator defined by its series expansion) is canonical. In what follows, the form of the Lie transform in the presence of slow variables will be derived, again following closely the derivation for the simpler case when all y_i vary on the same scale.[17] Let W be expandable in ϵ

$$W = \sum \epsilon^k W^{(k)}(\mathbf{y}) \tag{97}$$

and let operators $L_W{}^{(k)}$ be defined through

$$\epsilon L_W = \sum \epsilon^k L_W^{(k)}. \tag{98}$$

If one defines [compare Ref. 17, Eq. (35)]

$$\chi^{(k)} = -W^{(k-1)}, \tag{99}$$

then

$$L_W = -\sum \epsilon^{k+1} \sum_s \frac{\partial \chi^{(k)}}{\partial \bar{y}_s} \frac{\partial}{\partial y_s}$$

$$= \sum \epsilon^k \sum_s \frac{\partial \chi^{(k-\Gamma_s-\Delta_s)}}{\partial \bar{Y}_s} \frac{\partial}{\partial Y_s}, \tag{100}$$

where the lower limit of k in the last summation may be chosen as 1 if quantities with negative or zero index are understood to be zero. This gives

$$L_W^{(k)} = \sum_s \frac{\partial \chi^{(k-\Gamma_s-\Delta_s)}}{\partial \bar{Y}_s} \frac{\partial}{\partial Y_s}. \tag{101}$$

Expanding a typical component of (96) gives

$$z_i = [1 + (\sum \epsilon^k L_W^{(k)}) + \tfrac{1}{2}(\sum \epsilon^k L^{(k)})^2 + \cdots] * y_i$$

$$= \sum \epsilon^k M^{(k)} * (\epsilon^{-\Gamma_i} Y_i)$$

$$= \sum \epsilon^k M^{(k+\Gamma_i)} * Y_i, \tag{102}$$

where the $M^{(k)}$ all have the form

$$M^{(k)} = L_W^{(k)} + N^{(k)}, \tag{103}$$

with $N^{(k)}$ some operator involving lower orders. One gets

$$\mathbf{z}_i^{(k)} = \zeta_i^{(k+\Delta_i)}$$

$$= M^{(k+\Gamma_i+\Delta_i)} * Y_i$$

$$= L_W^{(k+\Gamma_i+\Delta_i)} * Y_i + N^{(k+\Gamma_i+\Delta_i)} * Y_i$$

$$= \frac{\partial \chi^{(k)}}{\partial \bar{Y}_i} + N^{(k+\Gamma_i+\Delta_i)} * Y_i, \tag{104}$$

which again is the sum of a gradient in $\bar{\mathbf{Y}}$ space and an expression involving lower orders which (presumably) is a particular solution of (78).

THE KRYLOV-BOGOLIUBOV-KRUSKAL METHOD WITH SLOW VARIABLES

Krylov and Bogoliubov[9-11] investigated the solution of a set of n equations vectorially represented by

$$\frac{d\mathbf{y}}{dt} = \sum_{k=0} \epsilon^k \mathbf{g}^{(k)}(\mathbf{y}), \tag{105}$$

with

$$\mathbf{g}^{(0)} = (0, 0, \ldots, 0, g_n^{(0)}) \tag{106}$$

ensuring that in the "unperturbed" limit $\epsilon \to 0$, y_n alone varies and all other components of $\mathbf{y}$ (to be collectively denoted by $\bar{\mathbf{y}}$) are constant. It is further assumed that the unperturbed system is periodic and that y_n is an angle variable appearing only in the angle argument of periodic functions. The

zero-order growth of y_n is then assumed to be linear, from which follows that $g_n^{(0)}$ may depend on $\tilde{\mathbf{y}}$ but not on y_n.

To eliminate the periodicity from this motion, Krylov and Bogoliubov used a near-identity transformation to new variables $\mathbf{z}$, given in a direct form inverse to that of (37)

$$\mathbf{y} = \mathbf{z} + \sum \epsilon^k \boldsymbol{\eta}^{(k)}(\mathbf{z}). \tag{107}$$

The new variables, which can be derived by a suitable recursive method, have the property that the equations by which they evolve do not contain the transformed angle variable z_n on the right-hand side, but have the form

$$\frac{d\mathbf{z}}{dt} = \sum \epsilon^k \mathbf{h}^{(k)}(\tilde{\mathbf{z}}). \tag{108}$$

The first $(n-1)$ equations of this set, representing $d\tilde{\mathbf{z}}/dt$, then form an independent set not involving z_n which can be solved separately.

If $\mathbf{y}$ represents a perturbed periodic canonical system with a Hamiltonian of the form (50), then the canonical equations of motion have the form (105) and the Krylov–Bogoliubov method can be used to eliminate the angle variable $y_n = \Omega$. Unfortunately, unless precautions are taken,[22,23] the $\mathbf{z}$ variables will in general *not* be canonical, so that the transformed variable corresponding to the canonical conjugate of y_n will in general *not* be a constant of the perturbed motion, as is automatically achieved by the Poincaré–Von Zeipel method.

On the other hand, the Krylov–Bogoliubov method has a much wider validity and can be used in non-Hamiltonian systems. A similar elimination procedure, which derives the transformation in the form (37), has been devised by Kruskal,[8,12] who followed it by the derivation (for canonical systems only) of a constant J of the motion, obtained by an ingenious application of integral invariants (it is the same constant as is obtained by the Poincaré–Von Zeipel method).

Here the Krylov–Bogoliubov method will be generalized for the case when slow variables are present. As with the Poincaré–Von Zeipel method, this allows the restrictions on the form of the zero-order equations—embodied in $\mathbf{g}^{(0)}$—to be eased. Specifically, some variables other than y_n are now allowed to have a zero-order variation and this variation (as in the canonical method) is passed intact to the "reduced" equations involving $\mathbf{z}$. The calculation will be done for the transformation (107); the treatment of Kruskal's method, using (37), follows practically identical steps and will therefore be omitted.

Following the notation of (89), let $\mathbf{R}$ and $\mathbf{r}$ denote the slow and normal components of $\mathbf{z}$, and let $\boldsymbol{\eta}_R^{(k)}$ and $\boldsymbol{\eta}_r^{(k)}$ be corresponding components of $\boldsymbol{\eta}^{(k)}$. Substituting (107) in the left-hand side of (105) gives, with the definitions (67) and (68) extended to $\mathbf{z}$ variables,

$$\begin{aligned}\frac{d\mathbf{y}}{dt} &= \frac{d\mathbf{z}}{dt} + \sum_{k,s} \epsilon^k \frac{\partial \boldsymbol{\eta}^{(k)}}{\partial z_s}\frac{dz_s}{dt} \\ &= \frac{d\mathbf{z}}{dt} + \sum_{k,s} \epsilon^k \frac{\partial \boldsymbol{\eta}^{(k-\Gamma_s)}}{\partial Z_s} \sum_{m=0} \epsilon^m h_s^{(m)} \\ &= \frac{d\mathbf{z}}{dt} + \sum \epsilon^k \sum_{m=0}^{k-1} \frac{\partial \boldsymbol{\eta}^{(k-m-\Gamma_s)}}{\partial Z_s} h_s^{(m)}.\end{aligned} \tag{109}$$

Expressing a typical term of the right-hand side of (105) in term of $\mathbf{z}$ gives, in a manner similar to (88),

$$\begin{aligned}\mathbf{g}^{(k)}(\mathbf{y}) &= \mathbf{g}^{(k)}(\mathbf{z} + \sum \epsilon^m \boldsymbol{\eta}^{(m)}) \\ &= \exp\left(\sum_{m,s} \epsilon^m \boldsymbol{\eta}_s^{(m)} \cdot \frac{\partial}{\partial z_s}\right) * \mathbf{g}^{(k)}(\mathbf{z}) \\ &= \exp\left(\sum_{m,s} \epsilon^m \eta_s^{(m-\Gamma_s)} \cdot \frac{\partial}{\partial Z_s}\right) * \mathbf{g}^{(k)}(\mathbf{z}) \\ &= \sum \epsilon^m K^{(m)} * \mathbf{g}^{(k)}(\mathbf{z}).\end{aligned} \tag{110}$$

The operators $K^{(m)}$ resemble those of (89) but with $\boldsymbol{\eta}^{(m)}$ everywhere replacing $\boldsymbol{\pi}^{(m)}$. Substituting preceding results in (105) gives

$$\begin{aligned}\frac{d\mathbf{z}}{dt} = \sum_{k=0} \epsilon^k \Bigg[&\sum_{m=1}^{k-1} \left(K^{(m)} * \mathbf{g}^{(k-m)} - \sum_s \frac{\partial \boldsymbol{\eta}^{(k-m-\Gamma_s)}}{\partial Z_s} \cdot h_s^{(m)} \right) \\ &+ (K^{(k)} * \mathbf{g}^{(0)}) - \sum_s \frac{\partial \boldsymbol{\eta}^{(k-\Gamma_s)}}{\partial Z_s} \cdot g_s^{(0)} + (1 - \delta_{k0}) \mathbf{g}^{(k)} \Bigg],\end{aligned} \tag{111}$$

where the factor preceding the last term denotes that it be omitted for $k = 0$ (in that case it is already counted as the term involving $K^{(k)}$) and where in the summation preceding this term $\mathbf{h}^{(0)}$, has been replaced by $\mathbf{g}^{(0)}$, which equals it since in the limit $\epsilon \to 0$, Eqs. (105) and (108) coincide.

Comparison with (108) shows that the expression in the square brackets equals $\mathbf{h}^{(k)}(\tilde{\mathbf{z}})$, and this equality forms the basis of the recursive derivation of $\mathbf{h}^{(k)}$ and $\boldsymbol{\eta}^{(k)}$.

The situation now resembles that of (63): In order that the recursion be at all possible, unknown components of $\boldsymbol{\eta}^{(k)}$ must appear in (111) only once, otherwise the result is a partial differential equation and cannot be easily integrated. One term which always contains $\boldsymbol{\eta}^{(k)}$ is contributed by the last summation in (111) and equals

$$\frac{\partial \boldsymbol{\eta}^{(k)}}{\partial z_n} g_n^{(0)},$$

since $g_n^{(0)}$ does not vanish and z_n, the transformed angle variable, is normal. No other appearance is permitted; hence

$$g_s^{(0)} = 0 \quad \text{for} \quad \Gamma_s = 0 \tag{112}$$

or, stated in words, only slow variables and the principal angle variable are allowed to have a zero-order variation.

In addition, $\boldsymbol{\eta}^{(k)}$ could enter through the term containing $K^{(k)}$, which has the form [compare Eq. (89)]

$$K^{(k)}*\mathbf{g}^{(0)} = \left(\frac{\eta_r^{(k)}\cdot\partial}{\partial_r} + N^{(k)}\right)*\mathbf{g}^{(0)}, \tag{113}$$

with $N^{(k)}$ containing lower orders. No problem arises here provided $\boldsymbol{\eta}_r^{(k)}$ is derived first and $\boldsymbol{\eta}_R^{(k)}$ only afterwards: Because of (112), this term is absent in the first part of the derivation, while in the second part those components of $\boldsymbol{\eta}^{(k)}$ that appear in it are already known. In either case one gets

$$\frac{\partial\boldsymbol{\eta}^{(k)}}{\partial z_n} g_n^{(0)} + \mathbf{h}^{(k)}(\mathbf{z}) = \boldsymbol{\lambda}^{(k)}, \tag{114}$$

where $(k > 0)$

$$\boldsymbol{\lambda}^{(k)} = \sum_{m=1}^{k-1}\left(K^{(m)}*\mathbf{g}^{(k-m)} - \sum_s \frac{\partial\boldsymbol{\eta}^{(k-m_s-\Gamma_s)}}{\partial Z_s} h_s^{(m)}\right.$$

$$\left. + (K^{(k)}*\mathbf{g}^{(0)}) + \mathbf{g}^{(k)} - \sum_{s\neq n}\frac{\partial\boldsymbol{\eta}^{(k-\Gamma_s)}}{\partial Z_s} g_s^{(0)}\right) \tag{115}$$

depends only on lower orders. The solving of (114) then resembles that of (21).

CONCLUSION

In the preceding sections, the main methods of classical perturbation theory have been extended to slowly (or adiabatically) perturbed systems with a single zero-order periodicity. At the same time, the basic concepts associated with such systems (e.g., adiabatic invariance and implicit ϵ) were examined and clarified.

The restrictions on the forms of the zero-order equations for slow perturbations have been derived and are generally less severe than for small perturbations. The extension of the methods themselves is relatively straightforward, involving mainly the shifting of indices for quantities corresponding either to slow variables or (as in the case of $\mathcal{J}^{(k)}$) to variables with slow conjugates. With the use of expansion operators the treatment is only slightly more complicated than for small perturbations.

---

[1] This term has occasionally been applied (even by this author, in Ref. 12) to denote approximate constants of motion derived from $\oint \mathbf{p}\cdot d\mathbf{q}$. Actually such constants may belong to either of the types discussed here.

[2] G. Contopoulos, J. Math. Phys. **7**, 788 (1966).

[3] H. Poincaré, *Les Methodes Nouvelles de la Mecanique Celeste*, (Gauthier-Villars, Paris, 1893), Vol. II; reprint (Dover, New York, 1957); NASA translation TT F-451, 1967.

[4] H. Von Zeipel, Ark. Astr. Mat. Fys. **11**, No. 1 (1916).

[5] H. C. Corben and P. Stehle, *Classical Mechanics*, 2nd ed. (Wiley, New York, 1960), Sec. 75.

[6] D. Ter Haar, *Elements of Hamiltonian Mechanics* (North-Holland, Amsterdam, 1961), Chap. 7.

[7] P. Musen, J. Astr. Sci. **12**, 129 (1965).

[8] M. Kruskal, J. Math. Phys. **3**, 806 (1962).

[9] N. M. Krylov and N. N. Bogoliubov, *Introduction to Non-Linear Mechanics* (Kiev, 1936).

[10] N. N. Bogoliubov and D. N. Zubarev, Ukr. Math. J. **7**, 5 (1955); (English transl.: Space Technology Laboratories, 1960).

[11] N. N. Bogoliubov and Y. A. Mitropolski, *Asymptotic Methods in the Theory of Nonlinear Oscillations*, (Gordon and Breach, New York, 1961), Sec. 25.

[12] D. Stern, J. Math. Phys. **11**, 2771 (1970).

[13] J. E. Littlewood, Ann. Phys. (N.Y.) **21**, 223 (1963).

[14] M. Born, *The Mechanics of the Atom*, 2nd. Ed. (Frederick Ungar, 1960), Sec. 10. Also in Ref. 6, Sec. 6-3.

[15] C. Gardner, Phys. Rev. **115**, 791 (1959).

[16] May be omitted at first reading.

[17] D. Stern, J. Math. Phys. **11**, 2776 (1970).

[18] D. Stern, "Hamiltonian Formulation of Guiding Center Motion," Goddard Space Flight Center Publication No. X-641-71-56 (1971).

[19] S. Lie, *Theorie der Transformationgruppen* (Teubner, Leipzig, 1888), Vol I.

[20] G. Hori, Pub. Astron. Soc. Japan **18**, 287 (1966).

[21] A. Deprit, Celestial Mech. **1**, 12 (1969).

[22] E. L. Burshtein and L. S. Solov'ev, Dokl. Akad Nauk SSSR. **139**, 855 (1961) [Sov. Phys. Dokl., **6**, 731 (1962)].

[23] D. Stern, J. Math. Phys. **12**, 2226 (1971).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 10 OCTOBER 1971

On the Inverse Scattering Problem at Fixed Energy For the Tensor and Spin-Orbit Potentials

M. A. Hooshyar
Physics Department, Indiana University, Bloomington, Indiana 47401
(Received 26 April 1971)

The main purpose of this work is to find an analog of the Regge–Newton equation for the problem of finding tensor and spin–orbit potentials, acting among particles of spin $\frac{1}{2}$, from the knowledge of the S-matrix as a function of angular momentum at a fixed energy. We first find a transformation which makes the Schrödinger equation containing a central, a spin–orbit, and a tensor potential the matrix analog of the Schrödinger equation for spin-$\frac{1}{2}$ particle in a central and a spin–orbit potential. Next, guided by the work of Sabatier for the case of central and spin–orbit potentials, we are able to find the desired integral equation. The necessary existence and uniqueness of the solution to this integral equation is studied, and it is shown that indeed the wavefunctions for tensor and spin–orbit potentials can be represented in terms of the solution of this integral equation. A series representation of the wavefunctions, which are necessary for finding the potentials from the S matrix, is found, but the actual construction of the potentials from the S matrix is not considered in this work.

1. INTRODUCTION

Information about nuclear forces is mostly obtained through scattering experiments, but the quantities which are found experimentally do not give us the interparticle forces directly. The problem of finding the interaction from the experimental data is called the inverse scattering problem. Because of the complicated nature of this problem, it it usually broken up into two separate parts. The first part is concerned with the problem of finding the S matrix from the experimental results, that is, cross sections, polarizations, etc., and the second part deals with the determination of the forces involved from the knowledge of the S matrix. If we assume that the energy range which we are considering is such that nonrelativistic quantum mechanics is valid and the corresponding potentials are spherically symmetric, so that we can make use of partial wave analysis, then the problem of finding the interaction from the S matrix can be conveniently separated into two domains. The first is called "inverse problem at fixed angular momentum." It deals with construction of potentials from the knowledge of the S matrix, at a fixed angular momentum, as a function of energy. Because of the completeness and orthogonality of the radial wavefunctions of one angular momentum and all energies, and the existence of the Gel'fand–Levitan integral equation[1] and its analogs, there exist a great many results for the inverse scattering problem at fixed angular momentum when the potential is independent of energy. Indeed one such result, specially relevant to this work, is on the subject of constructing the tensor force from the S matrix at fixed angular momentum which was considered in detail by Newton.[2]

The second inverse scattering problem deals with construction of potentials from the knowledge of the S matrix at one energy as a function of angular momentum. Because of the lack of completeness and orthogonality of the radial wavefunction of one energy and all angular momenta, the inverse scattering problem at fixed energy for the case of spinless particles in a central potential resisted solution for some time until it was finally solved by Newton[3] and later on this solution was extended by Sabatier.[4] In lieu of already existing well-established results for constructing potentials from the S matrix at fixed angular momentum, one justifies the need to do inverse scattering analysis at fixed energy on physical grounds. That is, although the inverse scattering problem at fixed angular momentum is mathematically elegant, from the physical point of view, the assumptions made are not consistent. In other words, the assumption that the energies involved are such that nonrelativistic quantum mechanics is valid and the assumption that, for large values of distance, the nonrelativistic wavefunctions, the S matrix, are given even for energies as high as infinity are not compatible.

The important point to be noticed is that all the results mentioned above were possible because one could at least formally deduce the analog of the Gel'fand–Levitan equation for the problem at hand, without any difficulty. But this is not the case for the inverse scattering problem at fixed energy for two spin-$\frac{1}{2}$ particles in a tensor and a spin–orbit potential. This lack of existence of analog of the Gel'fand–Levitan, or equivalently of the Regge–Newton,[5,3] equation has been the major reason why this inverse problem has resisted solution. Indeed it is the main task of this work to find such an analog. As we shall see, the key for finding such an analog is the remarkable fact that the Schrödinger equation for the tensor force[6] problem can be made to satisfy a matrix analog of the differential equation satisfied by a particle of spin $\frac{1}{2}$ in a central and a spin–orbit field.[7] Having found this analog, then the work of Sabatier[8,9] for the inverse scattering problem for the case of spin–orbit and central potentials, gives us the necessary indications on how to go about finding the analog of the Gel'fand–Levitan equation for the case of two identical spin-$\frac{1}{2}$ particles in a central, spin–orbit, and tensor potential.

In Sec. 2 we first review the relevant parts of the inverse scattering problems for the case of central potentials and for the case of central and spin–orbit potentials. Next we introduce the transformation which makes the Schrödinger equation for central, spin–orbit, and tensor potentials the matrix analog of the Schrödinger equation for central and spin–orbit potentials. In that section we also introduce the relevant equations which we expect to be the formal analog of the equations found by Sabatier for the case of central and spin-orbit potentials. But since this analogy between the Schrödinger equation for cental, spin–orbit, and tensor potentials and the Schrödinger equation for only central and spin–orbit potentials is purely formal, this

extension of the work of Sabatier for central and spin–orbit potentials to the case of central, spin–orbit, and tensor potentials is not straightforward. Aside from matrix complication, indeed some of the functions needed in this formalism will not exist, and, therefore, the formal method must be modified and proofs changed. These necessary changes and proofs are given in the remaining sections. In Sec. 3 we consider analytic tensor and spin–orbit potentials and find analytic properties of the wavefunctions and prove that they can be used in representing other analytic functions. This fact will be essential in the necessary existence proofs. Section 4 deals with the proposed integral equation which we expect to be the analog of the Gel'fand–Levitan equation. We prove the uniqueness and existence of the solution to this integral equation in Appendices A, B, and C. In Sec. 4, we also show that the regular solution of the tensor and spin–orbit differential equation can be represented in terms of the solution to the mentioned integral equation. The generalization of Gell-Mann "nonsense term,"[10] which is introduced by our new transformation of the Schrödinger equation, is amusing but it is also a source of some difficulty in this work. Section 5 is concerned with a series representation of the wavefunction, which is needed when one tries to connect the S matrix to the kernel of the above mentioned integral equation. Of course, knowing this kernel, we then are able to find the corresponding potentials. In this work we will not be directly concerned with the problem of constructing the tensor and spin–orbit potentials from the knowledge of the S matrix, but in Sec. 6 we will deal with the problem of constructing the potentials when this mentioned kernel is already given. An example is also given to demonstrate the construction of potentials from a given kernel.

Let us again point out that the inverse problem which we have in mind is for the case of two identical particles of spin $\frac{1}{2}$ whose interaction depends only on their positions and spins. A general form of such an interaction is[6]

$$V(\mathbf{r}) = V_c(r) + V_\sigma(r)\sigma_1\cdot\sigma_2 + V_t(r)\cdot S_{12}, \qquad (1.1)$$

where $S_{12} = 3\sigma_1\cdot\hat{r}\sigma_2\cdot\hat{r} - \sigma_1\cdot\sigma_2$, $\frac{1}{2}\sigma_1$ and $\frac{1}{2}\sigma_2$ are the spin operators $\hat{r}$ is the unit vector in the direction of $\mathbf{r}$. Use of partial wave analysis decomposes the potential matrix into three blocks,[11] two of which correspond to the singlet state and the triplet state of parity $(-)^j$. The last block corresponds to the triplet state of parity $(-)^{j+1}$. The Schrödinger equation for the latter can be written as[12]

$$-r^2\frac{d^2}{dr^2}\psi_j + L(L+1)\psi_j + r^2V_j(r)\psi_j = r^2k^2\psi_j, \qquad (1.2)$$

where

$$V_j(r) = \frac{1}{2j+1}\times\begin{bmatrix}(2j+1)V_d - 2(j-1)V_t & 6[j(j+1)]^{1/2}V_t\\ 6[j(j+1)]^{1/2}V_t & (2j+1)V_d - 2(j+1)V_t\end{bmatrix},$$

$$V_d = V_c + V_\sigma, \qquad L = \begin{bmatrix}j-1 & 0\\ 0 & j+1\end{bmatrix}$$

and ψ_j is a 2×2 matrix whose first row indicates the components of ψ_j belonging to the orbital angular momentum $j-1$ for each of the two possible boundary conditions, and likewise the second row of ψ_j corresponds to orbital angular momentum $j+1$. The boundary condition of the first column of ψ_j is that, at time $t\to-\infty$, the orbital angular momentum of the particles is $j-1$. For the second column the initial orbital angular momentum is $j+1$.

Since no restriction on the S matrix is known which assures that it corresponds to the potential given in (1.1), one also needs to introduce the spin–orbit force in the above interaction. The spin–orbit potential is found to make the following contribution to the potential matrix[12] for the case of triplet state of parity $(-)^{j+1}$:

$$V_{LS}^j(r) = \begin{bmatrix}j-1 & 0\\ 0 & -j-2\end{bmatrix}V_0(r). \qquad (1.3)$$

The form of differential equations (1.2) can be considerably simplified if we subject the solutions of (1.2) to the transformation[13]

$$\Xi_j = \begin{bmatrix}(j+1)^{1/2} & j^{1/2}\\ j^{1/2} & -(j+1)^{1/2}\end{bmatrix}. \qquad (1.4)$$

Applying transformation (1.4) to the Schrödinger equation (1.2), one obtains

$$-r^2\frac{d^2}{dr^2}\bar{\psi}_\lambda + r^2\bar{W}(r)\bar{\psi}_\lambda + D(\lambda)\bar{\psi}_\lambda = r^2k^2\bar{\psi}_\lambda, \qquad (1.5)$$

where

$$\bar{W}(r) = \begin{bmatrix}V_d(r)+2V_t(r) & 0\\ 0 & V_d(r)-4V_t(r)\end{bmatrix},$$

$$\psi_\lambda = \Xi_{\lambda-1/2}\psi_{\lambda-1/2}, \qquad \lambda = j+\tfrac{1}{2},$$

and

$$D(\lambda) = \begin{bmatrix}\lambda^2-\frac{1}{4} & -2(\lambda^2-\frac{1}{4})^{1/2}\\ -2(\lambda^2-\frac{1}{4})^{1/2} & \lambda^2+7/4\end{bmatrix}. \qquad (1.6)$$

Under transformation (1.4), the spin–orbit force (1.3) takes the form

$$\Xi_jV_{LS}^j(r)\Xi_j^{-1} = \begin{bmatrix}-1 & (\lambda^2-\frac{1}{4})^{1/2}\\ (\lambda^2-\frac{1}{4})^{1/2} & -2\end{bmatrix}V_0(r). \qquad (1.7)$$

The inverse scattering problem at fixed energy then reduces to the question of whether it is possible to find the potentials V_d, V_t, and V_0 given the S matrix for all angular momenta, and, if it is possible, what is the procedure to be followed in order to construct these three spin-independent potentials from the information on the S matrix?

2. A REVIEW OF PREVIOUS RESULTS AND A REDUCTION OF THE PROBLEM

The main tool in the construction of central potentials from the S matrix at fixed energy is the Regge–Newton equation. That is, from the information on the asymptotic behavior of the regular solutions one is able to define an integral equation whose solution is directly related to the wavefunction and the central potential of the scattering problem. More specifically let us consider the inverse problem at fixed energy for a spinless particle in a central potential $\tilde{u}(r)$. The wavefunction $\tilde{v}_\lambda(r)$ for this particle satisfies the following equation:

$$r^2\frac{d^2}{dr^2}\tilde{v}_\lambda(r) + r^2[1-\tilde{u}(r)]\tilde{v}_\lambda(r) = (\lambda^2 - \tfrac{1}{4})\tilde{v}_\lambda(r), \tag{2.1}$$

where we have chosen units in which $k^2 = 1$, $\lambda = l + \frac{1}{2}$, and $\tilde{v}_\lambda(r)$ is assumed to be the regular wavefunction satisfying the boundary condition

$$\lim_{r\to 0}\ (r/2)^{-\lambda}\Gamma(1+\lambda)(\tfrac{1}{2}\pi r)^{-1/2}\tilde{v}_\lambda(r) = 1. \tag{2.2}$$

If we now let $v_\lambda(r)$ be the regular solution to Eq. (2.1) when $\tilde{u}(r)$ is replaced by $u(r)$, then, following the method constructed by Newton[3] and its generalization by Sabatier,[4] we define the following input function:

$$f(r,r') = \sum_{\lambda\in\Omega} v_\lambda(r)v_\lambda(r')c_\lambda, \tag{2.3}$$

where the set Ω depends on the nature of the potentials being considered and the coefficients c_λ are left arbitrary for the moment and are to be found later from the information on the asymptotic behavior of $\tilde{v}_\lambda(r)$. Next the Regge–Newton equation is defined:

$$k(r,r') = f(r,r') - \int_0^r d\rho\rho^{-2}k(r,\rho)f(\rho,r'). \tag{2.4}$$

Then it can be shown that $\tilde{v}_\lambda(r)$, the regular solution to (2.1), satisfies the following equation,

$$\tilde{v}_\lambda(r) = v_\lambda(r) - \int_0^r d\rho\rho^{-2}k(r,\rho)v_\lambda(\rho), \qquad \lambda \geq 0, \tag{2.5}$$

if the coefficients c_λ are chosen in such a way that

$$\tilde{u}(r) - u(r) = -2r^{-1}\frac{d}{dr}\{r^{-1}k(r,r)\}. \tag{2.6}$$

Since we are interested in the inverse problem, in other words, since $\tilde{u}(r)$ is our unknown function and the asymptotic behavior of the $\tilde{v}_\lambda(r)$ are the only things which are specified, therefore Eq. (2.6) can be viewed as the definition of $\tilde{u}(r)$ if the coefficients c_λ are chosen in such a way that the right-hand side of Eq. (2.5) has the proper asymptotic behavior.

At this point we should realize that the question of finding the potential $\tilde{u}(r)$ from the S matrix is now reduced to finding a set of coefficients, which are dependent on our choice of the comparison potential $u(r)$. In other words, having chosen a comparison potential $u(r)$ and a set of coefficients c_λ that make Eq. (2.4) have an acceptable solution, Eq. (2.5) gives the regular solution to the differential equation (2.1) in which the potential $\tilde{u}(r)$ is defined by Eq. (2.6). The only condition which is then to be realized is that the set $\{c_\lambda\}$ should have been chosen in such a way that the regular solution given by Eq. (2.5) has the desired asymptotic behavior. In order to satisfy the asymptotic condition on the regular solution defined by Eq. (2.5), one needs to note that the substitution of (2.3) in (2.4) and use of (2.5) implies that

$$k(r,r') = \sum_{\lambda\in\Omega}\tilde{v}_\lambda(r)v_\lambda(r')c_\lambda. \tag{2.7}$$

Next, substitution of (2.7) in (2.5) will give us a representation of the regular solutions which is essential for relating the coefficients c_λ to the S matrix,

$$\tilde{v}_\lambda(r) = v_\lambda(r) - \sum_{\alpha\in\Omega}\tilde{v}_\alpha(r)c_\alpha\int_0^r d\rho\rho^{-2}v_\alpha(\rho)v_\lambda(\rho). \tag{2.8}$$

It is shown in Ref. 3 that if the comparison potential is suitably chosen, in the limit as $r\to\infty$, Eq. (2.8) gives the coefficients c_λ in terms of the phase shifts. Clearly, the coefficients c_λ having been found from the asymptotic form of Eq. (2.8), then the regular solutions defined by (2.5) will have the desired asymptotic behavior and therefore the potential $\tilde{u}(r)$ defined by (2.6) is the desired potential.

For the sake of completeness let us also note that it was found by Sabatier[4,8] that when the potentials under consideration are analytic, then the set Ω in Eq. (2.3) can contain only the positive integers and half-integers and, for example, for the case when $\tilde{u}(r) = 1$ and $u(r) = 0$, the coefficients c_λ are given as[8]

$$\begin{aligned} c_\lambda &= -2\lambda/\pi, && \text{for positive integers,}\\ &= 0 && \text{otherwise.}\end{aligned} \tag{2.9}$$

From Ref. 8 we also note that, given the potentials $\tilde{u}(r), u(r)$ and the coefficients c_λ such that Eq. (2.6) holds, $v_\lambda(r)$ can be represented in terms of $\tilde{v}_\lambda(r)$ in a form similar to Eq. (2.5):

$$v_\lambda(r) = \tilde{v}_\lambda(r) + \int_0^r d\rho\rho^{-2}k(\rho,r)\tilde{v}_\lambda(\rho), \qquad \lambda \geq 0. \tag{2.10}$$

Having seen the basic approach to the inverse problem for spinless particles in central potentials,

let us now briefly review the work of Sabatier[9] for the case of scattering of a spin-$\frac{1}{2}$ particle by a central and a spin–orbit potential.

It is shown in Ref. 9 that the differential equations of interest are

$$r^2\frac{d^2}{dr^2}\varphi_\lambda^+ + \{r^2U(r) - 2\lambda r^2Q(r)\}\varphi_\lambda^+ = (\lambda^2 - \tfrac{1}{4})\varphi_\lambda^+, \tag{2.11a}$$

$$r^2\frac{d^2}{dr^2}\varphi_\lambda^- + \{r^2U(r) + 2\lambda r^2Q(r)\}\varphi_\lambda^- = (\lambda^2 - \tfrac{1}{4})\varphi_\lambda^-, \tag{2.11b}$$

where φ_λ^+ and φ_λ^- are the regular solutions to (2.11) and $U(r)$ and $Q(r)$ are related to the central and the spin–orbit potentials. Again one tries to reduce the problem of finding the potentials from the S matrix to the problem of finding a set of coefficients from which the potentials and the wavefunctions can be easily defined. In Ref. 9, this is accomplished by introducing the following input functions:

$$f^\pm(r,r') = \sum_{\lambda\epsilon S} s_\lambda(r)s_\lambda(r')c_\lambda^\pm, \tag{2.12}$$

where $c_\lambda^\pm$ are the set of coefficients which are to be found from the S matrix, the set S contains positive integers and half-integers if the potentials under consideration are analytic, and $s_\lambda(r)$ is the regular solution to Eq. (2.1) for the case when $\tilde{u}(r) = 1$. In other words

$$s_\lambda(r) = (\tfrac{1}{2}\pi r)^{1/2}(r/2)^\lambda\{\Gamma(1+\lambda)\}^{-1}. \tag{2.13}$$

Next the analog of the Regge–Newton equation is defined as

$$k^\pm(r,r') = F^\mp(r)f^\pm(r,r') - \int_0^r d\rho\rho^{-2}k^\mp(r,\rho)f^\pm(\rho,r') \tag{2.14}$$

with

$$F^\pm(r) = \exp\{\pm\textstyle\int_0^r d\rho\rho Q(\rho)\}. \tag{2.14'}$$

$k^\pm(r,r')$ having been defined, it is again possible to show that the regular solutions to Eqs. (2.11) satisfy the following equations:

$$\varphi_\lambda^\pm(r) = F^\pm(r)s_\lambda(r) - \int_0^r d\rho\rho^{-2}k^\pm(r,\rho)s_\lambda(\rho), \tag{2.15}$$

if the coefficients $c_\lambda^\pm$ are chosen in such a way that[14]

$$2\dot{k}^\pm(r,r) - 2r^{-1}[1 \pm r^2Q(r)]k^\pm(r,r) = \{r^2U(r) \pm 2r^2Q(r) \pm r^3\dot{Q}(r) + r^4Q^2(r)\}F^\pm(r). \tag{2.16}$$

Again we note that from the inverse scattering point of view, Eqs. (2.16) are to be used as the definition of our unknown functions $U(r)$ and $Q(r)$ and the coefficients $c_\lambda^\pm$ are to be chosen in such a way that the right-hand sides of Eqs. (2.15) have the desired asymptotic behavior. Although the analogs of Eqs. (2.4) and (2.6), that is, Eqs. (2.14) and (2.16), are now coupled, it is still possible[9] to find $U(r)$, $Q(r)$, and $\varphi_\lambda^\pm$ if the coefficients $c_\lambda^\pm$ are given, and therefore the problem is again reduced to finding the right set of coefficients $c_\lambda^\pm$ which gives the desired asymptotic behavior of the regular solutions.

For the inverse scattering problem of two identical particles of spin $\frac{1}{2}$ in a central, a spin–orbit, and a tensor potential, we also would like to reduce the problem in the above manner. But, having seen the complication which arises when one tries to introduce even the spin–orbit potential alone in the Schrödinger equation, we first would like to make Eqs. (1.5) as similar to Eqs. (2.11) as possible. It is remarkable that the following simple transformation makes the Schrödinger equation containing a tensor force become the matrix analog of the Schrödinger equation containing only a spin–orbit force:

$$T_\lambda = \begin{bmatrix}(\lambda+\frac{1}{2})^{1/2} & 0\\ 0 & (\lambda-\frac{1}{2})^{1/2}\end{bmatrix}. \tag{2.17}$$

Subjecting solutions of Eqs. (1.5), when the spin–orbit force is also present, to T_λ for values of $\lambda \neq \frac{1}{2}$, we obtain

$$r^2\frac{d^2}{dr^2}\Phi_\lambda + \{-r^2W(r) - 1 - N(1+M)[1 - r^2V(r)] + 2\lambda M[1 - r^2V(r)]\}\Phi_\lambda = (\lambda^2 - \tfrac{1}{4})\Phi_\lambda, \tag{2.18}$$

where

$$W(r) = -\mathbb{1} + \bar{W}(r) - \tfrac{3}{2}V_0(r)\mathbb{1}, \quad N = \begin{bmatrix}-1 & 0\\ 0 & 1\end{bmatrix},$$

$$M = \begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}, \quad V(r) = \tfrac{1}{2}V_0(r)$$

and $\Phi_\lambda = T_\lambda\bar{\psi}_\lambda$.

Special attention is needed for the case of $\lambda = \frac{1}{2}$, because for that value of the angular momentum the transformation matrix (2.17) is singular and therefore there is no reason why Eqs. (2.18) should be related to (1.5). Of course, at $\lambda = \frac{1}{2}$ even the Eqs. (1.5) deserve special consideration,[15] because at $\lambda = \frac{1}{2}$ the differential equations (1.5) are uncoupled and the elements[16] ${}^{11}\bar{\psi}_{1/2}$ and ${}^{12}\bar{\psi}_{1/2}$ correspond to states with the orbital angular momentum -1 and ${}^{21}\bar{\psi}_{1/2}$ and ${}^{22}\bar{\psi}_{1/2}$ correspond to states with the orbital angular momentum 1. If we now remember the definition of ψ at time $t \to -\infty$, we note that the only term in $\bar{\psi}_{1/2}$ which has a physical meaning is the function ${}^{22}\bar{\psi}_{1/2}$ and, as far as the physics is concerned, we lose nothing by assuming the other elements of $\bar{\psi}_{1/2}$ are identically zero. On the other hand, we can certainly take the nonphysical elements of $\bar{\psi}_{1/2}$ not to be identically zero, as long as we accept them only as mathematical functions and the defined $\psi_{1/2}$ is a solution to (1.5). At this point

let us note that at $\lambda = \frac{1}{2}$ the element ${}^{22}\Phi_{1/2}$ of $\Phi_{1/2}$ satisfies the same differential equation as the one satisfied by ${}^{22}\bar{\psi}_{1/2}$, and, therefore, they are identical if we choose the same boundary condition for them at $r = 0$. It follows that $\Phi_{1/2}$ and $\bar{\psi}_{1/2}$ are physically equivalent and the fact that they can differ in their nonphysical elements is immaterial. Therefore, in this work, Eqs. (2.18) are considered as the Schrödinger equations for the interaction containing the tensor force and the spin–orbit force even at $\lambda = \frac{1}{2}$.

Inspection of (2.18) reveals that it is the matrix analog of (2.11b), where $-W(r) - 1/r^2 - [N(1+M)/r^2][1 - r^2V(r)]$ is the matrix analog of $U(r)$, $(M/r^2)(1 - r^2V)$ is the analog of $Q(r)$, and Φ_λ is the analog of φ_λ^-. We can make the analogy complete, if we introduce the matrix function Φ_λ^+ which is a solution to the differential equations obtained by changing the value of λ to $-\lambda$ in Eqs. (2.18). At this point we would like to point out that, although we have found a matrix analog to differential equations (2.11), there is no reason to believe that the method developed by Sabatier for representation of solutions to Eqs. (2.11) should generalize to the case when these equations include potentials which are not multiples of the identity matrix, because, besides the possible commutation difficulties, the proposed analogs to potentials $U(r)$ and $Q(r)$ are such that, at $r = 0$, the functions $r^2U(r)$ and $r^2Q(r)$ need to be different from zero and this is exactly the opposite of the assumption made for the case of only central and spin–orbital potentials.

Disregarding the mathematical questions for the moment, if we apply the method Sabatier developed for the case of central and spin–orbit potential to Eqs. (2.18) and modify the method in such a way as to circumvent some mathematical difficulties,[17] we expect the analog of Sabatier's equations for the case of central, spin–orbit, and tensor potentials to be as follows:

$$\mathbf{K}_-(r,r') = \mathbf{R}_- f_-(r,r') - \int_0^r d\rho\rho^{-2}\mathbf{K}_-(r,\rho)f_-(\rho,r'), \tag{2.19}$$

$$\mathbf{K}_+(r,r') = \mathbf{R}_+ f_+(r,r') - \int_0^r d\rho\rho^{-2}\mathbf{K}_+(r,\rho)f_+(\rho,r'),$$

with

$$f_\pm(r,r') = \sum_{\lambda\epsilon S_1}(-\tfrac{1}{2} \pm \lambda)\gamma_\lambda(rr')^{\lambda+1/2},$$

$$\mathbf{R}_\pm = \tfrac{1}{2}(rg_- \mp r^{-1}g_+, rg_- \pm r^{-1}g_+),$$

$$g_\pm = \exp\{\pm \int_0^r d\rho\rho V(\rho)\},$$

$S_1 = \{1, 3/2, 2, \cdots\}$, and for now γ_λ are a set of arbitrary constant 2×2 matrices. Next let us use the row vectors $\mathbf{K}_-(r,r') = ({}^{11}K(r,r'), {}^{12}K(r,r'))$ and $\mathbf{K}_+(r,r') = ({}^{21}K(r,r'), {}^{22}K(r,r'))$ to define the 2×2 matrix

$$K(r,r') = \begin{bmatrix}\mathbf{K}_-(r,r')\\ \mathbf{K}_+(r,r')\end{bmatrix}. \tag{2.20}$$

Having defined $K(r,r')$ by analogy with the work of Sabatier, if the γ_λ are chosen in such a way that

$$2\dot{K}(r,r) - 2r^{-1}\{1 + M[1 - r^2V(r)]\}K(r,r) = \{-N(1+M)[1 - r^2V(r)] - r^2[W(r) + 2V(r)1] - 2r^2V(r)M - r^3\dot{V}(r)M + r^4V^2(r)1\}R(r), \tag{2.21}$$

where

$$R(r) = \begin{bmatrix}\mathbf{R}_-\\ \mathbf{R}_+\end{bmatrix}, \tag{2.22}$$

then we expect that the regular solutions to Eqs. (2.18) have the following representation:

$$\Phi_\lambda = T'_\lambda\{R(r)r^{\lambda+1/2} - \int_0^r dr'r'^{-2}K(r,r')r'^{\lambda+1/2}\}, \tag{2.23}$$

with

$$T'_\lambda = (-\tfrac{1}{2}1 + \lambda N). \tag{2.24}$$

It is our main task in the remaining sections to show that the functions Φ_λ and $K(r,r')$ introduced in this section are indeed well defined and that Φ_λ is a solution to the differential equations (2.18) and, therefore, the inverse scattering problem for tensor and spin–orbit potentials can be completely solved if one is able to find the set $\{\gamma_\lambda\}$ such that the functions Φ_λ of (2.23) have the desired asymptotic behavior.

3. ANALYTIC PROPERTIES OF THE WAVE-FUNCTIONS

From the point of view of the inverse scattering problem, Eqs. (2.21) indicate that not any arbitrary $K(r,r')$ defined through Eqs. (2.19) is acceptable, because, in derivation of the equations in Sec. 2, it was necessary to assume that $W(r)$ is diagonal and $V(r)$ is a multiple of the identity. In other words, the set $\{\gamma_\lambda\}$ must be such that not only Eqs. (2.19) have a solution, but also this solution, that is, $K(r,r)$, must be such that when it is substituted in Eqs. (2.21) it will give us a diagonal potential $W(r)$ and a potential $V(r)$ that is a multiple of the identity. Indeed, it must be shown that such a set of $\{\gamma_\lambda\}$ exists.

A step toward this proof is to assume that the potentials are analytic and to see whether there can exist a set $\{\gamma_\lambda\}$ such that Eqs. (2.21) are satisfied if we replace $K(r,r)$ by

$$\sum_{\lambda\epsilon S}\tilde{\Phi}_\lambda\gamma_\lambda r^{\lambda+1/2},$$

where $\tilde{\Phi}_\lambda$ is the regular solution to (2.18), with $W(r)$ a diagonal potential and $V(r)$ a multiple of the identity, and $S = \{\frac{1}{2}, 1, \frac{3}{2}, \cdots\}$. In here we are not assuming that $\gamma_{1/2} = 0$, because we would like to find out what condition the potentials must satisfy so that $\gamma_{1/2}$ can be put equal to zero. Also, understanding of any condition that (2.21) may put on

$\gamma_{1/2}$ will be of use in extending the class of potentials with which we can deal.

To study the analytic properties of $\tilde{\Phi}_\lambda$, the regular solution to (2.18), we need to define a boundary condition on $\tilde{\Phi}_\lambda$ at $r = 0$. In the absence of any well-established boundary condition on the regular solution to the Schrödinger equations when a tensor force $V_t(r)$ is present and in the desire that the regular solution to (2.18) be representable in a form given by (2.23), we find that it is convenient to define the boundary condition through the following function for the case when $V_t = r^2 V_d = r^2 V_0 = 0$ at $r = 0$:

$$\tilde{\Omega}_n = \tfrac{1}{2}\{(1 + M)z^{-1} + (1 - M)z\}z^{-(n+1)/2} \times [(2 + 2nN)/(n^2 - 1)]\tilde{\Phi}_{n/2} \quad \text{for } n > 1. \tag{3.1}$$

$\tilde{\Omega}_n(z)$ satisfies the differential equation

$$z^2 \frac{d^2}{dz^2}\tilde{\Omega}_n(z) + (2M + n + 1)z\frac{d}{dz}\tilde{\Omega}_n(z) = W_n(z)\tilde{\Omega}_n, \tag{3.2}$$

where $W_n(z) = z^2\{W_+(z)1 - nV(z)M\} + \frac{1}{2}W_-(z) \times (1 + M)N + z^2\{1 - z^2V(z) + (z^2/2)W_-(z)\}N(1 + M)$, $W(z) = W_+(z)1 + W_-(z)N$ and W_+ and W_- are multiples of the identity.

If we assume the potentials are analytic in a circle of radius $R_0 > 0$, $|z| < R_0$, then we can represent them as

$$z^2W_+(z) = \sum_{m=1}^{\infty} a_m z^m, \quad z^2V(z) = \sum_{m=1}^{\infty} f_m z^m,$$
$$W_-(z) = \sum_{m=1}^{\infty} b_m z^m, \quad \text{and} \quad z^2W_n(z) = \sum_{m=1}^{\infty} A_{n,m}z^m. \tag{3.3}$$

Now the boundary condition on $\tilde{\Phi}_\lambda$ in terms of $\tilde{\Omega}_{2\lambda}$ can be represented by the following conditions:

$$\tilde{\Omega}_n(0) = C_{n,0} \equiv 1 - [(1 + M)/(n + 2)] \cdot\{1 + \tfrac{1}{2}(a_2 - a_1^2)N\} \quad \text{for} \quad n > 1. \tag{3.4}$$

The exact value of $\tilde{\Omega}_n$ at $z = 0$ given by (3.4) was motivated by (2.23).

Having defined the boundary condition for the solution to differential equations (3.2), we now would like to show that it can be represented by the following relations[18]:

$$\tilde{\Omega}_n(z) = \sum_{m=0}^{\infty} C_{n,m}z^m, \tag{3.5}$$

where

$$C_{n,m} = \frac{m + n - 2M}{m(m + 2 + n)(m - 2 + n)}\sum_{p=0}^{m-1} A_{n,m-p}C_{n,p}, \quad n > 1,$$

and $C_{n,0}$ is given by (3.4).

To prove that the above representation is valid, let us define $\|A\| = \max_{i,j=1,2}\|^{ij}A\|$. Clearly there exists a number $b(R)$ such that $\|(d/dz)W_n(z)\| < (1 + n)b(R)$ for $|z| \le R < R_0$. Consequently, $\|A_{n,m+1}\| \le (1 + n)b(R)/R^m$ and an upper bound for $C_{n,m}$ can be defined through the equation

$$d_{n,m+1} = \frac{2(1 + n)b(R)}{(m + 1)(m + n - 1)}\{d_{n,m} + d_{n,m-1}R^{-1} + \cdots d_{n,0}R^{-m}\}, \quad \text{with } d_{n,0} = \|C_{n,0}\|. \tag{3.6}$$

The so-defined numbers are clearly larger than the corresponding $\|C_{n,m}\|$, and they can be rewritten as

$$d_{n,m+1} = \left(\prod_{p=0}^{p=m} \frac{2(n + 1)b(R) + R^{-1}p(p + n - 2)}{(1 + p)(n + p - 1)}\right)d_{n,0}. \tag{3.7}$$

Following the method used in Ref. 9, we find that

$$d_{n,m+1} \le \left(\frac{2}{R}\right)^m \tilde{e}_n, \quad \tilde{e}_n = \max_{\zeta=0,\cdots,p_0}\left(\prod_{p=0}^{p=\zeta}\frac{2Rb(R)}{(1 + p)}\right)d_{n,0}, \tag{3.8}$$

where $p_0 = \max\{p; p < 2Rb(R)$ and p is an integer$\}$.

Therefore $\tilde{\Omega}_n$, $n > 1$, is analytic in a circle of radius $R/2$.

Calling $\tilde{\Omega}_1 = \tilde{\Phi}_{1/2}$, we find it is also analytic in the circle of radius $R/2$, if $a_1 + f_1 = 0$. In this case the differential equations satisfied by elements of $\tilde{\Omega}_1$ are given as

$$z^2\frac{d^2}{dz^2}{}^{11}\tilde{\Omega}_1 - z^2\{W_+(z) - W_-(z) + V(z)\}{}^{11}\tilde{\Omega}_1 = 0, \tag{3.9a}$$

$$z^2\frac{d^2}{dz^2}{}^{22}\tilde{\Omega}_1 - z^2\{W_+(z) + W_-(z) - V(z)\}{}^{22}\tilde{\Omega}_1 = 2\,{}^{22}\tilde{\Omega}_1, \tag{3.9b}$$

$$z^2\frac{d^2}{dz^2}{}^{12}\tilde{\Omega}_1 - z^2\{W_+(z) - W_-(z) + V(z)\}{}^{12}\tilde{\Omega}_1 = -2\{1 - z^2V(z)\}{}^{22}\tilde{\Omega}_1, \tag{3.9c}$$

where we have assumed ${}^{21}\tilde{\Omega}_1$, which satisfies the differential equation (3.9b), is identically zero. The values of ${}^{11}\tilde{\Omega}_1$, ${}^{12}\tilde{\Omega}_1$ and of their derivatives at $z = 0$ are assumed to be ${}^{11}C_{1,0}$, ${}^{12}C_{1,0}$, ${}^{11}C_{1,1}$ and ${}^{12}C_{1,1}$, which at present, except for the assumption that ${}^{11}C_{1,0} \neq 0$, are arbitrary numbers. $z^{-2}\,{}^{22}\tilde{\Omega}_1(z)$ is assumed to have the value of $\frac{2}{3}$ at $z = 0$. Needless to say, these boundary conditions on $\tilde{\Omega}_1$ are again motivated by (2.23). The solutions to (3.9) will have the following representations:

$$
{}^{11}\tilde{\Omega}_1(z)=\sum_{m=0}^{\infty}{}^{11}C_{1,m}z^m \quad \text{with} \quad {}^{11}C_{1,m}=\frac{1}{m(m-1)}\sum_{m=0}^{m-2}A^-_{m-p}\,{}^{11}C_{1,p} \quad \text{for} \quad m\geq 2,
$$

$$
{}^{22}\tilde{\Omega}_1(z)=\sum_{m=0}^{\infty}{}^{22}C_{1,m}z^m \quad \text{with} \quad {}^{22}C_{1,m}=\frac{1}{(m-2)(m+1)}\sum_{p=0}^{m-1}A^+_{m-p}\,{}^{22}C_{1,p} \quad \text{for} \quad m\geq 3, \tag{3.10}
$$

$$
{}^{12}\tilde{\Omega}_1(z)=\sum_{m=0}^{\infty}{}^{12}C_{1,m}z^m \quad \text{with} \quad {}^{12}C_{1,m}=\frac{1}{m(m-1)}\left\{\sum_{p=0}^{m-2}A^-_{m-p}\,{}^{12}C_{1,p}-2\sum_{m=0}^{m}F_{m-p}\,{}^{22}C_{1,p}\right\} \quad \text{for} \quad m\geq 2,
$$

where

$$
{}^{22}C_{1,0}={}^{22}C_{1,1}=0, \quad {}^{22}C_{1,2}=\tfrac{2}{3}, \quad F_0=1,
$$
$$
F_n=-f_n \quad \text{for} \quad n\geq 1,
$$

$$
A^{\pm}_1=a_1\mp f_1, \quad A^{\pm}_2=a_2\mp f_2 \quad \text{and} \quad A^{\pm}_n=a_n\mp f_n\pm b_{n-2}
$$
$$
\text{for} \quad n\geq 3.
$$

We would like to point out that ${}^{22}\tilde{\Omega}_1$ is usually referred to as the "sense" solution and that it is customary to give the name "nonsense" solution to ${}^{11}\tilde{\Omega}_1$. By analogy, ${}^{12}\tilde{\Omega}_1$ is a generalization of a "nonsense" solution. Of course, in order to define ${}^{11}\tilde{\Omega}_1$ and ${}^{12}\tilde{\Omega}_1$ uniquely, Eqs. (3.10) inform us that we must specify ${}^{11}C_{1,0}$, ${}^{12}C_{1,0}$, ${}^{11}C_{1,1}$ and ${}^{12}C_{1,1}$.

Equation (3.10) enables us to give a power series solution for $\tilde{\Omega}_1$:

$$
\tilde{\Omega}_1(z)=\sum_{m=0}^{\infty}C_{1,m}z^m \quad \text{where} \quad C_{1,m}=\begin{bmatrix}{}^{11}C_{1,m} & {}^{12}C_{1,m}\\ {}^{12}C_{1,m} & {}^{22}C_{1,m}\end{bmatrix}. \tag{3.11}
$$

Having seen the analyticity properties of the solution to (2.18), let us see whether the desired expansion is possible, that is, find the set $\{\gamma_\lambda\}$ which satisfies the following equations:

$$
\begin{aligned}
&\tfrac{1}{2}\{(1+M)z^{-2}+(1-M)\}\sum_{\lambda\in S}\tilde{\Phi}_\lambda\gamma_\lambda z^{\lambda+1/2}\\
&\quad\overset{?}{=}\sum_{n=0}^{\infty}B_nz^n\equiv g_-B_0+\tfrac{1}{4}g_-(1+M)\int_0^z dz'z'^{-1}\\
&\quad\times\{-z'^2(W_++2V)\mathbb{1}-z'M\frac{d}{dz'}(z'^2V)+z'^4V^2\mathbb{1}\\
&\quad-W_-g_+^2N\}+\tfrac{1}{4}g_+(1-M)\int_0^z dz'z'^{-1}\\
&\quad\times\Big(-z'^2(W_++2V)\mathbb{1}-z'M\frac{d}{dz'}(z'^2V)\\
&\quad+z'^4V^2\mathbb{1}-2(1-z'^2V+\frac{z'^2}{2}W_-)z'^2g_-^2N\Big).
\end{aligned}
\tag{3.12}
$$

Equations (3.12) are nothing but Eqs. (2.21) written in integral form. In this equation B_0 is not given, but is to be found. Since we would like $K(0,0)=0$, we have, however, assumed that it is of the form $B_0=(1+M)B'$. The other B_n can, of course, be found using the values of a_n, b_n, f_n, and B_0. Replacing $\tilde{\Phi}_\lambda$ by $\tilde{\Omega}_{2\lambda}$ and using the representations given in (3.5) and (3.11) in the left side of (3.12), one finds

$$
\begin{aligned}
&\tfrac{1}{4}\sum_{n=0}^{\infty}z^n\sum_{m=0}^{n}\{N(\mathbb{1}-M)(m+2)-2z^2\mathbb{1}\\
&\quad+N(\mathbb{1}+M)(m+2)z^4\}C_{m+2,n-m}\bar{\gamma}_{m+2}\\
&\quad+\tfrac{1}{2}\{(\mathbb{1}-M)z+(\mathbb{1}+M)z^{-1}\}\sum_{n=0}^{\infty}C_{1,n}\bar{\gamma}_1z^n\\
&\quad\overset{?}{=}\sum_{n=0}^{\infty}B_nz^n,\ \bar{\gamma}_n\equiv\gamma_{n/2}.
\end{aligned}
\tag{3.13}
$$

Comparing coefficients in (3.13), we are led to the following relations:

$$
\begin{aligned}
&B_0=(\mathbb{1}+M)\{\tfrac{1}{2}\mathbb{1}+\tfrac{1}{4}(a_2-a_1^2)N+(\tfrac{1}{2}\mathbb{1}+N)C_{1,1}\bar{\gamma}_1\},\\
&\gamma_1=-\tfrac{1}{16}(a_3-f_3-2a_2a_1+2a_1^3+2a_1f_2+b_1)\\
&\quad\times\{(\mathbb{1}+\bar{\alpha})\mathbb{1}+(\mathbb{1}-\bar{\alpha})N\}(\mathbb{1}-M),\\
&\bar{\alpha}={}^{12}C_{1,0}/{}^{11}C_{1,0}, \quad C_{2,0}\bar{\gamma}_2=\bar{B}_2,\\
&C_{n,0}\bar{\gamma}_n=\bar{B}_n+\sum_{m=0}^{n-3}\bar{C}_{m+2,n-m}\bar{\gamma}_{m+2}, \quad \text{for} \quad n\geq 3,
\end{aligned}
\tag{3.14}
$$

where

$$
\begin{aligned}
\bar{B}_n=&(\mathbb{1}+M)\frac{n}{(n^2-1)}\Big\{NB_{n+2}+\frac{1}{n+2}\\
&\times\Big[B_n-\frac{a_1B_{n-1}}{n+1}+p_nB_{n-2}\Big]-\Big[\Big(N+\frac{1}{n+2}\Big)C_{1,n+1}\\
&-\frac{a_1C_{1,n}}{(n+1)(n+2)}+p_n\frac{C_{1,n-1}}{n+2}\Big]\bar{\gamma}_1\Big\}\\
&+\frac{(\mathbb{1}-M)N(B_{n-2}-C_{1,n-1}\bar{\gamma}_1)}{n}
\end{aligned}
$$

and

$$
\begin{aligned}
\bar{C}_{m+2,n-m}=&(\mathbb{1}+M)\frac{n}{2(n^2-1)}\Big\{(1-\frac{m+2}{n+2})NC_{m+2,n-m}\\
&+\frac{(m+2)Na_1C_{m+2,n-1-m}}{(n+1)(n+2)}\\
&-\Big[\frac{-1}{n+2}+m+2+\frac{m+2}{n+2}Np_n\Big]C_{m+2,n-2-m}\\
&-\frac{a_1^2C_{m+2,n-3-m}}{(n+1)(n+2)}+\frac{p_{n,m}}{n+2}p_nC_{m+2,n-4-m}\Big\}\\
&+\frac{(\mathbb{1}-M)[p_{n,m}NC_{m+2,n-4-m}-(m+2)C_{m+2,n-2-m}]}{2n},
\end{aligned}
$$

with

$$
p_n=\Big[\frac{a_2}{n}+f_2-\frac{a_1^2}{n(n+1)}\Big]\Big/n
$$

and

$$p_{n,m} = \begin{cases} 1 & \text{if } m < n-3 \\ 0 & \text{if } m = n-3 \end{cases}.$$

Since the inverse of $C_{n,0}$, which is given by

$$C_{n,0}^{-1} = (1 + 2/n)\,1 - (1/n)\{1 - \tfrac{1}{2}(a_2 - a_1^2)N\}(1 - M), \tag{3.15}$$

exists for the relevant values of n, (3.14) will give us the desired values of γ_λ. To see whether this defined set of γ_λ will indeed make the left side of (3.13) analytic in a circle of nonvanishing radius, so that relations (3.13) can be satisfied, let us find some bounds on $\bar{\gamma}_m$. Because of the bounds on B's and C's, clearly in a circle of radius $\eta < R/2$, we can find a finite constant $q(\eta)$ such that $\|\bar{B}_n\| < \eta^{-n+2}$. $q(\eta)$ and $\|\bar{C}_{n,m}\| < \eta^{-m+2} \cdot \frac{1}{2}q(\eta)$. Calling

$$q'(\eta) = 2q(\eta)\,[\sup_{n\geq 2}\|C_{n,0}^{-1}\|],$$

we can define $\bar{\Gamma}_n \geq \|\bar{\gamma}_n\|$ with $\bar{\Gamma}_2 = \|\bar{\gamma}_2\|$ by the following relations,

$$\bar{\Gamma}_n = q'(\eta)[\eta^{-n+2} + \eta^{-1}\bar{\Gamma}_{n-1} + \cdots \eta^{-n+2}\bar{\Gamma}_2], \tag{3.16}$$

from which it follows that

$$\bar{\Gamma}_n = \eta^{-n+2}\{1 + q'(\eta)\}^{n-2}\bar{\Gamma}_2 \tag{3.17}$$

or

$$\|\bar{\gamma}_n\| \leq \{q'(\eta)\}^{-n}\bar{b}.$$

Since

$$\|\tilde{\Omega}_n\| = \sum_{m=0}^{\infty}\|C_{n,m}\| \cdot |z|^m \leq \bar{q}(\eta)\sum_{m=0}^{\infty}\frac{|z|^m}{\eta^m},$$

it follows that $\sum_{m=2}^{\infty}\tilde{\Omega}_m\bar{\gamma}_m z^m$ is absolutely bounded, uniformly in z, by the convergent series

$$\bar{q}(\eta)\{1 - |z|/\eta\}^{-1}\sum_{n=2}^{\infty}|z|^n\{\bar{q}(\eta)\}^{-n}. \tag{3.18}$$

This implies that the left-hand side of Eqs. (3.13) is analytic, the necessary interchange of summations is justified, and therefore we can conclude that given the set of potentials analytic in a circle of radius R, we can find an acceptable set of γ_λ, which makes the relations (3.13) true in a circle of nonvanishing radius less than R.

4. INTEGRAL EQUATIONS

In this section our aim is to show that there exist many sets of $\{\gamma_\lambda\}$ such that the corresponding $W(r)$ is diagonal and $V(r)$ is a multiple of the identity. In other words we can find sets of $\{\gamma_\lambda\}$ such that the function $K(r,r')$ defined through Eqs. (2.19) and (2.20) exists and the corresponding potentials defined through Eqs. (2.21) are such that the potential $W(r)$ is diagonal and $V(r)$ is a multiple of the identity. We will also show that the regular solution to Eqs. (2.18) is given by Eq. (2.23). But, in order to state the above purpose in a more precise way, we find it convenient to break up the problem into two cases.

Case 1: The diagonal analytic matrix function $W(r)$ and the analytic function $V(r)$, which is not a matrix, are such that, at $r = 0$, the functions $r^2W_+(r) = r^2V(r) = W_-(r) = rW_+(r) + rV(r) = 0$ and the defined numbers a_n, b_n, and f_n are such that $a_3 - f_3 - 2a_2a_1 + 2a_1^3 + 2a_1f_2 + b_1 = 0$. In this case, there exists a set of constant matrices γ_λ, $\gamma_{1/2} = 0$, such that the following is true: Let us define

$$f_\pm(r,r') = \sum_{\lambda\in S_1}(-\tfrac{1}{2} \pm \lambda)(rr')^{\lambda+1/2}\gamma_\lambda,$$

$$S_1 = \{1, 3/2, 2, 5/2, \cdots\},\quad R(r) = \begin{bmatrix} \mathbf{R}_-(r) \\ \mathbf{R}_+(r) \end{bmatrix},$$

$\mathbf{R}_\pm(r) = \frac{1}{2}\{rg_-(r) \mp r^{-1}g_+(r), rg_-(r) \pm (r)^{-1}g_+(r)\}$ and $g_\pm(r) = \exp\{\pm\int_0^r dr'r'V(r')\}$, and suppose $\mathbf{K}_\pm$ are the unique row-vector solutions of the integral equations

$$\mathbf{K}_\pm(r,r') = \mathbf{R}_\pm(r)f_\pm(r,r') - \int_0^r d\rho\rho^{-2}\mathbf{K}_\pm(r,\rho)f_\pm(\rho,r'). \tag{4.1}$$

Form the square matrix $K(r,r')$:

$$K(r,r') = \begin{bmatrix} \mathbf{K}_-(r,r') \\ \mathbf{K}_+(r,r') \end{bmatrix}.$$

Then

$$K(r,0) = 0$$

and

$$\{-N(1+M)[1 - r^2V(r)] - r^2[W(r) + 2V(r)1]$$
$$- 2r^2V(r)M - r^3\dot{V}(r)M + r^4V^2(r)1\}R(r)$$
$$= 2\dot{K}(r,r) - 2r^{-1}K(r,r) - 2r^{-1}$$
$$\times [1 - r^2V(r)]MK(r,r) \tag{4.2}$$

and $K(r,r')$ solves the partial differential equations

$$\left\{r^2\frac{d^2}{dr^2} - r^2W(r)\right\}K(r,r')$$
$$= \left\{r'^2\frac{d^2}{dr'^2} - 2r'[1 - r^2V(r)]M\frac{d}{dr'} + 1\right.$$
$$\left. + (N + M + NM)[1 - r^2V(r)]\right\}K(r,r'). \tag{4.3}$$

The proof of the above statement is given in Appendices A and B.

Application of the differential operator defined in the left-hand side of Eqs. (2.18) to

$$\Phi_\lambda = (-\tfrac{1}{2}1 + \lambda N)\{R(r)r^{\lambda+1/2} - \int_0^r dr'r'^{-2}K(r,r')$$
$$\times r'^{\lambda+1/2}\}, \tag{4.4}$$

together with two integrations by parts and use of (4.3) and (4.2) and noting that (4.4) satisfies the boundary conditions given in (3.4), verifies that

(4.4) is the regular solution of Eqs. (2.18) for the case of $\lambda > \frac{1}{2}$. Although, from the point of view of the inverse problem, we do not need the form of $\Phi_{1/2}$ for Case 1, since $\gamma_{1/2} = 0$ in this case and $\Phi_{1/2}$ is not needed when one tries to find γ_λ from scattering information, for the sake of completeness and possible physical use let us mention that the "sense" solution and a "nonsense" solution can be given as follows:

$$^{22}\Phi_{1/2} = r^2 g_-(r) - \int_0^r dr' r'^{-1}\{^{21}K(r,r') + {}^{22}K(r,r')\}, \tag{4.5a}$$

$$^{11}\Phi_{1/2} = \tfrac{1}{2} g_+(r) - \tfrac{1}{2}\int_0^r dr' r'^{-1}\{^{11}K(r,r') - {}^{12}K(r,r')\}. \tag{4.5b}$$

The proof is the same as that of (4.4), except that for proof of the fact that (4.5a) is the "sense" solution, one needs to observe that the function

$$\eta_0(r) = r^2 g_-(r) - \int_0^r dr' r'^{-1}\{^{11}K(r,r') + {}^{12}K(r,r')\} \tag{4.5c}$$

satisfies the same differential equation as the one satisfied by a "nonsense" solution, that is Eq. (3.9a). But since $\eta_0(0) = \dot{\eta}_0(0) = 0$, this implies that $\eta_0(r)$ is identically zero. For this case we are unable to give a form of $^{12}\Phi_{1/2}(r)$ in terms of $K(r,r')$. Again from the point of view of the scattering problem we do not need to know the form of $^{12}\Phi_{1/2}(r)$ in this case, but as we shall see for the general case, when $\gamma_{1/2} \neq 0$, the lack of having a simple relation between $^{12}\Phi_{1/2}$ and $K(r,r')$ will unfortunately complicate the form of the integral equations defining $K(r,r')$.

We conclude this case by noting that we can find the analog of equation (2.7) for the case when the Schrödinger equation contains a central, a spin-orbit, and a tensor potential if we again substitute the representations of $f_\pm(r,r')$ in terms of γ_λ in Eqs. (4.1) and then simplify by making use of Eqs. (4.4):

$$K(r,r') = \sum_{\lambda\in S_1} \Phi_\lambda(r)\gamma_\lambda r'^{\lambda+1/2}. \tag{4.6}$$

Having found the desired relations for the case when $\gamma_{1/2} = 0$, or equivalently when $a_3 - f_3 - 2a_2a_1 + 2a_1^3 + 2a_1f_2 + b_1 = 0$, let us now consider the case when $\gamma_{1/2} \neq 0$.

Case 2: The diagonal analytic matrix function $W(r)$ and the analytic function $V(r)$, which is not a matrix, are such that $r^2W_+ = r^2V = W_- = rW_+ + rV = 0$, when $r = 0$, *and* the defined numbers a_n, b_n, and f_n are such that $a_3 - f_3 - 2a_2a_1 + 2a_1^3 + 2a_1f_2 + b_1 \neq 0$. In this case, again there exists a set of constant matrices γ_λ, with $\gamma_{1/2} \neq 0$, with the following properties. Let

$$S_1 = \{1, \tfrac{3}{2}, 2, \cdots\}, \quad f(x) = \sum_{\lambda\in S_1} x^{\lambda+1/2}\gamma_\lambda,$$

$$f_-(x) = -xf(x), \bar{f}_+(x) = (1+M)\gamma_{1/2}x - f(x) - f_-(x),$$

$\gamma_{1/2} = -(\gamma_0/2)(1+N)(1-M)$, and $\gamma_0 \neq 0$, a number. Let φ satisfy the differential equation $r^2\ddot{\varphi} - r^2 \times (W_+ - W_- + V)\varphi = 0$ and the boundary conditions $\varphi(0) = \frac{1}{2}$ and $\dot{\varphi}(0) = -\frac{1}{2}a_1$, let $\mathbf{R}_\pm(r)$ be defined as before, let $P' = \frac{1}{2}(1-N)(1-M)$, let the row-vector $\boldsymbol{\varphi} = (\varphi, \varphi)$, and define

$$f_1(r,r') = \mathbf{R}_-(r)r' - \boldsymbol{\varphi} r^{-1}r'P';$$

then there exist $\mathbf{K}_\pm(r,r')$, the row-vector components of a 2×2 matrix $K(r,r')$, which solve the equations

$$\begin{aligned}\mathbf{K}_-(r,r') &= \mathbf{R}_-(r)\{f_-(rr') + (r'/r)f(r^2)\} + f_1(r,r') \\ &\quad - \int_0^r d\rho\rho^{-2}\mathbf{K}_-(r,\rho)\{f_-(\rho r') + (r'/r)f(\rho r)\}, \\ \mathbf{K}_+(r,r') &= \mathbf{R}_+(r)\bar{f}_+(rr') - \int_0^r d\rho\rho^{-2}\,\mathbf{K}_+(r,\rho)\bar{f}_+(\rho r').\end{aligned} \tag{4.7}$$

Furthermore,

$$\begin{aligned}&K(r,0) = 0, \\ &\quad 2K(r,r) - 2r^{-1}\{1 + [1 - r^2V(r)]M\}K(r,r) \\ &\quad = \{-N(1+M)[1 - r^2V(r)] - r^2[W(r) + 2V(r)1] \\ &\quad - 2r^2V(r)M - r^3\dot{V}(r)M + r^4V^2(r)1\}R(r)\end{aligned} \tag{4.8}$$

and $K(r,r')$ is a solution to the partial differential equations

$$\begin{aligned}r^2\left\{\frac{d^2}{dr^2} - W(r)\right\}K(r,r') &= \left\{r'^2\frac{d^2}{dr'^2}\right. \\ &\quad - 2r'[1 - r^2V(r)]M\frac{d}{dr'} + 1 \\ &\quad \left. + (N + M + NM)[1 - r^2V(r)]\right\}K(r,r').\end{aligned} \tag{4.9}$$

Proof of existence of the set $\{\gamma_\lambda\}_{\lambda\in S}$ with $\gamma_{1/2} \neq 0$, such that the defined function $K(r,r')$ exists and satisfies Eqs. (4.8) and (4.9), is similar to that of the case when $\gamma_{1/2} = 0$ and is given in Appendix C. By using $K(r,r')$, the regular solution to Eqs. (2.18), for values of $\lambda > \frac{1}{2}$, the "sense" solution for $\lambda = \frac{1}{2}$, and a solution to (3.9c) can be written in the form

$$\begin{aligned}\Phi_\lambda &= (-\tfrac{1}{2}1 + \lambda N)\{R(r)r^{\lambda+1/2} - \int_0^r dr' r'^{-2} \\ &\quad \times K(r,r')r'^{\lambda+1/2}\}, \\ ^{22}\Phi_{1/2} &= r^2 g_-(r) - \int_0^r dr' r'^{-1}\{^{22}K(r,r') \\ &\quad + {}^{21}K(r,r')\}, \\ ^{12}\Phi_{1/2} &= \frac{r^{-1}}{\gamma_0}\{\tfrac{1}{2}r^2 g_-(r) + \tfrac{1}{2}g_+(r) - \varphi(r) \\ &\quad - \int_0^r dr' r'^{-1}\,{}^{11}\bar{K}(r,r')\},\end{aligned} \tag{4.10}$$

with

$$\bar{K}(r,r') = \sum_{\lambda\in S_1}\Phi_\lambda(r)\gamma_\lambda r'^{\lambda+1/2}.$$

Proofs of above statements are the same as those of Case 1, and need not be carried through. Following the same procedure used for Case 1, we find

$$K(r,r') = \sum_{\lambda\in S} \Phi_\lambda(r)\gamma_\lambda r'^{\lambda+1/2}, \tag{4.11}$$

where

$$S = \{\tfrac{1}{2}, 1, 3/2, \cdots\} \quad \text{and} \quad \Phi_{1/2} = \begin{bmatrix} \varphi & {}^{12}\Phi_{1/2} \\ 0 & {}^{22}\Phi_{1/2} \end{bmatrix}.$$

From the point of view of the inverse scattering problem, Eqs. (4.7) as written down are not in a desirable form, because given the set $\{\gamma_\lambda\}$, in order to solve for $K(r,r')$, we have to know both $V(r)$ and $(W_+ - W_-)$, and we must solve for $\varphi(r)$. A possible way out is to note that, given the set $\{\gamma_\lambda\}$ and $V(r)$, one is able to find $\mathbf{K}_+(r,r')$ from Eq. (4.7) without having to know the function $(W_+ - W_-)$. Having found $\mathbf{K}_+(r,r')$, one then forms

$$\tau(r) = -\tfrac{1}{2}\{g_+(r) + \int_0^r dr' r'^{-1} \times [{}^{21}K(r,r') - {}^{22}K(r,r')]\}. \tag{4.12}$$

It is shown in Appendix C that the so-defined τ satisfies the differential equations

$$\left\{r^2\frac{d^2}{dr^2} - r^2(W_+ + W_- - V) - 2\right\}\tau(r) = 2(1 - r^2V)\varphi(r). \tag{4.13}$$

In general, Eqs. (4.12) and (4.13) enable us to find $\varphi(r)$, if we know $W_+ + W_-$, without having to solve a differential equation. $W_+ + W_-$ can be defined in terms of $K(r,r)$ if we consider the elements of Eqs. (4.8):

$$r^2(W_+ - W_-) + 3r^2V - r^4V^2 - 1 = r^{-2}g_+(K_1 + EK_2) + g_-(K_{1-} + EK_{2-}) - r^{-1}g_+\dot{K}_1 - rg_-\dot{K}_{1-}, \tag{4.14a}$$

$$3r^2V + r^3\dot{V} - 1 = r^{-2}g_+(K_1 + EK_2) - g_-(K_{1-} + EK_{2-}) - r^{-1}g_+\dot{K}_1 + rg_-\dot{K}_{1-}, \tag{4.14b}$$

$$r^2V + r^3\dot{V} + 1 = r^{-2}g_+(EK_1 + K_2) + g_- \times (EK_{1-} + K_{2-}) - r^{-1}g_+\dot{K}_2 - rg_-\dot{K}_{2-}, \tag{4.14c}$$

$$r^2(W_+ + W_-) + r^2V - r^4V^2 + 1 = r^{-2}g_+(EK_1 + K_2) - g_-(EK_{1-} + K_{2-}) - r^{-1}g_+\dot{K}_2 + rg_-\dot{K}_{2-}, \tag{4.14d}$$

where

$$K_1(r) = {}^{11}K(r,r) + {}^{12}K(r,r),$$
$$K_{1-}(r) = {}^{11}K(r,r) - {}^{12}K(r,r),$$
$$K_2(r) = {}^{21}K(r,r) + {}^{22}K(r,r),$$
$$K_{2-}(r) = {}^{21}K(r,r) - {}^{22}K(r,r),$$

and

$$E(r) = 1 - r^2V(r).$$

From (4.14) we find that

$$r^2(W_+ + W_-) = r^3\dot{V} + r^4V^2 - 2g_-(EK_{1-} + K_{2-}) + 2rg_-\dot{K}_{2-}. \tag{4.15}$$

Substituting (4.15) in (4.13) gives us the desired representation for the row-vector φ:

$$\varphi = \theta(r)\mathbf{K}_-(r,r)(\mathbf{1} - M)N + [\bar{\theta}(r), \bar{\theta}(r)], \tag{4.16}$$

with

$$\theta(r) = -g_-(r)\tau(r),$$
$$\bar{\theta} = \left(\left\{r^2\frac{d^2}{dr^2} - r^2[r\dot{V} - V + r^2V^2 - 2r^{-2}g_-(K_{2-} - r\dot{K}_{2-})] - 2\right\}\tau\right)\Big/2E.$$

Substituting (4.16) in (4.7), we find that $\mathbf{K}_-(r,r')$ satisfies the following integral equations:

$$\mathbf{K}_-(r,r') = \mathbf{R}_-f(r,r,r') + f_2(r,r') + \mathbf{K}_-(r,r)P - \int_0^r d\rho\rho^{-2}\mathbf{K}_-(r,\rho)f(\rho,r,r'), \tag{4.17}$$

where

$$f(\rho,r,r') = f_-(\rho r') + (r'/r)f(\rho r),$$
$$f_2(r,r') = \mathbf{R}_-(r)r' - (\bar{\theta},\bar{\theta})r^{-1}r'p',$$
$$P = r^{-1}r'\theta(\mathbf{1} - M)P'.$$

Inspection of (4.17) indicates that we no longer need to know the function $(W_+ - W_-)$ in order to find $\mathbf{K}_-(r,r')$.

We conclude this section by pointing out that the above analysis proves that if the set $\{\gamma_\lambda\}$ indeed corresponds to a tensor and a spin-orbit potential, then the integral equations (4.17) have at least one solution and that this solution, together with $\mathbf{K}_+(r,r')$, satisfies the differential equations (4.9) with the boundary conditions defined by (4.8).

5. A SERIES REPRESENTATION OF THE WAVE-FUNCTIONS

In this section we intend to derive a series representation of the wavefunctions of the central, spin-orbit, and tensor potentials, similar to Eq. (2.8), in terms of the wavefunctions of spinless particles. We assume the central, spin-orbit, and tensor potentials satisfy the conditions stated in Case 1, or Case 2, and the comparison potential $u(r)$, which is associated with $v_\lambda(r)$—the wavefunction of the spinless particle—is such that the coefficients c_λ corresponding to this potential $u(r)$ and the potential $\tilde{u}(r) = 1$ are such that $c_{1/2} = c_{3/2} = 0$. In other words $u(r)$ is chosen in such a way that the set Ω, appearing in Eq. (2.3) for the case when $\tilde{u}(r) = 1$, is identical with the set $\{1, 2, 5/2, 3, 7/2, \cdots\}$. We should note that indeed such potentials exist and $u(r) = 0$ is an example of such potentials, because from Eq. (2.9) we note that c_λ corresponding to potentials $u(r) = 0$ and $\tilde{u}(r) = 1$ are such that $c_{1/2} = c_{3/2} = 0$.

Before considering the representation of the

wavefunctions corresponding to two spin-$\frac{1}{2}$ particles interacting via a central, a tensor, and a spin-orbit force in terms of the wavefunctions $v_\lambda(r)$ of a spinless particle in a central field, let us conform with notations used in Sec. 2 and renormalize the wavefunctions of the two spin-$\frac{1}{2}$ particles, which are given by Eq. (4.4) or (4.10) in the following way:

$$\bar{\Phi}_\lambda(r) \equiv \Phi_\lambda(r)\tilde{E}_\lambda, \tag{5.1}$$

where $\tilde{E}_\lambda = e_{\lambda+1}(\mathbb{1} + M) + e_{\lambda-1}(\mathbb{1} - M)$ for $\lambda \epsilon S_1$ and $\tilde{E}_{1/2} = e_{3/2}$, with $e_\lambda = (\pi/2)^{1/2}(\frac{1}{2})^\lambda[\Gamma(1+\lambda)]^{-1}$.

Since the results hold true for potentials W and V satisfying either conditions of Case 1 or of Case 2, we do not need to distinguish between the wavefunctions related to potentials of Case 1 or of Case 2 in this section. Therefore, unless specifically stated otherwise, what follows is applicable to either case.

Multiplication of (4.4) or (4.10) by $\tilde{E}_\lambda$ on the right, and use of (2.13), implies that, for $\lambda \geq 1$,

$$\begin{aligned}\bar{\Phi}_\lambda(r) = T'_\lambda\{&g_-(r)s_{\lambda+1}(r)(\mathbb{1} + M) + g_+(r)s_{\lambda-1}(r)(\mathbb{1} - M) \\ &- \int_0^r dr' r'^{-2}K(r,r')[r'^{-1}s_{\lambda+1}(r')(\mathbb{1} + M) \\ &+ r's_{\lambda-1}(r')(\mathbb{1} - M)]\},\end{aligned} \tag{5.2}$$

where T'_λ is given by (2.24) and $s_\lambda(r)$ by (2.13). Substituting Eq. (2.5), for the case when $u(r) = 1$, as a definition of $s_\lambda(r)$ in (5.2), and performing the necessary change of order of integrations in the resulting equations, we get

$$\begin{aligned}\bar{\Phi}_\lambda(r) = T'_\lambda\{&g_-(r)v_{\lambda+1}(r)(\mathbb{1} + M) + g_+(r)v_{\lambda-1}(r)(\mathbb{1} - M) \\ &- \int_0^r dr' r'^{-2}\tilde{K}^-(r,r')v_{\lambda+1}(r') \\ &- \int_0^r dr' r'^{-2}\tilde{K}^+(r,r')v_{\lambda-1}(r)\}\end{aligned} \tag{5.3}$$

where

$$\begin{aligned}\tilde{K}^\pm(r,r') = \{&g_\pm(r)k(r,r')\mathbb{1} + K(r,r')r^{\pm 1} \\ &- \int_{r'}^r d\rho\rho^{-2}k(\rho,r')K(r,\rho)\rho^{\pm 1}\}(\mathbb{1} \mp M)\end{aligned} \tag{5.4}$$

and $k(r,r')$ is given by (2.7), with $\tilde{u}(r) = 1$.

In order to give an expression relating $W(r)$, $V(r)$, $u(r)$, and $\tilde{K}^\pm(r,r)$, we first find $K(r,r)$ in terms of $\tilde{K}^\pm(r,r)$ and $k(r,r)$ from Eqs. (5.4). Next substitute this expression of $K(r,r)$ in Eqs. (4.2) or (4.8) and note that $u(r)$ and $k(r,r)$ obey Eq. (2.6), with $u(r) = 1$. The resulting equations will have the following form:

$$\begin{aligned}\{&-N(\mathbb{1} + M)[1 - r^2V(r)] - r^2[W(r) + \mathbb{1} - u(r)\mathbb{1} + 2V(r)\mathbb{1}] - 2r^2V(r)M - r^3\dot{V}(r)M + r^4\dot{V}^2(r)\mathbb{1}\}R(r) \\ &= \frac{d}{dr}\{r^{-1}\tilde{K}^+(r,r) + r\tilde{K}^-(r,r)\} - r^{-1}\{\mathbb{1} + [1 - r^2V(r)]M\}\{r^{-1}\tilde{K}^+(r,r) + r\tilde{K}^-(r,r)\}.\end{aligned} \tag{5.5}$$

Having seen the relations between $W(r)$, $V(r)$, and $u(r)$, and also between their corresponding wavefunctions $\bar{\Phi}_\lambda(r)$ and $v_\lambda(r)$ in terms of $\tilde{K}^\pm(r,r')$, let us next try to eliminate these auxiliary functions, $\tilde{K}^\pm(r,r')$, in the mentioned relations. A step toward this purpose is to find $\tilde{K}^\pm$ in terms of $\bar{\Phi}_\lambda(r)$ and $v_\lambda(r)$, which can be accomplished by first rewriting the integrals $\int_{r'}^r$ in (5.4) as $\int_0^r - \int_0^{r'}$ and then substituting the definition of $k(r,r')$, (2.7), and that of $K(r,r')$, (4.6) or (4.11), in the resulting equations. Next perform the necessary change of order of summations and integrations and make use of the existing relations between the wavefunctions corresponding to a central, a spin-orbit, and a tensor potential and $K(r,r')$, (4.4) or (4.10), and between $v_\lambda(r)$ and $k(r,r')$, (2.10). The resulting equations give us the desired representations of $\tilde{K}^\pm(r,r')$, which are

$$\begin{aligned}\tilde{K}^\pm(r,r') = &\sum_{\lambda\in S}\bar{\Phi}_\lambda(r)\tilde{\gamma}_\lambda v_{\lambda\pm1}(r')(\mathbb{1} \mp M) \\ &+ \sum_{\lambda\in S_2}\tilde{T}_{\lambda\pm1}\bar{\Phi}_{\lambda\pm1}(r)\tilde{c}_\lambda v_\lambda(r')(\mathbb{1} \mp M),\end{aligned} \tag{5.6}$$

where $S = \{\frac{1}{2}, 1, \frac{3}{2}, 2, \cdots\}$, $S_2 = \{1, 2, \frac{5}{2}, 3, \frac{7}{2}, \cdots\}$, $\tilde{c}_\lambda = \frac{1}{2}c_\lambda$, with c_λ corresponding to coefficients associated with potentials $u(r)$ and $\tilde{u}(r) = 1$.

We have, also,

$$\begin{aligned}&\tilde{\gamma}_\lambda = \tilde{E}_\lambda^{-1}\gamma_\lambda[(1/2e_{\lambda-1})(\mathbb{1} + M) + (1/2e_{\lambda+1})(\mathbb{1} - M)], \\ &\tilde{T}_\lambda = (\lambda^2 - \tfrac{1}{4})^{-1}(\tfrac{1}{2}\mathbb{1} + \lambda N), \\ &\bar{\Phi}_0(r) = -\tfrac{1}{2}\{g_-(r)s_1(r) - \int_0^r dr' r'^{-3}K(r,r') \\ &\qquad \times s_1(r')\}(\mathbb{1} + M).\end{aligned} \tag{5.7}$$

In above derivations use was made of the fact that $\gamma_{1/2}$ must always be of the form $\gamma_{1/2} = \gamma(\mathbb{1} - M)$. Therefore we never needed to consider the function $v_\lambda(r)$ with negative values of λ, which we have not defined in this work. The function $v_{-1/2}(r)$ which appears in (5.6) is a symbol introduced only for the sake of simplicity of the form of Eqs. (5.6); it may be assumed to be identically zero. The necessary interchange of order of integrations and summations, in the above derivations, is justified if γ_λ satisfies Eq. (3.17) and c_λ is like (2.9).

Representation (5.7) of $\bar{\Phi}_0$ can be changed so that $\bar{\Phi}_0$ is given directly in terms of $\tilde{K}^-(r,r')$. The procedure to be followed is exactly the same as the one used in deriving (5.3):

$$\begin{aligned}\bar{\Phi}_0(r) = -\tfrac{1}{2}\{&g_-(r)v_1(r)(\mathbb{1} - M) - \int_0^r dr' r'^{-2}\tilde{K}^-(r,r') \\ &\times v_1(r')].\end{aligned} \tag{5.8}$$

Substituting (5.6) in (5.3) and (5.8) and doing the necessary change of the order of summations and integrations, we obtain the desired series representation for $\bar{\Phi}_\lambda$:

$$\bar{\Phi}_\lambda(r) = T'_\lambda\Big(g_-(r)v_{\lambda+1}(r)(\mathbb{1}+M) + g_+(r)v_{\lambda-1}(r)(\mathbb{1}-M - \sum_{\alpha\in S}\bar{\Phi}_\alpha(r)\tilde{\gamma}_\alpha[L^{\lambda+1}_{\alpha-1}(r)(\mathbb{1}+M) + L^{\lambda-1}_{\alpha+1}(r)(\mathbb{1}-M)] - \sum_{\alpha\in S_2}\tilde{T}_{\alpha-1}\bar{\Phi}_{\alpha-1}(r)\tilde{c}_\alpha L^{\lambda+1}_\alpha(r)(\mathbb{1}+M) - \sum_{\alpha\in S_2}\tilde{T}_{\alpha+1}\bar{\Phi}_{\alpha+1}(r)\tilde{c}_\alpha L^{\lambda-1}_\alpha(r)(\mathbb{1}-M)\Big) \quad \text{for } \lambda\in S_1, \tag{5.9}$$

$$\bar{\Phi}_0(r) = \tfrac{1}{2}\{g_-(r)v_1(r) - \sum_{\alpha\in S_1}\bar{\Phi}_\alpha(r)\tilde{\gamma}_\alpha L^1_{\alpha-1}(r) - \sum_{\alpha\in S_2}\tilde{T}_{\alpha-1}\bar{\Phi}_{\alpha-1}(r)\tilde{c}_\alpha L^1_\alpha(r)\}(\mathbb{1}+M), \tag{5.10}$$

where $L^\lambda_\alpha(r) = \int_0^r dr' r'^{-2} v_\lambda(r')v_\alpha(r')$.

Let us apply the procedures used for deriving (5.9) and (5.10) to the "sense" solution of the wavefunction at $\lambda = \frac{1}{2}$; we find that

$$\tilde{\Psi}_{1/2}(r) = P_0\Big(g_-(r)v_{3/2}(r) - \sum_{\alpha\in S_1}\bar{\Phi}_\alpha(r)\tilde{\gamma}_\alpha L^{3/2}_{\alpha-1}(r) - \sum_{\alpha\in S_2}\tilde{T}_{\alpha-1}\bar{\Phi}_{\alpha-1}(r)\tilde{c}_\alpha L^{3/2}_\alpha(r)\Big)(\mathbb{1}+M)P_0,$$

where

$$\tilde{\Psi}_{1/2} = \begin{bmatrix} 0 & 0 \\ 0 & {}^{22}\Phi_{1/2}(r)e_{3/2} \end{bmatrix} \quad \text{and} \quad P_0 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}. \tag{5.11}$$

In the above derivations, as before, use was made of the fact that $\gamma_{1/2}$ must always be of the form $\gamma_{1/2} = \gamma(\mathbb{1} - M)$. So, for the sake of simplicity of the form of Eq. (5.9), we introduced the symbol $L^{\lambda+1}_{-1/2}(r)$ which may be taken to be identically zero.

We should note that Eqs. (5.9), (5.10), and (5.11) are the analogs of Eq. (2.8) and, in principle, one should be able to use Eqs. (5.9), (5.10), and (5.11), in a manner similar to the method used in Ref. 3, for finding the set $\{\gamma_\lambda\}$ from the scattering information. In other words, by considering the asymptotic behavior of Eqs. (5.9), (5.10), and (5.11), as $r\to\infty$, we should be able to find the set $\{\gamma_\lambda\}$ in terms of eigenphase shifts and mixture parameters[19] of the wavefunctions corresponding to a central, a spin–orbit, and a tensor potential. However, in this work we will not consider the problem of finding the coefficients γ_λ from eigenphase shifts and mixture parameters.

We would like to point out that since $\gamma_{1/2}$ used in (4.7) is such that ${}^{11}\gamma_{1/2} = {}^{12}\gamma_{1/2} = 0$, one does not need to know the form of the "nonsense" solution in the problem of finding the constants γ_λ from eigenphase shifts and mixture parameters. Therefore, we have not given a representation for it either. For the case $\gamma_{1/2} = 0$, of course, we do not need to know the form of ${}^{12}\Phi_{1/2}$ in order to find $\{\gamma_\lambda\}$ from the S matrix. For the case $\gamma_{1/2} \neq 0$, it appears that we need a representation for ${}^{12}\Phi_{1/2}$ similar to (5.9) in order to be able to find $\{\gamma_\lambda\}$ from the S matrix. But since it seems there is no such simple representation of ${}^{12}\Phi_{1/2}$, and, since any other representation of ${}^{12}\Phi_{1/2}$ should be in a form which is convenient for the method to be used for finding $\{\gamma_\lambda\}$ from the S matrix, in this work we have not tried to give a representation of ${}^{12}\Phi_{1/2}$ in terms of $v_\lambda(r)$.

6. CONSTRUCTION OF POTENTIALS FROM THE INTEGRAL EQUATIONS

Up to now we have always assumed that the central, the spin–orbit, and the tensor potentials are given, and have proved that if they satisfy certain conditions, then we can find a set $\{\gamma_\lambda\}$, such that solution to (4.1) or (4.7) exists and satisfies (4.2) or (4.8). Bounds on $\{\gamma_\lambda\}$ are given by (3.17). In this section we would like to answer the reverse question for Case 1: If a set $\{\gamma_\lambda\}_{\lambda\in S_1}$ obeying the bounds (3.17) is given, then how are we going to find out whether this set $\{\gamma_\lambda\}_{\lambda\in S_1}$ indeed corresponds to a central, a spin–orbit, and a tensor potential, that is, to potentials $W(r)$ and $V(r)$ where the former is diagonal and the latter is a multiple of the identity. We will also find the functions $W(r)$ and $V(r)$ corresponding to the given set $\{\gamma_\lambda\}_{\lambda\in S_1}$, if the set corresponds to a central, a spin–orbit, and a tensor potential.

A possible way to answer the above question is to define the following integral equations:

$$H_\pm(r,r') = \tfrac{1}{2}[(\mathbb{1}+M)r + (\mathbb{1}-M)r^{-1}]f_\pm(r,r') - \int_0^r d\rho\rho^{-2}H_\pm(r,\rho)f_\pm(\rho,r'), \tag{6.1}$$

where the $f_\pm(r,r')$ are the same as those given in (4.1). In order to relate $H_\pm$ to the function K defined by Eqs. (4.1), let us choose an arbitrary function $V(r)$, which is a multiple of the identity and is to be defined later, in Eqs. (4.1), and define $G(r)$ to be

$$G(r) = \begin{bmatrix} \mathbf{G}_-(r) \\ \mathbf{G}_+(r) \end{bmatrix} = \tfrac{1}{2}(\mathbb{1}+M)g_-(r) + \tfrac{1}{2}(\mathbb{1}-M)g_+(r), \tag{6.2}$$

where $g_\pm(r)$ are defined in (4.1). Then multiplying (6.1) by $\mathbf{G}_\pm(r)$, respectively, we find

$$\mathbf{G}_\pm(r)H_\pm(r,r') = \mathbf{R}_\pm(r)f_\pm(r,r') - \int_0^r d\rho\rho^{-2}\mathbf{G}_\pm(r)H_\pm(r,\rho)\times f_\pm(\rho,r'). \tag{6.3}$$

The observation that Eqs. (6.3) are nothing but Eqs. (4.1) leads us to the conclusion that

$$\mathbf{K}_\pm(r,r') = \mathbf{G}_\pm(r)H_\pm(r,r'). \tag{6.4}$$

Since the functions $H_\pm$ are independent of $V(r)$ and are known for the given set $\{\gamma_\lambda\}_{\lambda\in S_1}$, the problem of finding $\mathbf{K}_\pm$ is then equivalent to the problem of finding $G(r)$. To find the potential $V(r)$, or the function

$G(r)$, we add and subtract the analog of Eqs. (4.14b) and (4.14c) for case $\gamma_{1/2} = 0$. Noting that one of the two resulting equations is a perfect differential integration of that equation and finding from the analysis of Sec. 3 that, at $r = 0$, $r^{-2}\{K_1(r,r) + K_2(r,r)\} = 2$, we conclude that the analogs of Eqs. (4.14b) and (4.14c), for the case $\gamma_{1/2} = 0$, are equivalent to the following equations:

$$2r^2V(r) - 2 = -r^{-2}g_+(r)\{K_1(r,r) + K_2(r,r)\} + g_-(r)\{K_{1-}(r,r) - K_{2-}(r,r)\},$$

$$2r^2V(r) - 2 = -r^{-1}g_+^2(r)\frac{d}{dr}g_-(r)[K_1(r,r) - K_2(r,r)] + g_-^2(r)^3\frac{d}{dr}g_+(r)r^{-2} \times [K_{1-}(r,r) + K_{2-}(r,r)]. \quad (6.5)$$

Now it is only a matter of finding K_1, K_2, K_{1-}, and K_{2-} in terms of G and $H_\pm$, substituting them in Eqs. (6.5), and then multiplying Eqs. (6.5) by $g_-^2(r)$ and using the facts that $\dot{g}_\pm = \pm rVg_\pm$, to arrive at the following equations:

$${}_0H(r)t^2(r) + {}_1H(r)t(r) + r\dot{t}(r) = {}_2H(r),$$
$${}_3H(r)t^2(r) + {}_4H(r)t(r) + {}_5H(r)\dot{t}(r) = {}_6H(r), \quad (6.6)$$

where

$${}_0H(r) = H_-^{+-}(r) - H_+^{+-}(r),$$
$${}_1H(r) = 2 + H_-^{--}(r) + H_+^{--}(r) - r^{-2}[H_-^{++}(r) + H_+^{++}(r)],$$
$${}_2H(r) = r^{-2}[H_-^{-+}(r) - H_+^{-+}(r)],$$
$${}_3H(r) = -2[H_-^{+-}(r) + H_+^{+-}(r)] + r[\dot{H}_-^{+-}(r) + \dot{H}_+^{+-}(r)],$$
$${}_4H(r) = -2[H_-^{--}(r) - H_+^{--}(r)] - r^{-1}[\dot{H}_-^{++}(r) - \dot{H}_+^{++}(r)] + r[H_-^{--}(r) - \dot{H}_+^{--}(r)] + 2,$$
$${}_5H(r) = r - r^{-1}[H_-^{+-}(r) - H_+^{+-}(r)] - r[H_-^{--}(r) - H_+^{-+}(r)],$$
$${}_6H = r^{-1}\{\dot{H}_-^{-+}(r) + \dot{H}_-^{-+}(r)\},$$

with

$$t(r) = g_-^2(r) \text{ and } H_\pm^{\alpha\beta}(r) = \tfrac{1}{2}\{{}^{11}H_\pm(r,r)\alpha\ {}^{21}H_\pm(r,r)\} \times \beta\tfrac{1}{2}\{{}^{12}H_\pm(r,r)\alpha\ {}^{22}H_\pm(r,r)\}$$
for $\alpha = +,-$ and $\beta = +,-$.

So we conclude that a set $\{\gamma_\lambda\}_{\lambda\in S_1}$ corresponds to a diagonal potential W and a potential V which is a multiple of the identity if there exists a nonnegative function $t(r)$, with $t(0) = 1$, which is a common solution to both of the nonlinear equations given by (6.6).

For the cases when ${}_0H$, ${}_3H$, or (and) ${}_5H$ is (are) zero, it is easy to check whether such a common solution exists directly from (6.6). For the general case, we can reduce the nonlinear differential equations (6.6) in an obvious way to a first order differential equation and an algebraic equation:

$$[r\,{}_3H(r) - {}_5H(r)\,{}_0H(r)]t^2(r) + [r\,{}_4H(r) - {}_5H(r)\,{}_1H(r)] \times t(r) = r\,{}_6H(r) - {}_5H(r)\,{}_2H(r), \quad (6.7a)$$

$$[r\,{}_3H(r) - {}_5H(r)\,{}_0H(r)]\dot{t}(r) + [{}_1H(r)\,{}_3H(r) - {}_4H(r)\,{}_0H(r)]t(r) = {}_2H(r)\,{}_3H(r) - {}_6H(r)\,{}_0H(r). \quad (6.7b)$$

So in order to answer the question whether a given set $\{\gamma_\lambda\}_{\lambda\in S_1}$ corresponds to a tensor and a spin-orbit potential, we can, at present, only calculate $H_\pm$ for the given set $\{\gamma_\lambda\}_{\lambda\in S_1}$ and then see if there exists a nonnegative function $t(r)$, with $t(0) = 1$, satisfying *both* of the equations given in (6.7). Of course, the analysis of Sec. 3 and 4 indicates that there exist many different sets $\{\gamma_\lambda\}_{\lambda\in S_1}$ such that the corresponding $t(r)$ satisfies both of the equations in (6.7), but this knowledge is of little comfort when we are faced with a specific set of γ_λ; we still have to go through all the steps mentioned in this section and at the end we may still find out that the set in question does not correspond to a tensor and spin-orbit potential. Indeed it is unfortunate that, for the case of the inverse scattering problem at fixed energy when a tensor force is present, we are not able to find a simple test for the set $\{\gamma_\lambda\}_{\lambda\in S_1}$ which can inform us whether it corresponds to tensor and spin-orbit potentials or not. The only simple test that we have on the set $\{\gamma_\lambda\}_{\lambda\in S_1}$ is obtained from Eq. (4.5c). Power series analysis of Eq. (4.5c) indicates that, for potentials W and V satisfying conditions of case 1, γ_1 must have the following property:

$${}^{11}\gamma_1 - {}^{21}\gamma_1 + {}^{12}\gamma_1 - {}^{22}\gamma_1 = -2. \quad (6.8)$$

So, given a set $\{\gamma_\lambda\}_{\lambda\in S_1}$, if it does not satisfy (6.8), then we can be sure that it does not correspond to a tensor and a spin-orbit force.

If a set $\{\gamma_\lambda\}_{\lambda\in S_1}$ is such that Eqs. (6.7) have an acceptable common solution, then the solution $t(r)$ will define $G(r)$ and $V(r)$. Substitution of this $G(r)$ in (6.4) will give us $K(r,r')$. Then, using the analog of Eqs. (4.14a) and (4.14d), for the case when $\gamma_{1/2} = 0$, we arrive at the functions $W_+(r)$ and $W_-(r)$, in other words, the diagonal potential $W(r)$. Thus we have completed the procedure of obtaining V and W from the set $\{\gamma_\lambda\}_{\lambda\in S_1}$.

To give an example of the above procedure, let us assume that we are given the following set of coefficients[17]

$$\gamma_1 = \tfrac{1}{2}N(1 + M), \text{ and } \gamma_\lambda = 0 \text{ for all } \lambda > 1, \quad (6.9)$$

and we would like to find out if this set, $\{\gamma_\lambda\}_{\lambda\in S_1}$, corresponds to a diagonal potential $W(r)$ and a multiple of the identity potential $V(r)$. In this case, Eqs. (6.1) reduce to

$$H_-(r,r') = -\tfrac{3}{2}\{\tfrac{1}{2}(1 + M)r^{5/2} + \tfrac{1}{2}(1 - M)r^{1/2} - \int_0^r d\rho\rho^{-1/2}H_-(r,\rho)\}\gamma_1 r'^{3/2},$$

$$H_+(r,r') = \tfrac{1}{2}\{\tfrac{1}{2}(\mathbb{1}+M)r^{5/2} + \tfrac{1}{2}(\mathbb{1}-M)r^{1/2} - \int_0^r d\rho\rho^{-1/2}H_+(r,\rho)\}\gamma_1 r'^{3/2}.$$

Writing $H_\pm(r,r')$ as $H_\pm(r,r') = X_\pm(r)\gamma_1 r'^{3/2}$, we find that

$$X_-(r) = -\tfrac{3}{4}\{(\mathbb{1}+M)r^{5/2} + (\mathbb{1}-M)r^{1/2} - X_-(r)r^2\gamma_1\}$$

and

$$X_+(r) = \tfrac{1}{4}\{(\mathbb{1}+M)r^{5/2} + (\mathbb{1}-M)r^{1/2} - X_+(r)r^2\gamma_1\}.$$

Solving for $X_\pm$ from above equations and substituting them in the definitions of $H_\pm$, we find

$$H_+(r,r') = \tfrac{1}{4}r^{1/2}r'^{3/2}N(\mathbb{1}+M),$$
$$H_-(r,r') = \tfrac{3}{4}r^{1/2}r'^{3/2}N(\mathbb{1}+M). \qquad (6.10)$$

Next, we evaluate the functions ${}_iH$ defined by Eqs. (6.6) and, substituting them in Eqs. (6.7), we find that in this case the two equations (6.7a) and (6.7b) are identical and are given as

$$2t(r) + r\dot{t}(r) = 2.$$

It is easy to note that $t(r) = 1$ is the solution to the above equation. It is acceptable and it corresponds to $V(r) = 0$ or $G(r) = 1$. (6.4) and (6.10) imply

$$K(r,r') = \tfrac{1}{2}(\mathbb{1} - \tfrac{1}{2}N)(\mathbb{1}+M)r^{1/2}r'^{3/2}.$$

From (4.14a) and (4.14d) it follows that $W_+(r) = W_-(r) = 0$ Next we find Φ_λ from Eq. (4.4):

$$\Phi_\lambda = \tfrac{1}{2}(-\tfrac{1}{2}\mathbb{1} + \lambda N)[(\mathbb{1}-M)r^{-1} + [r/(\lambda+1)] \times (\lambda\mathbb{1} - \tfrac{1}{2}N)(\mathbb{1}+M)]r^{\lambda+1/2}.$$

And, as expected, one can show directly that the above function is indeed the solution to (2.18) for the case $V(r) = 0$ and $W(r) = 0$, and it also satisfies the boundary conditions (3.4).

At this point we would like to point out that, as in the work of Sabatier[9], the method constructed here can be easily generalized to potentials which are analytic functions of a rational power of z. Of course, the set S_1 will now contain certain positive rational numbers greater than $\frac{1}{2}$. But the more important point to be noticed is that in this work the analyticity of potentials was extensively used only for the proof that there exist many sets of $\{\gamma_\lambda\}_{\lambda\in S_1}$ which make the corresponding $K(r,r')$ such as to define a diagonal potential W and a V that is a multiple of the identity and which satisfy a certain condition. In other words, we can find many functions $t(r)$ satisfying the necessary conditions. We can generalize the class of potentials which we can deal with if we require that $f_\pm(r,r')$ be such that the corresponding $t(r)$ satisfies the necessary conditions, that, at r or $r' = 0$, $r^{-1}r'^{-1}$ $f_\pm(r,r')$ be finite, and that $f_\pm(r,r')$ should satisfy Eqs. (A3). The last condition is needed for the proof that K satisfies (4.3) and that Φ_λ is a solution of (2.18). Clearly the class of $f_\pm$ defined by the above conditions is not empty, but it would be very interesting if one could prove that the above class defining $f_\pm$ is larger than the class of $f_\pm$ corresponding to the class of potentials which are analytic functions of z, or analytic function of a rational power of z.

We will not consider here the general case of construction of potentials from a given set $\{\gamma_\lambda\}_{\lambda\in S}$ for the case when $\gamma_{1/2} \neq 0$. Of course, if both $\{\gamma_\lambda\}_{\lambda\in S}$ and $V(r)$ are given and if indeed they correspond to a diagonal potential $W(r)$, the tensor force, then we can find a function $K(r,r)$, from Eqs. (4.7) and (4.17), such that Eqs. (4.14b) and (4.14c) are satisfied and $K(r,r')$ is a solution to (4.9) with $W(r)$ defined through (4.14a) and (4.14d). The so-defined $W(r)$ is the tensor potential corresponding to the set $\{\gamma_\lambda\}_{\lambda\in S}$ and $V(r)$. Extension of the method to a larger class of potentials, similar to Case 1, is again possible.

ACKNOWLEDGMENTS

It is a great pleasure to thank Professor R. G. Newton for his original suggestion of this problem, for many pertinent discussions, and for his continued assistance during the course of this work. The author also wishes to thank Professor P. C. Sabatier for suggesting some improvements in the final version of the manuscript.

APPENDIX A

For the case when $\gamma_{1/2} = 0$, the proof of the statement that if there exists a $K(r,r')$ satisfying relations (4.1) and (4.2), then it is a solution to differential Eqs. (4.3) is substantially the same as the one given in Refs. 3 and 4. That is, the problem of showing that $K(r,r')$ is a solution to the differential Eqs. (4.3) is reduced to the question of existence of a nontrivial solution to the homogeneous version of Eqs. (4.1). So, to prove $K(r,r')$ satisfies Eqs. (4.3), one first writes (4.3) for the two vector components of $K(r,r')$:

$$\left(r^2\frac{d^2}{dr^2} - r^2(W_+ - W_- + V)\right)\mathbf{K}_-(r,r') = r'^2\frac{d^2}{dr'^2}\mathbf{K}_-(r,r') - 2(1-r^2V)r'\frac{d}{dr'}\mathbf{K}_+(r,r'),$$

$$\left(r^2\frac{d^2}{dr^2} - r^2(W_+ + W_- - V)\right)\mathbf{K}_+(r,r') = \left(r'^2\frac{d^2}{dr'^2} + 2\right)\mathbf{K}_+(r,r') - 2(1-r^2V) \times \left(r'\frac{d}{dr'} - 1\right)\mathbf{K}_-(r,r'). \qquad (A1)$$

Next define the following auxiliary functions:

$$\xi_-(r,r') = \left\{r^2\frac{d^2}{dr^2} - r^2(W_+ - W_- + V) - r'^2 \times \frac{d^2}{dr'^2}\right\}\mathbf{K}_-(r,r') + 2(1-r^2V)r'\frac{d}{dr'} \times \mathbf{K}_+(r,r').$$

$$\xi_+(r,r') = \left\{r^2\frac{d^2}{dr^2} - r^2(W_+ + W_- - V) - 2r'^2 \times \frac{d^2}{dr'^2}\right\}\mathbf{K}_+(r,r') + 2(1-r^2V)\left(r'\frac{d}{dr'} - 1\right) \times \mathbf{K}_-(r,r'). \tag{A2}$$

Our purpose is to show that the $\xi_\pm$ are identically zero, in other words, $\mathbf{K}_\pm$ indeed satisfy Eqs. (A1), which is equivalent to the statement that $K(r,r')$ is a solution to Eqs. (4.3). Application of the operations defined on the right-hand side of (A2) to the integral equation (4.1), defining $\mathbf{K}_\pm$, integrating by parts, and using the following identities,

$$\begin{aligned} &f_\mp(r,0) = f_\mp(0,r') = 0,\\ &r\frac{d}{dr}f_\mp(r,r') = r'\frac{d}{dr'}f_\mp(r,r'),\\ &r^2\frac{d^2}{dr^2}f_\mp(r,r') = r'^2\frac{d^2}{dr'^2}f_\mp(r,r'),\\ &r\frac{d}{dr}[f_-(r,r') + f_+(r,r')] = f_-(r,r'),\\ &r^2\frac{d}{dr}[r^{-1}f_-(r,r') + r^{-1}f_+(r,r')] = -f_+(r,r'), \end{aligned} \tag{A3}$$

shows that $\xi_\pm$ are solutions of integral equations of the form

$$\xi_\pm(r,r') = -\int_0^r d\rho\rho^{-2}\xi_\pm(r,\rho)f_\pm(\rho,r'). \tag{A4}$$

In other words, proof of $K(r,r')$ being the solution of (4.3) reduces to the uniqueness question about the solutions of (4.1). In order to consider the uniqueness question, let us rewrite (4.1) in the following form[20]:

$$\mathbf{K}_r^\pm(r') = \mathbf{Z}_r^\pm(r') + \int_0^r d\rho\mathbf{K}_r^\pm(\rho)S^\pm(\rho,r') \tag{A5}$$

where

$$\begin{aligned} \mathbf{K}_r^\pm(r') &= (rr')^{-1}\mathbf{K}_\pm(r,r'),\\ \mathbf{Z}_r^\pm(r') &= (rr')^{-1}\mathbf{R}_\pm(r)f_\pm(r,r'),\\ S^\pm(r,r') &= (rr')^{-1}f_\pm(r,r'). \end{aligned}$$

Applying the generalization of the Fredholm method for coupled equations[21] to our matrix integral equations (A5), we find

$$K_r^\pm(r') = \mathbf{Z}_r^\pm(r') + \frac{1}{\Delta^\pm(r)}\int_0^r d\rho\mathbf{Z}_r^\pm(\rho)Y_r^\pm(\rho,r') \tag{A6}$$

if $\Delta^\pm(r) \neq 0$. The Fredholm determinant and minor have the following representations[16]:

$$\Delta^\pm(r) = 1 + \sum_{m=1}^{\infty}\frac{(-1)^m}{m!}\sum_{\alpha_1,\cdots,\alpha_m=1}^{2}\int_0^r\cdots\int_0^r Y^\pm\begin{bmatrix}\alpha_1,\alpha_2,\ldots,\alpha_m\\ r_1,r_2,\ldots,r_m\\ r_1,r_2,\ldots,r_m\\ \alpha_1,\alpha_2,\ldots,\alpha_m\end{bmatrix}dr_1\cdots dr_m,$$

$$\alpha\beta Y_r^\pm(\rho,r') = \sum_{m=0}^{\infty}\frac{(-1)^m}{m!}\sum_{\alpha_1,\ldots\alpha_m=1}^{2}\int_0^r\cdots\int_0^r Y^\pm\begin{bmatrix}\alpha,\alpha_1,\alpha_2,\ldots,\alpha_m\\ \rho,r_1,r_2,\ldots,r_m\\ r',r_1,r_2,\ldots,r_m\\ \beta,\alpha_1,\alpha_2,\ldots,\alpha_m\end{bmatrix}dr_1\cdots dr_m, \tag{A7}$$

where the symbol $Y^\pm$ stands for the determinant

$$Y^\pm\begin{bmatrix}\alpha,\alpha_1,\ldots,\alpha_n\\ x,x_1,\ldots,x_n\\ y,y_1,\ldots,y_n\\ \beta,\beta_1,\ldots,\beta_n\end{bmatrix} = \begin{vmatrix}[\alpha,\beta,x,y]^\pm & [\alpha,\beta_1,x,y_1]^\pm\cdots & [\alpha,\beta_n,x,y_n]^\pm\\ [\alpha_1,\beta,x_1,y]^\pm & [\alpha_1,\beta_1,x_1,y_1]^\pm\cdots & [\alpha_1,\beta_n,x_1y_n]^\pm\\ \vdots & \vdots & \vdots\\ [\alpha_n,\beta,x_n,y]^\pm & [\alpha_n,\beta_1,x_n,y_1]^\pm\cdots & [\alpha_n,\beta_n,x_n,y_n]^\pm\end{vmatrix}$$

with $[\alpha,\beta,x,y]^\pm = {}^{\alpha\beta}S^\pm(x,y)$, $\alpha,\beta = 1,2$. (A8)

If each element of $S^\pm(\rho,r')$ is bounded in the region $0 \leq \rho,r' \leq r$ by a number $S(r)$, then convergence of the series in (A7) follows from the Hadamard's theorem,[22] i.e.,

$$\begin{aligned} |\Delta^\pm(r)| &\leq 1 + \sum_{m=1}^{\infty}(m!)^{-1}\{2rS(r)\}^m m^{m/2}\\ \|Y_r^\pm(\rho,r')\| &\leq \sum_{m=0}^{\infty}(m!)^{-1}\{2rS(r)\}^m S(r)(m+1)^{(m+1)/2} \end{aligned} \tag{A9}$$

Now it is only a matter of repeating the argument given in Ref. 4, to come to the conclusion that if we let r and r' take complex values z and z', then Eqs. (A6) can be analytically continued in the domain where $f_\pm(z,z')$ are analytic. The so-defined $\mathbf{K}_z^\pm(z')$ are meromorphic functions of z, with poles corresponding to zeroes of $\Delta^\pm(z)$ and independent of z'. Furthermore, since the Neumann series for (4.1) converges inside a nonvanishing circle centered at the origin, none of these mentioned poles can be at the origin. Therefore, in the domain of analyticity of $\mathbf{K}_\pm(z,z')$, $\Delta^\pm(z) \neq 0$, Eqs. (4.1) have unique solutions and the homogeneous version of Eqs. (4.1) have only the trivial solutions. It follows that $\xi_\pm(z,z')$, defined by (A4), are zero and from

(A2) we can conclude that $K(z,z')$ is a solution of (4.3) in its domain of analyticity.

APPENDIX B

To prove that there indeed exists a set $\{\gamma_\lambda\}_{\lambda\in S_1}$, with $\gamma_{1/2}=0$, corresponding to the analytic potentials $W(r)$ and $V(r)$, such that the integral equations (4.1) have a solution which satisfies the conditions (4.2), we recall the functions $\tilde{\Phi}_\lambda$ and $\tilde{\Omega}_{2\lambda}$, which are solutions of (2.18) and (3.2), respectively. By assuming that $\tilde{\Omega}_{2\lambda}$ satisfies the boundary conditions (3.4), then (3.14) will give us a set of $\{\gamma_\lambda\}_{\lambda\in S_1}$ with bounds given by Eqs. (3.17). Next define the following function from the above set of $\{\gamma_\lambda\}_{\lambda\in S_1}$:

$$\tilde{G}(r,r')=\sum_{\lambda\in S_1}\tilde{\Phi}_\lambda(r)\gamma_\lambda r'^{\lambda+1/2}. \tag{B1}$$

By construction $\tilde{G}(r,r')$ satisfies the conditions (4.2) and the differential equations (4.3). Therefore, $\tilde{\psi}_\lambda(r)$ defined by the equations

$$\tilde{\psi}_\lambda(r)=(-\tfrac{1}{2}1+\lambda N)\{R(r)r^{\lambda+1/2}-\int_0^r d\rho\rho^{-2}\tilde{G}(r,\rho)\times\rho^{\lambda+1/2}\} \tag{B2}$$

is a solution to the differential equations (2.18), and the corresponding $\tilde{\Omega}_{2\lambda}$ satisfies the boundary conditions (3.4). By uniqueness, we have $\tilde{\psi}_\lambda=\tilde{\Phi}_\lambda$. It is now only a matter of multiplying the Eqs. (B2) by $\gamma_\lambda r'^{\lambda+1/2}$ and summing over the sets S_1 in order to arrive at the conclusion that $\tilde{G}(r,r')$ not only satisfies relations (4.2), but it is also a solution to the integral equations (4.1). Because of the existence of absolute and uniform bounds on the functions involved, if the bounds on γ_λ are given by (3.17), the necessary exchange of summation and integration, and taking the derivatives inside the summation and integration, are all justified.

APPENDIX C

For the case when $\gamma_{1/2}\neq 0$, the proof that if $K(r,r')$ satisfies Eqs. (4.7) and (4.8), then it must be a solution to the differential equation (4.9) is exactly similar to the procedure of Appendix A. But the algebra in the demonstration that $\zeta_\pm(r,r')$, the analog of $\xi_\pm(r,r')$ of Case 1, satisfies the integral equations

$$\zeta_-(r,r')=-\int_0^r d\rho\rho^{-2}\zeta_-(r,\rho)\{f_-(\rho r')+\left(\frac{r'}{r}\right)f(\rho r\}$$

and

$$\zeta_+(r,r')=-\int_0^r d\rho\rho^{-2}\zeta_+(r,\rho)\bar{f}_+(\rho r')$$

becomes more tedious. Again, using the same argument as before, we find that $\zeta_\pm(r,r')=0$. Therefore, $K(r,r')$ is a solution to (4.9).

The existence of proof in Case 2 is also similar to that given in Appendix B. In this case the boundary condition for the solution to Eqs. (2.18) when $\lambda>\frac{1}{2}$ is given by (3.4) and for $\lambda=\frac{1}{2}$ is

$$^{11}C_{1,0}=\tfrac{1}{2},\quad {}^{11}C_{1,1}=-\tfrac{1}{2}a_1,$$
$$^{12}C_{1,0}={}^{12}C_{1,1}=0,\quad \text{and } {}^{22}C_{1,2}=\tfrac{2}{3}.$$

To prove Eq. (4.13), one first subjects Eq. (4.12) to the differential operator defined by the left-hand side of Eq. (4.13). Integration by parts and use of the fact that $K_+(r,0)=0$ leads us to the following equation:

$$\left\{r^2\frac{d^2}{dr^2}-2-r^2(W_++W_--V)\right\}\tau(r)=2(1-r^2V)\bar{\tau}(r),$$

where

$$\bar{\tau}(r)=\tfrac{1}{2}g_+(r)-\tfrac{1}{2}\int_0^r dr'r'^{-1}\{{}^{11}K(r,r')-{}^{12}K(r,r')\}.$$

Next one subjects $\bar{\tau}(r)$ to the differential operator defined by the left-hand side of (3.9a). Again, integration by parts and use of the fact that $K_-(r,0)=0$ implies that $\bar{\tau}(r)$ is indeed a "nonsense" solution, satisfying the boundary conditions $\bar{\tau}(0)=\frac{1}{2}$ and $\dot{\bar{\tau}}(0)=-\frac{1}{2}a_1$. Therefore, by uniqueness we conclude that $\bar{\tau}(r)=\varphi(r)$ and Eq. (4.13) follows.

---

1 I. M. Gel'fand and B. M. Levitan, Isv. Akad. Nauk USSR **15**, 309 (1951) [Am. Math. Soc. Transl. **1**, 253 (1955)].
2 R. G. Newton, Phys. Rev. **100**, 412 (1955).
3 R. G. Newton, J. Math. Phys. **3**, 75 (1962).
4 P. C. Sabatier, J. Math. Phys. **8**, 905 (1967).
5 T. Regge, Nuovo Cimento **14**, 951 (1959).
6 J. M. Blatt and V. G. Weisskopf, *Theoretical Nuclear Physics* (Wiley, New York, 1963), p. 97.
7 E. U. Condon and G. H. Shortley, *The Theory of Atomic Spectra* (Cambridge U. P., Cambridge, 1967), p. 122.
8 P. C. Sabatier, J. Math. Phys. **8**, 1957 (1967).
9 P. C. Sabatier, J. Math. Phys. **9**, 1241 (1968).
10 M. Gell-Mann, Proceedings of the 1962 International High Energy Physics Conference at CERN, Geneva, 533 (1962).
11 R. G. Newton, *Scattering Theory of Waves and Particles* (McGraw-Hill, New York, 1966). We have closely followed the notation used in this reference.
12 See Ref. 11, p. 462.
13 This transformation first appeared in a paper by B. P. Desai and R. G. Newton, Phys. Rev. **129**, 1437 (1963).
14 A dot on top of a function, of a single variable, indicates the derivative of that function. We are also presenting a differentiated version of Eq. (2.49) of Ref. 9.
15 R. G. Newton, *The Complex J-Plane* (Benjamin, New York, 1964), Chap. 16.
16 We represent the element (ij) of a matrix A, by writing the number (ij) on the upper left-hand side of A, that is ${}^{ij}A$, instead of the customary notion in which (ij) is written on the lower right-hand side.
17 For more detail, see M. A. Hooshyar, thesis (Indiana University, 1970).
18 We are giving a proof similar to that of Ref. 9.
19 See p. 457 of Ref. 11.
20 The proof given here is similar to that of Ref. 4.
21 R. G. Newton, J. Math. Phys. **2**, 188 (1961).
22 F. Riesz and B. Sz.-Nagy, *Functional Analysis* (Ungar, New York, 1965), p. 176.

Duality in Generalized Ising Models and Phase Transitions without Local Order Parameters*

Franz J. Wegner†
Department of Physics, Brown University, Providence, Rhode Island 02912
(Received 29 March 1971)

It is shown that any Ising model with positive coupling constants is related to another Ising model by a duality transformation. We define a class of Ising models M_{dn} on d-dimensional lattices characterized by a number $n = 1, 2, \ldots, d$ ($n = 1$ corresponds to the Ising model with two-spin interaction). These models are related by two duality transformations. The models with $1 < n < d$ exhibit a phase transition without local order parameter. A nonanalyticity in the specific heat and a different qualitative behavior of certain spin correlation functions in the low and the high temperature phases indicate the existence of a phase transition. The Hamiltonian of the simple cubic dual model contains products of four Ising spin operators. Applying a star square transformation, one obtains an Ising model with competing interactions exhibiting a singularity in the specific heat but no long-range order of the spins in the low temperature phase.

1. INTRODUCTION

This paper deals with a general concept of duality and with phase transitions without a local order parameter.

Duality[1-5] is an inherent symmetry of the two-dimensional Ising model without crossing interaction bonds. This symmetry relates the partition function and the correlation functions[6-8] of a two-dimensional Ising model at temperature T to the partition function and the correlation functions of its dual Ising model at temperature T^*, where T^* is a decreasing function of T. In this paper the duality transformation is generalized to arbitrary Ising models with positive interaction constants (Sec. 2). This concept of duality is applied to a class of Ising models M_{dn} on d-dimensional lattices (Sec. 3). To obtain the Hamiltonian of the model M_{dn}, one takes the product of all spins located at the two ends of lines ($n = 1$), at the perimeter of surfaces ($n = 2$), and so on. Therefore, $n = 1$ describes the usual Ising model with two-spin interactions. The systems M_{dn} and $M_{d\,d-n}$ on dual lattices without external magnetic field are connected by a duality relation (Sec. 3A). For even dimensions $d = 2n$, one obtains self-dual models (models which are identical with their dual models). If there is only one singularity in the partition function of a self-dual model, then it must occur at $T = T^*$. Self-duality implies a symmetric singularity of the specific heat around the critical temperature (Sec. 3C). If an external magnetic field is present, the systems M_{dn} and $M_{d\,d-n+1}$ on dual lattices are connected by duality relations (Sec. 3A, 3C).

Most known phase transitions can be described by a local order parameter.[9-14] The models M_{dn} with $1 < n < d$ exhibit a phase transition without a local order parameter (Sec. 3B). The existence of a phase transition is indicated by a singularity in the specific heat (at least for $n = d - 1$) and by a qualitatively different asymptotic behavior of certain correlation functions at high and at low temperatures (Sec. 3B). For $n > 1$ the Hamiltonian consists of products of more than two spins. Applying the decoration,[15,16] the star triangle[3-5,17] and/or the star square[18] transformations, one reduces these models to Ising models with two-spin interactions (Sec. 2D). Thus the simple cubic dual model can be transformed to an Ising model with competing two-spin interactions. This model exhibits a singularity in the specific heat, but below the critical temperature there is no long range ordering of the spins (Sec. 4).

2. THE DUALITY TRANSFORMATION

The duality transformation for general Ising models is derived in this section. First (Sec. 2A) the Ising models with general interactions are defined, and some properties, like the degeneracy of the ground state and the spin correlation functions which vanish for all temperatures, are discussed. In Sec. 2B the duality relation for the partition function is stated and proved. The dislocation correlation functions are expressed both in terms of spin correlation functions of the original model and of the dual model in Sec. 2C. We show that a dual model exists for any Ising model (with positive interactions) and that this model can be reduced to an Ising model with only two-spin interactions and an external magnetic field (Sec. 2D).

A. The Model

The most general interaction of a system of N_s Ising spins $S(r) = \pm 1$, located at sites r of a lattice, is

$$H = -\sum_b I(b) R(b), \tag{2.1}$$

in which $I(b)$ is the coupling constant of the interaction bond labeled by the index b and

$$R(b) = \Pi_r S(r)^{\theta(r,b)}, \quad \theta(r,b) \in \{0,1\}. \tag{2.2}$$

We express all quantities which may assume two values by the two elements of the set $\{0,1\}$,

$$S(r) = (-1)^{\sigma(r)}, \quad \sigma(r) \in \{0,1\}, \tag{2.3}$$

$$R(b) = (-1)^{\rho(b)}, \quad \rho(b) \in \{0,1\}. \tag{2.4}$$

We define the field operations of addition (modulo 2)

$$0 \oplus 0 = 1 \oplus 1 = 0, \quad 0 \oplus 1 = 1 \oplus 0 = 1 \tag{2.5}$$

and multiplication (modulo 2)

$$0 \cdot 0 = 0 \cdot 1 = 1 \cdot 0 = 0, \quad 1 \cdot 1 = 1 \tag{2.6}$$

for the set $\{0,1\}$.

Then Eq. (2.2) can be written

$$\rho(b) = \oplus_r \theta(r,b)\,\sigma(r). \tag{2.7}$$

The operation symbol $\oplus$ with an index denotes summation (2.5) over this index. Let N_θ be the rank of the matrix $\theta(r,b)$. Then there are 2^{N_θ} different configurations[19] $\{\rho(b)\}$. We now restrict ourselves to systems with positive interaction constants, $I(b) > 0$. The ground states of the system (2.1) are defined by $R(b) = 1$ for all b. Therefore, the ground states are determined by the solutions $\{\sigma_0(r)\}$ of the homogeneous equations

$$\oplus_r \theta(r,b)\,\sigma_0(r) = 0 \quad \text{for all } b. \tag{2.8}$$

This system of equations has 2^{N_g} solutions with

$$N_g = N_s - N_\theta. \tag{2.9}$$

Therefore, the ground state is 2^{N_g}-fold degenerate. We associate the unitary operators

$$U\{\sigma_0\} = \Pi_r S_x(r)^{\sigma_0(r)} \tag{2.10}$$

with all ground states $\{\sigma_0(r)\}$. The operator $S_x(r)$ flips the spin at site r,

$$S_x^2(r) = 1, \qquad S_x(r)\,S(r)\,S_x(r) = -S(r). \tag{2.11}$$

The operators U commute with all operators R:

$$U\{\sigma_0\}\,R(b)\,U\{\sigma_0\}^{-1} = R(b)(-1)^{\oplus_r \theta(r,b)\sigma_0(r)} = R(b). \tag{2.12}$$

Therefore, all the operators U commute with the Hamiltonian:

$$U\{\sigma_0\}\,H\,U\{\sigma_0\}^{-1} = H. \tag{2.13}$$

A product of spins $\Pi_r S(r)^{\psi(r)}$, $\psi(r) \in \{0,1\}$, is transformed by $U\{\sigma_0\}$ into

$$U\{\sigma_0\}\,\Pi_r S(r)^{\psi(r)} U\{\sigma_0\}^{-1} = \Pi_r S(r)^{\psi(r)} (-1)^{\oplus_r \psi(r)\sigma_0(r)} \tag{2.14}$$

This product of spins commutes with all operators U if and only if

$$\oplus_r \sigma_0(r)\,\psi(r) = 0 \tag{2.15}$$

for all configurations $\{\sigma_0(r)\}$. There are $2^{N_s - N_g} = 2^{N_\theta}$ solutions $\{\psi(r)\}$, since the configurations $\{\sigma_0(r)\}$ form an N_g-dimensional linear manifold.

The product of operators $\Pi_b R(b)^{\varphi(b)}$, $\varphi(b) \in \{0,1\}$, can be expressed as a product of spin operators,

$$\Pi_b R(b)^{\varphi(b)} = \Pi_r S(r)^{\psi(r)} \tag{2.16}$$

with

$$\psi(r) = \oplus_b \theta(r,b)\,\varphi(b). \tag{2.17}$$

Since the rank of the matrix $\theta(r,b)$ is N_θ, the products of $R(b)$ in Eq. (2.16) represent 2^{N_θ} different products of spin operators characterized by the sets $\{\psi(r)\}$ of Eq. (2.17). The products of Eq. (2.16) commute with all operators U. Since there are only 2^{N_θ} different spin products which commute with all U, it follows that a product of spin operators commutes with all operators U if and only if it is a product of operators R. A product of spin operators which does not commute with all operators $U\{\sigma_0\}$ vanishes, since from

$$U \Pi_r S(r)^{\psi(r)} U^{-1} = -\Pi_r S(r)^{\psi(r)} \tag{2.18}$$

and from Eq. (2.13) it follows that

$$\langle \Pi_r S(r)^{\psi(r)} \rangle = \langle \Pi_r S(r)^{\psi(r)} U^{-1} U \rangle = \langle U \Pi_r S(r)^{\psi(r)} U^{-1} \rangle = -\langle \Pi_r S(r)^{\psi(r)} \rangle = 0. \tag{2.19}$$

Therefore, the expectation value of a product of spin operators vanishes if this product cannot be represented by a product of operators R.

It follows from Eqs. (2.16) and (2.17) that those products of operators R which are unity for each spin configuration $\{S(r)\}$ are determined by the $2^{N_b - N_\theta}$ solutions $\{\phi_0(b)\}$ of the system of homogeneous equations

$$\oplus_b \theta(r,b)\,\phi_0(b) = 0 \quad \text{for all } r. \tag{2.20}$$

Hereafter we will call any product of operators which is unity for each spin configuration the unit element.

B. The Duality Relation for the Partition Functions

We call two Ising models which are characterized by matrices $\theta(r,b)$, $\theta^*(r^*,b)$ and coupling parameters $K(b) = \beta I(b)$, $K^*(b) = \beta^* I^*(b)$ (β and β^* are the inverse temperatures of these systems) dual to each other if they fulfill these three conditions:

(a) the closure condition

$$\oplus_b \theta(r,b)\,\theta^*(r^*,b) = 0 \tag{2.21}$$

for all pairs of r, r^*,

(b) the completeness relation

$$N_\theta + N_\theta^* = N_b, \tag{2.22}$$

in which N_θ and N_θ^* are the ranks of the matrices θ and θ^* and N_b is the number of bonds b, and

(c)

$$\tanh K(b) = e^{-2K^*(b)} \tag{2.23}$$

for all bonds b.

The symmetric partition functions $Y\{K\}$ and $Y^*\{K^*\}$,

$$Y\{K\} = Z\{K\}\,2^{-(N_s+N_g)/2} \Pi_b [\cosh 2K(b)]^{-1/2}, \tag{2.24}$$

$$Z\{K\} = \sum_{\{S(r)\}} e^{-\beta H\{S\}}, \tag{2.25}$$

(and similarly for $Y^*\{K^*\}$) of two dual Ising models obey

$$Y\{K\} = Y^*\{K^*\}. \tag{2.26}$$

For the particular case of a planar Ising model without crossing bonds, this relation was proved by Wannier.[3] We prove now Eq. (2.26) for the general case, comparing the high-temperature expansion for Z with the low-temperature expansion for Z^*. From

$$Z\{K\} = \sum_{\{S(r)\}} e^{-\beta H\{S\}} = \sum_{\{S(r)\}} \Pi_b e^{K(b)R(b)} \tag{2.27}$$

one obtains

$$Z\{K\} = \Pi_b \cosh K(b) \sum_{\{\phi(b)\}} \Pi_b \tanh K(b)^{\phi(b)} \times \sum_{\{S(r)\}} \Pi_b R(b)^{\phi(b)}, \tag{2.28}$$

since

$$e^{K(b)R(b)} = \cosh K(b)[1 + R(b)\tanh K(b)] = \cosh K(b) \sum_{\{\phi(b)\}} \tanh K(b)^{\phi(b)} R(b)^{\phi(b)} \tag{2.29}$$

follows from $R(b) = \pm 1$.

If the product of the operators R in Eq. (2.28) is the unit element, then the sum over all spin configurations yields 2^{N_s}; otherwise the sum vanishes. The product of the operators R is the unit element for all sets $\{\phi_0(b)\}$ of Eq. (2.20) and only these sets. Therefore, it follows that

$$Z\{K\} = 2^{N_s} \Pi_b \cosh K(b) \sum_{\{\varphi_0(b)\}} \Pi_b \tanh K(b)^{\varphi_0(b)}. \tag{2.30}$$

The partition function $Z^*\{K^*\}$ can be written

$$Z^*\{K^*\} = \sum_{\{S(r^*)\}} e^{-\beta H^*\{S\}} = \sum_{\{S(r^*)\}} \Pi_b e^{K^*(b)R^*(b)} = \Pi_b e^{K^*(b)} \sum_{\{S(r^*)\}} \Pi_b e^{-2K^*(b)\rho^*(b)}, \tag{2.31}$$

since $R^*(b) = 1 - 2\rho^*(b)$. From the closure condition (2.21), one obtains

$$\oplus_b \theta(r, b)\, \rho^*(b) = \oplus_b \oplus_{r^*} \theta(r, b)\theta^*(r^*, b)\sigma(r^*) = 0. \tag{2.32}$$

Therefore, each set $\{\rho^*(b)\}$ obeys Eq. (2.20) with $\rho^*(b) = \phi_0(b)$. It follows that

$$Z^*\{K^*\} = \Pi_b e^{K^*(b)} \sum_{\{\phi_0(b)\}} N\{\phi_0\}\ \Pi_b e^{-2K^*(b)\phi_0(b)}. \tag{2.33}$$

Here $N\{\phi_0\}$ denotes the number of configurations $\{S(r^*)\}$ which obey

$$\phi_0(b) = \oplus_{r^*} \theta(r^*, b)\sigma(r^*) \quad \text{for all } b. \tag{2.34}$$

If for a given set $\{\phi_0(b)\}$ Eq. (2.34) has no solutions, then $N\{\phi_0\} = 0$; otherwise $N\{\phi_0\} = 2^{N_s^* - N_\theta^*}$. In particular, for $\beta^* = 0$ it follows that

$$Z^* = 2^{N_s^*} = \sum_{\{\phi_0(b)\}} N\{\phi_0\}. \tag{2.35}$$

There are $2^{N_b - N_\theta} = 2^{N_\theta^*}$ sets $\{\phi_0(b)\}$ [we used the completeness relation (2.22)]. Therefore all N obey $N\{\phi_0\} = 2^{N_s^* - N_\theta^*}$. From Eq. (2.33) one obtains

$$Z^*\{K^*\} = 2^{N_g^*} \Pi_b e^{K^*(b)} \sum_{\{\phi_0(b)\}} \Pi_b e^{-2K^*(b)\phi_0(b)}. \tag{2.36}$$

From Eqs. (2.30) and (2.36) one obtains $Y\{K\} = Y^*\{K^*\}$, Eq. (2.26), using Eq. (2.23).

If the completeness relation (2.22) is not fulfilled, but

$$N_b - N_\theta - N_\theta^* = N_m > 0, \tag{2.37}$$

and if all $K(b)$ and $K^*(b)$ are positive, then it follows from Eq. (2.33) that

$$Z^*\{K^*\} \le 2^{N_g^*} \Pi_b e^{K^*(b)} \sum_{\{\phi_0(b)\}} \Pi_b e^{-2K^*(b)\phi_0(b)}. \tag{2.38}$$

Using the analogous inequality for $Z\{K\}$, one obtains the inequality

$$2^{-N_m/2}\, Y\{K\} \le Y^*\{K^*\} \le 2^{N_m/2}\, Y\{K\}. \tag{2.39}$$

C. Dislocations

We now consider systems with magnetic dislocations. Let the operator $M(b)$ change the sign of the interaction constant $I(b)$ in the Hamiltonian. Then one obtains

$$\langle \Pi_b M(b)^{\phi^*(b)} \rangle \{K\} = \langle \Pi_b e^{-2\phi^*(b)K(b)R(b)} \rangle = \langle \Pi_b [\cosh 2K(b) - R(b)\sinh 2K(b)]^{\phi^*(b)} \rangle \tag{2.40}$$

and

$$\langle \Pi_b M(b)^{\phi^*(b)} \rangle \{K\} = Z\{(-1)^{\phi^*} K\}/Z\{K\} = Y\{(-1)^{\phi^*} K\}/Y\{K\} \tag{2.41}$$

with $\phi^*(b) \in \{0, 1\}$. From $\tanh K = e^{-2K^*}$, Eq. (2.23), it follows that

$$\tanh(-1)^{\phi^*} K = e^{-2K^* - i\pi\phi^*}. \tag{2.42}$$

Substituting Eq. (2.26) into (2.41) and using (2.42), one obtains

$$\langle \Pi_b M(b)^{\phi^*(b)} \rangle \{K\} = Y^*\{K^* + \tfrac{1}{2} i\pi\phi^*\}/Y^*\{K^*\} = i^{-\Sigma_b \phi^*(b)} \langle \Pi_b e^{i\pi\phi^*(b)R^*(b)/2} \rangle \{K^*\} = \langle \Pi_b R^*(b)^{\phi^*(b)} \rangle \{K^*\}. \tag{2.43}$$

Therefore, the expectation value of a product of dislocation operators equals the expectation value of the corresponding product of operators R^* in the dual lattice. Since R^* is a product of spin operators $S(r^*)$, one may introduce corresponding operators $\mu(r^*)$ in the original model and represent $M(b)$ by

$$M(b) = \Pi_{r^*} \mu(r^*)^{\theta^*(r^*, b)}, \qquad \mu^2(r^*) = 1. \tag{2.44}$$

Then one obtains

$$\langle \Pi_{r^*} \mu(r^*)^{\psi(r^*)} \rangle \{K\} = \langle \Pi_{r^*} S(r^*) \rangle^{\psi(r^*)} \rangle \{K^*\}. \tag{2.45}$$

For the particular case of the two-dimensional Ising model without crossing interactions this was derived by Kadanoff and Ceva.[7]

D. Construction of a Dual Ising Model: Reduction to Two-Spin Interactions

A dual Ising model exists for any given Ising model (with positive interactions) of Eq. (2.1). To obtain this dual model, one has to find a complete set of solutions $\{\phi_0(b)\}$ of Eq. (2.20). This set is complete if each solution $\{\phi_0(b)\}$ of Eq. (2.20) is a linear combination of the solutions of that set. Associate with each solution of the set a point $r^*\{\phi_0\}$. Then the lattice which is defined by the matrix

$$\theta^*(r^*\{\phi_0\}, b) = \phi_0\{b\} \tag{2.46}$$

is dual to the original lattice.

The Hamiltonian of the dual lattice may contain products of a large number of spins $S(r^*)$. We list three transformations[20] which reduce these systems with many-spin interactions to Ising models with two-spin interactions and possibly a magnetic field.

The Decoration Transformation[15,16]

The interaction $-IR_1R_2$, in which R_1 and R_2 are products of spins, can be reduced to an interaction $-I_1R_1S - I_2R_2S$, in which S is a new spin or a product of new spins

$$e^{KR_1R_2} = \tfrac{1}{2} f \sum_S e^{K_1R_1S + K_2R_2S} \tag{2.47}$$

with

$$f^2 = [\cosh(K_1 + K_2)\cosh(K_1 - K_2)]^{-1}, \tag{2.48}$$

$$\tanh K = \tanh K_1 \tanh K_2. \tag{2.49}$$

This transformation reduces products of more than three spins in the Hamiltonian to products of three spins.

A Generalized Triangle Transformation[3-5,17]

An interaction $-IS_1S_2S_3$ can be reduced to two-spin interactions and an interaction with a magnetic field by the transformation

$$\exp(KS_1S_2S_3) = \tfrac{1}{2} f \sum_S \exp[K_0S + (K_1S + K_2) \times (S_1 + S_2 + S_3) + K_3(S_1S_2 + S_1S_3 + S_2S_3)] \tag{2.50}$$

with

$$f^8 = f_0^{-1} f_1^{-3} f_2^{-3} f_3^{-1}, \quad e^{8K} = f_0^{-1} f_1^3 f_2^{-3} f_3, \tag{2.51}$$

$$e^{8K_2} = f_0 f_1 f_2^{-1} f_3^{-1}, \quad e^{8K_3} = f_0^{-1} f_1 f_2 f_3^{-1},$$
$$f_n = \cosh[K_0 + (2n-3)K_1]. \tag{2.52}$$

For the particular choice $K_1 = -K_0$ the equations simplify to

$$f^8 = [\cosh^4(2K_0)\cosh(4K_0)]^{-1},$$
$$e^{8K} = \cosh^4(2K_0)/\cosh(4K_0),$$
$$e^{8K_2} = e^{-8K_3} = \cosh(4K_0). \tag{2.53}$$

A Star Square Transformation[18]

If the Hamiltonian is invariant under flipping of all spins, then one may prefer to conserve this invariance. Products of more than four spins in the interaction can be reduced to four-spin interactions by the decoration transformation (2.47). The four-spin interactions are reduced to two-spin interactions by a star square transformation

$$\exp(KS_1S_2S_3S_4) = \tfrac{1}{2} f \sum_S \exp[K_0S(S_1 + S_2 + S_3 + S_4) + K_1(S_1S_2 + S_1S_3 + S_1S_4 + S_2S_3 + S_2S_4 + S_3S_4)] \tag{2.54}$$

with

$$e^{8K} = \cosh(4K_0)/\cosh^4(2K_0), \tag{2.55a}$$

$$e^{-8K_1} = \cosh(4K_0), \tag{2.55b}$$

$$f^8 = 1/[\cosh(4K_0)\cosh^4(2K_0)]. \tag{2.55c}$$

For real K_0 the right-hand side of Eq. (2.55a) is less than or equal to 1. Therefore, K must be negative or zero. To obtain negative K's, one may apply the decoration transformation with negative K_1 and K_2, Eq. (2.49).

Therefore, we have shown that there exists a dual Ising model (2.46) to any Ising model and that this can be reduced to an Ising model with only two-spin interactions and possibly an interaction with a magnetic field.

3. THE MODELS M_{dn} AND THEIR PROPERTIES

In this section we consider the models M_{dn}. In Sec. 3A we define the models and derive the duality relations which relate the systems M_{dn} and $M_{d\,d-n+1}$ in an external magnetic field and the duality relation between the systems M_{dn} and $M_{d\,d-n}$ without an external magnetic field. The behavior of the spin correlation functions at high and low temperatures is discussed in Sec. 3B. We prove that there is no local order parameter in the systems with $n > 1$. In Sec. 3C we discuss the thermodynamic properties of the systems.

A. The Models, Duality

We consider a d-dimensional hypervolume divided into C_d hypercells $B^{(d)}$. These are bounded by $(d-1)$-dimensional hypercells $B^{(d-1)}$ (total number C_{d-1}), these again by $(d-2)$-dimensional hypercells

$B^{(d-2)}$ (total number C_{d-2}), and so on, until we arrive at 0-dimensional hypercells which are simply the C_0 corners $B^{(0)}$ of the d-dimensional hypercells. For this original lattice L, one may construct a dual lattice L^* by placing one dual corner $B^{(0)*}$ in each original hypercell $B^{(d)}$, then connecting the dual corners by dual edges $B^{(1)*}$, each of which intersects one hypercell $B^{(d-1)}$, then connecting these dual edges by dual faces $B^{(2)*}$, each of which intersects one hypercell $B^{(d-2)}$, and so on, until we obtain the d-dimensional dual hypercells $B^{(d)*}$, each of which contains one original corner $B^{(0)}$. Denoting the number of the m-dimensional hypercells by C^*_m, we obtain

$$C^*_m = C_{d-m}, \tag{3.1}$$

since by construction there is a one-to-one correspondence of the m-dimensional dual hypercells to the $(d-m)$-dimensional original hypercells. Let us denote the intersection point of $B^{(m)}$ and its dual hypercell $B^{(d-m)*}$ by $r^{(m)} = r^{(d-m)*}$. Then a hypercell $B(r^{(m)})$ and a dual hypercell $B^*(r^{(m)})$ is associated with each point $r^{(m)}$.

Let us consider some examples. A *linear chain* ($d = 1$, Fig. 1) of points $r^{(0)} = i$ (black circles) (we denote integers by i, j, k) divides the line into one-dimensional cells (segments). The dual lattice consists of the segments between the points $r^{(1)} = i + \frac{1}{2}$ (open circles). The *square lattice* ($d = 2$, Fig. 2) consists of the squares bounded by the continuous lines; its dual lattice consists of the squares bounded by the broken lines. The corners $r^{(0)} = (i, j)$ of the original lattice are denoted by black circles, the corners $r^{(2)} = (i + \frac{1}{2}, j + \frac{1}{2})$ of the dual lattice are denoted by open circles and the edges of the original lattice intersect the edges of the dual lattice at the points $r^{(1)}$ (triangles). In Fig. 3 a cube of the original *cubic lattice* ($d = 3$) and a cube of the dual lattice are drawn. The corners $r^{(0)} = (i, j, k)$ of the original lattice are denoted by black circles and the corners $r^{(3)} = (i + \frac{1}{2}, j + \frac{1}{2}, k + \frac{1}{2})$ by open circles. The edges (continuous lines) of the original lattice and the faces of the dual lattice intersect at points $r^{(1)}$ (open squares), whereas the faces of the original lattice and the edges (broken lines) of the dual lattice intersect at points $r^{(2)}$ (black squares). We have considered only self-dual lattices, that is, lattices which are topologically equivalent to their dual lattice. Not all lattices are self-dual.

Now let us return to the general case and introduce the functions Θ and Θ^*. Let $\Theta(r^{(m-1)}, r^{(m)}) = 1$, if $r^{(m-1)}$ lies on the boundary of $B(r^{(m)})$; otherwise, $\Theta(r^{(m-1)}, r^{(m)}) = 0$. Let $\Theta^*(r^{(m)}, r^{(m-1)} = 1$ if $r^{(m)}$ lies on the boundary of $B^*(r^{(m-1)})$; otherwise

$$\Theta^*(r^{(m)}, r^{(m-1)}) = \Theta(r^{(m-1)}, r^{(m)}), \tag{3.2}$$

that is, if $r^{(m-1)}$ lies on the boundary of $B(r^{(m)})$, then $r^{(m)}$ lies on the boundary of $B^*(r^{(m-1)})$.

Since the m-dimensional boundaries of $B^{(m+1)}$ form a *closed* m-dimensional hypersurface, two m-dimensional boundaries $B^{(m)}$ of $B^{(m+1)}$ meet in each $(m-1)$-dimensional hypercell at the boundary of $B^{(m+1)}$. Therefore, one obtains

$$\oplus_{r^{(m)}} \Theta(r^{(m-1)}, r^{(m)})\Theta(r^{(m)}, r^{(m+1)}) = 0. \tag{3.3}$$

The Ising model M_{dn} on the lattice L with n dimensional bonds consists of Ising spins $S(r) = \pm 1$ at all sites $r = r^{(n-1)}$ interacting via

$$-\beta H_{dn} = K\sum_{r^{(n)}} \prod_r S(r)^{\theta(r, r^{(n)})} + h\sum_r S(r). \tag{3.4}$$

The product in the first term of the Hamiltonian runs over all spins $S(r)$ lying on the boundary of the n-dimensional hypercell $B(r^{(n)})$. For $n = 1$, the model (3.4) describes the Ising model with two-spin interactions between spins lying at the two ends of an edge and an external magnetic field

$$B = k_B \mathrm{Th}/\mu_B. \tag{3.5}$$

FIG. 1. The linear chain.

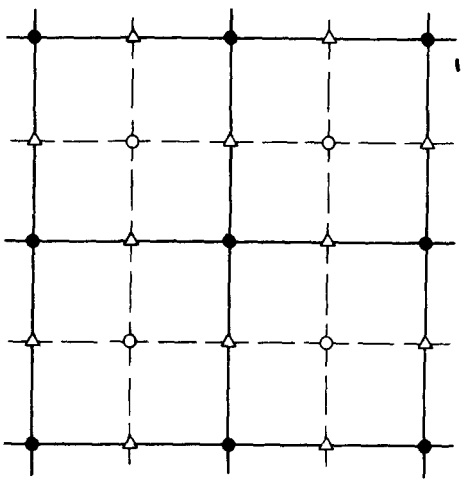


FIG. 2. The square lattice and its dual.

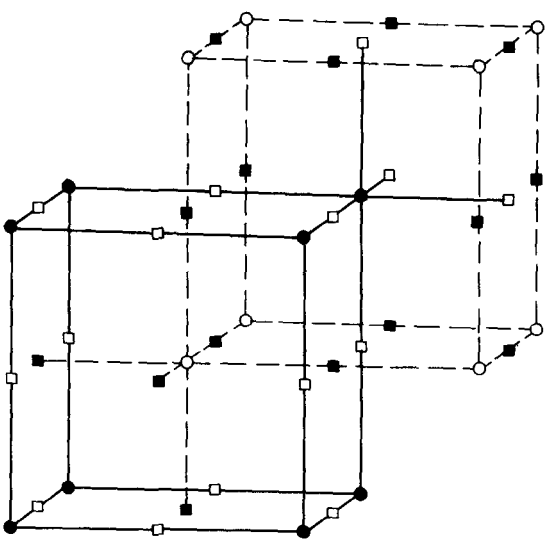


FIG. 3. The simple cubic lattice and its dual.

For the lattices considered above we obtain the Hamiltonians

$$-\beta H_{11} = K\sum_i S(i)S(i+1) + h\sum_i S(i), \tag{3.6}$$

$$-\beta H_{21} = K\sum_{ij} S(i,j)[S(i+1,j) + S(i,j+1)] + h\sum_{ij} S(i,j), \tag{3.7}$$

$$-\beta H_{31} = K\sum_{ijk} S(i,j,k)[S(i+1,j,k)+S(i,j+1,k) + S(i,j,k+1)] + h\sum_{ijk} S(i,j,k). \tag{3.8}$$

For $n = 2$ the Hamiltonian contains the products of spins lying at the boundary of the faces $B^{(2)}$,

$$-\beta H_{22} = K\sum_{ij} S(i,j+\tfrac{1}{2})S(i+1,j+\tfrac{1}{2}) \times S(i+\tfrac{1}{2},j)S(i+\tfrac{1}{2},j+1) + h\sum_{ij}[S(i,j+\tfrac{1}{2}) + S(i+\tfrac{1}{2},j)], \tag{3.9}$$

$$\begin{aligned}-\beta H_{32} = {} & K\sum_{ijk}[S(i,j+\tfrac{1}{2},k)S(i+1,j+\tfrac{1}{2},k) \\ & \times S(i+\tfrac{1}{2},j,k)S(i+\tfrac{1}{2},j+1,k) \\ & + S(i,j,k+\tfrac{1}{2})S(i+1,j,k+\tfrac{1}{2})S(i+\tfrac{1}{2},j,k) \\ & \times S(i+\tfrac{1}{2},j,k+1) + S(i,j,k+\tfrac{1}{2}) \\ & \times S(i,j+1,k+\tfrac{1}{2})S(i,j+\tfrac{1}{2},k)S(i,j+\tfrac{1}{2},k+1)] \\ & + h\sum_{ijk}[S(i,j,k+\tfrac{1}{2}) + S(i,j+\tfrac{1}{2},k) \\ & + S(i+\tfrac{1}{2},j,k)].\end{aligned} \tag{3.10}$$

For $n = 3$ the Hamiltonian contains the products of spins lying at the boundary of the volumes $B^{(3)}$:

$$\begin{aligned}-\beta H_{33} = {} & K\sum_{ijk} S(i,j+\tfrac{1}{2},k+\tfrac{1}{2})S(i+1,j+\tfrac{1}{2},k+\tfrac{1}{2}) \\ & \times S(i+\tfrac{1}{2},j,k+\tfrac{1}{2})S(i+\tfrac{1}{2},j+1,k+\tfrac{1}{2}) \\ & \times S(i+\tfrac{1}{2},j+\tfrac{1}{2},k)S(i+\tfrac{1}{2},j+\tfrac{1}{2},k+1) \\ & + h\sum_{ijk}[S(i,j+\tfrac{1}{2},k+\tfrac{1}{2}) \\ & + S(i+\tfrac{1}{2},j,k+\tfrac{1}{2}) + S(i+\tfrac{1}{2},j+\tfrac{1}{2},k)].\end{aligned} \tag{3.11}$$

In general, the model M_{dn} on a hypercubic lattice consists of $N_s = \binom{d}{n-1}N$ Ising spins located at the centers of the $(n-1)$-dimensional hypercubes. (N is the number of the d-dimensional hypercubes). The Hamiltonian consists of the sum of the products of the $2n$ spins at the $(n-1)$-dimensional hypersurfaces of the $N_b = \binom{d}{n}N$ hypercubes $B^{(n)}$. Let us denote a subset of n unit vectors e_i along the main axes by E_n; then the model M_{dn} for the d-dimensional hypercubic lattice is defined by

$$-\beta H_{dn} = K\sum_{r^{(0)},E_n} R(r^{(0)},E_n) + h\sum_{r^{(0)},E_{n-1}} S(r^{(0)} + v(E_{n-1})) \tag{3.12}$$

with

$$v(E_n) = \tfrac{1}{2}\sum_{e\in E_n} e, \tag{3.13}$$

$$R(r^{(0)},E_n) = \prod_{e\in E_n} S(r^{(0)} + v(E_n) - \tfrac{1}{2}e) \times S(r^{(0)} + v(E_n) + \tfrac{1}{2}e). \tag{3.14}$$

Similarly one defines the Ising model M^*_{dn} on the dual lattice L^*. The Ising spins $S(r^*) = \pm 1$ are located at the sites $r^* = r^{(n-1)*} = r^{(d-n+1)}$ and interact via

$$-\beta^* H^*_{dn} = K^*\sum_{r^{(n)}*}\prod_{r^*} S(r^*)^{\mathcal{O}^*(r^*,r^{(n)*})} + h^*\sum_{r^*} S(r^*). \tag{3.15}$$

We now show that the models M_{dn} and $M^*_{d,d-n+1}$ are related by the duality relation

$$Y_{dn}(K,h) = Y^*_{d,d-n+1}(K^*,h^*) \tag{3.16}$$

with

$$\tanh K = e^{-2h^*}, \qquad \tanh h = e^{-2K^*}. \tag{3.17}$$

If we label the interaction of the spin $S(r^{(n-1)})$ with the external magnetic field by $b(r^{(n-1)})$ and the interaction of the spins on the boundary of $B(r^{(n)})$ by $b(r^{(n)})$, then we have

$$\theta(r^{(n-1)},b(r^{(n)})) = \Theta(r^{(n-1)},r^{(n)}), \tag{3.18a}$$

$$\theta(r^{(n-1)},b(r^{(n-1)'})) = \delta_{r^{(n-1)}r^{(n-1)'}}, \tag{3.18b}$$

$$\theta^*(r^{(n)},b(r^{(n)'})) = \delta_{r^{(n)}r^{(n)'}}, \tag{3.19a}$$

$$\theta^*(r^{(n)},b(r^{(n-1)})) = \Theta^*(r^{(n)},r^{(n-1)}). \tag{3.19b}$$

Substituting Eqs. (3.18) and (3.19) into Eq. (2.21) and using Eq. (3.2), we find that the closure condition is fulfilled. From Eqs. (3.18b) and (3.19a) it follows that $N_\theta = N_s$ and $N^*_\theta = N^*_s$. Since $N_b = N_s + N^*_s$, the completeness relation is fulfilled.

We now compare the models M_{dn} and $M^*_{d\,d-n}$ without external magnetic fields. Then the bonds are connected with the sites $r^{(n)}$ by Eq. (3.18a) and

$$\theta^*(r^{(n+1)},b(r^{(n)})) = \Theta^*(r^{(n+1)},r^{(n)}) = \Theta(r^{(n)},r^{(n+1)}). \tag{3.20}$$

From Eq. (3.3) it follows that the closure condition is fulfilled. We now discuss the completeness relation (2.22). In the Appendix we derive relations between the N's, C's, and the topology of the lattice. Here we summarize the results: The exponents N_g and N^*_g of the orders of the degeneracy 2^{N_g} and $2^{N^*_g}$ of the models (3.4) and (3.15) are

$$N_g = t_g + \sum_{m=0}^{n-2}(-1)^{n-m}C_m, \tag{3.21}$$

$$N^*_g = t^*_g + \sum_{m=n+2}^{d}(-1)^{m-n}C_m, \tag{3.22}$$

in which t_g and t^*_g depend only on the boundary conditions and on n. From a generalization of Euler's theorem[21]

$$\sum_{m=0}^{d}(-1)^m C_m = t, \tag{3.23}$$

in which t depends only on the topology (boundary conditions), from

$$N_s = C_{n-1}, \quad N_b = C_n, \quad N_s^* = C_{n+1} \tag{3.24}$$

and from

$$N_m = N_g - N_s + N_b - N_s^* + N_g^*, \tag{3.25}$$

which is derived from Eqs. (2.9) and (2.37), it follows that

$$N_m = t_g + t_g^* + (-)^{d-n} t. \tag{3.26}$$

Therefore, N_m depends only on the topology of the system and on n. For a d-dimensional hypersurface wrapped on a $(d+1)$-dimensional hypersphere, one obtains $N_m = 0$ for $1 \leq n \leq d-1$. Therefore, the duality relation

$$Y_{dn}(K, 0) = Y_{d\,d-n}^*(K^*, 0), \tag{3.27}$$

with

$$\tanh K = e^{-2K^*}, \tag{3.28}$$

holds for this boundary condition. For the two-dimensional Ising model ($d = 2, n = 1$) this was shown in Ref. 3. For systems with periodic boundary conditions one obtains $N_m = \binom{d}{n}$. In the thermodynamic limit the factors $2^{Nm/2}$ in Eq. (2.39) can be neglected, and, using Eqs. (2.24), (2.25), and

$$-\beta F(K) = \ln Z(K), \tag{3.29}$$

we obtain for the free energy

$$\beta^* F_{d\,d-n}^*(K^*) = \beta F_{dn}(K) - \tfrac{1}{2}(N_g^* + N_s^* - N_g - N_s) \times \ln 2 + \tfrac{1}{2} N_b \ln \sinh 2K. \tag{3.30}$$

B. Correlation Functions

In this section we discuss the behavior of the spin correlation functions of the systems M_{dn} without an external magnetic field. We showed in Sec. 2 that an operator

$$U\{\sigma_0\} = \Pi_r S_x(r)^{\sigma_0(r)} \tag{2.10}$$

commutes with the Hamiltonian H and all operators R if

$$\oplus_r \theta(r, b)\,\sigma_0(r) = 0 \quad \text{for all } b. \tag{2.8}$$

The only solution for $n = 1$ besides the trivial solution $\sigma_0(r) = 0$ is

$$\sigma_0(r) = 1. \tag{3.31}$$

For $n > 1$ we obtain solutions

$$\sigma_0(r) = \Theta(r^{(n-2)}, r), \tag{3.32}$$

which can be verified using Eqs. (3.3) and (3.18a). Therefore, each operator R is invariant under flipping of all spins lying on the $(n-1)$-dimensional hypercells $B(r)$, which meet in the hypercell $B(r^{(n-2)})$. This leads to the high degeneracy 2^{N_g}, where N_g is given by Eq. (3.21). Since, for $r \neq r'$, there exists a neighbor $r^{(n-2)}$ of r with $\Theta(r^{(n-2)}, r) = 1$ and $\Theta(r^{(n-2)}, r') = 0$, we obtain from Eq. (2.19)

$$\langle S(r) S(r') \rangle = \delta_{rr'}. \tag{3.33}$$

Therefore, there is no long-range spin autocorrelation at any temperature. The only products of spins whose expectation values do not vanish can be represented by a product of operators R. These products are the products of all spins lying on the $(n-1)$-dimensional boundary of an n-dimensional hypervolume which consists of n-dimensional hypercells.

We now consider the long-range behavior of $\langle \Pi_r S(r) \rangle$ of the model M_{dn}, Eq. (3.12), where the spins of the product lie on the boundary of an n-dimensional hypercube. From the high temperature expansion it follows that

$$\langle \Pi_r S(r) \rangle = [\tanh K + 2(d-n)(\tanh K)^{1+2n} + \cdots]^v \quad \text{for } n > 1. \tag{3.34}$$

and

$$\langle \Pi_r S(r) \rangle = \tfrac{1}{2}\{\tanh K + [2(d-1)]^{1/2} \times (\tanh K)^2 + \cdots\}^v + \tfrac{1}{2}\{\tanh K - [2(d-1)]^{1/2} (\tanh K)^2 + \cdots\}^v \quad \text{for } n = 1, \tag{3.35}$$

where v is the volume of the hypercube (for $n = 1$, v is the distance between the two spins; for $n = 2$, v is the area of the square spanned by the spins). From the low temperature expansion one obtains

$$\langle \Pi_r S(r) \rangle = (1 - e^{-4(d-n+1)K} + \cdots)^f \quad \text{for } n < d, \tag{3.36}$$

$$\langle \Pi_r S(r) \rangle = (1 - 2e^{-2K} + \cdots)^v \quad \text{for } n = d, \tag{3.37}$$

where f is the hyperarea of the hypercube (for $n = 1$, f is the number of the ends of the line, that is, $f = 2$; for $n = 2$, f is the perimeter of the square). Therefore, we deduce that the behavior of these correlation functions in the limit of large hypercubes is different in the low and the high temperature phases, and we expect

$$\langle \Pi_r S(r) \rangle \propto \begin{cases} \exp[-v/v_0(T)] & \text{for } T > T_c, n < d, \end{cases} \tag{3.38}$$

$$\begin{cases} \exp[-f/f_0(T)] & \text{for } T < T_c, n < d. \end{cases} \tag{3.39}$$

We attribute the qualitatively different assymptotic behavior in both temperature regions to different states of the system above and below a critical temperature T_c.

For $n = d - 1$ the different behavior in both temperature regions becomes more evident if one makes use of the duality relation for dislocations, Eq. (2.43). One obtains

$$\langle \Pi_r S(r) \rangle \{K\} = \langle \Pi_b R(b) \rangle \{K\} = \langle \Pi_b M^*(b) \rangle \{K\}, \tag{3.40}$$

where the product runs over all b's in the $(d-1)$-dimensional hypercube. The expectation value on the right-hand side of Eq. (3.40) is to be taken in the model M^*_{d1}. The logarithm of this expectation value is proportional to the change in free energy due to the dislocations. This free energy is proportional to the $(d-2)$-dimensional hyperarea of the boundary in the disordered state ($T^* > T_c^*$, that is, for $T < T_c$), and it is proportional to the $(d-1)$-dimensional hypervolume in the ordered state of the dual system ($T^* < T_c^*$, that is, for $T > T_c$). This is in agreement with Eqs. (3.38) and (3.39).

We now compare the systems M_{dn} and $M_{d+1,n}$ on a hypercubic lattice. From the theorem of Griffiths generalized by Kelly and Sherman[22] it follows that any expectation value $\langle \Pi_r S(r) \rangle$ in the system M_{dn} is less or equal to the expectation value in the system $M_{d+1,n}$,

$$\langle \Pi_r S(r) \rangle_d \leq \langle \Pi_r S(r) \rangle_{d+1}, \tag{3.41}$$

since the $(d+1)$-dimensional system consists of layers of the system M_{dn} plus an additional interaction between the layers. Therefore, if this expectation value shows the long-range behavior, Eq. (3.39), for M_{dn}, then this long-range behavior is also apparent in $M_{d+1,n}$, and we obtain $T_{c,dn} \leq T_{c,d+1,n}$, that is,

$$K_{c,d+1,n} \leq K_{c,dn}. \tag{3.42}$$

The systems M_{dn} with $n > 1$ exhibit an unusually high ground-state entropy $S_0 \propto N$. Taking

$$S(r^{(0)} + v(E_{n-1})) = 1, \quad \text{if } e_d \in E_{n-1}, \tag{3.43}$$

in the hypercubic models (3.12), we may eliminate all spins with half-valued r_d-component. These systems (we denote them by M'_{dn}) consist of $N_s = N\binom{d-1}{n-1}$ spins and have a much smaller degeneracy,

$$N_g = \frac{N}{N_d}\binom{d-2}{n-2} + \binom{d-2}{n-1},$$

N_d denoting the length of the periodicity in the r_d direction.

For $n = d$ the system disintegrates into linear chains. For $n = 1$ the system is unchanged. For $d = 3, n = 2$ one obtains the Hamiltonian

$$\begin{aligned} -\beta H'_{32} = K \sum_{ijk} [&S(i, j+\tfrac{1}{2}, k) S(i+1, j+\tfrac{1}{2}, k) \\ &\times S(i+\tfrac{1}{2}, j, k) S(i+\tfrac{1}{2}, j+1, k) \\ &+ S(i+\tfrac{1}{2}, j, k) S(i+\tfrac{1}{2}, j, k+1) \\ &+ S(i, j+\tfrac{1}{2}, k) S(i, j+\tfrac{1}{2}, k+1)]. \end{aligned} \tag{3.44}$$

These systems obey the closure condition (2.21) if one chooses the model $M^*_{d,d-n}$ on the hypercubic lattice as the dual model. One obtains

$$N_m = \frac{N}{N_d}\binom{d-2}{n-2} + \binom{d-1}{n} + \binom{d-2}{n-1}. \tag{3.45}$$

Therefore, in the thermodynamic limit $N \to \infty$, $N_d \to \infty$, the duality relation (3.30) holds, and the free energies of M_{dn} and M'_{dn} show the same nonanalyticities. In the systems M'_{dn} the spins separated by a vector pointing in the e_d direction are correlated. At high temperatures one obtains

$$\begin{aligned} \langle S(r) S(r + r_d e_d) \rangle = [&\tanh K + 2(d-n) \\ &\times (\tanh K)^{1+2n} + \cdots]^{|r_d|} \quad \text{for } n > 1, \end{aligned} \tag{3.46}$$

and for low temperatures it follows that

$$\begin{aligned} \langle S(r) S(r + r_d e_d) \rangle = (1 - e^{-4(d-n+1)K} \\ + \cdots)^{2+2(n-1)|r_d|} \quad \text{for } n < d. \end{aligned} \tag{3.47}$$

Therefore, we expect an exponential decay of the correlation function for large r_d at all temperatures if $n > 1$. Here again we find no long range order.

Absence of a Local Order Parameter

A second-order phase transition with local order parameter is characterized as follows: Let us add local operators $\psi_i(r)$ to the Hamiltonian H_0,

$$-\beta H = K H_0 + \sum_{ir} h_i \psi_i(r).$$

Then there is a discontinuity of the first-order derivatives with respect to h of the free energy $F(K, \{h\})$ along a ν'-dimensional hypervolume known as a first-order transition line in the ν-dimensional (K, h_i) space. This hypervolume is bounded by a $(\nu'-1)$-dimensional λ-hypersurface commonly known as λ-point or λ-line, where the second-order phase transition takes place. Any local operator $\psi(r) = \sum h_i \psi_i(r)$ with a discontinuity of $\sum_r \langle \psi(r) \rangle = -\beta \Sigma_i h_i \partial F / \partial h_i$ along the first-order transition can be considered as an order parameter. In the homogeneous phase the limit

$$\lim_{r \to \infty} [\langle \psi(0) \psi(r) \rangle - \langle \psi(0) \rangle \langle \psi(r) \rangle] \tag{3.48}$$

vanishes. If the expectation values in expression (3.48) are averaged over *all* states along a first-order transition, then the limit (3.48) does not vanish. In the Ising model ($n = 1$) with ferromagnetic interactions $\psi(r) = S(r)$ is such a local operator. For $T = 0$ we have $\langle S(0) S(r) \rangle = 1$, whereas $\langle S(r) \rangle = 0$. In the models M_{dn} with $n > 1$ there is no first-order transition for $T < T_{c,dn}$, $h_i = 0$ associated with a local order

parameter $\langle\psi(r)\rangle$ if we confine ourselves to operators which are polynomials of spin operators located in a finite region about r. We can see this as follows: Any product of spins $S(r)$ which do not lie on a closed $(n-1)$-dimensional hypersurface of hypercells $B^{(n-1)}$ gives vanishing contributions for a sufficiently large distance r. Therefore, we may confine ourselves to expressions for ψ, which are polynomials of $R(b)$,

$$\psi(r) = P(r; R(b)). \tag{3.49}$$

Applying Eqs. (2.40) and (2.43), one obtains

$$[\langle\psi(0)\psi(r)\rangle - \langle\psi(0)\rangle\langle\psi(r)\rangle](K) = [\langle\psi^*(0)\psi^*(r)\rangle - \langle\psi^*(0)\rangle\langle\psi^*(r)\rangle](K^*) \tag{3.50}$$

with

$$\psi^*(r) = P(r; \cosh 2K^* - R^*(b)\sinh 2K^*). \tag{3.51}$$

Therefore, the correlation of the ψ's in the model M_{dn} below T_c is related to the correlation of the ψ^*'s in the dual model $M^*_{d\,d-n}$ above T_c^*. According to the cluster property of the Ising model, proved rigorously by Ruelle[23] for Ising models with $n = 1$, the right-hand side of Eq. (3.49) vanishes for $r \to \infty$. Therefore, there is no first-order transition characterized by a local order parameter in the models M_{dn} with $n > 1$ along the K axis.[24]

C. Thermodynamic Properties

In this section we consider the thermodynamic properties of the systems M_{dn}.

The Linear Chain M_{11}

The partition function of the linear chain (3.6) of N Ising spins with nearest neighbor interaction and the periodic boundary condition $S(N+1) = S(1)$ can be calculated[5] explicitly:

$$Z(K,h) = \mathrm{tr}\begin{pmatrix} e^{K+h} & e^{-K} \\ e^{-K} & e^{K-h} \end{pmatrix}^N. \tag{3.52}$$

With Eq. (2.24),

$$Y(K,h) = f^N Z(K,h), \quad f = (2\cosh 2K \cosh 2h)^{-1/2}, \tag{3.53}$$

one obtains

$$Y(K,h) = \mathrm{tr}\begin{pmatrix} fe^{K+h} & fe^{-K} \\ fe^{-K} & fe^{K-h} \end{pmatrix} = F_N(fe^K\cosh h, f^2(e^{2K} - e^{-2K})). \tag{3.54}$$

The first argument of F_N is half the trace of the 2×2 matrix; the second argument is its determinant. It follows that

$$F_N(t,d) = (t + \sqrt{t^2 - d})^N + (t - \sqrt{t^2 - d})^N, \tag{3.55}$$

with

$$t = [2(1 + e^{-4K})(1 + \tanh^2 h]^{-1/2}, \tag{3.56}$$

$$d = \tfrac{1}{2}(1 - e^{-4K})(1 - \tanh^2 h)(1 + e^{-4K})^{-1} \times (1 + \tanh^2 h)^{-1}. \tag{3.57}$$

Since the linear chain is a self-dual lattice, one obtains from Eq. (3.16) that

$$Y_{11}(K,h) = Y_{11}(K^*,h^*) \tag{3.58}$$

with Eq. (3.17), which is fulfilled since t and d, Eqs. (3.56) and (3.57), are invariant under this transformation.

The Models M_{dd}

The partition functions $Z_{dd}(K,0)$ of the models M_{dd} without external magnetic field can be calculated from the duality relation (3.16), (3.17):

$$Y_{dd}(K,0) = Y^*_{d1}(\infty, h^*), \quad \tanh h^* = e^{-2K}. \tag{3.59}$$

Since in the model M^*_{d1} all spins are coupled by a two-spin interaction with infinite K^*, only the two configurations $\{S(r^*) = 1\}$ and $\{S(r^*) = -1\}$ contribute

$$Z^*_{d1}(K^*,h^*) \sim \exp(N_b K^* + N_s^* h^*) + \exp(N_b K^* - N_s^* h^*). \tag{3.60}$$

It follows that

$$Z_{dd}(K,0) = 2^{N_s}[(\cosh K)^{N_b} + (\sinh K)^{N_b}]. \tag{3.61}$$

The partition functions of the models M_{dd} are analytic in K for all finite K and $h = 0$. Since the Ising model $M^*_{d1}, d > 1$, shows a phase transition for $h^* = 0$ at $K^* = K^*_{c,d1}$, a nonanalyticity is apparent in the partition function $Y_{dd}(K,h)$ for $K \to \infty$ at $h = -\frac{1}{2}\ln[\tanh(K^*_{c,d1})]$.

The Models M_{dn} with $n < d$ without External Magnetic Field

The nonanalyticity which is apparent in the free energy F_{d1} at the critical $K_{c,d1} = \beta_c I$ (β_c is the inverse critical temperature) also occurs in the free energy F^*_{dd-1} [Eq. (3.30)]. Since the correlation functions, Eqs. (3.38) and (3.39), show a qualitatively different assymptotic behavior at low and high temperatures, we expect a phase transition for all infinite systems M_{dn} with $1 \le n < d$ at some $K = K_{c,dn}$ accompanied by a nonanalyticity of the free energy. The critical K's of the model and its dual model are related by Eq. (3.28), which can be cast in the symmetric form

$$\sinh 2K_{c,dn}\sinh 2K^*_{c,d\,d-n} = 1. \tag{3.62}$$

In particular, for self-dual lattices like the hypercubic lattice, one obtains

$$\sinh 2K_{c,dn}\sinh 2K^*_{c,d\,d-n} = 1. \tag{3.63}$$

TABLE I. Critical parameters of some three-dimensional Ising models and their dual models.

original lattice	diamond	simple cubic	body-centered cubic	face-centered cubic
K_c	0.3698	0.2217	0.1575	0.1021
K_c^*	0.5195	0.7613	0.9284	1.1426
E_c/E_0	0.432	0.3284	0.270	0.245
E_c^*/E_0^*	0.937	0.9495	0.964	0.971
S_c/S_∞	0.737	0.808	0.845	0.853
$(S_c^* - S_0^*)/(S_\infty^* - S_0^*)$	0.100	0.092	0.072	0.063

For a self-dual model ($d = 2$, self-dual lattice) it follows that

$$K_{c,d,d/2} = \tfrac{1}{2}\ln(\sqrt{2} + 1). \tag{3.64}$$

We derived the inequality $K_{c,d+1,n} \leq K_{c,dn}$ for hypercubic lattices, Eq. (3.42). From Eq. (3.63) one obtains

$$K_{c,d,n-1} \leq K_{c,d+1,n}\,, \tag{3.65}$$

and from Eq. (3.42) and (3.65) it follows that

$$K_{c,d,n-1} \leq K_{c,d,n}\,. \tag{3.66}$$

The critical temperature of the hypercubic systems is a decreasing function of n.

Since the duality relation (3.30) relates the free energy F_{dn} at high temperatures to the free energy $F^*_{d\,d-n}$ of its dual model at low temperatures, we deduce that the critical exponent $\alpha_{d\,d-n}$ of the specific heat of the model $M^*_{d\,d-n}$ above T_c^* is given by the critical exponent α'_{dn} of the model M_{dn} below T_c and vice versa:

$$\alpha_{d\,d-n} = \alpha'_{dn}\,. \tag{3.67}$$

Therefore, any asymetry in the specific heat of the model M_{dn} near T_c is also apparent in the specific heat of the dual model, but the high temperature and the low temperature regions are interchanged. Self-dual systems exhibit a symmetric singularity of the specific heat around the critical temperature.

From the thermodynamic relation

$$E = \frac{\partial(\beta F)}{\partial \beta} = \frac{\partial(KF)}{\partial K}, \tag{3.68}$$

it follows that

$$E^*(K^*)/E_0^* = \cosh(2K) - \sinh(2K)\,E(K)/E_0, \tag{3.69}$$

in which E_0 denotes the ground state energy $E_0 = -IN_b$. Therefore, using Eqs. (3.30), (3.68), and

$$F = E - TS, \tag{3.70}$$

one is able to calculate the energy E^* and the entropy S^* of the dual model from E and S. From the critical parameters[13] of the Ising model on the diamond, the simple cubic, the body-centered cubic, and the face-centered cubic lattice, we have calculated the critical parameters of their dual models. The results are listed in Table I. The binding energies of the dual models at critical temperature are unusually large [for example, 95% of the ground state energy for the model (3.10)]. This is in agreement with the unusually low critical entropy. For the model (3.10) we obtain $S_c^*/k_BN = 0.82$ which is to be compared with the zero temperature entropy $S_0^*/k_BN = \ln 2 = 0.69$ and the entropy at infinite temperature $S_\infty^*/k_BN = 3 \ln 2 = 2.08$.

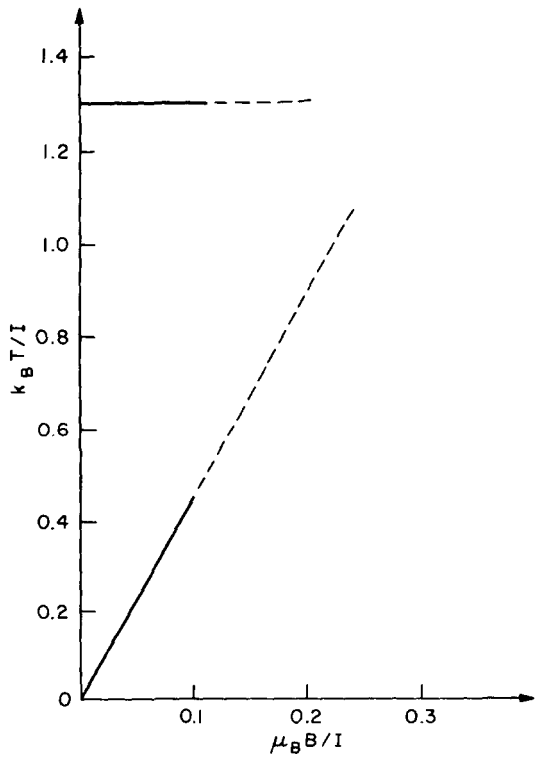


FIG. 4. Phase diagram of the system (3.10).

The Systems M_{dn} with $1 < n < d$ in an External Magnetic Field

Near the critical temperature the Ising models M_{d1} are very sensitive to an external magnetic field, since the spins exhibit a long range correlation. This does not apply to systems M_{dn} with $n > 1$. Therefore, a phase transition line $K = K_c(h)$

is expected. This function $K_c(h)$ can be calculated for small h if one assumes that the nonanalytic part of the free energy depends on $K - K_c(h)$ only[25,26]:

$$F_{\text{sing}}(K,h) = F_{\text{sing}}(K - K_c(h)). \tag{3.71}$$

From $KF = -I \ln Z$, one obtains Eq. (3.68). The mth derivative of KF with respect to h can be expressed as spin correlation functions involving up to m spins. Only products of spins which can be expressed as products of $R(b)$ yield nonvanishing expectation values. Since $R(b)$ is a product of $2n$ spins for the hypercubic systems, one obtains for $m < 2n$ only constant contributions of the type $\langle S^2(r)S^2(r')\cdots\rangle$. For $m = 2n$ expectation values $\langle R(b)\rangle$ also occur yielding $(2n)!\,E$:

$$\frac{\partial^m (KF)}{\partial h^m}\Big|_{h=0} = \begin{cases} \text{const} & \text{for } m < 2n \\ \text{const} + (2n)!\,E & \text{for } m = 2n. \end{cases} \tag{3.72}$$

From Eq. (3.71) and (3.72) it follows for the hypercubic systems that

$$K_{c,dn}(h) = K_{c,dn}(0) - h^{2n} + \cdots. \tag{3.73}$$

From the duality relation, Eqs. (3.16), (3.17), one obtains for large K the phase transition line

$$h_{c,dn}(K) = K_{c,d\,d-n}(0) - \sinh 2K_{c,d\,d-n}(0)e^{-4nK} + \cdots. \tag{3.74}$$

The reduced critical temperature $K_c^{-1} = k_B T_c/I$ is plotted as a function of the reduced magnetic field $h/K_c = \mu_B/I$ in Fig. 4 for the cubic model M_{32}, Eq. (3.10).

In an external magnetic field the systems M_{dn} with $d = 2n - 1$ on self-dual lattices are self-dual [Eq. (3.16)].

4. PHASE TRANSITION IN AN ISING MODEL WITH COMPETING INTERACTIONS

In this section we describe an Ising model with competing two-spin interactions. For special values of temperature and interaction parameters this model is related to the model (3.10) by the star square transformation (2.54). This system shows a singularity in the specific heat, but it shows no long range order below the critical temperature.

As in model (3.10) spins are located at the centers $r^{(1)}$ of the edges of the cubes (open squares in Fig. 3). Moreover, spins are located at the centers $r^{(2)}$ of the faces of these cubes (black squares in Fig. 3). We assume an interaction strength I_1 for nearest neighbor pairs $S(r^{(1)})$ and $S(r^{(2)})$, an interaction strength I_2 for next nearest neighbor pairs of spins $S(r^{(1)})$ and $S(r^{(1)\prime})$, and an interaction strength I_3 for pairs of spins $S(r^{(1)})$ lying opposite a spin $S(r^{(2)})$ (Fig. 5). Denoting the central spin $S(r^{(2)})$ of a face by S_5 and its four nearest neighbor spins by S_1, S_2, S_3, S_4, then the Hamiltonian H' of our model is the sum over all faces

$$\begin{aligned} H' = -\sum [&I_1 S_5(S_1 + S_2 + S_3 + S_4) \\ &+ I_2(S_1S_2 + S_2S_3 + S_3S_4 + S_4S_1) \\ &+ I_3(S_1S_3 + S_2S_4)]. \end{aligned} \tag{4.1}$$

We discuss the ground state of this model. The system is invariant under simultaneous reversal of I_1 and S_5. We assume I_1 to be positive. The ground state depends on the ratios I_2/I_1 and I_3/I_1. In Fig. 6 we plot the phase diagram at zero temperature. In region 0 the system is ferromagnetic, that is, all spins $S(r^{(1)})$ point in the same direction. In region 1 one of the four spins $S(r^{(1)})$ of a face points in one direction, all three other spins of the face point in the opposite direction. In region 2 a

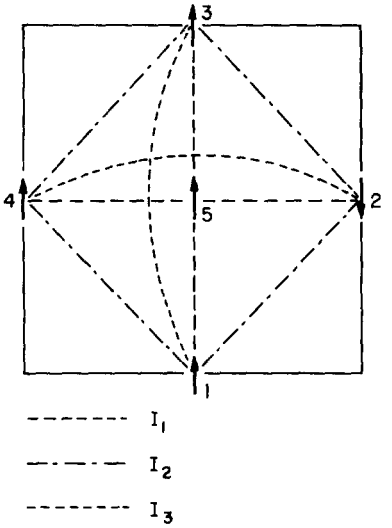


FIG. 5. The interactions in a face of the Ising model (4.1).

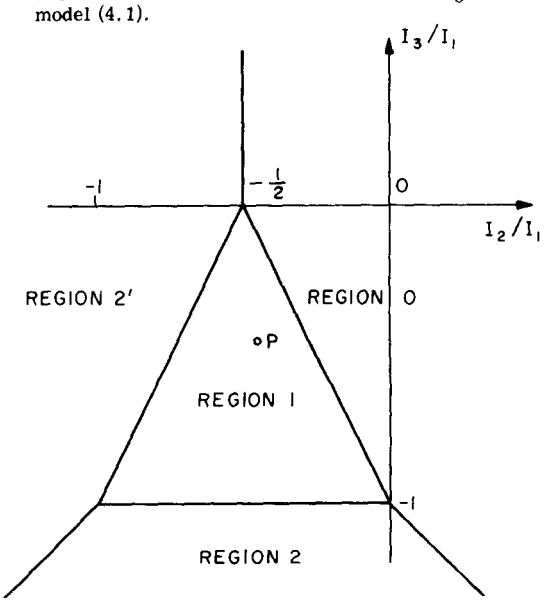


FIG. 6. Phase diagram for the model (4.1) at zero temperature.

pair of neighbored spins $S(r^{(1)})$ points in one direction, the other pair of spins $S(r^{(1)})$ of this face points in the other direction. In region 2′ two opposite spins $S(r^{(1)})$ at a face should point in one direction, the other pair in the other direction, but such an ordering is not possible in three dimensions. Therefore, in this region the ground state cannot be determined by looking merely for the ground state of one face.

Here we are interested in region 1. The ground states of system (4.1) and (3.10) for negative I are the same. The spins S_5 are determined by the surrounding spins S_1, S_2, S_3, S_4. Since the partition function of system (3.10) is invariant under change of sign of I, we obtain the ground state entropy of the system (4.1) in region 1,

$$S_0 = Nk_B \ln 2. \tag{4.2}$$

From the star square transformation (2.54) we find that the partition function Z' of system (4.1) and the partition function $\hat{Z}$ of the Hamiltonian

$$\hat{H} = -\sum [\hat{I}_1 S_1 S_2 S_3 S_4 + \hat{I}_2 (S_1 S_2 + S_2 S_3 + S_3 S_4 + S_4 S_1) + \hat{I}_3 (S_1 S_3 + S_2 S_4)] \tag{4.3}$$

are related by

$$Z'(K_1, K_2, K_3) = (2e^{2\hat{K}-\hat{K}_1})^{3N} \hat{Z}(\hat{K}_1, \hat{K}_2, \hat{K}_3), \tag{4.4}$$

where

$$\hat{K}_2 = K_2 + \hat{K}, \quad \hat{K}_3 = K_3 + \hat{K}, \tag{4.5}$$

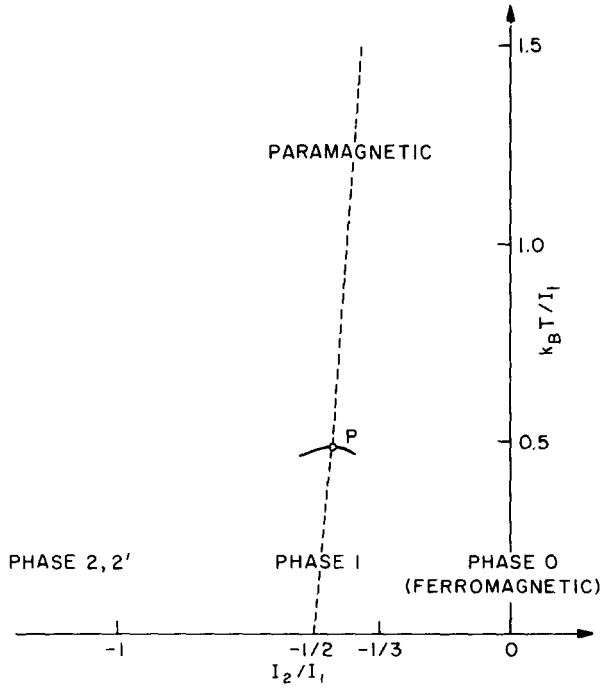


FIG. 7. Phase diagram for the model (4.1) for $I_2 = I_3$. Along the broken line the free energy can be calculated from that of the simple cubic Ising model. The heavy line denotes the phase transition line from Eq. (4.14).

$$\cosh 4K_1 = e^{+8\hat{K}}, \cosh 2K_1 = e^{2\hat{K}-2\hat{K}_1}. \tag{4.6}$$

Along the line $\hat{K}_2 = \hat{K}_3 = 0$ (broken line in Fig. 7) the partition function can be expressed in terms of that of the simple cubic Ising model. In particular from the critical singularity of the partition function of the simple cubic Ising model[13] at $I/k_B T_c = 0.2217$ we obtain a singularity of the partition function Z' at $K_{1c} = 2.039, K_{2c} = K_{3c} = -0.9344$, that is, for $I_2/I_1 = I_3/I_1 = -0.4582$, $k_B T_c/I_1 = 0.4904$ (point P of Figs. 6 and 7).

Now let us expand $\hat{Z}(\hat{K}_1, \hat{K}_2, \hat{K}_3)$ into powers of $\hat{K}_2$ and $\hat{K}_3$:

$$\ln \hat{Z}(\hat{K}_1, \hat{K}_2, \hat{K}_3) = \ln Z_{32}(\hat{K}_1) + \sum_{ij} a_{ij} \hat{K}_2^i \hat{K}_3^i. \tag{4.7}$$

The coefficients a_{ij} can be expressed in terms of the spin correlation functions of system (3.10):

$$a_{10} = a_{01} = 0, \tag{4.8}$$

$$\begin{aligned} a_{20} &= \tfrac{1}{2} \partial^2 \ln \hat{Z}/\partial \hat{K}_2^2 \\ &= \tfrac{1}{2} \sum \langle (S_1 S_2 + S_2 S_3 + S_3 S_4 + S_4 S_1)^2 \rangle \\ &= 2 \sum (1 + \langle S_1 S_2 S_3 S_4 \rangle) = 6N - 2E/I, \end{aligned}$$

$$a_{02} = 3N - E/I, \quad a_{11} = 0. \tag{4.9}$$

In general a $2n$th or $(2n+1)$th derivative of $\ln Z$ can be expressed in terms of cumulants of at most n operators $R(b)$ or products of $R(b)$'s. We expect that such a cumulant shows a singularity at the critical temperature of the form $\epsilon^{2-\alpha-n}$ [with $\epsilon = (T - T_c)/T_c$], since such cumulants occur in the nth derivative of the free energy with respect to the interaction constants in a system with Hamiltonian $-\hat{I} \sum R(b) - \hat{I}_2 \sum RR - \cdots$. The nth derivative with respect to $\hat{I}$ is proportional to the nth temperature derivative of the free energy and is thus proportional to $\epsilon^{2-\alpha-n}$. We assume that the cumulants of products of R's show no stronger singular behavior. Since $\hat{K}_2$ and $\hat{K}_3$ are regular functions of T for fixed $I_2/I_1, I_3/I_1$, we obtain for $I_2/I_1 = I_3/I_1 = -0.4582$

$$\ln Z' = \ln Z_{32}(\hat{K}_1) + \text{regular terms} + O(\epsilon^{3-\alpha}). \tag{4.10}$$

Therefore, we obtain for the singular part of the specific heat $c'_{\text{sing}}(T)$,

$$\begin{aligned} c'_{\text{sing}}(T_c(1+\epsilon)) &= q^{-2} c_{\text{sing},31}(T_{c,31}(1 - q\epsilon)) \\ &\cdot [1 + O(\epsilon)] \\ &= 2.088\, c_{\text{sing},31}(T_{c,31}(1 - 0.6924\epsilon)) \\ &\cdot [1 + O(\epsilon)], \quad q = (\partial \ln K_1/\partial \ln K)_{Tc}. \end{aligned} \tag{4.11}$$

If we assume that the singular part of the free energy depends only on $T\text{-}T_c(I_1, I_2, I_3)$ [compare

Eq. (3.71)], then from Eqs. (4.8) and (4.9) and from

$$\frac{\partial \ln \hat{Z}}{\partial \hat{K}_1} = -\frac{E}{I} \tag{4.12}$$

we obtain

$$\hat{F}_{\mathrm{sing}}(\hat{K}_1, \hat{K}_2, \hat{K}_3) = F_{\mathrm{sing},32}(\hat{K}_1 + 2\hat{K}_2^2 + \hat{K}_3^2 + \cdots). \tag{4.13}$$

Therefore, we may expand the critical temperature in powers of $I_2/I_1 + 0.4582$ and $I_3/I_1 + 0.4582$ (Fig. 7):

$$k_B T_c(I_1, I_2, I_3)/I_1 = 0.4904 - 4.00(I_2/I_1 + 0.4582)^2 - 2.00(I_3/I_1 + 0.4582)^2 - \cdots. \tag{4.14}$$

Since the ground state of this system is the same as for the model (3.10), the two-spin correlations at $T = 0$ vanish and no long range order is expected below T_c.

5. CONCLUSION

In 1966 Mermin and Wagner[27] proved that there is no spontaneous magnetization in the two-dimensional isotropic Heisenberg model. On the other hand, there is evidence from high temperature expansions of the magnetic susceptibility[28] that this system undergoes a phase transition. This raises the question of whether or not it is possible to have a phase transition without a local order parameter. In this paper we have exhibited systems which undergo phase transitions but which do not have a local order parameter. The specific systems were certain classes of Ising models. It would be of some interest to generalize this concept to other types of systems.

ACKNOWLEDGMENTS

It is a pleasure to thank Professor L. P. Kadanoff and Professor H. Wagner for helpful discussions.

APPENDIX

In this appendix we derive the Eqs. (3.21)–(3.23). Any lattice can be created from another lattice with the same boundary conditions by applying one of the following steps (see Fig. 8) as many times as needed.

Step 1: Divide an m-dimensional hypercell into two parts by creating an $(m+1)$-dimensional hypercell.

Step 2: Collapse an $(m+1)$-dimensional hypercell by merging two m-dimensional hypercells with the same boundary together into one.

By applying any of these steps, the left-hand side of

$$\sum_{m=0}^{d} (-1)^{d-m} C_m = t \tag{3.23}$$

remains unchanged. Therefore, t depends only on the boundary conditions. This is a generalization of Euler's theorem.

Next we consider the change of N_g resulting from the application of Step 1. If $m > n$, then the Hamiltonian does not change. For $m = n$ one interaction is effectively duplicated, since one boundary $B^{(n)}$ is duplicated. For $m = n - 1$ one spin is replaced by two spins, but for the ground state both must be equal. For $m = n - 2$ there is also one additional spin. Taking this spin aligned upwards, one obtains a one-to-one correspondence with the ground state of the original system. But changing the signs of all spins lying on bonds adjacent to one $(n-1)$-dimensional hypercell at the boundary of the new bond, we obtain another ground state. Therefore, the new system has twice the degeneracy of the original system. For $m < n - 2$ the Hamiltonian does not change. Therefore, we obtain (Step 2 is just the inverse of Step 1)

$$N_g = t_g + \sum_{m=0}^{n-2} (-1)^{n-m} C_m . \tag{3.21}$$

Since this expression changes only by + 1 after application of Step 1 and by − 1 after application of Step 2 for $m = n - 2$, t_g depends only on the boundary conditions and on n. Similarly, we obtain

$$N_g^* = t_g^* + \sum_{m=n+2}^{d} (-1)^{m-n} C_m . \tag{3.22}$$

Therefore, N_m, Eq. (3.16), depends only on the topology of the system and on n.

We consider two topologies: first, a lattice which is topologically equivalent to a d-dimensional hypersurface wrapped on a $(d+1)$-dimensional hypersphere. As a representative we choose the $(d+1)$-dimensional simplex, which is the generalization of the triangle and the tetrahedron. It has

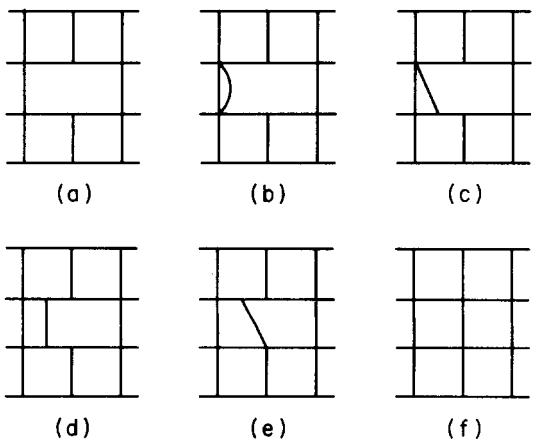


FIG. 8. Example for changing a lattice by applying the Steps 1 and 2. From the lattice (a) the lattice (f) is created by applying once the Step 1 with $m = 1$ (b), twice Step 1 with $m = 0$ (c), (d), twice the Step 2 with $m = 0$ (e), (f). The number $C_2 - C_1 + C_0$ remains unchanged.

$C_0 = d + 2$ corners. Any two corners are connected by an edge. Any three edges span a face and so on. It follows that $C_m = \binom{d+2}{m+1}$. Using Eq. (3.13), one obtains

$$t = 1 + (-1)^d. \tag{A1}$$

We may number the spins of the model M_{dn} on this lattice by n indices $1 \le i_1 < i_2 \cdots < i_n \le d + 2$. For the ground state all spins with $i_n = d + 2$ can be chosen arbitrarily. Then all other spins are given by $S(i_1 \cdots i_n) = S(i_2 \cdots i_n, d + 2) \cdot S(i_1 i_3 \cdots i_n, d + 2) \cdots S(i_1 \cdots i_{n-1}, d + 2)$. Therefore it follows that $N_g = \binom{d+1}{n-1}$. From the Eqs. (3.21), (3.22), (3.26), and (A1) one obtains

$$t_g = (-1)^{n+1}, \quad t_g^* = (-1)^{d+n+1} \quad \text{for } n \le d-1, \tag{A2}$$

$$N_m = 0 \quad \text{for } n \le d-1. \tag{A3}$$

Since $N_m = 0$, the duality relation (3.17) holds for all lattices wrapped on a $(d + 1)$-dimensional hypersphere.

Secondly, we consider a lattice with periodic boundary conditions. As a representative we choose the d-dimensional hypercube. Then it follows that $C_m = \binom{d}{m}$, and one obtains from Eq. (3.23)

$$t = 0. \tag{A4}$$

Because of the periodic boundary conditions, all spins occur twice in the products R. Therefore, all spins can be chosen arbitrarily, $N_g = \binom{d}{n-1}$. Then one obtains from the Eqs. (3.21), (3.22), (3.26), and (A1)

$$t_g = \binom{d-1}{n-1}, \quad t_g^* = \binom{d-1}{n}, \quad N_m = \binom{d}{n}. \tag{A5}$$

---

* Work supported in part by the National Science Foundation.

† On leave from the Kernforschungsanlage Jülich, Germany, where part of this work was done.

1 H. A. Kramers and G. H. Wannier, Phys. Rev. **60**, 252 (1941).

2 L. Onsager, Phys. Rev. **65**, 117 (1944).

3 G. H. Wannier, Rev. Mod. Phys. **17**, 50 (1945).

4 Compare also the review articles G. F. Newell and E. W. Montroll, Rev. Mod. Phys. **25**, 353 (1953) and Ref. 5.

5 C. Domb, Advan. Phys. **9**, 149 (1960).

6 L. P. Kadanoff, Phys. Rev. Letters **23**, 1430 (1969).

7 L. P. Kadanoff and H. Ceva, Phys. Rev. B 3, 3918 (1971).

8 Compare also M. E. Fisher and A. E. Ferdinand, Phys. Rev. Letters **19**, 169 (1967).

9 R. Brout, *Phase Transitions* (Benjamin, New York, 1965).

10 P. W. Anderson, Rev. Mod. Phys. **38**, 298 (1966).

11 L. D. Landau and E. M. Lifshitz, *Statistical Physics* (Pergamon, London, 1958), Chap. 14.

12 L. P. Kadanoff *et al.*, Rev. Mod. Phys. **39**, 395 (1967).

13 M. E. Fisher, Rept. Progr. Phys. **30**, 615 (1967).

14 P. Heller, Rept. Progr. Phys. **30**, 731 (1967).

15 I. Syozi, Progr. Theoret. Phys. (Kyoto) **6**, 306 (1951).

16 S. Naya, Progr. Theoret. Phys. (Kyoto) **11**, 53 (1954).

17 L. Onsager, Ref. 2.

18 A. Pais, Proc. Natl. Acad. Sci. (U.S.) **49**, 34 (1963).

19 Since the algebra on the set $\{0, 1\}$ with the operations (2.5) and (2.6) is a field, all theorems for vectors and matrices over a field apply. Compare, e.g., G. Birkhoff and S. MacLane, *A Survey of Modern Algebra* (MacMillan, New York, 1959), Chap. VII, Secs. 3–8 and 12, Chap. VIII, Secs. 1–9. Note that n linear independent vectors span a space of 2^n vectors.

20 M. E. Fisher, Phys. Rev. **113**, 969 (1959).

21 D. M. Y. Sommerville, *An Introduction to the Geometry of N Dimensions* (Dover, New York, 1958), Chap. IX.

22 D. G. Kelly and S. Sherman, J. Math. Phys. **9**, 466 (1968).

23 D. Ruelle, Rev. Mod. Phys. **36**, 580 (1964), proves for Ising models with two-spin interactions that any cumulant $\chi(r_1, r_2, \cdots)$ vanishes in the limit $|r_1 - r_2| \to \infty$ for $0 \le \beta \le \beta_0$, where β_0 depends on the system. (This proof can also be applied to lattices). Since the expectation value (3.50) can be expressed by a sum of products of cumulants and any product contains at least one cumulant $\chi(r_1^0, r_2^0 + r, \cdots)$, it can be proven exactly that the right-hand side of Eq. (3.50) vanishes for $n = d - 1$ for $0 \le \beta \le \beta_0$.

24 Assuming that a product of operators R can be expressed by the sum of a constant, an energy density, and operators with a less critical behavior in the sense of the operator algebra due to Kadanoff (Refs. 6 and 7), we find from the assumption $F_{\text{sing}}(K, \{h\}) = F_{\text{sing}}(K - K_c\{h\})$ (Refs. 25 and 26) finite derivatives $\partial K_c\{h\}/\partial h_i$. This gives some evidence that a $(\nu - 1)$-dimensional λ-hypersurface $K = K_c\{h\}$ exists which separates the $(K, \{h\})$ space into two phases (in some region around $h_i = 0$) and that this λ-hypersurface is not the boundary of a first-order phase transition.

25 M. E. Fisher, Phil. Mag. **7**, 1731 (1962).

26 E. Riedel and F. Wegner, Z. Physik 225, 195 (1969).

27 N. D. Mermin and H. Wagner, Phys. Rev. Letters **17**, 1133 (1966).

28 H. E. Stanley and T. A. Kaplan, Phys. Rev. Letters **17**, 913 (1966).