JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 7 JULY 1968

Cauchy Problem in the Scalar–Tensor Gravitational Theory

W. J. COCKE AND JEFFREY M. COHEN
Institute for Space Studies, Goddard Space Flight Center, National Aeronautics and Space Administration, New York, New York

(Received 17 November 1967)

We show that the scalar–tensor gravitational field equations can be split up into a set of initial-value equations and a set of time-evolution equations and that the initial-value equations preserve their form as they propagate forward in time. The proper amount of freedom for the stress–energy tensor is maintained, and the Cauchy problem is thus posed reasonably well.

I. INTRODUCTION

A very useful way of treating a physical problem is to consider it as a Cauchy problem. This consists in general of two parts: (a) determining what sort of initial data (Cauchy data) must be given in order to determine completely the future evolution of the system and finding what constraints must be satisfied by the initial data, and (b) determining the future evolution itself. The Cauchy formulation has been very useful for studying causality and is also of great utility in time-dependent machine calculations

In addition, a physical theory must be well posed in terms of the Cauchy problem if the theory can be said to represent physical reality.

In this paper we extend the work of Lichnerowicz,[1] Fourès-Bruhat,[2] and Brill[3] on the Cauchy problem in general relativity to the Brans–Dicke scalar–tensor gravitational theory.[4] We use methods similar to those of Brill[3] and show how, although the introduction of the scalar field Φ complicates the calculations, it is still possible to split the scalar–tensor field equations into a set of initial-value equations on an initial hypersurface and a set of evolution equations which continue the solution out of the initial hypersurface.[5]

We also show that any solution thus generated continues to satisfy the initial-value equations in some neighborhood of the initial hypersurface.

II. FORMULATION OF CAUCHY PROBLEM

We write the scalar–tensor field equations as[4]

$$X_{\mu\nu} \equiv \Phi G_{\mu\nu} - 8\pi T_{\mu\nu} - \frac{\omega}{\Phi}(\Phi_{,\mu}\Phi_{,\nu} - \tfrac{1}{2}g_{\mu\nu}\Phi_{,\lambda}\Phi^{,\lambda}) + g_{\mu\nu}\Phi_{,\lambda}{}^{;\lambda} - \Phi_{,\mu;\nu} = 0 \quad (1)$$

and

$$\Phi_{,\lambda}{}^{;\lambda} = \frac{8\pi}{3+2\omega}T, \quad (2)$$

where $G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R$ is the Einstein tensor, and

$$R_{\mu\nu} \equiv R^{\alpha}{}_{\mu\alpha\nu}.$$

Without loss of generality we now choose a normal Gaussian coordinate system,[6] which has the property $g_{0\mu} = g^{0\mu} = -\delta_{\mu}{}^{0}$. We then pose the initial-value problem on the hypersurface whose local equation is given by $x^0 \equiv t = 0$.

Now the field equations (1) and (2) contain no time derivatives higher than second, and thus the Cauchy data to be given on the surface $x^0 = 0$ will include g_{kl}, Φ, $T_{\mu\nu}$, $g_{kl,0}$, and $\Phi_{,0}$. We now show that the equations $X_{\mu}{}^{0} = 0$ are, as in Einstein's theory, the initial-value equations on the Cauchy data. These initial-value equations are also known as constraint equations.

We write the Bianchi identities as $G_{\mu}{}^{0}{}_{;0} = -G_{\mu}{}^{k}{}_{;k}$ and note that the right-hand side of this expression cannot contain time derivatives higher than second order. Therefore, since this expression is an identity, $G_{\mu}{}^{0}$ contains only first time derivatives. Note also that the expression $\delta_{\mu}{}^{0}\Phi_{,\lambda}{}^{;\lambda} - \Phi_{,\mu}{}^{;0}$ seems to contain second time derivatives if $\mu = 0$. In this case, however, the expression becomes $\Phi_{,0}{}^{;0} + \Phi_{,k}{}^{;k} - \Phi_{,0}{}^{;0} = \Phi_{,k}{}^{;k}$, and thus we conclude that $X_{\mu}{}^{0}$ contains no second time derivative at all. Therefore, the equations

$$X_{\mu}{}^{0} = 0 \quad (3)$$

[1] A. Lichnerowicz, *Théories relativistes de la gravitation et de l'électromagnetisme* (Masson et Cie., Paris, 1955); J. Math. Pure Appl. **23**, 37 (1944); Compt. Rend. Acad. Sci. Paris **260**, 3291 (1965).

[2] Y. Fourès-Bruhat, Acta Math. **88**, 141 (1952); J. Ratl. Mech. Anal. **5**, No. 6, 951 (1956).

[3] D. R. Brill, Ph.D. thesis, Princeton University, 1959; Ann. Physics (N.Y.) **7**, 469 (1959). This technique is also discussed in J. Weber, *General Relativity and Gravitational Waves* (Interscience Publishers, Inc., New York, 1961), pp. 107–109.

[4] C. Brans and R. H. Dicke, Phys. Rev. **124**, 925 (1960). For a discussion of more general scalar–tensor theories, see P. Jordan, *Schwerkraft und Weltall* (F. Vieweg, Braunschweig, 1955), 2nd ed. See also D. R. Brill, *Proceedings of the E. Fermi School, Course XX* (Academic Press Inc., New York, 1962).

[5] For an example of a solution of the Cauchy problem in general relativity see, e.g., J. M. Cohen, in *Lectures in Applied Mathematics* **8**, *Relativity Theory and Astrophysics*, J. Ehlers, Ed. (American Mathematical Society, Providence, R.I., 1967).

[6] J. L. Synge and A. Schild, *Tensor Calculus* (University of Toronto Press, Toronto, 1949), pp. 62–65.

are conditions on the initial values, which correspond to the situation in general relativity. Likewise the equations

$$X_k^{\ l} = 0, \tag{4}$$

together with Eq. (2), are the equations of time evolution. The quantities $X_0^{\ k}$ are, in the normal Gaussian coordinate system, linear combinations of $X_l^{\ 0}$ and need not be considered explicitly.

We now demonstrate that if the conservation laws

$$T_\mu^{\ \nu}{}_{;\nu} = 0 \tag{5}$$

are satisfied, then any solution found by integrating the time-evolution equations (2), (4), and (5) will always fulfill the initial-value equations (3) in the future. The time-development equations (2), (4), and (5) are considered to hold over all of that part of space–time where the normal Gaussian coordinate system is valid.

Equation (5), written $T_\mu^{\ \nu}{}_{;\nu} = T_\mu^{\ 0}{}_{;0} + T_\mu^{\ k}{}_{;k} = 0$, is a time-evolution equation for the $T_\mu^{\ 0}$ part of the stress–energy tensor, and we show later on that we are then left with just enough freedom to specify the type of stress–energy tensor corresponding to the physical system under consideration. These physical specifications then completely determine the problem. Remember that the arbitrariness of the coordinate system has been removed by choosing the coordinates to be normal Gaussian.

We note that it may be shown by the methods of Lichnerowicz[1] (as in general relativity) that the characteristic surfaces of the systems of Eqs. (1) and (2) are null surfaces and that these equations constitute a hyperbolic system.

We now use a variant of Brill's method[3] and show that if $X_\mu^{\ 0} = 0$ at $x^0 = 0$, then the propagation equations (2), (4), and (5) imply $0 = X_\mu^{\ 0}{}_{,0} = X_\mu^{\ 0}{}_{,00} = \cdots$, at $x^0 = 0$. Thus $X_\mu^{\ 0}$ will be shown to vanish identically inside some nonzero interval about $x^0 = 0$.

We take the divergence of Eq. (1) and use Eqs. (5) to obtain

$$X_\mu^{\ \nu}{}_{;\nu} = \Phi_{,\nu} R_\mu^{\ \nu} - \tfrac{1}{2}\Phi_{,\mu} R - \frac{\omega}{\Phi}\Phi_{,\nu}^{\ ;\nu}\Phi_{,\mu} + \tfrac{1}{2}\frac{\omega}{\Phi^2}\Phi_{,\mu}\Phi_{,\nu}\Phi^{,\nu} + \Phi_{,\nu}^{\ ;\nu}{}_{,\mu} - \Phi_{,\mu;\nu}^{\ \ ;\nu}. \tag{6}$$

Since $\Phi_{,\mu;\nu}^{\ \ ;\nu} = \Phi_{,\nu;\mu}^{\ \ ;\nu} = \Phi_{,\nu}^{\ ;\nu}{}_{;\mu} + R_\mu^{\ \nu}\Phi_{,\nu}$, Eq. (6) becomes

$$X_\mu^{\ \nu}{}_{;\nu} = \tfrac{1}{2}\Phi_{,\mu}\left(R + \frac{2\omega}{\Phi}\Phi_{,\nu}^{\ ;\nu} - \frac{\omega}{\Phi^2}\Phi_{,\nu}\Phi^{,\nu}\right) = \Phi_{,\mu} X/2\Phi. \tag{7}$$

We can now evaluate this expression at $x^0 = 0$. By taking the trace of Eqs. (1) and eliminating T by means of Eq. (2), we obtain $X_\mu^{\ \nu}{}_{;\nu} = 0$ at $x^0 = 0$.

Thus we obtain $X_\mu^{\ 0}{}_{,0} = \Gamma_{\mu 0}^{\ \ \nu} X_\nu^{\ 0} - \Gamma_{0\nu}^{\ \ 0} X_\mu^{\ \nu} - X_\mu^{\ k}{}_{;k}$ at $x^0 = 0$. But the right-hand side of this equation vanishes at $x^0 = 0$, and therefore $X_\mu^{\ 0}{}_{,0} = 0$ there.

Now consider again Eqs. (7). By differentiating partially with respect to x^0, one can show, with the use of the relations $X_\mu^{\ 0}{}_{,0} = 0$ (at $x^0 = 0$), that $X_\mu^{\ 0}{}_{,00} = 0$ at $x^0 = 0$. This process can be continued to any order, and thus $X_\mu^{\ 0}$ vanishes identically in some neighborhood of the hypersurface $x^0 = 0$. This neighborhood is a region at least as large as the region in which the normal Gaussian coordinate system is usable.

This proves that the Cauchy problem is well posed if the initial-value equations (3) can be satisfied on the initial hypersurface.

III. CONTINUATION OF PHYSICAL QUANTITIES FROM INITIAL HYPERSURFACE

We now discuss in detail how the metric, the scalar field, and the stress–energy tensor can be continued uniquely as far as the coordinate system is usable. If one writes out Eqs. (2), (4), and (5), one sees by inspection that $g_{kl,00}$, $\Phi_{,00}$, and $T_\mu^{\ 0}{}_{,0}$ are given as functions of g_{kl}, Φ, their space derivatives and first time derivatives, and of $T_\mu^{\ \nu}$. This allows us to continue g_{kl}, $g_{kl,0}$, Φ, $\Phi_{,0}$, and $T_\mu^{\ 0}$ from the hypersurface $x^0 = 0$ to a neighboring hypersurface $x^0 = \delta t$, an infinitesimal distance from the initial hypersurface.

But how is one to continue $T_k^{\ l}$? To answer this question we must go into the physics of the situation and show that we have the correct amount of freedom. This must be so if we are to treat physical problems sensibly.

Let us first note that if $T_k^{\ l}$ is given arbitrarily over all space–time, then the problem is completely solved: we will have "continued" $T_k^{\ l}$ by brute force, as in Synge's treatment.[7] This circumstance does not usually arise in physical problems, but we see that any physical method of continuing $T_k^{\ l}$ can be used.

We consider two general cases which are of physical interest and which include important special cases.

(A) $T_k^{\ l} = f_k^{\ l}(T_\mu^{\ 0})$. In this case we continue $T_k^{\ l}$ to the neighboring hypersurface by knowing it as a function of $T_\mu^{\ 0}$, which we have already shown how to continue. The perfect fluid, for which $T_{\mu\nu} = (p + \rho)\, U_\mu U_\nu + p g_{\mu\nu}$ with a given equation of state $p = p(\rho)$, is an example of this type. The four independent quantities ρ and U_k (remember that $U_\lambda U^\lambda = -1$) can be found from $T_\mu^{\ 0}$ and then substituted into $T_k^{\ l}$.

[7] J. L. Synge, *Relativity: The General Theory* (North-Holland Publishing Co., Amsterdam, 1960), pp. 213–16.

(B) $T_\mu^{\ \nu} = f_\mu^{\ \nu}(F)$, where the tensor F satisfies another set of differential equations for which the Cauchy problem may be correctly posed. The conservation equations (5) must then be compatible with the F equations, and we can say that F is continued by its own set of time-evolution equations in such a way that the conservation laws, Eqs. (5), hold automatically. The electromagnetic field *in vacuo* is of this type.

Another case which may be of interest is $T_k^{\ l} = f_k^{\ l}(g_{\mu\nu})$. In some special cases the stress tensor can be reduced to this form.

IV. DISCUSSION

We conclude that the Cauchy problem for the scalar–tensor gravitational theory can be posed in a manner similar to that for general relativity.

One can also use a coordinate-free method involving orthonormal Cartan frames,[8] as in general relativity.[1,2] The present method, employing normal Gaussian coordinates, was used without loss of generality because of its simplicity.

ACKNOWLEDGMENTS

The authors would like to thank Mme. Y. Choquet-Bruhat and Professor R. H. Dicke for reading the manuscript and for helpful suggestions.

This work was completed while the authors held National Academy of Sciences–National Research Council Postdoctoral Research Associateships, supported by the National Aeronautics and Space Administration.

[8] E. Cartan, *Leçons sur la géométrie des éspaces de Riemann* (Gauthier-Villars, Paris, 1946). For a review of the Cartan calculus, see, e.g., D. R. Brill and J. M. Cohen, J. Math. Phys. **7**, 238 (1966).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 7 JULY 1968

New Methods for Reduction of Group Representations. II*

JOHN GABRIEL
Argonne National Laboratory, Argonne, Illinois

(Received 9 June 1967)

A better algorithm for reduction of a group representation than that given in an earlier paper [J. R. Gabriel, J. Math. Phys. **5**, 494 (1964)] is derived.

1. INTRODUCTION

It has been shown in a previous publication,[1] which will be referred to as I for the rest of this paper, that, in order to reduce a given group representation, it is sufficient to reduce the commutator algebra[2] of the group and that the commutator algebra is the algebra of all matrices which commute with the representatives of the group generators[3] (the proof is in I, Lemmas 1 and 2). Paper I did not give a usable method for constructing and reducing a set of basic elements for the commutator algebra, and the purpose of this paper is to present easily usable techniques. Readers who would like a concise discussion of group algebras are referred to Weyl's book on group theory and quantum mechanics.[3] The construction of irreducible representations of the commutator algebra depends on the construction of irreducible representations of the commutator algebra of two given matrices U and V, which will be treated first in Secs. 2 and 3. The reduction of a group representation is discussed in Sec. 4.

2. COMMUTATOR ALGEBRA OF ONE MATRIX

Let U be unitary or Hermitian. Let one of its eigenvalues be α, and let the number of linearly independent eigenvectors belonging to α be n_α.

Call the eigenvectors $|\alpha i\rangle$, $i = 1, \cdots, n_\alpha$. It may be assumed with no loss of generality that

$$\langle \alpha i \mid \alpha' i' \rangle = \delta\alpha\alpha'\delta ii'. \tag{2.1}$$

Define

$$\epsilon^\alpha = \sum_{i=1}^{n_\alpha} |\alpha i\rangle\langle \alpha i|, \tag{2.2}$$

then

$$U = \sum_\alpha \alpha\epsilon^\alpha, \tag{2.3}$$

$$\epsilon^\alpha\epsilon^{\alpha'} = \delta\alpha\alpha'\epsilon^\alpha, \tag{2.4}$$

and

$$1 = \sum_\alpha \epsilon^\alpha. \tag{2.5}$$

* Work performed under the auspices of the U.S. Atomic Energy Commission.

[1] J. R. Gabriel, J. Math. Phys. **5**, 494 (1964).

[2] H. Weyl, *The Classical Groups* (Princeton University Press, Princeton, N.J., 1946), p. 93.

[3] H. Weyl, *Group Theory and Quantum Mechanics* (Methuen and Company Ltd., London, 1931; Dover Publications, Inc., New York), pp. 165–170 and 302–309.

(B) $T_\mu^{\ \nu} = f_\mu^{\ \nu}(F)$, where the tensor F satisfies another set of differential equations for which the Cauchy problem may be correctly posed. The conservation equations (5) must then be compatible with the F equations, and we can say that F is continued by its own set of time-evolution equations in such a way that the conservation laws, Eqs. (5), hold automatically. The electromagnetic field *in vacuo* is of this type.

Another case which may be of interest is $T_k^{\ l} = f_k^{\ l}(g_{\mu\nu})$. In some special cases the stress tensor can be reduced to this form.

IV. DISCUSSION

We conclude that the Cauchy problem for the scalar–tensor gravitational theory can be posed in a manner similar to that for general relativity.

One can also use a coordinate-free method involving orthonormal Cartan frames,[8] as in general relativity.[1,2] The present method, employing normal Gaussian coordinates, was used without loss of generality because of its simplicity.

ACKNOWLEDGMENTS

The authors would like to thank Mme. Y. Choquet-Bruhat and Professor R. H. Dicke for reading the manuscript and for helpful suggestions.

This work was completed while the authors held National Academy of Sciences–National Research Council Postdoctoral Research Associateships, supported by the National Aeronautics and Space Administration.

[8] E. Cartan, *Leçons sur la géométrie des éspaces de Riemann* (Gauthier-Villars, Paris, 1946). For a review of the Cartan calculus, see, e.g., D. R. Brill and J. M. Cohen, J. Math. Phys. **7**, 238 (1966).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 7 JULY 1968

New Methods for Reduction of Group Representations. II*

JOHN GABRIEL
Argonne National Laboratory, Argonne, Illinois

(Received 9 June 1967)

A better algorithm for reduction of a group representation than that given in an earlier paper [J. R. Gabriel, J. Math. Phys. **5**, 494 (1964)] is derived.

1. INTRODUCTION

It has been shown in a previous publication,[1] which will be referred to as I for the rest of this paper, that, in order to reduce a given group representation, it is sufficient to reduce the commutator algebra[2] of the group and that the commutator algebra is the algebra of all matrices which commute with the representatives of the group generators[3] (the proof is in I, Lemmas 1 and 2). Paper I did not give a usable method for constructing and reducing a set of basic elements for the commutator algebra, and the purpose of this paper is to present easily usable techniques. Readers who would like a concise discussion of group algebras are referred to Weyl's book on group theory and quantum mechanics.[3] The construction of irreducible representations of the commutator algebra depends on the construction of irreducible representations of the commutator algebra of two given matrices U and V, which will be treated first in Secs. 2 and 3. The reduction of a group representation is discussed in Sec. 4.

2. COMMUTATOR ALGEBRA OF ONE MATRIX

Let U be unitary or Hermitian. Let one of its eigenvalues be α, and let the number of linearly independent eigenvectors belonging to α be n_α.

Call the eigenvectors $|\alpha i\rangle$, $i = 1, \cdots, n_\alpha$. It may be assumed with no loss of generality that

$$\langle \alpha i \mid \alpha' i' \rangle = \delta\alpha\alpha'\delta ii'. \tag{2.1}$$

Define

$$\epsilon^\alpha = \sum_{i=1}^{n_\alpha} |\alpha i\rangle\langle \alpha i|, \tag{2.2}$$

then

$$U = \sum_\alpha \alpha\epsilon^\alpha, \tag{2.3}$$

$$\epsilon^\alpha\epsilon^{\alpha'} = \delta\alpha\alpha'\epsilon^\alpha, \tag{2.4}$$

and

$$1 = \sum_\alpha \epsilon^\alpha. \tag{2.5}$$

* Work performed under the auspices of the U.S. Atomic Energy Commission.

[1] J. R. Gabriel, J. Math. Phys. **5**, 494 (1964).

[2] H. Weyl, *The Classical Groups* (Princeton University Press, Princeton, N.J., 1946), p. 93.

[3] H. Weyl, *Group Theory and Quantum Mechanics* (Methuen and Company Ltd., London, 1931; Dover Publications, Inc., New York), pp. 165–170 and 302–309.

Suppose that

$$CU = UC. \tag{2.6}$$

By (2.5)

$$C = 1C1 = \sum_{\alpha\alpha'} \epsilon^{\alpha} C \epsilon^{\alpha'} \tag{2.7}$$

for any C.

Write

$$\epsilon^{\alpha} C \epsilon^{\alpha'} = C^{\alpha\alpha'}. \tag{2.8}$$

From (2.4), (2.7) may be written

$$C = \sum_{\alpha\alpha'} \epsilon^{\alpha} C^{\alpha\alpha'} \epsilon^{\alpha'}. \tag{2.9}$$

Substituting (2.3) and (2.9) in (2.6), and using (2.4), it may be shown that (2.6) implies

$$C^{\alpha\alpha'} = 0, \quad \text{unless} \quad \alpha = \alpha\alpha'; \tag{2.10}$$

i.e., that if C commutes with U,

$$C = \sum_{\alpha} \epsilon^{\alpha} C^{\alpha} \epsilon^{\alpha}. \tag{2.11}$$

This condition, or (2.10), is also sufficient, as can be verified by substituting (2.11) and (2.3) in (2.6).

Note that (2.11) shows that

$$\epsilon^{\alpha} C = C \epsilon^{\alpha}, \tag{2.12}$$

for all α. This is another necessary and sufficient condition for C to commute with U.

The matrices ϵ^{α} are primitive idempotents or projection operators because of (2.4). The space S^{α} selected by ϵ^{α} is spanned by the n_{α} vectors $|\alpha i\rangle$ belonging to α.

Substituting (2.2) into (2.11) shows that

$$\begin{aligned} C &= \sum_{\alpha} \sum_{i} \sum_{j} |\alpha i\rangle\langle \alpha i| \, C \, |\alpha j\rangle\langle \alpha j| \\ &= \sum_{\alpha} \sum_{i} \sum_{j} C^{\alpha} ij \, |\alpha i\rangle\langle \alpha j|, \end{aligned} \tag{2.13}$$

where

$$C^{\alpha} ij = \langle \alpha i| \, C \, |\alpha j\rangle. \tag{2.14}$$

The numbers $C^{\alpha} ij$ are not restricted in any way, and so any vector of S^{α} may be mapped onto any other vector of S^{α} by C. Therefore S^{α} is irreducible with respect to (w.r.t.) the commutator algebra of U. Equation (2.11) or (2.12) shows that it is also invariant under the commutator algebra of U.

3. THE COMMUTATOR ALGEBRA OF TWO MATRICES U AND V

One of the important results of Sec. 2 was to show how the total space S was divided into subspaces S^{α}, each invariant and irreducible under the commutator algebra of U. Two subspaces S^{α} and $S^{\alpha'}$ may be transformed independently by elements C of the commutator algebra of U.

The requirement that C commute with V as well as with U can reasonably be expected further to subdivide the spaces S^{α}. It has one other effect. Using the notation of (2.8), it forces certain of the matrices $C^{\alpha} = \epsilon^{\alpha} C \epsilon^{\alpha}$ to be equivalent to each other if C is to commute with both U and V. Although it is necessary in the end to discuss this formally, a heuristic treatment of the condition shows its true causes much better than the formal discussion, and will be given first.

Let an eigenvalue of V be β, analogous to α in Sec. 2. Then

$$\sum_{\beta} \epsilon^{\beta} = 1, \tag{3.1}$$

$$|\alpha i\rangle = \sum_{\beta} \epsilon^{\beta} \, |\alpha i\rangle = \sum_{\beta} |\alpha\beta i\rangle, \tag{3.2}$$

where

$$|\alpha\beta i\rangle = \epsilon^{\beta} \, |\alpha i\rangle \tag{3.3}$$

is an eigenvector of V. And

$$\begin{aligned} C \, |\alpha i\rangle &= \sum_{j} C^{\alpha} ij \, |\alpha j\rangle \\ &= \sum_{j} C^{\alpha} ij \sum_{\beta} |\alpha\beta j\rangle \qquad (3.4) \\ &= \sum_{\beta} \sum_{j} C^{\alpha\beta} ij \, |\alpha\beta j\rangle, \qquad (3.5) \end{aligned}$$

where

$$C^{\alpha\beta} ij = C^{\alpha} ij, \quad \text{for all } \beta. \tag{3.6}$$

Let $S^{\beta\alpha}$ be the space spanned by the n_{α} vectors $|\alpha\beta i\rangle$, $i = 1, 2, \cdots, n_{\alpha}$. Equation (3.6) requires the same transformation in each space $S^{\beta\alpha}$.

Equations (3.4)–(3.6) appear deceptively simple. However, since the n_{α} vectors $|\alpha\beta i\rangle$ are not necessarily linearly independent and some of them may be null, (3.4)–(3.6) have to be analyzed in great detail to discover exactly what they imply about the subspaces of S^{α} or S^{β}, which are invariant under the algebra of all matrices that commute with U and V.

It turns out that the argument revolves around Lemma 3 of I, which is an extension of Schur's lemma, and that the best approach to the problem is not through (3.4)–(3.6), but as follows:

Consider

$$C = \sum_{\beta} \epsilon^{\beta} C^{\beta} \epsilon^{\beta}. \tag{3.7}$$

By (2.11) it commutes with V; by (2.12) it commutes with U, if and only if it commutes with each ϵ^{α}.

Suppose

$$\epsilon^{\alpha} C = C \epsilon^{\alpha}, \tag{3.8}$$

i.e.,

$$\sum_{\beta} \epsilon^{\alpha} \epsilon^{\beta} C^{\beta} \epsilon^{\beta} = \sum_{\beta'} \epsilon^{\beta'} C^{\beta'} \epsilon^{\beta'} \epsilon^{\alpha}. \tag{3.9}$$

Premultiply by $\epsilon^{\beta'}$ and postmultiply by ϵ^{β}:

$$\epsilon^{\beta'} \epsilon^{\alpha} \epsilon^{\beta} C^{\beta} \epsilon^{\beta} = \epsilon^{\beta'} C^{\beta'} \epsilon^{\beta'} \epsilon^{\alpha} \epsilon^{\beta}. \tag{3.10}$$

Equations (2.4) and (2.9) show that (3.10) implies

$$H^{\beta'\beta}C^{\beta} = C^{\beta'}H^{\beta'\beta}, \tag{3.11}$$

where

$$H^{\beta'\beta} = \epsilon^{\beta'}\epsilon^{\alpha}\epsilon^{\beta}. \tag{3.12}$$

Consider (3.11) in the case where $\beta' = \beta$:

$$H^{\beta\beta}C^{\beta} = C^{\beta}H^{\beta\beta}. \tag{3.13}$$

Since ϵ^{α} and ϵ^{β} are both Hermitian because they have real eigenvalues, so also is $H^{\beta\beta}$.

It follows that a basis for S^{β} can be found in which $H^{\beta\beta}$ is diagonal, and S^{β} can be divided into subspaces $S^{\beta\lambda}$ each spanned by the eigenvectors of $H^{\beta\beta}$ belonging to the eigenvalue λ. From Sec. 2 it follows that each $S^{\beta\lambda}$ is invariant w.r.t. the algebra of all matrices that commute with ϵ^{α} and V. The space $S^{\beta\lambda}$ is not necessarily irreducible w.r.t. the algebra, because the condition $H^{\beta'\beta}C^{\beta} = C^{\beta'}H^{\beta\beta}$ for all β' may further restrict C^{β}.

The condition that $S^{\beta\lambda}$ is invariant w.r.t. C requires that

$$C = \sum_{\beta\lambda} e^{\beta\lambda}C^{\beta\lambda}\epsilon^{\beta\lambda}, \tag{3.14}$$

and $CU = UC$ requires that

$$C^{\beta\lambda}H^{\beta\lambda\beta'\lambda'} = H^{\beta\lambda\beta'\lambda'}C^{\beta'\lambda'}, \tag{3.15}$$

where

$$H^{\beta\lambda\beta'\lambda'} = \epsilon^{\beta\lambda}\epsilon^{\alpha}\epsilon^{\beta'\lambda'}. \tag{3.16}$$

For the sake of convenience write the single symbol γ for the pair $\beta\lambda$; let n_{γ} be the dimension of S^{γ}, and r the rank of $H^{\gamma\gamma'}$.

There are five cases:

(1) $n_{\gamma} = n_{\gamma'} = r$,

(2) $n_{\gamma} > n_{\gamma'} = r$,

(3) $n_{\gamma'} > n_{\gamma} = r$,

(4) $n_{\gamma'} > n_{\gamma} > r$,

(5) $n_{\gamma} > n_{\gamma'} > r$.

Case 1: Equation (3.15) becomes

$$C^{\gamma}H^{\gamma\gamma'} = H^{\gamma\gamma'}C^{\gamma'}. \tag{3.17}$$

The matrices C^{γ}, $C^{\gamma'}$, and $H^{\gamma\gamma'}$, partitioned conformably to S^{γ}, $S^{\gamma'}$, and the rest of S, are as in Fig. 1, where in Case 1, H is a nonsingular $n_{\gamma} \times n_{\gamma}$ submatrix. Equation (3.17) shows that

$$HC' = CH \tag{3.18}$$

or

$$C' = H^{-1}CH, \tag{3.19}$$

i.e., that S^{γ} and $S^{\gamma'}$ are equivalent, and C is in no way restricted so that both spaces are irreducible, since

$$C^{\gamma} = \begin{bmatrix} C & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad C^{\gamma'} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & C' & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{matrix} \}S^{\gamma} \\ \}S^{\gamma'} \\ \}S - S^{\gamma} - S^{\gamma'} \end{matrix},$$

$$H^{\gamma\gamma'} = \begin{bmatrix} 0 & H & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad C^{\gamma}H^{\gamma\gamma'} = \begin{bmatrix} 0 & CH & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

$$H^{\gamma\gamma'}C^{\gamma'} = \begin{bmatrix} 0 & HC' & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

FIG. 1. Conformably partitioned matrices C^{γ}, $C^{\gamma'}$, and $H^{\gamma\gamma'}$.

C^{γ} can map any vector of S^{γ} onto any other vector of S^{γ}, but both must afford the same irreducible representation of the commutator algebra, i.e., they are equivalent.

In Case 2, H has n_{γ} rows and $n_{\gamma'}$ columns. Since $r = n_{\gamma'}$, the vectors of S^{γ} may be transformed (by pivotal condensation of H) to leave a nonsingular square $(n_{\gamma'} \times n_{\gamma'})$ submatrix in the leading position above $(n_{\gamma} - n_{\gamma'})$ rows of zeros.

Lemma 3 of I shows that the subspace of S^{γ} mapped onto $S^{\gamma'}$ by this square nonsingular submatrix is invariant and irreducible w.r.t. the algebra of all matrices C that obey (3.17). Thus S^{γ} is further split into two spaces, one equivalent to $S^{\gamma'}$, as in Case 1, and the other not restricted.

Case 3 is analogous, and Cases 4 and 5 simply divide both S^{γ} and $S^{\gamma'}$ into r-dimensional subspaces and remainders. The r-dimensional subspaces are equivalent and the remainders are not.

4. PROCESS OF REDUCTION FOR TWO MATRICES

Reduction is started by reducing the whole space w.r.t. the commutator algebra of U, i.e., finding the ϵ^{α}'s.

Next find the ϵ^{β}'s from V. Starting with the first α, form each matrix

$$H^{\beta\beta}(\alpha) = \epsilon^{\beta}\epsilon^{\alpha}\epsilon^{\beta} \tag{4.1}$$

and reduce it to diagonal form, as in Sec. 3, to obtain the spaces S^{γ}.

Take the first γ value (i.e., $\beta\lambda$ pair). Form $H^{\gamma\gamma'}$ for a γ' not equal to γ. If $H^{\gamma\gamma'}$ is not null, determine which of the Cases 1–5 of Sec. 3 applies and subdivide S^{γ} or $S^{\gamma'}$ accordingly, noting which subspaces are equivalent. Suppose $\gamma 1$ and $\gamma 2$ are the components of γ and $\gamma 1'$ and $\gamma 2'$ of γ'. Let $\gamma 1 \equiv \gamma 1'$ and $\gamma 2 \not\equiv \gamma 2'$. Choose a new γ—γ'', form $H^{\gamma 1\gamma''}$, $H^{\gamma 2\gamma''}$, $H^{\gamma 1'\gamma''}$, $H^{\gamma 2'\gamma''}$, and use

them to subdivide γ'', $\gamma 1$, $\gamma 2$, $\gamma 1'$, and $\gamma 2'$. Note that a subdivision of $\gamma 1$ implies a corresponding subdivision of $\gamma 1'$ through the mapping $H^{\gamma\gamma'}$ of $S^{\gamma'}$ onto $S^{\gamma'}$. This process may be continued for a given α until the spaces γ', γ'', γ''', etc., have been exhausted. When this has been done, proceed to the next α. When all the α's have been exhausted, the whole space has been reduced w.r.t. the commutator algebra of U and V.

Now a further matrix W can be introduced (if necessary) to generate the group whose representation is being studied, and its ϵ may take the place of the ϵ^α's while the ϵ's of the spaces already obtained take the place of the ϵ^β's or vice versa.

Thus the group generators may be added one at a time to the scheme, as proposed in Sec. 1 of I.

Suppose that S has been divided into subspaces $S(i, p)$ by previous steps. The label i takes the values $i = 1, \cdots, n_s$, p the values $p = 1, 2, \cdots, p \max (i)$.

All the spaces with the same i are equivalent. Let us analyze the effect of adding another ϵ^α in the sense of Sec. 3.

Start with $S(1, 1)$. Form $\epsilon(1, 1)\epsilon^\alpha\epsilon(1, 1)$ and subdivide $S(1, 1)$ according to its eigenvalues. If the division is proper [i.e., there are subspaces other than $S(1, 1)$ and the null space], suppose that there are n distinct subspaces. Replace n_s by $n_s + n - 1$, and label the subspaces $i = n_s + 1$, $n_s + 2$, $n_s + 3, \cdots$, $n_s + n - 1$. Subdivide each of the $[p \max (1) - 1]$ equivalent spaces correspondingly.

Now proceed to $S(1, 2)$. Although it is not the same space as $S(1, 2)$ was before $S(1, 1)$ was subdivided, the process used on $S(1, 1)$ can be used on it; i.e., form $\epsilon(1, 2)\epsilon^\alpha\epsilon(1, 2)$ and use its eigenvalues to divide $S(1, 2)$. Divide $S(1, 1)$, $S(1, 3)$, etc., correspondingly adding the newly formed spaces to the far end (large i) of the line as before. Because the whole space S is finite, eventually the process of forming $\epsilon(i, p)\epsilon^\alpha\epsilon(i, p)$ reaches the end of the line, and then may be repeated for the next α. This process may be continued until all the α's have been used up, and we have a set of spaces $S(i, p)$, $i = 1, \cdots, n_s$, $p = 1, \cdots, p \max (i)$. Each space is invariant but not necessarily irreducible, since (3.10) has only been applied with $\beta' = \beta$ so far. We now consider the case with $\beta' \neq \beta$. Start with $S(1, 1)$ and the first α. Form $\epsilon(1, 1)\epsilon^\alpha\epsilon(1, 2)$. Let it be of rank r. If r is equal to the dimension of $S(1, 1)$, then the two spaces are equivalent and no new information has been obtained, since they both belong to $i = 1$ and are therefore equivalent anyway.

If $0 < r < n(1, 1)$, where $n(1, 1)$ is the dimension of $S(1, 1)$, then Lemma 3 of I shows that $S(1, 1)$ has an r-dimensional subspace which is invariant w.r.t. the commutator algebra of all the ϵ's, and so do all the others of p max (1) equivalent spaces. Rename these spaces $(1, p)$, $p = 1, \cdots, p \max (i)$ and rename the other $p \max (1)$, $[n(1, 1) - r]$-dimensional spaces, $i = n_s + 1$.

If $r = 0$, i.e., if $\epsilon(1, 1)\epsilon^\alpha\epsilon(1, 1)$ is the null matrix, then the situation is just as it was when $r = n(1, 1)$.

This process may be repeated for each ϵ^α until the α's are exhausted. Then consider the first α again and $\epsilon(1, 1)\epsilon^\alpha\epsilon(i, p)$. The above may be repeated with $i = 2$ instead of $i = 1$.

This case is slightly different, since $S(1, 1)$ and $S(i, p)$ are not equivalent for $i \neq 1$. To deal with this, let r be the rank of $H = \epsilon(1, 1)\epsilon^\alpha\epsilon(i, p)$. If $r = 0$, then there is no new information. If $r \neq 0$, then H maps an r-dimensional subspace of $S(i, p)$ onto an r-dimensional subspace of $S(1, 1)$. In the case where r is less than the dimension of either $S(1, 1)$ or $S(i, p)$, rename the r-dimensional subspace of $S(1, 1)S(1, 1)$, and divide all the p max (i) spaces $S(i, p)$ correspondingly. Add the p max (i) spaces from $S(i, p)$ that are equivalent to $S(1, 1)$ to the list belonging with $i = 1$ and replace $p \max (1)$ with $p \max (1) + p \max (i)$. Rename the $p \max (1)$ subspaces of $S(i, p)$ which have dimension $[n(i, p) - r, S(i, p)]$, and add the $p \max (1)$ spaces with dimension $[n(1, 1) - r]$ to the end of the list replacing n_s by $(n_s + 1)$.

In the case where r is equal to the dimension of one space, it is not necessary to add to the end of the list; and in the case where r is equal to the dimension of both spaces, then the spaces $S(i, p), p = 1, \cdots, p \max (i)$ are renamed $S(1, p)$ with $p \max (1)$ replaced by $p \max (1) + p \max (i)$.

If r is equal to the dimension of both spaces, then there is no proper subspace left to belong with the label i, and the spaces $S(ns, p)$, $p = 1, \cdots, p \max (n_s)$ must be relabelled $S(i, p)$, and n_s replaced by $(n_s - 1)$.

5. CONCLUSIONS

Although I contained the essential ideas and results in Lemmas 1, 2, and 3, Lemmas 4 and 5 described a process that proved to be too complicated to code for a computer. In this paper, Sec. 4 describes a procedure which is more amenable to automatic processing, and work is now in progress at Argonne National Laboratory to write the necessary code.

Tamm–Dancoff Method in the $2V$ Sector of the Lee Model*

LEONARD M. SCARFONE
Department of Physics, University of Vermont, Burlington, Vermont

(Received 3 November 1967)

The $2V$ sector of the Lee model with boson sources at zero separation is considered from the Tamm–Dancoff point of view. This involves three singular integral equations which are solved by appealing to results previously gained in the Lehmann–Symanzik–Zimmermann method of solution. Two scattering amplitudes and a production amplitude are rederived via the state vectors and an equation for the determination of the two-meson exchange potential is discussed.

I. INTRODUCTION

In a recent communication,[1] hereafter referred to as I, the Lehmann–Symanzik–Zimmermann (LSZ) formalism was applied to the $2V$ sector of the Lee model. This sector embraces two elastic scattering amplitudes, a production amplitude, and a bound-state problem. Thus it has many features in common with the much discussed $V\theta$ sector and is more suggestive than the VN sector in that both sources undergo renormalization. In the LSZ method of solution, it has been shown previously[2,3] that bound-state parameters characteristic of the V and VN systems follow from the analytic properties of their corresponding propagators. Analogously, in I we developed an equation for the determination of the $2V$ binding energy (at zero separation) by setting the denominator function in the $2V$ propagator equal to zero at a bound-state pole. This function also appears in the amplitudes and incorporates the expected analytic structure, namely, the elastic and inelastic cuts.

The LSZ development of the $V\theta$ sector[4] shows that the solution to one singular integral equation solves the entire sector. In I the solutions to two such equations are required. On the other hand, the Tamm–Dancoff (TD) method in the former problem[5] deals with two distinct integral equations, one for each of the elastic-scattering amplitudes. In this note, which is intended to be a sequel to I, we shall see that the TD approach to the $2V$ sector leads to three singular integral equations. To solve these equations, we appeal to the results already gained in I. Thus in Sec. II we begin with the physical states describing the elastic scattering of a θ by the VN system and show how the corresponding integral equation and its solution are obtained by taking various limits of the appropriate expressions in I. In this way we are led to the TD bare-state expansion and then to the conventional T matrix by taking the energy-shell limit of the pertinent expansion coefficient.[5] A similar procedure is used in dealing with the elastic scattering of two θ's by two N's. In this case we encounter a new integral equation, and again the solution is found in terms of previous knowledge. The production amplitude is read off of the $VN\theta$ scattering states. A discussion of the bound-state problem and some closing remarks are given in Secs. III and IV, respectively.

II. SCATTERING STATES AND AMPLITUDES

The physical states describing the elastic scattering of a θ by the VN system may be expressed in the form

$$|VN\theta_{k,\mathrm{in}}\rangle = a_k^+ |VN\rangle + ZC(\omega)|2_V\rangle + |\chi\rangle^+. \quad (1)$$

In this expansion $|VN\rangle$ denotes a stationary eigenstate of the total Hamiltonian H with eigenvalue $E_0 = 2m + \omega_0$. For algebraic convenience we deal with the case in which the physical V and N have equal mass m. Their energy of interaction is ω_0. The scattering part of (1) is given by

$$|\chi\rangle^+ = Z_0 Z^{\frac{1}{2}} \sum_{k'} \psi_1(\omega;\omega')|1_V, 1_N, k'\rangle + \sum_{k'k''} \psi_2(\omega;\omega',\omega'')|2_N, k', k''\rangle, \quad (2)$$

with outgoing waves only in $\psi_1(\omega;\omega')$ and $\psi_2(\omega;\omega',\omega'')$. The real constant Z_0 normalizes the state $|VN\rangle$ to unity. The TD bare states

$$\begin{aligned} |2_V\rangle &= 2^{-\frac{1}{2}} Z \psi_V^+ \psi_V^+ |0\rangle, \\ |1_V, 1_N, k\rangle &= Z^{\frac{1}{2}} \psi_V^+ \psi_N^+ a_k^+ |0\rangle = a_k^+ |1_V, 1_N\rangle, \\ |2_N, k, k'\rangle &= 2^{-1} \psi_N^+ \psi_N^+ a_k^+ a_{k'}^+ |0\rangle = 2^{-\frac{1}{2}} a_{k'}^+ |2_N, k\rangle \end{aligned} \quad (3)$$

are all normalized to unity or Kronecker delta functions in a discrete momentum space. The state $|VN\rangle$

* This work was supported in part by the U.S. Army Research Office (Durham) under the contract DA-31-124-ARO(D)-172.

[1] L. M. Scarfone, J. Math. Phys. **9**, 246 (1968).
[2] M. S. Maxon and R. B. Curtis, Phys. Rev. **137**, B996 (1965).
[3] L. M. Scarfone, J. Math. Phys. **7**, 159 (1966).
[4] M. S. Maxon, Phys. Rev. **149**, 1273 (1966).
[5] A. Pagnamenta, J. Math. Phys. **6**, 955 (1965).

has the form

$$|VN\rangle = Z_0\Big[Z^{\frac{1}{2}}|1_V, 1_N\rangle - 2^{\frac{1}{2}}g\sum_k \frac{X(\omega)}{\omega - \omega_0}|2_N, k\rangle\Big], \tag{4}$$

where $X(\omega)$ is an abbreviation for the ratio $f(\omega)/(2\omega\Omega)^{\frac{1}{2}}$ in which $f(\omega)$ is the usual cutoff factor and Ω the volume of quantization. Noting that the Fourier transform of the VN propagator, evaluated at the energy $W + 2m$, is expressed in continuous space as the inverse of the function

$$G^+(W) \equiv G(W + i\epsilon) = W + \frac{W^2}{2\pi}\int_\mu^\infty \frac{\operatorname{Im} G^+(\omega)\, d\omega}{\omega^2(\omega - W - i\epsilon)} + \frac{1}{2\pi}\int_\mu^\infty \frac{\operatorname{Im} G^+(\omega)\, d\omega}{\omega - W - i\epsilon}, \tag{5}$$

we have

$$Z_0^{-2} = \left.\frac{dG^+(W)}{dW}\right|_{W=\omega_0} \equiv G'(\omega_0). \tag{6}$$

The positive number ϵ is to be treated as infinitesimally small. We recall that $\operatorname{Im} G^+(\omega) = (2\pi)^{-1}g^2kf^2(\omega)$, where g is the renormalized coupling constant. It will be assumed throughout that g is less than its critical value (no ghost condition) and thus ω_0 is a real (negative) and single-valued function of g. The condition $G(\omega_0) = 0$ determines ω_0 for the case where V and N have a zero spatial separation. It is useful to write

$$G^+(W) = (W - \omega_0)\alpha^+(W), \tag{7}$$

which defines

$$\alpha^+(W) = Z_0^{-2} + \frac{W - \omega_0}{\pi}\int_\mu^\infty \frac{\operatorname{Im} G^+(\omega)\, d\omega}{(\omega - \omega_0)^2(\omega - W - i\epsilon)} = Z + \frac{1}{\pi}\int_\mu^\infty \frac{\operatorname{Im} \alpha^+(\omega)\, d\omega}{\omega - W - i\epsilon} \tag{8}$$

and shows that $\alpha(\omega_0) \equiv \alpha_0 = Z_0^{-2}$. We also remember that

$$Z = 1 - (2\pi)^{-1}\int_\mu^\infty \frac{\operatorname{Im} G^+(\omega)\, d\omega}{\omega^2}. \tag{9}$$

Using the Hamiltonian and the quantization rules as given in Ref. 3 by Eqs. (1) and (3), we find that the eigenvalue problem

$$H\,|VN\theta_{k,\mathrm{in}}\rangle = (2m + \omega_0 + \omega)\,|VN\theta_{k,in}\rangle \tag{10}$$

yields the following set of equations for the determination of the expansion coefficients:

$$Z_0(\omega_0 + \omega - 2\delta m)C(\omega) = 2^{\frac{1}{2}}gZ_0X(\omega) + 2^{\frac{1}{2}}gZ_0\sum_{k'} X(\omega')\psi_1(\omega;\omega'), \tag{11a}$$

$$Z^{\frac{1}{2}}Z_0(\omega_0 + \omega - \omega' - \delta m)\psi_1(\omega;\omega') = \frac{-2Z^{-\frac{1}{2}}g^2Z_0X(\omega)X(\omega')}{\omega' - \omega_0} + (2Z)^{\frac{1}{2}}gX(\omega')C(\omega) + 2Z^{-\frac{1}{2}}g\sum_{k''} X(\omega'')\psi_2(\omega;\omega',\omega''), \tag{11b}$$

$$\psi_2(\omega;\omega',\omega'') = gZ_0\frac{[X(\omega')\psi_1(\omega;\omega'') + X(\omega'')\psi_1(\omega;\omega')]}{\omega_0 + \omega - \omega' - \omega'' + i\epsilon}. \tag{11c}$$

Clearly, $\psi_2(\omega;\omega',\omega'')$ is symmetric under interchange of ω' and ω''. The expression for ψ_2 incorporates the boundary condition stated earlier. By making the change $i\epsilon \to -i\epsilon$, we automatically get the boundary condition for plane waves plus incoming waves, as required for the out states. Combining (11c) and (11b) and using

$$\delta m = (2\pi Z)^{-1}\int_\mu^\infty \frac{\operatorname{Im} G^+(\omega)\, d\omega}{\omega}, \tag{12}$$

we obtain the singular integral equation

$$\Psi(\omega;\omega') = \frac{2^{-\frac{1}{2}}ZC(\omega)}{gZ_0X(\omega)} - \frac{1}{\omega' - \omega_0} - \frac{1}{\pi}\int_\mu^\infty \frac{\operatorname{Im}[G^+(\omega'')]\Psi(\omega;\omega'')\, d\omega''}{(\omega'' + \omega' - \omega - \omega_0 - i\epsilon)G^+(\omega_0 + \omega - \omega'')}, \tag{13}$$

where

$$\Psi(\omega;\omega') \equiv \frac{G^+(\omega_0 + \omega - \omega')\psi_1(\omega;\omega')}{2g^2X(\omega)X(\omega')}. \tag{14}$$

A comparison between this equation and I(41), namely,

$$T(W;Z',\omega) = L^-(W;\omega) + \frac{1}{W - Z' - \omega} - \frac{1}{\pi}\int_\mu^\infty \frac{\operatorname{Im}[G^+(\omega'')]T^-(W;\omega'',\omega)\, d\omega''}{(\omega'' + Z' - W)G^+(W - \omega'')}, \tag{15}$$

shows that

$$\Psi(\omega;\omega') = \lim_{\substack{W\to\omega+\omega_0\\ Z'\to\omega'-i\epsilon}} T(W;Z',\omega) = T^-(\omega + \omega_0;\omega',\omega) \equiv T(\omega + \omega_0;\omega' - i\epsilon,\omega). \tag{16}$$

This relation, together with (14) and I(40), leads to

$$\psi_1(\omega;\omega') = \frac{2g^2X(\omega)X(\omega')T^-(\omega + \omega_0;\omega',\omega)}{G^+(\omega_0 + \omega - \omega')} = \frac{X(\omega)X(\omega')}{G^+(\omega + \omega_0 - \omega')} \times \{G^+(W - \omega)[G^+(W - \omega')\hat{\tau}_6(W + 2m;\omega',\omega) - \delta_{kk'}X^2(\omega')]\}_{W=\omega+\omega_0}, \tag{17}$$

where $\hat{\tau}_6(W+2m;\omega',\omega)$ corresponds to the Green's function for $VN\theta$ scattering. In making the comparison indicated above, we equate the first term on the right-hand side of (13) with $L^-(\omega+\omega_0;\omega)$ of (15). According to I(39), this identification implies that

$$C(\omega)=2^{-\frac{1}{2}}Z_0X(\omega)G^+(W-\omega)\hat{\tau}_4(W+2m;\omega)\big|_{W=\omega+\omega_0}, \tag{18}$$

where $\hat{\tau}_4(W+2m;\omega)$ describes the vertex $2V|VN\theta$. The connection between this function and the $2V$ propagator $\hat{\tau}_1(W+2m)$ is given in I(24). We can use (18) with the known expressions for these $\hat{\tau}$'s, or we can substitute (17) into (11a) to find

$$Z_0C(\omega)=\frac{2^{\frac{3}{2}}gX(\omega)}{D(\omega+\omega_0)}, \tag{19}$$

where D, the denominator function in $\hat{\tau}_1$, is written as

$$D(W)=Z^2(W-2\delta m)[1+\alpha_0G^+(W-\omega_0)\times I_W^+(W-\omega_0)]+\alpha_0G^+(W-\omega_0). \tag{20}$$

The integral $I_W(Z')$ has the form

$$I_W(Z')=\frac{1}{\pi}\int_\mu^\infty \mathrm{Im}\left[\frac{1}{G^+(\omega)}\right]\frac{d\omega}{(\omega-Z')\alpha^+(W-\omega)}. \tag{21}$$

If (11a) is used to derive (19), it should be noted that $C(\omega)$ enters $\psi_1(\omega;\omega')$ through $\hat{\tau}_6(W+2m;\omega',\omega)$ and that the integral involving $I^+_{\omega+\omega_0}(\omega+\omega_0-\omega')$ has been treated in the Appendix to I. We complete the states (1) by writing

$$\psi_2(\omega;\omega',\omega'')=\frac{2g^3Z_0X(\omega)X(\omega')X(\omega'')}{\omega+\omega_0-\omega'-\omega''+i\epsilon}\times\left[\frac{T^-(W;\omega',\omega)}{G^+(W-\omega')}+\frac{T^-(W;\omega'',\omega)}{G^+(W-\omega'')}\right]_{W=\omega+\omega_0}. \tag{22}$$

At this point let us record [see I(40) and I(43)] the following:

$$\begin{aligned}\frac{T^-(\omega+\omega_0;\omega',\omega)}{G^+(\omega+\omega_0-\omega')}&=-\frac{d^+(\omega)}{(\omega-\omega'+i\epsilon)(\omega'-\omega_0)\alpha^+(\omega)}\\&\quad-\frac{\alpha_0I^+_{\omega+\omega_0}(\omega+\omega_0-\omega')}{\omega'-\omega_0}-\frac{\alpha_0}{\omega'-\omega_0}\left[\frac{d^+(\omega)}{d^-(\omega)}\right]\\&\quad\times\left[I^+_{\omega+\omega_0}(\omega+\omega_0-\omega')+\frac{(\omega-\omega_0)A(\omega)}{\omega-\omega'+i\epsilon}\right]\\&\quad+\frac{2^{-\frac{1}{2}}C(\omega)}{gZ_0X(\omega)(\omega-\omega'+i\epsilon)(\omega'-\omega_0)d^-(\omega)}\\&\quad\times[(\omega'-\omega_0)d^-(\omega)+(\omega-\omega')\alpha_0G^+(\omega)\\&\quad\times I^+_{\omega+\omega_0}(\omega+\omega_0-\omega')+(\omega-\omega_0)\alpha_0G^+(\omega)A(\omega)],\end{aligned} \tag{23}$$

with

$$A(\omega)\equiv -I_{\omega+\omega_0}(\omega)$$

and

$$d^\pm(\omega)\equiv 1\pm\alpha_0G^+(\omega)A(\omega).$$

The conventional T matrix for $VN\theta$ scattering, defined by

$$\begin{aligned}S^{VN\theta}_{kk'}&=\langle VN\theta_{k',\mathrm{out}}\,|\,VN\theta_{k,\mathrm{in}}\rangle\\&=\delta_{kk'}-2\pi i\delta(\omega-\omega')T_{VN\theta}(\omega'),\end{aligned} \tag{24}$$

is obtained in a straightforward way by computing the residue of $\psi_1(\omega;\omega')$ at the pole $\omega=\omega'$. The proof of this follows the methods of Ref. 5, Appendix II. Thus with (17) and (23) we have

$$\begin{aligned}T_{VN\theta}(\omega')&=\lim_{\omega\to\omega'}(\omega-\omega')\psi_1(\omega;\omega')\\&=-2g^2\alpha_0X^2(\omega')\left[\frac{d^+(\omega')}{\alpha_0G^+(\omega')d^-(\omega')}-\frac{2}{D(\omega'+\omega_0)d^-(\omega')}\right].\end{aligned} \tag{25}$$

As in I, we wish to remark that the denominator factor $d^-(\omega')$ cancels out when the terms in the brackets are combined. The ω' dependence in the remaining denominator is contained in the product $G^+(\omega')D(\omega'+\omega_0)$ which consists of a factor $G^+(\omega')$, describing a three-particle scattering, and a factor $D(\omega'+\omega_0)$, representing the full four-body interaction. For reference we recall the $2N\theta$ scattering element

$$S^{2N\theta}_{kk'}=\delta_{kk'}-2\pi i\delta(\omega-\omega')\frac{2g^2X^2(\omega)}{G^+(\omega)}. \tag{26}$$

Finally, we write, as in I,

$$S^{VN\theta}_{kk'}=\delta_{kk'}+4\pi ig^2X^2(\omega)\delta(\omega-\omega')\times\frac{[Z^2(\omega+\omega_0-2\delta m)d^+(\omega)-\alpha_0G^+(\omega)]}{G^+(\omega)D(\omega+\omega_0)}. \tag{27}$$

Next, we consider the scattering of two θ's by two N's. For this collision process we deal with states of the form

$$\begin{aligned}|2N\theta_k\theta_{k',\mathrm{in}}\rangle&=|2_N,k,k'\rangle+ZC(\omega,\omega')|2_V\rangle\\&\quad+Z^{\frac{1}{2}}\sum_q\Phi_1(\omega,\omega';\omega_q)|1_V,1_N,q\rangle\\&\quad+\sum_{qq'}\Phi_2(\omega,\omega';\omega_q,\omega_q')|2_N,q,q'\rangle,\end{aligned} \tag{28}$$

with eigenvalues $2m+\omega+\omega'$, and with outgoing

waves only in Φ_1 and Φ_2. The usual procedure gives

$$Z(\omega+\omega'-2\delta m)C(\omega,\omega') = 2^{\frac{1}{2}}g\sum_q X(\omega_q)\Phi_1(\omega,\omega';\omega_q), \quad (29a)$$

$$Z(\omega+\omega'-\omega_q-\delta m)\Phi_1(\omega,\omega';\omega_q) = g[\delta_{qk'}X(\omega)+\delta_{qk}X(\omega')] + 2^{\frac{1}{2}}gZC(\omega,\omega')X(\omega_q) + 2g\sum_{q'}X(\omega_{q'})\Phi_2(\omega,\omega';\omega_q,\omega_{q'}), \quad (29b)$$

$$\Phi_2(\omega,\omega';\omega_q,\omega_q') = \frac{g[X(\omega_q')\Phi_1(\omega,\omega';\omega_q)+X(\omega_q)\Phi_1(\omega,\omega';\omega_q')]}{\omega+\omega'-\omega_q-\omega_q'+i\epsilon}. \quad (29c)$$

Note that $\Phi_2(\omega,\omega';\omega_q,\omega_q') = \Phi_2(\omega,\omega';\omega_q',\omega_q)$. Eliminating Φ_2 from (29b) and using (9) and (12), we find the singular integral equation

$$M^-(\omega,\omega';\omega_q) = C_0(\omega,\omega') + \frac{G^+(\omega)}{\omega_q-\omega'-i\epsilon} + \frac{G(\omega')}{\omega_q-\omega-i\epsilon} - \frac{1}{\pi}\int_\mu^\infty \frac{\operatorname{Im}[G^+(\omega_q')]M^-(\omega,\omega';\omega_q')\,d\omega_q'}{(\omega_q+\omega_q'-\omega-\omega'-i\epsilon)G^+(\omega+\omega'-\omega_{q'})}, \quad (30)$$

with the definitions

$$C_0(\omega,\omega') = \frac{-2^{\frac{1}{2}}ZC(\omega,\omega')G^+(\omega)G^+(\omega')}{2g^2X(\omega)X(\omega')}, \quad (31)$$

$$M^-(\omega,\omega';\omega_q) \equiv M(\omega,\omega';\omega_q-i\epsilon) = \frac{G^+(\omega)G^+(\omega')}{2g^3X(\omega)X(\omega')X(\omega_q)}[g\delta_{qk'}X(\omega)+g\delta_{qk}X(\omega') - G^+(\omega+\omega'-\omega_q)\Phi_1(\omega,\omega';\omega_q)]. \quad (32)$$

We look upon this equation as determining the ω_q dependence of $M^-(\omega,\omega';\omega_q)$ for fixed ω and ω'. To solve it we could, for example, use the analytic methods of Ref. 5, or transform it into an inhomogeneous Hilbert problem as in Ref. 4 and I. In effect we adopt the latter method by making the identification

$$M^-(\omega,\omega';\omega_q) = -[G^+(\omega)T^-(\omega+\omega';\omega_q,\omega) + G^+(\omega')T^-(\omega+\omega';\omega_q,\omega')], \quad (33)$$

which means that

$$\Phi_1(\omega,\omega';\omega_q) = gX(\omega)X(\omega')X(\omega_q)[\hat\tau_6(W+2m;\omega_q,\omega) + \hat\tau_6(W+2m;\omega_q,\omega')]_{W=\omega+\omega'} \quad (34)$$

and

$$C(\omega,\omega') = 2^{-\frac{1}{2}}gX(\omega)X(\omega')[\hat\tau_4(W+2m;\omega) + \hat\tau_4(W+2m;\omega')]_{W=\omega+\omega'}. \quad (35)$$

According to (29c) and (34), the remaining coefficients become

$$\Phi_2(\omega,\omega';\omega_q,\omega_q') = \frac{g^2X(\omega)X(\omega')X(\omega_q)X(\omega_q')}{\omega+\omega'-\omega_q-\omega_q'+i\epsilon} \times [\hat\tau_6(W+2m;\omega_q,\omega)+\hat\tau_6(W+2m;\omega_q,\omega') + \hat\tau_6(W+2m;\omega_q',\omega) + \hat\tau_6(W+2m;\omega_q',\omega')]_{W=\omega+\omega'}, \quad (36)$$

and with this we have completely determined the state vectors (28).

Now we obtain the corresponding T matrix defined by

$$S^{2N2\theta}_{kk';qq'} = \langle 2N\theta_q\theta_{q',\mathrm{out}} \mid 2N\theta_k\theta_{k',\mathrm{in}}\rangle = \frac{\delta_{qk}\delta_{q'k'}+\delta_{qk'}\delta_{q'k}}{2} - 2\pi i\delta(\omega+\omega'-\omega_q-\omega_q')T_{2N2\theta}(\omega,\omega';\omega_q). \quad (37)$$

Applying the residue method as in the previous case, and using (17), (23), and I(23), we find

$$T_{2N2\theta}(\omega,\omega';\omega_q) = \lim_{\omega_{q'}\to\omega+\omega'-\omega_q}(\omega'+\omega-\omega_q-\omega_q')\Phi_2(\omega,\omega';\omega_q,\omega_q') = g^2X(\omega)X(\omega')X(\omega_q)X(\omega_q') \times \left\{\left[\frac{X^{-2}(\omega_q)}{G^+(\omega_q')}(\delta_{qk}+\delta_{qk'})+\frac{X^{-2}(\omega_q')}{G^+(\omega_q)}(\delta_{q'k}+\delta_{q'k'})\right] \frac{-4g^2Z^2(\omega_q+\omega_q'-2\delta m)\alpha_0G^+(\omega_q+\omega_q'-\omega_0)}{G^+(\omega)G^+(\omega')G^+(\omega_q)G^+(\omega_q')D(\omega_q+\omega_q')}\right\}, \quad (38)$$

with the understanding that $\omega_q' = \omega+\omega'-\omega_q$. The completely connected term in (38) contains the factors $[G^+(\omega)G^+(\omega')]^{-1}$ and $[G^+(\omega_q)G^+(\omega_q')]^{-1}$, which represent the three-body interactions in the initial and final states, respectively. The factor $D^{-1}(\omega_q+\omega_q')$ represents, of course, the four-body interaction. Thus the interpretation here parallels that of $T_{N2\theta}$, which has similar factors for the initial and final two-body interactions and a factor attributed to the full three-particle interaction.

To conclude this section we introduce the production amplitude P defined by

$$S_{k;k'k''} = \langle 2N\theta_{k'}\theta_{k'',\mathrm{out}} \mid VN\theta_{k,\mathrm{in}}\rangle = -2\pi i\delta(\omega_0+\omega-\omega'-\omega'')P(\omega;\omega'). \quad (39)$$

We have only to observe that P is given by that part of $\psi_2(\omega;\omega',\omega'')$ containing $\delta(\omega_0+\omega-\omega'-\omega'')$.

Hence from (22) we get

$$P(\omega;\omega') = 2g^3 Z_0 X(\omega)X(\omega')X(\omega'') \times \left[\frac{T^-(\omega+\omega_0;\omega',\omega)}{G^+(\omega+\omega_0-\omega')} + \frac{T^-(\omega+\omega_0;\omega'',\omega)}{G^+(\omega+\omega_0-\omega'')}\right]_{\omega''=\omega_0+\omega-\omega'}, \quad (40)$$

which, with (23), leads to

$$S_{k;k'k''} = 8\pi i\alpha_0^{\frac{1}{2}} X(\omega')X(\omega'')X(\omega'+\omega''-\omega_0) \times \delta(\omega_0+\omega-\omega'-\omega'') \times \frac{Z^2(\omega'+\omega''-2\delta m)}{G^+(\omega')G^+(\omega'')D(\omega'+\omega'')}. \quad (41)$$

Here we have exhibited the symmetry in the ω variables of the outgoing θ's. The factor $[G^+(\omega')G^+(\omega'')]^{-1}$ describes the final state interaction between the $2N$ particles and each of the θ's.

III. BOUND STATE

Some well-known examples of bound states in the higher sectors of the Lee model have been discussed in the literature from several points of view. Thus the dynamical $V\theta$ bound state has been investigated via the TD scheme,[6] the N-quantum approximation[7] (NQA), the LSZ formulation,[4] and the techniques of dispersion theory.[8] The VN system has also been studied along these lines[3,9,10] (except for the NQA) and with the related Bethe–Salpeter equation.[11] In this section we derive the general $2V$ bound state in the TD framework, and once again the development exploits the connection between the expansion coefficients and $\hat{\tau}$ functions. Unlike the VN case, we are here involved with an exactly soluble two-meson exchange problem.

We seek a stationary eigenstate of the total Hamiltonian of the general form

$$|B\rangle = Z_B\Big[Z|2_V\rangle + Z^{\frac{1}{2}}\sum_k \varphi_1(\omega)|1_V,1_N,k\rangle + \sum_{kk'}\varphi_2(\omega,\omega')|2_N,k,k'\rangle\Big], \quad (42)$$

where Z_B is a normalization constant. Denoting the corresponding eigenvalue by $E_B = 2m + \omega_B$, with $\omega_B < \omega_0 + \mu < 2\mu$ for stability, and comparing terms on both sides of $H|B\rangle = E_B|B\rangle$, we come up with the following set of equations:

$$Z(\omega_B - 2\delta m) = 2^{\frac{1}{2}} g\sum_k X(\omega)\varphi_1(\omega), \quad (43a)$$

$$Z^{\frac{1}{2}}(\omega_B-\omega-\delta m)\varphi_1(\omega) = (2Z)^{\frac{1}{2}} gX(\omega) + 2Z^{-\frac{1}{2}} g\sum_{k'} X(\omega')\varphi_2(\omega,\omega'), \quad (43b)$$

$$\varphi_2(\omega,\omega') = \frac{g[X(\omega)\varphi_1(\omega') + X(\omega')\varphi_1(\omega)]}{\omega_B-\omega-\omega'}. \quad (43c)$$

Using the last relation in (43b) to eliminate φ_2, we obtain the integral equation

$$\varphi(\omega) = 1 - \frac{1}{\pi}\int_\mu^\infty \frac{\mathrm{Im}\,[G^+(\omega')]\varphi(\omega')\,d\omega'}{(\omega'+\omega-\omega_B)G(\omega_B-\omega')}, \quad (44)$$

in which

$$\varphi(\omega) \equiv \frac{2^{-\frac{1}{2}} G(\omega_B-\omega)\varphi_1(\omega)}{gZX(\omega)}. \quad (45)$$

Appealing to I(19) we note that

$$\varphi(\omega) = \lim_{\substack{W\to\omega_B \\ Z'\to\omega+i\epsilon}} J(W;Z') = J(\omega_B;\omega), \quad (46)$$

where, according to I(21),

$$J(\omega_B;\omega) = \frac{G(\omega_B-\omega)}{Zd^-(x)}\left\{\frac{1}{x-\omega} + \frac{\alpha_0 G(x)}{\omega-\omega_0} \times [I_{\omega_B}(\omega_B-\omega) + A(x)]\right\}. \quad (47)$$

We let x be an abbreviation for $\omega_B - \omega_0 = E_B - E_0$. Thus

$$\varphi_1(\omega) = \frac{2^{\frac{1}{2}} gZX(\omega)J(\omega_B;\omega)}{G(\omega_B-\omega)}, \quad (48)$$

or, in terms of $\hat{\tau}$ functions, we have

$$\varphi_1(\omega)\hat{\tau}_1(\omega_B+2m) = 2^{\frac{1}{2}} X(\omega)\hat{\tau}_4(\omega_B+2m;\omega). \quad (49)$$

From (43c) and (48), we conclude that

$$\varphi_2(\omega,\omega') = \frac{2^{\frac{1}{2}} Zg^2 X(\omega)X(\omega')}{\omega_B-\omega-\omega'} \times \left[\frac{J(\omega_B;\omega)}{G(\omega_B-\omega)} + \frac{J(\omega_B;\omega')}{G(\omega_B-\omega')}\right]. \quad (50)$$

At this point we make use of (47) and (48) in (43a) to obtain an equation which determines ω_B. As expected this equation also follows from the vanishing of $D(W)$ at $W = \omega_B$ and can be expressed in the form

$$F(x) = 2m_0 - E_0 - x, \quad (51)$$

where $m_0 = m + \delta m$ is the bare mass parameter of the V particle and $F(x)$ is defined by

$$F(x) = \alpha_0 G(x)/Z^2 d^-(x). \quad (52)$$

Due to the restrictions on ω_B and ω_0, the functions

[6] A. Pagnamenta, J. Math. Phys. **7**, 356 (1966).
[7] A. Pagnamenta, Ann. Phys. (Paris) **39**, 453 (1966).
[8] T. Muta, Progr. Theoret. Phys. (Kyoto) **33**, 666 (1965).
[9] S. Weinberg, Phys. Rev. **102**, 285 (1956).
[10] L. M. Scarfone, Nucl. Phys. **39**, 658 (1962).
[11] N. Mugibayashi, Progr. Theoret. Phys. (Kyoto) **25**, 803 (1961).

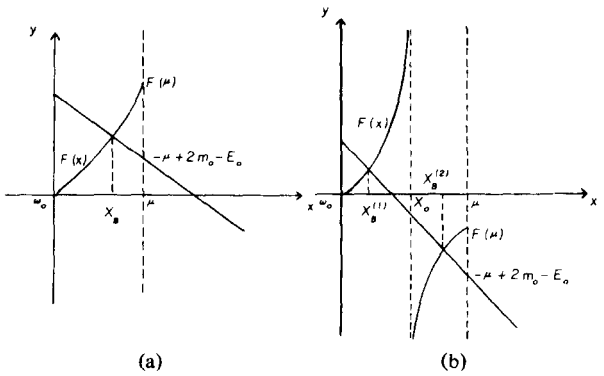


FIG. 1. Schematic representation of the eigenvalue equation (51) in the segment $\omega_0 \leq x \leq \mu$ for the case (a) that $d^-(x)$ has no zero, and the case (b) that $d^-(x)$ has a zero.

$G(x)$ and $A(x)$ are real and $x < \mu$. To make plausible the existence of a stable $2V$ bound-state root ω_B, let us first note that both $G(x)$ and $A(x)$ are positive and have positive slopes in the segment $\omega_0 < x < \mu$. Thus in this segment the product $\alpha_0 G(x)A(x)$ is a monotonically increasing function of x which starts from the value zero at $x = \omega_0$. Also $G(\mu) > 0$ and $A(\mu) > 0$, so that $\alpha_0 G(\mu)A(\mu) > 0$. For $|x| \to \infty$ we find that $\alpha_0 G(x)A(x)$ approaches the negative number $1 - Z^{-1}\alpha_0$. The inequality $\alpha_0 > Z$ follows from (8) and the "no ghost" condition $0 < Z < 1$. If we assume that the function $d^-(x)$ does not vanish, or, equivalently, that $d^-(\mu) > 0$, then $F(x)$ is monotonically increasing for $x < \mu$ and the value of ω_B is determined by the intersection of the curve $y = F(x)$ and the line $y = -x + 2m_0 - E_0$. An intersection is possible if m_0 is chosen such that the inequality $-\mu + 2m_0 - E_0 < F(\mu)$ is satisfied [Fig. 1(a)]. We also remark that $[F(x)/x] \to 1$ as $|x| \to \infty$. Since $A(x)$ is proportional to g^2, we expect the condition $d^-(\mu) > 0$ to hold for small g^2 provided that the cutoff factor is chosen judiciously. For g^2 large and still less than its critical value, $d^-(x)$ may have a zero x_0 in the region $\omega_0 < x < \mu$. This will be so if $d^-(\mu) < 0$. In such a case $F(x)$ would have the behavior shown in Fig. 1(b). It then appears that, however large m_0 is made, there is always a root $x_B^{(1)}$ between ω_0 and x_0. Here it may not always be possible to fulfill the inequality $-\mu + 2m_0 - E_0 < F(\mu) < 0$, in which case the root $x_B^{(2)}$ is nonexistent. The actual demonstration of the possible roots of (52) would, of course, require a specific choice for the cutoff function. The situations that could arise for $x > \mu$ are beyond our present considerations; however, if the V particle sector serves as a guide, it may be possible to have both bound states and resonances simultaneously present.[12]

[12] L. Fonda, G. C. Ghirardi, and A. Rimini, Phys. Rev. 133, B196 (1964).

Also we have assumed from the beginning that $0 < Z < 1$, or, equivalently, that $G(\omega)$ has only one root and that $\alpha(\omega)$ has none. Negative values of Z would lead to further zeros of $G(\omega)$ corresponding to stable VN bound states with real energy and negative norm, or with complex energy and zero norm.[9]

The normalization constant Z_B may be determined by calculating the residue of $\hat{\tau}_1(W + 2m)$ at the pole $W = \omega_B$:

$$2Z_B^2 = \lim_{W \to \omega_B} (W - \omega_B)\hat{\tau}_1(W + 2m). \tag{53}$$

This leads to

$$2Z_B^{-2} = \left\{\frac{d}{dW}\left[\frac{D(W)}{d^-(W - \omega_0)}\right]\right\}_{W=\omega_B}$$
$$= d^{-2}(x)[Z^2 d^2(x) + \alpha_0 G'(x) + \alpha_0^2 G^2(x)A'(x)], \tag{54}$$

which shows that $Z_B^{-2} > 0$; thus we can choose $Z_B > 0$, and $|B\rangle$ is normalizable. The same conclusion follows from the scalar product $\langle B \mid B\rangle = 1$.

IV. CONCLUDING REMARKS

The results obtained in this paper and in I are of methodological interest in that they extend the solved part of the ordinary Lee model. We have found that the mathematical structure of the $2V$ sector is slightly more complicated than the $V\theta$ sector, but their close correspondence is very evident. It is clear that solutions in the $V\theta$ sector have made it possible to obtain solutions in the present case. Basically, this is because we had only to add the trivial N particle to V and θ in order to get the $2V$ system. In carrying out the TD treatment of this system, we have exploited the connection between expansion coefficients and $\hat{\tau}$ functions. Had we proceeded independently of previous results, we could have solved the TD singular integral equations with other techniques such as the analytic methods used in Ref. 5. The existence of a $2V$ bound state is, of course, less surprising than the possibility of a $V\theta$ bound state; but, having solved the $2V$ sector, we could next inquire into the scattering of a θ by both renormalizable sources and study the likelihood of a dynamical three-body bound state.

The Bethe–Salpeter equation for the VN bound state has already been studied as an exactly soluble example with no redundant states.[11] This equation gives the same total energy for the VN system as that predicted by other methods of solution. In this case the so-called ladder approximation is exact. As it is not entirely clear whether the existence of these states is inherent in the conventional field-theoretic

treatment of the bound-state problem or whether it is due to the approximation employed, it is very desirable to have completely soluble examples. This suggests another interesting possibility, namely, that we study the $2V$ problem for the purpose of showing whether or not the usual field-theoretic procedure gives an integral equation for the Bethe–Salpeter amplitude $\langle 0|\, T(\psi_V(t_1)\psi_V(t_2))\,|B\rangle$ that reproduces the results found in our present work. In this case the ladder approximation is inadequate as we must consider an interaction which is transmitted by two θ's simultaneously. This problem is presently under investigation.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 7 JULY 1968

Derivation of Newton's Law of Gravitation from General Relativity*

AMNON KATZ†
Argonne National Laboratory, Argonne, Illinois

(Received 18 September 1967)

A static situation of two objects held apart by a strut is considered within the framework of general relativity, and the gravitational attraction between the objects is inferred from the stress in the strut. In the Newtonian limit—the objects are well separated and the field weak everywhere except in their immediate vicinity—Newton's law of gravitation is reproduced. This check goes well beyond verifications of the Newtonian limit based on considerations of "test particles," since arbitrarily strong self-fields are not excluded for either object. It is also an explicit verification of the equality of active and passive gravitational masses.

1. INTRODUCTION

The purpose of this paper is to compute the static gravitational attraction of two objects within the framework of general relativity and to verify that Newton's formula is indeed obtained under Newtonian conditions. By Newtonian conditions it is meant that the separation between the two objects is very much larger than the Schwarzschild radius of either, so that space is practically flat everywhere except in the immediate vicinity or interior of the objects. Each object is in the weak field of the other. Most of the space between them is in the weak-field region of both. However, self-fields of arbitrary strength are not excluded for either object, and the active gravitational masses of the objects may reflect a concentration of stress and not necessarily energy. In this respect, our check of the Newtonian limit goes far beyond considerations involving fictitious test particles that follow geodesics and produce no field. Even if the geodesic hypothesis is taken for granted, our check may be regarded as a verification that the passive gravitational mass is equal to the active gravitational mass.

As gravitational fields of arbitrary strength cannot be excluded for the interior and immediate vicinity of elementary particles, these are the conditions under which general relativity must reproduce Newtonian results.

We shall consider only situations with cylindrical symmetry around the line joining the two objects. The objects themselves will be confined to regions much smaller than their separation so that the asymptotic field of each object at the position of the other and over most of the space between them is spherically symmetric. (This situation seems well suited for applications to the gravitational attraction between elementary particles.)

For the actual computation we shall set up a "strut" holding the two objects apart. The force of attraction will be computed from the stress in the strut. The construction of the strut will be the most general possible, subject only to the conditions of cylindrical symmetry and weightlessness. This last condition means that the strut will be kept light enough so that the weight of the strut itself is prevented from making the attraction between the two objects ambiguous. The strut need not be visualized as a rigid body; it may be any static concentration of stress capable of balancing the gravitational attraction.

2. TWO OBJECTS HELD APART BY A STRUT

We wish to study a static, cylindrically symmetric situation in which two bodies of revolution are situated on the z axis and held apart against their

* Work performed under the auspices of the U.S. Atomic Energy Commission.

† On leave from the Weizmann Institute, Rehovoth, Israel. Present address: Physics Department, University of Washington, Seattle, Washington.

treatment of the bound-state problem or whether it is due to the approximation employed, it is very desirable to have completely soluble examples. This suggests another interesting possibility, namely, that we study the $2V$ problem for the purpose of showing whether or not the usual field-theoretic procedure gives an integral equation for the Bethe–Salpeter amplitude $\langle 0|\, T(\psi_V(t_1)\psi_V(t_2))\,|B\rangle$ that reproduces the results found in our present work. In this case the ladder approximation is inadequate as we must consider an interaction which is transmitted by two θ's simultaneously. This problem is presently under investigation.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 7 JULY 1968

Derivation of Newton's Law of Gravitation from General Relativity*

AMNON KATZ†
Argonne National Laboratory, Argonne, Illinois

(Received 18 September 1967)

A static situation of two objects held apart by a strut is considered within the framework of general relativity, and the gravitational attraction between the objects is inferred from the stress in the strut. In the Newtonian limit—the objects are well separated and the field weak everywhere except in their immediate vicinity—Newton's law of gravitation is reproduced. This check goes well beyond verifications of the Newtonian limit based on considerations of "test particles," since arbitrarily strong self-fields are not excluded for either object. It is also an explicit verification of the equality of active and passive gravitational masses.

1. INTRODUCTION

The purpose of this paper is to compute the static gravitational attraction of two objects within the framework of general relativity and to verify that Newton's formula is indeed obtained under Newtonian conditions. By Newtonian conditions it is meant that the separation between the two objects is very much larger than the Schwarzschild radius of either, so that space is practically flat everywhere except in the immediate vicinity or interior of the objects. Each object is in the weak field of the other. Most of the space between them is in the weak-field region of both. However, self-fields of arbitrary strength are not excluded for either object, and the active gravitational masses of the objects may reflect a concentration of stress and not necessarily energy. In this respect, our check of the Newtonian limit goes far beyond considerations involving fictitious test particles that follow geodesics and produce no field. Even if the geodesic hypothesis is taken for granted, our check may be regarded as a verification that the passive gravitational mass is equal to the active gravitational mass.

As gravitational fields of arbitrary strength cannot be excluded for the interior and immediate vicinity of elementary particles, these are the conditions under which general relativity must reproduce Newtonian results.

We shall consider only situations with cylindrical symmetry around the line joining the two objects. The objects themselves will be confined to regions much smaller than their separation so that the asymptotic field of each object at the position of the other and over most of the space between them is spherically symmetric. (This situation seems well suited for applications to the gravitational attraction between elementary particles.)

For the actual computation we shall set up a "strut" holding the two objects apart. The force of attraction will be computed from the stress in the strut. The construction of the strut will be the most general possible, subject only to the conditions of cylindrical symmetry and weightlessness. This last condition means that the strut will be kept light enough so that the weight of the strut itself is prevented from making the attraction between the two objects ambiguous. The strut need not be visualized as a rigid body; it may be any static concentration of stress capable of balancing the gravitational attraction.

2. TWO OBJECTS HELD APART BY A STRUT

We wish to study a static, cylindrically symmetric situation in which two bodies of revolution are situated on the z axis and held apart against their

* Work performed under the auspices of the U.S. Atomic Energy Commission.

† On leave from the Weizmann Institute, Rehovoth, Israel. Present address: Physics Department, University of Washington, Seattle, Washington.

gravitational attraction by a strut running between them along the z axis. This situation has been studied by Synge,[1] who shows that the empty-space metric may be brought to the form

$$ds^2 = e^{2\lambda}c^2\,dt^2 - e^{-2\lambda}[e^{2\nu}(d\rho^2 + dz^2) + \rho^2\,d\phi^2], \quad (2.1)$$

where $\rho^2 = x^2 + y^2$ and λ and ν depend on ρ and z, and satisfy the equations

$$\Delta\lambda = 0 \quad (2.2)$$

(where Δ stands for the Laplacian), and

$$\nu = \int \rho[(\lambda_\rho^2 - \lambda_z^2)\,d\rho + 2\lambda_\rho\lambda_z\,dz]. \quad (2.3)$$

In the last expression subscripts denote derivatives and the integral is independent of the path of integration.

A particular solution of this nature corresponding to two objects is[1]

$$\lambda_0 = -\mu'/r' - \mu''/r'', \quad (2.4)$$

$$\nu_0 = -\tfrac{1}{2}\frac{\mu'^2\rho^2}{r'^4} - \tfrac{1}{2}\frac{\mu''^2\rho^2}{r''^4} + \frac{2\mu'\mu''}{(z-z'')^2} \times \left[\frac{\rho^2 + (z-z')(z-z'')}{r'r''} - 1\right]. \quad (2.5)$$

In these equations z' and z'' are given positions on the z axis related to the objects; r' and r'' are defined as

$$r' \equiv [\rho^2 + (z-z')^2]^{\frac{1}{2}}, \quad (2.6)$$

$$r'' \equiv [\rho^2 + (z-z'')^2]^{\frac{1}{2}}, \quad (2.7)$$

and μ' and μ'' are parameters with the dimension of length.

It should be noticed that, contrary to superficial appearance, the two objects are not point particles, nor are they spherically symmetric. In any case Eqs. (2.4) and (2.5) represent a free-space solution and apply only outside the objects proper. As pointed out by Synge,[1] the segment of the z axis between z' and z'' must also be excluded because "elementary flatness" is violated there (i.e., for a circle centered on a point in this segment, the ratio of circumference to radius does not tend to 2π as they both tend to zero). It is expected that the two objects cannot be separated by free space in a static situation. A strut is necessary to hold them apart.

We shall consider the above solution only for values of r' and r'' large enough so that the gravitational fields are weak and space is almost flat. We shall assume that $|z' - z''|$ is much larger still, so that the field of each object is weak at the other object and over most of the strut. Although the intrinsic objects described by Eqs. (2.4) and (2.5) are not spherically symmetric, asymptotically for large[2] values of r' and r'' they do generate a spherically symmetric field corresponding to active gravitational masses $m' = \mu'c^2/G$ and $m'' = \mu''c^2/G$ (G is Newton's constant of gravitation).

We now introduce the strut. It will be quite general in structure, subject only to two conditions: cylindrical symmetry and weightlessness. Weightlessness will be explained presently.

The most general cylindrically symmetric static metric[1] may be brought to the form

$$ds^2 = e^{2\lambda}c^2\,dt^2 - e^{-2\sigma}[e^{2\nu}(d\rho^2 + dz^2) - \rho^2\,d\phi^2], \quad (2.8)$$

with λ, σ, and ν functions of ρ and z. For the particular metric of Eqs. (2.1), (2.4), and (2.5), we had $\lambda = \sigma = \lambda_0$ and $\nu = \nu_0$. Because of the presence of a strut, deviations from these values are now permitted both over the volume of the strut and outside it. We therefore put

$$\lambda = \lambda_0 + f(\rho, z), \quad (2.9)$$

$$\sigma = \lambda_0 + g(\rho, z), \quad (2.10)$$

$$\nu = \nu_0 + h(\rho, z). \quad (2.11)$$

Outside the strut an empty-space solution must again be obtained, and this implies $f = g$ there. Asymptotically for distances much larger than $|z' - z''|$ (for which the whole of our configuration appears point-like) we want overall spherical symmetry to be reached. This implies that h must go zero at infinity. Finally, in order to restore "elementary flatness," we must require ν to vanish on the z axis. Since

$$\nu_0(0, z) = \begin{cases} -4\mu'\mu''/(z'-z'')^2 & \text{when } z \text{ lies between } z' \text{ and } z'', \\ 0 & \text{otherwise,} \end{cases} \quad (2.12)$$

we must have

$$h(0, z) = \begin{cases} 4\mu'\mu''/(z'-z'')^2 & \text{when } z \text{ lies between } z' \text{ and } z'', \\ 0 & \text{otherwise.} \end{cases} \quad (2.13)$$

The gravitational attraction is to be evaluated for $\mu', \mu'' \ll |z' - z''|$, so that Euclidean geometry holds

[1] J. L. Synge, *Relativity: the General Theory* (North-Holland Publishing Co., Amsterdam; and Interscience Publishers, New York, 1960), Chap. VIII, Sec. 1.

[2] These radii should be large in comparison with the Schwarzschild radii μ' and μ'', but not necessarily large in comparison with $|z' - z''|$.

over most of the strut. We shall evaluate the attraction only to lowest order[3] in μ' and μ'', which will turn out to be the order of $\mu'\mu''$. It will turn out that T_{zz} is of that order; therefore all other quantities will be taken in zero order and all geometrical considerations will be Euclidean. We shall therefore consider $|z' - z''|$ as the distance between z' and z''. The total force that the strut supports will be computed by integrating T_{zz} (the pressure) over a plane perpendicular to the z axis.

Before computing the force, however, some thought must be given to the weight of the strut itself, which may render the calculation ambiguous. To get around this we want the stress–energy distribution in the strut to be itself of order $\mu'\mu''$. This will presumably make the force of attraction between parts of the strut and other parts of the system of higher order and it will disappear from our calculation. (Our results justify this assumption.) Since, as will be shown below, the stress energy of the strut is of the order of the functions f, g, and h, we may achieve a "weightless" strut by restricting those functions to be small (of order $\mu'\mu''$ or smaller). This is the condition of weightlessness mentioned above. Notice that this condition is consistent with the condition (2.13) already imposed on h.

When the line element of Eq. (2.8) is processed through Einstein's equations, the zz component of the stress–energy tensor comes out as

$$T_{zz} = -\frac{c^4}{8\pi G}\Big[\sigma_{\rho\rho} + \frac{\sigma_\rho}{\rho} - \lambda_{\rho\rho} - \frac{\lambda_\rho}{\rho} + \frac{\nu_\rho}{\rho} - \nu_\rho(\sigma_\rho - \lambda_\rho) + (\nu_z - \sigma_z)(\sigma_z - \lambda_z) + \lambda_z^2 - \lambda_\rho^2\Big]. \tag{2.14}$$

When Eqs. (2.9)–(2.11) are substituted into this, the terms that are zero order in f, g, and h cancel out, since in the absence of these functions we have an empty-space solution. Concentrating on terms in which f, g, and h do appear, we find that all quadratic terms are of order higher than $\mu'\mu''$ (λ_0 and ν_0 are themselves of order μ' or μ''). Our result is therefore to be computed from the contributions of f, g, and h to the linear terms in Eq. (2.14). These terms (like f, g, and h) are themselves of order $\mu'\mu''$, so that, as mentioned above, all other quantities may be taken in zero order.

[3] In all considerations of order, μ' and μ'' are considered to be small quantities, meaning small in comparison with the only other length in the problem, namely, $|z' - z''|$. The expansion is thus in powers of the small pure numbers $\mu'/|z' - z''|$ and $\mu''/|z' - z''|$.

The force F supported by the strut thus becomes

$$F = \int_0^\infty 2\pi\rho\, d\rho T_{zz} \approx -\frac{c^4}{4G}\int_0^\infty d\rho\{\rho g_{\rho\rho} + g_\rho - \rho f_{\rho\rho} - f_\rho + h_\rho\} = -\frac{c^4}{4G}\int_0^\infty d\rho \frac{d}{d\rho} \times \{\rho g_\rho - \rho f_\rho + h\} = -\frac{c^4}{4G}[\rho g_\rho - \rho f_\rho + h]_0^\infty = \frac{c^4}{4G}h(0, z). \tag{2.15}$$

When $h(0, z)$ is substituted from Eq. (2.13) and use is made of $\mu' = Gm'/c^2$ and $\mu'' = Gm''/c^2$, we find

$$F = \begin{cases} \dfrac{Gm'm''}{(z' - z'')^2} & \text{when } z \text{ lies between } z' \text{ and } z'', \\ 0 & \text{otherwise.} \end{cases} \tag{2.16}$$

Thus the force supported by the strut between z' and z'' is equal to the Newtonian attraction between masses m' and m'' a distance $|z' - z''|$ apart. The fact that this force is independent of the shape and details of the strut and of the particular plane of integration (so long as it passes between z' and z'') shows that we were justified in expecting the strut (which is subject to our weightlessness condition) to be indeed weightless (to the order $\mu'\mu''$).

3. GENERALITY OF THE RESULTS

The free-space metric of Eqs. (2.1), (2.4), and (2.5) is a particular free-space solution. However, we have used it only in the asymptotic region where it becomes identical to any other solution corresponding to active masses m' and m'' at z' and z''. Therefore the use of this solution results in no loss of generality.

The strut is in turn constructed in the most general way consistent with cylindrical symmetry and weightlessness. It need not be visualized as a rigid strut. It merely reflects the concentration of stress and energy of any nongravitational dynamical entity (e.g., the electromagnetic field) capable of counterbalancing the gravitational attraction.[4] We did, however, limit the strut to being static.

This calculation, which appears to apply well to the static gravitational attraction between two elementary particles, establishes that general relativity does lead to Newton's law of gravitation and to the equality of active and passive mass in this case.

ACKNOWLEDGMENT

The author is pleased to acknowledge valuable comments from Dr. C. Rosenkilde.

[4] If the width of the strut is not finite, we must require that $\rho f_\rho - \rho g_\rho$ tend to zero as $\rho \to \infty$ for Eq. (2.15) to hold.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 7 JULY 1968

Analyticity Properties of the Scattering Amplitude for Singular Potentials*

NORMAN DOMBEY
School of Mathematical and Physical Sciences, University of Sussex, England
AND
R. H. JONES†
Faculty of Natural Sciences, University of Kent at Canterbury, England

(Received 27 January 1967)

A modified Sommerfeld–Watson transformation is established for the singular potential $V(r) = g^2r^{-4}$. The positions of the Regge poles and their residues are obtained for a general class of singular potentials and the analyticity properties of the scattering amplitude $f(k, \cos\theta)$ are discussed.

1. INTRODUCTION

In the last few years nonrelativistic potential theory has been studied in order to gain insight into the analytic properties of scattering amplitudes in field theory. For example, the Mandelstam representation can be proved for potentials which are of the Yukawa form.[1] More recently, it has been conjectured that the scattering amplitude may have less restrictive analytic properties.[2] Scattering off singular repulsive potentials ("singular" here means more repulsive than the centrifugal barrier at small distances) provides a mathematical model which does lead to a scattering amplitude with a more general analytic structure and in which these analytic properties can be explicitly demonstrated.

Martin[3] has shown that, for purely repulsive potentials which are less singular at the origin than some inverse power of r, the forward scattering amplitude satisfies a dispersion relation in the energy k^2. In this paper we investigate the analytic properties of the scattering amplitude as a function of $\cos\theta$ (θ is the scattering angle) for fixed real positive k for a class of singular potentials.

We follow Regge's approach[4] of considering the partial-wave scattering amplitude $S(\lambda, k)$, $\lambda = l + \frac{1}{2}$, as a function of angular momentum and then converting the scattering amplitude $f(k, \cos\theta)$ into a contour integral in the λ plane by means of a suitable Sommerfeld–Watson transform. This has been a very powerful technique in the case of regular potentials.

For a general singular potential several properties of $S(\lambda, k)$ in the λ plane are known. The reflection property[5]

$$e^{-i\pi\lambda}S(\lambda, k) = e^{i\pi\lambda}S(-\lambda, k) \tag{1.1}$$

follows from the behavior of the wavefunction at the origin. With unitarity

$$S(\lambda, k)S^*(\lambda^*, k^*) = 1, \tag{1.2}$$

one obtains[6]

$$|S(-i\Lambda, k)| = e^{\pi\Lambda} \tag{1.3}$$

for real Λ and k. So $S(\lambda, k)$ is exponentially small along the positive imaginary λ axis and exponentially large along the negative imaginary λ axis. Furthermore, it has been shown[7] that an infinite number of Regge poles lie in an arbitrarily small angle to the right of the positive imaginary λ axis for real k.

It follows from these results that the usual Sommerfeld–Watson transformation involving an integral along the imaginary axis in the λ plane cannot be carried out.

In Sec. 2 we study in detail $S(\lambda, k)$ for the potential g^2r^{-4}. This potential is convenient as the Schrödinger equation in this case may be reduced to the Mathieu equation.[8] In Sec. 3 we produce a modified contour in the λ plane to demonstrate analyticity in $\cos\theta$ for this potential. In Sec. 4 we introduce the phase-integral method,[9,10] connecting solutions of the Schrödinger equation between different regions of the r plane. This yields the positions and residues of the Regge poles of $S(\lambda, k)$ for more general singular potentials. We then discuss the application of these results to the analyticity properties of the scattering amplitude $f(k, \cos\theta)$.

* This research has been sponsored in part by the European Office of Aerospace Research (OAR), United States Air Force, under Grant No. AF EOAR 66-51.

† Work based on a thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy to the University of Sussex in December, 1966.

[1] R. Blankenbecler, M. L. Goldberger, N. N. Khuri, and S. B. Treiman, Ann. Phys. (N.Y.) **10**, 62 (1960).

[2] For example, A. Martin, Nuovo Cimento **42**, 930 (1966).

[3] A. Martin, *Preludes in Theoretical Physics* (North-Holland Publishing Co., Amsterdam, 1966).

[4] For example, T. Regge, *Theoretical Physics* (IAEA, Vienna, 1963).

[5] E. Predazzi and T. Regge, Nuovo Cimento **24**, 518 (1962).

[6] J. E. Bowcock and A. P. Contogouris, Nuovo Cimento **33**, 873 (1964).

[7] B. Jaksic and N. Limic, Commun. Math. Phys. **2**, 94 (1966).

[8] G. H. Wannier, Quart. Appl. Math. **11**, 33 (1953); E. Vogt and G. H. Wannier, Phys. Rev. **95**, 1190 (1954).

[9] W. H. Furry, Phys. Rev. **71**, 360 (1947).

[10] J. Heading, *Phase-Integral Methods* (Methuen and Company Ltd., London, 1962).

2. THE POTENTIAL g^2r^{-4}

In this section it is assumed that $|\arg k| \leq \pi/2$. The following approach has been previously used in papers by Challifour and Eden[11] and Yuan-Ben.[12]

In the Schrödinger equation

$$\frac{d^2\psi}{dr^2} + \left(k^2 - \frac{\lambda^2 - \frac{1}{4}}{r^2} - \frac{g^2}{r^4}\right)\psi = 0, \tag{2.1}$$

substitute

$$\psi(r) = r^{\frac{1}{2}}\phi(x), \quad x = i \ln r/r_0, \quad r_0^2 = ig/k \tag{2.2}$$

to obtain the Mathieu equation

$$\frac{d^2\phi}{dx^2} + (\lambda^2 - 2igk \cos 2x)\phi = 0. \tag{2.3}$$

This equation has been studied by Wannier,[8] who defines four solutions: $he^{(i)}x$, $i = 1, 2, 3, 4$. These solutions are defined by their asymptotic behavior at $x = \pm i\infty$, which corresponds to $r = 0$ and $r = \infty$. They may thus be identified with the regular, irregular, and Jost solutions of Eq. (2.1). In his paper Wannier obtains the linear dependence of these solutions (connection formulas) which yield the expression for $S(\lambda, k)$:

$$S(\lambda, k) = \frac{e^{i\pi\lambda}}{\cos \pi\beta + i(e^{-2\Phi} + \sin^2 \pi\beta)^{\frac{1}{2}}} \tag{2.4}$$

$$= \frac{e^{i\pi\lambda}}{1 + e^{-2\Phi}}[\cos \pi\beta - i(e^{-2\Phi} + \sin^2 \pi\beta)^{\frac{1}{2}}], \tag{2.5}$$

where

$$\cos \pi\beta = 1 - 2\Delta(0) \sin^2 \frac{\pi\lambda}{2}, \tag{2.6}$$

$$\Delta(0) = \begin{vmatrix} & \vdots & & & \\ \cdots & \frac{ikg}{2^2 - \lambda^2} & 1 & \frac{ikg}{2^2 - \lambda^2} & \\ & \frac{ikg}{-\lambda^2} & 1 & \frac{ikg}{-\lambda^2} & \\ & & \frac{ikg}{2^2 - \lambda^2} & 1 & \frac{ikg}{2^2 - \lambda^2} \quad \cdots \\ & & & \vdots & \end{vmatrix}, \tag{2.7}$$

and Φ is a function of λ and k which is found by Wannier in a WKB approximation

$$\Phi = \Phi_0 + O[(\lambda^2 + 2ikg)^{-\frac{1}{2}}], \tag{2.8}$$

where

$$\Phi_0 = \frac{\pi(\lambda^2 - 2ikg)}{2(\lambda^2 + 2ikg)^{\frac{1}{2}}} F\left(\tfrac{3}{2}, \tfrac{1}{2}; 2; \frac{\lambda^2 - 2ikg}{\lambda^2 + 2ikg}\right). \tag{2.9}$$

Hence the approximation is a good one if $|\lambda^2 + 2ikg|$ is large.

We are particularly interested in the positions and residues of the Regge poles of $S(\lambda, k)$, and also its asymptotic behavior as $|\lambda| \to \infty$ throughout the right half-plane; i.e., for $|\arg \lambda| \leq \pi/2$.

From Eq. (2.5), $S(\lambda, k)$ has a pole if

$$1 + e^{-2\Phi} = 0,$$

i.e.,

$$\Phi = (m + \tfrac{1}{2})\pi i, \quad m = 0, \pm 1, \pm 2, \cdots. \tag{2.10}$$

The Regge poles thus lie along a line $L(\lambda)$ defined by

$$\operatorname{Re} \Phi(\lambda, k) = 0. \tag{2.11}$$

Equation (2.10) may be solved for small and large values of m. For small positive m (with the extra condition $|gk|^{\frac{1}{2}} \gg m + \frac{1}{2}$) we find the poles in the first quadrant[13]:

$$\lambda_m(k) = (2igk)^{\frac{1}{2}} + \sqrt{2}\,(m + \tfrac{1}{2})i\left\{1 + O\left(\frac{m + \frac{1}{2}}{(gk)^{\frac{1}{2}}}\right)\right\}. \tag{2.12}$$

Now $S(\lambda, k)$ must satisfy the unitarity condition, Eq. (1.2). Hence, for real k, we expect to find zeros of $S(\lambda, k)$ in the fourth quadrant at the points conjugate to Eq. (2.12):

$$\lambda_m^*(k) = (-2igk)^{\frac{1}{2}} - \sqrt{2}\,(m + \tfrac{1}{2})i\left\{1 + O\left(\frac{m + \frac{1}{2}}{(gk)^{\frac{1}{2}}}\right)\right\}. \tag{2.13}$$

Equation (2.4) will only have a zero if $e^{\Phi} = 0$, i.e., $\operatorname{Re} \Phi = -\infty$. Wannier,[8] however, has found his connection formulas under the assumption that $e^{\Phi} \neq 0$. By comparison of Eq. (2.13) and Eq. (2.8), it can be seen that the approximation to Φ is invalid at λ_m^*. If $|\arg k| \leq \pi/2$, the zeros of $S(\lambda, k)$ will always lie in the second and fourth quadrants of the λ plane. We shall therefore use Eq. (2.4) and Eq. (2.5) for λ in the first quadrant only.

It is possible to find another expression for $S(\lambda, k)$, valid for λ in the fourth quadrant. In Eq. (2.1) we substitute

$$\psi(r) = r^{\frac{1}{2}}\phi(x), \quad x = i \ln \frac{r}{r_0}, \quad r_0^2 = \frac{g}{ik} \tag{2.14}$$

[11] J. Challifour and R. J. Eden, J. Math. Phys. **4**, 359 (1963).
[12] D. Yuan-Ben, Sci. Sinica (Peking) **13**, 1319 (1964).
[13] Compare Ref. 11. Further inspection, however, shows that negative m corresponds to Regge poles in the third quadrant and not in the first.

to obtain

$$\frac{d^2\phi}{dx^2} + (\lambda^2 + 2igk\cos 2x)\phi = 0. \quad (2.15)$$

As before, we find an expression for $S(\lambda, k)$, which now reads

$$S(\lambda, k) = e^{i\pi\lambda}[\cos\pi\tilde{\beta} - i(e^{-2\tilde{\Phi}} + \sin^2\pi\tilde{\beta})^{\frac{1}{2}}] \quad (2.16)$$

$$= \frac{e^{i\pi\lambda}(1 + e^{-2\tilde{\Phi}})}{\cos\pi\tilde{\beta} + i(e^{-2\tilde{\Phi}} + \sin^2\pi\tilde{\beta})^{\frac{1}{2}}}, \quad (2.17)$$

where the introduction of the tilde implies that ikg has been replaced throughout by $-ikg$; the opposite sign of the square root must also be taken. Zeros of $S(\lambda, k)$ are given by

$$1 + e^{-2\tilde{\Phi}} = 0.$$

Now it is easy to prove the unitarity condition Eq. (1.2) directly from Eqs. (2.4) and (2.16).

We shall require the asymptotic behavior of $S(\lambda, k)$ in the λ plane. We know that, at least for real k, the poles lie in the vicinity of $\arg\lambda = \pi/2$. Let us therefore consider Eqs. (2.4) to (2.10) for $\operatorname{Im}\lambda \to \infty$, $\operatorname{Re}\lambda \geq 0$, and fixed k.

In Eq. (2.7) the off-diagonal terms will all be small and

$$\Delta(0) = 1 - \frac{\pi i k^2 g^2}{4\lambda^3}\left\{1 + O\left(\frac{k^2g^2}{\lambda^4}\right)\right\}. \quad (2.18)$$

Substitute this in Eq. (2.6) and note that $\cos\pi\lambda \simeq \frac{1}{2}e^{-i\pi\lambda}$:

$$\beta - \lambda = \frac{k^2g^2}{4\lambda^3}\left\{1 + O\left(\frac{k^2g^2}{\lambda^4}\right)\right\}. \quad (2.19)$$

Using the value of the hypergeometric function near unity in Eq. (2.8), we get

$$\Phi(\lambda, k) = 2\lambda\ln\left\{\frac{2\lambda}{e(igk)^{\frac{1}{2}}}\right\} + O(\lambda^{-1}). \quad (2.20)$$

It may thus be seen that $|\sin^2\pi\beta| \gg |e^{-2\Phi}|$, and hence, from Eq. (2.5),

$$S(\lambda, k) = \frac{e^{i\pi(\lambda-\beta)}}{1 + e^{-2\Phi}}. \quad (2.21)$$

When k is real, Tiktopoulos[14] has shown that, for large λ,

$$S(\lambda, k) \to \exp(2i\delta_{\mathrm{WKB}}); \quad |\arg\lambda| < \pi/4, \quad (2.22)$$

where

$$\delta_{WKB} = -\frac{\pi k^2 g^2}{8\lambda^3}\left\{1 + O\left(\frac{k^2g^2}{\lambda^3}\right)\right\}. \quad (2.23)$$

Now Eq. (2.21) has been obtained for large $\operatorname{Im}\lambda$, and so, from Eqs. (2.19), (2.21), and (2.23), we may write

$$S(\lambda, k) \to \frac{\exp(2i\delta_{\mathrm{WKB}})}{1 + e^{-2\Phi}}; \quad \frac{\pi}{4} \leq \arg\lambda \leq \frac{\pi}{2}. \quad (2.24)$$

For the remaining sector of the right half-plane we may follow an analogous procedure using Eq. (2.16) and obtain

$$S(\lambda, k) \to \exp(2i\delta_{\mathrm{WKB}})(1 + e^{-2\tilde{\Phi}}); \quad -\frac{\pi}{2} \leq \arg\lambda \leq -\frac{\pi}{4}. \quad (2.25)$$

The above equations now give the asymptotic behavior of $S(\lambda, k)$ throughout the right half-plane.

Writing $\lambda = \lambda_1 + i\lambda_2$ in Eq. (2.20), we have

$$\operatorname{Re}\Phi = 2\lambda_1\ln\left\{\frac{2|\lambda|}{e|gk|^{\frac{1}{2}}}\right\} - 2\lambda_2\left(\arg\lambda - \frac{\pi}{4} - \tfrac{1}{2}\arg k\right), \quad (2.26)$$

$$\operatorname{Im}\Phi = 2\lambda_1\left(\arg\lambda - \frac{\pi}{4} - \tfrac{1}{2}\arg k\right) + 2\lambda_2\ln\left\{\frac{2|\lambda|}{e|gk|^{\frac{1}{2}}}\right\}. \quad (2.27)$$

These equations were first obtained by Yuan-Ben[12]; when substituted in Eq. (2.10), they yield the positions of the Regge poles for large values of m. From Eq. (2.26), $\operatorname{Re}\Phi$ is large and positive unless $\lambda_1 \ll \lambda_2$; i.e., $e^{-2\Phi}$ is exponentially small unless $\arg\lambda \simeq \pi/2$. Thus it may be seen from the preceding equations that

$$S(\lambda, k) \to \exp(2i\delta_{\mathrm{WKB}}); \quad |\arg\lambda| \leq \frac{\pi}{2} - \epsilon, \quad \epsilon > 0, \quad (2.28)$$

i.e., the region in which Eq. (2.22) holds may be extended.

If $\arg k = \pi/2$, the line of poles $L(\lambda)$ defined by $\operatorname{Re}\Phi = 0$ lies along the imaginary axis. If $\pi/2 > \arg k \geq -\pi/2$, then $\lambda \to \infty$ along $L(\lambda)$ implies

$$\lambda_1, \lambda_2, \lambda_2/\lambda_1 \to \infty, \quad \arg\lambda \to \pi/2. \quad (2.29)$$

Hence there are an infinite number of Regge poles lying in an arbitrarily small angle to the right of the positive imaginary axis such that their real parts are unbounded. To the left of $L(\lambda)$, i.e., between $L(\lambda)$ and the positive imaginary λ axis, $\operatorname{Re}\Phi$ becomes increasingly negative as we approach the axis and hence $S(\lambda, k) \to e^{2\Phi}$. For real k, $S(\lambda, k)$ satisfies Eq. (1.3) along the axis.

[14] G. Tiktopoulos, Phys. Rev. **138**, B1550 (1965).

In the fourth quadrant of the λ plane there are an infinite number of zeros of $S(\lambda, k)$ lying along a line $\tilde{L}(\lambda)$ defined by Re $\tilde{\Phi} = 0$. When k is real, $\tilde{L}(\lambda)$ is conjugate to $L(\lambda)$ and $S(\lambda, k)$ satisfies Eq. (1.3) along the negative imaginary λ axis.

The value of the residue γ_m of $S(\lambda, k)$ at each pole may be calculated from Eq. (2.4);

$$\gamma_m = \lim_{\lambda \to \lambda_m} (\lambda - \lambda_m) S(\lambda, k) = \frac{e^{i\pi\lambda_m} \cos \pi\beta(\lambda_m)}{\Phi'(\lambda_m)}.$$

When m is large,

$$\gamma_m = \left[4 \ln \left\{ \frac{2\lambda_m}{(igk)^{\frac{1}{2}}} \right\} \right]^{-1}. \tag{2.30}$$

3. TOTAL SCATTERING AMPLITUDE

We now consider the scattering amplitude $f(k, \cos\theta)$ for the potential $V(r) = g^2 r^{-4}$ for real k. As is well known, the partial-wave sum for the scattering amplitude may be expressed as a contour integral in the complex λ plane:

$$f(k, \cos\theta) = \frac{1}{2ik} \sum_{l=0}^{\infty} (2l + 1) \times P_l(\cos\theta)[S_l(k) - 1] \tag{3.1}$$

$$= -\frac{1}{2k} \int_C \frac{\lambda \, d\lambda}{\cos \pi\lambda} P_{\lambda - \frac{1}{2}}(-\cos\theta)[S(\lambda, k) - 1], \tag{3.2}$$

where C is a contour enclosing the positive real axis (Fig. 1). Equation (3.1) is convergent for physical $\cos\theta$.

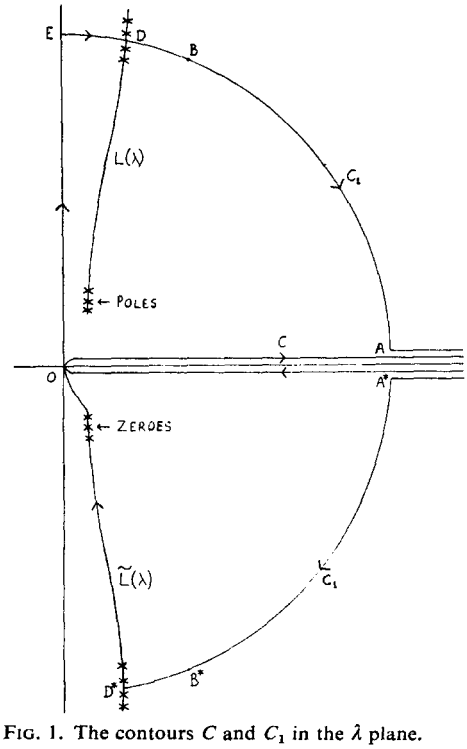


FIG. 1. The contours C and C_1 in the λ plane.

In order to consider $f(k, \cos\theta)$ as an analytic function of $\cos\theta$, we take Eq. (3.2) and open out the contour C as much as possible. For regular potentials it is possible to open out C until it runs along the imaginary axis. In the singular case this is not permitted. Let us consider the situation shown in Fig. 1; θ is, for the moment, restricted to $0 < \theta < 2\pi$. We have

$$f(k, \cos\theta) = -\frac{1}{2k} \int_{C_1} \frac{\lambda \, d\lambda}{\cos \pi\lambda} \times P_{\lambda - \frac{1}{2}}(-\cos\theta)[S(\lambda, k) - 1] - \frac{\pi i}{k} \sum_m \frac{\lambda_m}{\cos \pi\lambda_m} P_{\lambda_m - \frac{1}{2}}(-\cos\theta)\gamma_m, \tag{3.3}$$

where the sum extends over the poles enclosed by C_1. In the first quadrant C_1 is taken along the positive imaginary axis and in the fourth quadrant along $\tilde{L}(\lambda)$. AE and A^*D^* are arcs of large radius with A^*D^* conjugate to AD. We wish to prove that, as AE and A^*D^* are moved out to infinity, the parts of the integral in Eq. (3.3) taken along these arcs tend to zero.

In Sec. 2 we have seen that $S(\lambda, k) \to \exp(2i\delta_{\mathrm{WKB}})$ as $\lambda \to \infty$ in the range $|\arg\lambda| \leq \pi/2 - \epsilon$, where $\epsilon > 0$ is an arbitrarily small fixed quantity. We may therefore choose a point B on C_1 such that $\arg B = \pi/2 - \epsilon$. If AE is at a sufficiently large distance from the origin, B will lie to the right of $L(\lambda)$. As AB moves out to infinity, thus, along AB and A^*B^*, we have

$$|S(\lambda, k) - 1| \to \left| \frac{\pi k^2 g^2}{4\lambda^3} \right|.$$

Also,

$$\frac{P_{\lambda - \frac{1}{2}}(-\cos\theta)}{\cos \pi\lambda} \leq O(|\lambda^{-\frac{1}{2}}|).$$

Hence,

$$\left\{ \int_{A^*}^{B^*} + \int_B^A \right\} \frac{\lambda \, d\lambda}{\cos \pi\lambda} P_{\lambda - \frac{1}{2}}(-\cos\theta)[S(\lambda, k) - 1] \to 0, \tag{3.4}$$

as AB and A^*B^* move out to infinity.

Define EB by the condition Im $\Phi = M\pi$, where M is a large positive integer. Then $e^{-2\Phi} = e^{-2 \mathrm{Re}\, \Phi}$ is real and nonnegative along EB. Hence,

$$|S(\lambda, k)| = \frac{|\exp(2i\delta_{\mathrm{WKB}})|}{|1 + e^{-2\Phi}|} \leq |\exp(2i\delta_{\mathrm{WKB}})|.$$

As δ_{WKB} is small for large λ, $S(\lambda, k)$ is bounded along EB. Along DB we know that Re $\Phi \geq 0$; then

$$\frac{1}{2} \leq \left| \frac{S(\lambda, k)}{\exp(2i\delta_{\mathrm{WKB}})} \right| \leq 1.$$

By the unitarity condition it may be seen that $S(\lambda, k)$ is bounded along B^*D^*. Now $P_{\lambda-\frac{1}{2}}(-\cos\theta)/\cos\pi\lambda$ is exponentially small when $|\mathrm{Im}\,\lambda|$ is large. Hence

$$\left\{\int_{B^*}^{D^*} + \int_{E}^{B}\right\} \frac{\lambda\, d\lambda}{\cos\pi\lambda} P_{\lambda-\frac{1}{2}}(-\cos\theta)[S(\lambda, k) - 1] \to 0, \tag{3.5}$$

as EB and B^*D^* move out to infinity.

From Eqs. (3.3), (3.4), and (3.5) we obtain

$$f(k, \cos\theta) = -\frac{1}{2k}\int_{C_2} \frac{\lambda\, d\lambda}{\cos\pi\lambda} P_{\lambda-\frac{1}{2}}(-\cos\theta)[S(\lambda, k) - 1] - \frac{\pi i}{k}\sum_{m=0}^{\infty} \frac{\lambda_m}{\cos\pi\lambda_m} P_{\lambda_m-\frac{1}{2}}(-\cos\theta)\gamma_m, \tag{3.6}$$

where C_2 lies along $\tilde{L}(\lambda)$ in the fourth quadrant and along the positive imaginary axis.

So far we have restricted $\cos\theta$ by $0 < \theta < 2\pi$. Equation (3.6) reduces to Eq. (3.1) for these values of θ. However, Eq. (3.6) is convergent for more general complex values of $\cos\theta$, and thus it is an analytic continuation of $f(k, \cos\theta)$ into this larger domain.

For large λ we have

$$\frac{P_{\lambda-\frac{1}{2}}(-\cos\theta)}{\cos\pi\lambda} = O\left(\left|\frac{\lambda^{-\frac{1}{2}}\exp[\pm i\lambda(\pi-\theta)]}{\exp[\pm i\pi\lambda]}\right|\right),$$

where the signs are chosen to give the larger terms in both the numerator and the denominator. In the infinite sum over poles, $\lambda_m = \lambda_{m1} + i\lambda_{m2}$ where $\lambda_{m2}/\lambda_{m1} \to \infty$ as $m \to \infty$. Hence

$$\frac{P_{\lambda_m-\frac{1}{2}}(-\cos\theta)}{\cos\pi\lambda_m} = O\left(\frac{|\lambda_m^{-\frac{1}{2}}|\exp\{\pm(\lambda_{m1}\theta_2 - \lambda_{m2}[\pi - \theta_1])\}}{\exp[\pi\lambda_{m2}]}\right).$$

Since $\lambda_{m2}/\lambda_{m1} \to \infty$, it follows that this expression is exponentially decreasing at infinity for any finite value of θ_2 if $0 < \theta_1 < 2\pi$. Also γ_m decreases as m increases from Eq. (2.30). Hence the infinite sum over Regge poles is convergent if $0 < \mathrm{Re}\,\theta < 2\pi$ and $0 \leq \mathrm{Im}\,\theta < \infty$. From Eqs. (2.24) and (2.25), $|S - 1| \to 1$ as $\lambda \to \infty$ along either branch of C_2, and hence an analogous argument for the integral over C_2 in Eq. (3.6) shows that it converges for the same values of θ. This strip of the θ plane maps into the whole $\cos\theta$ plane with a cut along the positive axis between $+1$ and $+\infty$. Thus Eq. (3.6) is an analytic representation of $f(k, \cos\theta)$ in this cut plane.

4. PHASE-INTEGRAL METHODS

In Secs. 2 and 3 we have considered the potential g^2r^{-4}. In this section we shall find the positions and residues of the Regge poles of $S(\lambda, k)$ for a more general class of singular potentials. Our results will enable us to conjecture that the results of Sec. 3 hold also in the more general case.

The Schrödinger equation may be written

$$\psi''(r) + q(r)\psi(r) = 0. \tag{4.1}$$

If $q(r)$ contains a large parameter, this equation has the WKB asymptotic solutions

$$\psi_1(r) = q^{-\frac{1}{4}}\exp\left\{i\int^r [q(r')]^{\frac{1}{2}}\, dr'\right\},$$
$$\psi_2(r) = q^{-\frac{1}{4}}\exp\left\{-i\int^r [q(r')]^{\frac{1}{2}}\, dr'\right\}. \tag{4.2}$$

In our case the large parameter will be λ. These solutions are valid everywhere in the complex r plane except near the zeros of $q(r)$, provided that the potential has an analytic continuation in the region of the r plane under consideration.

We take $V(r)$ in the form

$$V(r) = \frac{g^2}{r^n} f(r), \quad n \text{ integer} \geq 4, \tag{4.3}$$

where $f(r)$ is a bounded analytic function in the right half-plane and $f(0) = 1$. In particular, we can take [for $f(r) \not\equiv 1$]

$$f(r) = \int_{\mu}^{\infty} \sigma(\alpha)e^{-\alpha r}\, d\alpha, \quad \mu \geq 0, \tag{4.4}$$

where $\int_{\mu}^{\infty} |\sigma(\alpha)|\, d\alpha < \infty$. It will be convenient to assume, in addition, that the region of analyticity includes the origin. Now

$$q(r) = k^2 - \frac{\lambda^2}{r^2} - \frac{g^2}{r^n} f(r), \tag{4.5}$$

where $\lambda^2 - \frac{1}{4}$ has been replaced by λ^2, as is normal in a WKB approximation; in any case, we are concerned with large λ.

Under the conditions on the potential it may be seen that, for large enough λ, $q(r)$ has $n - 2$ zeros near the origin, which are situated near to the roots of the equation $r^{n-2} = -g^2/\lambda^2$. Write $\lambda = \Lambda e^{i\omega}$ and these roots are

$$\alpha_p = \left(\frac{g}{\Lambda}\right)^{2/n-2} \exp\left\{\frac{(2p+1)\pi - 2\omega}{n-2}\, i\right\}, \quad p = 0, 1, \cdots, n-3, \tag{4.6}$$

where the first factor is taken as real and positive. Denote the zero of $q(r)$ near α_0 by α. Apart from these $n - 2$ zeros, $q(r)$ will have another zero near λ/k;

we denote this zero by β, and in the following it will always lie in the right half-plane. If we were considering a regular potential, β would be the only zero of $q(r)$ in the right half-plane for large λ.[15] It is the presence of extra zeros near the origin that enables us to use phase-integral methods to find the Regge poles.

The regular solution for a potential given by Eq. (4.3) has the behavior near the origin of the form

$$\phi(r) = u(r) \exp\left\{ - \frac{2g}{n-2} r^{-(n-2)/2} + h(r) \right\}, \quad (4.7)$$

where $u(r)$ is a power series in r, and $h(r)$ is less singular than the first term in the exponential. Hence $\phi(r) \to 0$ as $r \to 0$ only if $|\arg r| < \pi/n - 2$. In the consideration of the WKB solutions around the origin we shall thus restrict ourselves to this wedge.

The WKB solution $\psi_2(r)$ represents the regular solution in the region between the origin and α. If there is an anti-Stokes line (see Appendix A for details) joining α and β, we may follow the rules of the phase-integral method[10] and continue this solution around the zeros at α and β to the region to the right of β. It then becomes a linear combination of ψ_1 and ψ_2, which are then multiples of the Jost solutions. This linear combination yields

$$S(\lambda, k) = \frac{\exp(2i\delta_{\mathrm{WKB}})}{1 + e^{-2\Phi}}, \quad (4.8)$$

where

$$\Phi(\lambda, k) = -i \int_\alpha^\beta [q(r)]^{\frac{1}{2}} \, dr. \quad (4.9)$$

We shall see later that this Φ is a generalization of the one occurring in Sec. 2. The requirement that an anti-Stokes line joins α and β implies that $\mathrm{Re}\ \Phi = 0$.

Now Eq. (4.8) has a pole when $1 + e^{-2\Phi} = 0$, i.e., if

$$\Phi = (m + \tfrac{1}{2})\pi i, \quad m \text{ positive integer}. \quad (4.10)$$

As α and β are both square-root branch points of the integrand in Eq. (4.9), Eq. (4.10) may be written

$$\oint [q(r)]^{\frac{1}{2}} \, dr = -2\pi(m + \tfrac{1}{2}), \quad (4.11)$$

where the integral is taken around a suitable contour enclosing α and β. The negative sign arises from our choice of square root in the integrand. This is the Bohr–Sommerfeld quantization condition for bound states in a semiclassical approximation; i.e., we are interpreting Regge poles as bound states for complex angular momentum.[14]

By using phase-integral methods, however, we obtain more than the quantization condition: we also obtain an expression for the residue γ_m at each Regge pole. From Eq. (4.8)

$$\gamma_m = \frac{\exp[2i\delta_{\mathrm{WKB}}(\lambda_m, k)]}{2\Phi'(\lambda_m, k)}, \quad (4.12)$$

where the prime denotes differentiation with respect to λ.

Let us determine Φ as a function of λ:

$$\Phi(\lambda, k) = -i \int_\alpha^\beta \left(k^2 - \frac{\lambda^2}{r^2} - \frac{g^2}{r^n} f(r) \right)^{\frac{1}{2}} dr.$$

Calculation of this integral when $f(r) = 1$ is fairly straightforward. As $|\lambda|$ is considered large, $q(r) \simeq -\lambda^2/r^2$ over most of the range of integration, i.e., except near α and β. Near α,

$$q(r) \simeq -\frac{\lambda^2}{r^2}\left(1 - \frac{\alpha^{n-2}}{r^{n-2}}\right);$$

and near β,

$$q(r) \simeq -\frac{\lambda^2}{r^2}\left(1 - \frac{r^2}{\beta^2}\right).$$

Using these approximations in the appropriate intervals, Φ may be calculated to give

$$\Phi(\lambda, k) = \frac{n\lambda}{n-2} \ln\left\{ \frac{2\lambda}{e(ig)^{2/n} k^{1-2/n}} \right\} + O(\lambda^{-\frac{1}{2}}). \quad (4.13)$$

If $f(r) = 1 - ar^p$ for small r and some positive integer p, then Φ will have extra terms which can be large but are small in comparison with the first term of Eq. (4.13). An explicit calculation is made in Appendix B for $n = 4$ and $p = 1$ [e.g., $V(r) = (g^2/r^4)e^{-ar}$], where it is found that the extra term is a constant.

The position of the mth Regge pole, at $\lambda_m = \lambda_{m1} + i\lambda_{m2}$, may now be found from Eqs. (4.10) and (4.13):

$$\lambda_{m1} \ln\left\{ \frac{2|\lambda_m|}{e|g^{2/n} k^{1-2/n}|} \right\} - \lambda_{m2}\left\{ \arg \lambda_m - \frac{\pi}{n} - \left(1 - \frac{2}{n}\right) \arg k \right\} = 0, \quad (4.14)$$

$$\lambda_{m1}\left\{ \arg \lambda_m - \frac{\pi}{n} - \left(1 - \frac{2}{n}\right) \arg k \right\} + \lambda_{m2} \ln\left\{ \frac{2|\lambda_m|}{e|g^{2/n} k^{1-2/n}|} \right\} = \frac{n-2}{n}(m + \tfrac{1}{2})\pi. \quad (4.15)$$

Thus we have found the position of the Regge poles for large values of m. For large k but small m, $S(\lambda, k)$ is still given by Eq. (4.8) along the line of poles, but the expression for Φ is different.

So we see that $S(\lambda, k)$ has an infinite number of poles lying on a line $L(\lambda)$ defined by $\mathrm{Re}\ \Phi = 0$. $S(\lambda, k)$ is

[15] A. Bottino, A. M. Longoni, and T. Regge, Nuovo Cimento **23**, 954 (1962).

given by Eq. (4.8) along this line of poles for large enough $|\lambda|$.

If $\arg k = \pi/2$, Eqs. (4.14) and (4.15) are satisfied by $\lambda_{m1} = 0$, i.e., the poles lie on the imaginary axis.[16]

If $0 \leq \arg k < \pi/2$, then

$$m \to \infty \quad \text{implies} \quad \lambda_{m1},\ \lambda_{m2},\ \lambda_{m2}/\lambda_{m1} \to \infty, \quad \arg \lambda_m \to \pi/2. \quad (4.16)$$

Thus an infinite number of Regge poles lie eventually within a small angle to the right of the imaginary axis, although λ_{m1} itself tends to infinity. This is in agreement with the result found by Jaksic and Limic.[7]

An analogous procedure may be followed to find the zeros of $S(\lambda, k)$ in the fourth quadrant. If $-\pi/2 \leq \arg k \leq 0$, then along the line of zeros $\tilde{L}(\lambda)$ we have

$$S(\lambda, k) = \exp(2i\delta_{\mathrm{WKB}})(1 + e^{-2\tilde{\Phi}}), \quad (4.17)$$

where

$$\tilde{\Phi}(\lambda, k) = -i\int_{\tilde{\alpha}}^{\beta} [q(r)]^{\frac{1}{2}}\, dr \quad (4.18)$$

and $\tilde{\alpha}$ is the zero of $q(r)$ near the origin corresponding to $p = n - 3$ in Eq. (4.6).

The residues of $S(\lambda, k)$ at the Regge poles are (δ_{WKB} is small for large $|\lambda_m|$)

$$\gamma_m = (n-2)/2n \ln\left\{\frac{2\lambda_m}{(ig)^{2/n}k^{1-2/n}}\right\}. \quad (4.19)$$

It may be seen for $n = 4$ that the above equations reduce to those found in Sec. 2.

We have proved Eq. (4.8) when $0 \leq \arg k \leq \pi/2$ and Eq. (4.17) when $-\pi/2 \leq \arg k \leq 0$. Hence, when k is real, we know the position and residues of the Regge poles and the position of the zeros of $S(\lambda, k)$.

We now consider the scattering amplitude $f(k, \cos\theta)$ for real θ in terms of the contour integral discussed in Sec. 3. The integral along the arcs A^*B^* and BA of Fig. 1 still tends to zero as its radius increases because $S(\lambda, k) \to 1$ as $|\lambda| \to \infty$, provided that $|\arg \lambda| \leq (\pi/2) - \epsilon$.[7] The integrals along the positive imaginary axis and along the line of zeros and the sum over the poles in the upper half-plane converge just as in the previous case. As the angle ϵ is arbitrarily small, it is sufficient to show that $S(\lambda, k)$ is bounded (a polynomial bound would suffice) along some arc EB crossing the line of poles $L(\lambda)$ and some arc B^*D^* up to the line of zeros $\tilde{L}(\lambda)$.

Unfortunately, our method is too crude to guarantee this, as we only have the expression Eq. (4.8) for λ on $L(\lambda)$ and the corresponding expression Eq. (4.17) for λ on $\tilde{L}(\lambda)$. There is always a point between successive poles on $L(\lambda)$ satisfying $\mathrm{Im}\,\Phi = M\pi$, where M is an integer for which $S(\lambda, k) \to \frac{1}{2}$. But we are not allowed to say anything about the behavior of $S(\lambda, k)$ near the line $L(\lambda)$. Similar arguments apply to the lower half-plane.

Physically, we would not expect the scattering amplitude for the potential g^2r^{-4} to be substantially different from that of the potential g^2r^{-n} for other values of $n > 3$. We have shown in this section that the zeros and poles of all such inverse power potentials are given by Eqs. (4.8) and (4.17), where Φ and $\tilde{\Phi}$ have a similar form for any value of $n > 3$, and we know that these expressions can be continued away from $L(\lambda)$ and $\tilde{L}(\lambda)$ for $n = 4$. So it is very reasonable to assume that at least inverse-power potentials have no incongruous behavior along EB and B^*D^*. Similarly, we can consider the potential $V(r) = (g^2/r^4)e^{-\mu r}$. We have shown that the function Φ for this potential for large imaginary λ is asymptotically equal to the Φ of Sec. 2 (Appendix B). So here again it would be surprising if $S(\lambda, k)$ along EB and B^*D^* were not bounded for large enough $|\lambda|$.

Hence we expect, but have not proved, that there exists a class of singular potentials for which $S(\lambda, k)$ is bounded along EB and B^*D^* and, therefore, for which the modified Sommerfeld–Watson transformation of Sec. 3 can be carried out. This allows the analytic continuation of $f(k, \cos\theta)$ for fixed real positive k in $\cos\theta$.

5. CONCLUSIONS

We have shown that, for at least one singular potential, the scattering amplitude $f(k, \cos\theta)$ is analytic in the $\cos\theta$ plane, cut from $\cos\theta = 1$ to infinity along the real axis for real positive k. Following the discussion at the end of the last section, it seems likely that there is a more general class of singular potentials with inverse-power behavior at the origin for which this is true. Indeed, if we consider potentials of this class which are exponentially bounded by $e^{-\mu r}$ as $r \to \infty$, we also have analyticity in $\cos\theta$ within the Lehmann ellipse,[17] and so the region of analyticity can be extended to the cut $\cos\theta$ plane with the cut running from $\cos\theta = 1 + (\mu^2/2k^2)$ to infinity. This is the same region of analyticity as for exponentially bounded regular potentials. In the singular case, however, we are unable to prove that there is a polynomial bound of $f(k, \cos\theta)$ as $\cos\theta \to \infty$. This is indeed unlikely, for $\mathrm{Re}\,\lambda_m \to \infty$ as $m \to \infty$, where λ_m is the position of the mth Regge pole.

[16] It is easy to show that, as in the case of regular potentials, for $k^2 < 0$, λ^2 is real.

[17] This follows immediately from the asymptotic behavior of the phase shift $\delta(\lambda, k)$ for real λ and k as $\lambda \to \infty$ [given by B. Jaksic and N. Limic, J. Math. Phys. 7, 88 (1966)].

We do not know anything about the analytic properties of $f(k, \cos\theta)$ in the energy k^2 except for the case of forward scattering.[3] With regular potentials the analyticity properties in k can be investigated away from the forward direction by using methods based on the integral equation for scattering which is iterated to give the Born series.[18] The variable $t = -2k^2(1 - \cos\theta)$ appears naturally in this approach. Then it can be shown that, for fixed t, $f(k, t)$ is analytic in the upper-half k plane.

For singular potentials the Born series plays no role and so t is not a natural variable. We suspect that a new variable $\tau = \tau(k, 1 - \cos\theta)$ would have to be used in its place in order to obtain interesting analyticity properties of $f(k, \tau)$ in k for fixed τ and that τ would depend on the degree of singularity n of the potential.[19]

We thus see that the scattering amplitude $f(s = k^2, \cos\theta)$ has a more complex analytic structure in singular-potential theory than in the regular case. In this respect the analytic structure so far obtained seems to be similar to that which can now be proved in axiomatic field theory.[2] But, as t probably plays no role in singular potential theory,[20] it is difficult to consider them as mathematical models of relativistic field theory where crossing is an essential requisite.

Our method does not help to extend the region of analyticity of $f(k, \cos\theta)$ in k. Only real k can be considered with the modified Sommerfeld–Watson transformation we use. But, as dispersion relations in k^2 exist, at least for the forward scattering amplitude, it should be possible to exhibit the analytic continuation to complex k directly without recourse to the regularization procedures used by Martin.[3]

There is also the mathematical problem of Sec. 4 to overcome. Wannier[8] was able to obtain a WKB expression for $S(\lambda, k)$ valid for λ in the first quadrant for the g^2r^{-4} potential by using the periodicity properties of the Mathieu equation together with its special symmetry (i.e., under the transformation $r \to 1/r$). The periodicity properties can be generalized to all singular potentials which are single-valued functions of r,[21] but the symmetry breaks down. If the phase-integral approach is combined with the method of Fubini and Stroffolini,[21] it should be possible to obtain less restrictive asymptotic expansions for $S(\lambda, k)$ in the general case.

APPENDIX A

A general solution of the Schrödinger equation, Eq. (4.1), may be written as a linear combination of the WKB solutions:

$$\psi(r) = A_1\psi_1(r) + A_2\psi_2(r). \tag{A1}$$

But in asymptotic expansions A_1 and A_2 suffer discontinuous changes on going to different regions of the r plane. This is the Stokes phenomenon,[22] and phase-integral techniques enable us to trace solutions of Eq. (4.1) around the singular points of the WKB solutions, i.e., around the zeros of $q(r)$, where

$$q(r) = k^2 - (\lambda^2/r^2) - (g^2/r^n)f(r). \tag{A2}$$

It is known that, for large λ and real k, the Regge poles lie within a small angle to the right of the positive imaginary λ axis. We shall therefore consider $\arg\lambda$ nearly equal to $\pi/2$, $|\lambda|$ large, and k real. The two significant zeros of $q(r)$ are then given by

$$\alpha \simeq \left(\frac{g}{|\lambda|}\right)^{2/n-2} \exp\left\{\frac{\pi - 2\arg\lambda}{n-2}\, i\right\};$$

$$\beta \simeq \frac{\lambda}{k}.$$

So $\arg\alpha$ is nearly zero and $\arg\beta$ is nearly $\pi/2$.

The problem of finding $S(\lambda, k)$ is to express the regular solution of the Schrödinger equation, whose behavior near the origin is given by Eq. (4.7), as a linear combination of the Jost solutions defined by their asymptotic behavior at infinity. Thus we wish to trace a WKB solution from the region between 0 and α to the region between β and infinity. It is possible to do this in a straightforward manner if there is an anti-Stokes line joining α and β; the condition for this is

$$\operatorname{Im}\int_\alpha^\beta [q(r)]^{\frac{1}{2}}\, dr = 0. \tag{A3}$$

We shall therefore assume that this condition holds and see later what restrictions it places on our results. The r plane is illustrated in Fig. 2. There are three Stokes lines, defined by $\operatorname{Re}\int_\alpha^r q^{\frac{1}{2}}\, dr = 0$, radiating from α and marked as dotted lines in Fig. 2. There are also three anti-Stokes lines, defined by $\operatorname{Im}\int_\alpha^r q^{\frac{1}{2}}\, dr = 0$, lying between the Stokes lines and marked as full lines in Fig. 2. A similar situation occurs at β. In region 1 of Fig. 2 (the r plane is divided into regions

[18] N. N. Khuri, Phys. Rev. **107**, 1148 (1957).

[19] For example, in diffraction scattering at high energies by the potential $V(r) = g^2r^{-4}$ [or $V(r) = g^2r^{-4}\exp(-\mu r)$], the quantity which defines the diffraction peak is $k^{\frac{4}{3}}(1 - \cos\theta)$ [R. H. Jones, Ph.D. thesis, University of Sussex (1966)].

[20] It is possible that t is important for potentials which are exponentially singular at the origin as $t^{\frac{1}{2}} = 2k\sin(\theta/2)$ does occur in the amplitude for scattering off a hard sphere. But neither the results of this paper nor Martin's results (Ref. 3) apply to this case.

[21] S. Fubini and R. Stroffolini, Nuovo Cimento **37**, 1812 (1965).

[22] G. G. Stokes, Trans. Camb. Phil. Soc. **10**, 106 (1857).

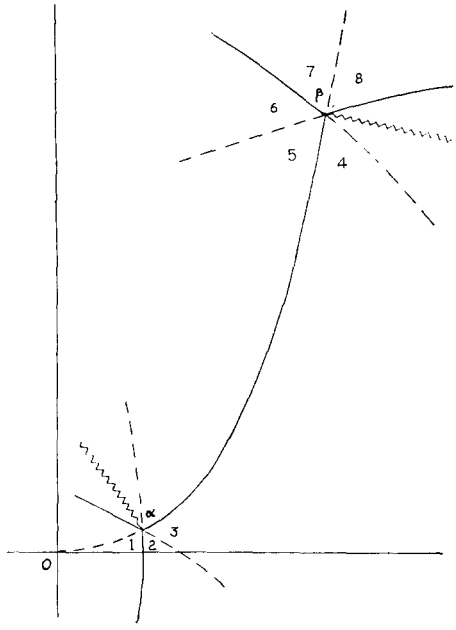


FIG. 2. Stokes and anti-Stokes lines in the r plane.

by the Stokes and anti-Stokes lines) the regular solution $\phi(r)$ may be identified with $\psi_2(r)$. In the notation of Heading[10] we write

$$\phi(r) \equiv (r, \alpha)_s, \tag{A4}$$

where

$$(r, \alpha) = q^{-\frac{1}{4}} \exp\left\{-i\int_\alpha^r q^{\frac{1}{2}}\, dr\right\}$$

and the suffix s denotes that the solution is subdominant or exponentially small. The regular solution may now be expressed in terms of the WKB solutions in other regions as follows:

$$\begin{aligned}
&\text{region 1.} \quad (r, \alpha)_s \equiv \phi(r),\\
&\text{region 2.} \quad (r, \alpha)_d,\\
&\text{region 3.} \quad (r, \alpha)_d + i(\alpha, r)_s,\\
&\text{region 4.} \quad [\beta, \alpha](r, \beta)_d + i[\alpha, \beta](\beta, r)_s,\\
&\text{region 5.} \quad [\beta, \alpha](r, \beta)_s + i[\alpha, \beta](\beta, r)_d,\\
&\text{region 6.} \quad ([\alpha, \beta] + [\beta, \alpha])(r, \beta)_s + i[\alpha, \beta](\beta, r)_d,\\
&\text{region 7.} \quad ([\alpha, \beta] + [\beta, \alpha])(r, \beta)_d + i[\alpha, \beta](\beta, r)_s,\\
&\text{region 8.} \quad ([\alpha, \beta] + [\beta, \alpha])(r, \beta)_d - i[\beta, \alpha](\beta, r)_s,
\end{aligned} \tag{A5}$$

where

$$(\alpha, r) = q^{-\frac{1}{4}} \exp\left\{i\int_\alpha^r q^{\frac{1}{2}}\, dr\right\}, \quad \text{etc.},$$

$$[\alpha, \beta] = \exp\left\{i\int_\alpha^\beta q^{\frac{1}{2}}\, dr\right\}, \quad \text{etc.},$$

and the suffix d denotes that the solution is dominant or exponentially large in the region considered. It may be seen from the above tabulation that: (i) on crossing an anti-Stokes line the subdominant term becomes dominant and vice versa; (ii) on crossing a Stokes line the coefficient of the subdominant term changes by i times the coefficient of the dominant term.

In region 8 of the r plane the WKB solutions are multiples of the Jost solutions:

$$\begin{aligned}
(r, \beta)_d &= \text{const} \times f_+(r),\\
(\beta, r)_s &= \text{const} \times f_-(r),
\end{aligned} \tag{A6}$$

where

$$f_\pm(r)_{r\to\infty} \exp(\mp ikr).$$

The multiplicative factors may be found as follows:

$$\begin{aligned}
&\lim_{r\to\infty} (r, \beta)_d \exp(ikr)\\
&= k^{-\frac{1}{2}} \lim_{r\to\infty} \exp\left\{ikr - i\int_\beta^r q^{\frac{1}{2}}\, dr'\right\}\\
&= k^{-\frac{1}{2}} \exp\left\{ik\beta - i\int_\beta^\infty \left[\left(k^2 - \frac{\lambda^2}{r^2} - \frac{g^2}{r^2} f(r)\right)^{\frac{1}{2}} - k\right] dr\right\}.
\end{aligned}$$

Now β is a generalization of the classical turning point, and so

$$\delta_{\mathrm{WKB}} = \frac{\lambda\pi}{2} - k\beta + \int_\beta^\infty \left\{\left(k^2 - \frac{\lambda^2}{r^2} - \frac{g^2}{r^n} f(r)\right)^{\frac{1}{2}} - k\right\} dr.$$

Thus

$$(r, \beta)_d = k^{-\frac{1}{2}} \exp\left(\frac{i\lambda\pi}{2} - i\delta_{\mathrm{WKB}}\right) f_+(r). \tag{A7}$$

Similarly, it can be shown that

$$(\beta, r)_s = k^{-\frac{1}{2}} \exp\left(-\frac{i\lambda\pi}{2} + i\delta_{\mathrm{WKB}}\right) f_-(r). \tag{A8}$$

Substituting these expressions in Eq. (A5),

$$S(\lambda, k) = \frac{\exp(2i\delta_{\mathrm{WKB}})}{1 + [\alpha, \beta]^2}, \tag{A9}$$

where we have used the relation $[\alpha, \beta][\beta, \alpha] = 1$. Equation (A9) has only been obtained when there is an anti-Stokes line joining α and β. Define a function $\Phi(\lambda, k)$ by

$$\Phi(\lambda, k) = -i\int_\alpha^\beta q^{\frac{1}{2}}\, dr. \tag{A10}$$

The above condition on the validity of Eq. (A9) may be summed up by the equation

$$S(\lambda, k) = \frac{\exp(2i\delta_{\mathrm{WKB}})}{1 + e^{-2\Phi}}, \quad \mathrm{Re}\,\Phi = 0. \tag{A11}$$

APPENDIX B

Here we shall calculate $\Phi(\lambda, k)$ for a potential given by Eq. (4.3) with $n = 4$:

$$\Phi(\lambda, k) = -i\int_\alpha^\beta \left(k^2 - \frac{\lambda^2}{r^2} - \frac{g^2}{r^4} f(r)\right)^{\frac{1}{2}} dr. \tag{B1}$$

As $f(r)$ has been defined as a bounded analytic function in the right half-plane and at the origin, we have

$$|f(r)| < A, \quad \text{for some finite } A, \quad \text{Re } r \geq 0, \tag{B2}$$

and

$$f(r) = 1 - ar + (b/2)r^2 + O(r^3),$$

in some region around the origin

Hence $\exists\ \zeta > 0$ such that $|r| < \zeta$ implies that

$$\begin{aligned} |f(r) - 1 + ar| &< |br^2|, \quad b \neq 0, \\ |f(r) - 1 + ar| &< |r^2|, \quad b = 0. \end{aligned} \tag{B3}$$

We shall assume in the following that $a \neq 0$ and $b \neq 0$ (e.g., the potential $g^2 r^{-4} e^{-ar}$). We note that ζ and A are now fixed quantities, i.e., they do not depend on λ. As we are considering the asymptotic behavior of Φ as $\lambda \to \infty$, we may thus write $\zeta = O(1)$ and $A = O(1)$.

α and β are zeros of $q(r)$ near the points ig/λ and λ/k, respectively; hence

$$\alpha = \frac{ig}{\lambda}\left\{1 - \frac{iag}{2\lambda} + O\left(\frac{1}{\lambda^2}\right)\right\}, \tag{B4}$$

$$\beta = \frac{\lambda}{k}\left\{1 + \frac{g^2k^2}{2\lambda^4} f(\beta) + \cdots\right\}. \tag{B5}$$

We split the range of integration in Eq. (B1) into four parts:

$$\begin{aligned} \Phi &= -i\left(\int_{\infty}^{\alpha_1} + \int_{\alpha_1}^{r_0} + \int_{r_0}^{\beta_1} + \int_{\beta_1}^{\beta}\right) q^{\frac{1}{2}}\, dr \\ &= I_1 + I_2 + I_3 + I_4 \quad \text{(say)}, \end{aligned} \tag{B6}$$

where

$$\alpha_1 = \frac{ig}{\lambda}\left\{1 + \frac{ag}{\lambda} e^{i\theta_1}\right\}, \tag{B7}$$

$$r_0 = \zeta e^{i\theta_2}, \tag{B8}$$

$$\beta_1 = \frac{\lambda}{k}\{1 + e^{i\theta_3}/\lambda_3\}, \tag{B9}$$

and θ_1, θ_2, θ_3 are phase angles chosen such that all three points lie on the contour of integration joining α and β. Hence

$$r \in [\alpha, r_0] \Rightarrow f(r) = 1 - ar + Cw(r)r^2, \quad |w(r)| < 1, \tag{B10}$$

$$r \in [r_0, \beta] \Rightarrow |f(r)| < A. \tag{B11}$$

The notation $r \in [\alpha, r_0]$ is here taken to mean that r lies on the contour of integration between the points α and r_0.

(i) $r \in [\alpha, \alpha_1] \Rightarrow |q(r)| < O(\lambda^3)$. Hence

$$I_1 \leq O(\lambda^{\frac{3}{2}})\,|\alpha_1 - \alpha| = O(\lambda^{-\frac{1}{2}}). \tag{B12}$$

(ii) $r \in [\alpha_1, r_0] \Rightarrow \left|\frac{\lambda^2}{r^2} + \frac{g^2}{r^4} - \frac{ag^2}{r^3}\right|$

$$\geq O(\lambda)\left|k^2 - \frac{Cw(r)g^2}{r^2}\right|.$$

$$I_2 = \int_{\alpha_1}^{r_0}\left(\frac{\lambda^2}{r^2} + \frac{g^2}{r^4} - \frac{ag^2}{r^3}\right)^{\frac{1}{2}} dr - \frac{1}{2}\left\{1 + O\left(\frac{1}{\lambda}\right)\right\}$$

$$\times \int_{\alpha_1}^{r_0}\left(\frac{\lambda^2}{r^2} + \frac{g^2}{r^4} - \frac{ag^2}{r^3}\right)^{-\frac{1}{2}}\left(k^2 - \frac{Cw(r)g^2}{r^2}\right) dr.$$

The expression under the square-root sign is a quadratic in r and may be factorized. Its roots are close to α and $-\alpha$. Using this property, the integrals may be calculated to obtain

$$I_2 = \lambda \ln\left\{\frac{2\lambda r_0}{eig}\right\} + \frac{iag\pi}{4} + O(\lambda^{-\frac{1}{2}}). \tag{B13}$$

(iii) $r \in [r_0, \beta_1] \Rightarrow \left|k^2 - \frac{\lambda^2}{r^2}\right| \geq O(\lambda)\left|\frac{g^2}{r^4} f(r)\right|.$

$$I_3 = \int_{r_0}^{\beta_1}\left(\frac{\lambda^2}{r^2} - k^2\right)^{\frac{1}{2}} dr$$

$$+ \frac{g^2}{2}\left\{1 + O\left(\frac{1}{\lambda}\right)\right\}\int_{r_0}^{\beta_1}\left(\frac{\lambda^2}{r^2} - k^2\right)^{\frac{1}{2}} \frac{f(r)}{r^4}\, dr.$$

By splitting up the interval of integration of the second part at $O(\lambda^{\frac{1}{2}})$ and $O(\lambda^{\frac{3}{4}})$, it may easily be bounded by $O(\lambda^{-\frac{1}{2}})$. Hence

$$I_3 = \lambda \ln\left\{\frac{2\lambda}{ekr_0}\right\} + O(\lambda^{-\frac{1}{2}}). \tag{B14}$$

(iv) $r \in [\beta_1, \beta] \Rightarrow |q(r)| \leq O(\lambda^{-3})$. Hence

$$I_4 \leq O(\lambda^{-\frac{3}{2}})\,|\beta_1 - \beta| = O(\lambda^{-\frac{7}{2}}). \tag{B15}$$

Adding together Eqs. (B12) to (B15),

$$\Phi(\lambda, k) = 2\lambda \ln\left\{\frac{2\lambda}{e(igk)^{\frac{1}{2}}}\right\} + \frac{iag\pi}{4} + O(\lambda^{-\frac{1}{2}}). \tag{B16}$$

Particlelike Solutions to Nonlinear Complex Scalar Field Theories with Positive-Definite Energy Densities*

GERALD ROSEN
Drexel Institute of Technology, Philadelphia, Pennsylvania

(Received 15 September 1967)

It is shown that a self-interacting complex scalar field theory with a positive-definite energy density can admit spatially localized singularity-free particlelike solutions. A condition on the self-interaction energy density, sufficient to guarantee the existence of such solutions, is that its derivative should be nonincreasing and not identically constant as the squared absolute value of the field increases from zero.

Consider a generic self-interacting complex scalar field theory derived from a Lorentz-invariant Lagrangian density of the form

$$\mathfrak{L} = |\dot{\psi}|^2 - |\nabla\psi|^2 - U(|\psi|^2), \tag{1}$$

where $U(\chi^2)$ is a nonnegative self-interaction energy density, a real continuous piecewise C^1 function for all χ^2, of the order χ^2 as $\chi^2 \to 0$, but otherwise arbitrary: $U(\chi^2) \geq 0$, $U(0) = 0$, $U'(\chi^2)$ piecewise continuous, $0 < U'(0) < \infty$.[1] Let us seek spatially localized singularity-free solutions to the field equation derived from (1) by putting $\psi = \chi e^{-i\omega t}$ with χ a real C^1 piecewise C^2 function of $\mathbf{x}$ alone and ω a real constant. Since the reduced field equation

$$\nabla^2\chi + \omega^2\chi - U'(\chi^2)\chi = 0 \tag{2}$$

follows consistently from the variational principle

$$\delta\int \mathfrak{L}\, d^3x = 0, \quad \mathfrak{L} = \omega^2\chi^2 - |\nabla\chi|^2 - U(\chi^2), \tag{3}$$

we have the pseudovirial theorem[2] applicable for a singularity-free C^1 piecewise C^2 solution χ such that

$$\lim_{|\mathbf{x}|\to\infty} [|\mathbf{x}|^3\chi(\mathbf{x})] = 0. \tag{4}$$

A statement of the pseudovirial theorem

$$\int [3\omega^2\chi^2 - |\nabla\chi|^2 - 3U(\chi^2)]d^3x = 0 \tag{5}$$

in combination with Eq. (2) produces the global condition

$$\int [2\,|\nabla\chi|^2 - 3U(\chi^2) + 3\chi^2U'(\chi^2)]d^3x = 0, \tag{6}$$

which implies that

$$U(\chi^2) > \chi^2U'(\chi^2), \tag{7}$$

for some values of χ^2, is necessary for the existence of such a particlelike solution.[3] Equation (7) precludes the existence of localized singularity-free particlelike solutions for a large class of simple nonnegative $U(\chi^2)$. Notwithstanding the absence of any patently negative "binding energy," it is however possible to formulate conditions on a nonnegative $U(\chi^2)$ compatible with Eq. (7) and sufficient to guarantee the existence of particlelike solutions, as shown in the following.

First observe that the condition (7) is satisfied strongly if $U'(\chi^2)$ is nonincreasing and not identically constant as χ^2 increases from zero, for then

$$U(\chi^2) - \chi^2U'(\chi^2) = \int_0^{\chi^2} [U'(\xi) - U'(\chi^2)]\, d\xi > 0 \tag{8}$$

for some values of χ^2. Particlelike solutions to Eq. (2) are in general obtainable with $U'(\chi^2)$ nonincreasing and not identically constant as χ^2 increases from zero, positive spherically symmetric solutions described in first approximation by

$$\chi_1 \equiv \begin{cases} \dfrac{\alpha \sin\{[\omega^2 - U'(\alpha^2)]^{\frac{1}{2}}\,|\mathbf{x}|\}}{[\omega^2 - U'(\alpha^2)]^{\frac{1}{2}}\,|\mathbf{x}|}, & |\mathbf{x}| \leq R, \\ \dfrac{\alpha \exp\{[U'(0) - \omega^2]^{\frac{1}{2}}(R - |\mathbf{x}|)\}}{[U'(0) - U'(\alpha^2)]^{\frac{1}{2}}\,|\mathbf{x}|}, & |\mathbf{x}| \geq R, \end{cases} \tag{9}$$

where

$$R = [\omega^2 - U'(\alpha^2)]^{-\frac{1}{2}}\left[\pi - \tan^{-1}\left\{\frac{[\omega^2 - U'(\alpha^2)]^{\frac{1}{2}}}{[U'(0) - \omega^2]^{\frac{1}{2}}}\right\}\right]$$

$$= [U'(0) - U'(\alpha^2)]^{-\frac{1}{2}}\,\frac{\alpha}{\beta} \tag{10}$$

with β the smallest positive value of χ for which $U'(\chi^2) \leq \omega^2$. Requiring

$$U'(0) > \omega^2 \geq U'(\beta^2) > U'(\alpha^2),$$

the second and third members of Eq. (10) fix α for

* Work supported by a National Science Foundation grant.

[1] See Appendix A.

[2] G. Rosen, J. Math. Phys. 7, 2066 (1966).

[3] It can be shown that Eq. (7) is necessary for the existence of any spatially localized and temporally periodic singularity-free particle-like solution derived from (1). For the method of proof, see: G. Rosen, J. Math. Phys. 7, 2071 (1966).

prescribed ω^2, and so the solutions (9) are parametrized completely by ω^2, free to range continuously over a certain open interval.[4] For the specific self-interaction energy density

$$U(\chi^2) = \begin{cases} m_0^2\chi^2, & \chi^2 \leq \beta^2, \\ m_1^2\chi^2 + (m_0^2 - m_1^2)\beta^2, & \chi^2 \geq \beta^2, \end{cases} \quad (11)$$

the solutions (9) are exact with admissible values of ω^2 such that $U'(0) = m_0^2 > \omega^2 > m_1^2 = U'(\alpha^2)$, the second and third members of Eq. (10) producing

$$\alpha = \beta\left[\frac{(m_0^2 - m_1^2)}{(\omega^2 - m_1^2)}\right]^{\frac{1}{2}}\left\{\pi - \tan^{-1}\left[\frac{(\omega^2 - m_1^2)^{\frac{1}{2}}}{(m_0^2 - \omega^2)^{\frac{1}{2}}}\right]\right\}.$$

To verify the approximate validity[5] and improve upon the accuracy of (9) for general $U(\chi^2)$, we evoke the integral form of Eq. (2):

$$\begin{aligned}\chi(\mathbf{x}) &= \Omega(\chi; \mathbf{x}) \\ &\equiv \frac{1}{4\pi}\int \frac{[\exp -[U'(0 - \omega^2]^{\frac{1}{2}}|\mathbf{x} - \mathbf{y}|]}{|\mathbf{x} - \mathbf{y}|} \\ &\quad \times [U'(0) - U'(\chi(\mathbf{y})^2)]\chi(\mathbf{y})d^3y \end{aligned} \quad (12)$$

and set up the iterative approximation series

$$\chi_{n+1}(\mathbf{x}) = \Omega(\chi_n; \mathbf{x}), \quad (13)$$

with Eq. (9) serving as the first approximation. For spherically symmetric $\chi(\mathbf{x}) = \bar{\chi}(|\mathbf{x}|)$, the integral transform in Eq. (12) is expressible as

$$\Omega(\chi; \mathbf{x}) = \int_0^\infty G(|\mathbf{x}|, s)[U'(0) - U'(\bar{\chi}(s)^2)]\bar{\chi}(s)\,ds,$$

with

$$G(r, s) \equiv \begin{cases} \frac{s}{\mu r}e^{-\mu r}\sinh \mu s, & s \leq r, \\ \frac{s}{\mu r}e^{-\mu s}\sinh \mu r, & r \leq s, \end{cases}$$

where $\mu \equiv [U'(0) - \omega^2]^{\frac{1}{2}}$. Because (9) satisfies the linear integral equation

$$\begin{aligned}\chi_1(\mathbf{x}) = \frac{1}{4\pi}\int_{|\mathbf{y}|\leq R} &\frac{[\exp -[U'(0) - \omega^2]^{\frac{1}{2}}|\mathbf{x} - \mathbf{y}|]}{|\mathbf{x} - \mathbf{y}|} \\ &\times [U'(0) - U'(\alpha^2)]\chi_1(\mathbf{y})d^3y, \end{aligned} \quad (14)$$

it follows that

$$\begin{aligned}\chi_2(\mathbf{x}) - \chi_1(\mathbf{x}) = \frac{1}{4\pi}&\left\{-\int_{|\mathbf{y}|\leq R}[U'(\chi_1(\mathbf{y})^2) - U'(\alpha^2)]\right. \\ &\left. + \int_{|\mathbf{y}|\geq R}[U'(0) - U'(\chi_1(\mathbf{y})^2)]\right\} \\ &\times \frac{[\exp -[U'(0) - \omega^2]^{\frac{1}{2}}|\mathbf{x} - \mathbf{y}|]}{|\mathbf{x} - \mathbf{y}|}\chi_1(\mathbf{y})d^3y. \end{aligned} \quad (15)$$

Specialized analysis applied to the right side of Eq. (15) can be used to bound the left side [generally small in absolute magnitude compared to $\chi_1(\mathbf{x})$] for any $U(\chi^2)$ such that $U'(\chi^2)$ is nonincreasing for $\chi^2 \leq \alpha^2$; a crude but general lower bound on the second approximation is obtained immediately from Eqs. (13) and (14) as

$$\chi_2(\mathbf{x}) > \left[\frac{U'(0) - U'(\beta^2)}{U'(0) - U'(\alpha^2)}\right]\chi_1(\mathbf{x}). \quad (16)$$

For certain $U(\chi^2)$ it is more convenient to use the first approximation[6]

$$\chi_1 \equiv \frac{\alpha \tanh \mu|\mathbf{x}|}{\mu|\mathbf{x}|\cosh \mu|\mathbf{x}|}, \quad (17)$$

where $\mu^2 \equiv \frac{1}{6}[U'(0) - U'(\alpha^2)]$ and

$$\omega^2 = \tfrac{1}{6}[5U'(0) + U'(\alpha^2)]. \quad (18)$$

The condition (18) associated with (17) has the virtue of fixing α for prescribed ω^2 by an analytical relationship more explicit than the condition (10) associated with (9). Note, however, that the first approximation (17) would be inappropriate for a self-interaction energy density such that

$$[U'(0) - U'(\chi^2)] \leq [U'(0) - U'(\alpha^2)](\chi/\alpha)^2$$

for $\chi^2 \leq \alpha^2$; if the latter (rather stringent) condition is satisfied by the nonincreasing $U'(\chi^2)$, then an upper bound on the second approximation follows from Eqs. (13) and (17), $\chi_2(\mathbf{x}) < \chi_1(\mathbf{x})$ or $\chi_2(\mathbf{x}) \leq K\chi_1(\mathbf{x})$ for a positive constant $K < 1$, since (17) satisfies the linear integral equation and associated inequality

$$\chi_1(\mathbf{x}) = \frac{1}{4\pi}\int \frac{[\exp -\mu|\mathbf{x} - \mathbf{y}|]}{|\mathbf{x} - \mathbf{y}|}\frac{6\mu^2\chi_1(\mathbf{y})}{[\cosh \mu|\mathbf{y}|]^2}d^3y \quad (19)$$

$$> \frac{1}{4\pi}\int \frac{[\exp -[U'(0) - \omega^2]^{\frac{1}{2}}|\mathbf{x} - \mathbf{y}|]}{|\mathbf{x} - \mathbf{y}|} \times \alpha^{-2}[U'(0) - U'(\alpha^2)]\chi_1(\mathbf{y})^3d^3y. \quad (20)$$

Hence, by mathematical induction based on Eq. (13) we have $\chi_n(\mathbf{x}) \leq K^{n-1}\chi_1(\mathbf{x})$, and so

$$\chi(\mathbf{x}) = \lim_{n\to\infty}\chi_n(\mathbf{x}) = 0$$

[4] Usually the critical case $\omega^2 = U'(0)$ does not admit a particlelike solution of finite total energy, but there are certain special nonnegative piecewise C^1 self-interaction energy densities with $U'(\chi^2)$ nonincreasing for which a physically acceptable particlelike solution is obtainable with $\omega^2 = U'(0)$.

[5] Rigorous existence and convergence theorems for the particlelike solutions can be based on the iterative approximation series, Eq. (13). For their construction with Eq. (13), see: M. A. Krasnosel'skii, *Topological Methods in the Theory of Nonlinear Integral Equations* (The Macmillan Company, New York, 1964), pp. 123–168; *Positive Solutions of Operator Equations* (P. Noordhoff Ltd., Groningen, The Netherlands, 1964), pp. 230–245.

[6] See Appendix B.

is the trivial solution with (17) as the first approximation if $[U'(0) - U'(\chi^2)] \leq [U'(0) - U'(\alpha^2)](\chi/\alpha)^2$ for $\chi^2 \leq \alpha^2$.

As an example of a physically interesting nonnegative self-interaction energy density $U(\chi^2)$ with $U'(\chi^2)$ nonincreasing, we have the form (expressed in convenient physical units)

$$U(\chi^2) = g(1 - e^{-\chi^2}) \tag{21}$$

with the composition-saturation property $U(\xi + \eta) = U(\xi) + U(\eta) - g^{-1}U(\xi)U(\eta)$. A first approximation to the positive spherically symmetric particlelike solutions to Eq. (2) with (21) is given by Eq. (17) where $\mu^2 = \frac{1}{6}g(1 - e^{-\alpha^2})$, condition (18) implying that the solutions are obtainable for ω^2 $[= \frac{1}{6}g(5 + e^{-\alpha^2})]$ such that $g > \omega^2 > \frac{5}{6}g$.

That $U'(\chi^2)$ be nonincreasing and not identically constant as χ^2 increases from zero apears to be the simplest general requirement on the self-interaction energy density sufficient to guarantee the existence of spatially localized singularity-free particlelike solutions to Eq. (2). However, the latter requirement is by no means necessary for the existence of particlelike solutions, and one can easily find nonnegative piecewise C^1 $U(\chi^2)$ of a more complicated character which also admit particlelike solutions. One obvious example is provided by the form (expressed in convenient physical units)

$$U(\chi^2) = \begin{cases} m_0^2\chi^2, & \chi^2 \leq 1, \\ m_0^2\chi^2 - g(|\chi| - 1), & \chi^2 \geq 1, \end{cases} \tag{22}$$

where g is a positive constant; particlelike solutions to Eq. (2) with (22), obtainable if ω^2 is such that $m_0^2 > \omega^2 > (m_0^2 - g/4)$, are given exactly by

$$\chi = \tfrac{1}{2}g\mu^{-2}J(|\mathbf{x}|), \tag{23}$$

$$J(|\mathbf{x}|) \equiv \frac{\mu^2}{4\pi}\int_{|\mathbf{y}|\leq\kappa/\mu} \frac{e^{-\mu|\mathbf{x}-\mathbf{y}|}}{|\mathbf{x} - \mathbf{y}|}\, d^3y$$
$$= \begin{cases} \left(1 - (1 + \kappa)e^{-\kappa}\dfrac{\sinh \mu\,|\mathbf{x}|}{\mu\,|\mathbf{x}|}\right), & |\mathbf{x}| \leq \kappa/\mu, \\ (\kappa \cosh \kappa - \sinh \kappa)\dfrac{e^{-\mu|\mathbf{x}|}}{\mu\,|\mathbf{x}|}, & |\mathbf{x}| \geq \kappa/\mu, \end{cases} \tag{24}$$

where $\mu^2 \equiv m_0^2 - \omega^2$ and κ is the positive root of the equation $(1 + \kappa^{-1})(1 - e^{-2\kappa}) = 2 - (4\mu^2/g)$. It would appear to be very difficult to establish conditions that are both necessary and sufficient on a generic nonnegative piecewise C^1 $U(\chi^2)$ for the existence of particlelike solutions to Eq. (2).

APPENDIX A

It should be remarked that the finiteness of $U'(0)$ is *not* necessary for the existence of a particlelike solution, but rather imposed here as a desirable physical property of the self-interaction energy density. An example of a real continuous nonnegative piecewise C^1 self-interaction density with $U'(0) = \infty$ admitting particlelike solutions is provided by

$$U(\chi^2) = \begin{cases} g\chi^2 \ln(\chi^{-2}), & \chi^2 \leq 1, \\ 0, & \chi^2 \geq 1, \end{cases}$$

where g is a positive constant and convenient physical units are employed. Associated particlelike solutions to Eq. (2), with

$$U'(\chi^2) = \begin{cases} g[-1 + \ln(\chi^{-2})], & \chi^2 \leq 1, \\ 0, & \chi^2 \geq 1, \end{cases}$$

take the form

$$\chi = \exp\left(1 - \frac{\omega^2}{2g} - \frac{g\,|\mathbf{x}|^2}{2}\right)$$

for all finite positive $\omega^2 \geq 2g$.

APPENDIX B

Equation (2) is satisfied about $\mathbf{x} = 0$ and asymptotically as $|\mathbf{x}| \to \infty$ by either first approximation (9) or (17). A third alternative first approximation to the particlelike solutions, generally in close correspondence with the latter two forms for $\mu\,|\mathbf{x}| \lesssim 1$, is given by

$$\chi_1 = 6\alpha I(|\mathbf{x}|),$$
$$I(|\mathbf{x}|) \equiv \frac{\mu^2}{4\pi}\int_{|\mathbf{y}|\leq\tau/\mu} \frac{e^{-\mu|\mathbf{x}-\mathbf{y}|}}{|\mathbf{x} - \mathbf{y}|}\, d^3y$$
$$= \begin{cases} \left(1 - \dfrac{5\sinh \mu\,|\mathbf{x}|}{6\mu\,|\mathbf{x}|}\right), & |\mathbf{x}| \leq \tau/\mu, \\ (\tfrac{3}{5}\tau^2 - \tfrac{11}{60})\dfrac{e^{-\mu|\mathbf{x}|}}{\mu\,|\mathbf{x}|}, & |\mathbf{x}| \geq \tau/\mu, \end{cases}$$

where $\mu^2 \equiv \frac{1}{6}[U'(0) - U'(\alpha^2)]$, $\tau \cong 0.732$ is the positive real root of the equation $(1 + \tau)e^{-\tau} = \frac{5}{6}$, and ω^2 is again related to α by Eq. (18); we obtain this third alternative first approximation by evoking Eq. (13) and the simple zeroth approximation

$$\chi_0 = \begin{cases} \alpha, & |\mathbf{x}| \leq \tau/\mu, \\ 0, & |\mathbf{x}| \geq \tau/\mu. \end{cases}$$

Charged Particlelike Solutions to Nonlinear Complex Scalar Field Theories*

GERALD ROSEN
Drexel Institute of Technology, Philadelphia, Pennsylvania

(Received 15 September 1967)

It is shown that spatially localized singularity-free particlelike solutions exist for Lorentz-covariant complex scalar field theories with minimal gauge-invariant electromagnetic coupling, a positive-definite energy density, and suitably prescribed nonlinear self-interaction. Such a theory provides a perfectly consistent structural model on the classical level for a charged elementary particle of finite positive energy.

Although there have been many attempts to formulate a classical field theory for the structure of a charged elementary particle,[1] none of the proposed Lorentz-invariant local field theories has featured minimal gauge-invariant electromagnetic coupling, a positive-definite energy density, and singularity-free particlelike solutions of finite energy. As observed previously by several authors, a divergence-free quantum field theory for a charged elementary particle may follow from a Lorentz-invariant local classical field theory with minimal gauge-invariance, positive-definite energy, and singularity-free particlelike solutions. Our purpose in the present paper is to exhibit a generic classical field theory which possesses all of these desirable physical features.

We consider a field theory associated with a Lorentz-invariant and gauge-invariant Lagrangian density of the form[2]

$$\mathfrak{L} = -(\partial^\mu\psi^* - i\epsilon a^\mu\psi^*)(\partial_\mu\psi + i\epsilon a_\mu\psi) - U - \tfrac{1}{4}f^{\mu\nu}f_{\mu\nu}, \tag{1}$$

where ψ is a complex charge-carrying scalar field, a_μ is the real electromagnetic four-potential, $U = U(|\psi|^2)$ (≥ 0) is a generic nonnegative self-interaction energy density, and $f_{\mu\nu} = \partial_\nu a_\mu - \partial_\mu a_\nu$ is the electromagnetic-field tensor; physical units are such that $\hbar$ and c equal unity, so we have $\epsilon^2/4\pi \cong \frac{1}{137}$ with $|\epsilon|$ the fundamental unit of electric charge. A positive-definite canonical energy density follows from (1),[3]

$$\begin{aligned}\mathfrak{E} = {} & (\partial_0\psi^* - i\epsilon a_0\psi^*)(\partial_0\psi + i\epsilon a_0\psi) \\ & + (\partial_k\psi^* - i\epsilon a_k\psi^*)(\partial_k\psi + i\epsilon a_k\psi) + U \\ & + \tfrac{1}{2}f_{0k}f_{0k} + \tfrac{1}{4}f_{jk}f_{jk},\end{aligned} \tag{2}$$

while the field equations are

$$(\partial^\mu + i\epsilon a^\mu)(\partial_\mu + i\epsilon a_\mu)\psi - U'\psi = 0, \tag{3}$$

$$\partial_\nu f^{\mu\nu} - i\epsilon(\psi\partial^\mu\psi^* - \psi^*\partial^\mu\psi) - 2\epsilon^2 a^\mu\psi^*\psi = 0. \tag{4}$$

Despite the positive-definite character of the energy density (2) with $U \geq 0$ and the absence of any patently negative "binding energy," the field equations (3) and (4) admit spatially localized singularity-free particlelike solutions of finite energy for a suitably prescribed nonnegative self-interaction energy density, as shown by the analysis which follows.

By putting $\psi = \chi e^{-i\omega x^0}$ with χ a real function of $\mathbf{x} = (x_1, x_2, x_3)$ alone, ω a real constant, and $a_\mu = (\phi, 0, 0, 0)$ with ϕ a function of $\mathbf{x}$ alone, the Lagrangian density (1) reduces to

$$\mathfrak{L} = (\omega - \epsilon\phi)^2\chi^2 - |\nabla\chi|^2 - U(\chi^2) + \tfrac{1}{2}|\nabla\phi|^2, \tag{5}$$

and the field equations (3) and (4) simplify to

$$\nabla^2\chi + (\omega - \epsilon\phi)^2\chi - U'(\chi^2)\chi = 0, \tag{6}$$

$$\nabla^2\phi + 2\epsilon(\omega - \epsilon\phi)\chi^2 = 0. \tag{7}$$

Since the latter field equations follow consistently from the variational principle

$$\delta\int\mathfrak{L}\,d^3\mathbf{x} = 0, \tag{8}$$

with $\mathfrak{L}$ the constrained Lagrangian density (5), we have the pseudovirial theorem[4] producing the global condition

$$\int[3(\omega - \epsilon\phi)^2\chi^2 - |\nabla\chi|^2 - 3U(\chi^2) + \tfrac{1}{2}|\nabla\phi|^2]d^3\mathbf{x} = 0 \tag{9}$$

for a spatially localized singularity-free solution such that

$$\lim_{|\mathbf{x}|\to\infty}[|\mathbf{x}|^{\frac{3}{2}}\chi(\mathbf{x})] = 0, \quad \lim_{|\mathbf{x}|\to\infty}[|\mathbf{x}|^{\frac{1}{2}}\phi(\mathbf{x})] = 0, \tag{10}$$

and with $U(\chi^2)$ of the order χ^2 as $\chi^2 \to 0$. On the other hand, the field equations (6) and (7) produce the

* Work supported by a National Science Foundation grant.

[1] M. Born and L. Infeld, Proc. Roy. Soc. (London) **A144**, 425 (1934); N. Rosen, Phys. Rev. **55**, 94 (1939); A. C. Menius, Jr. and N. Rosen, Phys. Rev. **62**, 436 (1942); R. J. Finkelstein, Phys. Rev. **75**, 1079 (1949); R. Finkelstein, R. LeLevier, and M. Ruderman, Phys. Rev. **83**, 326 (1951); P. A. M. Dirac, Proc. Roy. Soc. (London) **A257**, 32 (1960); C. Gilbert, Proc. Phys. Soc. (London) **83**, 181 (1964); M. Wakano, Progr. Theoret. Phys. (Kyoto) **35**, 1117 (1966).

[2] Tensor indices run 0, 1, 2, 3, the Minkowski metric which raises indices has signature +2, and the summation convention is understood here.

[3] See, for example: G. Wentzel, *Quantum Theory of Fields* (Interscience Publishers, Inc., New York, 1949), pp. 66–69.

[4] G. Rosen, J. Math. Phys. **7**, 2066 (1966).

integral conditions

$$\int [-|\nabla\chi|^2 + (\omega - \epsilon\phi)^2\chi^2 - \chi^2 U'(\chi^2)]d^3\mathbf{x} = 0, \quad (11)$$

$$\int [-\tfrac{1}{2}|\nabla\phi|^2 + \epsilon\phi(\omega - \epsilon\phi)\chi^2]d^3\mathbf{x} = 0 \quad (12)$$

for a singularity-free solution localized in space according to Eqs. (10). It follows from Eq. (12) that ω cannot vanish, excluding a purely static solution, and that the quantity $\int \epsilon\phi\chi^2 d^3\mathbf{x}$ must have the same sign as ω. By subtracting $\frac{1}{3}$ of Eq. (9) from Eq. (11), we obtain

$$\int [-\tfrac{2}{3}|\nabla\chi|^2 - \tfrac{1}{6}|\nabla\phi|^2 + U(\chi^2) - \chi^2 U'(\chi^2)]d^3\mathbf{x} = 0, \quad (13)$$

a relation which implies that

$$\int [U(\chi^2) - \chi^2 U'(\chi^2)]d^3\mathbf{x} > 0 \quad (14)$$

and hence that

$$U(\chi^2) > \chi^2 U'(\chi^2), \quad (15)$$

for some χ^2, is necessary for the existence of such a solution.[5] A useful global necessary condition for the existence of a particlelike solution is derived by using Eqs. (11) and (12) to eliminate the gradient terms in Eq. (9),

$$\int [(2\omega - \epsilon\phi)(\omega - \epsilon\phi)\chi^2 - 3U(\chi^2) + \chi^2 U'(\chi^2)]d^3\mathbf{x} = 0. \quad (16)$$

By employing Eqs. (2), (11), (12), and (16), the total energy associated with a particlelike solution is expressed as

$$\begin{aligned} E &\equiv \int \mathcal{E}\, d^3\mathbf{x} \\ &= \int [(\omega - \epsilon\phi)^2\chi^2 + |\nabla\chi|^2 + U(\chi^2) + \tfrac{1}{2}|\nabla\phi|^2]d^3\mathbf{x} \\ &= \int [4U(\chi^2) - 2\chi^2 U'(\chi^2)]d^3\mathbf{x}. \end{aligned} \quad (17)$$

The total charge associated with a particlelike solution follows immediately from Eq. (7) as

$$Q \equiv \lim_{|\mathbf{x}|\to\infty} [4\pi|\mathbf{x}|\,\phi(\mathbf{x})] = 2\epsilon \int (\omega - \epsilon\phi)\chi^2 d^3\mathbf{x}. \quad (18)$$

Now by application of well-known general theorems,[6] it follows that the solution functions $\chi = \chi(\mathbf{x}; \epsilon)$ and $\phi = \phi(\mathbf{x}; \epsilon)$ are analytic in ϵ about $\epsilon = 0$ for a singularity-free solution to Eqs. (6) and (7).[7] Hence, let us regard the small quantity ϵ as a perturbation parameter and suppose that $|\omega|$ is large compared to the maximum value of $|\epsilon\phi|$. Then by substituting the perturbative expansions

$$\chi = \chi_{(0)} + \epsilon^2\chi_{(2)} + O(\epsilon^4), \quad (19)$$

$$\phi = \epsilon\phi_{(1)} + O(\epsilon^3) \quad (20)$$

into Eqs. (6) and (7) and equating to zero coefficients of powers of ϵ, we find a series of simpler equations to be solved in succession, the first three equations in the series being

$$\nabla^2\chi_{(0)} + \omega^2\chi_{(0)} - U'(\chi_{(0)}^2)\chi_{(0)} = 0, \quad (21)$$

$$\nabla^2\phi_{(1)} + 2\omega\chi_{(0)}^2 = 0, \quad (22)$$

$$[\nabla^2 + \omega^2 - U'(\chi_{(0)}^2) - 2\chi_{(0)}^2 U''(\chi_{(0)}^2)]\chi_{(2)} = 2\omega\chi_{(0)}\phi_{(1)}. \quad (23)$$

Equation (21) has been encountered and studied previously in the context of self-interacting complex scalar field theories *without* electric charge and electromagnetic coupling.[8] The observation has been made that condition (15) is satisfied strongly if $U'(\chi^2)$ is nonincreasing and not identically constant as χ^2 increases from zero, and it has been shown that Eq. (21) admits singularity-free spatially localized particlelike solutions for such a $U'(\chi^2)$ [and also for certain $U'(\chi^2)$ of a more complicated character]. Obtainable by means of an iterative approximation procedure (but not generally expressible in closed form), the positive spherically symmetric particlelike solutions to Eq. (21) are parametrized completely by ω^2, free to range continuously over a certain open interval with $\omega^2 < U'(0)$.[9] This one-parameter dependency of spherically symmetric $\chi_{(0)}(\mathbf{x}) = \bar{\chi}_{(0)}(|\mathbf{x}|)$ solutions to Eq. (21) is communicated to the singularity-free (necessarily spherically symmetric) solutions to Eq. (22),

$$\phi_{(1)} = 2\omega\left(\frac{1}{|\mathbf{x}|}\int_0^{|\mathbf{x}|} r + \int_{|\mathbf{x}|}^{\infty}\right)\bar{\chi}_{(0)}(r)^2 r\, dr, \quad (24)$$

parametrized completely by ω. For a certain value of ω we may have *charge consistency* of the particlelike solution, that is, asymptotic convergence of ϕ to the Coulomb law $\epsilon/4\pi|\mathbf{x}|$ for a charge ϵ, so that

[5] It can be shown that Eq. (15) is necessary for the existence of any spatially localized and temporally periodic, singularity-free, particlelike solution derived from the more general reduced Lagrangian density

$$\mathcal{L} = |\partial_0\psi + i\epsilon\psi|^2 - |\nabla\psi|^2 - U + \tfrac{1}{2}|\nabla\phi|^2.$$

For the method of proof, see: G. Rosen, J. Math. Phys. 7, 2071 (1966).

[6] For example, see: E. Kamke, *Differentialgleichungen Lösungsmethoden und Lösungen* (Akademische Verlagsgesellschaft, Leipzig, 1944), pp. 62ff., 102ff., 137ff., and 213ff.

[7] See Appendix A.

[8] G. Rosen, J. Math. Phys. 9, 996 (1968).

[9] See Appendix B.

Eq. (18) produces $Q = \epsilon$. By solving the equation

$$\lim_{|\mathbf{x}| \to \infty} [4\pi |\mathbf{x}| \phi_{(1)}(\mathbf{x})] = 1, \tag{25}$$

we obtain a first approximation to the discrete value or values of ω for charge consistency, a second approximation to order ϵ^2 being obtained by solving the equation

$$2\int [\omega(\chi_{(0)}^2 + 2\epsilon^2 \chi_{(0)}\chi_{(2)}) - \epsilon^2 \chi_{(0)}^2 \phi_{(1)}] d^3x = 1. \tag{26}$$

The solution $\chi_{(2)}$ to the linear inhomogeneous equation (23) appears in Eq. (26) and is also needed to compute the total energy to order ϵ^2,

$$E = \int \{[4U(\chi_{(0)}^2) - 2\chi_{(0)}^2 U'(\chi_{(0)}^2)] + 4\epsilon^2[U'(\chi_{(0)}^2) - \chi_{(0)}^2 U''(\chi_{(0)}^2)]\chi_{(0)}\chi_{(2)}\} d^3\mathbf{x} + O(\epsilon^4), \tag{27}$$

which follows from the final member in Eq. (17). Also note that Eq. (18) can be used to write the total energy (17) as

$$E = \frac{\omega Q}{\epsilon} + \int [-\omega^2\chi^2 + \epsilon^2\phi^2\chi^2 + |\nabla\chi|^2 + U(\chi^2) + \tfrac{1}{2}|\nabla\phi|^2] d^3\mathbf{x}.$$

Then with $Q = \epsilon$ and ω determined by Eq. (26), we find

$$E = \omega + \int [U(\chi_{(0)}^2) - \chi_{(0)}^2 U'(\chi_{(0)}^2) + \epsilon^2 \omega \chi_{(0)}^2 \phi_{(1)}] d^3\mathbf{x} + O(\epsilon^4),$$

by recalling Eqs. (19)–(22).

The nature of particlelike solutions to Eqs. (21)–(23) can be illuminated semiqualitatively for generic piecewise C^2 nonnegative $U(\chi^2)$ with $U'(\chi^2)$ nonincreasing for $\chi^2 \leq \alpha^2$, not identically constant for $\chi^2 \leq \alpha^2$, and such that $[U'(0) - U'(\chi^2)]$ is greater than $[U'(0) - U'(\alpha^2)](\chi/\alpha)^2$ for some $\chi^2 \leq \alpha^2$. First we evoke the approximate general spherically symmetric particlelike solution to Eq. (21)[8],

$$\chi_{(0)} \cong \frac{\alpha \tanh \mu |\mathbf{x}|}{\mu |\mathbf{x}| \cosh \mu |\mathbf{x}|}, \tag{28}$$

where $\mu^2 \equiv \frac{1}{6}[U'(0) - U'(\alpha^2)]$ and

$$\omega^2 = \tfrac{1}{6}[5U'(0) + U'(\alpha^2)], \tag{29}$$

the latter condition fixing the real constant α for a prescribed admissible value of ω. Next, we use Eq. (28) to obtain the singularity-free solution to (22) in terms of a quadrature,

$$\phi_{(1)} \cong \frac{2\omega\alpha^2}{3\mu^3} \int_{|\mathbf{x}|}^{\infty} (\tanh \mu r)^3 \frac{dr}{r^2}. \tag{30}$$

Finally, by putting Eqs. (28) and (30) into Eq. (23), we approximate the singularity-free solution as

$$\chi_{(2)} \cong -\frac{2\omega^2\alpha^3 \sinh^{-1} \nu |\mathbf{x}|}{3\mu^5 |\mathbf{x}| \cosh \mu |\mathbf{x}|}, \tag{31}$$

where ν is the real positive root of the equation

$$\nu^3 + [4\mu^2 + 2\alpha^2 U''(\alpha^2)]\nu - 2A\mu^3 = 0,$$

the numerical constant $A \equiv \int_0^\infty (\tanh \lambda)^3 \lambda^{-2}\, d\lambda$. Applied to Eq. (30), the charge consistency condition (25) produces the approximate relation

$$\omega = 3\mu^3/8\pi\alpha^2, \tag{32}$$

which in combination with Eq. (29) yields a pair (or pairs) of discrete values for α and ω. Thus, no undetermined constant of integration appears in a particlelike solution satisfying the charge consistency condition. It is evident that a field theory based on the Lagrangian density (1) with suitably prescribed nonlinear self-interaction leads to an unambiguous structural model on the classical level for a charged elementary particle of finite positive energy.

APPENDIX A

No claim of analyticity in ϵ about $\epsilon = 0$ for χ and ϕ could be made if $U(\chi^2)$ were to exhibit a general analytic dependency on ϵ. As an academic example of a nonnegative ϵ-dependent self-interaction energy density which is analytic in ϵ about $\epsilon = 0$ and admits singularity-free particlelike solutions which are *not* analytic in ϵ about $\epsilon = 0$, we mention the form

$$U(\chi^2) = \begin{cases} 4m_0^2[\tfrac{1}{3}|\chi|^{\frac{3}{2}} - \tfrac{1}{2}|\chi| + |\chi|^{\frac{1}{2}} - \ln(1 + |\chi|^{\frac{1}{2}})] \\ \qquad + \dfrac{32\epsilon^2}{9}(|\chi|^3 - 3\chi^4) \quad \text{for } |\chi| \leq 1, \\ \text{piecewise } C^2 \text{ and positive for } |\chi| \geq 1, \end{cases}$$

where m_0 is a positive constant and a convenient physical length unit is employed, the positivity of such a $U(\chi^2)$ effected by requiring

$$m_0^2 > \tfrac{32}{3}(5 - 6\ln 2)^{-1}\epsilon^2.$$

With

$$U'(\chi^2) = m_0^2(1 + |\chi|^{\frac{1}{2}})^{-1} + \frac{16\epsilon^2}{3}(|\chi| - 4\chi^2)$$

for $|\chi| \leq 1$, the spherically symmetric particlelike solution to the coupled nonlinear equations (6) and (7) is obtainable in closed form for the case $\omega = m_0$ and expressed exactly by

$$\chi = (1 + \xi)^{-1} \quad \left[\xi \equiv \frac{8\epsilon^2}{3}|\mathbf{x}|^2\right],$$

$$\phi = \frac{m_0}{\epsilon}\{1 - [1 + \xi^{-1} - (\xi^{-1} + \xi^{-2})^{\frac{1}{2}}]^{\frac{1}{2}}\}.$$

The latter solution function is not analytic in ϵ about $\epsilon = 0$, for it gives

$$\lim_{\epsilon \to 0} [\epsilon \phi(\mathbf{x}; \epsilon)] = (1 - 1/\sqrt{2}) m_0 .$$

APPENDIX B

Certain nonnegative $U(\chi^2)$ admit a particlelike solution of finite total energy for the critical case $\omega^2 = U'(0)$, but generally such a particlelike solution is not stable with respect to the perturbation method described here, in the sense that it does not admit a perturbed particlelike solution of finite total energy. As an example, consider a self-interaction energy density of the form

$$U(\chi^2) = \begin{cases} m_0^2 \chi^2 + g |\chi|^3 - \lambda \chi^4 & \text{for } |\chi| \leq 3g/\lambda, \\ \text{piecewise } C^2 \text{ and positive} & \text{for } |\chi| \geq 3g/\lambda, \end{cases}$$

where m_0 ($> \sqrt{6}\, g\lambda^{-\frac{1}{2}}$), g, and λ are positive constants. By substituting the derivative of this self-interaction density into Eq. (21) and taking $\omega^2 = m_0^2$, we have

$$\nabla^2 \chi_{(0)} - \tfrac{3}{2} g |\chi_{(0)}| \chi_{(0)} + 2\lambda \chi_{(0)}^3 = 0$$

for $|\chi_{(0)}| \leq 3g/\lambda$. The spherically symmetric particle-like solution to the latter equation and the associated singularity-free solution to Eq. (22) are given exactly by

$$\chi_{(0)} = \left(\tfrac{3}{4} g |\mathbf{x}|^2 + \frac{\lambda}{3g} \right)^{-1},$$

$$\phi_{(1)} = \pm \left(\frac{8 m_0}{3 g \lambda^{\frac{1}{2}}} \right) |\mathbf{x}|^{-1} \tan^{-1} (\tfrac{3}{2} g \lambda^{-\frac{1}{2}} |\mathbf{x}|),$$

where the sign is that of $\omega = \pm m_0$. By putting the latter expressions into Eq. (23), it follows that $\chi_{(2)}$ is asymptotic to $\mp 8\pi m_0^2 / 9\lambda^{\frac{1}{2}} g^2 |\mathbf{x}|$ as $|\mathbf{x}| \to \infty$, and hence the energy term

$$\int \omega^2 \chi^2 d^3\mathbf{x} \cong \int \omega^2 (\chi_{(0)}^2 + 2\epsilon^2 \chi_{(0)} \chi_{(2)}) d^3\mathbf{x}$$

and the total perturbed energy Eq. (27) are logarithmically divergent quantities.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 7 JULY 1968

Variational Bounds for the Potential in Terms of the *S*-Wave Phase Shift

F. CALOGERO
Istituto di Fisica dell'Università di Roma, Istituto Nazionale di Fisica Nucleare, Sezione di Roma, Rome, Italy

AND

O. D. CORBELLA,* A. DEGASPERIS,† AND M. B. DE STEFANO†
Istituto di Fisica dell'Università di Roma, Rome, Italy

(Received 15 July 1967)

We establish a variational expression involving the *S*-wave phase shift and bound-state parameters and the potential, which appears particularly suited to evince from the *S*-wave-scattering data information (i.e., bounds or variational approximations) on the potential. An application of this approach to solvable potentials yields a number of inequalities involving special functions.

1. INTRODUCTION

This paper is the second one of a series devoted to the derivation of explicit relationships between the parameters which describe a scattering process and the interaction which causes it. In the first paper we derived expressions for the values of the potential and of its derivative at the origin in terms of the *S*-wave phase shift and bound-state parameters.[1] In this paper we establish a variational expression involving the *S*-wave phase shift and bound state parameters, the potential, and a trial function. The simplest choice of this function yields the simple lower bound on the Laplace transform of the potential

$$\int_0^\infty dr\, V(r) e^{-2pr} > \frac{-1}{\pi} \int_{-\infty}^{+\infty} dk\, \frac{k^2}{k^2 + p^2} [1 - |f(k)|^{-2}], \qquad p \geq 0, \quad (1.1)$$

where the "spectral function" $|f(k)|^{-2}$ is given, in terms of the *S*-wave phase shift $\eta(k)$, by the explicit formula

$$|f(k)|^{-2} = \exp \left[\frac{4}{\pi} P \int_0^\infty dq\, \frac{q\eta(q)}{q^2 - k^2} \right]. \quad (1.2)$$

* Supported by a fellowship of the Consejo Nacional de Investigaciones Científicas y Técnicas, Argentina; on leave from Instituto de Matemática, Astronomía y Física, Córdoba, Argentina.
† Supported by a CNR fellowship for 1967.
[1] F. Calogero and A. Degasperis, J. Math. Phys. **9**, 90 (1968); hereafter referred to as I.

The latter solution function is not analytic in ϵ about $\epsilon = 0$, for it gives

$$\lim_{\epsilon \to 0} [\epsilon \phi(\mathbf{x}; \epsilon)] = (1 - 1/\sqrt{2}) m_0 .$$

APPENDIX B

Certain nonnegative $U(\chi^2)$ admit a particlelike solution of finite total energy for the critical case $\omega^2 = U'(0)$, but generally such a particlelike solution is not stable with respect to the perturbation method described here, in the sense that it does not admit a perturbed particlelike solution of finite total energy. As an example, consider a self-interaction energy density of the form

$$U(\chi^2) = \begin{cases} m_0^2 \chi^2 + g |\chi|^3 - \lambda \chi^4 & \text{for } |\chi| \leq 3g/\lambda, \\ \text{piecewise } C^2 \text{ and positive} & \text{for } |\chi| \geq 3g/\lambda, \end{cases}$$

where m_0 ($> \sqrt{6}\, g\lambda^{-\frac{1}{2}}$), g, and λ are positive constants. By substituting the derivative of this self-interaction density into Eq. (21) and taking $\omega^2 = m_0^2$, we have

$$\nabla^2 \chi_{(0)} - \tfrac{3}{2} g |\chi_{(0)}| \chi_{(0)} + 2\lambda \chi_{(0)}^3 = 0$$

for $|\chi_{(0)}| \leq 3g/\lambda$. The spherically symmetric particle-like solution to the latter equation and the associated singularity-free solution to Eq. (22) are given exactly by

$$\chi_{(0)} = \left(\tfrac{3}{4} g |\mathbf{x}|^2 + \frac{\lambda}{3g} \right)^{-1},$$

$$\phi_{(1)} = \pm \left(\frac{8 m_0}{3 g \lambda^{\frac{1}{2}}} \right) |\mathbf{x}|^{-1} \tan^{-1} (\tfrac{3}{2} g \lambda^{-\frac{1}{2}} |\mathbf{x}|),$$

where the sign is that of $\omega = \pm m_0$. By putting the latter expressions into Eq. (23), it follows that $\chi_{(2)}$ is asymptotic to $\mp 8\pi m_0^2 / 9\lambda^{\frac{1}{2}} g^2 |\mathbf{x}|$ as $|\mathbf{x}| \to \infty$, and hence the energy term

$$\int \omega^2 \chi^2 d^3\mathbf{x} \cong \int \omega^2 (\chi_{(0)}^2 + 2\epsilon^2 \chi_{(0)} \chi_{(2)}) d^3\mathbf{x}$$

and the total perturbed energy Eq. (27) are logarithmically divergent quantities.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 7 JULY 1968

Variational Bounds for the Potential in Terms of the *S*-Wave Phase Shift

F. CALOGERO
Istituto di Fisica dell'Università di Roma, Istituto Nazionale di Fisica Nucleare, Sezione di Roma, Rome, Italy

AND

O. D. CORBELLA,* A. DEGASPERIS,† AND M. B. DE STEFANO†
Istituto di Fisica dell'Università di Roma, Rome, Italy

(Received 15 July 1967)

We establish a variational expression involving the *S*-wave phase shift and bound-state parameters and the potential, which appears particularly suited to evince from the *S*-wave-scattering data information (i.e., bounds or variational approximations) on the potential. An application of this approach to solvable potentials yields a number of inequalities involving special functions.

1. INTRODUCTION

This paper is the second one of a series devoted to the derivation of explicit relationships between the parameters which describe a scattering process and the interaction which causes it. In the first paper we derived expressions for the values of the potential and of its derivative at the origin in terms of the *S*-wave phase shift and bound-state parameters.[1] In this paper we establish a variational expression involving the *S*-wave phase shift and bound state parameters, the potential, and a trial function. The simplest choice of this function yields the simple lower bound on the Laplace transform of the potential

$$\int_0^\infty dr\, V(r) e^{-2pr} > \frac{-1}{\pi} \int_{-\infty}^{+\infty} dk\, \frac{k^2}{k^2 + p^2} [1 - |f(k)|^{-2}], \qquad p \geq 0, \quad (1.1)$$

where the "spectral function" $|f(k)|^{-2}$ is given, in terms of the *S*-wave phase shift $\eta(k)$, by the explicit formula

$$|f(k)|^{-2} = \exp \left[\frac{4}{\pi} P \int_0^\infty dq\, \frac{q\eta(q)}{q^2 - k^2} \right]. \quad (1.2)$$

* Supported by a fellowship of the Consejo Nacional de Investigaciones Científicas y Técnicas, Argentina; on leave from Instituto de Matemática, Astronomía y Física, Córdoba, Argentina.
† Supported by a CNR fellowship for 1967.
[1] F. Calogero and A. Degasperis, J. Math. Phys. **9**, 90 (1968); hereafter referred to as I.

These equations apply to the case without bound states; their generalization to the case with bound states is given below. As in I, in this paper the consideration is restricted to holomorphic potentials.

The derivation of the variational formula is given in Sec. 2; it is based on the expression, first given by Bellman,[2] of the solution of a differential equation (belonging to a certain class) through a maximum principle. This same mathematical trick had been previously used to obtain variational expressions and bounds for the phase shift in terms of the potential.[3]

It should be emphasized that the variational and maximum principles for the phase shift just mentioned, as any other variational formula for the scattering phase shift,[4] also qualify as exact variational expressions involving the measured scattering parameters, the potential, and a trial function. However, the formulas which we derive below appear much more appropriate than those, for the purpose of deriving information on the potential from the knowledge of the phase shift. In fact, the expression we derive comes somewhat closer to the goal of writing a variational expression suitable for the direct transfer of information from the scattering phase shift to the interaction, than any previously obtained.[5]

In Sec. 3 we apply the results to a few cases of exactly solvable potentials, thereby obtaining nontrivial variational expressions and bounds for certain special functions of mathematical physics.

Although throughout we draw on the conclusions and the notation of I, we have reported in this paper all the results we use, so as to make the presentation self-contained. Units are such that $\hbar = 2m = 1$.

2. DERIVATION OF THE VARIATIONAL EXPRESSION

Let $f(k, r)$ be the Jost solution of the S-wave radial Schrödinger equation

$$f''(k, r) + [k^2 - V(r)]f(k, r) = 0, \qquad (2.1)$$

namely, that solution of this equation characterized by the asymptotic boundary condition

$$\lim_{r\to\infty} [e^{ikr}f(k, r)] = 1. \qquad (2.2)$$

The potential $V(r)$ is by assumption holomorphic for real nonnegative r, and vanishes asymptotically faster than the first inverse power of r.

Define the function $g(k, r)$ through

$$g(k, r) = ik + f'(k, r)/f(k, r), \qquad (2.3)$$

which implies

$$g'(k, r) = V(r) + 2ikg(k, r) - g^2(k, r) \qquad (2.4)$$

and

$$g(k, \infty) = 0, \qquad (2.5)$$

and set

$$g(k) \equiv g(k, 0). \qquad (2.6)$$

We know from I that the function

$$\tilde{g}(k) = g(k) + \sum_n C_n(k^2 - E_n)^{-1} \qquad (2.7)$$

is holomorphic in the lower half k plane, and that it is given there and on the real axis by the formula[6]

$$\tilde{g}(k) = \frac{-1}{\pi}\int_{-\infty}^{+\infty} dq \frac{q}{q - k + i\epsilon}[1 - |f(q)|^{-2}]. \qquad (2.8)$$

The parameters C_n and E_n are, respectively, the (positive) normalization coefficients[7] and negative binding energies of the S-wave bound states associated with the potential $V(r)$; the "spectral function" $|f(k)|^{-2}$ is given, in terms of the S-wave phase shift $\eta(k)$ and the binding energies E_n, by the formula

$$|f(k)|^{-2} = \left[\prod_n (1 - k^{-2}E_n)^{-2}\right] \times \exp\left[\frac{4}{\pi} P\int_0^{\infty} dq \frac{q}{q^2 - k^2}\eta(q)\right]. \qquad (2.9)$$

In particular, for

$$k = -ip, \qquad p > 0, \qquad (2.10)$$

we get from Eq. (2.8)

$$\tilde{g}(-ip) = \frac{-1}{\pi}\int_{-\infty}^{+\infty} dq \frac{q^2}{q^2 + p^2}[1 - |f(q)|^{-2}]. \qquad (2.11)$$

To obtain this equation we have used the fact that

[2] R. E. Bellman, Proc. Natl. Acad. Sci. U.S. **41**, 743 (1955); R. E. Kalaba, J. Math. Mech. **8**, 519 (1959); F. Calogero, J. Math. Phys. **4**, 427 (1963); R. E. Bellman and R. E. Kalaba, *Quasilinearization and Nonlinear Boundary-Value Problems* (Elsevier Publ. Co., New York, 1965).

[3] F. Calogero, Nuovo Cimento **27**, 261 (1963); **27**, 947 (1963); **28**, 320 (1963); *Variable Phase Approach to Potential Scattering* (Academic Press Inc., New York, 1967), Chaps. 9 and 14 and Appendix II.

[4] Such variational expressions, originally introduced by L. Hulthén and J. Schwinger, are now discussed in most textbooks on scattering theory. See, for instance, T. Y. Wu and T. Ohmura, *Quantum Theory of Scattering* (Prentice-Hall, Inc., Englewood Cliffs, N. J., 1962); M. L. Goldberger and K. M. Watson, *Collision Theory* (John Wiley & Sons, Inc., New York, 1964); R. G. Newton, *Scattering Theory of Waves and Particles*, McGraw-Hill Book Co., New York, 1965).

[5] T. Ohmura, Progr. Theoret. Phys. (Kyoto) **16**, 231 (1956).

[6] This formula follows trivially from Eqs. (2.22)–(2.24) of I.

[7] The bound-state normalization coefficient C_n is defined by

$$C_n^{-1} = \int_0^{\infty} dr\, \varphi^2(E_n, r),$$

where $\varphi(E_n, r)$ is the radial wavefunction of the bound state with energy E_n, normalized by the condition at the origin $\varphi'(E_n, 0) = 1$.

$|f(q)|$ is an even function of q. Note that this equation implies that $\tilde{g}(-ip)$, and therefore also $g(-ip)$, is real.

In fact, setting $k = -ip$ in Eq. (2.4), we obtain

$$g'(-ip, r) = V(r) + 2pg(-ip, r) - g^2(-ip, r), \tag{2.12}$$

which, together with Eq. (2.5), implies the reality of $g(-ip, r)$. On the other hand it is easily seen that $g(-ip, r)$ is a continuous function of r on the positive real axis, provided p is so large that $-p^2$ is smaller than the (negative) energies of all S-wave bound states,

$$p^2 > |E_n|. \tag{2.13}$$

In fact, the singularities of $g(-ip, r)$ might arise only from the zeros of the function $f(-ip, r)$, which is by definition that solution of the radial Schrödinger equation behaving asymptotically as $\exp(-pr)$, and therefore it has no zeros on the positive real axis if p^2 is large, the first zero occurring, as p^2 is decreased, at $r = 0$ just at the value $p^2 = -E_1$, where E_1 is the energy of the lowest S-wave bound state. Of course if there are no S-wave bound states, $g(-ip, r)$ is a continuous function of r, on the positive real axis, for all positive values of p.

We are therefore entitled to solve the differential equation (2.12) by Bellmann's maximum principle,[2] writing

$$g(-ip, r) = \max_w \left\{ -\int_r^\infty ds[V(s) + w^2(p, s)] \times \exp\left[2\int_s^r dt[p - w(p, t)]\right]\right\}, \tag{2.14}$$

which implies

$$g(-ip) \geq -\int_0^\infty dr[V(r) + w^2(p, r)] \times \exp\left[-2pr + 2\int_0^r ds\, w(p, s)\right], \tag{2.15a}$$

or, equivalently

$$\int_0^\infty dr V(r) \exp\left[-2pr + 2\int_0^r ds\, w(p, s)\right] \geq -g(-ip) - \int_0^\infty dr\, w^2(p, r) \times \exp\left[-2pr + 2\int_0^r ds\, w(p, s)\right]. \tag{2.15b}$$

The equality sign in these equations corresponds to the optimal choice for the trial function

$$w(p, r) = w_{\text{opt}}(p, r) \equiv g(-ip, r). \tag{2.16}$$

From these equations and Eqs. (2.11) and (2.7) we obtain our final result:

$$\int_0^\infty dr\, V(r) \exp\left[-2pr + 2\int_0^r ds\, w(p, s)\right] \geq \frac{1}{\pi}\int_{-\infty}^{+\infty} dq\, q^2(q^2 + p^2)^{-1}[1 - |f(q)|^{-2}] - \sum_n C_n(p^2 + E_n)^{-1} - \int_0^\infty dr\, w^2(p, r) \times \exp\left[-2pr + 2\int_0^r ds\, w(p, s)\right], \tag{2.17}$$

with $|f(q)|^{-2}$ defined by Eq. (2.9). This inequality holds for all the nonnegative values of p consistent with the inequality (2.13); of course, if there are bound states, as $-p^2$ approaches the energy of the lowest bound state, the inequality becomes less and less stringent. But in any case whenever the trial function, which must always be a continuous function of r, happens to coincide with the optimal-choice equation (2.16), then the equality sign prevails in these equations.

It is in fact easily seen that, substituting in this equation in place of $w(p, s)$ the optimal-choice equation (2.16), and performing an asymptotic expansion in p, one re-obtains the results of I.

The choice

$$w(p, r) = w(p) < p \tag{2.18}$$

yields the explicit lower bound for the Laplace transform of the potential

$$\int_0^\infty dr\, V(r) e^{-2[p - w(p)]r} > \frac{1}{\pi}\int_{-\infty}^{+\infty} dq\, q^2(q^2 + p^2)^{-1}[1 - |f(q)|^{-2}] - \sum_n C_n(p^2 + E_n)^{-1} - \tfrac{1}{2}w^2(p)(p - w(p))^{-1}. \tag{2.19}$$

The special choice $w(p) = 0$ yields, in the case without bound states, Eq. (1.1).

3. EXAMPLES

In this section we report some results which obtain from a straightforward application of the conclusions of the previous section to potentials such that the corresponding radial Schrödinger equation may be solved in terms of standard special functions,[8] so that the function $g(-ip)$ may be evaluated in closed form. In this manner we obtain, from Eq. (2.15a), nontrivial expressions and bounds involving special functions. A more systematic exploitation of this technique to obtain results for special functions is, however, beyond the scope of the research reported in this paper.

[8] A. K. Bose, Phys. Letters 7, 245 (1963).

To the potential

$$V(r) = V_1 e^{-\mu r} + V_2 e^{-2\mu r} \tag{3.1}$$

there corresponds the Jost solution

$$f(k, r) = \exp\left[-ikr - V_2^{\frac{1}{2}}\mu^{-1}e^{-\mu r}\right]\Phi(\tfrac{1}{2}c + \lambda, c; ze^{-\mu r}), \tag{3.2}$$

where Φ is the confluent hypergeometric function[9] and

$$\lambda = \tfrac{1}{2}V_1 V_2^{-\frac{1}{2}}\mu^{-1}, \tag{3.3a}$$

$$c = 1 + 2ik\mu^{-1}, \tag{3.3b}$$

$$z = 2V_2^{\frac{1}{2}}\mu^{-1}. \tag{3.3c}$$

A sufficient condition to exclude the presence of bound states is the Bargmann rule[10]

$$\int_0^\infty dr\, rV(r) \geq -1, \tag{3.4}$$

which becomes in this case

$$V_1 + \tfrac{1}{4}V_2 \geq -\mu^2. \tag{3.5}$$

A straightforward application of the results of the previous section yields the formula

$$\frac{z\Phi'(\frac{1}{2}c + \lambda, c; z)}{\Phi(\frac{1}{2}c + \lambda, c; z)} - \frac{z}{2} \leq \int_0^\infty dx[\lambda z e^{-cx} + \tfrac{1}{4}z^2 e^{-(c+1)x} + w^2(x)e^{-(c-1)x}] \times \exp\left[2\int_0^x dy\, w(y)\right]. \tag{3.6}$$

Sufficient conditions for the validity of this formula are

$$c \geq 1 \tag{3.7a}$$

and

$$z\lambda + \tfrac{1}{16}z^2 \geq -1. \tag{3.7b}$$

While c must be real, z and λ may be either both real or both imaginary.[11] The first condition arises, through Eq. (3.3b), from the requirement that p be positive, Eq. (2.10); the second condition, through Eqs. (3.3a) and (3.3c), from the requirement that no bound states be present, Eq. (3.5). Of course these are not the most stringent conditions for the validity of Eq. (3.6). The equality sign in this equation applies provided

$$w(x) = \frac{-ze^{-x}\Phi'[\frac{1}{2}c + \lambda, c; ze^{-x}]}{\Phi[\frac{1}{2}c + \lambda, c; ze^{-x}]} + \tfrac{1}{2}ze^{-x}, \tag{3.8}$$

and from this formula one may also conclude that a less stringent condition for the validity of Eq. (3.6) (besides the requirement that c be real and λ and z be either both real or both imaginary), is that the values of c, λ, and z be such that the function $\Phi[\frac{1}{2}c + \lambda, c; zy]$ does not vanish for $0 < y \leq 1$. On the other hand the choice

$$w(x) = w < \tfrac{1}{2}(c - 1)$$

yields the simple bound

$$\frac{z\Phi'(\frac{1}{2}c + \lambda, c; z)}{\Phi(\frac{1}{2}c + \lambda, c; z)} - \frac{z}{2} \leq \lambda z(c - 2w)^{-1} + \tfrac{1}{4}z^2(c + 1 - 2w)^{-1} + w^2(c - 1 - 2w)^{-1}. \tag{3.9}$$

Choosing for w a convenient dependence on z and integrating over z, one may obtain from this equation an upper bound for the confluent hypergeometric function itself, provided of course the validity of the two conditions (3.7) is maintained throughout the integration.

The particularly simple choice $w = 0$, which corresponds to the bound of Eq. (1.1), yields

$$\frac{z\Phi'(\frac{1}{2}c + \lambda, c; z)}{\Phi(\frac{1}{2}c + \lambda, c; z)} - \frac{z}{2} \leq \frac{\lambda z}{c} + \frac{1}{4}\frac{z^2}{c + 1}. \tag{3.10}$$

A check of the stringency of this inequality is particularly easy in the case

$$\lambda = \tfrac{1}{2}c, \tag{3.11}$$

because then the confluent hypergeometric function becomes an exponential, and Eq. (3.10) becomes simply

$$\tfrac{1}{4}z^2(c + 1)^{-1} > 0, \tag{3.12}$$

whose validity is implied by Eq. (3.7a) and the fact that z must be real, since λ is real.

In the special case $\lambda = 0$ the confluent hypergeometric function reduces to a Bessel function and Eq. (3.6) may be written in the form

$$\frac{+\frac{1}{2}izJ_{\nu+1}(-\frac{1}{2}iz)}{J_\nu(-\frac{1}{2}iz)} \leq \int_0^\infty dx[\tfrac{1}{4}z^2 e^{-2(\nu+1)x} + w^2(x)e^{-2\nu x}] \times \exp\left[2\int_0^x dy w(y)\right], \tag{3.13}$$

sufficient conditions for the validity of this inequality being that ν be real and nonnegative and that z be either pure imaginary with $|z| < 4$ or be real.

A more general example is provided by the potential

$$V(r) = V_1(1 - ge^{\mu r})^{-1} + V_2(1 - ge^{\mu r})^{-2}, \tag{3.14}$$

which depends on three (dimensionless) parameters

[9] For all special functions we adopt the notation of *Higher Transcendental Functions*, A. Erdélyi, Ed. (McGraw-Hill Book Co., New York, 1953).

[10] V. Bargmann, Proc. Natl. Acad. Sci. U.S. **38**, 961 (1952).

[11] In fact, it is easily seen that the left-hand side of Eq. (3.6) is real even if z and λ are imaginary, since by the Kummer transformation it may be written in the form

$$z\left[\frac{\Phi'(\frac{1}{2}c + \lambda, c; z)}{\Phi(\frac{1}{2}c + \lambda, c; z)} - \frac{\Phi'(\frac{1}{2}c - \lambda, c; -z)}{\Phi(\frac{1}{2}c - \lambda, c; -z)}\right].$$

rather than two. We exclude in the following the values

$$0 \leq g \leq 1, \tag{3.15}$$

to guarantee its regularity. A sufficient condition to exclude that this potential possesses bound states is

$$0 > gV_1 \geq V_1 + V_2; \tag{3.16}$$

in fact this condition guarantees that the potential be everywhere repulsive (positive).

To this potential there corresponds the Jost solution

$$f(k, r) = e^{-ikr}(1 - g^{-1}e^{-\mu r})^{\rho+\frac{1}{2}} \times F(\tfrac{1}{2}c + \rho + \sigma, \tfrac{1}{2}c + \rho - \sigma, c; g^{-1}e^{-\mu r}), \tag{3.17}$$

where

$$\rho = \tfrac{1}{2}(1 + 4V_2\mu^{-2})^{\frac{1}{2}}, \tag{3.18a}$$

$$\sigma = \mu^{-1}(V_1 + V_2 - k^2)^{\frac{1}{2}}, \tag{3.18b}$$

$$c = 1 + 2ik\mu^{-1}, \tag{3.18c}$$

and F is the hypergeometric function.[9]

A straightforward application of the technique of Sec. 2 yields the following inequality:

$$\frac{zF'(\frac{1}{2}c + \rho + \sigma, \frac{1}{2}c + \rho - \sigma; c; z)}{F(\frac{1}{2}c + \rho + \sigma, \frac{1}{2}c + \rho - \sigma, c; z)} \leq \tfrac{1}{2}(2\rho + 1)z(1 - z)^{-1} + \int_1^\infty dx\{[\tfrac{1}{2}c(1 - \tfrac{1}{2}c) + \sigma^2 - \rho^2]z(z - x)^{-1} + (\rho - \tfrac{1}{2})(\rho + \tfrac{1}{2})z^2(z - x)^{-2} + w^2(x)\}x^{-c} \times \exp\left[2\int_1^x dy y^{-1}w(y)\right]. \tag{3.19}$$

In this inequality, and always in the following, c and z are real, while ρ and σ may be either real or imaginary. Conditions sufficient for the validity of Eq. (3.19) are

$$c \geq 1, \tag{3.20a}$$

$$z \leq 1, \tag{3.20b}$$

$$z^{-1}[\tfrac{1}{2}c(1 - \tfrac{1}{2}c) + \sigma^2 - \rho^2] < 0, \tag{3.20c}$$

$$\sigma^2 - \tfrac{1}{4}(c - 1)^2 \geq z^{-1}[\tfrac{1}{2}c(1 - \tfrac{1}{2}c) + \sigma^2 - \rho^2]. \tag{3.20d}$$

The equality sign prevails in Eq. (3.19) provided

$$w(x) = w_{\text{opt}}(x) \equiv \tfrac{1}{2}(2\rho + 1)z(x - z)^{-1} - (z/x)F'(\tfrac{1}{2}c + \rho + \sigma, \tfrac{1}{2}c + \rho - \sigma, c; z/x)/F(\tfrac{1}{2}c + \rho + \sigma, \tfrac{1}{2}c + \rho - \sigma, c; z/x). \tag{3.21}$$

From this equation one may infer that the less stringent requirement for the validity of Eq. (3.19), which replaces Eqs. (3.20c, d), is the condition that ρ and σ be such that $F(\frac{1}{2}c + \rho + \sigma, \frac{1}{2}c + \rho - \sigma, c; y)$ has no zeros for $y \leq 1$.

To obtain more explicit bounds we have inserted in Eq. (3.19) the two trial functions

$$w(x) = w < \tfrac{1}{2}(c - 1) \tag{3.22a}$$

and

$$w(x) = wz(x - z)^{-1}, \tag{3.22b}$$

obtaining, respectively,

$$zF'(\tfrac{1}{2}c + \rho + \sigma, \tfrac{1}{2}c + \rho - \sigma, c; z)/F(\tfrac{1}{2}c + \rho + \sigma, \tfrac{1}{2}c + \rho - \sigma, c; z) \leq \tfrac{1}{4}(2\rho + 1)^2 z(1 - z)^{-1} + \tfrac{1}{4}(c - 1 - \beta)^2 - z\{(\beta + 1)^{-1}[\tfrac{1}{2}c(1 - \tfrac{1}{2}c) + \sigma^2 - \rho^2] + (\rho - \tfrac{1}{2})(\rho + \tfrac{1}{2})\}F(1, \beta + 1, \beta + 2, z), \tag{3.23a}$$

with $\beta = c - 1 - 2w$ arbitrary except for the requirement

$$\beta > 0 \tag{3.24}$$

and

$$zF'(\tfrac{1}{2}c + \rho + \sigma, \tfrac{1}{2}c + \rho - \sigma, c; z)/F(\tfrac{1}{2}c + \rho + \sigma, \tfrac{1}{2}c + \rho - \sigma, c; z) \leq -z(1 - z)^{-1}[(2w - 1)^{-1}(\rho^2 + w^2 - \tfrac{1}{4}) - \tfrac{1}{2}(2\rho + 1)] + z(1 - z)^{-2w}\{[(\tfrac{1}{2}c + \rho)^2 - \sigma^2]c^{-1} - \tfrac{1}{2}(2\rho + 1) + (2w - 1)^{-1}(\rho^2 + w^2 - \tfrac{1}{4})\}F(1 - 2w, c, c + 1; z), \tag{3.23b}$$

where w is arbitrary (of course real).

This last formula takes a particularly simple form if in the right-hand side we set

$$w = -\tfrac{1}{2}c, \tag{3.25}$$

because one may then use the simple formula[9]

$$F(c, c + 1, c + 1; z) = (1 - z)^{-c} \tag{3.26}$$

to obtain

$$zF'(\tfrac{1}{2}c + \rho + \sigma, \tfrac{1}{2}c + \rho - \sigma, c; z)/F(\tfrac{1}{2}c + \rho + \sigma, \tfrac{1}{2}c + \rho - \sigma, c; z) \leq zc^{-1}[(\tfrac{1}{2}c + \rho)^2 - \sigma^2] + z^2(1 - z)^{-1} \times \{[\rho^2 - \tfrac{1}{4} + \tfrac{1}{4}c^2](c + 1)^{-1} + \tfrac{1}{2}(2\rho + 1)\}. \tag{3.27}$$

A test of the accuracy of this bound may be made in the case

$$\rho - \sigma = \tfrac{1}{2}c, \tag{3.28}$$

when it becomes

$$z^2(1 - z)^{-1}(c + 1)^{-1}(\sigma - \tfrac{1}{2})^2 \geq 0, \tag{3.29}$$

which is of course consistent with the conditions (3.20a) and (3.20b).

ACKNOWLEDGMENT

One of us (O. D. C.) wishes to thank the Istituto di Fisica "G. Marconi" dell'Università di Roma for its hospitality.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 7 JULY 1968

Weak-Graph Method for Obtaining Formal Series Expansions for Lattice Statistical Problems

JOHN F. NAGLE*
Department of Chemistry, Cornell University, Ithaca, New York

(Received 26 April 1967)

A unified exposition of the weak-graph method for obtaining formal series expansions for lattice statistical problems is presented. The prototype of this method is the derivation of the hyperbolic-tangent high-temperature expansion for the spin-$\frac{1}{2}$ Ising model. Also, recent expansions of the monomer–dimer problem and various hydrogen-bonded problems have been treated by essentially the same method. In this paper the method is further illustrated by obtaining series expansions for the low-temperature spin-$\frac{1}{2}$ Ising problem, the low-density hard-core lattice-gas problem, the high-temperature spin-1 Ising problem, the k-color problem, and two new model problems, the ramrod model and a special ternary model. The weak-graph method enables one to obtain especially useful series expansions for a certain class of problems, including the spin-$\frac{1}{2}$ Ising problem and the monomer–dimer problem, which have essentially a binary nature.

1. GENERAL INTRODUCTION

A. Background

Formal series expansions have been very useful in statistical mechanics.[1,2] Occasionally, they have revealed and clarified underlying principles, such as the duality principle of the two-dimensional spin-$\frac{1}{2}$ Ising model.[3] More important, they have provided systematic ways of obtaining the coefficients for the beginning portion of exact series for various physical quantities. These series are then extrapolated by one means or another to give estimates for quantities of interest.[4,5]

The method of deriving the formal series expansions for continuum systems has been quite well developed and formalized by Ursell, Mayer, Uhlenbeck, and others.[6,7] Of course, any continuum method may also be applied to lattice systems as particular cases.[8–10] However, by taking into account the special features of lattice problems, it has often been possible to obtain formal series expansions which are less complicated than those obtained by the continuum method.[1,11]

Various ways have been found for obtaining formal series expansions for various lattice problems. Some of these ways may be conveniently grouped together and thought of as particular cases of one or other general method. One such method, which may be defined in a general way, will be called the direct method in this paper. A simple example of the direct method is the low-temperature spin-$\frac{1}{2}$ Ising expansion in powers of exp $(-J/kT)$ which will be described in Sec. 2.[12] The direct method is generally applicable when the first term in the expansion represents some ordered state and the higher terms represent simple perturbations from that ordered state. The procedure is to "look and see" how the perturbations may be represented graphically.

Another method is called the finite-cluster method.[1,13–17] This method is based on the fact that a formal series expansion for thermodynamic quantities may always be written involving only graphs which are connected or star graphs, depending on the series in question. The first task is to compute the contribution or weight of each (connected or star) graph. The finite cluster method solves this through solving the actual problem on small finite clusters, i.e., graphs. Each new finite graph solution gives an additional relation among different graph weights. Thus, the weights may be deduced recursively by starting with the smallest cluster or graph and building up to more complex graphs. (Of course, when this is finished, one still must count the number of ways in which each graph may occur as a subgraph of the underlying lattice.) This is a very general method for finding the beginning portion of an exact series

* Present address: Dept. of Phys., Carnegie-Mellon University, Pittsburgh, Pa. 15213.

[1] C. Domb, Advan. Phys. **9**, 149 (1960).
[2] M. E. Fisher, *Lectures in Theoretical Physics* (University of Colorado Press, Boulder, Colorado, 1965), Vol. VII C.
[3] G. F. Newell and E. W. Montroll, Rev. Mod. Phys. **25**, 353 (1953).
[4] C. Domb and M. F. Sykes, J. Math. Phys. **2**, 63 (1961).
[5] G. A. Baker, Jr., Advan. Theoret. Phys. **1**, 1 (1965).
[6] G. E. Uhlenbeck and G. W. Ford, Studies Statistical Mech. **1**, 119 (1962).
[7] J. E. Mayer and M. G. Mayer, *Statistical Mechanics* (John Wiley & Sons, Inc., New York, 1940).
[8] S. Katsura, Progr. Theoret. Phys. (Kyoto) **20**, 192 (1958).
[9] G. S. Rushbrooke, H. I. Scoins, and A. J. Wakefield, Discussions Faraday Soc. **15**, 57 (1953).
[10] G. S. Rushbrooke and H. I. Scoins, Proc. Roy. Soc. (London) **230**, 74 (1955).
[11] J. F. Nagle, Phys. Rev. **152**, 190 (1966).
[12] B. L. van der Waerden, Z. Physik **118**, 473 (1941).
[13] C. Domb and D. W. Wood, Proc. Phys. Soc. (London) **86**, 1 (1965).
[14] G. S. Rushbrooke, J. Math. Phys. **5**, 1106 (1964).
[15] G. A. Baker *et al.*, Phys. Letters **20**, 146 (1966).
[16] G. S. Joyce and R. G. Bowers, Proc. Phys. Soc. (London) **88**, 1053 (1966).
[17] J. W. Essam and M. F. Sykes, J. Math. Phys. **7**, 1573 (1966).

and is often all that is available for the more difficult problems.

The first purpose of this paper is to call attention to another method for obtaining formal series expansions which is especially useful for the simpler lattice problems. This method will be called the weak-graph method. The prototype of the weak-graph method is the very fruitful hyperbolic-tangent expansion for the spin-$\frac{1}{2}$ Ising model.[3] Other applications of the method have been to the monomer–dimer problem[11] and to various simple hydrogen-bonded lattice problems.[18]

Rather than give an unintelligible qualitative introduction to the weak-graph method at this point, let us proceed to review the familiar hyperbolic-tangent expansion in the next subsection. The formalization of the salient features of this expansion technique, which is accomplished in the last subsection of this extended introduction, provides a natural introduction to the weak-graph method.

B. The Hyperbolic-Tangent Expansion for the Spin-$\frac{1}{2}$ Ising Problem

The partition function for the spin-$\frac{1}{2}$ Ising problem may be written as

$$Z(K, L) = \sum_{\sigma} \prod_{\text{edges}} [\cosh K + \sigma_i\sigma_j \sinh K] \times \prod_{i=1}^{N} \exp(\sigma_i L), \quad (1.1)$$

where $K = J/kT$ and $L = Hm/kT$, q is the coordination number of the underlying lattice L_q which has N vertices, and $\sigma_i = \pm 1$ is the spin value of the ith vertex. The summation is over all 2^N combinations of spin values, the first product is over all edges ij in L_q, where vertices i and j are nearest neighbors and the second product is over all vertices in L_q. The multiple product in (1.1) is expanded into its $2^{qN/2}$ terms. Each term T may be represented by a subgraph $G \subseteq L_q$ according to the rule that the edge ij of L_q is contained in G if and only if the $(\sigma_i\sigma_j \sinh K)$ part of the edge factor $(\cosh K + \sigma_i\sigma_j \sinh K)$ was used in forming T.

Thus, the very form of (1.1) dictates that each subgraph G may be thought of as a collection of edges. In particular, isolated vertices, i.e., vertices of degree zero, are not allowed in G; otherwise no restrictions have been made on G. Such subgraphs are called *weak graphs*.[19]

The summation in (1.1) is performed for each term T. The result is called the weight w(G) of the subgraph G corresponding to T. Formally,

$$w(G) = \sum_{\sigma} (\cosh K)^{qN/2} \prod_{\text{edges in G}} \sigma_i\sigma_j \tanh K \times \prod_{i=1}^{N} \exp(\sigma_i L). \quad (1.2)$$

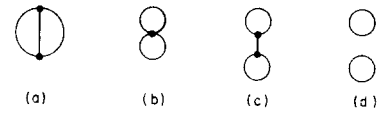


FIG. 1. Examples of closed graphs with no vertices of degree one in which vertices of degree two are suppressed: (a) Star graph; (b) articulated graph; (c) graph with a bridge; (d) unconnected graph.

Because each product $\sigma_i\sigma_j \tanh K$ is factorable into a factor depending only on spin σ_i and a factor depending only on spin σ_j, the cooperative summation $\sum_{\sigma}$ may be performed independently for each spin σ_i and (1.2) becomes

$$w(G) = (\cosh K)^{qN/2} \prod_{i=1}^{N} \sum_{\sigma_i=\pm 1} \sigma_i^{s_i} (\tanh K)^{s_i/2} \exp(\sigma_i L), \quad (1.3)$$

where s_i is the degree of vertex i in G, i.e., s_i is the number of edges in G incident to the ith vertex. Each summation in (1.3) may be designated as $w(i)$, which is easily seen to be

$$w(i) = (\tanh K)^{s_i/2}(2 \cosh L), \quad \text{if } s_i \text{ is an even integer,}$$
$$= (\tanh K)^{s_i/2}(2 \sinh L), \quad \text{if } s_i \text{ is an odd integer.}$$

Thus,

$$w(G) = (\cosh K)^{qN/2} \prod_{i=1}^{N} w(i) = (\cosh K)^{qN/2}(2 \cosh L)^N (\tanh K)^e \times (\tanh L)^{v(\text{odd})}, \quad (1.4)$$

where $e = \frac{1}{2}\sum_i s_i$ is the number of edges in G and $v(\text{odd})$ is the number of vertices in G with odd degrees, $s_i = 1, 3, 5, \cdots$.

It will be noticed that the above series expansion simplifies considerably in zero field ($L = 0$) because from (1.4) all graphs with any vertices of odd degree then have zero weight and drop out of the expansion. It has also been shown recently that the combinatorial information concerning the number of subgraphs of particular type in a lattice needed for the $L \neq 0$ expansion may be obtained by algebraic transformation from the subgraph combinatorial information for a relatively small subset of the weak graphs.[20] This subset is called the *closed weak graphs* and is restricted to weak graphs with no vertices of degree one (see Fig. 1). This algebraic transformation will be called

[18] J. F. Nagle, J. Math. Phys. 7, 1484, 1492 (1966).
[19] M. F. Sykes *et al.*, J. Math. Phys. 7, 1557 (1966).
[20] J. F. Nagle and H. N. V. Temperley, J. Math. Phys. 9, 1020 (1968), following paper.

the *weak-closed transformation* in the remainder of this paper. Many of the series expansions obtained by the weak-graph method have required only combinatorial information for the closed weak graphs, either directly or via the weak-closed transformation; this is a considerable simplification compared to needing the combinatorial information for all the weak graphs.

Before generalizing this Ising procedure, let us review briefly the reasons that this $\tanh K$ series expansion has been the best one for the spin-$\frac{1}{2}$ Ising-model problem. First, this expansion may be used to explain the duality principle in zero field for two-dimensional lattices.[3] Secondly, with the same amount of effort longer exact high-temperature series may be obtained using this expansion than by using other expansions.[1] (The length may be measured in powers of $1/kT$.) In particular, the direct method of obtaining series expansions is not applicable to the high-temperature disordered phase. The finite-cluster method actually requires less graph-configurational information than the $\tanh K$ expansion, which requires articulated, bridged, and separated closed subgraphs (see Fig. 1) as well as the stars or connected graphs. However, the finite-cluster method still requires more labor because the related weight problem is nontrivial, whereas the weight problem is completely solved for the $\tanh K$ expansion. It might also be mentioned that there are various $1/kT$ expansions.[21,22] These expansions are at a disadvantage for the spin $\frac{1}{2}$-Ising problem because they require graphs with multiple edges which become far more numerous than the closed graphs. The Ursell–Mayer continuum method when applied to lattice problems also has the similar disadvantage of formal multiple occupancy.[11]

C. The Formulations of the Weak-Graph Method

Let us now write the abstract form of (1.1):

$$Z = \sum_{C} \prod_{\text{edges}} [a + c(\xi_i)c(\xi_j)] \prod_{i=1}^{N} b(\xi_i). \quad (1.5)$$

Here ξ_i is defined to be the vertex configuration of the ith vertex. In the previous Ising case ξ_i is the simple two-valued variable which denotes the state of the spin on the ith vertex in L_q. The summation is over all the different combinations of the vertex configurations at all N vertices; each of these combinations is conveniently called a vertex complection. There is a vertex factor $b(\xi_i)$ for each vertex i, and an edge factor for each edge ij. Further, the edge factor is the sum of a constant and a product of two functions; the latter depend only on the vertex configurations of the vertices at opposite ends of the edge. Therefore, if any problem can be put in the form of (1.5), then it can be expanded in terms of weak graphs and the weight of each graph will be known explicitly as in (1.6):

$$Z = \sum_{G \subseteq L_q} w(G) = \sum_{G \subseteq L_q} a^{(qN/2)} \prod_{i=1}^{N} \sum_{\xi_i} a^{-s_i/2} c(\xi_i)^{s_i} b(\xi_i), \quad (1.6)$$

where s_i is the degree of the ith vertex in G.

In Sec. 2 it is shown how the low-temperature spin-$\frac{1}{2}$ Ising configurational problem deduced by the direct method can be put in the form (1.5). The straightforward way of doing this recovers a weak graph expansion due to Temperley,[23] as is shown in Sec. 2B. It is also shown how one can obtain a closed weak-graph expansion directly in Sec. 2C. Each of these expansions reduces in a certain limit to a hard-core lattice-gas expansion at low density.

However, the form (1.5) is generally limited to a particular class of simple problems. To see this let there be n different vertex configurations at any vertex i (i.e., ξ_i may take on n values). Then, $c(\xi_i)$ may take on at most n values. Remembering that the variable a is also at our disposal in (1.5), there are then $n + 1$ free variables in (1.5), if we ignore the $b(\xi_i)$ for the moment. On the other hand, the problem may demand a general interaction between the state of vertex i and the state of vertex j, which may take on as many as $n(n + 1)/2$ values. Thus, for (1.5) to apply generally requires $(n + 1) \geq n(n + 1)/2$, i.e., $n \leq 2$. [Actually, as we shall see later on, there is an additional relation involving the $b(\xi_i)$ which effectively gives one more variable. This generally enables us to reduce each weak-graph expansion to a closed weak-graph expansion, but is not sufficient to enable the general $n = 3$ case to be put in the form (1.5). For this reason form (1.5) will be called the *binary vertex form* of the weak-graph expansion method.] However, it may also happen that the binary vertex form (1.5) may apply for special cases when $n > 2$. For example, in Sec. 2E a ternary model ($n = 3$) is proposed for which form (1.5) applies at one particular temperature.

There are a number of interesting problems which have been treated by what may now be considered to be a generalization of (1.5). This generalization is

$$Z = \sum_{C} \prod_{\text{edges}} [a + c_{ij}(\xi_i)c_{ji}(\xi_j)] \prod_{i=1}^{N} b(\xi_i). \quad (1.7)$$

Here the ξ_i vertex configurations have an edge dependence and accordingly form (1.7) will be called

[21] G. Horwitz and H. B. Callen, Phys. Rev. **124**, 1757 (1961).
[22] R. Brout, Phys. Rev. **118**, 1009 (1960).
[23] H. N. V. Temperley, Proc. Phys. Soc. (London) **74**, 183 and 432 (1959).

the *binary-edge form* of the weak-graph expansion method. Thus, in general $c_{ij}(\xi_i) \neq c_{ik}(\xi_i)$. It will be be noticed that even though the $c_{ij}(\xi_i)$ may in the general case be only two-valued as before in the binary-vertex form (1.5), there may now be as many as 2^q different vertex configurations since the presence or absence of an edge in the vertex configuration may vary independently on the q different edges. Furthermore, just as for problems which can be put in the binary-vertex form (1.5), any problem which can be put in the binary-edge form (1.7) may be expanded in terms of weak graphs and the weight of each graph will be known explicitly.

The binary-edge form (1.7) has been applied to the monomer–dimer problem[11] and is at the basis of the algebraic weak-closed transformation mentioned in connection with the Ising $\tanh K$ expansion in a field.[20] The binary-edge form is also the proper starting point for deriving the series expansions for various simple hydrogen-bonded problems.[18] Since these latter expansions were derived in a more complicated way before the general method was clarified, the residual entropy-of-ice series expansion will be rederived in Sec. 3 as a simple illustration of the application of (1.7). Further, in Sec. 3 a new and interesting model is devised which has a phase transition and which may be expanded via the binary-edge form. This model is named the ramrod model.

It is also possible to generalize the binary-vertex form (1.5) to apply to problems where the number n of vertex configurations (with no edge dependence) is greater than two as follows:

$$Z = \sum_{\mathcal{C}} \prod_{\text{edges}} \left[a + \sum_{r=1}^{h} c_r(\xi_i) c_r(\xi_j) \right] \prod_{i=1}^{N} b(\xi_i). \quad (1.8)$$

The form (1.8) will be called the higher-order vertex form. The weights of graphs obtained via the form (1.8) may still be determined. However, the graphs themselves usually become more complicated in that generally h different kinds of edges are required in the weak-graph expansion. For example, the spin-1 Ising model treated in Sec. 4 requires two different kinds of edges in the graphs. On the other hand the weak-graph expansion of the k-color problem also given in Sec. 4 recovers a classical graph-theoretical expansion due to Birkhoff,[24] which only requires one kind of edge.[25] However, the weak-graph weights are in such a form that the graph-combinatorial information cannot be easily obtained from the closed-graph combinatorial information. For these and more complicated problems it is likely that one could more easily obtain finite exact series with the finite-cluster method or the $1/kT$ expansions.

Thus, the weak-graph method is not a panacea to cure all series-expansion problems. Rather, its success depends on being able to use to advantage the special properties of simple lattice problems which more general methods overlook.

2. APPLICATIONS OF THE WEAK-GRAPH METHOD USING THE BINARY-VERTEX FORM

In this section the weak-graph method using the binary-vertex form is applied to three problems, the low-temperature spin-$\frac{1}{2}$ Ising problem, the hard-core lattice-gas problem, and a special ternary-model problem. The first two problems are explained in Sec. 2A. In Sec. 2B a straightforward weak-graph series expansion due originally to Temperley[23] is rederived for these two problems. It is also possible to derive a series expansion in terms of only the closed weak graphs as is shown in Sec. 2C. In Sec. 2D these expansions are compared, from the point of view of efficiently obtaining finite exact series, with previous expansion techniques. Finally, in Sec. 2E a special ternary model is introduced and a weak-graph series expansion is derived for it.

A. The Configurational Problems for the Low-T Ising Model and the Hard-Core Lattice-Gas Model

The usual low-temperature spin-$\frac{1}{2}$ Ising series expansion is one of the simplest applications of the direct method of obtaining series expansions. The ground state of the system (ferromagnetic) has all spins aligned, say $\sigma_i = -1$, $i = 1, \cdots, N$. The first perturbed state has one overturned $(+1)$ spin and this costs $2qJ + 2mH$ in energy, where $\sigma_i\sigma_j J$ is the nearest-neighbor pair interaction energy and H is the field. Usually, overturning two spins costs $4qJ + 4mH$ in energy. However, if the two spins are nearest neighbors, it only costs $(4q - 4)J + 4mH$. In general, the energy to overturn v spins is $(2vq - 4e)J + vmH$, where e is the number of nearest-neighbor pairs among the v overturned spins. This leads one to consider subgraphs of the lattice for which the vertices correspond to overturned spins. One should notice that these subgraphs may have isolated vertices of degree zero, but that they are restricted in the sense that whenever two overturned spins are nearest neighbors on the lattice, then the subgraph must contain the edge (s) between these two vertices (see Fig. 2).

If G is an abstract graph, then $[G \mid L_q]$ will denote the number of distinct ways in which G may occur as

[24] G. D. Birkhoff, Ann. Math. **14**, 42 (1912).
[25] O. Ore, "Theory of Graphs," Am. Math. Soc. Colloq. Publ., Vol. 37 (1962).

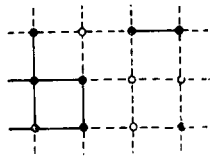


FIG. 2. An example of a strongly embedded subgraph (solid vertices and solid edges).

a subgraph of L_q under the previous restriction that any two vertices of G (G considered now as a particular subgraph of L_q) which are joined by an edge in L_q must also be joined by an edge in G. The quantity $[G \mid L_q]$ is called the number of strong embeddings of G in L_q.[19] (Often it is convenient to just refer to strong graphs. However, the lattice L_q must always be remembered. For example, a six-sided polygon may be a strong subgraph in the honeycomb lattice but not in the square lattice.) Thus, the direct-method formulation of the low-T spin-$\frac{1}{2}$ Ising series expansion is

$$Z(K, L) = \sum_{G \subseteq L_q} [G_{ve} \mid L_q] z^v u^{-e}, \tag{2.1}$$

where

$$z = \exp -(2mH + 2qJ)/kT \quad \text{and} \quad u = \exp(-4J/kT)$$

and v and e are, respectively, the number of vertices and number of edges in G.

The hard-core lattice gas is defined as follows:[26] Each gas atom sits on a vertex of the lattice L_q. Further, two gas atoms may not occupy either the same or adjacent (neighboring) vertices which are joined by an edge in L_q. Thus, each gas atom has an infinitely repulsive core consisting of its own lattice vertex and its neighboring vertices. Now let z be the activity of a gas atom. Clearly, taking $u = 0$ in (2.1) results in a low-activity (or low-density) expansion for the hard-core lattice-gas problem. Thus, any weak-graph series expansion which we derive for (2.1) includes a low-density expansion for the hard-core lattice-gas problem as a special case.

B. The Weak-Graph Expansion for the Low-T Ising Problem

Let us define the vertex configuration ξ_i to take on the value 0 when the vertex i of the lattice L_q is unoccupied and the value 1 when the vertex i is occupied. Each occupied vertex contributes a factor z and each unoccupied vertex contributes a factor 1 in (2.1). It is easy to accomplish the same thing in the binary-vertex form (1.5) by setting

$$\begin{aligned} b(\xi_i) &= 1, \quad \xi_i = 0, \\ &= z, \quad \xi_i = 1. \end{aligned} \tag{2.2}$$

To put (2.1) into the form (1.5) it is next necessary to require

$$a + [c(0)]^2 = a + c(1)c(0) = 1 \tag{2.3}$$

and

$$a + [c(1)]^2 = u^{-1}. \tag{2.4}$$

Equations (2.3) and (2.4) can be satisfied if and only if $a = 1$ and

$$\begin{aligned} c(\xi_i) &= 0, \quad \xi_i = 0, \\ &= y \equiv (u^{-1} - 1)^{\frac{1}{2}}, \quad \xi_i = 1. \end{aligned} \tag{2.5}$$

Now let us expand (1.5) into a weak-graph series where $b(\xi_i)$ and $c(\xi_i)$ are defined by (2.2) and (2.5). The weight of each graph is

$$\begin{aligned} w(G_{ve}) &= \sum_{G} \prod_{\text{edges in G}} c(\xi_i)c(\xi_j) \prod_{i=1}^{N} b(\xi_i) \\ &= \prod_{i=1}^{N} \sum_{\xi_i} [c(\xi_i)]^{s_i} b(\xi_i) \\ &= (1 + z)^N [z/(1 + z)]^v y^{2e}. \end{aligned} \tag{2.6}$$

Here s_i is the degree of the ith vertex in G_{ve} and v and $e = \frac{1}{2}\sum_i s_i$ are, respectively, the number of vertices and edges in G_{ve}. Hence,

$$Z(K, L) = (1 + z)^N \sum_{G_{ve} \subseteq L_q} y^{2e} \left(\frac{z}{1 + z}\right)^v, \tag{2.7}$$

where the summation is over all weak graphs G_{ve} which are contained in the lattice L_q.

At this point it will be convenient to introduce some notation for weak graphs contained in the q-coordination lattice L_q. Let p_s be the number of vertices which have degree s in G_{ve}, i.e., which are joined by s edges to other vertices in G_{ve}. Let us partially describe G_{ve} as a subgraph of L_q by the vector or q-tuple $\mathbf{p} = (p_1, \cdots, p_s, \cdots, p_q)$. In particular,

$$v = \sum_{s=1}^{q} p_s$$

gives the number of vertices in G_{ve} and

$$e = \frac{1}{2} \sum_{s=1}^{q} s p_s$$

gives the number of edges in G_{ve}. Also, let $g(\mathbf{p})$ be the number of subgraphs of L_q with partial description $\mathbf{p}$. Then, (2.7) may be rewritten [using (2.1) as well]

$$\begin{aligned} (1 + z)^N \sum_{\mathbf{p}} g(\mathbf{p}) y^{2e} \left(\frac{z}{1 + z}\right)^v \\ = \sum_{G_{ve} \subseteq L_q} [G_{ve} \mid L_q] z^v (1 + y^2)^e. \end{aligned} \tag{2.8}$$

Equation (2.8) was essentially described by Temperley.[23]

[26] D. S. Gaunt and M. E. Fisher, J. Chem. Phys. **43**, 2840 (1965).

Now, by the weak-closed algebraic transformation mentioned earlier, the g(**p**) for weak graphs may be obtained from the g(**p**) for the closed weak graphs according to various patterns which are described elsewhere.[20] Therefore, (2.8) plus the weak-closed transformation gives a two-step transformation by which the low-temperature Ising combinatorial problem may be reduced to the closed weak-graph combinatorial problem. However, as is shown in the next subsection, it is possible to achieve this (strong–closed-weak) transformation in just one step. Although it turns out that these two ways of doing the strong–closed-weak transformation are for all practical purposes entirely equivalent, the derivation of the direct method illustrates some additional flexibility of the weak-graph method. It will also be easier to see just what closed-graph information is needed to obtain particular terms in the low-temperature Ising series.

C. The Closed Weak-Graph Expansion

In the last subsection our rather obvious choice (2.2) for $b(\xi_i)$ determined $c(\xi_i)$ via (2.3) and (2.4). Consequently, in order to eliminate graphs with any vertices of degree one in the final expansion, that is, to require

$$\sum_{\xi_i} c(\xi_i)b(\xi_i) = 0, \tag{2.9}$$

will obviously require a different choice of $b(\xi_i)$ and $c(\xi_i)$.

It is easiest to start by allowing $c(\xi_i)$ to be more generally defined as

$$\begin{aligned} c(\xi_i) &= -x \quad \xi_i = 0, \\ &= y, \quad \xi_i = 1. \end{aligned} \tag{2.10}$$

Let us for the moment retain (2.2) for $b(\xi_i)$ and compute

$$Q = \sum_{\mathcal{C}} \prod_{\text{edges}} [1 + c(\xi_i)c(\xi_i)] \prod_{i=1}^{N} b(\xi_i), \tag{2.11}$$

where $c(\xi_i)$ is given by (2.10) and $b(\xi_i)$ is given by (2.2). Then,

$$Q = \sum_{\mathcal{C}} (1 + x^2)^{n_{00}}(1 - xy)^{n_{10}+n_{01}}(1 + y^2)^{n_{11}}z^{N_1}, \tag{2.12}$$

where N_α is the number of vertices with vertex configuration $\alpha = 0$ or 1 and $n_{\alpha\beta}$ is the number of edges with vertex configurations $\alpha = 0$ or 1 at one end and $\beta = 0$ or 1 at the other end. The summation is over all vertex complections. Now divide out $(1 + x^2)^{qN/2}$ in (2.12) and then use the simple relation

$$2n_{11} + n_{10} + n_{01} = qN_1 \tag{2.13}$$

to eliminate $n_{10} + n_{01}$ in (2.12). The result is

$$Q = (1 + x^2)^{qN/2} \sum_{\mathcal{C}} [(1 + y^2)(1 + x^2)/(1 - xy)^2]^{n_{11}} \times [z((1 - xy)/(1 + x^2))^q]^{N_1}. \tag{2.14}$$

Now, remembering that $n_{11} = e$ and $N_1 = v$, we see that (2.14) is in the form of (2.1). To obtain an exact equivalence we need only divide out $(1 + x^2)^{qN/2}$, identify $u = (1 + y^2)(1 + x^2)/(1 - xy)^2$, and redefine

$$\begin{aligned} b(\xi_i) &= 1 \qquad\qquad\qquad \xi_i = 0, \\ &= z[(1 + x^2)/(1 - xy)]^q, \quad \xi_i = 1. \end{aligned} \tag{2.15}$$

The reason for these gyrations is to enable us to satisfy (2.9) which requires from (2.15) and (2.10) that

$$x = yz[(1 + x^2)/(1 - xy)]^q. \tag{2.16}$$

Although we cannot solve (2.16) explicitly for x in terms of y and z for general q, we can expand it in a series as follows:

$$x = zy + z^2qy^3 + z^3qy^3[1 + (3q + 1)(y^2/2)] + \cdots. \tag{2.17}$$

Thus, for the purpose of obtaining exact finite series expansions, x may be considered to be a known function of y and z.

Now we may use (2.10) and (2.15) in (1.5) in the same way as in Secs. 1B and 2B to obtain the following closed weak-graph series expansion:

$$\begin{aligned} &(1 + xy^{-1})^N \sum_{\mathbf{p}} \mathrm{g}(\mathbf{p})y^{2e}z^v[(1 + x^2)/(1 - xy)]^{qv} \\ &\qquad \times \prod_{s=1}^{q} \{[1 - (-xy^{-1})^{s-1}]/(1 + xy^{-1})\}^{p_s} \\ &= (1 + x^2)^{qN/2} \\ &\qquad \times \sum_{\mathrm{G} \subseteq \mathrm{L}_q} [\mathrm{G} \mid \mathrm{L}_q]z^v[(1 + y^2)(1 + x^2)/(1 - xy)^2]^e, \end{aligned} \tag{2.18}$$

and x is related to z and y by (2.16). It may be verified that if **p** is not closed, i.e., $p_1 \neq 0$, then g(**p**) is multiplied by a zero factor on the left-hand side of (2.18), so only closed weak g(**p**) need be considered.

Let us consider how (2.18) enables us to obtain the strong $[\mathrm{G}_{ve} \mid \mathrm{L}_q]$ for subgraphs G_{ve} with v vertices and e edges from the closed weak g(**p**). First, we observe that the expansion (2.17) for x has a general term $z^v\psi_v(y)$ where $\psi_v(y)$ is a polynomial in y with highest power $2v - 1$. Next, we observe that the coefficient of either g(**p**) or $[\mathrm{G}_{ve} \mid \mathrm{L}_q]$ in (2.18) is a polynomial

$$z^v\phi_v(y) + z^{v+1}\phi_{v+1}(y) + \cdots, \tag{2.19}$$

where the highest power of y which occurs in each $\phi_{v+k}(y)$ is y^{2e+2k}. Again, v and e are, respectively, the

numbers of vertices and edges of either the weak **p** in g(**p**) or the strong G_{ve} in $[G_{ve} \mid L_q]$. Now, we see that equating the coefficients of $z^v y^{2e}$ on both sides of (2.18) gives a relation between only those g(**p**) and $[G_{v_p e_p} \mid L_q]$ for which the number of vertices v_p in **p** or $G_{v_p e_p}$ satisfies

$$v_p \leq v \tag{2.20}$$

and the number or edges e_p in the **p** or G satisfies

$$e_p - v_p \geq e - v. \tag{2.21}$$

Thus, if one knows all the g(**p**) and $[G_{v_p e_p} \mid L_q]$ for **p** and $G_{v_p e_p}$ satisfying (2.20) and (2.21) with the single exception that one does not know $[G_{v_p e_p} \mid L_q]$ for $G_{v_p e_p}$ such that $v_p = v$ and $e_p = e$, then one may compute this last $[G_{v_p e_p} \mid L_q]$ from (2.18). By induction, then, one may compute the $[G_{v_p e_p} \mid L_q]$ for $G_{v_p e_p}$ satisfying (2.20) and (2.21) from the corresponding g(**p**).

It may occur to the reader that finding the $[G_{ve} \mid L_q]$ from the g(**p**) via (2.18) requires a great deal of algebra; in fact, it is easier to obtain the first few $[G_{ve} \mid L_q]$ by straightforward counting than via (2.18). However, the labor involved in counting graphs increases roughly exponentially in adding more terms to series whereas the labor of doing the algebra increases roughly linearly. Furthermore, an algebraic computer program works for all lattices whereas computer programs to count subgraphs are not as flexible. Therefore, the complicated algebra involved in (2.18) is a relatively small price to pay in extending exact finite series when compared to straightforward subgraph counting.

D. On Obtaining Exact Finite Series via (2.18) Compared to Other Ways

First, let us briefly discuss which g(**p**) are required to obtain two different kinds of Ising series from (2.18). First, there is the usual low-temperature expansion in powers of $u = \exp(-4J/kT)$.[1] From (2.1) the nth term in this expansion requires all $[G_{ve} \mid L_q]$ for G_{ve} such that

$$(q/2)v - e \leq n. \tag{2.22}$$

From (2.20) and (2.21) we see that from (2.18) we also need only the closed weak g(**p**) where the **p** also satisfy (2.22). Second, there is the high-field expansion in powers of $\mu = \exp(2mH/kT)$.[1] Again from (2.1) the vth term in this expansion requires all $[G_{v_p e_p} \mid L_q]$ for $G_{v_p e_p}$ with $v_p \leq v$. Also, from (2.20) and (2.21) we see that only the closed weak g(**p**) where **p** satisfies $v_p \leq v$ are needed to give the vth term in the high-field expansion via (2.18).

The $[G_{ve} \mid L_q]$ are difficult to obtain directly because many separated components occur for small graphs due to the possibility of vertices of degree zero. The strong embedding restriction does not reduce the labor very much. It is easier to obtain the g(**p**) for general open **p** because there can be no isolated vertices of degree zero and therefore fewer separated components than for the corresponding $[G_{ve} \mid L_q]$. However, the additional closed-graph requirement that there be no vertices of degree one makes it several orders of magnitude easier to supply the closed g(**p**) in (2.18) than to supply the corresponding $[G_{ve} \mid L_q]$.

Previously, other ways have been developed for obtaining the $[G_{ve} \mid L_q]$ for the Ising spin-$\frac{1}{2}$ low-temperature or high-field expansions. The first way uses what will be called the high–low Ising transformation in this paper.[27] This transformation was discovered by Domb because the coefficients in the μ-high-field expansion seem to follow a certain form. Domb suggested that this form follows because the high-temperature hyperbolic-tangent expansion may also be rearranged in a high-field expansion. (This requires very plausible assumptions concerning the nonexistence of phase boundaries in unlikely regions of the phase diagram for the Ising model.) The actual transformation equation seems never to have been written down in the literature so we present it as follows:

$$\begin{aligned}(1/N)\log Z_N(K, L) &= \mu a u^{q/2} + \mu^2(b_0 u^q + b_1 u^{q-1} + \cdots) \\ &\quad + \mu^3(c_0 u^{3q/2} + \cdots) + \cdots \\ &= \log(1+\mu) + \sum_{r=1}^{\infty}[\varphi_r(\tau) + (-1)^r(q/2r)] \\ &\quad \times \left[\sum_{n=1}^{\infty}((1-u^{\frac{1}{2}})/2)^n\right]^r,\end{aligned} \tag{2.23}$$

where $\tau = (1-\mu)/(1+\mu)$. The right-hand side comes from rearranging the hyperbolic-tangent expansion. Thus,

$$\varphi_r(\tau) = \sum_{\substack{\mathbf{p} \\ e=r}} g_1(\mathbf{p})\tau^{v(\text{odd})}, \tag{2.24}$$

where the summation is over all weak **p** with $e = r$ edges and v (odd) is the number of vertices of odd degree in **p**. Also, $g_1(\mathbf{p})$ is the coefficient of N in g(**p**). Since no easy way has existed until recently to obtain the weak g(**p**), only the form of the high–low Ising transformation has been used. In this way it has been possible to eliminate all $[G_{ve} \mid L_q]$ with $e \leq v + 1$, which include the most difficult separated graphs.

[27] C. Domb, Proc. Roy. Soc. (London) **A199**, 199 (1949).

Although this is a very great help, obtaining the remaining $[G_{ve} \mid L_q]$ still poses a counting problem which is more difficult than the corresponding closed-weak-graph counting problem.

Now that the weak-closed transformation is known, it may also be asked if the $[G_{ve} \mid L_q]$ in the high-field expansion can be computed more efficiently via (2.23) and the closed weak g(**p**) than via (2.18). The answer seems to be no. To obtain the vth term in (2.23) requires all g(**p**) for **p** with e^* or fewer edges, where e^* is the maximum number of edges which any graph of v vertices may have. This set of closed graphs with e^* or fewer edges generally contains the set of closed-weak graphs with v or fewer vertices so the high–low transformation requires more g(**p**) than the weak-graph transformation (2.18).

To extend the low-temperature and high-field Ising spin-$\frac{1}{2}$ series, Sykes, Essam, and Gaunt have ingeniously applied the direct series-expansion method in a way which takes advantage of the two-sublattice structure of loose-packed lattices.[28] It is difficult to make efficiency comparisons between this way of obtaining the $[G \mid L]$ and (2.18) because the kinds of graphs are so different. [Also, of course, extensive calculations have not yet been done using (2.18).] However, it seems to this author that for the loose-packed lattices probably the Sykes–Essam–Gaunt direct approach is slightly easier for the μ-high-field expansion and the weak-graph approach may be slightly easier for the low-temperature expansion in u.

It should also be kept in mind that large portions of the closed weak g(**p**) needed to obtain the low-temperature–high-field Ising series via (2.18) are also needed to obtain the high-temperature hyperbolic-tangent expansion. Since the high-temperature series is not efficiently obtainable in any other way, (2.18) would allow a more unified graphical approach to obtaining Ising series coefficients.

Finally, it should be mentioned that Gaunt and Fisher[26] have developed the direct series-expansion method for the hard-core lattice-gas problem in a way similar to Sykes *et al.*[28] The hard-core lattice-gas low-density expansion is analogous to the μ-high-field Ising expansion and, as has been mentioned, the direct approach is probably easier for this type of series for loose-packed lattices. Furthermore, Gaunt and Fisher find a connection between their low-density expansion and a high-density expansion. Thus far, the author has not succeeded in fruitfully applying the weak-graph method to the high-density phase. Thus, the direct method seems superior to the weak-graph method for the problem of obtaining exact series for the hard-core lattice-gas problem.

[28] M. F. Sykes, J. W. Essam, and D. S. Gaunt, J. Math. Phys. **6**, 283 (1965).

E. A Special Ternary Model

This section will show how the binary-vertex weak-graph form (1.5) may be applied under special conditions to models with more than two vertex configurations at each vertex. Let us consider a ternary model where each vertex must be occupied by one atom which may be of three types, a, b, or c. Let us suppose that an a atom and a b atom may not be placed on vertices which are joined by an edge in L_q (see Fig. 3). Further, let the c atoms be neutral in the sense that each c atom has zero energy of interaction with neighboring a, b, or c atoms. Thus, the nearest-neighbor pair-interaction energies are

$$J_{ab} = \infty, \quad J_{ac} = J_{bc} = J_{cc} = 0, \tag{2.25}$$

and also let us require that $J_{aa} = J_{bb}$. Thus, there remains only one unspecified interaction energy.

Now, let us write

$$\begin{aligned} b(\xi_i) = z_a, &\quad \text{if} \quad \xi_i = a, \\ z_b, &\quad \text{if} \quad \xi_i = b, \\ z_c, &\quad \text{if} \quad \xi_i = c, \end{aligned} \tag{2.26}$$

where the z's are the activities of the atoms. In order that (1.5) may represent the grand canonical partition function under the conditions in the preceding paragraph, we require that $a = 1$ and

$$\begin{aligned} c(\xi_i) = \quad 1, &\quad \text{if} \quad \xi_i = a, \\ -1, &\quad \text{if} \quad \xi_i = b, \\ 0, &\quad \text{if} \quad \xi_i = c, \end{aligned} \tag{2.27}$$

and this further requires that $\exp(J_{aa}/kT) = 2$. Thus, for a given interaction energy J_{aa} in this ternary model, the weak-graph expansion method in the form (1.5) applies for one particular temperature, namely, $kT = J_{aa}/\log 2$.

We may obtain the weak-graph expansion of the grand canonical partition function in the usual way [see (1.6)]:

$$\begin{aligned} &\Xi_N(z_a, z_b, z_c) \\ &\quad = \prod_{i=1}^{N} \sum_{\xi_i} [c(\xi_i)]^{s_i} b(\xi_i) \\ &\quad = (z_a + z_b + z_c)^N \sum_{\mathbf{p}} \mathrm{g}(\mathbf{p}) \\ &\qquad \times \prod_{s=1}^{q} [(z_a + (-1)^s z_b)/(z_a + z_b + z_c)]^{p_s}, \end{aligned} \tag{2.28}$$

where p_s and g(**p**) were defined in Sec. 2B.

The series expansion (2.28) applies to the high z_c, single-phase region which has as an "end point"

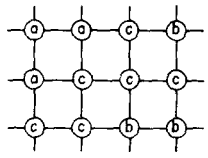


FIG. 3. An example of an allowed configuration for the special ternary model discussed in Sec. 2E.

$z_a = 0 = z_b$ and $z_c > 0$. The two-phase region will be described by $z_a = z_b \neq 0$ and z_c smaller than some critical activity, $z_{c,\text{crit}}$. By analogy to the hard-core lattice-gas problem[29] or the antiferromagnetic Ising problem, it seems that the best way to find the value for the critical activity $z_{c,\text{crit}}$ from the one-phase series expansion would be from a study of a "staggered" compressibility which essentially measures the fluctuations in the average number density difference $\langle N_a - N_b \rangle$.[23] For this ternary problem this staggered susceptibility may be defined as

$$\chi_s = (d^2/dx^2)(1/N) \log \Xi_N(z_a, z_b, z_c)|_{x=0}, \quad (2.29)$$

where $z_a = z_0 + x$ and $z_b = z_0 - x$. Then, from (2.28)

$$\chi_s = 2z_0/(2z_0 + z_c) + \sum' g(\mathbf{p}) 2[2z_0/(2z_0 + z_c)]^v + \sum'' g(\mathbf{p}) v[z_c/(2z_0 + z_c)][2z_0/(2z_0 + z_c)]^v, \quad (2.30)$$

where the first sum is over all $\mathbf{p}$ which have two vertices of odd degree and the second sum is over all $\mathbf{p}$ which have no vertices of odd degree. An analysis of the series for this problem will be the object of further study.

3. APPLICATIONS OF THE WEAK-GRAPH METHOD USING THE BINARY-EDGE FORM

A. Introduction

In this paper the weak-graph method has been presented as a generalization of the Ising spin-$\frac{1}{2}$ hyperbolic-tangent expansion. Historically, it seems that no one realized the underlying generality of the hyperbolic-tangent expansion and the weak-graph method has been developed in an entirely different way. The first step of clearly defining vertex configurations and the notion of a compatibility function $a(\xi_i, \xi_i)$, which has since been circumvented, was performed by DiMarzio and Stillinger for the residual entropy-of-ice problem.[30] It was proved by this author that the resulting graph weights could be explicitly determined,[18] but without realizing the significance of the factorability of the compatibility function $a(\xi_i, \xi_i) = c_{ij}(\xi_i)c_{ji}(\xi_i)$. However, this point was discovered when the former technique was applied to the monomer–dimer problem.[11] Then further applications via form (1.7) became apparent, such as the one suggested by Temperley which led to the weak-closed transformation.[20] Once the binary-edge form (1.7) was revealed, the simplification to the binary-vertex form (1.5) was easy.

As a result of this somewhat reverse order of development, applications of the binary-edge form (1.7) are already represented in the literature whereas the applications in Sec. 2 are new. However, it seems appropriate first to illustrate the use of form (1.7) by rederiving in Sec. 3B the simple residual entropy-of-ice series expansion (for loose-packed lattices); this was originally derived in the more complicated way.[18] Then, form (1.7) will be applied to the ramrod model introduced in Sec. 3C.

It might also be mentioned in passing that the series expansions derived via the binary-edge form are relatively more valuable than the expansions derived via the binary-vertex form in Secs. 2B and 2C. This is largely because they have no competition from the direct method of obtaining series expansions, which appears not to be readily applicable to these more complex problems. Previously, the high–low Ising transformation has been used to obtain series, such as the Ising-model series in a field at high temperatures[31] and the monomer–dimer series,[32] which can most likely be obtained more easily via the weak-closed transformation and the weak-graph monomer–dimer series expansion.

B. The Residual Entropy-of-Ice Series Expansion

In order to unify this presentation with other work (e.g., Sec. 3C and Nagle[11]), it is advantageous to discuss a formulation of the ice problem which is equivalent to the usual formulation for loose-packed lattices only. (A loose-packed lattice has no cycles with an odd number of edges, e.g., square or simple cubic lattices.) The allowed states of the ice problem on loose-packed lattices are in one–one correspondence with subgraphs of a certain type on the lattice as illustrated in Fig. 4. Each of these subgraphs, which will be called icelike subgraphs, contains all the vertices in the lattice L_q (with q an even integer) and obeys the restriction that each vertex is of degree $(q/2)$ in the subgraph. The number of icelike subgraphs will be called W_N; then the residual entropy of ice is

[29] It might be mentioned that the hard-core lattice gas on loose-packed lattices may be thought of as a ternary model, by identifying component a with occupied vertices on the A sublattice, component b with occupied vertices on the B sublattice, and component c with empty vertices.

[30] E. A. DiMarzio and F. H. Stillinger, Jr., J. Chem. Phys. **40**, 1577 (1964).

[31] A. Bienenstock, J. Appl. Phys. **37**, 1459 (1966).

[32] D. S. Gaunt (private communication).

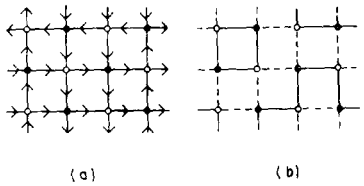


FIG. 4. (a) An allowed configuration for the ice problem. (b) The corresponding icelike subgraph. Notice that the vertices designated by open circles are on a different sublattice than those designated by closed circles. The correspondence between (a) and (b) is that those arrows directed toward an open-terminal vertex become edges in the icelike subgraph.

$(k/N)\log W_N$. (Of course, this is a generalization of the residual entropy-of-ice problem, since ice-I requires just one particular three-dimensional lattice which is four-coordinated.)[18]

Now let us formulate the weak-graph expansion for W_N. First, the vertex configurations are characterized as shown in Fig. 5. Each of the solid lines incident to a vertex may be thought of as a half-edge in one of the W_N icelike subgraphs. Each of the dotted lines incident to a vertex may be thought of as a half-edge of the lattice which is not included in the icelike subgraph (see Fig. 4).

Let us consider a general vertex complection on the lattice where each vertex assumes one of its possible vertex configurations. If both vertex configurations ξ_i and ξ_j have a half-edge on the edge ij connecting vertices i and j, then the edge ij is analogous to the edges in the icelike subgraph in Fig. 4. If both vertex configurations ξ_i and ξ_j do not have half-edges on the edge ij, then the edge ij is analogous to one of the edges in the lattice not included in the ice-like subgraph in Fig. 4. Finally, when just one of ξ_i or ξ_j has a half-edge on the edge ij, then there is a mismatch which is not analogous to anything in Fig. 4. Clearly, counting only those vertex complexions for which the last possibility does not occur for any edge ij gives W_N. This will be accomplished by using the binary-edge form (1.7) with the proper choice of $c_{ij}(\xi_i)$.

It is perhaps worth describing what has been accomplished here. The original formulation of the problem involved a q-body vertex restriction which depended on the "states" of the q edges incident to that vertex. However, by reformulating the problem in terms of vertex configurations, the restriction involves only the binary (or 2-body) matching or mismatching of half-edges.

The mechanics of applying form (1.7) are now quite simple. First, one requires that

$$c_{ij}(\xi_i')c_{ji}(\xi_j'') = -a,$$
$$c_{ij}(\xi_i')c_{ji}(\xi_j') = c_{ij}(\xi_i'')c_{ji}(\xi_j'') \neq -a, \qquad (3.1)$$

where ξ_i' or ξ_j' is a vertex configuration which has a half-edge on edge ij in the lattice, and ξ_i'' or ξ_j'' is a vertex configuration which does not have a half-edge on edge ij in the lattice. The solution of (3.1) is

$$c_{ij}(\xi_i) = a^{\frac{1}{2}}, \quad \text{if } \xi_i \text{ has a half-edge on edge } ij,$$
$$= -a^{\frac{1}{2}}, \quad \text{if } \xi_i \text{ does not have a half-edge on edge } ij. \qquad (3.2)$$

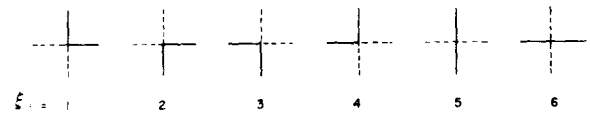


FIG. 5. The vertex configurations for the ice problem for $q = 4$.

From (3.2) it is easy to confirm that

$$W_N = \sum_{\mathcal{C}} \prod_{\text{edges}} (1/2a)[a + c_{ij}(\xi_i)c_{ji}(\xi_j)]. \qquad (3.3)$$

This may be compared to (1.7) for which the $b(\xi_i)$ is taken to have the value 1 for all ξ_i.

The weak-graph series expansion for (3.3) now follows from the usual manipulations. Each graph weight is given by

$$\mathrm{w(G)} = (\tfrac{1}{2})^{qN/2} \sum_{\mathcal{C}} \prod_{\text{edges in G}} (1/a)c_{ij}(\xi_i)c_{ji}(\xi_j)$$
$$= (\tfrac{1}{2})^{qN/2} \prod_{i=1}^{N} \sum_{\xi_i} (1/a)^{s_i/2} \prod_{\substack{k \\ ik \text{ in G}}} c_{ik}(\xi_i). \qquad (3.4)$$

The value of the summation over each ξ_i in (3.4) depends on how many $c_{ik}(\xi_i)$ factors there are in the product. For general q (even), all these sums vanish when there is an odd number of $c_{ik}(\xi_i)$ factors, that is, when vertex i has odd degree in G. Now let us particularize to the case $q = 4$. Then, the sum is easily calculated to be 6 when there are zero or four $c_{ik}(\xi_i)$ factors and it equals (-2) when there are two $c_{ik}(\xi_i)$ factors. Thus, for $q = 4$ we have

$$\mathrm{w(G)} = \mathrm{w(p)} = (\tfrac{1}{2})^{2N}(6)^{p_0+p_4}(-2)^{p_2}, \qquad (3.5)$$

where $\mathbf{p}$ has no vertices of odd degree. Otherwise, $\mathrm{w}(\mathbf{p}) = 0$. Finally, we have for $q = 4$

$$W_N = \sum_{\mathbf{p}} \mathrm{w(p)} = (\tfrac{3}{2})^N \sum_{\mathbf{p}} \mathrm{g(p)}(-\tfrac{1}{3})^{p_2}, \qquad (3.6)$$

where the sum is only over those $\mathbf{p}$ for which $p_1 = 0 = p_3$. (3.6) agrees with the previous expansion[18] except for the factor $(-1)^{p_2}$. However, for loose-packed lattices p_2 is always an even number, so the expansions are equivalent for this case only, just as one would expect.

C. The Ramrod Model

It seems likely that the monomer–dimer model does not undergo a proper phase transition.[11] However, it seems more likely that a model consisting of

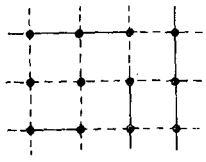


FIG. 6. An example of configuration of ramrods on the square lattice.

FIG. 7. The vertex configurations for the ramrod model for $q = 4$.

somewhat longer rigid rods on a lattice, such as linear trimers, would undergo a proper phase transition. Since the weak-graph series expansion works so well for the monomer–dimer problem, one naturally would like to apply the method to the trimer problem. However, this would require a more complex form than (1.7) which leads to complications such as the ones to be discussed in the next section.

Nevertheless, the attempt to apply the binary-edge form (1.7) to the trimer problem naturally leads one to consider a new lattice-model problem for which a weak-graph series expansion can be derived via form (1.7). This model also consists of rigid linear rods, called ramrods. Thus, a random state of the model would consist of a mixture of dimers, linear trimers, linear quadrimers, and so on. The activity of each linear r-mer or ramrod is z^{r-1} where $r-1$ is the number of edges in the ramrod (see Fig. 6).

Restricting the discussion to the square lattice for simplicity, let us qualitatively consider the behavior of the ramrod model. When the activity z is small, the most probable state will consist of only a relatively few short ramrods scattered on the lattice. Since there is little excluded volume effect at low density of ramrods, these ramrods will be isotropically oriented with respect to the x or y axes of the lattice. At the other extreme when z is very large, say infinity, the most probable state must have a maximum length of total ramrod. This can only be accomplished by lining up very long (infinite) ramrods parallel to the x axis or by lining up long ramrods parallel to the y axis. These two "ground" states are extremely anisotropic with respect to the x or y axes. Clearly, somewhere in between the high-z and the low-z extremes there must be a phase transition from the isotropic phase to the anisotropic phase.

The ramrod model may also be discussed for other lattices, including the honeycomb and diamond lattices, if obvious modifications are made. Also, the model may be specialized to the case when each ramrod must be infinitely long, although this is an uninteresting case for two-dimensional lattices. This latter model is similar to the continuum model of oriented long thin rods treated by Zwanzig[33] who followed up some work of Onsager[34]; references to physical problems of interest may be found in these two papers.

It seems likely to this author that the lattice ramrod model has a continuous phase transition similar to the hard-core lattice-gas phase transition[26] and the ternary-model transition conjectured in Sec. 2E. If this is the case, then the critical activity will be most accurately located from a study of a suitably defined staggered susceptibility. Although the staggered susceptibility is not difficult to obtain via the weak-graph method, it does provide one additional complication which will be left to accompany later work on obtaining and analyzing the thermodynamic series of interest.

For simplicity in presentation, the following derivation of the weak-graph series expansion for the ramrod model will be restricted to the square lattice. The vertex configurations are shown in Fig. 7. Using the ideas of Sec. 3B, it is now easy to write down a general weak-graph series expansion. However, this does not accomplish much, especially since the original conformations of ramrods are themselves weak graphs already. The problem then is to obtain a closed weak-graph series expansion. Following Sec. 2C it is easiest to start by defining a general half-edge function

$$\begin{aligned} c_{ij}(\xi_i) &= x_1, \text{ if } \xi_i \text{ does not have a half-edge on edge } ij, \\ &= -x_2, \text{ if } \xi_i \text{ has a half-edge on edge } ij, \end{aligned} \quad (3.7)$$

and then $a = x_1 x_2$. However, the constant a may just as generally be divided out (or set $1 = x_1 x_2$) so that we will take

$$\begin{aligned} c_{ij}(\xi_i) &= y, \text{ if } \xi_i \text{ does not have a half-edge on edge } ij, \\ &= -y^{-1}, \text{ if } \xi_i \text{ has a half-edge on edge } ij. \end{aligned} \quad (3.8)$$

Now, it is easy to see that the grand canonical partition function will be

$$\Xi_N(z) = (1+y^2)^{-2N} \sum_{\mathcal{C}} \prod_{\text{edges}} [1 + c_{ij}(\xi_i) c_{ji}(\xi_j)] \prod_{i=1}^{N} b(\xi_i), \quad (3.9)$$

where

$$\begin{aligned} b(\xi_i) = 1, &\quad \xi_i = 1, \\ z^{\frac{1}{2}}y, &\quad \xi_i = 2, \cdots, 5, \quad \text{(see Fig. 7)} \\ zy^2, &\quad \xi_i = 6, 7. \end{aligned} \quad (3.10)$$

[33] R. Zwanzig, J. Chem. Phys. **39**, 1714 (1963).

[34] L. Onsager, Ann. N.Y. Acad. Sci. **51**, 627 (1949).

The usual procedure for obtaining the closed weak-graph series expansion may now be carried out for (3.10) (see Sec. 2C):

$$\Xi_N(z) = (1 + y^2)^{-2N} \sum_{G \subseteq L_{sq}} w(G), \tag{3.11}$$

where

$$w(G) = \sum_{C} \prod_{\text{edges in G}} c_{ij}(\xi_i)c_{ji}(\xi_j) \prod_{i=1}^{N} b(\xi_i)$$
$$= \prod_{i=1}^{N} \sum_{\xi_i=1}^{7} b(\xi_i) \prod_{\substack{k \\ ik \text{ in G}}} c_{ik}(\xi_i). \tag{3.12}$$

First, let us perform one of the sums in (3.12) for a vertex of degree one:

$$v_1 = y + z^{\frac{1}{2}}y(3y - y^{-1}) + zy^2(y - y^{-1}). \tag{3.13}$$

In order for (3.9) to be a closed weak-graph series expansion, the sum v_1 must be set equal to zero and this determines y:

$$y = z^{\frac{1}{2}} - 2z^{\frac{3}{2}} + 9z^{\frac{5}{2}} + \cdots. \tag{3.14}$$

Next, for vertices of degree 0, 3, and 4 in (3.12), the sums v_0, v_3, and v_4 are easily computed to be

$$v_0 = 1 + 4z^{\frac{1}{2}}y + 2zy^2, \tag{3.15}$$
$$v_3 = y^3 + z^{\frac{1}{2}}y(y^3 - 3y) + zy^2(-y + y^{-1}), \tag{3.16}$$
$$v_4 = y^4 + z^{\frac{1}{2}}y(-4y^2) + zy^2(2). \tag{3.17}$$

However, it must be noticed that the sum in (3.12) is different for two different kinds of vertices of degree two. For vertices of degree two for which the incident edges in G form an angle of 180°, the sum is

$$v_{2,180} = y^2 + z^{\frac{1}{2}}y(2y^2 - 2) + zy^2(y^2 + y^{-2}), \tag{3.18}$$

and if the edges form an angle of 90° the sum is

$$v_{2,90} = y^2 + z^{\frac{1}{2}}y(2y^2 - 2) + zy^2(-2). \tag{3.19}$$

Using (3.12)–(3.19) gives

$$w(G) = v_0^N \sum_{G \subseteq L_{sq}} (v_{s,\theta}/v_0)^{p_{s,\theta}}, \tag{3.20}$$

where $p_{s,\theta}$ is the number of vertices of degree s in G and for $s = 2$, the angle θ must be distinguished. Because the lowest-order term in y in (3.14) is $z^{\frac{1}{2}}$, it can easily be verified from (3.15) through (3.19) that the lowest-order term in $v_{s,\theta}/v_0$ is $z^{s/2}$. Therefore, the lowest-order term in w(G) is z^e, where e is the number of edges in G. Thus, (3.20) gives a low-activity series expansion valid for a low density of ramrods on the square lattice.

4. APPLICATIONS OF THE WEAK-GRAPH METHOD USING HIGHER-ORDER VERTEX FORMS

A. The Spin-1 Ising Problem

Let the value of the spin σ_i (or vertex configuration ξ_i) at each vertex i of the lattice L_q take one of the three values $+1, 0, -1$. Then, the partition function is given by

$$Z(K, L) = \sum_{C} \prod_{\text{edges}} \exp(K\sigma_i\sigma_j) \prod_{i=1}^{N} \exp(L\sigma_i), \tag{4.1}$$

where $K = J/kT$ and $L = mH/kT$. It is impossible to write

$$\exp(K\sigma_i\sigma_j) = a + c(\sigma_i)c(\sigma_j), \tag{4.2}$$

so the binary-vertex form (1.5) cannot be applied to derive a weak-graph series expansion. However, it is possible to write

$$\exp(K\sigma_i\sigma_j) = 1 + \sigma_i\sigma_j \sinh K + \sigma_i^2\sigma_j^2(\cosh K - 1), \tag{4.3}$$

so the higher-order vertex form (1.8) may be applied where

$$c_1(\sigma_i) = \sigma_i(\sinh K)^{\frac{1}{2}} \tag{4.4}$$

and

$$c_2(\sigma_i) = \sigma_i^2(\cosh K - 1)^{\frac{1}{2}}, \tag{4.5}$$

$a = 1$, and $h = 2$ in (1.8).

Next, let us expand the multiple product

$$\prod_{\text{edges}} [1 + c_1(\sigma_i)c_1(\sigma_j) + c_2(\sigma_i)c_2(\sigma_j)]$$

into its $3^{qN/2}$ terms. Each term T corresponds to a subgraph G. If in T the factor $c_1(\sigma_i)c_1(\sigma_j)$ is taken from the ij-edge bracket

$$[1 + c_1(\sigma_i)c_1(\sigma_j) + c_2(\sigma_i)c_2(\sigma_j)]$$

in the multiple product, then edge ij is said to be a type-1 edge in G. If in T the factor $c_2(\sigma_i)c_2(\sigma_j)$ is taken from the ij-edge bracket in the multiple product, the edge ij is said to be a type-2 edge in G. Finally, if in T the factor 1 is taken from the ij-edge bracket, the edge ij is said not to be included in G. Thus, the subgraphs G may now have two different types of edges, type 1 and type 2.

Let us compute the weight w(G) of a weak subgraph G:

$$w(G) = \sum_{C} \prod_{\substack{\text{type 1} \\ \text{edges in G}}} c_1(\sigma_i)c_1(\sigma_j)$$
$$\times \prod_{\substack{\text{type 2} \\ \text{edges in G}}} c_2(\sigma_i)c_2(\sigma_j) \prod_{i=1}^{N} \exp(\sigma_i L)$$
$$= \prod_{i=1}^{N} \sum_{\sigma_i 0, \pm 1} \exp(\sigma_i L)[c_1(\sigma_i)]^{s_{1i}}[c_2(\sigma_i)]^{s_{2i}}, \tag{4.6}$$

where s_{1i} is the number of type-1 edges in G incident to the ith vertex and s_{2i} is the number of type-2 edges in G incident to the ith vertex. s_{1i} is conveniently called the type-1 degree of vertex i in G and s_{2i} is called the type 2 degree of vertex i in G. Using (4.4) and (4.5) in (4.6) gives us after some easy manipulation

$$w(G) = (1 + 2\cosh L)^N(\sinh K)^{e_1}(\cosh K - 1)^{e_2}$$
$$\times [2\cosh L/(1 + 2\cosh L)]^v(\tanh L)^{v(1,\text{odd})}, \tag{4.7}$$

where e_1 and e_2 are, respectively, the numbers of type-1 edges and type-2 edges in G, v is the total number of vertices in G, and v (1, odd) is the number of vertices in G with odd type-1 degree.

It should be noticed that, whereas sinh K has a leading power of $(1/kT)^1$, (cosh $K-1$) has $(1/kT)^2$ as its leading inverse power of temperature. Thus, the nth term in a $1/kT$ expansion may be obtained from knowing the subgraphs of L_q which have $e_1 + 2e_2 \leq n$. The subgraph weights in (4.7) simplify in zero field when $L = 0$; especially all weights vanish for graphs G such that $v(1, \text{odd}) \neq 0$. Thus, that part of the subgraph consisting only of type-1 edges must be closed. Nevertheless, the problem of counting the resulting subgraphs is not easy and it would appear quite difficult to obtain $Z(K, L)$ beyond the 11th term for the square lattice or the susceptibility series beyond the ninth term. This does not compare very well to the spin-$\frac{1}{2}$ case, where the reasonable goal is to obtain the 18th term in both series.[35]

B. The k-color Problem

A proper k-coloring of a lattice (or an arbitrary graph) is an assignment of one of k colors to each of the vertices of the graph under the restriction that no two vertices joined by an edge of the graph may have the same color.[25,36] The problem is to find the number of ways W_N to k-color a lattice (or graph). Of course, the outstanding problem of this type is the well-known four-color map conjecture. Although the general k-color problem is not of immediate interest in statistical mechanics, it seems worthwhile to consider because of its conceptual simplicity.

The case $k = 1$ is trivial. The case $k = 2$ may be expanded using closed graphs only via the binary-vertex form (1.5). This leads to the trivial bipartite or bicoloring theorem which states that a lattice (graph) can only be colored with two colors if there are no cycles of odd length in the lattice (graph),[25,36] i.e., the lattice must be loose packed.

For $k > 2$ one must resort to the higher-order-vertex form (1.8) in order to obtain a weak-graph expansion. However, this is very easily accomplished as follows: Let the vertex configurations be the colors $\xi_i = 1, 2, \cdots, k$. Then define

$$\begin{aligned} c_r(\xi_i) &= (-1)^{\frac{1}{2}} \quad \text{if} \quad \xi_i = r, \\ &= 0 \quad \text{otherwise}, \end{aligned} \tag{4.8}$$

for $r = 1, 2, \cdots, k$. Then, the number of proper k colorings W_N is given by

$$W_N = \sum_{\mathrm{C}} \prod_{\text{edges}} \left[1 + \sum_{r=1}^{k} c_r(\xi_i) c_r(\xi_j) \right]. \tag{4.9}$$

The graphs G in the weak-graph expansion of (4.9) will have k types of edges depending on which factor, $c_r(\xi_i)c_r(\xi_j)$ or 1, is used in the expansion of the term corresponding to G. But it will be noticed that $c_r(\xi_i)c_s(\xi_i) = 0$ unless $r = s$. Therefore, each connected component of G must have all the same kind of edges. Because of this it is reasonable to suppress the distinction between different types of edges so the subgraphs G may be thought of as weak graphs with only one kind of edge. But now we must keep in mind that each of the t-separated components of G could arise in k different ways from the series expansion. Thus, the weight of a weak graph will be

$$\begin{aligned} \mathrm{w(G)} &= k^t \sum_{\mathrm{C}} \prod_{\text{edges in G}} c_1(\xi_i) c_1(\xi_j) \\ &= k^t \prod_{i=1}^{N} \sum_{\xi_i=1}^{k} [c_1(\xi_i)]^{s_i} = (-1)^e k^{N-v+t}, \end{aligned} \tag{4.10}$$

where e, v, t, and s_i are the numbers of edges, vertices, separated components, and degree of vertex i in G. If we define $\mathrm{g}(\mathbf{p}, t)$ to be the number of subgraphs of L_q with partial description $\mathbf{p}$ and t separated components, then

$$W_N = k^N \sum_{\mathbf{p},t} \mathrm{g}(\mathbf{p}, t)(-1)^e (1/k)^{v-t}. \tag{4.11}$$

The result (4.11) was first obtained by Birkhoff.[24] Whitney[37] rederived (4.11) using the inclusion–exclusion principle. Whitney also simplified the subgraph counting somewhat by eliminating subgraphs which contained any broken cycles, such as the closed subgraphs. Thus, the subgraphs which Whitney counts are all nonclosed and quite frequently separated, and therefore hard to count. It would be preferable if the weak $\mathrm{g}(\mathbf{p}, t)$ could be obtained from the closed weak $\mathrm{g}(\mathbf{p}, t)$. Until this becomes possible, the series expansion (4.11) is not as valuable for obtaining lengthy series for lattice coloration estimates as one might wish.[38]

ACKNOWLEDGMENTS

The author wishes to thank Professor M. E. Fisher and Professor H. N. V. Temperley for very helpful discussion and communication.

Also, the support of the Advanced Research Agency through the Materials Science Center at Cornell University is gratefully acknowledged.

[35] M. F. Sykes (private communication).

[36] C. Berge, *The Theory of Graphs*, A. Doigt, trans. (John Wiley & Sons, Inc., New York, 1962).

[37] H. Whitney, Bull. Am. Math. Soc. **38**, 572 (1932).

[38] It may also be mentioned that it is possible to obtain a different closed-graph expansion which will enable us to obtain longer series, although the weight problem is not solved explicitly. This will be the subject of another paper.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 7 JULY 1968

Combinatorial Theorem for Graphs on a Lattice

J. F. NAGLE*
Baker Laboratory, Cornell University, Ithaca, New York
AND
H. N. V. TEMPERLEY
Department of Applied Mathematics, University College of Swansea, Swansea, Wales, U.K.

(Received 26 April 1967)

Several problems in lattice statistical mechanics, such as the spin-$\frac{1}{2}$ Ising problem and the monomer–dimer problem, can be formulated in terms of the **p**-generating function for the weak subgraphs of a regular lattice. This paper presents an algebraic transformation theorem which allows the **p**-generating function for the weak subgraphs of a lattice to be determined from the **p**-generating function for the far less numerous subset consisting of the closed weak subgraphs. This result will be especially useful in reducing the labor required to obtain exact finite series for various problems. The theorem also enables one to give a straightforward proof of the Ising susceptibility graph theorem due to Sykes.

1. INTRODUCTION

A series-expansion method for lattice statistical-mechanical problems has been developed by one of the authors. This method transforms the particular physical problem under consideration into a purely graph-combinatorial one of determining the number of specified subgraphs of an underlying graph or lattice. The prototype of this method is the hyperbolic-tangent expansion of the spin-$\frac{1}{2}$ Ising model.[1] More recent applications have been made to the residual entropy-of-ice problem[2] (and various other hydrogen-bonded problems[3]) and to the monomer–dimer problem.[4] In addition, a number of other model problems can be treated using this method. A discussion of these other problems and a more complete exposition of the general method, its limitations and modifications to suit particular problems, is given in another paper.[5]

The purpose of this paper is to present one of the more powerful and abstract results of the method. This result is a purely graph-combinatorial one in that the starting point is a graph-combinatorial problem and not a physical problem. This initial problem is to find what will be defined as the **p**-generating function for the weak subgraphs of a regular graph; weak subgraphs are composed of links (or edges) with no restrictions. (In Ref. 5 it is shown how many physical problems can be formulated as **p**-generating functions for weak subgraphs.) The closed weak subgraphs are a subset of the weak subgraphs under the restriction that they have no vertices of degree one, that is, no vertices which are joined to the subgraph by only one edge (or link). The closed weak subgraphs are far less numerous than the weak subgraphs. The main result of this paper is to show how the open weak-graph **p**-generating function can be obtained algebraically from the closed weak-graph **p**-generating function.

The theoretical significance of this result is to reinforce the view that the closed weak graphs are of fundamental importance for lattice problems. Some progress in this direction has been made by Sykes.[6] In particular, he showed that the self-avoiding random-walk generating function could be expressed entirely in terms of closed graphs, and he conjectured a similar theorem for the zero-field susceptibility of the spin-$\frac{1}{2}$ Ising series. The use of the theorem of this paper allows us to prove the Sykes susceptibility theorem in a straightforward way. The theorem of this paper is also related to some earlier work of Temperley.[7]

The practical significance of this result is that it makes it feasible to obtain more terms in series expansions for various physical problems which have required the numerous weak graphs. For example, the spin-$\frac{1}{2}$ high-temperature Ising-model series in a magnetic field is just a special case of the weak-graph **p**-generating function.[1] Previously, the most efficient method for obtaining this series has used an Ising transformation; eleven terms have been obtained in this way for the plane square lattice.[8] The method of this paper should allow one to obtain the square lattice series up to at least fifteen terms without a great deal of extra labor; eventually eighteen terms should be obtainable. This should permit a more detailed Padé-approximant study of the antiferromagnetic phase boundary for the spin-$\frac{1}{2}$ Ising model.

* Present address: Dept. of Physics, Carnegie-Mellon University, Pittsburgh, Pa. 15213.

[1] See, for example, C. Domb, Advan. Phys. **9**, 149 (1960), p. 175.
[2] J. F. Nagle, J. Math. Phys. **7**, 1484 (1966).
[3] J. F. Nagle, J. Math. Phys. **7**, 1492 (1966).
[4] J. F. Nagle, Phys. Rev. **152**, 190 (1967).
[5] J. F. Nagle, J. Math. Phys. **9**, 1007 (1968), preceding paper.
[6] M. F. Sykes, J. Math. Phys. **2**, 52 (1961); Physica **28**, 919 (1962).
[7] H. N. V. Temperley, Proc. Phys. Soc. **74**, 432 (1959).
[8] A. Bienenstock, J. Appl. Phys. **37**, 1459 (1966).

The outline of the paper is as follows. The basic notation, clarification of the problem, and the statement of the theorem and two corollaries are given in Sec. 2. The theorem and corollaries are derived in Sec. 3. To transform the weak-graph **p**-generating function straightforwardly to the closed weak-graph **p**-generating function requires solving a qth order algebraic equation explicitly. Section 4 shows how this difficulty may be circumvented so that the transformation may be effectively carried out using a simple algebraic iterative procedure. Finally, in Sec. 5 the transformation theorem of this paper is compared with earlier work of Temperley, and the Sykes susceptibility theorem is proved.

2. THE THEOREM

Let us clarify the problem and present some terminology. We are concerned with finding the number of ways that a graph G occurs as a subgraph of a certain underlying graph L_q which has N vertices. In statistical mechanics this underlying graph L_q is usually a regular lattice with periodic boundary conditions. However, the results apply also when L_q is an arbitrary regular graph in which each of the N vertices is joined by q edges to other vertices.[9] The symbol q is called the coordination number and, for example, is equal to four for the square lattice or the octahedron.

We also wish to partially describe the graphs G which occur as subgraphs of the graph L_q. Let **p** be a $(q+1)$-tuple or vector $\mathbf{p} = (p_0, p_1, \cdots, p_s, \cdots, p_q)$. We shall partially describe G by a particular **p** such that p_s is the number of vertices in G which have degree s; that is, p_s is the number of vertices of G which are joined by s edges to other vertices in G.

All the subgraphs G of L_q which are considered in this paper will be weak subgraphs. A weak subgraph G of L_q consists primarily of a set of edges of L_q. Secondly, it is necessary to specify the vertices of G. Normally a vertex of L_q is said to be contained in G if and only if it is incident to an edge in G; thus there would be no vertices of degree zero in a weak subgraph G. However, in this paper it is convenient to count all the vertices of L_q which are not contained in G (which have degree zero in G) as p_0. Thus

$$\sum_{s=0}^{q} p_s = N,$$

where G is described (partially) by **p**.

Next let g(**p**) be the number of graphs G with description **p** which are subgraphs of L_q. Then, we define the weak-graph **p**-generating function to be

$$\mathrm{F}(\mathbf{z}) = \sum_{\mathbf{p}} \mathrm{g}(\mathbf{p}) \prod_{s=0}^{q} z_s^{p_s} = \sum_{\mathbf{p}} \mathrm{g}(\mathbf{p})\mathrm{w}(\mathbf{p}, \mathbf{z}), \quad (2.1)$$

where $\mathbf{z} = (z_0, z_1, z_2, \cdots, z_q)$.

The **p**-generating function contains much useful information. For example, by setting

$$z_k = \tanh(H\mu/kT)[\tanh(J/kT)]^{k/2}$$

when k is odd, and $z_k = [\tanh(J/kT)]^{k/2}$ when k is even, F(**z**) becomes the high-temperature expansion for the spin-$\frac{1}{2}$ Ising problem in a magnetic field. However, the **p**-generating function does not contain all the information of interest in physical problems. For certain problems it would be helpful to decompose each g(**p**) according to the number of separated components in the **p** subgraphs, but our theorem is of no help here. (On the other hand, it should be remembered that some loss of knowledge is necessary to prevent the subgraph information from becoming extremely unwieldy.)

Now we may state the theorem.

Theorem:

$$\mathrm{F}(\mathbf{v}) = (1 + y^2)^{qN/2}\mathrm{F}(\mathbf{z}), \quad (2.2)$$

where

(i) y is a free variable,

(ii) $\mathbf{v} = (v_0, v_1, v_2, \cdots, v_q)$, analogous to **z**, and

(iii) $$v_s = \sum_{r=0}^{q} z_r \sum_{t=0}^{r} (-1)^t \binom{q-s}{r-t}\binom{s}{t} y^{r+s-2t}, \quad (2.3)$$

where

$$\binom{m}{n} = 0, \quad \text{if} \quad n < 0 \quad \text{or} \quad n > m.$$

It will be noticed that the same generating function F appears on both sides of (2.2). Because of this, it is not immediately obvious that it is possible to obtain the g(**p**) for the unrestricted weak graphs from the g(**p**) for the closed weak graphs, defined by $p_1 = 0$. However, it is always possible to fix the free variable y so that $v_1 = 0$. Then the left-hand side of (2.2) is the **p**-generating function just for the closed graphs from which F(**z**) for the unrestricted graphs may then be obtained. In practice it is not possible to solve $v_1(y) = 0$ explicitly in terms of **z** for large q. However, it is still possible to obtain the g(**p**) for the unrestricted graphs from the g(**p**) for the closed graphs by the simple iterative method described in Sec. 4.

It may also be noted that this transformation includes the monomer–dimer transformation as a special case.[4] This may be seen by setting $z_r = 0$ for

[9] The regularity may be dropped, but then a local coordination number must be introduced and the notation becomes too unwieldy to justify inclusion of this generalization here.

$r \geq 2$. In this case $v_1(y)$ is only a quadratic equation and may be solved explicitly.

The theorem (2.2) may also be thought of as describing what happens to a function $\mathrm{F}(\mathbf{z})$ when the space on which F is defined undergoes a transformation $\mathbf{T}_y$ where $\mathbf{v} = \mathbf{T}_y\mathbf{z}$. The transformation $\mathbf{T}_y$ (for arbitrary y) has the property that $\mathbf{T}_y^2 = (1 + y^2)^q\mathbf{I}$. Thus $\mathbf{T}_y$ may be redefined so that it is its own inverse. Further, a composition (multiplication) may be defined so that the set of $\mathbf{T}_y$ (plus a reflection $\mathbf{R}$) form a group: $\mathbf{T}_x \circ \mathbf{T}_y$ is given by $\mathbf{R}\mathbf{T}_w$ and $w = (y - x)/(1 + xy)$.

It is also useful to have the following corollaries to the theorem.

Corollary 1:

$$\sum_{s=0}^{q-1}(q - s)z_{s+1}\frac{\partial \mathrm{F}(\mathbf{z})}{\partial z_s} = \sum_{s=1}^{q} sz_{s-1}\frac{\partial \mathrm{F}(\mathbf{z})}{\partial z_s}. \qquad (2.4)$$

Dividing (2.4) by $N\mathrm{F}(\mathbf{z})$, we obtain another corollary.

Corollary 2:

$$\sum_{s=0}^{q-1}(q - s)z_{s+1}\frac{\partial}{\partial z_s}\frac{1}{N}\log \mathrm{F}(\mathbf{z}) = \sum_{s=1}^{q} sz_{s-1}\frac{\partial}{\partial z_s}(1/N)\log \mathrm{F}(\mathbf{z}), \qquad (2.5)$$

where

$$(1/N)\log \mathrm{F}(\mathbf{z}) = \log z_0 + \sum_{\mathbf{p}} \mathrm{g}_1(\mathbf{p})\prod_{s=1}^{q}\left(\frac{z_s}{z_0}\right)^{p_s}. \qquad (2.6)$$

When L_q is a lattice with periodic boundaries, $\mathrm{g}_1(\mathbf{p})$ is the coefficient of N^1 in $\mathrm{g}(\mathbf{p})$ (see Ref. 1).

These corollaries exhibit the result in a particularly simple and appealing way. In fact, one may prove Corollary 1 in a straightforward manner. The idea is to consider all graphs G with e edges in F. Then, the left-hand side of (2.4) tells one in how many ways one can add a "half-edge" to G. (A "half-edge" is a line ending at one vertex only. If we add a half-edge to a vertex of degree s, we change it to a vertex of degree $s + 1$, leaving all others unchanged. Of course, the half-edge may be attached to a vertex of degree zero.) Now consider graphs G′ in F which have $e + 1$ edges. The right-hand side of (2.4) tells one in how many ways one can delete a "half-edge" from a graph G′. Clearly the number of pseudographs with $e + \frac{1}{2}$ edges must be the same according to both computations. Although this proof is fully rigorous, we present a more systematic way of proving the theorem and the corollaries in the next section.

It may also be mentioned that Corollary 2 is in the most useful form for applications to statistical mechanics, since it is usually the free energy per vertex, $A/N = -(kT/N)\log \mathrm{F}(\mathbf{z})$ which is required.

3. PROOF OF THE THEOREM AND COROLLARIES

We suppose the graph to be decomposed into individual vertices by dividing each edge. Each vertex then has $0, 1, \cdots, q$ half-edges attached to it. For each vertex i there are 2^q *vertex configurations* ξ_i as shown in Fig. 1 for $q = 3$. The first vertex configuration ξ_i corresponds to a vertex of degree zero. The next q vertex configurations correspond to a vertex with one half-edge attached to it. The next $q(q - 1)/2$ vertex configurations correspond to a vertex with two half-edges, and so on. Next, we define a *labeling function*

$$b(\xi_i) = z_r y^r, \quad \text{if } \xi_i \text{ has } r \text{ half-edges.} \qquad (3.1)$$

Also, for each edge ij in L_q we define a *half-edge function* on ξ_i or ξ_j:

$$\begin{aligned} c_{ij}(\xi_i) &= y, \text{ if } \xi_i \text{ has no half-edge on edge } ij \text{ of } \mathrm{L}_q, \\ &= -y^{-1}, \text{ if } \xi_i \text{ has a half-edge on edge } ij \text{ of } \mathrm{L}_q. \end{aligned} \qquad (3.2)$$

Now consider

$$Q = \sum_{\mathrm{C}} \prod_{\text{edges}} [1 + c_{ij}(\xi_i)c_{ji}(\xi_j)] \prod_{i=1}^{N} b(\xi_i), \qquad (3.3)$$

where the first product is over all edges ij in L_q, the second product is over all vertices in L_q, and the summation is a 2^{qN}-fold summation over all combinations of vertex configurations for all N vertices. Each term in this summation is conveniently described as a *vertex complexion.*

First we observe that

$$\begin{aligned} 1 + c_{ij}(\xi_i)c_{ji}(\xi_j) &= 1 + y^2, && \text{if neither } \xi_i \text{ nor } \xi_j \text{ has a half-edge on edge } ij, \\ &= 1 + y^{-2}, && \text{if } \xi_i \text{ and } \xi_j \text{ each have a half-edge on edge } ij, \\ &= 0, && \text{otherwise.} \end{aligned} \qquad (3.4)$$

Thus a particular one of the 2^q vertex complexions gives a nonzero contribution to Q if and only if it corresponds to a permissible subgraph G of L_q. (That is, dangling half-edges are not allowed in Q.)

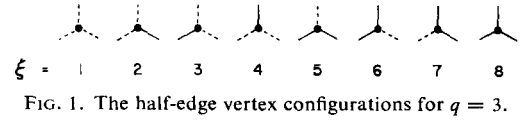


FIG. 1. The half-edge vertex configurations for $q = 3$.

Now, for a particular vertex complexion corresponding to a subgraph G with e edges and partial description $\mathbf{p}$ we have

$$\prod_{\text{edges}} [1 + c_{ij}(\xi_i)c_{ji}(\xi_j)] = (1 + y^{-2})^e(1 + y^2)^{(qN/2)-e}. \tag{3.5}$$

Also from (3.1) we have

$$\prod_{i=1}^{N} b(\xi_i) = y^{2e} \prod_{s=0}^{q} z_s^{p_s} = y^{2e}\mathrm{w}(\mathbf{p}, \mathbf{z}) \tag{3.6}$$

because there must be $2e$ half-edges in a graph with e edges. Therefore, summing over all subgraphs $G \subseteq L_q$, we have

$$Q = \sum_{G \subseteq L_q} (1 + y^2)^{qN/2}\mathrm{w}(\mathbf{p}, \mathbf{z}) = (1 + y^2)^{qN/2} \sum_{\mathbf{p}} \mathrm{g}(\mathbf{p})\mathrm{w}(\mathbf{p}, \mathbf{z}) = (1 + y^2)^{qN/2}\mathrm{F}(\mathbf{z}). \tag{3.7}$$

This is the first half of the theorem.

Next we obtain the second half of the theorem. Consider the expansion of $\prod_{i<j} [1 + c_{ij}(\xi_i)c_{ji}(\xi_j)]$ into its $2^{qN/2}$ terms. Each term corresponds to a weak subgraph G of L_q as follows: The subgraph G is said to contain the edge ij if and only if the term corresponding to G in the expansion has the factor

$$c_{ij}(\xi_i)c_{ji}(\xi_j).$$

Now we wish to obtain the contribution to Q of the term corresponding to G. This will be called the weight w(G) of G. Because each $c_{ij}(\xi_i)c_{ji}(\xi_j)$ is factorable the factors dependent on ξ_i may be grouped together for each i, and we may write

$$\mathrm{w}(\mathrm{G}) = \prod_{i=1}^{N} \sum_{\xi_i=1}^{2^q} b(\xi_i) \prod_j c_{ij}(\xi_i), \tag{3.8}$$

where the last product is over all j such that the edge ij is in G. The summation in Eq. (3.8) may be performed for an arbitrary vertex i of degree s in G; the result is v_s given by Eq. (2.3). To show this we first consider the vertex i to be of degree zero, i.e., $\xi_i = 1$. Then $b(\xi_i) = z_0$ and all s of the $c_{ij}(\xi_i) = y$. In general, for all $\binom{q}{r}$ of the ξ_i with r half-edges, $b(\xi_i) = z_r$. Also there are $\binom{q-s}{r-t}\binom{s}{t}$ of the ξ_i with r half-edges which have precisely t of their half-edges on the s edges of G incident to vertex i. This explains the general term in v_s in Eq. (2.3). Thus

$$\mathrm{w}(\mathrm{G}) = \prod_{s=0}^{q} v_s^{p_s} = \mathrm{w}(\mathbf{p}, \mathbf{v}), \tag{3.9}$$

where $\mathbf{p}$ describes G. By the definition of $\mathrm{w}(\mathbf{p}, \mathbf{v})$,

$$Q = \sum_{\mathbf{p}} \mathrm{g}(\mathbf{p})\mathrm{w}(\mathbf{p}, \mathbf{v}) = \mathrm{F}(\mathbf{v}), \tag{3.10}$$

and the theorem, Eq. (2.2), follows from Eqs. (3.10) and (3.7).

We next give a formal proof of the corollaries. Starting with Eq. (2.2), we take the derivative with respect to y, holding $\mathbf{z}$ fixed, and then take $y = 0$. Thus we have

$$\sum_{s=0}^{q} \frac{\partial v_s}{\partial y} \frac{\partial \mathrm{F}(\mathbf{v})}{\partial v_s}\bigg|_{y=0} = 0. \tag{3.11}$$

From Eq. (2.3) we find that

$$\frac{\partial v_s}{\partial y}\bigg|_{y=0} = (-1)^s[(q - s)z_{s+1} - sz_{s-1}] \tag{3.12}$$

and

$$v_s|_{y=0} = (-1)^s z_s.$$

Therefore, remembering that $\sum_{\text{odd } s} p_s$ is always even, Eq. (3.11) becomes

$$\sum_{s=0}^{q} [(q - s)z_{s+1} - sz_{s-1}] \frac{\partial \mathrm{F}(\mathbf{z})}{\partial z_s} = 0, \tag{3.13}$$

and Corollary 1 is proved.

Corollary 2 follows by dividing Eq. (3.13) by $N\mathrm{F}(\mathbf{z})$. Also, if L_q is a lattice with periodic boundaries so that every $\mathrm{g}(\mathbf{p})$ is a polynomial in N with lowest power N^1, then, by expanding the exponential function of N times the right-hand side of Eq. (2.6), one sees that $\mathrm{g}_1(\mathbf{p})$ is the coefficient of N^1 in $\mathrm{g}(\mathbf{p})$.

This method of proving the corollaries, especially when setting y equal to zero, suggests that they may be weaker than the theorem. However, it is possible to show that Eq. (2.5) holds also when $\mathrm{F}(\mathbf{z})$ is replaced by $\mathrm{F}(\mathbf{v})$ for *any* value of y. Then one may retrace one's steps to regain Eq. (2.2). Therefore Corollary 2 is actually equivalent to the theorem, and Corollary 1 undoubtedly contains the same amount of information.

4. ITERATIVE SOLUTION

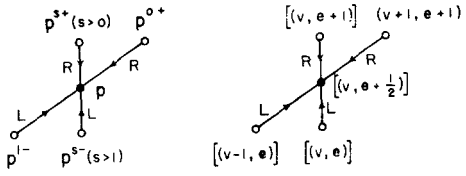
The purpose of this section is to describe a simple iterative procedure for obtaining numbers from the preceding theoretical results. As was mentioned in Sec. 2, the necessity for having some such procedure is due to our inability to solve $v_1(y) = 0$ for general $\mathbf{z}$ and q.

First let us define $v = \sum_{s=1}^{q} p_s$ to be the number of proper vertices in subgraphs G. (See the discussion at the beginning of Sec. 2.) Also let us rewrite Eq. (2.4) as

$$[L(\mathbf{z}) - R(\mathbf{z})]\mathrm{F}(\mathbf{z}) = 0, \tag{4.1}$$

where $R(\mathbf{z})$ and $L(\mathbf{z})$ are the differential operators which appear on the right- and left-hand sides, respectively, of Eq. (2.4). If we relate $\mathbf{p}^{s-}$ to $\mathbf{p}$ by

$$\mathbf{p}^{s-} = (p_0, \cdots, p_{s-1} + 1, p_s - 1, \cdots, p_q), \tag{4.2}$$



(2a) (2b)

FIG. 2. Figure 2(a) is a schematic representation of Eq. (4.6), and Fig. 2(b) is a schematic representation of Eq. (4.7).

then

$$R(\mathbf{z})\mathrm{w}(\mathbf{p}, \mathbf{z}) = \sum_{s=0}^{q} s p_s \mathrm{w}(\mathbf{p}^{s-}, \mathbf{z}). \tag{4.3}$$

Also let

$$\mathbf{p}^{s+} = (p_0, \cdots, p_s - 1, p_{s+1} + 1, \cdots, p_q). \tag{4.4}$$

Then

$$L(\mathbf{z})\mathrm{w}(\mathbf{p}, \mathbf{z}) = \sum_{s=0}^{q} (q - s) p_s \mathrm{w}(\mathbf{p}^{s+}, \mathbf{z}). \tag{4.5}$$

As was mentioned in Sec. 2, if $\mathbf{p}$ describes graphs with e edges, then $\mathbf{p}^{s+}$ and $\mathbf{p}^{s-}$ in Eqs. (4.2) and (4.4) describe "pseudographs" with $e + \frac{1}{2}$ and $e - \frac{1}{2}$ edges, respectively. Also, if $\mathbf{p}$ describes graphs with v (proper) vertices, then $\mathbf{p}^{s+}$ for $s \geq 1$ and $\mathbf{p}^{s-}$ for $s \geq 2$ describe pseudographs with v vertices; but $\mathbf{p}^{0+}$ and $\mathbf{p}^{1-}$ describe pseudographs with $v + 1$ and $v - 1$ (proper) vertices, respectively.

Let us consider a particular $\mathbf{p}$ which now is taken to describe pseudographs with $e + \frac{1}{2}$ edges and v vertices. By Eqs. (4.2)–(4.5), the coefficient of $\mathrm{w}(\mathbf{p}, \mathbf{z})$ in Eq. (4.1), which must be zero, comes from $L(\mathbf{z})$ operating on $\sum_s \mathrm{g}(\mathbf{p}^{s-})\mathrm{w}(\mathbf{p}^{s-}, \mathbf{z})$ and $R(\mathbf{z})$ operating on $\sum_s \mathrm{g}(\mathbf{p}^{s+})\mathrm{w}(\mathbf{p}^{s+}, \mathbf{z})$, where again $\mathbf{p}^{s-}$ and $\mathbf{p}^{s+}$ are given by Eqs. (4.2) and (4.4), but where they now describe graphs with e and $e + 1$ edges, respectively. Also $\mathbf{p}^{1-}$ has $v - 1$ vertices, $\mathbf{p}^{0+}$ has $v + 1$ vertices, and the other $\mathbf{p}^{s\pm}$ have v vertices. Figures 2(a) and 2(b) show this schematically. Now if we know all the $\mathrm{g}(\mathbf{p}^{s-})$ and all the $\mathrm{g}(\mathbf{p}^{s+})$ except for $\mathrm{g}(\mathbf{p}^{0+})$, then the latter can be calculated from $\mathrm{g}(\mathbf{p}^{0+}) = (p_1 + 1)^{-1}$ coeff of $\mathrm{w}(\mathbf{p}, \mathbf{z})$ in

$$L(\mathbf{z}) \sum_{s=0}^{q} \mathrm{g}(\mathbf{p}^{s-})\mathrm{w}(\mathbf{p}^{s-}, \mathbf{z}) - R(\mathbf{z}) \sum_{s=1}^{q} \mathrm{g}(\mathbf{p}^{s+})\mathrm{w}(\mathbf{p}^{s+}, \mathbf{z}). \tag{4.6}$$

It should be noticed that the g for closed weak graphs can never be obtained from the left-hand side of Eq. (4.6) because $\mathbf{p}^{0+}$ always has at least one vertex of degree one.

Since $\mathbf{p}^{0+}$ is an arbitrary nonclosed graph description within the set of $\mathbf{p}$'s with $v + 1$ vertices and $e + 1$ edges, it follows that all nonclosed $\mathrm{g}(\mathbf{p})$ where $\mathbf{p}$ has $v + 1$ vertices and $e + 1$ edges can be computed from Eq. (4.6), knowing only the $\mathrm{g}(\mathbf{p})$ where $\mathbf{p}$ has v vertices and e or $e + 1$ edges and the $\mathrm{g}(\mathbf{p})$ where $\mathbf{p}$ has $v - 1$ vertices and e edges. Let us represent the partial summation $\sum' \mathrm{g}(\mathbf{p})\mathrm{w}(\mathbf{p}, \mathbf{z})$ in the following ways, depending on the $\mathbf{p}$'s summed over:

(a) (v, e), if the summation is over all nonclosed $\mathbf{p}$ and where $\mathbf{p}$ has v (proper) vertices and e edges;

(b) $[v, e]$, if the summation is over all closed $\mathbf{p}$ where $\mathbf{p}$ has v vertices and e edges;

(c) $[(v, e)]$, if the summation is over all weak $\mathbf{p}$ where $\mathbf{p}$ has v vertices and e edges.

Using this notation, the preceding remark may be represented as

$$R(\mathbf{z})(v + 1, e + 1) = L(\mathbf{z})[(v - 1, e)] + L(\mathbf{z})[(v, e)] - R(\mathbf{z})[(v, e + 1)], \tag{4.7}$$

where the equality holds only for those coefficients of $\mathrm{w}(\mathbf{p}, \mathbf{z})$ where $\mathbf{p}$ has v vertices and $e + \frac{1}{2}$ edges. Knowing the right-hand side of Eq. (4.7) determines $(v + 1, e + 1)$ in exactly the same manner as $\mathrm{g}(\mathbf{p}^{0+})$ is determined from Eq. (4.6). The analogous relation is shown schematically in Fig. 2(b).

Now let us look at the broader picture. In Fig. 3 the possible weak-subgraph combinations of v and e for the triangular lattice are represented in coordinate form. Each filled circle, which represent pseudographs, is the terminus of four arrows [which have the same meaning as in Fig. 2(b)]. We now discuss three possible sequences which may be followed in determining the nonclosed $\mathrm{g}(\mathbf{p})$ or (v, e); the different sequences are appropriate to different problems.

(1) *The edge sequence.* Suppose we know $[(v_n, e)]$ and $[v_n, e + 1]$ for all v_n, and we wish to find the $(v_n, e + 1)$. One starts by computing the $(v_n, e + 1)$ for the smallest $v_n = v_{\min}$ which may always be done using Eq. (4.7). Then one computes $(v_{\min} + 1, e + 1)$ using $(v_{\min}, e + 1)$, and so on. This iteration may

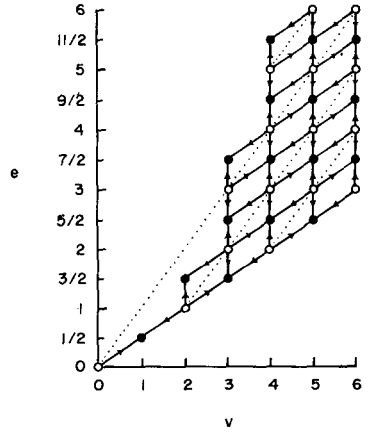


FIG. 3. The coordinate representation of the various $[(v, e)]$ which may occur on the triangular lattice (small v and e). The filled circles represent pseudographs and the open circles represent graphs. The arrows have the same meaning as in Fig. 2(b). The dotted lines show constant $e - v$ grouping.

always be followed in computing the $(v_n, e+1)$ since the right-hand side of Eq. (4.7) involves only sets $[(v_r, e_s)]$ which have either fewer than $e+1$ edges or fewer than v_n vertices.

(2) *The vertex sequence.* Suppose we know $[(v, e_n)]$ and $[v+1, e_n]$ for all e_n and we wish to find the $(v+1, e_n)$. One may compute the $(v+1, e_n)$ in any order, using Eq. (4.7), since the right-hand side depends only on sets with fewer than $v+1$ vertices.

(3) *The $e-v$ sequence.* Suppose we know $[(v_n, e_m)]$ and $[v_n - 1, e_m]$ for $e_m - v_n > e - v$, and we wish to find the (v_n, e_m) with $e_m - v_n = e - v$. One starts by computing (v_n, e_m) for the smallest v_n and e_m where $e_m - v_n = e - v$, using Eq. (4.7). (This is always trivially zero.) Then one computes $(v_{\min}+1, e_{\min}+1)$ using $(v_{\min}, e_{\min})$, and so on. This procedure may always be followed since the right-hand side of Eq. (4.7) involves only sets $[(v_r, e_s)]$ with $e_s - v_r \geq e_m - v_n$, and if the equality holds, then $v_r < v_n$.

To conclude this section it may be mentioned that this iterative procedure has been programmed for the computer. The input data consists of the coordination number q and the closed graph $\mathbf{g(p)}$. It is not necessary to specify the nonclosed $\mathbf{p}^{0+}$ since these are generated automatically by the iterative procedure. The largest case tested to date is for all $\mathbf{g(p)}$ of the square lattice with $e \leq 12$ or $v \leq 10$. The running time for this case is less than 15 sec.

5. RELATION TO EARLIER WORK

The use of vertex configurations and half-edges in this paper is reminiscent of the decoration transformation considered in connection with the spin-$\frac{1}{2}$ Ising problem by Temperley and others.[7,10] In the decoration transformation each edge on the lattice is decorated with a new spin which interacts only with the two spins connected by the edge. The original partition function can be transformed to a new partition function in which the configurations of the decorating spins are included also. This new partition function is given the following alternative definition. A $(+1)$ decorated edge spin is designated as an occupied edge and a (-1) decorated edge spin is designated as an unoccupied edge, thereby transforming the original partition function into a weak-graph series. This latter also depends upon a vertex activity-tuple $\mathbf{z}$ which is partially determined by the requirements of the original problem. For example, Temperley[7] showed that for $q = 3$

$$\exp(4J/kT) = 1 + (C/B^2), \tag{5.1}$$

and in connection with the present work it has been shown that

$$\cosh(2mH/kT) = (AC - 2B^3)/2(B^2 + C)^{\frac{3}{2}}. \tag{5.2}$$

The functions A, B, and C are defined as follows:

$$\begin{aligned} A &= z_0^2 + 3z_1^2 + 3z_2^2 + z_3^2, \\ B &= z_0z_2 - z_1^2 + z_1z_3 - z_2^2, \\ C &= (z_0z_3 - z_1z_2)^2 - 4(z_0z_2 - z_1^2)(z_1z_3 - z_2^2). \end{aligned} \tag{5.3}$$

It can easily be verified that A, B, and C are independent and that each of them can be substituted for F in Eq. (2.4). It follows from Lagrange's method that $\mathrm{F}(A, B, C)$ is the general solution of Eq. (2.4).

Since Eqs. (5.1) and (5.2) put two restrictions on $\mathbf{z}$, and since it is only the ratios $z_0:z_1:z_2:z_3$ which are important, there is one free relation in $\mathbf{z}$ which is analogous to the free variable y in the present paper. Temperley used this free relation in several ways; in particular, the usual weak-graph hyperbolic-tangent expansion and a weak-graph Mayer-like expansion were rederived. Another possibility not considered in that paper[7] is to set $z_1 = 0$, in which case the Ising problem is transformed in one step to a closed weak-graph expansion. This is in contrast to the two-step approach of this paper in which the hyperbolic tangent expansion first reduces the problem to the weak-graph combinatorial problem; then the method of this paper further reduces it to the closed weak-graph combinatorial problem. A difficulty with the one-step decoration transformation is the finding of restrictions analogous to Eqs. (5.1) and (5.2) for general q. The problem of solving for the z_s, even in a series expansion, is not simple either. Finally, the nonclosed weak $\mathbf{g(p)}$ provided by the method of this paper may be used for purposes other than the high-temperature spin-$\frac{1}{2}$ Ising problem. Thus the two-step approach of this paper seems to divide the problem in a useful way, in comparison to the earlier work.

A theorem due to Sykes transforms the particular weak-graph combinatorial problem associated with the Ising spin-$\frac{1}{2}$ high-temperature zero-field susceptibility into a closed weak-graph combinatorial problem.[6] This theorem was first conjectured by Sykes upon noticing certain regularities in the series; the only proof known is rather involved.[11] In contradistinction to this, the use of our theorem enables one to prove the susceptibility graph theorem in an exceedingly straightforward way, as we now show.

Let $w = \tanh K$ and $\tau = \tanh L$, where $K = J/kT$ and $L = mH/kT$. Then the zero-field susceptibility is proportional to

$$\chi_0(w) \propto (1/N)(\partial^2/\partial\tau^2) \log \mathrm{F}(\mathbf{z})\big|_{\tau=0}, \tag{5.4}$$

[10] M. E. Fisher, Phys. Rev. **113**, 969 (1959); S. Naya, Progr. Theoret. Phys. (Kyoto) **11**, 53 (1954).

[11] M. F. Sykes (private communication).

where

$$z_r = w^{r/2}, \quad \text{if } r \text{ is even},$$
$$= \tau w^{r/2}, \quad \text{if } r \text{ is odd}. \tag{5.5}$$

Let $\sigma = q - 1$. Then, from Eqs. (2.3) and (5.5), we have

$$v_1 = y + \tau w^{\frac{1}{2}}[\sigma y^2 - 1] + w[(\sigma(\sigma - 1)/2)y^3 - \sigma y] + \text{terms of order } y^2 \text{ or higher}. \tag{5.6}$$

We require that $v_1 = 0$ for all τ in order to have a closed weak-graph series-expansion. When $\tau = 0$, $y = 0$ solves Eq. (5.6). Expanding away from this solution for small τ, to first order in τ we have

$$y = \tau w^{\frac{1}{2}}/(1 - \sigma w). \tag{5.7}$$

Now let us write down the other v_s to order τ^2 or y^2, since this is all that is needed after taking $\partial^2/\partial\tau^2|_{\tau=0}$:

$$v_{2n+1} = w^n(2n + 1)y - \tau w^{n+\frac{1}{2}} - w^{n+1}(q - 2n - 1)y \tag{5.8}$$

and

$$v_{2n} = w^{n-1}\binom{2n}{2}y^2 - w^{n-\frac{1}{2}}2n\tau y + w^n(1 - 2n(q - 2n)y^2) + w^{n+\frac{1}{2}}(q - 2n)\tau y + w^{n+1}\binom{q - 2n}{2}y^2. \tag{5.9}$$

As τ and y approach 0, v_{2n+1} goes to zero linearly with τ or y, and v_{2n} approaches w^n.

Using Eq. (2.6), we have

$$(1/N)\log \mathrm{F}(\mathbf{z}) = -(q/2)\log(1 + y^2) + \log v_0 + (1/N)\log \mathrm{F}(\mathbf{x}), \tag{5.10}$$

where $\mathbf{x} = (x_0, \cdots, x_s, \cdots, x_q)$ and $x_s = v_s/v_0$. Performing the two differentiations on the first two terms on the right-hand side of Eq. (5.10) and setting $\tau = 0$ gives us

$$\chi_B = qw/(1 - \sigma w), \tag{5.11}$$

which is the Bethe approximation for the susceptibility. Next consider graphs in $\mathrm{F}(\mathbf{x})$ with vertices of odd degree. Since each vertex of odd degree goes to zero linearly as τ approaches 0, the contribution of such graphs to χ_0 will vanish unless each vertex factor x_{2n+1} is differentiated. Since there are only two differentiations available, there must be, at most, two vertices of odd degree. (Of course, there are no graphs with one vertex of odd degree.) Each of the differentiations yields

$$(\partial x_{2n+1}/\partial\tau)_{\tau=0} = w^{n+\frac{1}{2}}2n(1 + w)/(1 - \sigma w). \tag{5.12}$$

Remembering that we have two such derivatives to take, we find the weight for graphs with two odd vertices of order $2n_1 + 1$ and $2n_2 + 1$ to be

$$w^e(1 + w)^2 8n_1n_2/(1 - \sigma w)^2, \tag{5.13}$$

where the additional w factors come from the vertices of even degree.

Next consider the so-called nonmagnetic graphs with all vertices of even degree. Clearly, from Eq. (5.9), both differentiations must operate on the same factor x_{2n} to give a nonzero contribution to $\chi_0(w)$. Thus we compute

$$(\partial^2 x_{2n}/\partial\tau^2)_{\tau=0} = [w^n/(1 - \sigma w)^2] \times [2n(2n - 2)(1 + w)^2 - 4n(1 - w^2)]. \tag{5.14}$$

The contribution of all vertices must be considered to give us the following weight of nonmagnetic graphs:

$$[w^e/(1 - \sigma w)^2]\left[(1 + w)^2 \sum_{s=0}^{q} s(s - 2)p_s - 2(1 - w^2)e\right]. \tag{5.15}$$

We may now add together all the contributions to give

$$\chi_0(w) = (1 - \sigma w)^{-2}\left[qw(1 - \sigma w) - 2(1 - w^2)\sum_{\mathbf{p}}{}' \mathrm{g}(\mathbf{p})ew^e + (1 + w)^2\sum_{p}{}' \mathrm{g}(\mathbf{p})w^e \sum_{s} s(s - 2)p_s + 2(1 + w)^2 \sum_{\mathbf{p}}{}'' \mathrm{g}(\mathbf{p})w^e(s_i - 1)(s_j - 1)\right]. \tag{5.16}$$

The first term in the brackets comes from Eq. (5.11). The second term comes from the second part of Eq. (5.15); and, accordingly, the primed summation is only over nonmagnetic graphs; this term differs notationally from Sykes's $-2vU(v)$ by qw^2. The third term in Eq. (5.16) comes from the first part of Eq. (5.15). The final term in the brackets in Eq. (5.16) comes from Eq. (5.13), where now the summation is over all $\mathbf{p}$ with precisely two vertices of odd degrees s_i and s_j. It can now be verified that (5.16) is equivalent to the counting rules conjectured by Sykes, with the trivial difference of a constant term of unity.

It would also be possible to derive a similar theorem for higher zero-field derivatives with respect to τ. However, there is no longer much need for this kind of theorem, since the necessary graphical information is obtainable directly from the iterative solution described in Sec. 4.

ACKNOWLEDGMENT

The Materials Science Center of Cornell University is gratefully acknowledged for financial support of J. F. N. during the course of this work and for the use of the MSC Central Facilities for aid in preparing this paper for publication.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 7 JULY 1968

Rotation and Lorentz Groups in a Finite Geometry

E. G. BELTRAMETTI
Istituto di Fisica dell'Università di Genova, I.N.F.N., Sezione di Genova, Italy
AND
A. A. BLASI
Scuola di Perfezionamento in Fisica dell'Università di Genova, Italy

(Received 12 June 1967)

The introduction in physics of a finite geometry approximating the ordinary Euclidean one poses the problem of studying the relativity groups over such a geometry. We present a detailed analysis of the structure and irreducible representations of the rotation, Lorentz, and Poincaré groups. It is found that, besides the usual quantum numbers, a new two-valued label is necessary to specify the representations.

1. INTRODUCTION

It has been pointed out that, in principle, there are no fundamental objections to the idea of building up a description of physical phenomena by taking the space–time coordinates from a finite field instead of taking them, as usual, from the real-number field. The finite geometry so generated may contain a subset approximating, with an unlimited accuracy, the ordinary Euclidean geometry. The last has been experimentally tested from distances of about 10^{-14} cm to distances of about 10^{28} cm; the corresponding minimum order p of the finite field has been roughly estimated as $10^{10^{81}}$.[1]

The aim of the present paper is to investigate the general features of the simpler relativity groups in the framework of a finite geometry built over a finite field $GF(p)$. This problem has not been investigated in detail in the studies of various authors[2–4] who have attempted the introduction of finite geometries in the field of physics; an analysis of an extension of the four-dimensional orthogonal group has been done by Coish,[2] who tentatively suggested a geometrical interpretation of the electric-charge quantum number.

In Sec. 2 we quote some relevant properties of finite fields, also called Galois fields after Evariste Galois who first attempted their systematic study. Sections 3 and 4 are devoted to the analysis of the structure and representations of the rotation group, while in Secs. 5 and 6 the analysis is extended to the Lorentz and Poincaré groups.

[1] G. Jarnefelt, Ann. Acad. Sci. Fennicae, Ser. A, **I**, No. 96 (1951). The problem of defining a metric in a geometry over a finite field has been extensively examined by G. Jarnefelt and P. Kustaanheimo, Publ. Atron. Obs. Helsinki, No. 32, (1952); P. Kustaanheimo, Rendi. Mat. **16**, 286 (1957); Publ. Astron. Obs. Helsinki, No. 52 (1957) and Math. Scand. **5**, 197 (1957); Publ. Astron. Obs. Helsinki No. 54 (1957).

[2] H. R. Coish, Phys. Rev. **114**, 383 (1959).

[3] Y. Ahmavaara, J. Math. Phys. **6**, 87 (1965); **6**, 220 (1965); **7**, 197 (1966); **7**, 201 (1966).

[4] I. S. Shapiro, Nucl. Phys. **21**, 474 (1960).

2. GALOIS FIELDS[5]

Let us recall that a set of k elements $x_0, x_1, \cdots, x_{k-1}$ is said to form a finite field $GF(k)$ of order k if it is closed under two operations $x_i + x_j$ (addition) and $x_i x_j$ (multiplication), satisfying:

A. $x_i + x_j = x_j + x_i$,

$$x_i + (x_j + x_l) = (x_i + x_j) + x_l;$$

B. a unique element exists having the properties of "zero" under addition; without loss of generality we may identify it with x_0 and write $x_i + x_0 = x_i$, $\forall x_i$: the element x_0 will also be denoted by 0;

C. for every x_i, a unique element x_j exists such that $x_i + x_j = x_0$; the element x_j will also be denoted by $-x_i$;

D. $x_i x_j = x_j x_i$, $x_i(x_j x_l) = (x_i x_j) x_l$,

$$x_i(x_j \pm x_l) = x_i x_j \pm x_i x_l;$$

E. a unique element exists having the properties of "unity" under multiplication: without loss of generality, we may identify it with x_1 and write $x_i x_1 = x_i$, $\forall x_i$: the element x_1 will also be denoted by 1;

F. for every $x_i \neq x_0$ a unique element x_j exists such that $x_i x_j = x_1$: the element x_j will also be denoted by x_i^{-1} or $1/x_i$.

Finite fields exist only of order $k = p^n$, p being a prime, n an integer: moreover all finite fields of the same order are isomorphic so that a finite field is uniquely determined by the number p^n of its elements.

The elements of $GF(p^n)$ defined by the sequence

$$x_0, x_1, x_1 + x_1, \quad x_1 + x_1 + x_1, \cdots$$

are called integral marks and will be denoted by $0, 1, 2, 3, \cdots$; it may be shown that there are only p integral marks, namely $0, 1, 2, \cdots, p-1$. The notation mx, m being an integer and $x \in GF(p^n)$, used as a

[5] For an exhaustive treatment of the Galois field theory, see, e.g., L. Dickson, *Linear Groups* (Dover Publications Inc., New York, 1958).

short hand for $x + x + \cdots + x$ (m times), individuates the same element when m is considered an integral mark and mx is a product in $GF(p^n)$. The following relation holds:

$$px = 0, \quad \forall x \in GF(p^n). \tag{2.1}$$

Any element x of a finite field of order p^n satisfies the relation (Fermat's theorem)

$$x^{p^n} = x, \quad x \in GF(p^n). \tag{2.2}$$

Of course a relation $x^l = x$ with $l < p^n$ may be fulfilled for suitable x; however, in any $GF(p^n)$ there exists at least one element (called a primitive root) such that $1, \omega, \omega^2, \cdots, \omega^{p^n-1}$ are all different and span the whole field.

The elements of $GF(p^n)$ which are even powers, or respectively odd powers, of a primitive root are called "squares," or respectively, "not-squares"; the element x_0 is not recognized either as a square or a not-square. The number of squares is $\frac{1}{2}(p^n - 1)$ and the same is the number of not-squares.

Let us now briefly comment on the problem of building up an ordering in $GF(p^n)$, i.e., a notion of "greater than" and "smaller than" for the elements of a finite field. The property of being a square or a not-square is combined under multiplication according to the sign rule: in this respect, squares and not-squares behave like positive and negative numbers, respectively. Thus, for any two elements $x_i, x_j \in GF(p^n)$, we may state

$$\begin{aligned} &x_i \text{ is greater than } x_j \ (x_i > x_j), \\ &\qquad \text{if } x_i - x_j \text{ is a square,} \\ &x_i \text{ is smaller than } x_j \ (x_i < x_j), \\ &\qquad \text{if } x_i - x_j \text{ is a not-square.} \end{aligned} \tag{2.3}$$

The ordering so defined is in general not transitive; in fact from $x_i > x_j$, $x_j > x_l$ we cannot deduce, by Eq. (2.3), $x_i > x_l$. Though transitivity cannot be ensured for the whole field, it may be established for a subset E of $GF(p^n)$. This problem has been worked out in detail for the case $n = 1$, to which we shall mainly refer in the following since it avoids unessential complications and leads to the same kind of result (concerning the structure of the relativity groups and of their representations) which would be obtained for arbitrary n. It is easily seen that the field $GF(p)$ may be realized by the integers

$$-\frac{p-1}{2}, \ -\frac{p-3}{2}, \cdots, -1, 0, 1, \cdots, \frac{p-3}{2}, \ \frac{p-1}{2}, \tag{2.4}$$

assuming as addition and multiplication the usual arithmetical operations and taking the result "modulo p." Under suitable restrictions on the choice of the order p,[6] which are of the kind

$$p = 4l - 1, \quad l \text{ INTEGER}, \tag{2.5}$$

the transitivity of the ordering (2.3) may be ensured over a chain E of consecutive integers containing about $N = \ln p$ elements, and the element (-1) may be made not-square (this implies that if x is a square, $-x$ is a not-square and vice versa). The finite geometry constructed by taking the coordinates in $GF(p)$ will contain, corresponding to the subset E, a finite lattice which may approximate the observed physical space; by allowing N to be suitably large, the accuracy of this approximation could be made arbitrarily large.

We now want to point out that when the real number field is replaced by $GF(p)$, the complex number field is most naturally replaced by $GF(p^2)$. Of course $GF(p) \subset GF(p^2)$ and it may be shown that all elements of $GF(p)$ are squares in $GF(p^2)$. This implies that the integral marks are squares in $GF(p^2)$. In particular, the element (-1) which is not-square in $GF(p)$ becomes a square in $GF(p^2)$: thus there exists an element of $GF(p^2)$, formally denoted by i, such that

$$\begin{aligned} &i^2 = -1, \quad i \in GF(p^2), \\ &\qquad (-1) \text{ not-square in } GF(p). \end{aligned} \tag{2.6}$$

Now consider the set of p^2 pairs (x, y), $x \in GF(p)$, $y \in GF(p)$ and introduce the two operations:

$$\begin{aligned} (x_1, y_1) + (x_2, y_2) &= (x_1 + x_2, y_1 + y_2), \\ (x_1, y_1)(x_2, y_2) &= (x_1x_2 - y_1y_2, x_1y_2 + x_2y_1); \end{aligned} \tag{2.7}$$

it is easily verified that this set, with these operations, is a finite field, and it is actually $GF(p^2)$, owing to the uniqueness of Galois fields of given order. Due to Eq. (2.6), the elements of $GF(p^2)$ may also be written as

$$z = x + iy, \quad z \in GF(p^2), \quad x, y \in GF(p) \tag{2.8}$$

with the operations:

$$\begin{aligned} z_1 + z_2 &= x_1 + x_2 + i(y_1 + y_2), \\ z_1z_2 &= x_1x_2 - y_1y_2 + i(x_1y_2 + x_2y_1). \end{aligned} \tag{2.7'}$$

In order to define complex conjugation, one first remarks that from (2.2), (2.5), and (2.6), it follows that

$$i^p = -i. \tag{2.9}$$

[6] P. Kustaanheimo, Soc. Sci. Fennicae, Comm. Phys. Math. **15**, 19 (1950).

Furthermore, one gets

$$(x+iy)^p = x^p + (iy)^p + \sum_{m=1}^{p-1}\binom{p}{m}x^m(iy)^{p-m} = x - iy + \sum_{m=1}^{p-1} p\left[\frac{(p-1)\cdots(p-m+1)}{m!}\right] \times x^m(iy)^{p-m}, \quad (2.10)$$

where the square bracket appearing in the last sum is an integer q since p is prime and $\binom{p}{m}$ is integer: hence from Eq. (2.1),

$$(x+iy)^p = x - iy. \quad (2.11)$$

From Eq. (2.2), it follows that $(x^p)^p = z^{p^2} = z$ and one may state, according to Eq. (2.11), that in $GF(p^2)$ the complex conjugate z^* of an element z is its pth power:

$$z^* = z^p. \quad (2.12)$$

We finally quote some results about quadratic equations which will be needed in the next sections.

The number ν of sets of solutions $(x_1, x_2, \cdots, x_{2m})$ in $GF(p^n)$ of the equation,

$$c_1x_1^2 + c_2x_2^2 + \cdots + c_{2m}x_{2m}^2 = d, \quad m \text{ integer}, \quad (2.13)$$

where c_i and d are nonzero elements of $GF(p^n)$, is given by

$$\nu = p^{n(2m-1)} - \eta p^{n(m-1)}, \quad (2.14)$$

where η is $+1$ or -1 according as $(-1)^m c_1c_2\cdots c_{2m}$ is a square or a not-square in $GF(p^n)$.[7] In particular, we notice that the equation $x_1^2 + x_2^2 = -1$ has $p+1$ solutions in $GF(p)$: the possibility of solving such an equation is connected with the lack of ordering over the whole field $GF(p)$.

The number ν of sets of solutions $(x_1, x_2, \cdots, x_{2m+1})$ in $GF(p^n)$ of the equation,

$$c_1x_1^2 + c_2x_2^2 + \cdots + c_{2m+1}x_{2m+1}^2 = d, \quad m \text{ integer}, \quad (2.15)$$

where c_i and d are nonzero elements of $GF(p^n)$, is given by

$$\nu = p^{2nm} + \omega p^{nm}, \quad (2.16)$$

where ω is $+1$ or -1 according as

$$(-1)^m\, dc_1c_2\cdots c_{2m+1}$$

is a square or a not-square in $GF(p^n)$.[7]

3. PROPER ROTATION GROUP

In the three-dimensional space, whose points have coordinates $x_1, x_2, x_3 \in GF(p)$, consider the group of invertible linear substitutions,

$$x_i' = \sum_j r_{ij}x_j, \quad r_{ij} \in GF(p), \quad (3.1)$$

satisfying

$$x_1'^2 + x_2'^2 + x_3'^2 = x_1^2 + x_2^2 + x_3^2. \quad (3.2)$$

Denoting by r the 3×3 matrix whose elements are the coefficients r_{ij} introduced in (3.1), it is immediately seen that Eq. (3.2) implies and is implied by

$$r^T r = 1 \quad (T \text{ means transposition}). \quad (3.3)$$

From Eq. (3.3) we get $\det r = \pm 1$; the subgroup corresponding to $\det r = +1$ will be denoted by $R(3,p)$ and called proper rotation group. The order $\Omega_{R(3,p)}$ of $R(3,p)$ is found to be

$$\Omega_{R(3,p)} = p(p^2-1). \quad (3.4)$$

$R(3,p)$ may be generated by the transformations[8]:

$$O_{i,j}^{(a,b)} = \begin{cases} x_i' = ax_i + bx_j, \\ x_j' = -bx_i + ax_j, \\ x_l' = x_l, \quad l \neq i,j, \end{cases} \qquad a, b \in GF(p),\ a^2 + b^2 = 1, \quad (3.5)$$

in analogy with the usual rotation group. The last is most conveniently studied looking at its covering group SU_2; we now examine what is the corresponding situation for $R(3,p)$. Let $SU^{(\pm)}(2,p^2)$ be the set of the 2×2 matrices

$$u = a_0 1 + i\sum_l a_l\sigma_l : a_0, a_l \in GF(p), \quad \det u = \pm 1, \quad (3.6)$$

where $i \in GF(p^2)$ is given by Eq. (2.6), 1 is the 2×2 unit matrix, and σ_l are the Pauli matrices.

The matrices u thus have elements in $GF(p^2)$ and satisfy the relation

$$u^\dagger u = uu^\dagger = (\det u)1, \quad (3.7)$$

where $u^\dagger$ is the Hermitian conjugate of u, complex conjugation being given by Eqs. (2.11) and (2.12). $SU^{(\pm)}(2,p^2)$ is a group under matrix multiplication: by restriction to the matrices having $\det u = +1$, we get a subgroup $SU^{(+)}(2,p^2)$. The order of these groups follows from Eq. (2.14) (notice that $\det u = a_0^2 + a_1^2 + a_2^2 + a_3^2$):

$$\Omega_{SU^{(+)}(2,p^2)} = p(p^2-1),$$
$$\Omega_{SU^{(\pm)}(2,p^2)} = 2p(p^2-1). \quad (3.8)$$

We may now state:

Proposition 1: There is a 1 to 2 homomorphism of $R(3,p)$ onto $SU^{(\pm)}(2,p^2)$; the subgroup $SU^{(+)}(2,p^2)$

[7] See Ref. 5, p. 47, 48.

[8] See Ref. 5, Chap. VII.

individuates a subgroup $R^{(+)}(3,p)$ of $R(3,p)$ having order $\frac{1}{2}p(p^2-1)$ and being formed by all rotations r such that $1+Sp(r)$ is a square in $GF(p)$.

The proof of this proposition rests on the explicit construction of the homomorphism: let us first consider $R^{(+)}(3,p)$ and $SU^{(+)}(2,p^2)$. To every $r \in R^{(+)}(3,p)$ we may associate two matrices $\pm u \in SU^{(+)}(2,p^2)$ by means of the formulas

$$a_0 = \tfrac{1}{2}[1+Sp(r)]^{\frac{1}{2}}, \quad a_l = \tfrac{1}{2}[1+Sp(r)]^{-\frac{1}{2}} \sum_{jk} \in_{jkl} r_{jk}, \qquad r \in R^{(+)}(3,p), \quad (3.9)$$

where $\in_{jkl}$ is the Levi–Civita tensor and the square roots are defined in $GF(p)$ since $1+Sp(r)$ is assumed to be a square for every $r \in R^{(+)}(3,p)$. From (3.9) one easily checks $\det u = +1$, taking into account the identity $Sp(r^2) - [Sp(r)]^2 + 2Sp(r) = 0$ which holds for any $r \in R(3,p)$. Conversely, to every

$$u \in SU^{(+)}(2,p^2)$$

we may associate a rotation $r \in R^{(+)}(3,p)$ by means of

$$r_{jk} = \tfrac{1}{2}Sp(\sigma_j u \sigma_k u^\dagger) = \tfrac{1}{2}Sp(\sigma_j u \sigma_k u^{-1}), \qquad u \in SU^{(+)}(2,p^2); \quad (3.10)$$

of course u and $-u$ give rise to the same rotation.

From (3.10) one easily checks that $1+Sp(r) = (2a_0)^2$, i.e., a square in $GF(p)$.

We remark that $R^{(+)}(3,p)$ is generated by the transformations $(O^{a,b}_{i,j})^2$ and $O^{\bar{a}b}_{ij} \cdot O^{\bar{a}\bar{b}}_{kl}$ where a, b range over all solutions of $a^2+b^2=1$ [see Eq. (3.5)], while $\bar{a}$, $\bar{b}$ is a particular solution of this equation such that $O^{\bar{a}b}_{ij}$ extends the group generated by $(O^{ab}_{ij})^2$ into the group generated by O^{ab}_{ij}.[8] In fact, the rotations $(O^{ab}_{ij})^2$ and $O^{\bar{a}b}_{ij} \cdot O^{\bar{a}\bar{b}}_{kl}$ satisfy the condition $1+Sp(r)=1$ and generate a group of the same order as $R^{(+)}(3,p)$.

We now come to the subset $R^{(-)}(3,p)$ of $R(3,p)$ composed of the rotations for which $1+Sp(r)$ is a not-square in $GF(p)$, i.e., $-1-Sp(r)$ is a square, since (-1) is a not-square in $GF(p)$. To every $r \in R^{(-)}(3,p)$ we associate two matrices $\pm u$, belonging to the subset $SU^{(-)}(2,p^2)$ of $SU^{(\pm)}(2,p^2)$ composed of the matrices (3.6) for which $\det u = -1$, by means of

$$a_0 = \tfrac{1}{2}[-1-Sp(r)]^{\frac{1}{2}}, \quad a_l = -\tfrac{1}{2}[-1-Sp(r)]^{-\frac{1}{2}} \sum_{jk} \in_{jkl} r_{jk}, \quad r \in R^{(-)}(3,p). \quad (3.9')$$

Conversely, to every $u \in SU^{(-)}(2,p^2)$, the formula

$$r_{jk} = -\tfrac{1}{2}Sp(\sigma_j u \sigma_k u^\dagger) = \tfrac{1}{2}Sp(\sigma_j u \sigma_k u^{-1}), \qquad u \in SU^{(-)}(2,p^2), \quad (3.10')$$

associates a rotation belonging to $R^{(-)}(3,p)$: actually $1+Sp(r) = -(2a_0)^2$, i.e., a not-square in $GF(p)$. Of course u and $-u$ determine the same rotation.

It is easily verified that the correspondences (3.9), (3.10), (3.9′), and (3.10′) between $R(3,p)$ and $SU^{(\pm)}(2,p^2)$ satisfy the product rules $u(rr') = u(r)\cdot u(r')$ (apart from a sign ambiguity) and $r(uu') = r(u)\cdot r(u')$; this completes the proof of Proposition 1.

4. REPRESENTATIONS OF $R(3,p)$

To proceed further in the analysis of the group representations, we now find the number of equivalence classes in $R(3,p)$ making use of the homomorphism with $SU^{(\pm)}(2,p^2)$; by known properties of finite groups, this number is equal to that of inequivalent irreducible representations. Let $[u]$ be the class of equivalence associated to $u \in SU^{(\pm)}(2,p^2)$; it is formed from the matrices

$$v = tut^{-1}, \ \forall t \in SU^{(\pm)}(2,p^2). \quad (4.1)$$

Of course $Sp(v) = Sp(u)$ and $\det v = \det u$, so that the spur and the determinant are necessary to determine an equivalence class. Actually they are also sufficient; in fact, from $Sp(u') = Sp(u)$ and $\det u' = \det u$, one may prove the existence of a $\bar{t} \in SU^{(\pm)}(2,p^2)$ such that

$$u'\bar{t} = \bar{t}u, \quad (4.2)$$

thus ensuring $[u'] = [u]$. Putting $\bar{t} = b_0 1 + i\sum_l b_l\sigma_l$, one obtains from (4.2) a system of four homogeneous equations for b_0, b_1, b_2, b_3; for this system, the determinant of the coefficients and all third-order minors vanish, due to the conditions $Sp(u') = Sp(u)$ and $\det u' = \det u$. Thus one may give arbitrary values to two unknowns, say b_0 and b_1, expressing b_2, b_3 as a linear homogeneous combination of them. The further condition $\det \bar{t} = \pm 1$ now takes the form $c_{00}b_0^2 + c_{01}b_0b_1 + c_{11}b_1^2 = \pm 1$, $c_{ij} \in GF(p)$, which can be reduced by a linear homogeneous substitution belonging to $GF(p)$ to the form (2.13) with $m=1$; this proves the existence of matrices $\bar{t} \in SU^{(\pm)}(2,p^2)$, satisfying Eq. (4.2) [their number is given by Eq. (2.14)]. Therefore, an equivalence class $[u]$ in $SU^{(\pm)}(2,p^2)$ is uniquely determined by the values of $Sp(u)$ and $\det u$: the spur may clearly assume any value in $GF(p)$, i.e., p different values, while the determinant may independently assume the two values $+1$ and -1. Hence we conclude that $SU^{(\pm)}(2,p^2)$ has $2p$ inequivalent irreducible representations.

Since u and $-u$ correspond to the same rotation, while $[u] \neq [-u]$, we may conclude that $R(3,p)$ has p equivalence classes, i.e., p inequivalent irreducible representations.

In order to give an explicit realization of all inequivalent irreducible representations, we shall adopt the standard Weyl's procedure, by introducing as a basis of the carrier space the homogeneous monomials in $GF(p^2)$,

$$f_m^{(j)}(\zeta,\eta) = N_m^{(j)}\xi^{j+m}\eta^{j-m}: \quad N_m^{(j)}, \xi, \eta \in GF(p^2), \quad (4.3)$$

where $j+m$ and $j-m$ are both positive integer numbers, i.e., j and m are both integers or half-integers with $-j \le m \le j$, $j \ge 0$. Writing

$$u \in SU^{(\pm)}(2,p^2)$$

in the form

$$u = [\alpha,\beta] = \begin{pmatrix} \alpha & \beta \\ -\beta^* & \alpha^* \end{pmatrix}; \alpha,\beta \in GF(p^2),$$
$$\alpha\alpha^* + \beta\beta^* = \pm 1,$$

the transformation

$$T^{(0)}(u)f_m^{(j)}(\xi,\eta) = N_m^{(j)} \cdot (\alpha^*\xi - \beta\eta)^{j+m}(\beta^*\xi + \alpha\eta)^{j-m} = \sum_{m'} D_{m',m}^{(j,0)}(u) f_{m'}^{(j)}(\xi,\eta) \quad (4.4)$$

defines a representation of $SU^{(\pm)}(2,p^2)$ of order $2j+1$ with $D_{m',m}^{(j,0)}(u) \in GF(p^2)$ given by

$$D_{m',m}^{(j,0)}(u) = \frac{N_m^{(j)}}{N_{m'}^{(j)}} \sum_{k=\max(0,m-m')}^{\min(j+m,j-m')} \binom{j+m}{k}\binom{j-m}{k+m'-m} \times \alpha^{j-m'-k}\alpha^{*j+m-k}(-\beta)^k\beta^{*k-m+m'}. \quad (4.5)$$

Due to the property (2.1), some matrix elements $D_{m',m}^{(j,0)}(u)$ could be identically equal to the zero element of $GF(p^2)$ for those particular values of j, m', m which permit factorization of a binomial coefficient in (4.5) into the product of the prime p times an integer. (The binomial coefficients are integer numbers in the usual sense and mean iteration of addition in $GF(p^2)$. Alternatively, they may be thought as integral marks in accordance with what was said in Sec. 2.) Actually, it is easily seen that

$$D_{m',m}^{(j,0)}(u) \equiv 0, \quad \text{if} \quad j \ge \tfrac{1}{2}(p+1), \quad j > m \ge p-j, \quad -j+p-1 > m' \ge j-p+1, \quad (4.6)$$

while it is impossible to have $D_{m',m}^{(j,0)}(u) \equiv 0$ if $j \le \frac{1}{2}(p-1)$. By use of Schur's lemma, which holds for algebraically closed fields, we now show:

Proposition 2: The representation $D^{(j,0)}(u)$ given by Eq. (4.5) is irreducible if $j \le \frac{1}{2}(p-1)$ and reducible if $j \ge \frac{1}{2}(p+1)$.

Let $A^{(j)}$ be a square matrix of order $2j+1$ defined in $GF(p^2)$ which commutes with the representation $D^{(j,0)}(u)$; for the subgroup of $SU^{(\pm)}(2,p^2)$ formed by the diagonal matrices $u = [\alpha, 0]$ we get

$$\sum_{m''} A_{m',m''}^{(j)} D_{m'',m}^{(j,0)}([\alpha,0]) = \sum_{m''} D_{m',m''}^{(j,0)}([\alpha,0]) A_{m'',m}^{(j)}, \quad (4.7)$$

and by use of (4.5)

$$A_{m',m}^{(j)}\alpha^{j-m}\alpha^{*j+m} = A_{m',m}^{(j)}\alpha^{j-m'}\alpha^{*j+m'}$$

or, according to Eq. (2.12),

$$A_{m',m}^{(j)}(\alpha^{m-m'(p-1)} - 1) = 0.$$

In order to have nonvanishing values of $A_{m',m}^{(j)}$ for $m' \ne m$ we need, after Eq. (2.2), $(m-m')(p-1) = n(p^2-1)$, i.e., $m-m' = n(p+1)$, $n = \pm 1, \pm 2, \cdots$, and this is possible if and only if $j \ge \frac{1}{2}(p+1)$. Thus, Eq. (4.7) gives rise to the conditions

$$A_{m',m}^{(j)} = a_m\delta_{m'm}, \quad a_m \in GF(p^2), \quad \text{if} \quad j \le \tfrac{1}{2}(p-1),$$
$$A_{m',m}^{(j)} = a_m\delta_{m'm} + b_m\delta_{m'm-n(p+1)},$$
$$a_m, b_m \in GF(p^2), \quad \text{if} \quad j \ge \tfrac{1}{2}(p+1). \quad (4.8)$$

Consider first the case $j \le \frac{1}{2}(p-1)$; the commutation condition

$$\sum_{m''} A_{m',m''}^{(j)} D_{m'',m}^{(j,0)}(u) = \sum_{m''} D_{m',m''}^{(j,0)}(u) A_{m'',m}^{(j)} \; \forall u \in SU^{(\pm)}(2,p^2) \quad (4.9)$$

reads

$$a_m D_{m',m}^{(j,0)}(u) = a_{m'} D_{m',m}^{(j,0)}(u); \quad (4.10)$$

this ensures $a_m = a_{m'}$ since $D_{m',m}^{(j,0)}(u)$ is not identically zero.

Thus $A^{(j)}$ is a multiple of the identity and the first part of Proposition 2 is proved. Consider now the case $j \ge \frac{1}{2}(p+1)$: by choosing, e.g., $b_m = 0$ in (4.8), Eq. (4.10) is still obtained, but, due to (4.6), one can freely assume $a_m \ne a_{m'}$, for $j > m \ge p-j$, $-j+p-1 > m' \ge j-p+1$, so that a matrix $A^{(j)}$ exists, which commutes with the representation $D^{(j,0)}(u)$, but is not a multiple of the identity. This completes the proof of Proposition 2.

Thus, for $j = 0, \frac{1}{2}, 2, \frac{3}{2}, \cdots, \frac{1}{2}(p-1)$ the matrices $D^{(j,0)}(u)$ given by (4.5) provide p inequivalent irreducible representations of $SU^{(\pm)}(2,p^2)$, i.e., one half of the total number of representations we are looking for. On the contrary, these representations exhaust the inequivalent irreducible representations of the subgroup $SU^{(+)}(2,p^2)$ (which indeed has p equivalence classes); in fact, the same procedure leading to Proposition 2 shows that $D^{(j,0)}(u)$, with $j \le \frac{1}{2}(p-1)$, remains irreducible when restricted to $SU^{(+)}(2,p^2)$.

In order to construct all representations of $SU^{(\pm)}(2,p^2)$, let us put

$$D_{m',m}^{(j,1)}(u) = \det u \, D_{m',m}^{(j,0)}(u). \quad (4.11)$$

It may be shown that for $j \le \frac{1}{2}(p-1)$, $D^{(j,1)}(u)$ is not equivalent to $D^{(j,0)};(u)$ an equivalence relation would imply the existence of a matrix $B^{(j)}$ of order

$2j+1$, defined in $GF(p^2)$, satisfying the relation $B^{(j)}D^{(j,0)}(u) = D^{(j,1)}(u)B^{(j)}$, which becomes

$$B^{(j)}D^{(j,0)}(u) = D^{(j,0)}(u)B^{(j)}$$

when u is restricted to the subgroup $SU^{(+)}(2,p^2)$; therefore $B^{(j)}$ should be a multiple of the identity owing to the irreducibility of $D^{(j,0)}(u)$ thought of as a representation of $SU^{(+)}(2,p^2)$, and this makes it impossible to satisfy the equivalence relation

$$B^{(j)}D^{(j,0)}(u) = D^{(j,1)}(u)B^{(j)}, \forall u \in SU^{(\pm)}(2,p^2).$$

Furthermore, all we have said about the irreducibility or reducibility of $D^{(j,0)}(u)$ may be repeated without modifications for $D^{(j,1)}(u)$ so that Proposition 2 is still true when $D^{(j,0)}(u)$ is replaced by $D^{(j,1)}(u)$.

The $2p$ inequivalent irreducible representations of $SU^{(\pm)}(2,p^2)$ may thus be written in the form

$$D^{(j,e)}(u) = (\det u)D^{(j,0)}(u);$$
$$e = 0, 1, \quad j = 0, \tfrac{1}{2}, 1, \tfrac{3}{2}, \cdots, \tfrac{1}{2}(p-1), \quad (4.12)$$

where $D^{(j,0)}(u)$ is given by Eq. (4.5). The general representation $D^{(j,e)}(u)$ operates on the carrier functions (4.3) according to [see Eq. (4.4)]

$$T^{(e)}(u)f_m^{(j)}(\xi,\eta)$$
$$= N_m^{(j)}(\det u)^e \cdot (\alpha^*\xi - \beta\eta)^{j+m}(\beta^*\xi + \alpha\eta)^{j-m}. \quad (4.13)$$

Owing to the 1 to 2 homomorphism of the rotation group $R(3,p)$ onto $SU^{(\pm)}(2,p^2)$, representations of $R(3,p)$ may be deduced from $D^{(j,e)}(u)$ by use of Eqs. (3.9) and (3.9′). By inspection of Eqs. (4.12) and (4.5), it is seen that

$$D^{(j,e)}(-u) = (-1)^{2j}D^{(j,e)}(u); \quad (4.14)$$

since u and $-u$ correspond to the same rotation, one comes to the conclusion that all inequivalent irreducible representations of $R(3,p)$ are obtained from (4.12) for integer values of the spin label j. Corresponding to the half-integer values of j, one obtains matrices which behave as representations of $R(3,p)$ up to a sign ambiguity. The last property reproduces the situation occurring in the usual continuum case: the main novelty is the appearance of the two-valued label e. Hence two representations of the proper rotation group $R(3,p)$ are obtained for each value of spin j.

A closer analogy with the continuum case is obtained looking at the subgroup $R^{(+)}(3,p)$: its representations are deduced from (4.12) taking $e=0$ and are thus uniquely specified by spin j.

One may finally remark that the extension to the rotation–reflection group may be done in the usual way: the group has the structure of a direct product of $R(3,p)$ times the reflection group, which consists of the identity and of the inversion; the representations are specified by a further two-valued label.

5. PROPER LORENTZ GROUP

Let us now consider the group of linear substitutions in $GF(p)$:

$$x'_\mu = \sum_\nu l_{\mu\nu}x_\nu; \quad \mu,\nu = 0,1,2,3, \quad l_{\mu\nu} \in GF(p), \quad (5.1)$$

leaving invariant the quadratic form

$$x_0^2 - x_1^2 - x_2^2 - x_3^2. \quad (5.2)$$

When restricted to the case $\det l = 1$, these transformations form the proper Lorentz group $L(4,p)$ over $GF(p)$; its order is

$$\Omega_{L(4,p)} = p^2(p^4-1). \quad (5.3)$$

It is easily shown that the invariance of (5.2) means

$$l^T g l = g, \quad g = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \quad (5.4)$$

Notice that the usual splitting of the proper Lorentz group into orthochronus and nonorthochronus subsets does not hold any longer; this is because in a finite field the sum of two squares may be a not-square.

We also need to introduce the subset $S(4,p) \subset L(4,p)$ of the special (or pure) Lorentz transformations satisfying

$$s^T = s, \quad s \in S(4,p), \quad (5.5)$$

and we further split $S(4,p)$ into the subsets $S^{(+)}(4,p)$ and $S^{(-)}(4,p)$ formed by those special Lorentz transformations such that $4 - Sp(s^2) + [Sp(s)]^2$ is a square or, respectively, a not-square in $GF(p)$.

We may now set up an application of $S^{(+)}(4,p)$ over the set $H^{(+)}(2,p^2)$ of the 2×2 Hermitian matrices

$$h = c_0 1 + \sum_l c_l\sigma_l; \quad c_0, c_l \in GF(p), \quad \det h = 1, \quad (5.6)$$

by means of

$$c_0 = \{4 - Sp(s^2) + [Sp(s)]^2\}^{-\frac{1}{2}}Sp(s),$$
$$c_l = 2\{4 - Sp(s^2) + [Sp(s)]^2\}^{-\frac{1}{2}}s_{0l}, \; s \in S^+(4,p) \quad (5.7)$$

and making the pair $\pm h$ to correspond to the same s. Conversely, to every $h \in H^{(+)}(2,p^2)$ we may associate a special Lorentz transformation $s \in S^{(+)}(4,p^2)$ by means of

$$s_{\mu\nu} = \tfrac{1}{2}Sp(\sigma_\mu h \sigma_\nu h); \quad \sigma_0 = 1, \quad h \in H^{(+)}(2,p^2); \quad (5.8)$$

of course h and $-h$ give rise to the same s.

Similarly, we may set up a 1 to 2 application of $S^{(-)}(4,p)$ over the set $H^{(-)}(2,p^2)$ of the 2×2 Hermitian matrices

$$h = c_0 1 + \sum_i c_i \sigma_i; \quad c_0, c_i \in GF(p), \quad \det h = -1, \tag{5.6'}$$

by means of

$$c_0 = -\{-4 + Sp(s^2) - [Sp(s)]^2\}^{-\frac{1}{2}} Sp(s),$$
$$c_i = -2\{-4 + Sp(s^2) - [Sp(s)]^2\}^{-\frac{1}{2}} s_{0i}, \qquad s \in S^{(-)}(4,p), \tag{5.7'}$$

$$s_{\mu\nu} = -\tfrac{1}{2} Sp(\sigma_\mu h \sigma_\nu h), \quad h \in H^{(-)}(2,p^2), \tag{5.8'}$$

where h and $-h$ correspond to the same s.

Of course $S(4,p)$ and $H(2,p^2) = H^{(+)}(2,p^2) \cup H^{(-)}(2,p^2)$ do not have a group structure; nevertheless, one may check

$$h(ss') = h(s)h(s'), \quad \text{whenever} \quad s, s', ss' \in S(4,p),$$
$$s(hh') = s(h)s(h'), \quad \text{whenever} \quad h, h', hh' \in H(2,p^2) \tag{5.9}$$

(apart from a sign ambiguity in the first equality).

Taking into account that $\det h = c_0^2 - c_1^2 - c_2^2 - c_3^2$, it is seen, from Eqs. (2.13) and (2.14), that $H^{(+)}(2,p^2)$ and $H^{(-)}(2,p^2)$ both contain $p(p^2+1)$ matrices; hence, the number of special Lorentz transformations is

$$\Omega_{S(4,p)} = p(p^2+1). \tag{5.10}$$

By inspection of Eqs. (3.4), (5.3), and (5.10), the relation

$$\Omega_{L(4,p)} = \Omega_{S(4,p)} \Omega_{R(3,p)} \tag{5.11}$$

follows. Actually we may state:

Proposition 3: Every $l \in L(4,p)$ may be decomposed, in a unique way, into the product

$$l = s \cdot r; \quad s \in S(4,p), \quad r \in R(3,p). \tag{5.12}$$

[Of course, by $r \in R(3,p)$ we mean here a particular Lorentz transformation satisfying $r^T r = 1$, $r_{00} = +1$, i.e., a 4×4 matrix for which $r_{0\mu} = \delta_{0\mu}$ while the elements r_{ij} $(i,j = 1,2,3)$ form a 3×3 rotation in the sense of Sec. 3.]

It is apparent that for any $s \in S(4,p)$ and any $r \in R(3,p)$, the product $s \cdot r$ belongs to $L(4,p)$; the proof of Proposition 3 only requires giving a procedure leading, from any $l \in L(4,p)$, to one $s \in S(4,p)$ and one $r \in R(3,p)$ satisfying (5.12): the uniqueness of the decomposition is then guaranteed by (5.11). Such a procedure may be sketched as follows:

A. Put $ll^T = \tilde{s} \in S(4,p)$; by remarking that $4 - Sp(\tilde{s}^2) + [Sp(\tilde{s})]^2 = (4l_{00})^2$ it follows $\tilde{s} \in S^{(+)}(4,p)$ and by use of Eq. (5.7) the corresponding $\tilde{h} \in H^{(+)}(2,p^2)$ is determined.

B. Let

$$h_1 = \begin{cases} [2 + Sp(\tilde{h})]^{-\frac{1}{2}}(1 + \tilde{h}) \in H^{(+)}(2,p^2) & \text{if} \\ \qquad [2 + Sp(\tilde{h})] \text{ is a square in } GF(p), \\ [-2 - Sp(\tilde{h})]^{-\frac{1}{2}}(-1 - \tilde{h}) \in H^{(-)}(2,p^2) & \text{if} \\ \qquad [2 + Sp(\tilde{h})] \text{ is a not-square in } GF(p), \end{cases} \tag{5.13a}$$

$$h_2 = \begin{cases} [-2 + Sp(\tilde{h})]^{-\frac{1}{2}}(-1 + \tilde{h}) \in H^{(-)}(2,p^2) & \text{if} \\ \qquad [-2 + Sp(\tilde{h})] \text{ is a square in } GF(p), \\ [2 - Sp(\tilde{h})]^{-\frac{1}{2}}(1 - \tilde{h}) \in H^{(+)}(2,p^2) & \text{if} \\ \qquad [-2 + Sp(\tilde{h})] \text{ is a not-square in } GF(p), \end{cases} \tag{5.13b}$$

and denote by s_1 and s_2 the corresponding elements of $S(4,p)$ obtained by Eqs. (5.7), (5.8) or (5.7'), (5.8'); it is easily checked that $h_1^2 = (\det h_1)\tilde{h}$, $h_2^2 = (-\det h_2)\tilde{h}$: hence [see Eq. (5.9)]

$$s_1^2 = s_2^2 = \tilde{s} = ll^T.$$

Moreover, one has

$$(s_1)_{00} = Sp(\tilde{h}) = -(s_2)_{00}.$$

C. Put $(s_1)^{-1}l = r_1$, $(s_2)^{-1}l = r_2$ and remark that

$$r_i^T r_i = l^T (s_i^2)^{-1} l = l^T (ll^T)^{-1} l = 1, \quad (i = 1,2);$$

this implies $(r_i)_{0\mu} = \pm\delta_{0\mu}$ and one may conclude that r_i is an element of $R(3,p)$ only if the further condition $(r_i)_{00} = +1$ is met. Actually, $l_{00} = (r_1)_{00}(s_1)_{00} = (r_2)_{00}(s_2)_{00}$ and we have seen that $(s_1)_{00} = -(s_2)_{00}$; therefore $(r_1)_{00} = -(r_2)_{00}$, i.e., either r_1 or r_2 belongs to $R(3,p)$. Denoting by r such a rotation and by s the corresponding choice between s_1 and s_2, the decomposition (5.12) is accomplished.

We may now come to:

Proposition 4: There is a 1 to 2 homomorphism of $L(4,p)$ onto the group $SL^{(\pm)}(2,p^2)$ of the 2×2 matrices:

$$a = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}; \alpha, \beta, \gamma, \delta \in GF(p^2), \quad \det a = \pm 1. \tag{5.14}$$

The correspondence from $l \in L(4,p)$ to $a \in SL^{(\pm)}(2,p^2)$ is determined by use of Proposition 3 together with Eqs. (5.7) or (5.7'), and (3.9) or (3.9'). Conversely, to every $\pm a \in SL^{(\pm)}(2,p^2)$ we may associate a Lorentz transformation $l \in L(4,p)$ by first factorizing $a = h \cdot u$ with $h \in H^{(\pm)}(2,p^2)$,

$u \in SU^{(\pm)}(2, p^2)$,[9] and then using Eqs. (5.8) or (5.8′) and (3.10) or (3.10′); the result may be simply expressed in the form

$$l_{\mu\nu} = (\det a) \cdot \tfrac{1}{2} Sp(\sigma_\mu a \sigma_\nu a^\dagger) \quad a \in SL^{(\pm)}(2, p^2), \quad (5.15)$$

and the product rule

$$l(aa') = l(a)l(a')$$

is easily verified.

Writing the condition $\det a = \pm 1$ in the form

$$\tfrac{1}{4}[(\alpha + \delta)^2 - (\alpha - \delta)^2 - (\beta + \gamma)^2 + (\beta - \gamma)^2] = \pm 1,$$

and by use of Eqs. (2.13) and (2.14), the order of $SL^{(\pm)}(2, p^2)$ is found to be $2p^2(p^4 - 1)$, in agreement with Eq. (5.3). When restricted to the condition $\det a = +1$, the matrices (5.14) form a subgroup

$$SL^{(+)}(2, p^2)$$

of order $p^2(p^4 - 1)$, which determines, through the homomorphism, a subgroup $L^{(+)}(4, p)$ of $L(4, p)$ having order $\frac{1}{2}p^2(p^4 - 1)$. An element of $L^{(+)}(4, p)$ is factorized, according to (5.12), into $s \in S^{(+)}(4, p)$ and $r \in R^{(+)}(3, p)$ or alternatively into $s \in S^{(-)}(4, p)$ and $r \in R^{(-)}(3, p)$.

We now proceed in evaluating the number of equivalence classes of $SL^{(+)}(2, p^2)$. Two matrices a', $a \in SL^{(\pm)}(2, p^2)$ belong to the same equivalence class $[a]$ if they are connected by a relation $a' = bab^{-1}$, $b \in SL^{(\pm)}(2, p^2)$. The spur and the determinant are thus constants over the whole equivalence class; furthermore, $[a]$ is uniquely determined by the two quantities $Sp(a)$ and $\det a$. To show this, we have to prove that the equalities $Sp(a') = Sp(a)$ and $\det a' = \det a$ guarantee the existence of a matrix $\bar{b} \in SL^{(\pm)}(2, p^2)$ satisfying

$$a'\bar{b} = \bar{b}a, \quad (5.16)$$

i.e., a system of four homogeneous linear equations for the matrix elements $\bar{b}_{11}$, $\bar{b}_{12}$, $\bar{b}_{21}$, $\bar{b}_{22}$. Actually the determinant of the coefficients and all third-order minors vanish if $Sp(a') = Sp(a)$, $\det a' = \det a$, so that two unknowns, say $\bar{b}_{12}$ and $\bar{b}_{21}$, may be expressed as linear homogeneous combinations of $\bar{b}_{11}$, $\bar{b}_{22}$, and the further condition $\det \bar{b} = \pm 1$ takes the form $c_{11}\bar{b}_{11}^2 + c_{12}\bar{b}_{11}\bar{b}_{22} + c_{22}b_{22}^2 = \pm 1$, $c_{ij} \in GF(p^2)$, which can be reduced by a linear homogeneous substitution belonging to $GF(p^2)$ to the form (2.13): this shows the existence of $\bar{b} \in SL^{(\pm)}(2, p^2)$ satisfying $a'\bar{b} = \bar{b}a$. Clearly, $Sp(a)$ may assume any value in $GF(p^2)$, i.e., p^2 different values; for any given value of $Sp(a)$, one is still free to assume $\det a = +1$ or $\det a = -1$; therefore $SL^{(\pm)}(2, p^2)$ has $2p^2$ equivalence classes and $2p^2$ inequivalent irreducible representations.

Because of a and $-a$ corresponding to the same $l \in L(4, p)$, but $[a] \neq [-a]$, the number of inequivalent irreducible representations of $L(4, p)$ is p^2.

An explicit realization of these representations could now be given by a procedure similar to the one used in the previous section. We shall only quote the result, which may also be deduced from the general treatment of Brauer and Nesbitt[10]; the $2p^2$ irreducible representations of $SL^{(\pm)}(2, p^2)$ may be expressed as a direct product:

$$D^{(j,k;e)}(a) = (\det a)^e \cdot D^{(j)}(a) \otimes [D^{(k)}(a)]^*,$$
$$e = 0, 1, \quad j, k = 0, \tfrac{1}{2}, 1, \cdots, \tfrac{1}{2}(p - 1), \quad (5.17)$$

where $D^{(j)}(a)$ is obtained by Eq. (4.5) replacing $\det u$, $-\beta^*$, α^* by $\det a$, γ, δ, with reference to Eq. (5.14). By remarking that

$$D^{(j,k;e)}(-a) = (-1)^{2(j+k)} D^{(j,k;e)}(a),$$

the p^2 irreducible representations of $L(4, p)$ are obtained from (5.17) for integer values of $j + k$, by use of the homomorphism between $SU^{(\pm)}(2, p^2)$ and $L(4, p)$. For half-integer values of $j + k$, one obtains, from (5.17), matrices which behave as representations of $L(4, p)$ up to a sign ambiguity. Compared with the usual continuum case, one remarks on the appearance of the new label e: from Ref. 10, it could assume any integer value from zero to $p^2 - 2$, but its range is actually restricted to the values 0, 1 since in our case $\det a = \pm 1$. If the condition $\det a = \pm 1$ is dropped and replaced by $(\det a)^{p+1} = \pm 1$, one is led to the situation considered by Coish.[2] [In this reference the label e ranges from zero to $p^2 - 2$, and the quantity $Q = 2(e + j - k)$ is related to some gauge transformation; an interpretation of Q as the electric charge is suggested.]

If the subgroup $L^{(+)}(4, p)$ is considered, one may deduce the representations from (5.17), assuming $e = 0$, and a close analogy with the ordinary continuum case is obtained.

6. POINCARÉ GROUP OVER $GF(p)$

The Poincaré or inhomogeneous Lorentz group $P(4, p)$ is the group of linear transformations

$$x_\mu \to x'_\mu = \sum_\nu l_{\mu\nu} x_\nu + a_\mu; \quad l \in L(4, p),$$
$$x_\mu, a_\mu \in GF(p); \quad (6.1)$$

its elements are denoted by (a, l) with composition law

$$(a, l)(a', l') = (a + la', ll'),$$

where a stands for the 4-vector (a_0, a_1, a_2, a_3).

[9] Letting $aa^\dagger = \tilde{h} \in H^{(+)}(2, p^2)$, the matrix h is given by (5.13a) if $\det a = +1$, by (5.13b) if $\det a = -1$; the matrix u is then determined as $h^{-1}a$.

[10] R. Brauer and C. Nesbitt, Ann. Math. **42**, 556 (1941).

The subgroup of translations $T(4,p) \equiv \{(a,1), \forall a\}$ is Abelian invariant since

$$(a',l')^{-1}(a,1)(a',l') = (l'^{-1}a,1) \in T(4,p);$$

thus, $P(4,p)$ has a semidirect product structure, and its irreducible representations can be found by the standard method of induced representations.[11] The order of $P(4,p)$ is $p^6(p^4-1)$, i.e., the product of the order p^4 of $T(4,p)$ times the order of $L(4,p)$. [See Eq. (5.3).]

Let us sketch a brief outline of the steps leading to all irreducible representations of $P(4,p)$.

A. There are p^4 inequivalent characters of the Abelian subgroup $T(4,p)$. In order to determine them, consider the vector space V with elements $v(u)$, $u \in GF(p)$, satisfying the product rule $v(u) \cdot v(u') = v(u+u')$,[12] corresponding to any given 4-vector $k = (k_0, k_1, k_2, k_3)$, $k_\mu \in GF(p)$, one may define on V the one-dimensional representation $K(a)$ of $T(4,p)$,

$$K(a)v(k \cdot x) = v(k \cdot (x+a)) = v(k \cdot a)v(k \cdot x), \quad (6.2)$$

where $k \cdot x = k_0x_0 - k_1x_1 - k_2x_2 - k_3x_3$ and use has been made of Eq. (6.1). The character $v(k \cdot a)$ is individuated by the 4-vector k, thus obtaining p^4 of them.

B. As in the classical treatment, to every character $v(k \cdot a)$, or simply to every 4-vector k, one associates an orbit, i.e., the set of 4-vectors k' defined by $\{k' : k' = lk, \forall l \in L(4,p)\}$; one thus obtains four different classes of orbits specified by the particular choices

$$k^{(1)} = (k_0, 0, 0, 0), \quad k^{(2)} = (0, 0, 0, k_3),$$
$$k^{(3)} = (k_0, 0, 0, k_0), \quad k^{(4)} = (0, 0, 0, 0). \quad (6.3)$$

C. For each class, the corresponding little group is defined as the subgroup of $L(4,p)$, leaving invariant the orbit representative (6.3). For the timelike choice $k^{(1)}$, the little group is $R(3,p)$; for the spacelike choice $k^{(2)}$, the little group is the subgroup $L(3,p)$ of $L(4,p)$ which leaves unchanged the coordinate x_3[13]; for the lightlike choice $k^{(3)}$, the little group is the semidirect product $E(2,p)$ of translations and rotations in a two-dimensional geometry built over $GF(p)$; for the trivial choice $k^{(4)}$ the little group coincides with the whole $L(4,p)$.

According to standard procedures, each irreducible representation of the Poincaré group is now determined by the representation (6.2) of the Abelian subgroup $T(4,p)$ and by the representation of the little group associated to the orbit of k.

As in the classical treatment, one may try to associate a representation of $P(4,p)$ to a free stable-particle state, interpreting $k \cdot k$ (which is invariant over an orbit) as the square of the particle mass: this selects the choices $k^{(1)}$ and $k^{(3)}$ in (6.3) as the ones of physical relevance if one wants the mass to be defined on the Galois field $GF(p)$. (In Ref. 3, the consequences of a further requirement of $k_1^2 + k_2^2 + k_3^2$ being a square is discussed; in this way a notion of mass spectrum is introduced.)

The representations of the little group $R(3,p)$ have been previously discussed; the ones of $E(2,p)$ can be determined by the same method of induced representations used for $P(4,p)$. Looking, e.g., to the case of nonzero mass, the association between the $P(4,p)$ representation and the particle states would require an answer to the question: What is the physical meaning, if any, of the label e appearing in the $R(3,p)$ representations? We finally remark that restricting l in Eq. (6.1) to $L^{(+)}(4,p)$, a subgroup $P^{(+)}(4,p)$ is obtained whose representations exhibit a close analogy to the ones of the ordinary geometry case.

ACKNOWLEDGMENT

The authors wish to thank Professor P. Kustaanheimo for useful discussions.

[11] See, e.g., G. W. Mackey, Acta Math. **99**, 265 (1958); E. P. Wigner, Ann. Math. **40**, 149 (1939).

[12] Note that $v(u)$ cannot belong to any finite field.

[13] It coincides with the group $0_{-1}(3,p)$ in the notation of Ref. 5.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 7 JULY 1968

Model for Converging Detonations in Solid Explosives

ROBERT L. CONGER*
Naval Ordnance Laboratory, Corona, California

(Received 5 January 1967)

The fluid-dynamic equations which describe converging detonation fronts in high explosives are solved over a coordinate system that changes from circular to a pleated pattern as the detonation converges. There are three reasons for introducing such a coordinate system rather than making the usual assumption that when a spherical or cylindrical explosive charge is initiated simultaneously on its curved surface the detonation front maintains its symmetry as the detonation converges to the center: (1) Calculations made elsewhere show that a converging circular shock wave is unstable. (2) Calculations in this paper show that if spherical or cylindrical symmetry is maintained, density and velocity become infinite, but if the detonation front folds these quantities remain finite except at a few points which can be removed by cuts. The average fluid velocity remains almost constant over 99% of the detonation front. (3) Experiments indicate that a cylindrical detonation front becomes rather symmetrically deformed.

I. INTRODUCTION

This paper presents a mathematical model for a converging detonation front. Previous theories of converging detonations have assumed that when a spherical or cylindrical explosive charge is initiated simultaneously along a curved surface, the spherical or cylindrical symmetry is maintained by the detonation front as the detonation converges to the center. There are at least three reasons for thinking that this symmetry is not maintained: (1) Calculations have shown that a converging circular shock wave is unstable in that a small disturbance in the shock will grow larger with time. (2) Calculations presented here show that if the cylindrical or spherical symmetry of the detonation front were maintained, the density, detonation velocity, and detonation product velocity would become infinite at the center; but if the detonation front is allowed to collapse into a star-shaped pattern, these quantities remain finite as the detonation process goes to completion, except for material at the points of the collapsed star. These points, however, represent only a small part of the total charge. (3) When a cylindrical disk of explosive is placed in contact with a steel block and is detonated, the dent produced in the center of the block by the converging detonation is not circular, but is approximately star-shaped. Figure 1 is a drawing of a photomicrograph of such a dent.

The complete solution of this problem of a converging detonation involves the solution of fluid-dynamics equations in three-space coordinates and time. There is no way of obtaining such a solution. In this paper, the four independent variables are reduced to one space coordinate and time, with the space coordinate expressed in terms of general curvilinear coordinates so that the results can be applied to irrotational flow in any geometry for which metrical coefficients can be found. With the detonation products flowing along one curvilinear coordinate and the other two orthogonal curvilinear coordinates constant, the equations are then reduced to characteristic form.

The calculations in this paper are based on a mechanical model rather than a thermodynamic model. The justification for doing this for solid explosives is described in considerable detail by Zeldovich and Kompaneets.[1] Briefly, the arguments are the following: Two types of energy are present in detonations of solids. One is thermal energy which consists of the kinetic energy of the moving molecules. The other is the elastic energy arising from the binding forces between the atoms and molecules. It is possible to describe the behavior of condensed explosives with density greater than about 1 g/cm^3, by considering that the elastic energy and the elastic part of the pressure are predominant. This results in an equation of state of the form $p = A\rho^3$ which is independent of temperature. The Rayleigh and Hugoniot curves are close together. Zeldovich and Kompaneets, on p. 228, state that this simple equation of state can be used for calculating the motion of explosion products if it is remembered that about a third of the pressure actually is not elastic but of thermal origin, and if the density is greater than about 1 to 1.3 g/cm^3.

With the aid of the equation of state, the number of independent variables in the characteristic equation is reduced to two, but the equation is of a nonintegrable form. To convert the equation to a form that can be integrated, it is combined with the equation of motion for detonation products crossing the detonation front. The integration is performed for

* Present address: Naval Weapons Center, Corona Laboratories, Corona, Calif.

[1] I. B. Zeldovich and A. S. Kompaneets, *Theory of Detonation* (Academic Press Inc., New York, 1960), pp. 222–246.

general curvilinear coordinates and the results are applied to three geometries: the sphere, the circular cylinder, and a shape produced by a conformal transformation of an ellipse.

II. CHARACTERISTIC EQUATIONS IN CURVILINEAR COORDINATES

The equations for conservation of mass and conservation of momentum in vector form are[2]

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0, \tag{1}$$

$$\rho\left[\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u}\right] + c^2 \nabla \rho = 0, \tag{2}$$

where ρ is the fluid density, u is the fluid velocity, c is the speed of sound in the fluid, and t is time. In general curvilinear coordinates for one dimension, these two equations become[3]

$$\frac{\partial \rho}{\partial t} + \frac{1}{h_1 h_2 h_3}\left[\frac{\partial h_2}{\partial x} h_3 \rho u + h_2 \frac{\partial h_3}{\partial x} \rho u + h_2 h_3 u \frac{\partial \rho}{\partial x} + h_2 h_3 \rho \frac{\partial u}{\partial x}\right] = 0, \tag{3}$$

$$\rho \frac{\partial u}{\partial t} + \frac{\rho u}{h_1}\frac{\partial u}{\partial x} + \frac{c^2}{h_1}\frac{\partial \rho}{\partial x} = 0, \tag{4}$$

where u is now the velocity in the direction of flow (the x direction), and x is the independent curvilinear variable. The terms $h_1 h_2 h_3$ are the metrical coefficients. These coefficients are, in general, functions of the space coordinates.

Equations (3) and (4) will now be transformed into characteristic form.[4] One set of characteristic equations is

$$\frac{\partial u}{\partial \alpha} + \frac{c}{\rho}\frac{\partial \rho}{\partial \alpha} + \frac{cu}{h_1}\frac{\partial t}{\partial \alpha}\left[\frac{1}{h_2}\frac{\partial h_2}{\partial x} + \frac{1}{h_3}\frac{\partial h_3}{\partial x}\right] = 0 \tag{5}$$

and

$$\frac{\partial x}{\partial \alpha} \frac{u + c}{h_1} \frac{\partial t}{\partial \alpha}. \tag{6}$$

The other set is

$$\frac{\partial u}{\partial \beta} - \frac{c}{\rho}\frac{\partial \rho}{\partial \beta} - \frac{cu}{h_1}\frac{\partial t}{\partial \beta}\left[\frac{1}{h_2}\frac{\partial h_2}{\partial x} + \frac{1}{h_3}\frac{\partial h_3}{\partial x}\right] = 0 \tag{7}$$

and

$$\frac{\partial x}{\partial \beta} = \frac{u - c}{h_1}\frac{\partial t}{\partial \beta}. \tag{8}$$

Equations (5) and (6) represent a wave moving in the direction of increasing x with a velocity $(u + c)/h_1$. Equations (7) and (8) represent a wave moving in the direction of decreasing x with a velocity $(u - c)/h_1$. In this second case, u is negative.

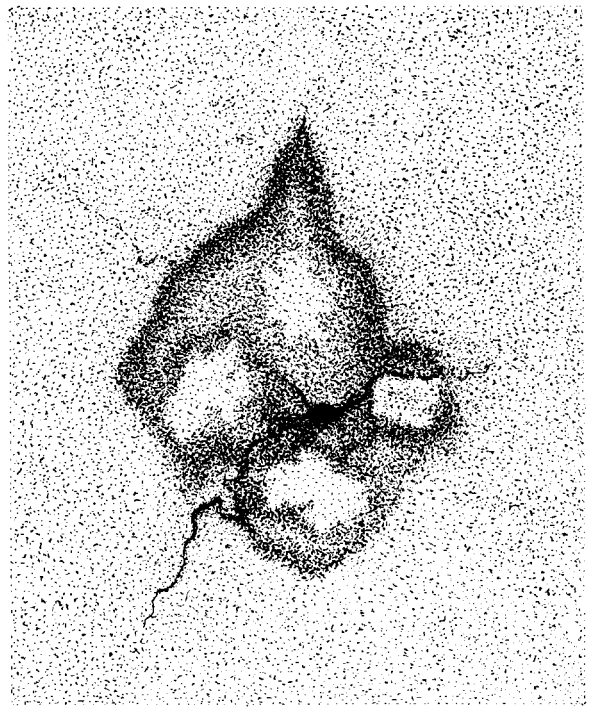


FIG. 1. Drawing of the pattern produced at the center of a steel block by the implosion of a $\frac{1}{4}$ inch by 3-in circular disk of PETN sheet explosive.

Combining Eqs. (7) and (8) gives

$$du - \frac{c}{\rho} d\rho - \frac{cu}{u - c}\left[\frac{dh_2}{h_2} + \frac{dh_3}{h_3}\right] = 0 \tag{9}$$

for the motion along the characteristic. The third term in Eq. (9) represents the effects of convergence.

Since the speed of sound is given by

$$c^2 = dp/d\rho, \tag{10}$$

where p is pressure, and the equation of state is

$$p = A\rho^3, \tag{11}$$

then

$$c = \rho(3A)^{\frac{1}{2}} \tag{12}$$

behind the detonation front.

Substitution of Eq. (12) into (9) gives

$$du - dc - \frac{cu}{u - c}\left[\frac{dh_2}{h_2} + \frac{dh_3}{h_3}\right] = 0. \tag{13}$$

Equation (13) is of the Pfaffian form.[5] In Eq. (13), h_2 and h_3 are functions of the one space variable ξ;

[2] R. Courant and K. O. Friedrichs, *Supersonic Flow and Shock Waves* (Interscience Publishers, Inc, New York, 1948), p. 14.

[3] P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill Book Co., Inc., New York, 1953), p. 37.

[4] Ref. 2, p. 45.

[5] A. R. Forsyth, *Theory of Differential Equations* (Dover Publications, Inc., New York, 1959), Vol. I.

therefore (13) can be written in the form

$$du - dc - \frac{cu}{u-c} f(\xi)\, d\xi = 0. \tag{14}$$

The coefficients of du, dc, and $d\xi$ can be considered as components of a vector $\mathbf{F}$. The condition that the Pfaffian differential equation be integrable is

$$\mathbf{F} \cdot \nabla \times \mathbf{F} = 0. \tag{15}$$

The coefficients of Eq. (14) do not satisfy (15) unless $u = c$, but $u = c$ gives a stationary condition.

Equation (13) represents any convergent flow with equation of state (11). To make (13) represent a converging detonation front, it must be combined with a detonation-front equation. The conditions across the detonation front will now be considered. The detonation front is a step discontinuity in pressure, density, fluid velocity, and sound speed. The chemical reaction is assumed to take place instantaneously at the discontinuity. In front of the detonation front, fluid velocity is zero and pressure is essentially zero compared to the pressure behind the detonation front. According to the Chapman–Jouguet hypothesis,[6] the detonation front moves at the speed of fluid flow plus the speed of sound, that is, $u + c$.

In this converging detonation model u, c, and the detonation velocity are variables. The detonation velocity is assumed to be $-u + c$, since u is negative. The equation for the conservation of momentum across the detonation front becomes[7]

$$(-u + c)^2 \rho_0 - c^2 \rho = p, \tag{16}$$

where the subscript 0 refers to conditions ahead of the detonation front.

From Eqs. (11), (12), and (16), it is found that

$$(-u + c)^2 c_0 = 4c^3/3, \tag{17}$$

where c_0 is defined by

$$c_0 = \rho_0 (3A)^{\frac{1}{2}}. \tag{18}$$

Solving (17) for u gives

$$u = \frac{4c_0}{3}\left(\psi^2 - \frac{4\psi^3}{3}\right), \tag{19}$$

where

$$\psi = \left(\frac{3c}{4c_0}\right)^{\frac{1}{2}}. \tag{20}$$

When (19) is substituted into (13), Eq. (13) can be integrated to give

$$h_2 h_3 / h_{2_\infty} h_{3_\infty} = (4\psi - 3)^{-3} e^{(4-4\psi)}, \tag{21}$$

[6] Reference 2, p. 212.
[7] Reference 2, p. 206.

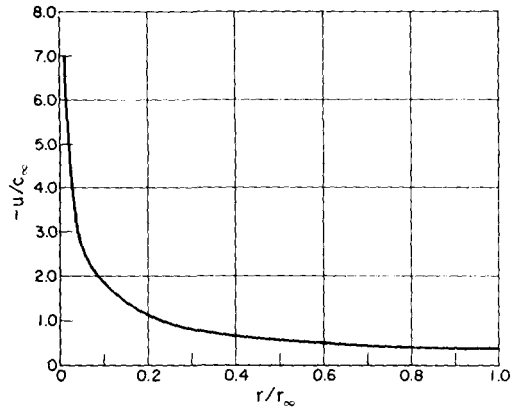


FIG. 2. Fluid velocity u as a function of radius r for a spherical implosion.

where h_{2_∞} and h_{3_∞} are the values of the metrical coefficients far enough from the convergent region so that $c = 4c_0/3$. Equation (21) now applies to the detonation front only.

III. APPLICATION TO SPHERICAL- AND CYLINDRICAL-COORDINATE SYSTEMS

Some specific geometries will now be considered. The first will be a sphere initiated simultaneously on the surface. The metrical coefficients are $h_2 = r$ and $h_3 = r \sin\theta$. The independent variable is the radius r. Substitution into Eq. (21) gives

$$r/r_\infty = (4\psi - 3)^{-\frac{3}{2}} e^{(2-2\psi)}. \tag{22}$$

Although h_3 is a function of θ, h_3/h_{3_∞} is not and Eq. (22) is a function of only one space variable r.

Equations (19) and (22) express u and r as a function of the parameter ψ. Figure 2 is a graph of u as a function of r as determined by the use of Eqs. (19) and (22). For cylindrical coordinates, $h_2 = r_0$, $h_3 = 1$, and Eq. (21) becomes

$$r/r_\infty = (4\psi - 3)^{-3} e^{(4-4\psi)}. \tag{23}$$

Since Eqs. (19), (22), and (23), as well as Fig. 1, indicate an infinite velocity at $r = 0$, the actual detonation cannot be described by Eqs. (22) and (23) at the center. Whitham[8] shows that a converging circular shock wave is unstable in that a small disturbance on the shock will grow larger with time. It is probable that the converging detonation is unstable also and that, as a result, a cylindrical detonation front becomes folded or pleated in some way as the detonation converges. Such a shape would increase the detonation surface and allow the detonation to proceed at a lower average velocity. This suggests that a coordinate system which changes from circular to a different shape be used in Eq. (21).

[8] G. B. Whitham, J. Fluid Mech. 2, 145 (1957), p. 170.

IV. FOLDED-COORDINATE SYSTEMS

The detonation front could be folded in many ways. Figure 1 is a drawing of the pattern produced by a converging detonation in about a $\frac{1}{4}''$ diameter area at the center of a steel block. This drawing was made from a color photograph. A black and white photograph did not show clearly the central pattern.

Coordinates are needed which change from circles to curves similar to Fig. 1 with the number of axes of symmetry arbitrary.

For two dimensions, orthogonal-coordinate systems can be created at will by the methods of conformal transformations.

The transformation

$$w = q + js = z^n, \tag{24}$$

where $z = x + jy$ and $j = (-1)^{\frac{1}{2}}$ transforms an ellipse in the w plane into an approximately star-shaped pattern in the z plane with the number of points equal to $2n$. The two limiting curves for the ellipse—the circle and the straight line—are transformed into another circle and a wheel spoke pattern, respectively. Figure 3 shows a computer-generated series of transformed curves for $n = 2$. The metrical coefficient for $n = 2$ is given by

$$h_\eta = \left[\frac{\partial x^2}{\partial \eta} + \frac{\partial y^2}{\partial \eta}\right]^{\frac{1}{2}} = \frac{C^{\frac{1}{2}}}{2(\xi^2 + \eta^2 - 1)^{\frac{1}{4}}}\left[\frac{\xi^2(1-\eta^2) + \eta^2(\xi^2 - 1)}{1 - \eta^2}\right]^{\frac{1}{2}}, \tag{25}$$

where ξ and η are the elliptic coordinates, $q = a\xi\eta$, $s = a[(\xi^2 - 1)\cdot(1 - \eta^2)]^{\frac{1}{2}}$, $\xi \geq 1.0$, $|\eta| \leq 1.0$. The curves ξ constant are ellipses and the curves η-constant are hyperbolas in the w plane. (For the curves of Fig. 3, ξ^2 has the following values starting at the center: 1.00, 1.001, 1.01, 1.05, 1.20, 2.00, 4.00, 10.0, 40.0, 100.0.) Equation (21), in which $h_2 = h_\eta$ and $h_3 = 1.0$, becomes

$$\left[\frac{\xi_\infty^2 + \eta^2 - 1}{\xi^2 + \eta^2 - 1}\right]^{\frac{1}{4}}\left[\frac{\xi^2(1-\eta^2) + \eta^2(\xi^2 - 1)}{\xi_\infty^2(1-\eta^2) + \eta^2(\xi_\infty^2 - 1)}\right]^{\frac{1}{2}} = (4\psi - 3)^{-3}e^{(4-4\psi)}. \tag{26}$$

Equation (26) is a function of both elliptic coordinates, ξ and η. Since this model is to be a function of only one space coordinate, η will be held constant along the flow path. This constraint means that at the detonation front, the detonation products are assumed to move in channels of constant η. η then becomes a parameter which identifies a flow channel. The justifications for this assumption are given in the introduction. This model says nothing about the flow of detonation products behind the detonation front. The term on the right of Eq. (26) becomes 0, and therefore y and the fluid velocity u become infinite only at $\xi = \eta = 1.0$. At $\xi = 1.0$, $\eta = 0$, $y = \frac{3}{4}$, and $u = 0$.

Figure 4 shows the fluid velocity as a function of the parameter η for $\xi = 1$, the limit of a converging detonation where the star pattern is collapsed to a cross.

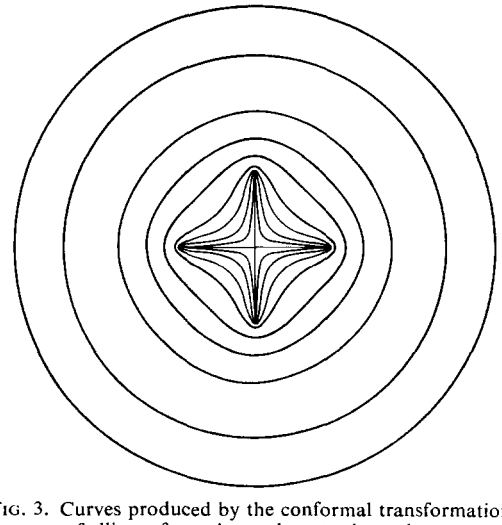


FIG. 3. Curves produced by the conformal transformations of ellipses from the w plane to the z plane.

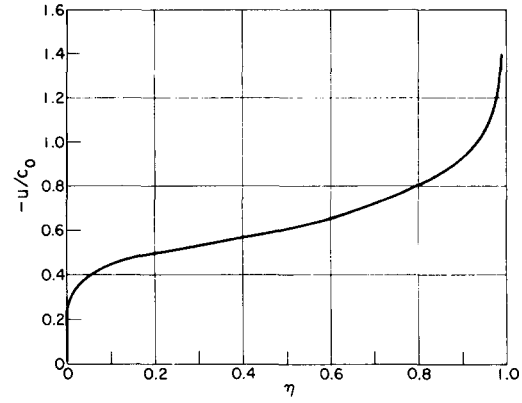


FIG. 4. Fluid velocity u as a function of elliptic coordinate parameter η at the limit of an implosion in the transformed coordinate system. $\xi_\infty = 10.0$.

V. DISCUSSION

Numeric integration gives an average velocity, between $\eta = 0$ and $\eta = 0.99$, of $0.625c_0$, which is less than twice the fluid velocity far enough from the converging region so that convergence has no effect. Cuts could be made to remove the points of the crosses. These cuts could remove 1 % of the explosive and the remaining explosive would have a detonation product velocity which increases only slightly as the detonation converges to a cross. It could be assumed, with fair accuracy, that the detonation product flow

velocity at the detonation front $u - c$ is almost constant also. This is in agreement with experiments which indicate a near constant detonation velocity in converging cylindrical detonations.

Perhaps the circle collapses to a pattern with only a small number of points because the total distance around such a pattern is only slightly larger than the distance around the circumscribing circle. With this distance about the same and the detonation velocity nearly constant, the rate of release of chemical energy is nearly constant during the final part of the detonation. This would suggest that the spherical detonation converges to a pattern corresponding to the collapse of an octahedron or perhaps a cube or a dodecahedron.

APPENDIX: A MORE GENERAL EQUATION OF STATE

To fit experimental data better, the equation of state

$$p = A\rho^3 e^{\delta\rho} \tag{A1}$$

can be used in place of Eq. (11).

Equation (A1) is substituted into Eqs. (9), (10), and (16), and for δ small, terms higher than the first power in δ can be dropped. The result for spherical coordinates is

$$\frac{2}{\omega_0}\left[\frac{3\omega}{\omega_0} - \frac{2\delta\omega^2}{3^{\frac{1}{2}}} + \frac{5\delta\omega^3}{2\omega_0}\right] d\omega \Big/ \left[3 - 2(3)^{\frac{1}{2}}\frac{\omega}{\omega_0} + \frac{3\delta\omega^2}{2} - \frac{11\delta\omega^3}{4(3\omega_0)^{\frac{1}{2}}}\right] = \frac{dr}{r}, \tag{A2}$$

where $\omega^2 = \rho$ and $\omega_0^2 = \rho_0$.

From Eqs. (10), (16), and (A1), for δ small, the fluid velocity becomes

$$u = A\omega^2\left(3^{\frac{1}{2}} - \frac{2\omega}{\omega_0} + \frac{2\omega^2\delta}{3^{\frac{1}{2}}} - \frac{5\omega^3\delta}{4\omega_0}\right). \tag{A3}$$

Equation (A2) can be integrated numerically by the use of a computer. Equation (A3) and the numerical integration of Eq. (A2) then express u and r in terms of the parameter ω. For $\delta = 0.1/\omega_0^2$, the resulting curve is very similar to Fig. 2.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 7 JULY 1968

Plethysm and the Theory of Complex Spectra*

P. R. SMITH† AND B. G. WYBOURNE
Physics Department, University of Canterbury, Christchurch, New Zealand

(Received 26 August 1967)

Notable progress in developing the theory of complex spectra has come from the exploitation of the symmetry properties of atomic wavefunctions using the theory of continuous groups. Many seemingly simple results have previously been obtained by what would appear to be unreasonably complex methods. Many of these complexities may be removed, and new results derived, using the algebra of *plethysm* first developed by Littlewood. Applications to three central problems in the theory of complex spectra are discussed: (1) the classification of the atomic states of n-electron configurations; (2) the analysis and classification of the N-particle operators that arise in the application of perturbation theory to atomic problems; (3) the derivation of selection rules for the matrix elements of operators.

I. INTRODUCTION

The analysis and interpretation of the spectra of atoms and nuclei has drawn heavily upon parallel developments in the theory of complex spectra. Notable progress in developing the theory of complex spectra came from Racah's[1,2] exploitation of the symmetry properties of atomic wavefunctions using the classical theory of continuous groups as developed by Lie,[3] Cartan,[4] Weyl,[5,6] and others. Later developments[7–9] have made considerable use of Young's

* Research sponsored in part by the Air Force Office of Scientific Research, Office of Aerospace Research, United States Air Force, under AFSOR Grant No. 1275-67.

† Present address: Physics Department, Cambridge University, Cambridge, England.

[1] G. Racah, Phys. Rev. **76**, 1352 (1949).

[2] G. Racah, *Group Theory and Spectroscopy, Ergebnisse der exacten Naturwissenschaften, Vol. 37* (Springer-Verlag, Berlin, 1965).

[3] S. Lie and G. Scheffers, *Vorlesungen über continuierliche Gruppen* (Teubner Verlagsgesellschaft, Leipzig, 1893).

[4] E. Cartan, "Sur la structure des groupes de transformation finis et continus," Ph.D. thesis, Nony, Paris (1894).

[5] H. Weyl, *Gruppentheorie und Quantenmechanike* (S. Hirzel Verlag, Leipzig, 1931); translated by H. P. Robertson, *The Theory of Groups and Quantum Mechanics* (reprinted by Dover Publications, Inc., New York).

[6] H. Weyl, *The Classical Groups* (Princeton University Press, Princeton, N.J., 1946).

[7] H. A. Jahn, Proc. Roy. Soc. (London) **A201**, 516 (1950).

[8] B. H. Flowers, Proc. Roy. Soc. (London) **A212**, 248 (1952).

[9] J. P. Elliott, Proc. Roy. Soc. (London) **A245**, 128 (1958).

velocity at the detonation front $u - c$ is almost constant also. This is in agreement with experiments which indicate a near constant detonation velocity in converging cylindrical detonations.

Perhaps the circle collapses to a pattern with only a small number of points because the total distance around such a pattern is only slightly larger than the distance around the circumscribing circle. With this distance about the same and the detonation velocity nearly constant, the rate of release of chemical energy is nearly constant during the final part of the detonation. This would suggest that the spherical detonation converges to a pattern corresponding to the collapse of an octahedron or perhaps a cube or a dodecahedron.

APPENDIX: A MORE GENERAL EQUATION OF STATE

To fit experimental data better, the equation of state

$$p = A\rho^3 e^{\delta\rho} \tag{A1}$$

can be used in place of Eq. (11).

Equation (A1) is substituted into Eqs. (9), (10), and (16), and for δ small, terms higher than the first power in δ can be dropped. The result for spherical coordinates is

$$\frac{2}{\omega_0}\left[\frac{3\omega}{\omega_0} - \frac{2\delta\omega^2}{3^{\frac{1}{2}}} + \frac{5\delta\omega^3}{2\omega_0}\right] d\omega \Big/ \left[3 - 2(3)^{\frac{1}{2}}\frac{\omega}{\omega_0} + \frac{3\delta\omega^2}{2} - \frac{11\delta\omega^3}{4(3\omega_0)^{\frac{1}{2}}}\right] = \frac{dr}{r}, \tag{A2}$$

where $\omega^2 = \rho$ and $\omega_0^2 = \rho_0$.

From Eqs. (10), (16), and (A1), for δ small, the fluid velocity becomes

$$u = A\omega^2\left(3^{\frac{1}{2}} - \frac{2\omega}{\omega_0} + \frac{2\omega^2\delta}{3^{\frac{1}{2}}} - \frac{5\omega^3\delta}{4\omega_0}\right). \tag{A3}$$

Equation (A2) can be integrated numerically by the use of a computer. Equation (A3) and the numerical integration of Eq. (A2) then express u and r in terms of the parameter ω. For $\delta = 0.1/\omega_0^2$, the resulting curve is very similar to Fig. 2.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 7 JULY 1968

Plethysm and the Theory of Complex Spectra*

P. R. SMITH† AND B. G. WYBOURNE
Physics Department, University of Canterbury, Christchurch, New Zealand

(Received 26 August 1967)

Notable progress in developing the theory of complex spectra has come from the exploitation of the symmetry properties of atomic wavefunctions using the theory of continuous groups. Many seemingly simple results have previously been obtained by what would appear to be unreasonably complex methods. Many of these complexities may be removed, and new results derived, using the algebra of *plethysm* first developed by Littlewood. Applications to three central problems in the theory of complex spectra are discussed: (1) the classification of the atomic states of n-electron configurations; (2) the analysis and classification of the N-particle operators that arise in the application of perturbation theory to atomic problems; (3) the derivation of selection rules for the matrix elements of operators.

I. INTRODUCTION

The analysis and interpretation of the spectra of atoms and nuclei has drawn heavily upon parallel developments in the theory of complex spectra. Notable progress in developing the theory of complex spectra came from Racah's[1,2] exploitation of the symmetry properties of atomic wavefunctions using the classical theory of continuous groups as developed by Lie,[3] Cartan,[4] Weyl,[5,6] and others. Later developments[7–9] have made considerable use of Young's

* Research sponsored in part by the Air Force Office of Scientific Research, Office of Aerospace Research, United States Air Force, under AFSOR Grant No. 1275-67.

† Present address: Physics Department, Cambridge University, Cambridge, England.

[1] G. Racah, Phys. Rev. **76**, 1352 (1949).

[2] G. Racah, *Group Theory and Spectroscopy, Ergebnisse der exacten Naturwissenschaften, Vol. 37* (Springer-Verlag, Berlin, 1965).

[3] S. Lie and G. Scheffers, *Vorlesungen über continuierliche Gruppen* (Teubner Verlagsgesellschaft, Leipzig, 1893).

[4] E. Cartan, "Sur la structure des groupes de transformation finis et continus," Ph.D. thesis, Nony, Paris (1894).

[5] H. Weyl, *Gruppentheorie und Quantenmechanike* (S. Hirzel Verlag, Leipzig, 1931); translated by H. P. Robertson, *The Theory of Groups and Quantum Mechanics* (reprinted by Dover Publications, Inc., New York).

[6] H. Weyl, *The Classical Groups* (Princeton University Press, Princeton, N.J., 1946).

[7] H. A. Jahn, Proc. Roy. Soc. (London) **A201**, 516 (1950).

[8] B. H. Flowers, Proc. Roy. Soc. (London) **A212**, 248 (1952).

[9] J. P. Elliott, Proc. Roy. Soc. (London) **A245**, 128 (1958).

theory of substitutional analysis[10,11] and Littlewood's[12] theory of S functions, emphasizing the close connection between the theory of the symmetric group and of the continuous groups. These later developments have been summarized by Judd.[13]

Recent studies by Rajnak and Wybourne,[14] Judd,[15,16] and Feneuille[17] of the representation of configuration interactions by effective operators have again made extensive use of the theory of continuous groups to discuss the structure and classification of N-particle scalar operators acting within a configuration of equivalent electrons. Further applications of group theory to the elucidation of selection rules using the conventional Wigner–Eckart theorem have been made by Judd and Wadzinski.[18] Inspection of these recent studies reveals that many seemingly simple results have required what would appear to be excessively complex derivations which conceal the underlying simplicities of the theory of complex spectra.

Many of the apparent complexities found in the theory of complex spectra may be removed by use of Littlewood's[12,19–23] algebra of *plethysm*, rather than the conventional applications of the theory of continuous groups. The basic operation of plethysm has been defined in a previous paper[24,25] and methods of calculation for plethysms in terms of the group characters of the full linear group, and those of restricted groups, was outlined. The notation used throughout this paper coincides with that of Littlewood.[12,19,20] The theory of plethysm forms the mathematical basis of a method for obtaining all the results previously found using the theory of continuous groups in more than three dimensions, without considering the detailed operations of these higher groups.

In the present paper three central problems in the theory of complex spectra are discussed from the mathematical viewpoint of plethysm. In order of presentation they are:

(1) the classification of the atomic states of n-electron configurations;
(2) the analysis and classification of the N-particle operators that arise in the application of perturbation theory to atomic problems;
(3) the derivation of selection rules for the matrix elements of operators.

II. CLASSIFICATION OF ATOMIC STATES

The conventional group-theoretical methods used to classify the atomic states of equivalent electron configurations l^n have been reviewed by Judd.[13] These methods usually commence by first decomposing a few simple representations of the $(2l + 1)$-dimensional unitary group U_{2l+1} into those of the three-dimensional rotation group R_3. Further branching rules are then established by resolving Kronecker products of these representations and then proceeding with a chain calculation. The resolution of these Kronecker products is seldom obvious, even in quite simple cases.

For example, in the classification of the states of the f^3 configuration one readily obtains the $U_7 \rightarrow R_3$ branching rule:

$$\{111\} + \{210\} \rightarrow SPD_3F_3G_3H_2I_2KL. \qquad (1)$$

It is by no means obvious which of the representations of R_3 are to be associated with the $\{111\}$ representation of U_7 and which with the $\{210\}$ representation. Judd avoided this difficulty by making use of the method of determinantal states to find the states of f^3 with spin $S = \frac{3}{2}$ to give

$$\{111\} \rightarrow SDFGI, \qquad (2)$$

and then obtained the decomposition

$$\{210\} \rightarrow PD_2F_2G_2H_2IKL. \qquad (3)$$

The above approach is unsatisfying because it requires the introduction of an additional method that is alien to the idea of using group theory to classify atomic states. The shortcomings of the conventional method may be completely avoided by the application of the method of plethysm.

A. Plethysm and the Classification of States

The total wavefunction for a single electron can be written as the product of an orbital function and a spin function. If the orbital function transforms according to the representation $[l]$ of R_3, then the set of all possible wavefunctions for an electron with orbital quantum number e spans the representation

[10] D. E. Rutherford, *Substitutional Analysis* (Edinburgh University Press, Edinburgh, 1948).
[11] G. de B. Robinson, *Representation Theory of the Symmetric Group* (Edinburgh University Press, Edinburgh, 1961).
[12] D. E. Littlewood, *The Theory of Group Characters* (Oxford University Press, Oxford, England, 1950), 2nd ed.
[13] B. R. Judd, *Operator Techniques in Atomic Spectroscopy* (McGraw-Hill Book Company, Inc., New York, 1963).
[14] K. Rajnak and B. G. Wybourne, Phys. Rev. **132**, 280 (1963).
[15] B. R. Judd, Phys. Rev. **141**, 4 (1966).
[16] B. R. Judd, *Second Quantization and Atomic Spectroscopy* (The Johns Hopkins Press, Baltimore, Md., 1967).
[17] S. Feneuille, Compt. Rend. **262**, 23 (1966); J. Phys. **28**, 315 (1967).
[18] B. R. Judd and H. T. Wadzinski, J. Math. Phys. **8**, 2125 (1967).
[19] D. E. Littlewood, Phil. Trans. Roy. Soc. (London) **A239**, 305 (1944).
[20] D. E. Littlewood, Phil. Trans. Roy. Soc. (London) **A239**, 287 (1944).
[21] D. E. Littlewood, Proc. London Math. Soc. (2) **50**, 349 (1948).
[22] D. E. Littlewood, J. London Math. Soc. **30**, 121 (1955).
[23] D. E. Littlewood, Can. J. Math., **10**, 17 (1958).
[24] P. R. Smith and B. G. Wybourne, J. Math. Phys. **8**, 2434 (1967).
[25] P. R. Smith, "The Use of Plethysm in the Study of Configurations of Equivalent Electrons," M.Sc. thesis, University of Canterbury, Christchurch, New Zealand (1967).

$[l][\frac{1}{2}]$ of R_3, where $[\frac{1}{2}]$ is the spin representation spanned by the two spin functions.

The functions which describe the n-particle configuration are simply the antisymmetric functions obtained by forming products of single-particle functions defined on each of the n sets of coordinates. These functions transform according to the representation

$$([l][\tfrac{1}{2}]) \otimes \{1^n\} \tag{4}$$

of R_3, which is generally reducible. The analysis of the separation of this plethysm provides the classification of the atomic states.

B. Seniority Classification

The seniority classification of the states may be obtained directly by expressing the symbol $\{1^n\}$ in terms of symplectic symbols[20]:

$$\{1^n\} = \sum_{a=0} \langle 1^{n-2a} \rangle, \tag{5}$$

where the a's are positive integers such that $n \geq 2a$. The plethysm of Eq. (4) may then be written as

$$\begin{aligned}([l][\tfrac{1}{2}]) \otimes \{1^n\} &= \sum_a ([l][\tfrac{1}{2}]) \otimes \langle 1^{n-2a} \rangle \\ &= \sum_a ([l][\tfrac{1}{2}]) \otimes (\{1^{n-2a}\} - \{1^{n-4a}\}).\end{aligned} \tag{6}$$

The representations appearing in the separation of the terms corresponding to each value of a may be classified either with the symbol $\langle 1^{n-2a} \rangle$, or, equivalently, with the seniority number

$$v = n - 2a. \tag{7}$$

The classification obtained at this stage is independent of both the coupling scheme and the value of n for a given $[l]$.

C. *LS* Coupling

In this coupling scheme the orbital and spin spaces of the wavefunction are considered to be uncoupled for the purposes of classification. As a result the single-particle wavefunctions are said to transform according to the representation $[l][\frac{1}{2}]'$, where the prime indicates that transformations of the orbital and spin functions are to be considered independently. Correspondingly, the n-particle functions transform according to the representation

$$([l][\tfrac{1}{2}]') \otimes \{1^n\}. \tag{8}$$

This plethysm may be expanded using the formulas given by Littlewood[19,26] for the plethysm of a product of functions defined on different sets of variables as follows:

$$([l][\tfrac{1}{2}]') \otimes \{1^n\} = \sum_{\alpha,\beta} K_{\alpha\beta\{1^n\}}([l] \otimes \{\alpha\})([\tfrac{1}{2}]' \otimes \{\beta\}), \tag{9}$$

where $K_{\alpha\beta\{1^n\}}$ is the coefficient of $\{1^n\}$ in the reduction of the inner product

$$\{\alpha\} \circ \{\beta\} = \sum_\gamma K_{\alpha\beta\gamma}\{\gamma\} \tag{10}$$

of the S functions $\{\alpha\}$ and $\{\beta\}$. However, $\{1^n\}$ appears once only in products of the type $\{\alpha\} \circ \{\tilde{\alpha}\}$ and in no other products.[27] Thus Eq. (9) reduces to

$$([l][\tfrac{1}{2}]') \otimes \{1^n\} = \sum_\alpha ([l] \otimes \{\alpha\})([\tfrac{1}{2}]' \otimes \{\tilde{\alpha}\}). \tag{11}$$

The isomorphism between the binary unitary group and the ternary orthogonal group[21] can now be used to solve the plethysm $[\frac{1}{2}]' \otimes \{\tilde{\alpha}\}$. Now, $[\frac{1}{2}]$ of $R_3 \leftrightarrow \{1\}$ of U_2, and hence $[\frac{1}{2}] \otimes \{\tilde{\alpha}\} \leftrightarrow \{1\} \otimes \{\tilde{\alpha}\} = \{\tilde{\alpha}\} = \{\tilde{\alpha}_1 - \tilde{\alpha}_2\} \leftrightarrow [\frac{1}{2}(\tilde{\alpha}_1 - \tilde{\alpha}_2)]$. Thus the plethysm is null whenever $\{\tilde{\alpha}\}$ corresponds to a partition into more than two parts. The classification is therefore given by the expression

$$([l][\tfrac{1}{2}]') \otimes \{1^n\} = \sum_a ([l] \otimes \{2^a 1^{n-2a}\})[\tfrac{1}{2}(n - 2a)]', \tag{12}$$

where the sum is taken over all positive integers a such that $n \geq 2a$.

D. *jj* Coupling

In this coupling scheme the representation spaces of the orbital and spin functions are taken to be coupled, and so the product $[l][\frac{1}{2}]$ in the plethysm $([l][\frac{1}{2}]) \otimes \{1^n\}$ giving the n-particle classification can be rewritten as the sum of two spin representations of the rotation group, where

$$[l][\tfrac{1}{2}] = [l + \tfrac{1}{2}] + [l - \tfrac{1}{2}]. \tag{13}$$

The plethysm giving the classification for the n-particle configuration can then be expanded by using the formula for the sum of two functions defined on the same variables; this gives

$$\begin{aligned}([l][\tfrac{1}{2}]) \otimes \{1^n\} &= ([l + \tfrac{1}{2}] + [l - \tfrac{1}{2}]) \otimes \{1^n\} \\ &= \sum_{a=0}^{n} ([l + \tfrac{1}{2}] \otimes \{1^a\})([l - \tfrac{1}{2}] \otimes \{1^{n-a}\}).\end{aligned} \tag{14}$$

The plethysms $[l + \frac{1}{2}] \otimes \{1^a\}$ can be solved using the isomorphism between R_3 and U_2 and considering the plethysms $\{2l + 1\} \otimes \{1^a\}$ in U_2. The states appearing in the *jj* coupling classification of the n-particle configuration are then given by the expansion of Eq. (14).

[26] D. E. Littlewood, J. London Math. Soc. **32**, 18 (1957).

[27] M. Hamermesh, *Group Theory* (Addison-Wesley Publishing Company, Reading, Mass., 1962).

E. *LL* Coupling

Judd[28] has introduced a new concept in the coupling of the spin and orbital representations in the product $[l][\frac{1}{2}]$ in the analysis of the l^n configurations. Two spaces are considered, one in which all the spins are up and the other in which all are down. This means that the product $[l][\frac{1}{2}]$ is rewritten as $([l] + [l'])$ where $[l]$ corresponds to the orbital functions in "*spin-up space*" and $[l]'$ to the set in "*spin-down spaces.*" An operation of the spin space as a whole, in this coupling scheme, must be interpreted as an operation which exchanges functions between "*spin-up*" and "*spin-down*" spaces.

The wavefunctions span the representation $([l] + [l]')$ of the imprimitive rotation group in two sets of variables. An imprimitive group is a group which contains all the operations of a group G on two sets of variables together with an operation which exchanges these two sets.[20] The plethysm for the n-particle configuration is, therefore,

$$([l] + [l']) \otimes \{1^n\}, \tag{15}$$

which can be expanded by using the rule for the plethysm of a sum[19,24] to yield

$$([l] + [l']) \otimes \{1^n\} = \sum_{a=0}^{n} ([l] \otimes \{1^a\})([l]' \otimes \{1^{n-a}\}). \tag{16}$$

The states obtained in this classification scheme may be labeled by the quantum numbers

$$|\tau\tau'\{1^a\}\{1^{n-a}\}'[L][L'']'LM_L\rangle$$

or, equivalently,

$$|l^n\tau\tau'[L][L'']'LM_LM_S\rangle, \tag{17}$$

since $M_S = \frac{1}{2}n - a$. The quantum numbers $[L]$ and $[L'']'$ are the orbital numbers of the "*spin-up*" and "*spin-down*" spaces, respectively. L is the total angular momentum obtained by coupling $[L]$ and $[L'']'$, M_L is the projection of the orbital quantum number L, and τ and τ' are used as auxiliary labels when required.

The purity of a state in *LL* coupling can be taken as an indication of the accuracy of the approximation where it is assumed that $[l][\frac{1}{2}]$ may be written as $([l] + [l]')$. This approximation only holds in fact when the operations of R_3 in the total spin space are restricted to two operations, the identity and 180° rotation about a line in the xy plane. It would seem that a pure state in *LL* coupling results from the restriction of the operations in the spin space to those of this simple finite group—a fact that would seem to be of considerable interest in the study of ions whose eigenfunctions are of this type.

[28] B. R. Judd, Bull. Am. Phys. Soc. **12**, 515 (1967).

F. Further Classification of States in *LS* Coupling

The introduction of seniority gives a classification of the states in both *LS* and *jj* coupling which reflects the symmetry of both the orbital and spin parts of the corresponding wavefunctions. In *LS* coupling a further classification is obtained by the detailed expansion of Eq. (6) to give the traditional description in terms of the quantum numbers S and L.

When $l \leq 2$, the quantum numbers vSL suffice to label uniquely the states of the l^n configuration. For $l \geq 3$, additional quantum numbers must be found. In these cases it is useful to express the symmetry symbol $\{\alpha\}$ which appears in Eq. (6) as a sum of symbols corresponding to the characters of the rotation group in $2l + 1$ dimensions; this gives

$$([l][\tfrac{1}{2}]') \otimes \{1^n\} = \sum_{\alpha,\gamma,\mu} \Gamma_{\nu\mu\alpha}([l] \otimes [\mu])[\tfrac{1}{2}(\tilde{\alpha}_1 - \tilde{\alpha}_2)]'. \tag{18}$$

The symbol $[\mu]$ may then be used to label the orbital states which appear in the plethysm $([e] \otimes [\mu])$.

When $l = 3$, a further classification of states is possible under the group G_2, a proper subgroup of the group R_7. In this case the symmetry symbols $[\mu]$ in the plethysm $[3] \otimes [\mu]$ may be rewritten in terms of symbols (λ) corresponding to characters of the representations of G_2. These characters are obtained by the decomposition rule

$$[\lambda_1, \lambda_2, \lambda_3] = \sum \{(i-k, j+k) + (j-k-1, i-j)\}, \tag{19}$$

the sum being taken over all i, j, k such that $\lambda_1 \geq i \geq \lambda_2$, $\lambda_2 \geq j \geq \lambda_3$, $\lambda_3 \geq k \geq -\lambda_3$. Improper symbols (where $u_2 \geq u_1$ or $u_2 < 0$) may be rewritten using the rules given by Judd[13]:

$$(\mu_1, \mu_2) = -(\mu_2 - 1, \mu_1 + 1); \quad (\mu_1, -1) = 0; \quad (\mu_1, -2) = (\mu_1 - 1, 0). \tag{20}$$

G. Nonstandard Symbols

One further point must be considered before completing the classification of the states of l^n configurations. When $n > l$, the plethysms involved in making the classification frequently yield nonstandard symbols on the right. In these cases modification rules[20,24] must be used to yield the standard classification. However, it is worth noting that a classification of n-particle states in terms of nonstandard symbols does not result in a loss of information about the functions they classify; they do, however, make this information less accessible.

When $n \geq 2l + 1$, the states of the configurations l^n and l^{4l+2-n} have identical classifications; for values of $n > 4l + 2$ all the plethysms are null.

H. Application to the f^3 Configuration

The states of the f^3 configuration in LS coupling are derived from the plethysm (see Eq. 12)

$$([3][\tfrac{1}{2}]') \otimes \{1^3\} = [3] \otimes \{1^3\}[\tfrac{3}{2}]' + [3] \otimes \{21\}[\tfrac{1}{2}]'. \quad (21)$$

Since in R_3

$$[3] \otimes \{1^3\} = [0] + [2] + [3] + [4] + [6]$$

and

$$[3] \otimes \{21\} = [1] + 2[2] + [3] + 2[4] + 2[5] + [6] + [7] + [8],$$

it is apparent that the states given by Eq. (1) can, in fact, be written as

$$^4(SDFGI)^2(PD_2F_2G_2H_2IKL). \quad (22)$$

Equation (18) may then be used to classify the above states according to the symmetry symbols $[\mu]$ of the group R_7. The plethysms $[3] \otimes [\mu]$ may then be rewritten in terms of symbols (λ) of the group G_2 to yield the final classification given in Table I, which is identical to that obtained by Judd[13] by the conventional method.

The seniority classification may be derived directly from Eq. (6). States of seniority $v = 3$ are given by $([3][\tfrac{1}{2}]') \otimes (\{1^3\} - \{1\})$ and those of seniority $v = 1$ by $([3][\tfrac{1}{2}]') \otimes \{1\} = [3][\tfrac{1}{2}]'$ corresponding to the ^2F states. Thus there is only one state of seniority $v = 1$, and the remainders are all of seniority $v = 3$.

In jj coupling the configuration f^3 is described by the states appearing in the separation of the plethysm [see Eq. (14)]:

$$\begin{aligned}([3][\tfrac{1}{2}]) \otimes \{1^3\} &= ([\tfrac{7}{2}] + [\tfrac{5}{2}]) \otimes \{1^3\} \\ &= [\tfrac{7}{2}] \otimes \{1^3\} + [\tfrac{5}{2}]([\tfrac{7}{2}] \otimes \{1^2\}) \\ &\quad + ([\tfrac{5}{2}] \otimes \{1^2\})[\tfrac{7}{2}] + [\tfrac{5}{2}] \otimes \{1^3\}.\end{aligned}$$

TABLE I. Classification of the states of the f^3 configuration.[a]

$\{\alpha\}$	$[\mu]$	(λ)	^{2S+1}L
$\{1^3\}$	$[1^3]$	(0)	^4S
		(1)	^4F
		(2)	$^4(DGI)$
$\{21\}$	[1]	(1)	^2F
	[21]	(11)	$^2(PH)$
		(2)	$^2(DGI)$
		(21)	$^2(DFGHKL)$

[a] $\{\alpha\}$ labels the symmetry symbols of U_7, $[\mu]$ those of R_7 and (λ) those of G_2.

TABLE II. Classification[a] of the states of $(d + s)^3$.

$\{\alpha\}$	$[\lambda]$	$[\mu]$	^{2S+1}L
$\{1^3\}$	[111]	[11]	$^4(PF)$
	[11 − 1]	[11]	$^4(PF)$
$\{21\}$	[21]	[1]	^2D
		[11]	$^2(PF)$
		[20]	$^2(DG)$
		[21]	$^2(PDFGH)$
	[1]	[0]	^2S
		[1]	^2D

[a] $\{\alpha\}$ labels the symmetry symbols for U_6, $[\lambda]$ those of R_6, and $[\mu]$ those of R_5.

The plethysms appearing in this expansion are

$$\begin{aligned}[\tfrac{7}{2}] \otimes \{1^3\} &= [\tfrac{3}{2}] + [\tfrac{5}{2}] + [\tfrac{7}{2}] + [\tfrac{9}{2}] + [\tfrac{11}{2}] + [\tfrac{15}{2}], \\ [\tfrac{7}{2}] \otimes \{1^2\} &= [0] + [2] + [4] + [6], \\ [\tfrac{5}{2}] \otimes \{1^3\} &= [\tfrac{3}{2}] + [\tfrac{5}{2}] + [\tfrac{9}{2}], \\ [\tfrac{5}{2}] \otimes \{1^2\} &= [0] + [2] + [4].\end{aligned}$$

Thus the J structure of the configuration follows immediately. The terms of seniority $v = 1$ are given by $[\tfrac{7}{2}] + [\tfrac{5}{2}]$.

I. Mixed Configurations

The methods of plethysm may also be readily applied to the classification of the states of mixed configuration of the general type $(\sum [l_i])^n$, a typical example being the states of the $(d + s)^n$ configurations. In this case the LS coupling states are found by expansion of the plethysm

$$\begin{aligned}&([2][\tfrac{1}{2}]') \otimes \{1^n\} \\ &\quad = \sum_{a=0} (\{2\} \otimes \{2^a 1^{n-2a}\})([\tfrac{1}{2}]' \otimes \{n - a, a\}) \\ &\quad = \sum_{a=0} (\{2\} \otimes \{2^a 1^{n-2a}\})[\tfrac{1}{2}(n - 2a)]'. \quad (23)\end{aligned}$$

Thus, for the particular case of $(d + s)^3$,

$$([2][\tfrac{1}{2}]') \otimes \{1^3\} = (\{2\} \otimes \{1^3\})[\tfrac{3}{2}]' + (\{2\} \otimes \{21\})[\tfrac{1}{2}]',$$

and since in R_3

$$\{2\} \otimes \{1^3\} = 2[3] + 2[1]$$

and

$$\{2\} \otimes \{21\} = [5] + 2[4] + 2[3] + 4[2] + 2[1] + [0],$$

the LS coupling states of $(d + s)^3$ are found to be

$$^4(P_2F_2)^2(SP_2D_4F_2G_2H).$$

This classification may then be extended by expressing $\{1^3\}$ and $\{21\}$ as sums of symmetry symbols $[\lambda]$ for R_6 and then $[\mu]$ of R_5; the results are given in Table II, in agreement with those found by Elliott,[9] Judd,[13] and Feneuille.[29]

[29] S. Feneuille, J. Phys. **28**, 61 (1967).

The states of the $(d+s)^n$ configurations may be uniquely classified by the set of quantum numbers

$$|\{\alpha\}[\lambda][\mu]SLM_SM_L\rangle.$$

Coupling schemes of this type are likely to be of some importance when the configurations d^n, $d^{n-1}s$, and $d^{n-2}s^2$ are energetically close together—as is indeed the case for the iron-group atoms. The development of classification schemes for mixed configurations is only useful when the n electron orbitals are all of the same parity.

III. CLASSIFICATION OF N-PARTICLE OPERATORS

The use of scalar N-particle operators to simplify the analysis of the effect of configuration interaction on the l^n configuration of n equivalent electrons has created a need for a systematic method of classifying their transformation properties. The method of plethysm provides a simple and unambiguous solution to this problem.

An N-particle operator is any operator formed from N single-particle operators which is symmetric with respect to the interchange of the coordinates of any two of the particles. An operator formed in this way is not necessarily a scalar under the operations of R_3.

A. Plethysm for N-Particle Scalar Operators

The single-particle operators $\mathbf{v}_i$ chosen to form the N-particle operator must have nonvanishing matrix elements between single-particle states of the l^n configuration. This restricts the set of operators $\mathbf{v}_i$ which may be used to construct valid N-particle operators for the l^n configuration to those that span the representation $[l][l]$ of R_3. If these operators are denoted by $\mathbf{v}^{(k)}$ where $[k]$ is a representation of R_3, then $2l+1 \geq k \geq 0$, and they may be conveniently identified with the single-electron tensor operators $\mathbf{v}^{(k)}$, whose components $v_q^{(k)}$ are defined by the equation[16]

$$(lm|\, v_q^{(k)}\,|lm') = (-1)^{l-m}(2k+1)^{\frac{1}{2}}\begin{pmatrix} l & k & l \\ -m & q & m' \end{pmatrix}. \tag{24}$$

If $\mathcal{A}$ represents a set of operators chosen from the set of $v_q^{(k)}$'s defined above, then it is clear that the representation of R_3 spanned by the completely symmetric linear combinations of products of degree N in the single-particle operators are given by the plethysm $\mathcal{A} \otimes \{N\}$. In particular, if the full set of operators transforming according to the representation $[l][l]$ of R_3 is used, the complete set of all possible N-particle operators is given by

$$[l][l] \otimes \{N\}. \tag{25}$$

Every operator in the set which is constructed so that it involves $\mathbf{v}^{(0)}$ α times is equivalent to an $(N-\alpha)$-particle operator. It is convenient to exclude these reducible operators and to consider only those operators which are not equivalent to operators involving a few number of particles. If the operator $\mathbf{v}_0^{(0)}$ is excluded, then the plethysm

$$\left(\sum_{k=1}^{2l+1}[k]\right) \otimes \{N\} \tag{26}$$

serves to classify the irreducible N-particle operators.

A prescription for the formation and classification of irreducible N-particle operators can be obtained by noting that

$$[l][l] = [l] \otimes (\{2\} + \{11\}), \tag{27}$$

where the symmetry symbol $\{2\}$ produces the representations $[k]$ where k is *even* and $\{11\}$ produces the representations $[k]$ where k is *odd*. If these symbols are rewritten in terms of rotation symbols, Eq. (27) becomes

$$[l][l] = [l] \otimes ([0] + [2] + [11]). \tag{28}$$

The operator $v_0^{(0)}$ may thus be excluded in the classification of irreducible N-particle operators by simply excluding the symmetry symbol [0] in the plethysm. When this is done, the plethysm appropriate to the classification of irreducible N-particle operators becomes

$$[l] \otimes ([2] + [11]) \otimes \{N\}. \tag{29}$$

If the plethysm

$$([2] + [11]) \otimes \{N\} \tag{30}$$

is performed, first a set of symmetry symbols $[\mu]$ is obtained which can be used to classify the operators that transform according to the representations appearing in the plethysm

$$[l] \otimes [\mu] \tag{31}$$

in R_3. In general, some of these symbols are nonstandard and have to be treated with modification rules, if standard symbols are desired.

B. Classification of N-Particle Scalar Operators

In a perturbation theory that is restricted to Coulomb interactions, only those operators that transform as scalars need to be considered. As a result only those plethysms $[l] \otimes [\mu]$ which yield a [0] representation in R_3 are selected from the plethysm of Eq. (29). For each one of these [0]'s there is an irreducible N-particle scalar operator that can be classified with the symmetry symbol $[\mu]$. In some cases it may be possible to obtain a finer classification by expanding the symmetry symbols $[\mu]$ of the

$(2l+1)$-dimensional rotation group into the characters of a subgroup.

If the plethysm of Eq. (29) is expanded as follows,

$$[l] \otimes ([2] + [11]) \otimes \{N\} = [l] \otimes \left[\sum_{\alpha=0}^{N} ([2] \otimes \{\alpha\})([11] \otimes \{N-\alpha\})\right], \quad (32)$$

it becomes clear that the operators which transform according to the representations given by the plethysm

$$[l] \otimes [([2] \otimes \{\alpha\})([11] \otimes \{N-\alpha\})] \quad (33)$$

involve α single-particle operators with k *even* and $(N-\alpha)$ with k *odd*. If the letter g is used to designate *even*-rank operators and u is used for *odd*-rank operators, then the operators arising from the above plethysm are said to be of type $g \cdots g$, $u \cdots u$ where there are α g's and $(N-\alpha)$ u's.

Irreducible N-particle operators involving products of an *odd* number of tensor operators of *odd* rank [i.e., $(N-\alpha)$ is *odd*] are non-Hermitian and, as such, would not be expected to appear in a perturbation expansion. While it is clear that only Hermitian operators can arise in second-order perturbation theory, it has not, as yet, been unequivocally demonstrated that effective non-Hermitian operators cannot arise in higher orders of perturbation theory. However, it should be possible to develop perturbation expansions that exclude the possibility of non-Hermitian operators.

The symmetry symbols occurring in the plethysm of Eq. (33) may be readily interpreted in terms of conventional group theory. The N-particle operators span a $4l(l+1)$-dimensional space and, in fact, belong to the symmetric $\{N\}$ representation of the unitary group $U_{4l(l+1)}$. The symmetry symbols $\{\alpha\}$ and $\{N-\alpha\}$ of Eq. (33) thus label the symmetric representations $\{\alpha\}$ of the unitary group $U^A_{l(2l+3)}$ and $\{N-\alpha\}$ of $U^B_{l(2l+1)}$, respectively. Since the direct-product group $U^A_{l(2l+3)} \times U^B_{l(2l+1)}$ is a subgroup of $U_{4l(l+1)}$, the plethysms $[2] \otimes \{\alpha\}$ and $[11] \otimes \{N-\alpha\}$ can be interpreted as particular decompositions of the representations $\{\alpha\}$ of $U^A_{l(2l+3)}$ and $\{N-\alpha\}$ of $U^B_{l(2l+1)}$ into those of R_{2l+1}. Schematically, the resolution of the plethysm may be interpreted as selecting certain of the representations derived from the group decomposition

$$\begin{array}{l} U_{4l(l+1)} \rightarrow U^A_{l(2l+3)} \times U^B_{l(2l+1)} \rightarrow \\ \qquad\qquad\qquad\qquad\qquad\qquad R_{2l+1} \\ \qquad\qquad\qquad\qquad\qquad\quad \nearrow \quad \searrow \\ R^A_{2l+1} \times R^B_{2l+1} \qquad\qquad\qquad R_3 \\ \qquad\qquad\qquad\qquad\qquad\quad \searrow \quad \nearrow \\ \qquad\qquad\qquad\qquad\qquad\qquad R^A_3 \times R^B_3 \end{array} \quad (34)$$

of the symmetric representation $\{N\}$ of $U_{4l(l+1)}$.

When $\alpha = N$ in Eq. (32), the N-particle operators (which will be constructed entirely from N-tuple products of *even* rank single-particle operators) may be labeled by the representations $[\lambda]$ of R^A_{2l+1} contained in the decomposition of the representation $\{N\}$ of $U^B_{l(2l+1)}$ according to the plethysm $[11] \otimes \{N\}$. In all other cases the operators are labeled by the sequence

$$([\lambda][\lambda'])[\mu]O,$$

where $[\lambda]$ is the representation of R^A_{2l+1} contained in the decomposition of $\{\alpha\}$ of $U^A_{l(2l+3)}$ according to the plethysm $[2] \otimes \{\alpha\}$, $[\lambda']$ that of R^B_{2l+1} contained in the decomposition of $\{N-\alpha\}$ of $U^B_{l(2l+1)}$ according to the plethysm $[11] \otimes \{N-\mathrm{a}\}$, $[\mu]$ the representation of R_{2l+1} contained in the reduction of the product representation $[\lambda][\lambda']$, and where O indicates the identity representation $D^{(0)}$ of R_3 obtained in performing the plethysm $[l] \otimes [\mu]$.

The typical plethysms $[2] \otimes \{n\}$ and $[11] \otimes \{n\}$ occurring in Eq. (32) may be readily evaluated by noting that

$$\{2\} \otimes \{n\} = \sum \{\mu\} \quad (35)$$

and

$$[11] \otimes \{n\} = \{11\} \otimes \{n\} = \sum \{\tilde{\mu}\}, \quad (36)$$

where the summation in Eq. (35) is over all partitions (μ) of $2n$ into even parts, while, in Eq. (36), the summation is over the partitions $(\tilde{\mu})$ conjugate to those arising in Eq. (35). The plethysm $[2] \otimes \{n\}$ is then obtained as

$$[2] \otimes \{n\} = \{2\} \otimes (\{n\} - \{n-1\}). \quad (37)$$

C. Classification of N-Particle Scalar Operators for d^n Electron Configurations

The plethysms of Eq. (32) for $l = 2$ may be readily performed for $N = 0$, 2, 3, and 4 to give the classification of N-particle scalar operators for d^n electron configurations; this is listed in Table III. This list contains 48 distinct N-particle scalar operators and represents the maximum number of effective scalar operators with $N \leq 4$ that can arise in a perturbation expansion limited to scalar Coulomb interactions.

Information obtained from decomposing each of the representations $[\lambda][\lambda']$ occurring in $R^A_5 \times R^B_5$ into those of $R^A_3 \times R^B_3$ and then into those of R_3, as indicated by the lower branch of Eq. (34), is useful in the explicit construction of the N-particle operators.

The three-particle scalar operator t_{10} transforms as $([22][11])[22]O$. The representations [2] and [11] of R_5 decompose as

$$[2] \rightarrow D + G \quad \text{and} \quad [11] \rightarrow P + F,$$

under restriction to R_3. In this case a three-particle

TABLE III. Classification of scalar N-particle operators ($N \leq 4$) for d^n electron configurations.

N	Plethysm	Label	Classification of Scalars	Type
0	$[2] \otimes ([0] \otimes \{0\})$	e_0	$[0]O$	
2	$[2] \otimes ([2] \otimes \{2\})$	e_1	$[0]O$	gg
		e_2	$[22]O$	
	$[2] \otimes ([11] \otimes \{2\})$	e_3	$[0]O$	uu
		e_4	$[22]O$	
3	$[2] \otimes ([2] \otimes \{3\})$	t_1	$[0]O$	ggg
		t_2	$[22]O$	
		t_3	$[42]O$	
		t_4	$[6]O$	
	$[2] \otimes [[2]([11] \otimes \{2\})]$	t_5	$([2][2])[0]O$	guu
		t_6	$([2][2])[22]O$	
		t_7	$([2][22])[22]O$	
		t_8	$([2][11])[3]O$	
		t_9	$([2][22])[42]O$	
	$[2] \otimes [([2] \otimes \{2\})[11]]$	t_{10}	$([22][11])[22]O$	ggu
4	$[2] \otimes ([2] \otimes \{4\})$	f_1, f_2	$[0]O$	$gggg$
		f_3, f_4	$[22]O$	
		f_5, f_6	$[42]O$	
		f_7	$[44]O$	
		f_8	$[6]O$	
	$[2] \otimes [([2] \otimes \{2\})([11] \otimes \{2\})]$	f_9	$([0][0])[0]O$	$gguu$
		f_{10}	$([2][2])[0]O$	
		f_{11}	$([22][22])[0]O$	
		f_{12}	$([22][1])[22]O$	
		f_{13}	$([22][2])[22]O$	
		f_{14}	$([22][22])[22]O$	
		f_{15}	$([22][0])[22]O$	
		f_{16}	$([2][22])[22]O$	
		f_{17}	$([0][22])[22]O$	
		f_{18}	$([4][1])[3]O$	
		f_{19}	$([2][1])[3]O$	
		f_{20}	$([22][22])[3]O$	
		f_{21}	$([4][2])[42]O$	
		f_{22}	$([22][2])[42]O$	
		f_{23}	$([4][22])[42]O$	
		f_{24}	$([22][22])[42]O$	
		f_{25}	$([2][22])[42]O$	
		f_{26}	$([22][22])[44]O$	
		f_{27}	$([4][22])[52]O$	
		f_{28}	$([4][2])[6]O$	
	$[2] \otimes ([11] \otimes \{4\})$	f_{29}, f_{30}	$[0]O$	$uuuu$
		f_{31}, f_{32}	$[22]O$	
		f_{33}	$[3]O$	
		f_{34}	$[42]O$	
		f_{35}	$[44]O$	
	$[2] \otimes [([2] \otimes \{3\})[11]]$	f_{36}	$([22][11])[22]O$	$gggu$
		f_{37}	$([31][11])[22]O$	
		f_{38}	$([31][11])[3]O$	
		f_{39}	$([31][11])[42]O$	
		f_{40}, f_{41}	$([42][11])[42]O$	
		f_{42}	$([42][11])[52]O$	
		f_{43}	$([6][11])[6]O$	
	$[2] \otimes [[2]([11] \otimes \{3\})]$	f_{44}	$([2][21])[22]O$	$guuu$
		f_{45}	$([2][31])[22]O$	
		f_{46}	$([2][21])[3]O$	
		f_{47}	$([2][31])[42]O$	
		f_{48}	$([2][33])[42]O$	

scalar operator can only be formed if $(D + G) \otimes \{2\} \supset P, F$. In fact,

$$(D + G) \otimes \{2\} = D \otimes \{2\} + G \otimes \{2\} + D \times G,$$

and since the singlets of the configurations d^2 and g^2 contain only *even* representations of R_3, it may be concluded that the three-particle operator t_{10} can only be constructed from operators of the type $(\mathbf{v}_i^{(2)}\,\mathbf{v}_j^{(4)})^{(3)} \cdot \mathbf{v}_h^{(3)}$.

However, the three-particle operator $(\mathbf{v}_i^{(2)}\,\mathbf{v}_j^{(4)})^{(3)} \cdot \mathbf{v}_h^{(3)}$ is not symmetric under interchange of the ith and jth coordinates, and hence cannot, by itself, constitute the operator t_{10}. The linear combination

$$[(\mathbf{v}_i^{(2)}\,\mathbf{v}_j^{(4)})^{(3)} + (\mathbf{v}_j^{(4)}\,\mathbf{v}_i^{(2)})^{(3)}]\mathbf{v}_h^{(3)}$$

is symmetric and has the proper transformation properties. But the matrix elements of these two operators differ only by a factor of -1, and hence must yield null matrix elements. For similar reasons the twelve non-Hermitian operators (f_{36} through f_{48}, inclusive) may be rejected as nonphysical operators; this leaves 36 Hermitian N-particle operators still to be considered.

The four operators (t_4, f_8, f_{27}, and f_{28}) all yield null matrix elements when evaluated between the states $[\lambda]$ and $[\lambda']$ of the d^n electron configuration, since the identity representation [0] does not occur in the Kronecker products $[\lambda] \times [\lambda'] \times [\mu]$ for $[\mu] = [6]$ or [52], i.e., $c([\lambda][\lambda'][\mu]) = 0$ for all $[\lambda]$ and $[\lambda']$.

Further simplifications for the remaining 32 operators can be found from an inspection of the classification of the states of d^n electron configurations for $n \leq 5$ given in Table IV, together with the tables of the numbers $c([\lambda][\lambda'][\mu])$ for $[\mu] = [22]$, [3], [42], and [44] given in Tables Va–d.

If the perturbation expansion is restricted to second-order, only the eight effective operators e_0 through t_3 need be considered. The four operators e_0, e_1, e_3, and t_1 all transform as $[0]O$, and thus their matrix elements are completely diagonal in $[\lambda]$. Since there are just three different representations $[\lambda]$ in d^3, only

TABLE IV. Classification of the states of d^n ($0 \leq n \leq 5$).

n	$\{\lambda\}$	$[\mu]$	^{2S+1}L
0	{0}	[0]	^1S
1	{1}	[1]	^2D
2	{11}	[11]	^3PF
	{2}	[2]	^1DG
		[0]	^1S
3	{111}	[11]	^4PF
	{21}	[21]	^2PDFGH
		[1]	^2D
4	{1111}	[1]	^5D
	{211}	[21]	^3PDFGH
		[11]	^3PF
	{22}	[22]	^1SDFGI
		[2]	^1DG
		[0]	^1S
5	{11111}	[0]	^6S
	{2111}	[2]	^4DG
		[11]	^4PF
	{221}	[1]	^2D
		[21]	^2PDFGH
		[22]	^2SDFGI

three independent operators transforming as $[0]O$ are required; hence t_1 can be absorbed in e_0, e_1, and e_3. Thus, to second-order, the energy levels of the d^n configuration may be expressed in terms of the seven operators e_0, e_1, e_2, e_3, e_4, t_2, and t_3, in agreement with the prior observations of Feneuille.[17]

Extension of the perturbation expansion to third order introduces five additional three-particle operators (t_5 through t_9) of type *guu*, seven four-particle operators (f_1 through f_7) of type *gggg*, and eighteen four-particle operators of type *gguu*. Judd[30] has observed that in third order there are coupling restrictions on the ranks of the tensor operators constituting the four-particle operators of type *gguu* and, as a result, it is not possible to parameterize freely these operators. However, four-particle terms arising from higher orders of perturbation are undoubtedly free of this restriction, making a complete parameterization then possible.

If the four-particle operators are neglected, then just two additional operator t_6 and t_8 are required to complete the parameterization of the three-particle effective operators, since the operators t_5, t_7, and t_9 may be absorbed by the operators of type *gg*, *uu*, and *ggg*. Thus the effects of two- and three-particle operators may be completely described in terms of the nine operators e_0, e_1, e_2, e_3, e_4, t_2, t_3, t_6, and t_8. In this case there is one more operator than there are terms in d^3.

Accounting for the effects of two- and three-particle operators together with the four-particle operators of type *gggg* requires a total of 19 operators ($6[0]O$, $8[22]O$, $[3]O$, $3[42]O$, and $[44]O$), whereas there are just 16 terms in d^4. If the effects of all possible four-particle Hermitian operators are considered, it is apparent, from inspection of Tables IV–VI, that a total of 21 operators would be required, compared with the expected number of 45 operators found by ignoring additional symmetry restrictions.

TABLE Va. The numbers $c([\lambda][\lambda'][22])$ for R_5.

$[\lambda]$ \ $[\lambda']$	[0]	[1]	[11]	[2]	[21]	[22]
[0]						1
[1]					1	1
[11]			1		1	1
[2]				1	1	1
[21]		1	1	1	2	1
[22]	1	1	1	1	1	1

TABLE Vb. The numbers $c([\lambda][\lambda'][3])$ for R_5.

$[\lambda]$ \ $[\lambda']$	[0]	[1]	[11]	[2]	[21]	[22]
[0]						
[1]				1		
[11]						
[2]		1			1	
[21]				1	1	1
[22]					1	1

TABLE Vc. The numbers $c([\lambda][\lambda'][42])$ for R_5.

$[\lambda]$ \ $[\lambda']$	[0]	[1]	[11]	[2]	[21]	[22]
[0]						
[1]						
[11]						
[2]						1
[21]					1	1
[22]				1	1	1

TABLE Vd. The numbers $c([\lambda][\lambda'][44])$ for R_5.

$[\lambda]$ \ $[\lambda']$	[0]	[1]	[11]	[2]	[21]	[22]
[0]						
[1]						
[11]						
[2]						
[21]						
[22]						1

TABLE VI. Number of scalar N-particle ($N \leq 4$) operators required to characterize the energy levels of d^N electron configurations.

	$[0]O$	$[22]O$	$[3]O$	$[42]O$	$[44]O$
$N = 0$	1	0	0	0	0
$N = 2$ *gg*	1	1	0	0	0
uu	1	1	0	0	0
$N = 3$ *ggg*	1	1	0	1	0
uug	1	2	1	1	0
$N = 4$ *gggg*	2	2	0	2	1
gguu	3	6	3	5	1
uuuu	2	2	1	1	1
Total number	12	15	5	10	3
Number required	6	9	2	3	1

D. Explicit Construction of the Three-Particle Operators

Shadmi[31] has studied the effect of the three-particle operators t_2 and t_3 on the energy-level structure of the

[30] B. R. Judd, "Effective Operators for Configurations of Equivalent Electrons," (NATO Summer Institute on *Correlations in Atoms and Molecules*, Frascati, 1967).

[31] Y. Shadmi, "The Use of a Complete Set of Parameters of Effective Interactions between 3*d* Electrons in the Spectra of the Iron Group," presented at the Atomic Spectroscopy Symposium, National Bureau of Standards, Gaithersburg, Maryland (1967).

$3d^n$ configurations of the iron group. It would be interesting to repeat his analysis including the additional three-particle operators t_6 and t_8 and thus to complete the parameterization of the three-particle operators.

The operator t_8 transforms as $([2][1])[3]O$ and is proportional to a linear combination

$$\sum_{i\neq j\neq h}(\mathbf{v}_i^{(2)}\cdot[a_{11}(\mathbf{v}_j^{(1)}\mathbf{v}_h^{(1)})^{(2)} + a_{33}(\mathbf{v}_j^{(3)}\mathbf{v}_h^{(3)})^{(2)} + a_{13}(\mathbf{v}_j^{(1)}\mathbf{v}_h^{(3)})^{(2)}]) \quad (38)$$

of the operators $\mathbf{v}^{(1)}$, $\mathbf{v}^{(2)}$, and $\mathbf{v}^{(3)}$. The coefficients $a_{kk'}$ of linear combination are directly proportional to the values assumed by the generalized Wigner coupling coefficient[13]

$$\begin{aligned}&\langle[11]kq[11]k'q'\,|\,[1]2\rangle\\ &=\sum_{[\lambda]K}\langle[11]k''\,|\,[1]2+d\rangle\langle[11]k+d\,|\,[\lambda]K\rangle\\ &\quad\times\langle[\lambda]K+d\,|\,[1]2\rangle\langle([11][1])[\lambda],[1],[1]\,|\,[11],\\ &\qquad([1][1])[11],[1]\rangle(-1)^k\{[K,k']\}^{\frac{1}{2}}\begin{Bmatrix}k & k' & 2\\ 2 & K & 2\end{Bmatrix}.\end{aligned} \quad (39)$$

However, $[\lambda]K$ is restricted to $[11]PF$; hence the dependence of the coupling coefficient on k and k' is proportional to just

$$\sum_K\langle[11]k+d\,|\,[11]K\rangle(-1)^{k+K}[K]\{[k']\}^{\frac{1}{2}}\begin{Bmatrix}k & k' & 2\\ 2 & K & 2\end{Bmatrix}. \quad (40)$$

Using Racah's tables[2] of $\langle[11]k+d\,|\,[11]K\rangle$, it is readily shown that the operator

$$\sum_{i\neq j\neq h}(\mathbf{v}_i^{(2)}\cdot[2^{\frac{1}{2}}(\mathbf{v}_j^{(1)}\mathbf{v}_h^{(1)})^{(2)} - 7^{\frac{1}{2}}(\mathbf{v}_j^{(3)}\mathbf{v}_h^{(3)})^{(2)} + 7^{\frac{1}{2}}(\mathbf{v}_j^{(1)}\mathbf{v}_h^{(3)})^{(2)}]) \quad (41)$$

does indeed transform as $([2][1])[3]O$, and hence may be taken for the operator t_8.

The explicit construction of the four-particle operators has yet to be made, but poses no special difficulties. Smith[25] and Judd[30] have given the corresponding classifications for f^n configurations, and Judd[15] has constructed operators of type *ggg* for these configurations.

E. Classification of Spin-Dependent Two-Particle Operators

The theory of plethysm provides a particularly useful basis for studying the properties of spin-dependent two-particle operators, such as those associated with the spin–spin and spin–other-orbit interactions. Judd[32] has discussed the classification of these operators for d^n electron configurations by conventional methods. The application of plethysm to the classification of these operators for f^n configurations is now taken up.

The spin–spin interaction for a configuration l^n may be taken as[2]

$$H_{ss} = -2\sum_{i\neq j}\sum_k[(k+1)(k+2)(2k+3)]^{\frac{1}{2}} \times M^k(l\,||\,C^{(k)}\,||\,l)(l\,||\,C^{(k+2)}\,||\,l)\{\boldsymbol{\omega}_i^{(1k)}\boldsymbol{\omega}_j^{(1k+2)}\}^{(22)O}, \quad (42)$$

where M^k is the radial integral defined by Marvin[33] and the notation generally follows that of Judd.[13]

The reduced matrix elements of $\mathbf{C}^{(k)}$ vanish unless k is even. Thus the operators $\boldsymbol{\omega}^{(1k)}$ and $\boldsymbol{\omega}^{(1k+2)}$ transform as the representations $\langle 2\rangle$, $[2]$, and (20) of Sp_{14}, R_7, and G_2, respectively. The spin–spin operator must then transform according to some of the representations of Sp_{14} contained in the plethysm

$$\langle 2\rangle\otimes\{2\} = \langle 0\rangle + \langle 1^2\rangle + \langle 2^2\rangle + \langle 4\rangle. \quad (43)$$

The representation $\langle 4\rangle$ may be discarded, since all the states of f^N transform as $\langle 1\cdots 10\cdots 0\rangle$ of Sp_{14}. The remaining representations $\langle\sigma\rangle$ may be rapidly decomposed into those of $SU_2\times R_7$ by performing the plethysms

$$([1][\tfrac{1}{2}]')\otimes\langle\sigma\rangle \quad (44)$$

and remembering that

$$[1]\otimes\{\lambda\} = \{\lambda\}. \quad (45)$$

As an example, the representation $\langle 2^2\rangle$ of Sp_{14} may be decomposed into those of $SU_2\times R_7$ as follows:

$$\begin{aligned}([1][\tfrac{1}{2}]')\otimes\langle 2^2\rangle &= ([1][\tfrac{1}{2}]')\otimes(\{2^2\}-\{1^2\})\\ &= \{2^2\}([\tfrac{1}{2}]'\otimes\{4\}) + (\{31\}+\{21^2\})\\ &\quad\times([\tfrac{1}{2}]'\otimes\{31\}) + (\{4\}+\{2^2\}\\ &\quad+\{1^4\})([\tfrac{1}{2}]'\otimes\{2^2\}) - \{1^2\}([\tfrac{1}{2}]'\\ &\quad\otimes\{2\}) - \{2\}([\tfrac{1}{2}]'\otimes\{1^2\})\\ &= {}^5\{2^2\} + {}^3(\{31\}+\{21^2\}-\{1^2\})\\ &\quad+{}^1(\{4\}+\{2^2\}+\{1^4\}-\{2\})\\ &= {}^5([2^2]+[2]+[0]) + {}^3([31]\\ &\quad+[21^2]+[2]+[11]) + {}^1([4]\\ &\quad+[2^2]+[2]+[1^3]+[0]).\end{aligned} \quad (46)$$

Proceeding in this manner, the branching rules of Table VII are readily established without recourse to the usual chain calculation.

The spin and orbital parts of H_{ss} transform identically to those of the state ${}^5D_0$, and hence only the representation $\langle 2^2\rangle$ of Sp_{14}, appearing in Eq. (43),

[32] B. R. Judd, Physica 33, 174 (1967).

[33] H. H. Maruin, Phys. Rev. 71, 102 (1947).

TABLE VII. Branching rules for the reduction $Sp_{14} \rightarrow SU_2 \times R_7$.

$\langle\sigma\rangle$	$^{2S+1}[\lambda]$
$\langle 0\rangle$	$^1[0]$
$\langle 1^2\rangle$	$^3[1^2]^1[2]$
$\langle 1^4\rangle$	$^5[1^3]^3[21^2]^1[2^2]$
$\langle 2\rangle$	$^3([2][0])^1[1^2]$
$\langle 2^2\rangle$	$^5([2^2][2][0])^3([31][21^2][2][1^2])^1([4][2^2][2][1^2][0])$

need be obtained. Using Eq. (46) and then decomposing the representations $[\lambda]$ of R_7 into the representations $(u_1\, u_2)$ of G_2, it is found that the spin–spin interaction may be expressed in terms of four operators transforming individually as

$$\langle 2^2\rangle[2^2](20),\quad \langle 2^2\rangle[2^2](21),\quad \langle 2^2\rangle[2^2](22),\quad \text{and}\quad \langle 2^2\rangle[2](20), \tag{47}$$

in agreement with the classification obtained by Judd and Wadzinski[18] by the conventional chain calculation.

The spin–other-orbit interaction H_{soo} for the configuration l^n may be written as[32]

$$\begin{aligned} H_{soo} = \sum_{i\neq j}\sum_{k} &[(k+1)(2l+k+2)(2l-k)]^{\frac{1}{2}} \\ &\times [\{\omega_i^{(0k+1)}\omega_j^{(1k)}\}^{(11)0} \\ &\times \{M^{k-1}(l\|C^{(k+1)}\|l)^2 + 2M^k(l\|C^{(k)}\|l)^2\} \\ &+ \{\omega_i^{(0k)}\omega_j^{(1k+1)}\}^{(11)0} \\ &\times \{M^k(l\|C^{(k)}\|l)^2 + 2M^{k-1}(l\|C^{(k+1)}\|l)^2\}]. \end{aligned} \tag{48}$$

The rank k may be even and odd; hence the plethysm

$$\langle 1^2\rangle \otimes \{2\} = \langle 2^2\rangle + \langle 1^4\rangle + \langle 1^2\rangle + \langle 0\rangle \tag{49}$$

must be considered as well as that of Eq. (43). Remembering that H_{soo} transforms identically to the state $^3P_0$ and making use of Table VII, the classification of Table VIII is obtained. However, the operators of type *uu* and *gg*, transforming as the $\langle 2^2\rangle$ representation of Sp_{14}, may be combined; hence the spin–other-orbit interaction may be expressed as the sum of eight operators, two transforming as $\langle 1^2\rangle$, one as $\langle 1^4\rangle$, and five as $\langle 2^4\rangle$.

F. Quasispin Classification of Operators

The classification of N-particle operators may be usefully extended by introducing the quasispin[32,34] quantum number K to label the representations $\langle\sigma\rangle$ of the symplectic group Sp_{4l+2}. This amounts to finding the decomposition of representations of the group R_{8l+4} into those of $SU_2 \times Sp_{4l+2}$. In the case of N-particle operators this involves the decomposition of representations $[11\cdots10\cdots0]$ containing not more than $2N$ symbols 1. If there are r symbols 1, then the decomposition is given by the plethysm

$$(\langle 1\rangle[\tfrac{1}{2}]') \otimes [1^r]. \tag{50}$$

As an example, the decomposition of the antisymmetric representation $[1^4]$ of R_{28} into those of $SU_2 \times Sp_{14}$ may be obtained as

$$\begin{aligned} (\langle 1\rangle[\tfrac{1}{2}]') \otimes [1^4] &= (\langle 1\rangle[\tfrac{1}{2}]') \otimes \{1^4\} \\ &= \{2^2\}([\tfrac{1}{2}]' \otimes \{2^2\}) + \{21^2\}([\tfrac{1}{2}]' \\ &\quad \otimes \{31\}) + \{1^4\}([\tfrac{1}{2}]' \otimes \{4\}) \\ &= {}^1(\langle 2^2\rangle + \langle 1^2\rangle + \langle 0\rangle) + {}^3(\langle 21^2\rangle + \langle 2\rangle \\ &\quad + \langle 1^2\rangle) + {}^5(\langle 1^4\rangle + \langle 1^2\rangle + \langle 0\rangle), \end{aligned}$$

in agreement with Judd's earlier chain calculation.[15]

The representation $\langle 2^2\rangle$ occurs in the decomposition of $[1^4]$ of R_{28} with quasispin $K = 0$ only, and hence the spin–spin interaction H_{ss} is diagonal in the seniority number v and for a given v independent of the number of particles n.

The interpretation of the role of quasispin for spin–other-orbit interactions H_{soo} is rather more complex and has been discussed by Judd[32] for the particular case of d^n configurations. The spin–other-orbit interaction for f^n configurations may be represented in terms of operators that preserve the quasispin quantum number K by constructing a set of ten operators having the transformation properties indicated by the classification scheme of Table IX. The operator transforming as $^3\langle 1^2\rangle[1^2](11)$ transforms identically to the spin–orbit interaction, and hence may be absorbed by the ordinary spin–orbit interaction.

While the group-theoretical description of the spin–other-orbit interaction does introduce striking

TABLE VIII. Classification of spin–other-orbit operators for f^n configurations.

Plethysm	Type	Classification
$\langle 1^2\rangle \otimes \{2\}$	*uu*	$\langle 1^2\rangle[1^2](11)$
		$\langle 1^4\rangle[21^2](11)$
		(30)
		$\langle 2^2\rangle[1^2](11)$
		$[21^2](11)$
		(30)
		[31](30)
		(31)
$\langle 2\rangle \otimes \{2\}$	*gg*	$\langle 1^2\rangle[1^2](11)$
		$\langle 2^2\rangle[1^2](11)$
		$[21^2](11)$
		(30)
		[31](30)
		(31)

[34] R. D. Lawson and M. H. Macfarlane, Nucl. Phys. **66**, 80 (1965).

TABLE IX. Quasispin representation of spin–other-orbit operators for f^n configurations.

$\langle\sigma\rangle$	K	$[\lambda]$	$(u_1 u_2)$
$\langle 1^4\rangle$	2	$[21^2]$	(11)
			(30)
$\langle 1^2\rangle$		$[1^2]$	(11)
$\langle 1^2\rangle$	1	$[1^2]$	(11)
$\langle 2^2\rangle$	0	$[1^2]$	(11)
		$[21^2]$	(11)
			(30)
		[31]	(30)
			(31)
$\langle 1^2\rangle$	0	$[1^2]$	(11)

simplifications such as selection rules and the seniority dependence of the matrix elements, the advantage gained is largely offset by the need to consider the matrix elements of all ten of the operators classified in Table IX, each of which is a particular linear combination of the two-particle spin-dependent operators appearing in Eq. (48).

G. Further Remarks on N-Particle Operators

The classification of N-particle operators for j^n configurations could be of particular interest in nuclear problems. In this case the symplectic group Sp_{2j+1} plays an important role, and plethysms of the type

$$[j] \otimes (\langle 2\rangle + \langle 11\rangle) \otimes \{N\}$$

have to be considered.

The classification of scalar N-particle operators for d^n configurations of neutrons or of protons is identical to that found for electrons, if the eigenfunctions of the states are in the LS basis. The isotopic spin would have to be considered for configurations involving both neutrons and protons, and the classification of scalar operators presented here would need some extension. If tensor forces are admitted, it is necessary to extend the classification to include nonscalar operators, but no new principles arise.

IV. SELECTION RULES

The explicit calculation of the matrix elements of operators having well-defined transformation properties may be greatly simplified by the judicious application of selection rules. If the bra, operator, and ket of a matrix element transforms individually according to the representations Γ_a, Γ_b, and Γ_c of some group G, then the number of times $c(\Gamma_a\Gamma_b\Gamma_c)$ the identity representation appears in the Kronecker product $\Gamma_a \times \Gamma_b \times \Gamma_c$ determines both whether or not the matrix elements vanish and the number of independent sets of matrix elements that correspond to a given $(\Gamma_a, \Gamma_b, \Gamma_c)$.

The Kronecker products may be readily evaluated for continuous groups using Littlewood's theorem.[23] If $c(\Gamma_a\Gamma_b\Gamma_c) = 0$, the vanishing of the matrix elements is assured. An inspection of Tables IV and Vb shows that the matrix elements of the three-particle scalar operator t_8 vanish in all d^n, except for the states labeled by the [21] and [22] representations of R_5. For the half-filled d^5 shell the diagonal matrix elements associated with these representations vanish, leaving just the matrix elements connecting the ^2DFG states of [21] to the ^2DFG states of [22]. Thus in d^5 there are just three nonvanishing matrix elements of t_8.

If $c(\Gamma_a\Gamma_b\Gamma_c) = \alpha$, then there are not more than α linearly independent sets of matrix elements associated with the labels Γ_a, Γ_b, and Γ_c. Since $c([21][21][3]) = 1$, it follows that the matrix elements of t_8 for d^4 diagonal in [21] are simply proportional to the corresponding matrix elements in d^3. Thus the complete evaluation of the matrix elements of t_8 for d^n configurations involves, at most, the explicit calculation of thirteen matrix elements.

Further selection rules may be obtained by resolving the Kronecker squares of irreducible representations into their symmetric and antisymmetric parts, as noted by Judd and Wadzinski[18] and elaborated upon by Smith and Wybourne,[24] using the method of plethysm.

V. CONCLUSION

Littlewood's algebra of plethysm plays a decisive role in the development of the theory of complex spectra. In particular, the ambiguities of the usual chain-calculational methods are entirely removed. The algebra of plethysm rests heavily upon the theory of induced matrices, and substantial progress in the development of the theory of complex spectra should be possible by focusing attention upon the inducing matrices associated with the construction of plethysms.

ACKNOWLEDGMENT

We are grateful to Professor B. R. Judd for communicating the contents of his Frascati Lectures prior to their publication and for several helpful comments.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 7 JULY 1968

Type-Null Vacuum Solutions in General Relativity

M. MISRA*
Mathematics Department, Gujarat University, Ahmedabad-9, India

(Received 29 March 1967)

It is shown that the necessary and sufficient conditions for the vacuum solutions of a certain metric to be of type null according to the Pirani–Petrov classification are that a certain other metric be conformally flat. The general solution is obtained.

In an earlier paper[1] the metric

$$ds^2 = e^{2w}d^*s^2 - \epsilon e^{2w}(dx^3)^2, \tag{1}$$

$$d^*s^2 = {}^*g_{ij}\,dx^i\,dx^j, \quad (i, j, \cdots, = 1, 2, 4), \tag{2}$$

where w and ${}^*g_{ij}$ are function of x^i only, was considered, and it was found that the necessary and sufficient conditions in order that the vacuum solutions corresponding to the metric (1) be of type null according to the Pirani–Petrov classification[2] are given by

$${}^*g^{ij}{}^*w_{;i}{}^*w_{;j} = 0, \tag{3}$$

$${}^*w_{;i}{}^*w_{;hj} - {}^*w_{;h}{}^*w_{;ij} = 0, \tag{4}$$

where an asterisk denotes quantities defined with respect to the metric (2) and a semicolon followed by a lower index implies covariant differentiation with respect to (2).

The object of this note is to show that Eqs. (4) are precisely the conditions for the metric (2) to be conformally flat. Hence the necessary and sufficient conditions for the vacuum solutions corresponding to the metric (1) to be of type null according to the Pirani–Petrov classification are that the metric (2) be conformally flat and that ${}^*w_{;i}$ be a null vector.

Proof: The empty space-time field equations for the metric (1) are[1]

$${}^*R_{ij} + 2{}^*w_{;i}{}^*w_{;j} = 0. \tag{5}$$

Also the necessary and sufficient conditions for a V_3 [and the metric (2) is a V_3] to be conformally flat are[3]

$${}^*R_{ij;k} - {}^*R_{ik;j} + \tfrac{1}{4}({}^*g_{ik}{}^*R_{;j} - {}^*g_{ij}{}^*R_{;k}) = 0. \tag{6}$$

Substituting from Eqs. (3) and (5) in (6) and rearranging the indices, we get precisely Eq. (4).

Hence the result follows.

Thus in the case when the vacuum solutions are of type null we can choose coordinates such that

$${}^*g_{ij} = e^{2\sigma}\eta_{ij}, \tag{7}$$

where σ is a function of x^i and

$$\eta_{ij} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & +1 \end{bmatrix}. \tag{8}$$

And from Eqs. (3), (5), and (7) we get the equations

$$\eta^{ij}w_{,i}w_{,j} = 0, \tag{9}$$

$$\sigma_{,ij} - \sigma_{,i}\sigma_{,j} + \eta_{ij}\eta^{km}(\sigma_{,km} + \sigma_{,k}\sigma_{,m}) + 2w_{,i}w_{,j} = 0, \tag{10}$$

where a comma denotes partial differentiation. From Eqs. (9) and (10) we get

$$\eta^{ij}(\sigma_{,ij} + \tfrac{1}{2}\sigma_{,i}\sigma_{,j}) = 0. \tag{11}$$

It follows that σ is a function of

$$u = \alpha_i x^i, \tag{12}$$

where α_i are arbitrary constants subject to the condition

$$\eta^{ij}\alpha_i\alpha_j = 0. \tag{13}$$

From Eqs. (10), (12), and (13) it follows that

$$(\sigma'' - \sigma'^2)\alpha_i\alpha_j + 2w_{,i}w_{,j} = 0. \tag{14}$$

Here a prime denotes ordinary differentiation with respect to u.

It is clear from Eqs. (9) and (14) that w also is a function of u. Hence we have

$$(\sigma'' - \sigma'^2 + 2w'^2)\alpha_i\alpha_j = 0, \tag{15}$$

and, since α_i is not zero, we must have

$$\sigma'' - \sigma'^2 + 2w'^2 = 0. \tag{16}$$

Equation (9) is satisfied identically and we are left with only one equation, viz., Eq. (16), for determining σ and w.

If we put

$$\sigma = -\log\varphi, \tag{17}$$

Eq. (16) becomes

$$2w'^2 - \frac{\varphi''}{\varphi} = 0,$$

or

$$w = \int\left[\frac{1}{2\varphi}\frac{d^2\varphi}{du^2}\right]^{\frac{1}{2}}du + \text{const.} \tag{18}$$

This is the well-known solution for plane gravitational waves. It was obtained earlier[4] by the present author from different considerations.

ACKNOWLEDGMENT

The author wishes to thank Professor P. C. Vaidya for helpful discussions.

* Present address: Dublin Institute for Advanced Studies, Dublin 2, Ireland.

[1] M. Misra, J. Math. Phys. 7, 155 (1966).

[2] F. A. E. Pirani, *Gravitation—An Introduction to Current Research*, L. Witten, Ed. (John Wiley & Sons, Inc., New York, 1962), Chap. 6.

[3] L. P. Eisenhart, *Riemannian Geometry* (Princeton University Press, Princeton, N.J., 1949), p. 92.

[4] M. Misra, Proc. Natl. Inst. Sci. India **29A**, 104 (1962).

Radiating Spheres in the Scalar–Tensor Theory of General Relativity*

WILLIAM J. KAUFMANN III†
Goethe Link Observatory, Indiana University, Bloomington, Indiana

(Received 15 November 1967)

Using Bondi's "radiation coordinates," the Brans–Dicke field equations describing the interior of a sphere consisting of a perfect fluid and a radiation field, part of which is purely isotropic and the remainder of which is purely radial, are stated. A method for seeking solutions to these equations in the radiation-filled space outside the sphere is given.

I. INTRODUCTION

Several years ago Brans and Dicke[1–3] proposed a new theory of gravitation, which differs from Einstein's theory by the presence of a scalar $\Phi(x^\alpha)$ ($\alpha = 0, 1, 2, 3$, where x^0 is the timelike coordinate) in the field equations. One of the important consequences of this new theory is that the gravitational "constant" k is inversely proportional to the strength of the scalar field $\Phi(x^\alpha)$, and therefore depends on the gravitational configuration of matter. In addition, Einstein's theory is a special case of the Brans–Dicke theory; if we let $\Phi(x^\alpha) = \text{const}$, and let the coupling constant $\omega \to \infty$, the field equations of the latter theory reduce to those of the former. Since the field equations of Brans and Dicke depend fundamentally on both tensors (such as the metric tensor $g_{\mu\nu}$ and the energy–momentum tensor $T_{\mu\nu}$) and scalars (such as the strength of the scalar field $\Phi(x^\alpha)$ and the contracted energy–momentum tensor $T \equiv T_\nu{}^\nu$) their theory is called the scalar–tensor theory.

At present there is no evidence to preclude the validity of the scalar–tensor theory. While this theory does not predict an anomalous gravitational redshift, it does give values for the gravitational deflection of light rays and the perihelion advance of planetary orbits different from those given by Einstein's theory.[4] But in view of the relatively large discrepancies in the measurements of the deflection of starlight near the sun's limb during a total eclipse[5] and the recent measurements of the oblateness of the sun,[6] we must conclude that the scalar–tensor theory is not in conflict with observations, provided that $\omega \geq 6$.

Only very recently, however, has this theory been applied to more interesting problems in astrophysics in order to appreciate fully the implications of the addition of a long-range scalar interaction. By comparing the predictions of this theory with those of Einstein's theory, we may hope to obtain important differences which might be used to decide between the two theories. For example, while Salmona[7] has shown that certain gross features of a cold neutron star remain unchanged by the presence or strength of the scalar field, Morganstern and Chiu[8] have shown that if a neutron star is observed to exhibit symmetric radial pulsations, then the existence of the scalar field may be ruled out.

The purpose of this paper is to make an attack on the time-dependent problem. We shall consider a massive sphere consisting of a perfect fluid (i.e., a fluid which is incapable of exerting tangential stresses) to which we shall add both an isotropic radiation field inside the sphere, and a radial flux of electromagnetic energy inside and outside the sphere. After giving explicit expression to the field equations inside the sphere, a method for obtaining solutions in the radiation-filled space outside the sphere is presented.

II. STRUCTURE EQUATIONS

The scalar–tensor field equations are

$$G_{\mu\nu} - \tfrac{1}{2}g_{\mu\nu}G = -\frac{8\pi}{c^4\Phi}T_{\mu\nu} - \frac{\omega}{\Phi^2}(\Phi_{,\mu}\Phi_{,\nu} - \tfrac{1}{2}g_{\mu\nu}g^{\alpha\beta}\Phi_{,\alpha}\Phi_{,\beta}) - \Phi^{-1}(\Phi_{,\mu;\nu} - g_{\mu\nu}\Box\Phi),$$

$$\Box\Phi = [-g]^{-\frac{1}{2}}[[-g]^{\frac{1}{2}}g^{\alpha\eta}\Phi_{,\eta}]_{,\alpha} = \frac{8\pi}{(3+2\omega)c^4}T, \quad (1)$$

where, in addition to the symbols previously defined, $G_{\mu\nu}$ is the contracted Riemann curvature tensor, $G \equiv G_\nu{}^\nu$, $A_{,\mu} \equiv \partial A/\partial x^\mu$, $A_{;\mu}$ means covariant differentiation with respect to the coordinate x^μ, and the coupling constant ω is dimensionless.

---

* Publications of the Goethe Link Observatory, Indiana University No. 83.

† Present address: Department of Astronomy, University of California, Los Angeles, Calif.

[1] C. Brans and R. H. Dicke. Phys. Rev. **124**, 925 (1961).

[2] C. Brans, Phys. Rev. **125**, 2194 (1962).

[3] R. H. Dicke, Phys. Rev. **125**, 2163 (1962).

[4] R. H. Dicke, *The Theoretical Significance of Experimental Relativity* (Gordon and Breach Science Publishers, New York, 1964).

[5] R. H. Dicke, Physics Today **20**, 55 (1967).

[6] R. H. Dicke and H. Mark Goldenberg, Phys. Rev. Letters **18**, 313 (1967).

[7] A. Salmona, Phys. Rev. **154**, 1218 (1967).

[8] R. E. Morganstern and Hong-Yee Chiu, Phys. Rev. **157**, 1228 (1967).

Since we wish to contemplate a system in which radiation plays a major role, it is advantageous to make use of the so-called "radiation coordinates" first introduced by Bondi[9] and more fully discussed by Bondi, van der Burg, and Metzner.[10] For the spherically symmetric case the metric in these coordinates takes the form

$$ds^2 = e^{2\beta}[(V/r)c^2\,du^2 + 2c\,du\,dr] - r^2\,d\Omega^2, \quad (2)$$

where $d\Omega^2 = d\theta^2 + \sin^2\theta\,d\phi^2$ and where V and β are functions of both u and r, and $(u, r, \theta, \phi) = (0, 1, 2, 3)$. Since this metric is both useful and not very well known, the nonvanishing Christoffel 3-index symbols are given in Appendix A.

In order to construct the energy–momentum tensor, we consider a local Minkowski coordinate system $(\bar{t}, \bar{x}, \bar{y}, \bar{z})$ which is at rest with respect to the local matter inside the sphere, but which is moving with velocity w in the radial $(\bar{x})$ direction with respect to the rest frame of the entire sphere. In this Minkowski system inside the sphere we now make use of the energy–momentum tensor, first given by Bondi,[11] whose covariant components $\bar{T}_{\mu\nu}$ may be written as the matrix

$$\begin{pmatrix} (c^2\bar{\rho} + 3\bar{\sigma} + \bar{\epsilon})c^2 & -\bar{\epsilon}c & 0 & 0 \\ -\bar{\epsilon}c & \bar{p} + \bar{\sigma} + \bar{\epsilon} & 0 & 0 \\ 0 & 0 & \bar{p} + \bar{\sigma} & 0 \\ 0 & 0 & 0 & \bar{p} + \bar{\sigma} \end{pmatrix}, \quad (3)$$

where $\bar{p}$ and $\bar{\sigma}$ are the macroscopic pressure and density of a perfect fluid, respectively, $3\bar{\sigma}$ is the energy density of an isotropic radiation field, and $\bar{\epsilon}$ is the energy density of radiation travelling in the radial direction.

It is clear that one may present a number of objections against using this energy–momentum tensor. Atkinson has pointed out to me that, in addition to the considerations given by him in an earlier paper,[12] the artificial division of the radiation field into a purely isotropic part and a purely radial part entails the unreasonable assumption that *the radial component is coherent*, which is not possible over the surface of a sphere. But in order to allow for a noncoherent flux, the precise form of the energy–momentum tensor would depend not only on the electromagnetic properties of the radiation, but also on the location of the local Minkowski coordinates with respect to Atkinson's "photon circle."[13] For the sake of simplicity, however, we shall continue to use Bondi's energy–momentum tensor and leave the aforementioned modifications as the subject of a future publication.

We now proceed to make a Lorentz transformation from our previous Minkowski coordinate system $(\bar{t}, \bar{x}, \bar{y}, \bar{z})$ to a second Minkowski coordinate system (t, x, y, z) which is at rest with respect to the rest frame of the sphere and is related to Bondi's radiation coordinates by

$$\begin{aligned} dt &= e^{\beta}[(V/r)^{\frac{1}{2}}\,du + (Vc^2/r)^{-\frac{1}{2}}\,dr], \\ dx &= e^{\beta}(V/r)^{-\frac{1}{2}}\,dr, \\ dy &= r\,d\theta, \quad dz = r\sin\theta\,d\phi. \end{aligned} \quad (4)$$

By the method described in detail by Atkinson,[12] we now transform our energy–momentum tensor from the first Minkowski coordinates to the radiation coordinates, and find in the latter coordinate system the nonvanishing components of the energy–momentum tensor are

$$\begin{aligned} T_{00} &= e^{2\beta}(V/r)[(c^4\rho + w^2p)/(1 - w^2/c^2) + c^2\epsilon], \\ T_{01} &= T_{10} = e^{2\beta}(c^3\rho - wp)/(1 + w/c), \\ T_{11} &= e^{2\beta}(r/V)(c^2\rho + p)(1 - w/c)/(1 + w/c), \\ T^2_2 &= T^3_3 = -p, \end{aligned} \quad (5)$$

where

$$\begin{aligned} c^2\rho &= c^2\bar{\rho} + 3\bar{\sigma}, \\ p &= \bar{\rho} + \bar{\sigma}, \\ \epsilon &= \bar{\epsilon}(1 + w/c)/(1 - w/c). \end{aligned} \quad (6)$$

Noting that $T \equiv T^{\nu}_{\nu} = -3p + c^2\rho$, we may now write out the scalar–tensor field equations as

$$\begin{aligned} &G_{00} - \tfrac{1}{2}g_{00}G \\ &= -\frac{c}{r^2}(V_0 - 2\beta_0 V) - \frac{Vc^2}{r^3}(e^{2\beta} - V_1 + 2\beta_1 V) \\ &= -\frac{8\pi}{c^4\Phi}(e^{2\beta}V/r)\left(c^2\epsilon + \frac{c^4\rho + w^2p}{1 - w^2/c^2} - \frac{c^4\rho - 3pc^2}{3 + 2\omega}\right) \\ &\quad - \frac{\Phi_{00}}{\Phi} + \frac{\omega}{\Phi^2}\left[\left(\frac{Vc}{r}\right)\Phi_1\Phi_0 - \frac{V^2c^2}{2r^2}(\Phi_1)^2 - (\Phi_0)^2\right] \\ &\quad + \frac{\Phi_0}{\Phi}\left[2\beta_0 + \frac{Vc}{2r^2} - \frac{V_1c}{2r} - \frac{\beta_1 Vc}{r}\right] \\ &\quad + \frac{\Phi_1}{\Phi}\left[\frac{cV_0}{2r} - \frac{c\beta_0 V}{r} - \frac{V^2c^2}{2r^3} + \frac{VV_1c^2}{2r^2} + \frac{\beta_1 V^2c^2}{r^2}\right], \end{aligned} \quad (7)$$

[9] H. Bondi, Nature **186**, 535 (1960).
[10] H. Bondi, M. G. J. van der Burg, and A. W. K. Metzner, Proc. Roy. Soc. (London) **A269**, 21 (1962).
[11] H. Bondi, Proc. Roy. Soc. (London) **A281**, 39 (1964).
[12] R. d'E. Atkinson, Proc. Natl. Acad. Sci. U.S. **51**, 723 (1965).
[13] R. d'E. Atkinson, Astron. J. **70**, 517 (1965).

$$G_{11} - \tfrac{1}{2}g_{11}G = -\frac{4\beta_1}{r}$$
$$= -\frac{8\pi}{c^4\Phi}(e^{2\beta}r/V)\left(\frac{1 - w/c}{1 + w/c}\right)(\rho c^2 + p)$$
$$- \omega(\Phi_1/\Phi)^2 - \Phi_{11}/\Phi + 2\beta_1\Phi_1/\Phi, \quad (8)$$

$$G_{01} - \tfrac{1}{2}g_{01}G = -\frac{c}{r^2}(e^{2\beta} - V_1 + 2\beta_1 V)$$
$$= -\frac{8\pi}{c^4\Phi}e^{2\beta}\left(\frac{c^3\rho - wp}{1 + w/c} - \frac{c^3\rho - 3cp}{3 + 2\omega}\right)$$
$$- \frac{Vc\omega}{2r}(\Phi_1/\Phi)^2 - \Phi_{01}/\Phi$$
$$+ \frac{\Phi_1}{\Phi}\left[\frac{\beta_1 Vc}{r} + \frac{V_1 c}{2r} - \frac{Vc}{2r^2}\right], \quad (9)$$

$$G_2^2 - \tfrac{1}{2}g_2^2G = G_3^3 - \tfrac{1}{2}g_3^3G$$
$$= -e^{-2\beta}\left\{\frac{2\beta_{01}}{c} - \frac{1}{2r^2}[V_{11}r - 2\beta_1 V + 2r(\beta_{11}V + \beta_1 V_1)]\right\}$$
$$= -\frac{8\pi}{c^4\Phi}\left(-p + \frac{3p - \rho c^2}{3 + 2\omega}\right)$$
$$+ \frac{\omega e^{-2\beta}}{\Phi^2}\left[\frac{\Phi_1\Phi_0}{c} - \frac{V}{2r}(\Phi_1)^2\right]$$
$$+ \frac{e^{-2\beta}}{\Phi}\left[-\frac{\Phi_0}{rc} + \frac{V\Phi_1}{r^2}\right], \quad (10)$$

$$\Box\Phi = \frac{e^{-2\beta}}{cr^2}[2r^2\Phi_{01} + 2r\Phi_0 - cV\Phi_1$$
$$- crV_1\Phi_1 - crV\Phi_{11}]$$
$$= \frac{8\pi(c^2\rho - 3p)}{c^4(3 + 2\omega)}, \quad (11)$$

where the subscripts 0 and 1 on the variables V, Φ, β, and Ψ (to be defined later) denote partial differentiation with respect to u and r, respectively. (Of course, the subscripts 0 and 1 on G, g, T, and Γ indicate components.)

III. METHOD FOR OBTAINING SOLUTIONS OUTSIDE THE SPHERE

It is clear that the equations describing the radiation-filled space outside the sphere are readily obtained from Eqs. (7) through (11) by setting everywhere $p = 0$ and $\rho = 0$. We immediately notice that we seem to have one more equation than unknowns. This apparent difficulty is easily explained by the fact that the vanishing divergence equation is contained in Eq. (1), and therefore one of the five equations obtained from Eqs. (7) through (11) is redundant.

[In order to avoid lengthening this paper unnecessarily, the five equations obtained from Eqs. (7), (8), (9), (10), and (11) by setting $p = 0$ and $\rho = 0$ will be designated Eqs. (7′), (8′), (9′), (10′), and (11′), respectively, and will not be written out explicitly.]

We now notice that if we define a new variable $\Psi(u, r)$ related to $\Phi(u, r)$ by

$$\Psi_{11} = \Phi, \quad (12)$$

then Eq. (11′) may be immediately integrated to obtain

$$V = 2[\Psi_0 - r\Psi_{01} + r^2\Psi_{011} + \Xi][cr\Psi_{111}]^{-1}, \quad (13)$$

where $\Xi = \Xi(u)$ is an arbitrary function of u only.

Also, in view of the definition of $\Psi(u, r)$, Eq. (8′) may be written as

$$2\beta_1 = [r\Psi_{1111}\Psi_{11} + \omega r\Psi_{111}^2][2\Psi_{11}^2 + r\Psi_{111}\Psi_{11}]^{-1}. \quad (14)$$

And finally, Eqs. (9′), (8′), (11′), and (12) may be combined to give

$$2\beta = \ln\left[V_1 - \frac{r}{c}\left(\frac{\Psi_{011}}{\Psi_{11}}\right) + V\left(\frac{\Psi_{111}}{\Psi_{11}}\right)\right]. \quad (15)$$

If we now substitute for V and V_1 from Eq. (13) into Eq. (15), differentiate the result to obtain a second expression for $2\beta_1$, and set this result equal to the right-hand side of Eq. (14), we obtain one equation in one unknown whose solution gives $\Psi(u, r)$. Substituting this solution back into Eqs. (13) and (15), we obtain $V(u, r)$ and $\beta(u, r)$, which may in turn be substituted into Eq. (7′) to give an expression for the energy density of the radiant flux from the sphere $\epsilon(u, r)$.

IV. CONDITIONS FOR PHYSICALLY MEANINGFUL SOLUTIONS

As would be expected, the fifth order, nonlinear partial differential equation for $\Psi(u, r)$ obtained by the method outlined in the previous paragraph is not easy to solve. In addition, the requirement that our solutions be physically meaningful adds another degree of difficulty to the problem. For example, the solution

$$\Psi = \frac{U}{(n + 1)(n + 2)}r^{n+2} + U'r + U'', \quad (16)$$

where U, U', and U'' are arbitrary functions of u only, and where

$$n = \frac{1 \pm (3 + 2\omega)^{\frac{1}{2}}}{\omega + 1},$$

may be dismissed as unphysical after a moment's consideration: e.g., it predicts that the ratio of the values of the gravitational "constant" at the earth's orbit to that at Neptune's orbit for the case $\omega = 11$ is either approximately $(30)^{+\frac{1}{2}}$ or $(30)^{-\frac{1}{2}}$. There are,

however, two general requirements we can place on the solution in order for it to be physically meaningful.

1. *The solution must reduce to Vaidya's solution in the Einsteinian limit.*

As was previously mentioned, in the Einsteinian limit where $\Phi = \text{const}$ and $\omega \to \infty$, the scalar–tensor field equations reduce to the field equations of ordinary general relativity. The solution to Einstein's equations for the physical situation under consideration was first given by Vaidya[14] and later elaborated upon by Bondi[11]; it is

$$\beta = 0, \quad V = r - 2m, \quad \epsilon = m_0 c^3/4\pi k r(r - 2m), \tag{17}$$

where $m = m(u)$ is an arbitrary nonincreasing function of u only. Since the properties of this solution have been considered in detail by Lindquist *et al.*,[15] we shall forego any such elaborations. Suffice it to say acceptable solutions must reduce to Vaidya's solution in the Einsteinian limit.

2. *In the limit that the flux (ϵ) vanishes, we must in principle be able to transform the resulting solution to one of the four empty space solutions of Brans.*

By way of analogy with ordinary general relativity, we notice from Eq. (17) that $\epsilon = 0$ in the limit as $m \to \text{const}$. As mentioned by Lindquist *et al.*,[15] Finkelstein has shown that, under the condition that $m = \text{const}$, Vaidya's metric

$$ds^2 = (1 - 2m/r)c^2\,du^2 + 2c\,du\,dr - r^2\,d\Omega^2 \tag{18}$$

reduces to the Schwarzschild metric

$$ds^2 = -(1 - 2m/r)^{-1}\,dr^2 - r^2\,d\Omega^2 + (1 - 2m/r)c^2\,dt^2, \tag{19}$$

if we take

$$cu = ct - r - 2m \ln [(r - 2m)/(R - 2M)], \tag{20}$$

where R and M are constants having the dimensions of length.

Now, Brans[2,16] has given four solutions to the scalar–tensor field equations in empty space outside a spherically symmetric distribution of matter for the case of the isotropic metric

$$ds^2 = e^{2\alpha}c^2\,dt^2 - e^{2\gamma}(dr_1^2 + r_1^2\,d\Omega^2). \tag{21}$$

It is usually argued (see, for example, Salmona[7]) that Brans's "Solution 1" is the physically meaningful solution of the four, namely,

$$\begin{aligned} e^{\alpha} &= D\left[\frac{r_1 - B}{r_1 + B}\right]^{1/\lambda}, \\ e^{\gamma} &= E\left[1 + \frac{B}{r_1}\right]^2\left[\frac{r_1 - B}{r_1 + B}\right]^{(\lambda - A - 1)/\lambda}, \\ \Phi &= F\left[\frac{r_1 - B}{r_1 + B}\right]^{A/\lambda}, \end{aligned} \tag{22}$$

where A, B, D, E, F are constants, and where

$$\lambda^2 = (A + 1)^2 - A(1 - \tfrac{1}{2}\omega A) > 0.$$

Having obtained an expression for ϵ from Eq. (7′) using the solution $\Psi(u, r)$ which we want to test, we place conditions on Ψ so that $\epsilon = 0$. Under these conditions V becomes V^* and β becomes β^*, and we may transform from our radiating metric

$$ds^2 = e^{2\beta^*}\{(V^*/r)c^2\,du^2 + 2c\,du\,dr\} - r^2\,d\Omega^2 \tag{23}$$

to the isotropic metric in Eqs. (21) and (22) by means of the equations

$$u = u(r_1, t) \quad \text{and} \quad r = r_1 e^{\gamma}. \tag{24}$$

Clearly, we have an acceptable solution if the transformation is successful, namely, if

$$\frac{V^*c^2}{r_1}e^{2\beta^* - \gamma}\left(\frac{\partial u}{\partial r_1}\right)^2 + 2ce^{2\beta^* + \gamma}\left(\frac{\partial u}{\partial r_1}\right)\left(1 + r_1\frac{d\gamma}{dr_1}\right) = -e^{2\gamma}, \tag{25}$$

$$\frac{V^*c^2}{r_1}\left(\frac{\partial u}{\partial r_1}\right) + e^{2\gamma}c\left(1 + r_1\frac{d\gamma}{dr_1}\right) = 0, \tag{26}$$

$$e^{2\beta^* - \gamma}\left(\frac{V^*}{r_1}\right)\left(\frac{\partial u}{\partial t}\right)^2 = e^{2\alpha}. \tag{27}$$

Unfortunately, all attempts to find analytical solutions $\Psi(u, r)$ which meet the above two conditions have so far been unsuccessful. Nevertheless, the obtaining of solutions to the problem of radiating spheres in the scalar–tensor theory may be expected to shed much light on the equation of deciding between the Brans–Dicke theory and the Einstein theory. And clearly this is one of the most important and intriguing questions facing the general relativist today.

ACKNOWLEDGMENTS

It is a pleasure to thank Professor Robert d'E. Atkinson for many helpful discussions and suggestions.

[14] P. C. Vaidya, Nature **171**, 260 (1953).
[15] R. W. Lindquist, R. A. Schwartz, and C. W. Misner, Phys. Rev. **137**, B1364 (1965).
[16] C. Brans, Ph.D. thesis, Princeton University (1961).

The author was supported by a NASA predoctoral fellowship during the course of this study.

APPENDIX A

The nonvanishing Christoffel 3-index symbols for the metric stated in Eq. (2) are

$$\Gamma^0_{00} = 2\beta_0 + \frac{Vc}{2r^2} - \frac{V_1 c}{2r} - \frac{\beta_1 Vc}{r},$$

$$\Gamma^0_{33} = \sin^2\theta\Gamma^0_{22} = (re^{-2\beta}/c)\sin^2\theta,$$

$$\Gamma^1_{00} = \frac{cV_0}{2r} - \frac{c\beta_0 V}{r} - \frac{V^2c^2}{2r^3} + \frac{VV_1c^2}{2r^2} + \frac{\beta_1 V^2c^2}{r^2},$$

$$\Gamma^1_{10} = \Gamma^1_{01} = \frac{\beta_1 Vc}{r} + \frac{V_1 c}{2r} - \frac{Vc}{2r^2},$$

$$\Gamma^1_{11} = 2\beta_1,$$

$$\Gamma^1_{33} = \sin^2\theta\Gamma^1_{22} = -Ve^{-2\beta}\sin^2\theta,$$

$$\Gamma^2_{21} = \Gamma^2_{12} = \Gamma^3_{31} = \Gamma^3_{13} = 1/r,$$

$$\Gamma^2_{33} = -\sin\theta\cos\theta,$$

$$\Gamma^3_{23} = \Gamma^3_{32} = \cot\theta,$$

where the subscripts 0 and 1 on the variables $V(u, r)$ and $\beta(u, r)$ indicate partial differentiation with respect to u and r, respectively.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 7 JULY 1968

Bases for the Representations of U_4 in the Chain $U_4 \supset U_2 + U_2$

V. SYAMALA DEVI AND T. VENKATARAYUDU
Andhra University, Waltair, India

(Received 3 March 1967)

The highest-weight polynomials of irreducible representations of U_3 occurring in the reduction of the direct product of two irreducible representations are constructed from Young diagrams. The different irreducible representations in the product are labeled by a parameter which distinguishes their multiplicities also. The method of this paper can be easily extended to any U_n.

I. INTRODUCTION

Moshinsky[1] discussed a procedure for deriving the Wigner coefficients for U_n and gave the details for U_3. He showed that the highest-weight polynomial P of the representation (h_1, h_2, h_3) occurring in the product $(h'_1, h'_2, 0) \times (h'_3, h'_4, 0)$ is a solution of the following equations:

$$C_{11}P = h'_1P, \quad C_{22}P = h'_2P, \quad C_{12}P = 0, \tag{1}$$

$$C_{33}P = h'_3P, \quad C_{44}P = h'_4P, \quad C_{34}P = 0, \tag{2}$$

$$C^{\mu}_{\mu}P = h_\mu P, \quad C^{\mu'}_{\mu}P = 0, \quad \mu < \mu' = 1, 2, 3, \tag{3}$$

where

$$C^{\mu'}_{\mu} = \sum_{s=1}^{4} a_{\mu s}a^{\mu'}_{s}, \quad C_{ss'} = \sum_{\mu=1}^{3} a_{\mu s}a^{\mu}_{s'},$$

and $a_{\mu s}$ are boson creation operators and a^{μ}_{s} are the corresponding annihilation operators. The solutions of Eqs. (1), (2), and (3) are polynomials in a basis for the irreducible representation $(h_1, h_2, h_3, 0)$ of U_4 in which the two subgroups U_2, whose generators are C_{11}, C_{12}, C_{21}, C_{22} and C_{33}, C_{34}, C_{43}, C_{44}, respectively, are explicitly reduced so that the polynomials are of highest weight in these subgroups. In this procedure there is the difficulty of the appearance of negative powers of functions of $a_{\mu s}$ in some forms of the solution (division problem). Suitable combinations have then to be worked out to get the required polynomial solutions. This involves considerable difficulty—even in the case of U_4. The object of this paper is to obtain these highest-weight polynomials in a more direct way, based on Littlewood's rules for building up the product representations. After getting the highest-weight polynomial, the entire basis for that representation can be obtained by applying the lowering operators[2] on the highest-weight polynomial. The generalization to the case of U_n is straightforward and gets rid of the division problem.

II. PRODUCT REPRESENTATIONS

Let $A = (h'_1, h'_2, 0)$ and $B = (h'_3, h'_4, 0)$ be two irreducible representations of U_3. (We can restrict ourselves to such representations where the last component is zero.) The various irreducible representations in the product $A \times B$ are obtained according to Littlewood's rules[3] by applying h'_3 symbols α and h'_4 symbols β to the rows of the Young diagram

[1] M. Moshinsky, J. Math. Phys. **4**, 1128 (1963).

[2] J. G. Nagel and M. Moshinsky, J. Math. Phys. **6**, 682 (1965).

[3] D. E. Littlewood, *Theory of Group Characters* (Clarendon Press, Oxford, Eng., 1950), p. 94.

The author was supported by a NASA predoctoral fellowship during the course of this study.

APPENDIX A

The nonvanishing Christoffel 3-index symbols for the metric stated in Eq. (2) are

$$\Gamma^0_{00} = 2\beta_0 + \frac{Vc}{2r^2} - \frac{V_1 c}{2r} - \frac{\beta_1 Vc}{r},$$

$$\Gamma^0_{33} = \sin^2\theta\Gamma^0_{22} = (re^{-2\beta}/c)\sin^2\theta,$$

$$\Gamma^1_{00} = \frac{cV_0}{2r} - \frac{c\beta_0 V}{r} - \frac{V^2c^2}{2r^3} + \frac{VV_1c^2}{2r^2} + \frac{\beta_1 V^2c^2}{r^2},$$

$$\Gamma^1_{10} = \Gamma^1_{01} = \frac{\beta_1 Vc}{r} + \frac{V_1 c}{2r} - \frac{Vc}{2r^2},$$

$$\Gamma^1_{11} = 2\beta_1,$$

$$\Gamma^1_{33} = \sin^2\theta\Gamma^1_{22} = -Ve^{-2\beta}\sin^2\theta,$$

$$\Gamma^2_{21} = \Gamma^2_{12} = \Gamma^3_{31} = \Gamma^3_{13} = 1/r,$$

$$\Gamma^2_{33} = -\sin\theta\cos\theta,$$

$$\Gamma^3_{23} = \Gamma^3_{32} = \cot\theta,$$

where the subscripts 0 and 1 on the variables $V(u, r)$ and $\beta(u, r)$ indicate partial differentiation with respect to u and r, respectively.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 7 JULY 1968

Bases for the Representations of U_4 in the Chain $U_4 \supset U_2 + U_2$

V. SYAMALA DEVI AND T. VENKATARAYUDU
Andhra University, Waltair, India

(Received 3 March 1967)

The highest-weight polynomials of irreducible representations of U_3 occurring in the reduction of the direct product of two irreducible representations are constructed from Young diagrams. The different irreducible representations in the product are labeled by a parameter which distinguishes their multiplicities also. The method of this paper can be easily extended to any U_n.

I. INTRODUCTION

Moshinsky[1] discussed a procedure for deriving the Wigner coefficients for U_n and gave the details for U_3. He showed that the highest-weight polynomial P of the representation (h_1, h_2, h_3) occurring in the product $(h'_1, h'_2, 0) \times (h'_3, h'_4, 0)$ is a solution of the following equations:

$$C_{11}P = h'_1P, \quad C_{22}P = h'_2P, \quad C_{12}P = 0, \tag{1}$$

$$C_{33}P = h'_3P, \quad C_{44}P = h'_4P, \quad C_{34}P = 0, \tag{2}$$

$$C^{\mu}_{\mu}P = h_\mu P, \quad C^{\mu'}_{\mu}P = 0, \quad \mu < \mu' = 1, 2, 3, \tag{3}$$

where

$$C^{\mu'}_{\mu} = \sum_{s=1}^{4} a_{\mu s}a^{\mu'}_{s}, \quad C_{ss'} = \sum_{\mu=1}^{3} a_{\mu s}a^{\mu}_{s'},$$

and $a_{\mu s}$ are boson creation operators and a^{μ}_{s} are the corresponding annihilation operators. The solutions of Eqs. (1), (2), and (3) are polynomials in a basis for the irreducible representation $(h_1, h_2, h_3, 0)$ of U_4 in which the two subgroups U_2, whose generators are C_{11}, C_{12}, C_{21}, C_{22} and C_{33}, C_{34}, C_{43}, C_{44}, respectively, are explicitly reduced so that the polynomials are of highest weight in these subgroups. In this procedure there is the difficulty of the appearance of negative powers of functions of $a_{\mu s}$ in some forms of the solution (division problem). Suitable combinations have then to be worked out to get the required polynomial solutions. This involves considerable difficulty—even in the case of U_4. The object of this paper is to obtain these highest-weight polynomials in a more direct way, based on Littlewood's rules for building up the product representations. After getting the highest-weight polynomial, the entire basis for that representation can be obtained by applying the lowering operators[2] on the highest-weight polynomial. The generalization to the case of U_n is straightforward and gets rid of the division problem.

II. PRODUCT REPRESENTATIONS

Let $A = (h'_1, h'_2, 0)$ and $B = (h'_3, h'_4, 0)$ be two irreducible representations of U_3. (We can restrict ourselves to such representations where the last component is zero.) The various irreducible representations in the product $A \times B$ are obtained according to Littlewood's rules[3] by applying h'_3 symbols α and h'_4 symbols β to the rows of the Young diagram

[1] M. Moshinsky, J. Math. Phys. **4**, 1128 (1963).

[2] J. G. Nagel and M. Moshinsky, J. Math. Phys. **6**, 682 (1965).

[3] D. E. Littlewood, *Theory of Group Characters* (Clarendon Press, Oxford, Eng., 1950), p. 94.

of A so that: (i) After the addition of each set of symbols α and β, we must have a regular Young diagram with no two α's or two β's in the same column. (ii) If the total added symbols are read from right to left in consecutive rows of the final diagram, we obtain a lattice permutation of $\alpha^{h_3'}\beta^{h_4'}$. We use the symbols 1 and 2 for the diagram of A and symbols 3 and 4 for α and β in the diagram of B. The diagram of (h_1, h_2, h_3) in $A \times B$ now takes the form

$$\begin{array}{l} \overbrace{1\,1\cdots 1}^{h_1'}\ \overbrace{3\,3\cdots 3}^{x_1} \\ \overbrace{2\,2\cdots 2}^{h_2'}\ \overbrace{3\,3\cdots 3}^{x_2}\ \overbrace{4\,4\cdots 4}^{y_1} \\ \overbrace{3\,3\cdots 3}^{x_3}\ \overbrace{4\,4\cdots 4}^{y_2}, \end{array} \tag{4}$$

where

$$h_1' + x_1 = h_1, \quad h_2' + x_2 + y_1 = h_2, \quad x_3 + y_2 = h_3,$$
$$x_1 + x_2 + x_3 = h_3', \quad \text{and} \quad y_1 + y_2 = h_4'. \tag{5}$$

The rules (i) and (ii) lead to the inequalities

$$x_1 \geq y_1, \qquad x_1 + x_2 \geq y_1 + y_2,$$
$$h_1' \geq h_2' + x_2 \geq x_3 + y_2, \quad h_2' \geq x_3.$$

Equations (5) are not all independent as they are subject to the condition $h_1 + h_2 + h_3 = h_1' + h_2' + h_3' + h_4'$. Thus there is only one independent variable which may be chosen as y_1. All the other quantities x and y can be expressed in terms of the parameter y_1 which distinguishes between the different representations occurring in the product and takes into account their multiplicities also.

III. CONSTRUCTION OF THE HIGHEST-WEIGHT POLYNOMIAL

We can associate a unique polynomial with each product diagram (4) as follows. We divide the diagram (4) columnwise into what may be called elementary permissible diagrams (EPD). A permissible diagram is one in which the symbols of A as well as those of B are separately in lattice order. A diagram is said to be elementary if it cannot be split into two permissible diagrams. This splitting of (4) into EPD is not unique. It can, however, be made unique by adopting the following convention: If a column in (4) is not an EPD, we construct an EPD containing this column by adjoining to it other columns of (4) as we proceed from left to right in (4). To associate a polynomial with each EPD, we first write down the polynomial

$$\psi = \pi\Delta_{1\,2\cdots r}^{s_1 s_2 \cdots s_r},$$

where $s_1, s_2, \cdots, s_r$ are symbols in a column of the EPD and π is the product taken over all the columns.[4] For brevity we write $\psi = \pi(s_1, s_2, \cdots, s_r)$. In the present case $r \leq 3$. This polynomial ψ satisfies Eqs. (1) and (3), but not necessarily (2). Any EPD may be regarded as corresponding to a certain representation A_1 or B_1 or $A_1 \times B_1$. In the first two cases the associated polynomial ψ satisfies Eqs. (3) also. In the third case we apply on ψ the Young antisymmetrizer[5] of B_1. This antisymmetrized ψ satisfies all the equations (1), (2), and (3). The polynomial associated with the diagram (4) is the product of the antisymmetrized polynomials associated with each EPD. The various EPD in (4) are as follows:

$$\begin{array}{llllllllllll} 1 & 1\,1 & 1\,3 & 1 & 1 & 1 & 1\,3 & 1 & 3 & 3 \\ 2, & 2\,3, & 2 & , \ 2, & 3, & 3, & 4 & , & , & 4, & . \\ 3 & 4 & 4 & & 4 & & & & & & \end{array}$$

Their frequencies are, respectively, x_3, a, $m_3 - a$, m_4, $y_2 - m_3$, $x_2 - y_2 + m_3 - a$, m_1, m_2, $y_1 - m_1$, $x_1 - y_1 - m_3 + a$, where

$$m_1 = \min(h_1' - h_2' - x_2, y_1), \quad m_2 = \max(0, h_1' - h_2),$$
$$m_3 = \min(h_2' - x_3, y_2), \quad m_4 = \max(0, h_2' - h_3),$$
$$a = \min(m_3, x_2 - y_2 + m_3).$$

The associated polynomials are, respectively, (123), (12)(134), (3)(124) − (4)(123), (12), (134), (13), (1)(34), (1), (34), and (3). The polynomial corresponding to the diagram (4) (on rearrangement) is

$$P = (123)^{x_3}[(12)(134)]^{a}[(3)(124) - (4)(123)]^{m_3 - a} \times (1)^{h_1' - h_2' - x_2}(134)^{y_2 - m_3}(12)^{m_4}(13)^{x_2 - y_2 + m_3 - a} \times (34)^{y_1}(3)^{x_1 - y_1 - m_3 + a}.$$

The polynomial obtained here coincides with the polynomials (6.7) and (6.15) given by Moshinsky.[1] This can be seen by setting $y_1 = q$ and observing that a becomes m_3 or $x_2 - y_2 + m_3$ respectively, as $x_2 - y_2 = h_2 - h_2' - h_4' \geq 0$ or < 0. The symbols in Ref. 1 are to be identified with the polynomials (3)(124) − (4)(123).

The results in the case of the chain $U_6 \supset U_3 + U_3$ will be given in a separate paper.

ACKNOWLEDGMENTS

The authors' thanks are due to Professor D. M. Dennison and Professor M. Moshinsky for their kind interest in this work.

[4] G. E. Baird and L. C. Biedenharn, J. Math. Phys. **4**, 1459 (1963).
[5] H. Weyl, *Classical Groups* (Princeton University Press, Princeton, N.J., 1946), p. 120.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 7 JULY 1968

Spherical Model as an Instance of Eigenvalue Degeneracy

COLIN J. THOMPSON*
Applied Mathematics Department, The University of New South Wales, Kensington, N.S.W., Australia

(Received 18 September 1967)

It is shown that the free energy of the spherical model can be expressed in terms of the largest eigenvalue of an integral equation. In three dimensions the spectrum of the integral equation becomes degenerate as the critical point is approached from the high-temperature side, thus heralding the onset of long-range order.

1. INTRODUCTION

In the absence of any generally satisfactory theory of phase transitions, one is led to study idealized models which yield to exact analysis. The hope is that in so doing one will gain some insight into the intricacies of the transition region and that a general underlying mathematical mechanism for phase transitions will emerge.

For the Ising model with nearest-neighbor interactions a mechanism for the transition is well known in the matrix formulation of the problem, where the free energy is related to the largest eigenvalue of a matrix,[1] and long-range order exists if and only if the largest eigenvalue is asymptotically degenerate[2] (in the limit of an infinite lattice). Eigenvalue degeneracy, then, may be said to provide a mathematical mechanism for the Ising-model phase transition.

Until recently it was generally thought that the mechanisms underlying the classical theories of phase transitions, which are based on the assumption of weak long-range forces, and the theories based on short-range forces must be essentially different. Kac, however, has proposed a class of Ising models with exponentially decaying interactions, which approach on the one hand, the classical theories (when the inverse range of interaction γ approaches zero) and, on the other hand, the short-range or Onsagarian theories[3] (γ approaches infinity). For general γ, Kac has shown that the free energy can be expressed in terms of the largest eigenvalue of an integral equation and that long-range order exists if and only if the largest eigenvalue is degenerate. In two and three dimensions it has been demonstrated (but not proved) that the largest eigenvalue is asymptotically degenerate at sufficiently low temperatures and for small γ, and it has been proved that in the classical limit ($\gamma \to 0$) eigenvalue degeneracy occurs in one and higher dimensions. Kac suggests that such a mechanism may well prove to be general. Thus, given a Hamiltonian, one might hope to construct a linear operator whose largest eigenvalue is related to the free energy of the system and that the degeneracy of the largest eigenvalue can be associated with a phase transition.

The purpose of this note is to show that the spherical-model transition[4] can be considered from this point of view. We will show in fact that the free energy can be expressed in terms of the largest eigenvalue of an integral equation and that the spectrum of the integral operator becomes degenerate as the critical point is approached from the high-temperature side.

2. ONE-DIMENSIONAL MODEL

To illustrate the procedure let us consider first the one-dimensional spherical model consisting of a chain of N spins x_j such that $-\infty < x_j < \infty$, $x_{N+1} = x_1$ and

$$\sum_{j=1}^{N} x_j^2 = N. \tag{1}$$

The partition function is given by

$$Z_N = \int \cdots \int_{\sum_{j=1}^{N} x_j^2 = N} dx_1 \cdots dx_N \exp\left[2K \sum_{j=1}^{N} x_j x_{j+1}\right], \tag{2}$$

where $K = J/kT$ and J is the coupling constant between neighboring spins. Introducing the constraint (1) as a delta function in the integrand, one obtains, after elementary manipulation,

$$Z_N = \frac{1}{2\pi i} \int_{\alpha_0 - i\infty}^{\alpha_0 + i\infty} e^{Ns} Q_N(s), \tag{3}$$

where

$$Q_N(s) = \int_{-\infty}^{\infty} \cdots \int dx_1 \cdots dx_N \times \exp\left[-s \sum_{j=1}^{N} x_j^2 + 2K \sum_{j=1}^{N} x_j x_{j+1}\right], \tag{4}$$

* Address from September 1968: Applied Mathematics Department, Massachusetts Institute of Technology, Cambridge, Mass. 02139.

[1] H. A. Kramers and G. H. Wannier, Phys. Rev. **60**, 252, 263 (1941).

[2] J. Askin and W. E. Lamb, Jr., Phys. Rev. **64**, 159 (1943). See also the review article by G. F. Newell and E. W. Montroll, Rev. Mod. Phys. **25**, 353 (1953).

[3] For a review see M. Kac, Brandeis Lectures, 1966.

[4] T. H. Berlin and M. Kac, Phys. Rev. **86**, 821 (1952).

and α_0 is to be chosen so that all the singularities of the integrand are to the left of the line $s = \alpha_0$.

The integral in (4) can be done trivially by diagonalizing the quadratic form in the exponent, and the final integral over s can be done by the method of steepest descents.[4] It is important to note that the saddle point is real (which is necessary to obtain a real partition function). We remark also, following Lewis and Wannier,[5] that $Q_N(s)$ may be considered as the grand-partition function with the conjugate variable s deing determined by the condition

$$N = -(\partial/\partial s)[\log Q_N(s)]. \tag{5}$$

Equation (5) is essentially equivalent to the equation determining the saddle point for (3), *and either way one can consider s in Eq.* (4) *as a real variable.* We will do so henceforth.

Instead of evaluating $Q_N(s)$ directly, we write the integral (4) in the form

$$\begin{aligned} Q_N(s) &= \int_{-\infty}^{\infty}\cdots\int dx_1\cdots dx_N \\ &\quad\times \exp\left[-K(zx_1^2 - 2x_1x_2 + zx_2^2)\right] \\ &\quad\times \exp\left[-K(zx_2^2 - 2x_2x_3 + zx_3^2)\right]\cdots \\ &\quad\times \exp\left[-K(zx_N^2 - 2x_Nx_1 + zx_1^2)\right] \\ &= \int_{-\infty}^{\infty} dx_1 K^{(N)}(x_1, x_1), \end{aligned} \tag{6}$$

where $K^{(n)}(x, y)$ denotes the nth iterate of the integral operator

$$K(x, y) = \exp\left[-K(zx^2 - 2xy + zy^2)\right], \tag{7}$$

and

$$s = 2Kz. \tag{8}$$

Consider now the integral equation

$$\int_{-\infty}^{\infty} K(x, y)\phi(y)\, dy = \lambda\phi(x). \tag{9}$$

It is not difficult to show that the kernel $K(x, y)$ is symmetric, positive-definite, and that it is of Hilbert–Schmidt type, provided that $z > 1$. That is,

$$\int_{-\infty}^{\infty}\int K^2(x, y)\, dx\, dy = \pi/2K(z^2 - 1)^{\frac{1}{2}} < \infty,$$

$$\text{provided that} \quad z > 1, \tag{10}$$

and it is essential now, of course, that z be real. It follows that, for $z > 1$, (9) has a discrete, nondegenerate set of positive eigenvalues $\lambda_0 > \lambda_1 > \cdots$, and a corresponding complete orthonormal set of eigenfunctions $\phi_j(x)$, and that

$$K^{(n)}(x, y) = \sum_{j=0}^{\infty} \lambda_j^n \phi_j(x)\phi_j(y). \tag{11}$$

[5] H. W. Lewis and G. H. Wannier, Phys. Rev. **88**, 682 (1952).

When $z = 1$, the integral equation (9) becomes

$$\int_{-\infty}^{\infty} \exp\left[-K(x - y)^2\right]\phi(y)\, dy = \lambda\phi(x), \tag{12}$$

which has a continuous spectrum $[(\pi/K)^{\frac{1}{2}}, \infty)$ for real functions $[\phi(x) = \exp(\alpha x)]$ and a continuous spectrum $[0, \infty)$ for complex functions $[\phi(x) = \exp(i\alpha x)]$. In both cases, of course, the eigenfunctions are not contained in $L^2(-\infty, \infty)$.

Substituting (11) into (6) gives

$$Q_N(s) = \sum_{j=0}^{\infty} \lambda_j^N \sim \lambda_0^N \quad \text{as} \quad N \to \infty,$$

$$\text{provided} \quad z = s/2K > 1, \tag{13}$$

and the problem is to find the largest eigenvalue λ_0 of the integral equation (9).

The complete spectrum of (9) can easily be obtained by noting that the integral operator $K(x, y)$ [Eq. (7)] commutes with the differential operator

$$L = -(d^2/dx^2) + 4K^2(z^2 - 1)x^2. \tag{14}$$

The normalized eigenfunctions of L, and consequently of (9), are the Weber functions

$$\begin{aligned} \phi_n(x) &= \frac{1}{(2^n n!)^{\frac{1}{2}}}\left[\frac{2K(z^2 - 1)^{\frac{1}{2}}}{\pi}\right]^{\frac{1}{4}} \\ &\quad\times \exp\left[-K(z^2 - 1)^{\frac{1}{2}}x^2\right] H_n\{[2K(z^2 - 1)^{\frac{1}{2}}]^{\frac{1}{2}}x\}, \end{aligned} \tag{15}$$

where $H_n(x)$ is the nth Hermite polynomial. The eigenvalues of (9) are then found by substituting (15) into (9) and putting $x = 0$ for n even, and differentiating and putting $x = 0$ for n odd. One then finds that

$$\lambda_n = (\pi/K)^{\frac{1}{2}}[z + (z^2 - 1)^{\frac{1}{2}}]^{-(n+\frac{1}{2})}, \tag{16}$$

and in particular from (13), provided $z > 1$,

$$\begin{aligned} \frac{1}{N}\log Q_N(s) \sim \log\lambda_0 &= \tfrac{1}{2}\log(\pi/2K) \\ &\quad - \tfrac{1}{2}\log\left[\frac{z + (z^2 - 1)^{\frac{1}{2}}}{2}\right], \end{aligned} \tag{17}$$

which is precisely the value obtained by integrating (4) directly. The equation determining the "saddle point" z_s [or, equivalently, the conjugate variable s from (5)] of the integral (3) is

$$4K = (z_s^2 - 1)^{-\frac{1}{2}} \quad \text{or} \quad z_s = [1 + (4K)^{-2}]^{\frac{1}{2}} > 1. \tag{18}$$

Notice from (16) that

$$\lim_{z\to 1+} \lambda_n = (\pi/K)^{\frac{1}{2}}, \quad \text{for all} \quad n = 0, 1, 2, \ldots. \tag{19}$$

That is, in the limit $z \to 1+$ the whole discrete spectrum collapses into a point, or, in other words, *the discrete spectrum becomes degenerate as* $z \to 1+$.

We will show that this degeneracy (i.e., the collapsing of the discrete spectrum into a point) heralds the transition in three dimensions. In one and two dimensions the saddle point z_s [see Eq. (18)] is such that the discrete nondegenerate spectrum is maintained at all finite temperatures.

3. TWO- AND THREE-DIMENSIONAL MODELS

We now discuss the two- and three-dimensional models. Consider first a two-dimensional $N \times M$ lattice wrapped on a torus, with spins located on the vertices. Then, by analogy with the transfer-matrix method for the Ising model [or a straightforward generalization of the reduction (6) above], we have that

$$Q_{N,M}(s) = \int_{-\infty}^{\infty} \cdots \int \prod dx_{ij} \exp\left[-s \sum_{i=1}^{N} \sum_{j=1}^{M} x_{i,j} + K \sum_{\substack{(i,j) \\ (k,l)}}^{*} x_{i,j} x_{k,l}\right]$$
$$= \int_{-\infty}^{\infty} \cdots \int dy_1 \cdots dy_M K^{(N)}(\mathbf{y}, \mathbf{y}), \tag{20}$$

where the starred sum is over nearest-neighbor sites (i,j) and (k,l), $\mathbf{y} = (y_1, \cdots, y_M)$ and $K^{(n)}(\mathbf{x}, \mathbf{y})$ denotes the nth iterate of the M-dimensional integral operator

$$K(\mathbf{x}, \mathbf{y}) = \exp\left[-K \sum_{i=1}^{M} (z x_i^2 - x_i x_{i+1} - 2x_i y_i - y_i y_{i+1} + z y_i^2)\right]. \tag{21}$$

The partition function is given by (3) with $Q_N(s)$ replaced by $Q_{N,M}(s)$.

Consider then the integral equation

$$\int_{-\infty}^{\infty} \cdots \int dy_1 \cdots dy_M K(\mathbf{x}, \mathbf{y}) \phi(\mathbf{y}) = \Lambda \phi(\mathbf{x}). \tag{22}$$

The change of variables

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} \to \sum_{r=0}^{M-1} a_{lr} \begin{pmatrix} x_r \\ y_r \end{pmatrix},$$

where

$$a_{lr} = \begin{cases} (1/M)^{\frac{1}{2}}, & \text{for } r = 0, \\ (2/M)^{\frac{1}{2}} \cos(2\pi ln/M), & \text{for } r = 2n, \\ (2/M)^{\frac{1}{2}} \sin(2\pi ln/M), & \text{for } r = 2n-1, \end{cases} \tag{23}$$

transforms the kernel $K(\mathbf{x}, \mathbf{y})$ into separable form, i.e.,

$$K(\mathbf{x}, \mathbf{y}) \to \sum_{s=0}^{M-1} \exp[-K(w_{\tilde{s}} x_s^2 - 2x_s y_s + w_{\tilde{s}} y_s^2)], \tag{24}$$

where

$$w_s = z - \cos(2\pi s/M) \tag{25}$$

and $\tilde{s}$ denotes the integral part of $(s+1)/2$. The eigenvalues of (22) are then products of eigenvalues of (9) with z replaced by $w_{\tilde{s}}$. In particular,

$$\Lambda_0 = \prod_{s=0}^{M-1} \{(K/\pi)[w_{\tilde{s}} + (w_{\tilde{s}}^2 - 1)^{\frac{1}{2}}]\}^{-\frac{1}{2}},$$
$$\phi_0(\mathbf{x}) = \prod_{s=0}^{M-1} \left[2K\left(\frac{w_{\tilde{s}}^2 - 1}{\pi^2}\right)^{\frac{1}{2}}\right]^{\frac{1}{4}} \exp[-K(w_{\tilde{s}}^2 - 1)^{\frac{1}{2}} x_s^2] \tag{26}$$

and

$$\Lambda_1 = [w_0 + (w_0^2 - 1)^{\frac{1}{2}}]^{-1} \Lambda_0,$$
$$\phi_1(\mathbf{x}) = x_0[4K(w_0^2 - 1)^{\frac{1}{2}}]^{\frac{1}{2}} \phi_0(\mathbf{x}). \tag{27}$$

In the limit $N, M \to \infty$, we have from (20) and (26), *provided* $z > 2$, that

$$Q_{N,M} \sim \Lambda_0^N \tag{28}$$

and

$$\Lambda_0 \sim (\pi/2K)^{M/2} \exp\left\{-\frac{M}{2} f_2(z)\right\}, \tag{29}$$

where $[w(\theta) = z - \cos\theta]$

$$f_2(z) = \frac{1}{2\pi} \int_0^{2\pi} d\theta \log\{[w(\theta) + (w^2(\theta) - 1)^{\frac{1}{2}}]/2\}$$
$$= \frac{1}{2\pi} \int_0^{2\pi} d\theta \cosh^{-1}(z - \cos\theta) - \log 2$$
$$= \left(\frac{1}{2\pi}\right)^2 \iint_0^{2\pi} \log(z - \cos\theta - \cos\phi)\, d\theta\, d\phi, \tag{30}$$

which is precisely the result obtained previously.[4]

The equation determining the saddle point is

$$4K = \left(\frac{1}{2\pi}\right)^2 \iint_0^{2\pi} \frac{d\theta\, d\phi}{z_s - \cos\theta - \cos\phi}, \tag{31}$$

and it is easy to show that for all $K > 0$ (i.e., for all temperatures) there is a unique solution for z_s such that $z_s > 2$.[4] Thus $w_0 = z - 1 > 1$, and from (27) the spectrum of the integral equation is nondegenerate for all temperatures.

For a three-dimensional $N \times M \times M$ lattice one proceeds exactly as above, and, omitting obvious detail, one finds that as $N, M \to \infty$

$$Q_{N,M,M} \sim \Lambda_0^N, \tag{32}$$

where Λ_0 is the largest eigenvalue of the M^2-dimensional integral equation

$$\int_{-\infty}^{\infty} \cdots \int \prod_{ij=1}^{M} dy_{i,j} \exp\left[-K \sum_{i,j=1}^{M} (z x_{i,j}^2 - x_{i,j} x_{i,j+1} - x_{i,j} x_{i+1,j} - 2x_{i,j} y_{i,j} - y_{i,j} y_{i,j+1} - y_{i,j} y_{i+1,j} + z y_{i,j}^2)\right] \phi(\{y_{ij}\}) = \Lambda \phi(\{x_{i,j}\}). \tag{33}$$

Transforming to diagonal form, one finds that

$$\Lambda_0 = \prod_{s,t=0}^{M-1} \{(K/\pi)[w_{\tilde{s},\tilde{t}} + (w_{\tilde{s},\tilde{t}}^2 - 1)^{\frac{1}{2}}]\}^{-\frac{1}{2}} \tag{34}$$

and

$$\Lambda_1 = [w_{0,0} + (w_{0,0}^2 - 1)^{\frac{1}{2}}]^{-1}\Lambda_0, \text{ etc.},$$

where

$$w_{s,t} = z - \cos\left(\frac{2\pi s}{M}\right) - \cos\left(\frac{2\pi t}{M}\right). \quad (35)$$

From (32), (34), and (35) one has, as $N, M \to \infty$ and provided that $z > 3$, that

$$Q_{N,M,M} \sim (\pi/2K)^{M^2/2} \exp\left[-\frac{M^2}{2} f_3(z)\right], \quad (36)$$

where

$$f_3(z) = \left(\frac{1}{2\pi}\right)^3 \int_0^{2\pi}\int\int\int \log\,[z - \cos\theta - \cos\phi - \cos\chi] \times d\theta\, d\phi\, d\chi. \quad (37)$$

The equation determining the saddle point is

$$4K = \left(\frac{1}{2\pi}\right)^3 \int_0^{2\pi}\int\int\int \frac{d\theta\, d\phi\, d\chi}{z_s - \cos\theta - \cos\phi - \cos\chi}. \quad (38)$$

The integral (38) converges now as $z_s \to 3+$, giving a critical point $4K_c = 0.50546$, and one can show that for $K > K_c$ the saddle point "sticks" at $z_s = 3$.[4] We now note from (34) that, as $z \to 3+$, $w_{0,0} = z - 2 \to 1+$ and hence $\Lambda_0/\Lambda_1 \to 1$, or, in other words, the *spectrum of the integral equation becomes degenerate as* $K \to K_c^-$.

4. EIGENVALUE DEGENERACY AND LONG-RANGE ORDER

To complete the analogy with the Ising model we derive expressions for the pair-correlation function in terms of the eigenvalues and eigenfunctions of the integral equation and show that long-range order exists if and only if there is eigenvalue degeneracy.

Consider the one-dimensional case. From (3), (4), and (6) we have that

$$\rho_N(r) = \langle x_j x_{j+r}\rangle = Z_N^{-1}(2\pi i)^{-1}\int_{-\infty}^{\infty} ds\, e^{Ns} \times \int_{-\infty}^{\infty}\cdots\int dx_1 \cdots dx_N x_j x_{j+r} \times \exp\left[\sum_{j=1}^{N}(-sx_j^2 + 2Kx_j x_{j+1})\right]$$

$$\sim \lambda_0^{-N}\int_{-\infty}^{\infty}\int\int\int dx\, dy\, du K^{(j-1)}(x, y) \times y K^{(r)}(y, u) u K^{(N-j-r+1)}(u, x), \quad (39)$$

where, as before, $K^{(n)}(x, y)$ denotes the nth iterate of the kernel (7) and z appearing in λ_0 and $K^{(n)}(x, y)$ is the saddle point value z_s, Eq. (18). Using the decomposition (11) of the iterated kernel in terms of the eigenvalues and eigenfunctions of the integral equation (9), (39) becomes

$$\rho_N(r) \sim \lambda_0^{-N}\sum_{l=0}^{\infty}\sum_{m=0}^{\infty}\lambda_l^{N-r}\lambda_m^r [A_{l,m}^{(1)}]^2, \quad (40)$$

where

$$A_{l,m}^{(1)} = \int_{-\infty}^{\infty} dx \phi_l(x) x \phi_m(x), \quad (41)$$

and finally, keeping r fixed and letting $N \to \infty$, we have that

$$\rho^{(1)}(r) = \lim_{N\to\infty}\rho_N(r) = \sum_{m=1}^{\infty}(\lambda_m/\lambda_0)^r [A_{0,m}^{(1)}]^2. \quad (42)$$

From (15) and the fact that

$$\frac{1}{(\pi)^{\frac{1}{2}}}\int_{-\infty}^{\infty} H_n(x)H_m(x)e^{-x^2}\,dx = \delta_{n,m}2^n n!, \quad (43)$$

we have from (41) and (18) that

$$A_{0,m}^{(1)} = \delta_{m,1}[4K(z_s^2 - 1)^{\frac{1}{2}}]^{-\frac{1}{2}} = \delta_{m,1}, \quad (44)$$

and hence from (16) and (41) that

$$\rho^{(1)}(r) = (\lambda_1/\lambda_0)^r = [z_s - (z_s^2 - 1)^{\frac{1}{2}}]^r, \quad (45)$$

which agrees with formula (39) of Ref. 4. Note that the correlation function is directly related to the ratio of the largest and next largest eigenvalue of the integral equation. Precisely the same is true in two and three dimensions. Thus, proceeding exactly as above, the pair, correlation function for two spins separated by a distance r in the "N direction" (take i ranging from 1 to N and j and k from 1 to M) is given in dimensions $d = 2$ and 3 by

$$\rho^{(d)}(r) = \sum_{m=1}^{\infty}(\Lambda_m/\Lambda_0)^r [A_{0,m}^{(d)}]^2, \quad (46)$$

where

$$\rho^{(d)}(r) = \begin{cases} \lim\limits_{N,M\to\infty}\langle x_{i,j}x_{i+r,j}\rangle & d = 2, \\ \lim\limits_{N,M\to\infty}\langle x_{i,j,k}x_{i+r,j,k}\rangle & d = 3, \end{cases} \quad (47)$$

$$A_{l,m}^{(d)} = \int_{-\infty}^{\infty}\cdots\int d\mathbf{x}\phi_1(\mathbf{x})x\phi_m(\mathbf{x}), \quad (48)$$

and for $d = 2$ and 3, respectively, $\mathbf{x} = (x_j)$, $\mathbf{x} = (x_{j,k})$, and Λ_m and $\phi_m(\mathbf{x})$ are the eigenvalues and eigenfunctions of the integral equations (22) and (33). It is not difficult to reduce (46) to the formulas given in Ref. 4. We merely remark that, in three dimensions as $T \to T_c+$, $\Lambda_1/\Lambda_0 \to 1$ and long-range order sets in. For $T > T_c$ $\lim\limits_{r\to\infty}(\Lambda_1/\Lambda_0)^r = 0$ and there is no long-range order.

ACKNOWLEDGMENTS

The author is grateful to Professor M. Kac for valuable discussions and correspondence.

The support of the Commonwealth Government through the award of a Queen Elizabeth II Fellowship is also gratefully acknowledged.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 7 JULY 1968

3-Sphere "Backgrounds" for the Space Sections of the Taub Cosmological Solution*

RICHARD A. MATZNER†
Department of Applied Mathematics and Theoretical Physics, University of Cambridge, and Department of Physics and Astronomy, University of Maryland

(Received 3 November 1967)

The cosmological solution due to Taub resembles a radiation filled Robertson–Walker solution which contains long-wavelength gravitational radiation rather than electromagnetic radiation. This paper presents several schemes for obtaining 3-sphere "backgrounds" for the t = const space slices $\mathcal{M}(t)$ in the Taub solution. [$\mathcal{M}(t)$ is characterized by being homogeneous for each t.] The reason for finding such backgrounds is to present models for defining wave-background separation in general. Unspecified averaging methods as suggested by Isaacson work for short-wavelength gravitational radiation. We attempt here to specify a method which works even when the radiation is of long wavelength. One background is obtained by averaging the metric by Lie transport along certain invariantly defined vector fields on $\mathcal{M}(t)$. This background is compared with two other (different) plausible definitions for a background, and reasons are given to suggest that the Lie-transport average is the preferable definition.

I. INTRODUCTION

Isaacson[1] has given a discussion of high-frequency gravitational radiation when the metric can be written in coordinates such that $g_{\alpha\beta} = \gamma_{\alpha\beta} + \epsilon h_{\alpha\beta}$; $\gamma_{\alpha\beta}$ is supposed to be a slowly varying function of position—the "background metric," and $h_{\alpha\beta}$ is a rapidly varying function of position.

Considering gravitational radiation from this viewpoint, several problems are suggested. One is to find a background if one exists in a space presumably containing a background plus wave. Isaacson assumes that if such separation is possible, it can be found by sufficiently intense inspection. But it would be much more satisfactory to have an invariant method of finding the background when one exists. We propose that the background can be found by averaging along certain invariantly defined vector fields. On the Taub space slices $\mathcal{M}(t)$, these fields are invariantly defined because of the topological equivalence of $\mathcal{M}(t)$ to a 3-sphere. These vector fields are also characterized by the fact that (a) they are Killing fields on a 3-sphere (in an obvious sense defined by identifying coordinates; see Sec. II) and (b) they are eigensolutions to the differential equation[2] which has been suggested as a generalization of Killing's equation

$$\xi^{(i,j)}{}_{,j} + \lambda\xi^i = 0, \tag{1}$$

where the comma denotes covariant derivative in $\mathcal{M}(t)$. This latter characterization is the one we propose in general. On the slices $\mathcal{M}(t)$, criteria (a) and (b) specify the only two vector fields which are Killing on the 3-sphere but not on $\mathcal{M}(t)$.

Of the remaining four Killing fields on $\mathcal{M}(t)$, three describe the spatial homogeneity of the space, while the fourth gives one axis of isotropy at each point. $\mathcal{M}(t)$ can in fact be characterized by its homogeneity.

II. THE TAUB SOLUTION

Taub[3] obtained a cosmological solution which is in a sense a generalization of the Robertson–Walker[4] (R.W.) metrics. It has $R_{\alpha\beta} = 0$ and homogeneous but not isotropic t = const space sections $\mathcal{M}(t)$. In one particular coordinate system[5]

$$\begin{aligned} ds^2 = {} & (t^2 + l^2)(d\theta^2 + \sin^2\theta\, d\phi^2) \\ & + U(t)(2l)^2(d\psi + \cos\theta\, d\phi)^2 \\ & - 2(2l)(d\psi + \cos\theta\, d\phi)\, dt. \end{aligned} \tag{2}$$

Here $\theta \in [0, \pi]$, $\phi \in [0, 2\pi\rangle$ and $\psi \in [0, 4\pi\rangle$ are coordinates on S^3; l is a constant length ($l > 0$), and

$$U(t) = -1 + \frac{2(mt + l^2)}{t^2 + l^2} = -\frac{(t - t_-)(t - t_+)}{t^2 + l^2},$$

where $t_\pm = m \pm (m^2 + l^2)^{\frac{1}{2}}$; m is another positive constant length. The function $U(t)$ is positive only for $t_- < t < t_+$; for t in this range, $\mathcal{M}(t)$ is spacelike. The surfaces $t = t_\pm$ are null surfaces which bound the solution from another region of 4-space described by the empty Newman–Unti–Tamburino (NUT) space,[6]

* Supported in part by NASA grant NSG 436 and by the Aerospace Research Lab., Office of Aerospace Research, U.S. Air Force. This work is based on part of a thesis submitted as partial fulfillment of Ph.D. requirements at the University of Maryland.

† Presently NSF Faculty Associate at the University of Texas, Austin, Texas.

[1] R. A. Isaacson, Phys. Rev. **166**, 1263 (1968); **166**, 1272 (1968).

[2] R. A. Matzner, J. Math. Phys. **9** (1968) (to be published).

[3] A. H. Taub, Ann. Math. **54**, 472 (1951).

[4] For a modern exposition of these metrics see S. W. Hawking, Astrophys. J. **145**, 544 (1966).

[5] C. W. Misner and A. H. Taub, paper submitted to J.E.T.P. (1967).

[6] E. Newman, L. Tamburino, and T. Unti, J. Math. Phys. **4**, 915 (1963).

which is given by the same metric in the region where $U(t) < 0$. In NUT space, t is a spacelike coordinate.

The homogeneity of the space slices is demonstrated by the three Killing vectors [which are Killing in $\mathcal{M}(t)$ for each t[5]]:

$$\begin{aligned}\xi_x &= -\sin\phi\partial_\theta - \cos\phi(\cot\theta\partial_\phi - \csc\theta\partial_\psi),\\ \xi_y &= \cos\phi\partial_\theta - \sin\phi(\cot\theta\partial_\phi - \csc\theta\partial_\psi),\\ \xi_z &= \partial_\phi.\end{aligned} \tag{3}$$

Although these look like rotations, they have no fixed points, since they have nonvanishing length.

There is also one more Killing vector in $\mathcal{M}(t)$:

$$\mathcal{N}_z = -\partial_\psi.$$

In addition, we may define two more vectors which are invariantly distinguished on the 3-sphere, but are *not* Killing in $\mathcal{M}(t)$:

$$\begin{aligned}\mathcal{N}_x &= -\sin\psi\partial_\theta + \cos\psi(\csc\theta\partial_\phi - \cot\theta\partial_\psi),\\ \mathcal{N}_y &= -\cos\psi\partial_\theta - \sin\psi(\csc\theta\partial_\phi - \cot\theta\partial_\psi).\end{aligned} \tag{4}$$

The topological equivalence mentioned in Sec. I is simply given by the coordinates θ, ϕ, ψ, which would be the Euler-angle coordinates on S^3 if $t^2 + l^2 = 4l^2U$.

The Taub solution is a generalization of the closed R.W. model in the sense that $\mathcal{M}(t)$ for each $t \in \langle t_-, t_+\rangle$ has only four Killing vectors. The corresponding slice in the R.W. metric is topologically S^3 also, but the metric there does not distinguish any direction on the 3-sphere, and the R.W. metrics have six Killing vectors in each time slice; the full set of generators for the symmetries of the 3-sphere. The Taub solution is analytic[5] for $t \in \langle t_-, t_+\rangle$.

As we mentioned above, $\mathcal{N}_x$ and $\mathcal{N}_y$ are eigensolutions to Eq. (1); their eigenvalues are, in fact, equal and equal to

$$\lambda_{x,y} = \tfrac{1}{2}B^2(B^{-2} - A^{-2})^2,$$

where $A^2(t) \equiv t^2 + l^2$ and $B^2(t) = 4l^2U(t)$. The eigenvalue vanishes when $A^2 = B^2$ in which case the 3-space $\mathcal{M}(t)$ is instantaneously isotropic as well as homogeneous. Thus at this instant $\mathcal{N}_x$ and $\mathcal{N}_y$ are Killing and we identify them with the "missing" Killing vectors in general $\mathcal{M}(t)$.

The vectors ξ_i and $\mathcal{N}_i$ are nowhere zero and so are well suited to the averaging by Lie transport along them which we use in the following section.

We shall have need of the following commutation relations for the invariant vector fields[5] (where these are Lie brackets):

$$\begin{aligned}[\xi_i, \xi_j] &= -\epsilon_{ijk}\xi_k,\\ [\xi_i, \mathcal{N}_j] &= 0,\\ [\mathcal{N}_i, \mathcal{N}_j] &= -\epsilon_{ijk}\mathcal{N}_k.\end{aligned} \tag{5}$$

III. THE ISOTROPIC BACKGROUND IN $\mathcal{M}(t)$

It has been suggested[7,8] that Taub space is an R.W.-type solution, similar to Brill's gravitational radiation filled one,[7] but with the longest-wavelength gravitational wave that will fit into it giving the energy density to close it in place of the matter or radiation in the usual R.W. forms. We show here a way of finding underlying closed constant-curvature spaces for $\mathcal{M}(t)$ by a procedure of iteratively averaging the metric for the Taub slice $\mathcal{M}(t)$ along the non-Killing members of the sets $\{\{\xi_i\}$ and $\{\mathcal{N}_i\}$. These are the vector fields that satisfy (a) and (b) of Sec. I.

In the 3-space, six averaging vector fields are necessary to obtain a completely homogeneous and isotropic background space. For one can, by averaging over three fields which describe homogeneity in a 3-space, obtain an anisotropic but completely homogeneous space. Further averaging along these vector fields, which are then Killing, will have no effect. More averaging fields are needed.

In this section, we shall make extensive use of the invariant definition of the vectors $\mathcal{N}_i$ and ξ_i. We shall average along these fields, and we note that $\mathcal{N}_i$ and ξ_i are everywhere nonzero in $\mathcal{M}(t)$. It clearly will be pointless to average along a Killing vector by the method of Lie transport, which we use here, so in this case we have initially only two vectors to consider averaging along $\mathcal{N}_x$ and $\mathcal{N}_y$. In at least some completely nonsymmetric situations—such as those obtained by small general perturbations from an R.W. form—we can have up to six such candidates for the vector fields in the 3-space. It is, of course possible that by averaging along one vector field we destroy the Killing nature of another. We shall see that after averaging along $\mathcal{N}_x$, say, in the space $\mathcal{M}(t)$, $\mathcal{N}_z$ is no longer Killing. (If, as we do here, we average along a complete trajectory of $\mathcal{N}_x$, then afterward $\mathcal{N}_x$ is clearly Killing.) Because of the commutation relations, Eq. (5), the ξ_i stay Killing under $\mathcal{N}_i$ averaging.

Now consider averaging the contravariant metric tensor g^{ab} of $\mathcal{M}(t)$. We do this by a process of Lie transport. Pick one of the averaging vector fields, say $\mathcal{N}_x$. We then average $g^{\cdot\cdot}(p)$ along $\mathcal{N}_x$ by carrying the space back to the point p along the trajectories of $\mathcal{N}_x$ and averaging the values of $g^{\cdot\cdot}$ there for the whole trajectory of $\mathcal{N}_x$.

We use the following formula due to Misner[9]:

$$g^{\cdot\cdot} = A^{-2}\xi_i \otimes \xi_i + (B^{-2} - A^{-2})\mathcal{N}_z \otimes \mathcal{N}_z, \tag{6}$$

[7] D. R. Brill, Nuovo Cimento Suppl. **2**, No. 1 (1964).

[8] J. A. Wheeler, "Geometrodynamics and the Issue of Final State" in *Relativity Groups and Topology* (Gordon and Breach Science Publishers, London, 1964), para. 9 and problem 55.

[9] C. W. Misner, J. Math. Phys. **4**, 924 (1963).

where we indicate by $g^{\cdot\cdot}$ the contravariant metric in $\mathcal{M}(t)$ (as distinct from its components g^{ab}).

After the averaging, $\mathcal{N}_x$ will clearly be a Killing vector of the resultant space. We denote the back-translated metric (a parameter length r along the curves) as $g^{\cdot\cdot}(p; rx)$. Formulas in Schouten[10] then give for this finite translation:

$$g^{\cdot\cdot}(p; rx) = e^{r\mathfrak{L}_x} g^{\cdot\cdot}(p) \equiv g^{\cdot\cdot}(p) + r(\mathfrak{L}_x g^{\cdot\cdot})(p) + \frac{r^2}{2}(\mathfrak{L}_x^2 g^{\cdot\cdot})(p) + \cdots . \quad (7)$$

The symbol $\mathfrak{L}_x$ denotes Lie differentiation along $\mathcal{N}_x$. Since $[\mathcal{N}_i, \xi_j] = 0$ and by Eqs. (5) we see that

$$\mathfrak{L}_x(\mathcal{N}_z \otimes \mathcal{N}_z) = \mathcal{N}_y \otimes \mathcal{N}_z + \mathcal{N}_z \otimes \mathcal{N}_y,$$
$$\mathfrak{L}_x^2(\mathcal{N}_z \otimes \mathcal{N}_z) = -2\mathcal{N}_z \otimes \mathcal{N}_z + 2\mathcal{N}_y \otimes \mathcal{N}_y,$$

and

$$\mathfrak{L}_x^3(\mathcal{N}_z \otimes \mathcal{N}_z) = -4\mathfrak{L}_x(\mathcal{N}_z \otimes \mathcal{N}_z);$$

we obtain [remember the analyticity of $\mathcal{M}(t)$]

$$g^{\cdot\cdot}(p; rx) = g^{\cdot\cdot}(p) + \tfrac{1}{2}(\mathfrak{L}_x g^{\cdot\cdot})(p) \sin 2r + \tfrac{1}{2}(\mathfrak{L}_x^2 g^{\cdot\cdot})(p)\tfrac{1}{2}(1 - \cos 2r). \quad (8)$$

The range of the path parameter r is 0 to 4π, as can be seen by considering the equivalent quaternion translation on the 3-space. (See Ref. 9, Appendix B; this whole discussion of translations could be done in terms of quaternions on S^3.) We will average over the whole range of r and use[9] $\mathcal{N}_i \otimes \mathcal{N}_i = \xi_i \otimes \xi_i$ to obtain

$$g^{\cdot\cdot}(p; \overline{rx}) = A^{-2}(\xi_i \otimes \xi_i) + \tfrac{1}{2}(B^{-2} - A^{-2})(\xi_i \otimes \xi_i - \mathcal{N}_x \otimes \mathcal{N}_x), \quad (9)$$

This can be put again into the form of a Taub slice $\mathcal{M}(t)$, but now with a smaller asymmetry, and now $\mathcal{N}_x$ is Killing, but $\mathcal{N}_z$ is not.

If we now average with a nowhere-zero non-Killing vector in this space, say $\mathcal{N}_y$, we will get by cyclic symmetry:

$$g^{\cdot\cdot}(p; \overline{rx}, \overline{sy}) = A^{-2}(\xi_i \otimes \xi_i) + \tfrac{1}{2}(B^{-2} - A^{-2}) \times [\xi_i \otimes \xi_i - \tfrac{1}{2}(\xi_i \otimes \xi_i - \mathcal{N}_y \otimes \mathcal{N}_y)]. \quad (10)$$

We see that if this cyclic averaging is continued, the anisotropic terms $\mathcal{N} \otimes \mathcal{N}$ have a coefficient which vanishes as 2^{-n} in the nth term. The coefficient of $\xi_i \otimes \xi_i$ approaches $A^{-2} + \tfrac{1}{2}(B^{-2} - A^{-2}) \sum_0^\infty (-2)^{-n} = \tfrac{2}{3}A^{-2} + \tfrac{1}{3}B^{-2}$. This metric, $g^{\cdot\cdot} = \text{const} \times \xi_i \otimes \xi_i$, is metrically S^3, and we have succeeded in finding a symmetric background.

[10] J. A. Schouten, *Ricci Calculus* (Springer-Verlag, Berlin, 1954).

We pointed out above that the ξ_i vector fields are completely invariant under this averaging, since the $\mathcal{N}_i$ fields all commute with the ξ_i. At each step in this iterative process, there are thus only two fields along which the averaging has any effect; the two $\mathcal{N}$ fields which do not appear in the metric at that step. (This is because we have averaged along the entire trajectory—all around the space—at each step.) A calculation for $g^{\cdot\cdot}(p; ry)$ (i.e., for the metric dragged back along the vector field $\mathcal{N}_y$ instead of along the field $\mathcal{N}_x$) analogous to that leading to Eq. (9) gives an equation which is the same as Eq. (9) but with $\mathfrak{L}(\mathcal{N}_y)$ in place of $\mathfrak{L}_x$. Recalling again that the $\mathcal{N}$ fields can act only among themselves, it is easy to verify that, as we carry out the next averaging for one of the choices of non-Killing field, we obtain an equation like Eq. (10), except possibly with a different subscript on the $\mathcal{N}$ field. The point of this is that the final result for the averaged metric is independent of the order of averaging, where we average along the whole trajectory of the vector field.

It can also be shown that if one averages using a weighting function which is symmetric and decreases as a function of the distance r from the origin (not a delta function), then the end result is still independent of the order of averaging. The reason for this is that an infinite number of smoothings, as we do here, will bring adjacent points to the same "level," even when a weighting function is used.

IV. COMPARISON WITH OTHER AVERAGING DEFINITIONS

We have been able to perform this averaging because of the invariant description of the vectors $\mathcal{N}_i$ and ξ_i. One invariant characterization of these vectors is that they are eigensolutions to $(\mathfrak{D} + \lambda)\xi = 0$, where $(\mathfrak{D}\xi)^i = \xi^{(i,j)}_{\ ,j}$ and are also the fields that would be obtained from the Killing fields on a spherical space which is continuously deformed into the Taub shape. We suspect that this characterization is the relevant one; in particular, we expect that averaging on a 2-sphere for instance can be carried out by averaging along continuous vector fields in the way we did in Sec. III, even though each such field has at least two zeros. (Neither ξ_i nor $\mathcal{N}_i$ has any zeros.) Further investigation of this question and of the specification of just what fields averaging can effectively be done along remains to be carried out. In particular, since the spectrum of $-\mathfrak{D}$ specifies a countable number of vector fields on the space, we apparently have all these as candidates for averaging fields.

One might wonder what the meaning of the averaging space is. A completely symmetric background for

a space which is topologically S^3 is S^3 itself. The only unknown quantity remaining is the radius of the sphere. It might be expected that a sphere of the same volume as $\mathcal{M}(t)$ would be the simplest such average. However, the average defined here does not lead to the sphere of this radius.

Since the $\mathcal{N}_i$ are orthogonal, we can write the covariant Taub metric in terms of σ^i, the 1-forms[9] dual to the $\mathcal{N}_i$ as

$$ds^2 = A^2(\sigma^x\sigma^x + \sigma^y\sigma^y) + B^2\sigma^z\sigma^z.$$

To obtain this form we used (again) $\mathcal{N}_i \otimes \mathcal{N}_i = \xi_i \otimes \xi_i$. The integral defining the volume of the space is $\int \sqrt{g}\, \sigma_i \wedge \sigma_j \wedge \sigma_k$.

The covariant form of the averaged metric, which has the same volume as $\mathcal{M}(t)$, is thus given by

$$ds^2 = (A^4B^2)^{\frac{1}{3}}(\sigma^x\sigma^x + \sigma^y\sigma^y + \sigma^z\sigma^z),$$

where we used the fact that the metric coefficients are constant in $\mathcal{M}(t)$.

The averaged metric we obtained in Sec. III by averaging along vector fields is

$$(\tfrac{2}{3}A^{-2} + \tfrac{1}{3}B^{-2})^{-1}(\sigma^x\sigma^x + \sigma^y\sigma^y + \sigma^z\sigma^z),$$

the ratio of the volume of this space to the volume of $\mathcal{M}(t)$ is thus

$$(\tfrac{2}{3}A^{-2} + \tfrac{1}{3}B^{-2})^{\frac{3}{2}}A^2B.$$

The ratio is equal to unity only when $A^2 = B^2$; then the space is spherical.

It would be encouraging if the averaged sphere had some desirable property as a background metric. However, it is easy to check that neither the equal-volume average, nor the average defined by the Lie transport given here, has a behavior like a radiation-dominated R.W. metric. We show this by writing the averaged space we have found in the full 4-dimensional form, which makes its (3-sphere) × (time) structure apparent,[9] as

$$ds^2 = -L^2(t)(2l)^2[d\chi^2 + \sin^2\chi(d\theta^2 + \sin^2\theta\, d\phi^2)] - \left[\frac{(t-t_-)(t-t_+)}{t^2+l^2}\right]^{-1} dt^2.$$

For the equal-volume averaging,

$$L_E^2 = \left[\frac{t^2+l^2}{(2l)^2}\frac{(t-t_-)}{(2l)}\frac{(t-t_+)}{(2l)}\right]^{\frac{1}{3}}; \qquad (11)$$

for the Lie-transport averaging, we have

$$L_L^2 = \left[\frac{2}{3}\frac{(2l)^2}{t^2+l^2} + \frac{1}{3}\frac{t^2+l^2}{(t-t_-)(t-t_+)}\right]^{-1}. \qquad (12)$$

In a radiation-dominated R.W. metric, the volume of the space sections $V(\tau)$ is $V_R(\tau) \sim \tau^{\frac{3}{2}}$. In the equal-volumes average, Eq. (11), the behavior near the singularity is $V_E \sim \tau$, while for the Lie-transport average Eq. (12) it is $V_L(\tau) \sim \tau^3$. (Here τ is the proper time measured from the singularity.) Thus, neither of these averaging schemes makes the gravitational radiation (i.e., the anisotropy) appear as a massless radiation field. (This result of nonvanishing trace does not contradict the result of Brill[7] that the effective averaged stress tensor for small-amplitude short-wavelength gravitational waves is traceless to the lowest order in the deviation from the background. The waves in Taub space are neither short wavelength nor weak.) It is easy to write the unique R.W. radiation-dominated metric which has the correct singular points $t_\pm$, and which must be considered as a third candidate for the background. If one's inclination to call the anisotropy in the Taub space radiation is strong enough, this background is the one to be preferred.

Some of this ambiguity in defining a background is due to the fact that Taub space contains only a few wavelengths of radiation. The Isaacson scheme, on the other hand, assumes *short*-wavelength radiation. We would expect that in such cases averaging over only a few vector fields—instead of the infinite number we needed for Taub space—would lead to unique results. In particular, for the high-frequency radiation problem of Isaacson, one would want to use only a few averagings with a smoothing function that averaged over many wavelengths, but did nothing to change a scale the size of the background. (The problem in this respect with the Taub slices is that they are not much bigger than one wavelength of radiation.) The "few" vector fields are necessary because one wants to average over a volume, not just along a line; sufficient averagings are necessary to span such a small volume. The averaging Isaacson requires is some unspecified averaging process over such a volume. For the Taub slices, as we showed above, two averaging ideas lead to two different definitions of the energy density, and both were different from the massless radiation solutions.

We point out, however, that entirely naïve considerations can lead to the Lie-transport average. For

$$\xi_i \otimes \xi_i = \mathcal{N}_x \otimes \mathcal{N}_x + \mathcal{N}_y \otimes \mathcal{N}_y + \mathcal{N}_z \otimes \mathcal{N}_z.$$

One would expect that, in a sum of squares, where nothing picks out a direction (and nothing does on the sphere we average over), we would have

$$\tfrac{1}{3}\xi_i \otimes \xi_i = \langle \mathcal{N}_k \otimes \mathcal{N}_k \rangle,$$

where we sum on i but not on k, and the brackets indicate the average. This is the result which we found by the Lie-transport average, and in fact is so simple that one is inclined to demand it of all candidates for averaging schemes. Neither the equal-volume nor the radiation-dominated background has this simple averaging property.

Further, this Lie-transport scheme is well defined and can be used for local averaging, and makes use of vector fields defined in an invariant way by the metric itself. Other schemes, like the constant volume one defined for the Taub space, do not seem to be applicable at all to local averaging.

ACKNOWLEDGMENTS

I would like to thank the members of the Department of Applied Mathematics, Cambridge, for many interesting and stimulating discussions during this past year. I also thank Professor C. W. Misner for very helpful advice and comment. This work was begun while I held a National Science Foundation graduate fellowship.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 7 JULY 1968

Construction of Orthonormal Angular-Momentum Operators

P. L. CORIO
Mobil Research and Development Corporation, Research Department
Central Research Division, Princeton, New Jersey

(Received 17 October 1967)

The problem of constructing an orthonormal basis for the $(2j+1)^2$-dimensional space of angular-momentum operators is reduced to the problem of constructing a sequence of orthogonal polynomials of a discrete real variable. A number of properties of the polynomials are obtained, including a pure recurrence formula and several mixed recurrence relations. Three explicit representations of the polynomials are given, together with a table of orthonormal angular-momentum operators.

1. THE RATIONAL BASIS

The $(2j+1)^2$ irreducible tensor operators $T_q^{(k)}$ of maximum rank $2j$ define an orthogonal basis for the vector space $\mathfrak{C}$ of all linear transformations on the $(2j+1)$-dimensional space of angular-momentum states.[1] The elements of this tensor basis can be expressed in terms of $J_0 \equiv J_z$ and $J_\pm = J_x \pm iJ_y$, but the computational labor increases rapidly with increasing j, as evidenced by the limited tabulations available in the literature.[2,3]

A more systematic procedure for constructing an orthogonal basis for $\mathfrak{C}$ commences with the *rational basis* defined in the following lemma.

Lemma: The $(2j+1)^2$ linear operators $\{J_\pm^r J_0^s\}$ $(s = 0, 1, 2, \cdots, 2j-r;\ r = 0, 1, 2, \cdots, 2j)$ constitute a basis for $\mathfrak{C}$.

Proof: The elements of $\{J_\pm^r J_0^s\}$ can be exhibited as a triangular array of $2j+1$ rows labeled by the values of r. The rth row is

$$J_\pm^r,\ J_\pm^r J_0,\ J_\pm^r J_0^2,\ \cdots,\ J_\pm^r J_0^{2j-r}. \tag{1.1}$$

Since

$$J_\pm^r\,|j, m\rangle = r!\left\{\binom{j \mp m}{r}\binom{j \pm m + r}{r}\right\}^{\frac{1}{2}} |j, m \pm r\rangle, \tag{1.2}$$

it follows that (1) for $r \neq 0$ the $+$ and $-$ signs decompose (1.1) into two mutually orthogonal subsets; (2) every operator in the rth row is orthogonal to every operator in the r'th row $(r \neq r')$. The proof of the lemma is thus reduced to a verification of the linear independence of $\{J_+^r J_0^s\}$ and of $\{J_-^r J_0^s\}$.

Suppose that $2j - r + 1$ scalar coefficients c_s exist such that

$$J_-^r(c_0 I + c_1 J_0 + c_2 J_0^2 + \cdots + c_{2j-r} J_0^{2j-r}) = 0,$$

where I is the $(2j+1)$-dimensional identity operator. As r does not exceed $2j$, the vector $J_-^r\,|j, m\rangle$ does not vanish for $m = j, j-1, \cdots, -j+r$. Hence the assumed linear relation implies that

$$c_0 + c_1 m + \cdots + c_{2j-r} m^{2j-r} = 0,$$

for $m = j, j-1, \cdots, -j+r$. The determinant of

[1] If $\mathbf{A}$ and $\mathbf{B}$ are elements of $\mathfrak{C}$, their scalar product is defined as the trace of $\mathbf{A}^\dagger\mathbf{B}$; and $\mathbf{A}$ and $\mathbf{B}$ are orthogonal if $\mathrm{Tr}\,\mathbf{A}^\dagger\mathbf{B} = 0$.

[2] See, for example, K. W. H. Stevens, Proc. Phys. Soc. (London) **A65**, 209 (1952); G. F. Koster and H. Staatz, Phys. Rev. **113**, 445 (1959); E. Ambler, J. C. Eisenstein, and J. F. Schooley, J. Math. Phys. **3**, 118 (1962); H. Watanabe, *Operator Methods in Ligand Field Theory* (Prentice-Hall, Inc., Englewood Cliffs, N.J., 1966), pp. 151–152.

[3] The simultaneous eigenvectors of $\mathbf{J}^2$ and J_0 will be denoted $\{|j, m\rangle\}$, $m = j, j-1, \cdots, -j+1, -j$, and $\hbar$ will be taken as the unit of angular momentum.

where we sum on i but not on k, and the brackets indicate the average. This is the result which we found by the Lie-transport average, and in fact is so simple that one is inclined to demand it of all candidates for averaging schemes. Neither the equal-volume nor the radiation-dominated background has this simple averaging property.

Further, this Lie-transport scheme is well defined and can be used for local averaging, and makes use of vector fields defined in an invariant way by the metric itself. Other schemes, like the constant volume one defined for the Taub space, do not seem to be applicable at all to local averaging.

ACKNOWLEDGMENTS

I would like to thank the members of the Department of Applied Mathematics, Cambridge, for many interesting and stimulating discussions during this past year. I also thank Professor C. W. Misner for very helpful advice and comment. This work was begun while I held a National Science Foundation graduate fellowship.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 7 JULY 1968

Construction of Orthonormal Angular-Momentum Operators

P. L. CORIO
Mobil Research and Development Corporation, Research Department
Central Research Division, Princeton, New Jersey

(Received 17 October 1967)

The problem of constructing an orthonormal basis for the $(2j+1)^2$-dimensional space of angular-momentum operators is reduced to the problem of constructing a sequence of orthogonal polynomials of a discrete real variable. A number of properties of the polynomials are obtained, including a pure recurrence formula and several mixed recurrence relations. Three explicit representations of the polynomials are given, together with a table of orthonormal angular-momentum operators.

1. THE RATIONAL BASIS

The $(2j+1)^2$ irreducible tensor operators $T_q^{(k)}$ of maximum rank $2j$ define an orthogonal basis for the vector space $\mathfrak{C}$ of all linear transformations on the $(2j+1)$-dimensional space of angular-momentum states.[1] The elements of this tensor basis can be expressed in terms of $J_0 \equiv J_z$ and $J_\pm = J_x \pm iJ_y$, but the computational labor increases rapidly with increasing j, as evidenced by the limited tabulations available in the literature.[2,3]

A more systematic procedure for constructing an orthogonal basis for $\mathfrak{C}$ commences with the *rational basis* defined in the following lemma.

Lemma: The $(2j+1)^2$ linear operators $\{J_\pm^r J_0^s\}$ $(s = 0, 1, 2, \cdots, 2j-r;\ r = 0, 1, 2, \cdots, 2j)$ constitute a basis for $\mathfrak{C}$.

Proof: The elements of $\{J_\pm^r J_0^s\}$ can be exhibited as a triangular array of $2j+1$ rows labeled by the values of r. The rth row is

$$J_\pm^r,\ J_\pm^r J_0,\ J_\pm^r J_0^2,\ \cdots,\ J_\pm^r J_0^{2j-r}. \tag{1.1}$$

Since

$$J_\pm^r\,|j, m\rangle = r!\left\{\binom{j \mp m}{r}\binom{j \pm m + r}{r}\right\}^{\frac{1}{2}} |j, m \pm r\rangle, \tag{1.2}$$

it follows that (1) for $r \neq 0$ the $+$ and $-$ signs decompose (1.1) into two mutually orthogonal subsets; (2) every operator in the rth row is orthogonal to every operator in the r'th row $(r \neq r')$. The proof of the lemma is thus reduced to a verification of the linear independence of $\{J_+^r J_0^s\}$ and of $\{J_-^r J_0^s\}$.

Suppose that $2j - r + 1$ scalar coefficients c_s exist such that

$$J_-^r(c_0 I + c_1 J_0 + c_2 J_0^2 + \cdots + c_{2j-r} J_0^{2j-r}) = 0,$$

where I is the $(2j+1)$-dimensional identity operator. As r does not exceed $2j$, the vector $J_-^r\,|j, m\rangle$ does not vanish for $m = j, j-1, \cdots, -j+r$. Hence the assumed linear relation implies that

$$c_0 + c_1 m + \cdots + c_{2j-r} m^{2j-r} = 0,$$

for $m = j, j-1, \cdots, -j+r$. The determinant of

[1] If $\mathbf{A}$ and $\mathbf{B}$ are elements of $\mathfrak{C}$, their scalar product is defined as the trace of $\mathbf{A}^\dagger\mathbf{B}$; and $\mathbf{A}$ and $\mathbf{B}$ are orthogonal if $\mathrm{Tr}\,\mathbf{A}^\dagger\mathbf{B} = 0$.

[2] See, for example, K. W. H. Stevens, Proc. Phys. Soc. (London) **A65**, 209 (1952); G. F. Koster and H. Staatz, Phys. Rev. **113**, 445 (1959); E. Ambler, J. C. Eisenstein, and J. F. Schooley, J. Math. Phys. **3**, 118 (1962); H. Watanabe, *Operator Methods in Ligand Field Theory* (Prentice-Hall, Inc., Englewood Cliffs, N.J., 1966), pp. 151–152.

[3] The simultaneous eigenvectors of $\mathbf{J}^2$ and J_0 will be denoted $\{|j, m\rangle\}$, $m = j, j-1, \cdots, -j+1, -j$, and $\hbar$ will be taken as the unit of angular momentum.

this linear homogeneous system for the c_s does not vanish, so that every $c_s = 0$. A similar argument applied to $\{J_+^r J_0^s\}$ completes the proof.

When $r = 0$, the lemma asserts that the first $2j + 1$ integral powers of J_0 are linearly independent. Since the trace of any odd power of J_0 vanishes, the odd and even powers of J_0 decompose the row $r = 0$ into two mutually orthogonal subsets.

2. ORTHOGONALIZATION OF THE RATIONAL BASIS

The rational basis for $\mathcal{T}$ can be orthogonalized by the Gram–Schmidt process, but this procedure is not practicable for arbitrary j. It is evident, however, that the orthogonalization of the operators (1.1) yields polynomials of degrees $n = 0, 1, 2, \cdots, 2j - r$ in J_0, multiplied from the left by $J_\pm^r$. Let $Z_n^{(r)}(jl + J_0)$ denote the polynomial of degree n in J_0 obtained by orthogonalizing the operators of (1.1) with the + signs.[4] The orthogonality of $J_+^r Z_n^{(r)}(jl + J_0)$ and $J_+^r Z_{n'}^{(r)}(jl + J_0)$ demands that

$$(r!)^2 \sum_{x=0}^{2j-r} \binom{2j - x}{r} \binom{x + r}{r} Z_n^{(r)}(x) Z_{n'}^{(r)}(x) = 0, \qquad n' < n, \quad (2.1)$$

where

$$x = j + m \qquad (2.2a)$$

and

$$Z_n^{(r)}(x) = \langle j, m | \, Z_n^{(r)}(jl + J_0) \, | j, m \rangle. \qquad (2.2b)$$

Thus the $Z_n^{(r)}(x)$ are orthogonal with respect to the weight function

$$w_r(x) = \begin{cases} (r!)^2 \binom{2j - x}{r} \binom{x + r}{r}, \\ \quad \text{for} \qquad x = 0, 1, 2, \cdots, 2j - r, \\ 0 \quad \text{for any} \quad x \neq 0, 1, 2, \cdots, 2j - r. \end{cases} \quad (2.3)$$

In the special case $r = 0$, $w_0(x)$ has jumps of one unit at $x = 0, 1, 2, \cdots, 2j$, so that the $Z_n^{(0)}(x)$ are equivalent to Tchebychef's orthogonal polynomials of a discrete real variable.[5]

The orthogonality conditions determine $Z_n^{(r)}(x)$ up to an arbitrary multiplicative constant $N(n, r, j)$; the polynomial $Z_n^{(r)}(x)$ is uniquely determined by further stipulating that

$$\text{Re}\,[N(n, r, j)] > 0, \quad \text{Im}\,[N(n, r, j)] = 0; \quad (2.4a)$$

$$\sum_{x=0}^{2j-r} w_r(x) [Z_n^{(r)}(x)]^2 = 1. \qquad (2.4b)$$

An explicit formula for $Z_n^{(r)}(x)$ may be obtained from a study of the function[6]

$$G_n^{(r)}(x) = (-1)^n [(n + r)!]^2 N(n, r, j) \times \binom{x + r}{n + r} \binom{2j + n - x}{n + r}, \quad (2.5)$$

whose μth difference is[7]

$$\Delta^\mu G_n^{(r)}(x) = (-1)^n [(n + r)!]^2 N(n, r, j) \times \sum_{k=0}^{\mu} (-1)^{\mu - k} \binom{\mu}{k} \binom{x + r}{n + r - k} \binom{2j + n - \mu - x}{n + r - \mu + k}. \quad (2.6)$$

The nth difference of $G_n^{(r)}(x)$ divided by $w_r(x)$ is a polynomial of degree n in x. In fact,

$$\Delta^n G_n^{(r)}(x) = w_r(x) Z_n^{(r)}(x). \qquad (2.7)$$

For upon inserting (2.7) into (2.1) and performing n partial summations, then, since

$$\Delta^{n-\nu} G_n^{(r)}(x + \nu - 1)$$

vanishes at $x = 0$ and at $x = 2j - r + 1$ for $\nu = 1, 2, \cdots, n$, one obtains

$$\sum_{x=0}^{2j-r} w_r(x) Z_n^{(r)}(x) Z_{n'}^{(r)}(x) = (-1)^n \sum_{x=0}^{2j-r} G_n^{(r)}(x + n) \Delta^n Z_{n'}^{(r)}(x). \quad (2.8)$$

If $n' < n$, $\Delta^n Z_{n'}^{(r)}(x) = 0$, so that the $Z_n^{(r)}(x)$ defined by (2.7) satisfy (2.1).

The preceding results yield the explicit representation

$$Z_n^{(r)}(x) = N(n, r, j) \times \sum_{\nu=0}^{n} \frac{(-1)^\nu n! [(n + r)!]^2}{(r + \nu)! (n + r - \nu)!} \binom{2j - x - r}{\nu} \binom{x}{n - \nu}. \quad (2.9)$$

Another expression for ${}_n^{(r)}Z(x)$ may be obtained by expanding the right-hand member of (2.9) in a binomial series about $x = 0$:

$$Z_n^{(r)}(x) = \frac{N(n, r, j) n! [(n + r)!]^2}{(n + 2r)!} \times \sum_{\nu=0}^{n} (-1)^{n+\nu} \binom{2j - r - \nu}{n - \nu} \binom{n + 2r + \nu}{n + r} \binom{x}{\nu}. \quad (2.10)$$

[4] The polynomials associated with J_-^r will be obtained from a symmetry property of the $Z_n^{(r)}(jl + J_0)$ in Sec. 3.

[5] (a) C. Jordan, *Calculus of Finite Differences* (Chelsea Publ. Co., New York, 1950), p. 437; (b) G. Szegö, *Orthogonal Polynomials* (American Mathematical Society, New York, 1959), p. 33.

[6] For the theory of polynomials of a discrete variable, orthogonal with respect to nontrivial weight functions, see, for example, E. H. Hildebrandt, Ann. Math. Statistics **2**, 379 (1931); M. Weber and A. Erdélyi, Am. Math. Monthly **59**, 163 (1952).

[7] The symbol Δ, as used here, denotes the finite-difference operator for unit increment in x: $\Delta f(x) = f(x + 1) - f(x)$, $\Delta^{n+1} f(x) = \Delta[\Delta^n f(x)]$.

A third expression for $Z_n^{(r)}(x)$ is given by

$$Z_n^{(r)}(x) = \frac{N(n,r,j)n!\,[(n+r)!]^2}{(n+2r)!} \times \sum_{v=0}^{n}(-1)^v\binom{2j-r-v}{n-v}\binom{n+2r+v}{n+r}\binom{2j-r-x}{v}, \tag{2.11}$$

which follows from (2.10) and Eq. (3.1) of the following section.

When $n = n'$, the right-hand member of (2.8) reduces to

$$(-1)^n\Delta^n Z_n^{(r)}(x)\sum_{x=0}^{2j-r} G_n^{(r)}(x+n) = [(n+r)!]^2 N(n,r,j)\Delta^n Z_n^{(r)}(x)\binom{2j+n+r+1}{2n+2v+1}. \tag{2.12}$$

From (2.10),

$$\Delta^n Z_n^{(r)}(x) = (n!)^2 N(n,r,j)\binom{2n+2r}{r}, \tag{2.13}$$

so that

$$N(n,r,j) = \frac{1}{n!\,(n+r)!} \times \left\{\binom{2n+2r}{n}\binom{2j+n+r+1}{2n+2r+1}\right\}^{-\frac{1}{2}}. \tag{2.14}$$

3. PROPERTIES OF THE POLYNOMIALS

The $Z_n^{(r)}(x)$ possess an important symmetry property which follows immediately from (2.9) upon replacing x with $2j - r - x$:

$$Z_n^{(r)}(2j-r-x) = (-1)^n Z_n^{(r)}(x). \tag{3.1}$$

This equation can be used to relate the polynomials in J_0 obtained by orthogonalizing the operators in (1.1) with the $-$ signs to the $Z_n^{(r)}(jI + J_0)$. Let

$$P_n^{(r)}[(j-r)I + J_0]$$

denote the polynomial of degree n in J_0 associated with J_-^r, and let $y = j - r + m$. A simple calculation shows that the $P_n^{(r)}(y)$ satisfy orthogonality conditions identical in form to (2.1); hence $P_n^{(r)}(y)$ must be proportional to $Z_n^{(r)}(y)$. Adopting the normalization (2.4) for the $P_n^{(r)}(y)$, and observing that $Z_n^{(r)}(y)$ satisfies (3.1) with x replaced by y, one concludes that

$$P_n^{(r)}(y) = Z_n^{(r)}(y) = (-1)^n Z_n^{(r)}(j-m). \tag{3.2}$$

Thus the operators obtained by orthonormalization of the operators (2.1) are

$$U_{\pm r}^{(n)} = (\pm 1)^n J_\pm^r Z_n^{(r)}(jI \pm J_0), \quad n = 0,1,2,\cdots,2j-r. \tag{3.3}$$

The complete orthonormal basis for $\mathfrak{C}$ follows upon setting $r = 0, 1, \cdots, 2j$. When $r = 0$, the operators in (3.3) with the $+$ signs are related to those with the $-$ signs by (3.1); in this case only the $2j + 1$ operators with the $+$ (or $-$) signs are taken.

The Hermitian adjoints of the $U_{\pm r}^{(n)}$ are given by

$$U_{\pm r}^{(n)\dagger} = U_{\mp r}^{(n)}, \tag{3.4}$$

which follows from (3.1), (3.2), and the identity

$$J_0^s J_\pm^r = J_\pm^r (J_0 \pm rI)^s. \tag{3.5}$$

The matrix elements of $U_{\pm r}^{(n)}$ relative to $\{|j,m\rangle\}$ are

$$\langle j,m'|\,U_{\pm r}^{(n)}\,|j,m\rangle = (\pm 1)^n r!\left\{\binom{j\mp m}{r}\binom{j\pm m+r}{r}\right\}^{\frac{1}{2}} Z_n^{(r)}(j\pm m) \times \delta_{m',m\pm r} \tag{3.6}$$

with the same reservation for the $\pm$ signs noted above when $r = 0$. Equations (2.9), (2.10), and (2.11) may be combined with (3.6) to give three equivalent forms for the matrix elements.

The $Z_n^{(r)}(x)$ satisfy the following mixed recurrence relations:

$$(2n+2r+1)R(n,r,j)Z_{n+1}^{(r)}(x) + (n+2r+1)(2j+n-r-2x)Z_n^{(r)}(x) + 2[(x+r+1)(2j-r-x)]\Delta Z_n^{(r)}(x) = 0, \tag{3.7}$$

$$(2n+2r+1)R(n,r,j)\Delta Z_{n+1}^{(r)}(x) - 2(n+1)(n+2r+1)Z_n^{(r)}(x) + (n+1)(2j-n-3r-2-2x)\Delta Z_n^{(r)}(x) = 0, \tag{3.8}$$

where

$$R(n,r,j) = \left[\frac{(n+1)(n+2r+1)(2j+n+r+2)\times(2j-n-r)}{(2n+2r+1)(2n+2r+3)}\right]^{\frac{1}{2}}. \tag{3.9}$$

It is easily verified that the explicit representations obtained for the $Z_n^{(r)}(x)$ in the preceding section satisfy (3.7) and (3.8).

By differencing (3.7) and eliminating $\Delta Z_{n+1}^{(r)}(x)$ with the help of (3.8), one finds that the $Z_n^{(r)}(x)$ satisfy the linear homogeneous second-order difference equation

$$(x+r+2)(2j-r-1-x)\Delta^2 Z_n^{(r)}(x) + [(r+1)(2j-r-2-2x) + n(n+2r+1)]\Delta Z_n^{(r)}(x) + n(n+2r+1)Z_n^{(r)}(x) = 0. \tag{3.10}$$

The $Z_n^{(r)}(x)$ also satisfy the following pure recurrence

relation:

$$R(n, r, j)Z^{(r)}_{n+1}(x) + (2j - r - 2x)Z^{(r)}_n(x) + R(n-1, r, j)Z^{(r)}_{n-1}(x) = 0, \quad (3.11)$$

which is valid for all nonnegative integers n, provided that $Z^{(r)}_{-1}(x) \equiv 0$. The derivation of (3.11) is given in the Appendix.

Finally, by differencing (2.10), one obtains the useful formula

$$\Delta Z^{(r)}_n(x) = [n(n + 2r + 1)]^{\frac{1}{2}} Z^{(r+1)}_{n-1}(x). \quad (3.12)$$

The preceding results may be used to derive the following operator identities:

$$2J_0 U^{(n)}_{\pm r} = R(n, r, j)U^{(n+1)}_{\pm r} \pm rU^{(n)}_{\pm r} + R(n-1, r, j)U^{(n-1)}_{\pm r}; \quad (3.13)$$

$$2U^{(n)}_{\pm r}J_{\pm} = \left[\frac{(n + 2r + 2)(n + 2r + 1)(2j + n + r + 2)(2j - n - r)}{(2n + 2r + 1)(2n + 2r + 3)}\right]^{\frac{1}{2}} U^{(n)}_{\pm(r+1)} \pm [n(n + 2r + 1)]^{\frac{1}{2}} U^{(n-1)}_{\pm(r+1)} - \left[\frac{n(n-1)(2j + n + r + 1)(2j - n - r + 1)}{(2n + 2r - 1)(2n + 2r + 1)}\right]^{\frac{1}{2}} U^{(n-2)}_{\pm(r+1)}, \quad r = 0, 1, 2, \cdots, 2j; \quad (3.14)$$

$$2U^{(n)}_{\pm r}J_{\mp} = -\left[\frac{(n + 1)(n + 2)(2j + n + r + 2)(2j - n - r)}{(2n + 2r + 1)(2n + 2r + 3)}\right]^{\frac{1}{2}} U^{(n+2)}_{\pm(r-1)} \pm [(n + 1)(n + 2r)]^{\frac{1}{2}} U^{(n+1)}_{\pm(r-1)} + \left[\frac{(n + 2r)(n + 2r - 1)(2j + n + r + 1)(2j - n - r + 1)}{(2n + 2r - 1)(2n + 2r + 1)}\right]^{\frac{1}{2}} U^{(n)}_{\pm(r-1)}, \quad r = 1, 2; \cdots, 2j. \quad (3.15)$$

These results, together with (3.4), yield the commutation relations

$$[J_0, U^{(n)}_{\pm r}] = \pm rU^{(n)}_{\pm r}, \quad (3.16)$$

$$[U^{(n)}_{\pm r}, J_{\pm}] = \pm[n(n + 2r + 1)]^{\frac{1}{2}} U^{(n-1)}_{\pm(r+1)}, \quad (3.17)$$

$$[U^{(n)}_{\pm r}, J_{\mp}] = \pm[(n + 1)(n + 2r)]^{\frac{1}{2}} U^{(n+1)}_{\pm(r-1)}. \quad (3.18)$$

Equations (3.16)–(3.18) reveal the relation of the $U^{(n)}_{\pm r}$ to the irreducible tensor operators $T^{(k)}_q$:

$$n + r = k, \quad r = |q|, \quad (3.19)$$

and $T^{(k)}_q$ is proportional to $U^{(n)}_{+r}$ or $U^{(n)}_{-r}$, accordingly as $q > 0$ or $q < 0$. The phase of the proportionality constant is $(\mp 1)^r$, according to Racah's definition[8] of $T^{(k)}_q$.

The polynomial operators $Z^{(r)}_n(jI + J_0)$ may be obtained by replacing x in (2.10) with $X = jI + J_0$. This procedure gives the $Z^{(r)}_n(jI + J_0)$ in terms of "factorial" operators

$$X^{(0)} = I,$$

$$X^{(\lambda)} = X[X - I] \cdots [X - (\lambda - 1)I], \quad \lambda = 1, 2, \cdots.$$

In applications, however, an expansion in powers of J_0 is often preferable. For this purpose, it is especially convenient to use the pure recurrence formula (3.11) and the initial operator

$$Z^{(r)}_0(jI + J_0) = \frac{1}{r!}\left\{\binom{2j + r + 1}{2r + 1}\right\}^{-\frac{1}{2}} I. \quad (3.20)$$

This procedure was used to construct the $U^{(n)}_{\pm r}$ given in Table I, where

$$\kappa = j(j + 1), \quad (3.21)$$

$$K_t = [2j(2j - t)(2j - t - 1)]^{\frac{1}{2}}\left\{\prod_{\nu=0}^{t+1}(4j^2 - \nu^2)\right\}^{-\frac{1}{2}}. \quad (3.22)$$

[8] G. Racah, Phys. Rev. 62, 438 (1942).

Table I[a,b]. The orthonormal angular-momentum operators $U^{(n)}_{\pm r}$.

$U^{(0)}_0$ $K_0 I$

$U^{(1)}_0$ $2(3)^{\frac{1}{2}}K_1 J_0$

$U^{(2)}_0$ $2(5)^{\frac{1}{2}}K_2\{3J_0^2 - \kappa I\}$

$U^{(3)}_0$ $4(7)^{\frac{1}{2}}K_3\{5J_0^3 - (3\kappa - 1)J_0\}$

$U^{(4)}_0$ $6K_4\{35J_0^4 - 5(6\kappa - 5)J_0^2 + 3\kappa(\kappa - 2)I\}$

$U^{(5)}_0$ $4(11)^{\frac{1}{2}}K_5\{63J_0^5 - 35(2\kappa - 3)J_0^3 + (15\kappa^2 - 50\kappa + 12)J_0\}$

$U^{(6)}_0$ $4(13)^{\frac{1}{2}}K_6\{231J_0^6 - 105(3\kappa - 7)J_0^4 + 21(5\kappa^2 - 25\kappa + 14)J_0^2 - 5\kappa(\kappa - 6)(\kappa - 2)I\}$

$U^{(7)}_0$ $8(15)^{\frac{1}{2}}K_7\{429J_0^7 - 231(3\kappa - 10)J_0^5 + 21(15\kappa^2 - 105\kappa + 101)J_0^3 - (35\kappa^3 - 385\kappa^2 + 882\kappa - 180)J_0\}$

$U^{(8)}_0$ $2(17)^{\frac{1}{2}}K_8\{6435J_0^8 - 6006(2\kappa - 9)J_0^6 + 1155(6\kappa^2 - 56\kappa + 81)J_0^4 - 6(210\kappa^3 - 3045\kappa^2 + 9898\kappa - 4566)J_0^2 + 35\kappa(\kappa - 12)(\kappa - 6)(\kappa - 2)I\}$

. . .

$U^{(0)}_{\pm 1}$ $(6)^{\frac{1}{2}}K_1 J_{\pm}$

$U^{(1)}_{\pm 1}$ $(30)^{\frac{1}{2}}K_2 J_{\pm}\{2J_0 \pm I\}$

$U^{(2)}_{\pm 1}$ $2(21)^{\frac{1}{2}}K_3 J_{\pm}\{5J_0^2 \pm 5J_0 - (\kappa - 2)I\}$

$U^{(3)}_{\pm 1}$ $6(5)^{\frac{1}{2}}K_4 J_{\pm}\{14J_0^3 \pm 21J_0^2 - (6\kappa - 19)J_0 \mp 3(\kappa - 2)I\}$

$U^{(4)}_{\pm 1}$ $2(330)^{\frac{1}{2}}K_5 J_{\pm}\{21J_0^4 \pm 42J_0^3 - 7(2\kappa - 9)J_0^2 \mp 14(\kappa - 3)J_0 + (\kappa - 6)(\kappa - 2)I\}$

$U^{(5)}_{\pm 1}$ $2(546)^{\frac{1}{2}}K_6 J_{\pm}\{66J_0^5 \pm 165J_0^4 - 60(\kappa - 6)J_0^3 \mp 15(6\kappa - 25)J_0^2 + 2(5\kappa^2 - 55\kappa + 117)J_0 \pm 5(\kappa - 6)(\kappa - 2)I\}$

$U^{(6)}_{\pm 1}$ $2(210)^{\frac{1}{2}}K_7 J_{\pm}\{429J_0^6 \pm 1287J_0^5 - 165(3\kappa - 23)J_0^4 \mp 495(2\kappa - 11)J_0^3 + 3(45\kappa^2 - 645\kappa + 1832)J_0^2 \pm 9(15\kappa^2 - 160\kappa + 332)J_0 - 5(\kappa - 12)(\kappa - 6) \times (\kappa - 2)I\}$

$U^{(7)}_{\pm 1}$ $6(34)^{\frac{1}{2}}K_8 J_{\pm}\{1430J_0^7 \pm 5005J_0^6 - 1001(2\kappa - 19)J_0^5 \mp 5005(\kappa - 7)J_0^4 + 385(2\kappa^2 - 36\kappa + 131)J_0^3 \pm 385(3\kappa^2 - 41\kappa + 112)J_0^2 - (70\kappa^3 - 1785\kappa^2 + 12488\kappa - 22356)J_0 \mp 35(\kappa - 12)(\kappa - 6)(\kappa - 2)I\}$

. . .

$U^{(0)}_{\pm 2}$ $(30)^{\frac{1}{2}}K_2J^2_{\pm}$
$U^{(1)}_{\pm 2}$ $2(210)^{\frac{1}{2}}K_3J\{J_0 \pm I\}$
$U^{(2)}_{\pm 2}$ $6(10)^{\frac{1}{2}}K_4J^2_{\pm}\{7J_0^2 \pm 14J_0 - (\kappa - 9)I\}$
$U^{(3)}_{\pm 2}$ $4(330)^{\frac{1}{2}}K_5J^2_{\pm}\{3J_0^3 \pm 9J_0^2 - (\kappa - 12)J_0 \mp (\kappa - 6)I\}$
$U^{(4)}_{\pm}$ $2(1365)^{\frac{1}{2}}K_6J^2_{\pm}\{33J_0^4 \pm 132J_0^3 - 3(6\kappa - 91)J_0^2 \mp 6(6\kappa - 47)J_0 + (\kappa - 20)(\kappa - 6)I\}$
$U^{(5)}_{\pm 2}$ $12(35)^{\frac{1}{2}}K_7J^2_{\pm}\{143J_0^5 \pm 715J_0^4 - 55(2\kappa - 37)J_0^3 \mp 55(6\kappa - 59)J_0^2 + (15\kappa^2 - 490\kappa + 2862)J_0 \pm 15(\kappa - 12)(\kappa - 6)I\}$
$U^{(6)}_{\pm 2}$ $12(595)^{\frac{1}{2}}K_8J^2_{\pm}\{143J_0^6 \pm 858J_0^5 - 143(\kappa - 22)J_0^4 \mp 572(\kappa - 12)J_0^3 + 11(3\kappa^2 - 119\kappa + 853)J_0^2 \pm 22(3\kappa^2 - 67\kappa + 333)J_0 - (\kappa - 35)(\kappa - 12)(\kappa - 6)I\}$

. . .

$U^{(0)}_{\pm 3}$ $2(35)^{\frac{1}{2}}K_3J^3_{\pm}$
$U^{(1)}_{\pm 3}$ $15(35)^{\frac{1}{2}}K_4J^3_{\pm}\{2J_0 \pm 3 \cdot I\}$
$U^{(2)}_{\pm 3}$ $2(385)^{\frac{1}{2}}K_5J^3_{\pm}\{9J_0^2 \pm 27J_0 - (\kappa - 24)I\}$
$U^{(3)}_{\pm 3}$ $2(1365)^{\frac{1}{2}}K_6J^3_{\pm}\{22J_0^3 \pm 99J_0^2 - (6\kappa - 179)J_0 \mp 3(3\kappa - 40)I\}$
$U^{(4}_{\pm 3}$ $6(70)^{\frac{1}{2}}K_7J^3_{\pm}\{143J_0^4 \pm 858J_0^3 - 11(6\kappa - 215)J_0^2 \mp 66(3\kappa - 49)J_0 + 3(\kappa^2 - 62\kappa + 600)I\}$
$U^{(5)}_{\pm 3}$ $2(39270)^{\frac{1}{2}}K_8J^3_{\pm}\{78J_0^5 \pm 585J_0^4 - 52(\kappa - 42)J_0^3 \mp 117(2\kappa - 39)J_0^2 + 2(3\kappa^2 - 223\kappa + 2595)J_0 \pm 3(3\kappa^2 - 106\kappa + 840)I\}$

. . .

$U^{(0)}_{\pm 4}$ $3(70)^{\frac{1}{2}}K_4J^4_{\pm}$
$U^{(1)}_{\pm 4}$ $6(770)^{\frac{1}{2}}K_5J^4_{\pm}\{J_0 \pm 2 \cdot I\}$
$U^{(2)}_{\pm 4}$ $6(182)^{\frac{1}{2}}K_6J^4_{\pm}\{11J_0^2 \pm 44J_0 - (\kappa - 50)I\}$
$U^{(3)}_{\pm 4}$ $12(770)^{\frac{1}{2}}K_7J^4_{\pm}\{13J_0^3 \pm 78J_0^2 - (3\kappa - 179)J_0 \mp 6(\kappa - 25)I\}$
$U^{(4)}_{\pm 4}$ $6(2618)^{\frac{1}{2}}K_8J^4_{\pm}\{65J_0^4 \pm 520J_0^3 - 13(2\kappa - 139)J_0^2 \mp 52(2\kappa - 59)J_0 + (\kappa^2 - 122\kappa + 2100)I\}$

. . .

$U^{(0)}_{\pm 5}$ $6(77)^{\frac{1}{2}}K_5J^5_{\pm}$
$U^{(1)}_{\pm 5}$ $6(1001)^{\frac{1}{2}}K_6J^5_{\pm}\{2J_0 \pm 5 \cdot I\}$
$U^{(2)}_{\pm 5}$ $6(770)^{\frac{1}{2}}K_7J^5_{\pm}\{13J_0^2 \pm 65J_0 - (\kappa - 90)I\}$
$U^{(3)}_{\pm 5}$ $6(34034)^{\frac{1}{2}}K_8J^5_{\pm}\{10J_0^3 \pm 75J_0^2 - (2\kappa - 209)J_0 \mp 5(\kappa - 42)I\}$

. . .

$U^{(0)}_{\pm 6}$ $2(3003)^{\frac{1}{2}}K_6J^6_{\pm}$
$U^{(1)}_{\pm 6}$ $12(5005)^{\frac{1}{2}}K_7J^6_{\pm}\{J_0 \pm 3 \cdot I\}$
$U^{(2)}_{\pm 6}$ $4(7293)^{\frac{1}{2}}K_8J_{+}\{15J_0^2 \pm 90J_0 - (\kappa - 147)I\}$

. . .

$U^{(0)}_{\pm 7}$ $6(1430)^{\frac{1}{2}}K_7J^7_{\pm}$
$U^{(1)}_{\pm 7}$ $6(24310)^{\frac{1}{2}}K_8J_{\pm}\{2J_0 \pm 7 \cdot I\}$

. . .

$U^{(0)}_{\pm 8}$ $3(24310)^{\frac{1}{2}}K_8J^8_{\pm}$

[a] The first eight K_t's are given in Table II.
[b] $\kappa = j(j+1)$.

TABLE II. The K_t of Eq. (3.22).

t	K_t
0	$(2j+1)^{-\frac{1}{2}}$
1	$\frac{1}{2}[j(j+1)(2j+1)]^{-\frac{1}{2}}$
2	$\frac{1}{2}[j(j+1)(2j+3)(4j^2-1)]^{-\frac{1}{2}}$
3	$\frac{1}{4}[j(j+2)(2j+3)(j^2-1)(4j^2-1)]^{-\frac{1}{2}}$
4	$\frac{1}{4}[j(j+2)(2j+5)(j^2-1)(4j^2-1)(4j^2-9)]^{-\frac{1}{2}}$
5	$\frac{1}{6}[j(j+3)(2j+5)(j^2-1)(j^2-4)(4j^2-1)(4j^2-9)]^{-\frac{1}{2}}$
6	$\frac{1}{6}[j(j+3)(2j+7)(j^2-1)(j^2-4)(4j^2-1)(4j^2-9) \times (4j^2-25)]^{-\frac{1}{2}}$
7	$\frac{1}{16}[j(j+4)(2j+7)(j^2-1)(j^2-4)(j^2-9)(4j^2-1) \times (4j^2-9)(4j^2-25)]^{-\frac{1}{2}}$
8	$\frac{1}{16}[j(j+4)(2j+9)(j^2-1)(j^2-4)(j^2-9)(4j^2-1) \times (4j^2-9)(4j^2-25)(4j^2-49)]^{-\frac{1}{2}}$

The first eight K_t's are given in Table II.

ACKNOWLEDGMENTS

The author wishes to express his gratitude to Dr. R. H. Albert for several informative discussions, and to B. Wichnoski for his assistance in the preparation of Table I.

APPENDIX

The orthogonality of the $Z_n^{(r)}(x)$ implies a pure recurrence relation of the form

$$xZ_n^{(r)}(x) = c_{n+1}Z_{n+1}^{(r)}(x) + c_nZ_n^{(r)}(x) + c_{n-1}Z_{n-1}^{(r)}(x), \tag{A1}$$

where the c's are independent of x. Inserting the values $x_1 = 0$ and $x_2 = 2j - r$, noting that

$$Z_n^{(r)}(2j - r) = (-1)^nZ_n^{(r)}(0),$$

and that zero is not a root of any $Z_n^{(r)}(x)$, one obtains

$$c_n = \tfrac{1}{2}(2j - r). \tag{A2}$$

The coefficient of $Z_{n+1}^{(r)}(x)$ may be determined by multiplying both members of (A1) by $w_r(x)Z_{n+1}^{(r)}(x) = \Delta^{n+1}G_{n+1}^{(r)}(x)$ and summing from $x = 0$ to $x = 2j - r$:

$$c_{n+1} = \sum_{x=0}^{2j-r} xZ_n^{(r)}(x)\Delta^{n+1}G_{n+1}^{(r)}(x)$$
$$= (-1)^{n+1}\Delta^{n+1}\{xZ_n^{(r)}(x)\}\sum_{x=0}^{2j-r} G_{n+1}^{(r)}(x + n + 1),$$

where the last form is obtained by performing $n + 1$ partial summations. Since

$$\Delta^{n+1}[xZ_n^{(r)}(x)] = (n + 1)\Delta^nZ_n^{(r)}(x),$$

it follows that

$$c_{n+1} = -\frac{1}{2}\left\{\frac{(n + 1)(n + 2r + 1) \times (2j + n + r + 2)(2j - n - r)}{(2n + 2r + 1)(2n + 2r + 3)}\right\}^{\frac{1}{2}}. \tag{A3}$$

The coefficient c_{n-1} is obtained by replacing n in the right member of (A3) with $n - 1$.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 7 JULY 1968

Solutions of Einstein's Equations for a Fluid Which Exhibit Local Rotational Symmetry

J. M. STEWART AND G. F. R. ELLIS
Department of Applied Mathematics and Theoretical Physics, University of Cambridge, England

(Received 14 August 1967)

All solutions of Einstein's equations for pressure-free matter which exhibit local rotational symmetry were classified in an earlier paper by one of us. This paper extends the earlier theory to the case of a general fluid, with an electromagnetic field possibly present. A classification of these solutions for a perfect fluid is given, and assuming a physically reasonable equation of state, some exact solutions of cosmological interest are obtained. Finally, the difficulties encountered when extending the treatment to a general fluid are discussed; the same general classification can be made.

1. INTRODUCTION

In several recent papers,[1] one of us has classified all the solutions of Einstein's field equations for pressure-free matter which have the property of local rotational symmetry (hereafter denoted by L.R.S.). This paper extends the classification to include all the L.R.S. solutions for a perfect fluid plus electromagnetic field, exhibiting some new anisotropic solutions for such a fluid. Finally, the difficulties involved in extending the scheme to more general fluids are considered.

As far as possible, the notation and methods used are identical with those of Ref. 1. Most proofs are only outlined. For further details, one should consult the fuller treatment in Ref. 1.

For the units used, $c = 1$, $G = \frac{1}{8}\pi$; Roman letters range over 0 to 3, Greek letters over 1 to 3; $i, j, k, \cdots$, denote coordinate indices and $a, b, c, \cdots$, denote tetrad indices. The metric tensor g_{ij} has signature $(- + + +)$. Covariant differentiation in the v^i direction is $\nabla_v = {}_{;k}v^k$; partial differentiation in the v^i direction is ${}_{,k}v^k$. Round and square brackets denote, respectively, symmetrization and antisymmetrization of indices; η^{ijkl} is a totally skew tensor, $\eta^{ijkl} = \eta^{[ijkl]}$.

The fluid-flow vector u^i $(u^i u_i = -1)$ determines the tensor $h_{ij} = g_{ij} + u_i u_j$ which projects into the instantaneous rest space of an observer comoving with the fluid. The acceleration of the fluid is $\dot{u}_i = u_{i;j}u^j$ (∇_u is denoted by $\cdot$). The velocity gradient $u_{i;j}$ can be written

$$u_{i;j} = \omega_{ij} + \sigma_{ij} + \tfrac{1}{3}\theta h_{ij} - \dot{u}_i u_j, \tag{1.1}$$

where

$$\omega_{ij} = u_{[i;j]} + \dot{u}_{[i}u_{j]} \tag{1.2}$$

is the vorticity tensor and

$$\sigma_{ij} = u_{(i;j)} + \dot{u}_{(i}u_{j)} - \tfrac{1}{3}\theta h_{ij} \tag{1.3}$$

is the trace-free shear tensor. The expansion is $\theta = u^i{}_{;i}$. The vorticity vector is

$$\omega^i = \tfrac{1}{2}\eta^{ijkl}u_j\omega_{kl} \quad (\omega^i u_i = 0). \tag{1.4}$$

The scalars $\dot{u} = (\dot{u}_i\dot{u}^i)^{\frac{1}{2}}$, $\omega = (\frac{1}{2}\omega_{ij}\omega^{ij})^{\frac{1}{2}}$, $\sigma = (\frac{1}{2}\sigma_{ij}\sigma^{ij})^{\frac{1}{2}}$ vanish if and only if the corresponding tensors vanish.

We consider a viscous fluid. The energy–momentum tensor can be written in the covariant form[2]

$$T_{ij} = \mu u_i u_j + p h_{ij} + 2u_{(i}q_{j)} + \pi_{ij}, \tag{1.5}$$

where u_i is the average 4-velocity of matter. In the rest frame of u_i, μ is the energy density, p is the scalar pressure, q_i is the energy flux density $(q_i u^i = 0)$, and π_{ij} is the anisotropic pressure tensor (e.g., viscosity). A noninteracting electromagnetic field can be included by adding to T_{ij} the extra term

$$\tau_{ij} = \tfrac{1}{4}g_{ij}(F_{kl}F^{kl}) - F_{ik}F_j{}^k, \quad \tau_i{}^i = 0,$$

where F_{ik} is the electromagnetic field tensor which satisfies Maxwell's equations. For a charged fluid,

$$F^{ij}{}_{;j} = \epsilon u^i,$$
$$F_{ij;k} = 0.$$

ϵ is the charge density (possibly zero). The conservation equations (contracted Bianchi identities) reduce to[2]

$$\dot{\mu} + (\mu + p)\theta + \pi_{ij}\sigma^{ij} + q^i{}_{;i} + q_i\dot{u}^i = 0, \tag{1.6}$$

$$(\mu + p)\dot{u}_i + h_i{}^j(q_j + p_{,j} + \pi_{jk}{}^{;k}) + (\omega_{ij} + \sigma_{ij})q^j + \tfrac{4}{3}\theta q_i + \epsilon F_{ij}u^j = 0. \tag{1.7}$$

At each point, introduce a local orthonormal tetrad of vectors $\{e_a\}$. This tetrad is, in general, not parallely transferred from point to point, and in all subsequent work we consider only a neighborhood of a point. e_a represents a directional derivative denoted by ∂_a. Introduce a local system of coordinates

[1] G. F. R. Ellis, Ph.D. thesis, University of Cambridge (1964); J. Math. Phys. **8**, 1171 (1967).

[2] J. Ehlers, Abh. Math. Nat. Kl. Mainz Akad. Wiss. u. Litt. Nr. 11 (1961).

$\{x^i\}$. The components of e_a referred to the basis $\partial/\partial x^i$ are $e_a{}^i$ given by

$$e_a = e_a{}^i \partial/\partial x^i, \quad \det(e_a{}^i) \neq 0,$$

so

$$\partial_a f = e_a f = f_{,i} e_a{}^i.$$

The components $e^a{}_i$ of $\partial/\partial x^i$ referred to basis $\{e_a\}$ are given by

$$\partial/\partial x^i = e^a{}_i e_a.$$

Since the tetrad is orthonormal, the tetrad components of the metric tensor are

$$g_{ab} = e_a \cdot e_b = \text{diag}(-1, 1, 1, 1),$$

where

$$e_a \cdot e_b = e_a{}^i e_b{}^j g_{ij} = e_a{}^i e_{bi}.$$

Then $g_a{}^b = \delta_a{}^b = e_a{}^i e^b{}_i$ and so $e^a{}_i$, $e_b{}^j$ are inverse matrices. The tensor η^{abcd} is normalized by the condition that $\eta^{0123} = 1$ holds for the tetrad components. The Ricci rotation coefficients are

$$\Gamma_{abc} = e_a \cdot \nabla_b e_c = e_a{}^i e_{ci;j} e_b{}^j.$$

Since $g_{ij;k} = 0$, it follows that $\Gamma_{abc} + \Gamma_{cba} = 0$. The Lie derivative of e_b with respect to e_a is $[e_a, e_b]$, where

$$[e_a, e_b]f = \partial_a(\partial_b f) - \partial_b(\partial_a f).$$

It is a vector with tetrad components $\gamma^c{}_{ab}$:

$$[e_a, e_b] = \gamma^c{}_{ab} e_c.$$

The $\gamma^c{}_{ab}$ and $\Gamma^c{}_{ab}$ are linearly dependent quantities:

$$\gamma^c{}_{ab} = \Gamma^c{}_{ab} - \Gamma^c{}_{ba}, \tag{1.8}$$

$$\Gamma_{abc} = \tfrac{1}{2}(\gamma_{abc} + \gamma_{cab} - \gamma_{bca}). \tag{1.9}$$

The Riemann-tensor and Ricci-tensor tetrad components can easily be found from the definitions

$$v^b{}_{;cd} - v^b{}_{;dc} = R^b{}_{ecd} v^e \quad \text{and} \quad R_{bd} = R^c{}_{bcd}. \tag{1.10}$$

The antisymmetry property of the curvature tensor $R^i{}_{[jkl]} = 0$ is equivalent to the Jacobi identity

$$[e_b, [e_c, e_d]] + [e_c, [e_d, e_b]] + [e_d, [e_b, e_c]] = 0,$$

which reduces to

$$\partial_{[d}\gamma^f{}_{bc]} + \gamma^f{}_{g[d}\gamma^g{}_{bc]} = 0. \tag{1.11}$$

Equations (1.11) are labeled $\binom{f}{bcd}$, and the field equations

$$R_{bd} - \tfrac{1}{2}R g_{bd} = -T_{bd} + \Lambda g_{bd}$$

are labeled (bd). The differential equations $\binom{a}{bcd}$ and (gf) are written in tetrad form and solved as far as possible for $\gamma^a{}_{bc}$, using (1.9). The $e_a{}^i$ are then determined, using (1.8). By inversion, $e^a{}_i$ can be found, and finally, the g_{ij} can be found from

$$g_{ij} = g_{ab} e^a{}_i e^b{}_j. \tag{1.12}$$

2. L.R.S. AND THE CHOICE OF THE TETRAD AND COORDINATE SYSTEMS

Space-time is said to be *locally rotationally symmetric* (L.R.S.) in the neighborhood $N(P_0)$ of a point P_0 if at each point P in $N(P_0)$ there exists a non-discrete subgroup g of the Lorentz group in the tangent space T_P which leaves invariant u^a, the curvature tensor, and their derivatives up to third order. Since μ, p, q^a, π^{ab} are defined uniquely by u^a and R_{ab}, they and their derivatives are also invariant under g. Thus g operates in the subspace of T_P orthogonal to u^a and so g is a one- or three-dimensional group of rotations in T_P.

If g is three-dimensional, then $q^a = \pi^{ab} = 0$, $\omega = \sigma = \dot{u} = 0$, and we have a Robertson–Walker model. These are included in the solutions for a one-dimensional group g discussed below.

Suppose g is one-dimensional. Choose e_0 to lie along u^a, and e_1 to lie along an axis of symmetry. L.R.S. implies that all covariantly defined spacelike vectors are parallel to e_1, and the spacelike parts of all covariantly defined tensors have diagonal form with (22) and (33) terms equal. Clearly the derivatives in the e_2, e_3 directions of covariantly defined scalars must be zero. Thus $\omega^a = \omega\delta_1{}^a$, $\dot{u}^a = \dot{u}\delta_1{}^a$, $q^a = q\delta_1{}^a$, $\tau_{ab} = \text{diag}(\tau, -\tau, \tau, \tau)$ [where $\tau = \frac{1}{2}(E^2 + B^2)$, and $(0, E, 0, 0)$, $(0, B, 0, 0)$ are the electric and magnetic fields in the rest frame of u^a, respectively],

$$\pi_{ab} = \text{diag}(0, 2\pi, -\pi, -\pi)$$

and

$$\sigma_{ab} + \tfrac{1}{3}\theta h_{ab} = \text{diag}(0, \alpha, \beta, \beta),$$

where α and β are the expansions of the fluid along and perpendicular to e_1. Then Eqs. (1.1)–(1.5) show that

$$\gamma^0{}_{01} = \dot{u}, \quad \gamma^0{}_{02} = \gamma^0{}_{03} = 0, \quad \gamma^1{}_{01} = -\alpha,$$
$$\gamma^2{}_{02} = \gamma^3{}_{03} = -\beta, \quad \gamma^0{}_{32} = 2\omega, \quad \gamma^0{}_{31} = \gamma^0{}_{12} = 0.$$

$\dot{e}_1 = \nabla_0 e_1$ is covariantly defined and so $0 = e_2 \cdot \dot{e}_1 = e_3 \cdot \dot{e}_1$. By orthonormality, $e_1 \cdot \dot{e}_2 = e_1 \cdot \dot{e}_3 = 0$. Therefore, $\gamma^1{}_{02} = \gamma^1{}_{03} = \gamma^2{}_{01} = \gamma^3{}_{01} = 0$. Similarly, $e_2 \cdot \nabla_1 e_1 = e_3 \cdot \nabla_1 e_1 = 0 = \gamma^1{}_{21} = \gamma^1{}_{31}$.

The integral lines of e_1 must have no shear in the rest space of u^a, and so $\gamma^2{}_{13} = 0$, $\gamma^2{}_{12} = \gamma^3{}_{13} = a$, say. If $\omega \neq 0$, $\omega^{-2}\eta^{abcd}u_a\omega_b\omega_{c;d}$ is an invariant which is up to a factor $\gamma^1{}_{32} = k$, say. (If $\omega = 0$, the corresponding expression for e_1 is an invariant.)

The following scalars are covariantly defined and so have zero derivatives in the e_2, e_3 directions: $\mu, p, q, \pi, \tau, \omega, \dot{u}, \alpha, \beta, a, k$. The Jacobi identities and

field equations with these restrictions on the γ's are given in Appendix A. The choice of the e_2, e_3 axes is arbitrary by a rotation within the e_2, e_3 2-plane in each tangent space. This remaining tetrad freedom can be used to simplify the remaining $\gamma^a{}_{bc}$.

Theorem 1: If, near a point P, a fluid-filled space-time has L.R.S., then an orthonormal tetrad can be chosen near P so that the $\gamma^a{}_{bc}$ are given by the following array:

$$\begin{aligned}
[e_0, e_1] &= \dot{u}e_0 - \alpha e_1, \\
[e_0, e_2] &= -\beta e_2, \\
[e_0, e_3] &= -\beta e_3, \\
[e_2, e_3] &= -2\omega e_0 - k e_1 + s e_3, \\
[e_3, e_1] &= -a e_3, \\
[e_1, e_2] &= a e_2,
\end{aligned} \tag{2.1}$$

where $s = \gamma^3{}_{23}$ and $\partial_3 s = 0$ everywhere.

For an outline of the proof see Appendix B.

Clearly from (2.1),

$$0 = [e_2, e_3]\omega = -2\omega\partial_0\omega - k\partial_1\omega. \tag{2.2}$$

From $\binom{0}{023}$, $\partial_0\omega = -2\omega\beta - \frac{1}{2}k\dot{u}$, and from $\binom{0}{123}$, $\partial_1\omega = \omega\dot{u} + 2a\omega$. Then, substituting in (2.2), if $\omega \neq 0$, $2\omega\beta = ka$. If $\omega = 0$, then $0 = [e_2, e_3]k = -k\partial_1 k$ and so $\partial_1 k = 0$. From $\binom{1}{123}$, $-\partial_1 k = -2ak$ so that $ak = 0$ in this case. Thus for any ω

$$2\omega\beta = ak. \tag{2.3}$$

Also,

$$0 = [e_2, e_3]\beta = -2\omega\partial_0\beta - k\partial_1\beta. \tag{2.4}$$

$\partial_0\beta$ can be found from (00) + (11) − (22) − (33) and $\partial_1\beta$ from (01). Using (2.3) and (2.4),

$$\omega(\Lambda - p - 2\pi + \tau - \beta^2 - 2\alpha\beta - 2a\dot{u} - \omega^2 + a^2 + \tfrac{1}{4}k^2 - \partial_2 s + s^2) - \tfrac{1}{2}kq = 0. \tag{2.5}$$

Similarly,

$$[e_2, e_3]a = 0 = -2\omega\partial_0 a - k\partial_1 a. \tag{2.6}$$

We can find $\partial_0 a$ from $\binom{2}{012}$ and (01), $\partial_1 a$ from (00) + (11) + (22) + (33). Then

$$k(\Lambda + \mu + \tau - \beta^2 - 2\alpha\beta - 2a\dot{u} - \omega^2 + a^2 + \tfrac{1}{4}k^2 - \partial_2 s + s^2) - 2\omega q = 0.$$

Using (2.5),

$$\omega k(\mu + p + 2\pi) = q(2\omega^2 + \tfrac{1}{2}k^2). \tag{2.7}$$

Theorem 2: If space-time contains a perfect fluid for which $p + \mu \neq 0$ and an electromagnetic field, there are three disjoint and exhaustive classes of solutions exhibiting L.R.S.:

Case I: $\omega \neq 0$, $k = \alpha = \beta = 0$;
Case II: $\omega = k = 0$;
Case III: $\omega = 0$, $k \neq 0$, $\dot{u} = a = \partial_1 k = \partial_1\alpha = \partial_1\beta = 0$.

Proof: For a perfect fluid $\pi = q = 0$. Then $\omega k = 0$ from (2.7). Case II is when $\omega = k = 0$. Case I is when $\omega \neq 0$, $k = 0$. From (2.3) $\beta = 0$, and from $\binom{1}{123}$, $\alpha = 0$. Case III is when $\omega = 0$, $k \neq 0$. From (2.3) $a = 0$. From $\binom{0}{023}$, $\binom{3}{013}$, and $\binom{1}{123}$, $\dot{u} = \partial_1\beta = \partial_1 k = 0$. Also $0 = [e_2, e_3]\alpha = -k\partial_1\alpha$ which implies $\partial_1\alpha = 0$.

Theorem 3: If space-time containing a perfect fluid and an electromagnetic field has L.R.S. near a point P, then the coordinate freedom can be used to set the metric in the form

$$\begin{aligned}
ds^2 = {} & [-(dx^0)^2/F^2(x^0, x^1)] + X^2(x^0, x^1)(dx^1)^2 \\
& + Y^2(x^0, x^1)[(dx^2)^2 + t^2(x^2)(dx^3)^2] \\
& + [y(x^2)/F^2(x^0, x^1)][2\,dx^0 - y(x^2)\,dx^3]\,dx^3 \\
& - X^2(x^0, x^1)h(x^2)[2\,dx^1 - h(x^2)\,dx^3]\,dx^3,
\end{aligned} \tag{2.8}$$

where $t(x^2)$, $y(x^2)$, $h(x^2)$ are functions defined by equations given in Appendix B. The following specializations are possible:

Class I: $X \equiv 1$, $Y = Y(x^1)$, $h \equiv 0$;
Class II: $h \equiv y \equiv 0$;
Class III: $F \equiv 1$, $X = X(x^0)$, $Y = Y(x^0)$, $y \equiv 0$.

For an outline of the proof see Appendix B.

3. CLASSIFICATION OF EXACT SOLUTIONS FOR A PERFECT FLUID WHICH EXHIBIT L.R.S.

Each of the classes defined by Theorem 2 is investigated in detail. Where possible, the equations are integrated, but in most cases it is necessary to assume an equation of state, and the choice of such an equation is deferred until the next section. In general, p, μ are independent thermodynamical quantities and the equation of state can be used to define the temperature $T = T(p, \mu)$. However, it is sometimes convenient to also consider the case of a *barotropic* fluid for which $\mu = \mu(p)$.

Case I: $\omega \neq 0$, $k = 0$. It is convenient to divide Case I into four subcases accordingly as a is zero or nonzero and as $\dot{u}$ is zero or nonzero.

Case Ia: $a = \dot{u} = 0$. Differentiating the equations $\partial_1 a = \partial_1\dot{u} = 0$ gives $\epsilon E = 0$. Then, p, μ, r, ω, τ are constants with

$$\begin{aligned}
p &= \Lambda - \tfrac{1}{2}r, \\
\mu &= 2\omega^2 + \tfrac{1}{2}r - \Lambda, \\
\tau &= \omega^2 + \tfrac{1}{2}r.
\end{aligned} \tag{3.1}$$

If $E = 0$, then

$$\epsilon, H \text{ are constants with } \epsilon = 2\omega H.$$

If $E \neq 0$, then $\epsilon = 0$, and

$$E = (2\tau)^{\frac{1}{2}} \cos 2\omega x^1,$$
$$H = (2\tau)^{\frac{1}{2}} \sin 2\omega x^1.$$

The solution (3.1) is a generalization of the Godel solution.[3] Unlike the dust case, there is a nontrivial solution when $\Lambda = \tau = 0$. The space–time is invariant under a G_5 of motions multiply transitive on space–time. The electromagnetic fields have sinusoidal form if $E \neq 0$, but all other physical quantities are homogeneous; if $E = 0$, the electromagnetic quantities are also homogeneous.

Case Ib: $a = 0$, $\dot{u} \neq 0$. For a barotropic fluid, if $\epsilon E = 0$, then $d\mu/dp = -3$, and the solution is

$$\begin{aligned} &r, \tau \text{ are const,} \\ &\omega = \omega_0/F(x^1), \quad \omega_0 \text{ a nonzero const,} \\ &p = \Lambda - r + \tau - {\omega_0}^2/F^2, \\ &\mu = -\Lambda + r - \tau + 3{\omega_0}^2/F^2, \\ &x^1 = \int [cF^4 + (2\tau - r)F^2 - {\omega_0}^2]^{-\frac{1}{2}}\, dF, \quad c \text{ const,} \end{aligned} \tag{3.2}$$

which can be integrated once the signs of the constants have been specified.

If $E \neq 0$, then $\epsilon = 0$, and $E = (2\tau)^{\frac{1}{2}} \cos\theta$, $H = (2\tau)^{\frac{1}{2}} \sin\theta$, where

$$\theta(x^1) = 2\int^{x^1} \omega(u)\, du.$$

If $E = 0$, $H = (2\tau)^{\frac{1}{2}}$, $\epsilon = 2\omega H$. Since ω varies, the electromagnetic waves now have variable frequency.

If $\epsilon E \neq 0$, an equation of state with $d\mu/dp \neq -1$ must be given ($d\mu/dp = -1$ implies $\dot{u} = 0$). Values of p, $\dot{u}$, E, and H may be specified at an initial value of x^1. Then, provided

$$\epsilon = \frac{(3 + d\mu/dp)(p + \mu)\dot{u}}{(1 + d\mu/dp)E},$$

the equations are integrable.

For a nonbarotropic fluid, ϵ can be specified as an arbitrary function of x^1. Then if ω, p, μ, $\dot{u}$, E, and H are specified at an arbitrary value of x^1, the equations are integrable.

Case Ic: $\dot{u} = 0$, $a \neq 0$. For a barotropic fluid, if $\epsilon E = 0$, the solution is

$$\begin{aligned} &r = K/Y^2(x^1), \quad K \text{ const,} \\ &\tau = \tau_0/Y^4(x^1), \quad \tau_0 \text{ const,} \\ &\omega = \omega_0/Y^2(x^1), \end{aligned} \tag{3.3}$$

where

$$2{\omega_0}^2 = \tau_0, \quad p = \text{const}, \quad \mu = 2\Lambda - 3p,$$

$$x^1 = \int [(p - \Lambda)Z^2 + KZ - {\omega_0}^2]^{-\frac{1}{2}}\, dZ,$$

where

$$Z = Y^2.$$

If $E = 0$, then $H = (2\tau_0)^{\frac{1}{2}}/Y^2$ and $\epsilon = 2\omega H$. If $E \neq 0$, then $\epsilon = 0$, and

$$E = [(2\tau_0)^{\frac{1}{2}} \cos\theta]/Y^2, \quad H = [(2\tau_0)^{\frac{1}{2}} \sin\theta]/Y^2,$$

where

$$\theta(x^1) = 2\int^{x^1} \omega(u)\, du.$$

Note that this solution is not possible unless electromagnetic fields are present. The integration can be performed once the constants are given.

If $\epsilon E \neq 0$, an equation of state with $d\mu/dp \neq -5$ must be specified, together with values of p, E, H, and a for an initial value of x^1. If $\epsilon = 8a(2\omega^2 - \tau_0)/[(5 + d\mu/dp)E]$, then the equations can be integrated.

For a nonbarotropic fluid, ϵ can be specified as an arbitrary function of x^1. Then if ω, p, μ, a, E, and H are given at an initial value of x^1, the equations are integrable.

Case Id: $\dot{u} \neq 0$, $a \neq 0$. In the barotropic case, if ϵ is specified everywhere together with an equation of state, the equations can be integrated, once initial values of E, H, a, $\dot{u}$, p, and ω are given. In the nonbarotropic case, ϵ and μ must be specified everywhere, together with initial values of the other quantities.

Solutions of Cases Ib, c, d are invariant under a G_4 of motions, multiply transitive in the 3-surfaces $x^1 = \text{const}$.

Case II: $\omega = k = 0$. Applying the commutation relations (2.1) to E gives $\dot{\epsilon} = -\epsilon(\alpha + 2\beta)$, an equation which expresses the conservation of charge.

If $\alpha = \beta$, then the solution is shearfree, and since it is already irrotational, it is a Robertson–Walker model, if $\dot{u} = 0$. (See, e.g., Ref. 2.) Again it is convenient to classify the solutions accordingly as a, $\dot{u}$ are zero or nonzero.

Case IIa: $a = 0$. $a = 0$ implies $\beta\dot{u} = 0$, and $\beta = 0$ then implies $p + \mu = 0$, which is unreasonable for a fluid. Thus

$$a = 0 \quad \text{leads to} \quad \dot{u} = 0.$$

Differentiating $\partial_1 a = 0$ gives an expression for $\partial_1 \alpha$. The commutation relations are satisfied by all the

[3] K. Godel, Rev. Mod. Phys. 21, 447 (1949).

quantities except α, ϵ, and p. In the nonbarotropic case, ϵ must be specified on an initial hypersurface, $x^1 = \text{const}$, and E, H, α, β, μ, and r must be given for some value of x^1 in this hypersurface. Also p must be specified for all x^0 along the world line of a particle, before the equations can be integrated. Even if $\epsilon E = 0$, the solutions need not be homogeneous, but they are homogeneous if either of $\partial_1\alpha$, $\partial_1\mu$ is zero.

In the barotropic case, p will only satisfy the commutation relation if either $\epsilon E = 0$, or there is a complicated restriction on the equation of state. (This latter possibility has not been investigated.) If $\epsilon E = 0$, then $\epsilon = 0$ and the solutions are homogeneous. These models include the generalization of the Einstein–de Sitter universe, and the Friedmann models, and have been considered by a variety of authors as possible cosmological models. (See the next section for details.)

There are no static models in this case, since $\beta \neq 0$. However, there are solutions with nonzero shear, where $\alpha = 0$, which implies $\epsilon E = 0$. They have the equation of state given by

$$\frac{d\mu}{dp} = \frac{p + \mu}{3p - \mu - 2\Lambda}. \tag{3.4}$$

The general solution of this equation is

$$\mu - p = -2\Lambda + p_0 \exp\,[-2(\mu + p + 2\Lambda)/(p - \Lambda)], \quad p_0 \text{ const},$$

which has no great physical significance, but there is a singular solution given by $\mu = p - 2\Lambda$. Then,

$$\begin{aligned} &p = p_0/Y^4(x^0) + \Lambda, \quad p_0 \text{ const},\\ &r = K/Y^2, \quad K \text{ const},\\ &\tau = 0, \end{aligned} \tag{3.5}$$

$$Y = \begin{cases} (p_0/K - x^{0^2})^{\frac{1}{2}}, & K > 0, \text{ which implies } p_0 > 0,\\ (x^{0^2} - p_0/K)^{\frac{1}{2}}, & K < 0,\\ (p_0)^{\frac{1}{4}}(x^0)^{\frac{1}{2}}, & K = 0,\ p_0 > 0. \end{cases}$$

Note that no electromagnetic fields are possible in this model.

The homogeneous models are invariant under a G_4 of motions multiply transitive in the 3-surfaces $x^0 = \text{const}$. If models exist with $\epsilon E \neq 0$, then the group is a G_3 multiply transitive in the 2-surfaces $x^0 = \text{const}$, $x^1 = \text{const}$.

Case IIb: $\dot{u} = 0$, $a \neq 0$. In the nonbarotropic case, ϵ, μ, and α must be specified on an initial hypersurface, with $\alpha \neq 0$, and p must be specified along a single world line. When the other quantities are given at a single point in the initial hypersurface, the equations can be integrated.

In the barotropic case, if $\epsilon E = 0$, then applying the commutation relations to p and α gives

$$\tau = \alpha - \beta = 0. \tag{3.6}$$

This is a shearfree case with $\dot{u} = 0$, and so gives a Robertson–Walker model. When an equation of state has been specified, it can be integrated. If $\epsilon E \neq 0$, there is a complicated restriction on the equation of state. This has not been investigated.

There is a static solution for this case, given by $\epsilon E = 0 = \alpha = \beta$. Then $\tau = 0$.

Also μ, p are constant and are related by

$$\begin{aligned} &2\Lambda = 3p + \mu,\\ &r = K/Y^2(x^1), \quad K \text{ const}, K > 0,\\ &Y(x^1) = [K/(\Lambda - p)]^{\frac{1}{2}} \sin\,[(\Lambda - p)x^1]. \end{aligned} \tag{3.7}$$

No electromagnetic fields are possible in this case, which is a generalization of the Einstein static solution for dust. If $\epsilon E \neq 0$, a static solution can be found once an equation of state has been specified, with $d\mu/dp \neq -5$. (See Sec. 4.)

For these models, space-time is invariant under a G_4 of motions multiply transitive in the 3-surface $x^1 = \text{const}$. There is no shear case with $\alpha = 0$, $\beta \neq 0$.

Case IIc: $\dot{u} \neq 0$, $a \neq 0$. In the nonbarotropic case, ϵ, μ, and α must be specified on an initial hypersurface, and p must be specified along a single world line. When the other quantities are given at a single point in the initial hypersurface, the equations can be integrated. In the barotropic case, the equations can be integrated once an equation of state has been given, α, $\dot{u}$, and ϵ specified as functions of x^1 for an initial value of x^0, and the other quantities given at the same x^0 and some fixed x^1. Space-time is invariant under a G_3 of motions multiply transitive in the 2-surfaces $x^0 = \text{const}$, $x^1 = \text{const}$. In the stationary case, the initial, values are subject to

$$\Lambda - p + \tau + a^2 - 2a\dot{u} - r = 0. \tag{3.8}$$

If $\alpha = 0$, $\beta \neq 0$, there is a complicated restriction on the equation of state, which has not been investigated.

Case III: $\omega = 0$, $k \neq 0$. It is convenient to classify the solutions accordingly as β is zero or nonzero. Applying (2.1) to E gives $0 = [e_2, e_3]E = k\partial_1 E$ and so $\epsilon = 0$.

Case IIIa: $\beta = 0$. Then $p + \mu = \frac{1}{2}k^2$ and $\alpha(3 + dp/d\mu) = 0$. If $\alpha = 0$, then p, μ, k, r, and τ are

constants, and

$$\begin{aligned} E &= (2\tau)^{\frac{1}{2}} \cos kx^0, \\ H &= (2\tau)^{\frac{1}{2}} \sin kx^0, \\ \Lambda &= \tau + \tfrac{1}{2}\mu + \tfrac{3}{2}p, \\ k^2 &= 2(p + \mu), \\ r &= p + \mu + 2\tau. \end{aligned} \tag{3.9}$$

There is no solution for $r \leq 0$. This is a generalization of the Einstein static solution and resembles the Godel solution, Case Ia. In fact, for the cases of empty space or space containing an electromagnetic field, the Case III solutions are analytic continuations of the Case I solution across a null hypersurface. If dust or a fluid is present, the solutions are very similar but are not equivalent. If $\alpha \neq 0$, then, by suitable normalization,

$$\begin{aligned} &Y \equiv 1, \\ &\tau, r \text{ are const}, \\ &k = X(x^0), \\ &\mu = c - \tfrac{1}{4}X^2, \\ &p = -c + \tfrac{3}{4}X^2, \qquad c \text{ const} \\ &E = (2\tau)^{\frac{1}{2}} \cos\theta, \\ &H = (2\tau)^{\frac{1}{2}} \sin\theta, \end{aligned} \tag{3.10}$$

where

$$\theta(x^0) = \int^{x^0} X(u)\, du,$$

$$x^0 = \int [b + (\Lambda - \tau + c)X^2 - \tfrac{1}{2}X^4]^{-\frac{1}{2}}\, dX, \quad b \text{ const}.$$

The integration can be performed once the constants have been specified. This solution is analogous to Case Ib.

Case IIIb: $\beta \neq 0$. Now $k \neq 0$ implies $\alpha \neq 0$. Once an equation of state is given (or p is specified along a world line in the nonbarotropic case) together with initial values satisfying the identity, the system is integrable.

All Class III solutions except the first are invariant under a G_4 of motions multiply transitive in the 3-surface $x^0 = \text{const}$. For Case IIIa, the group is a G_5.

We now have the following theorem.

Theorem 4: (a). If a space-time containing a perfect fluid satisfies L.R.S., there is a local group of motions G_r in each open neighborhood W of a point P_0 which is multiply transitive on some q-surface $(q < r)$ through each point P in W. Then at each point P in W there is a nondiscrete isotropy group; the subgroup g of the Lorentz group in the tangent space, which leaves invariant the curvature tensor and its derivatives up to third order, in fact leaves invariant the derivatives of all orders (see, e.g., Ref. 4).

(b). If a perfect fluid is L.R.S. in an open set U with $\omega \neq 0$ in U, then $\theta = \sigma = 0$ in U and u is collinear with a Killing vector. Nonrigid-body motion is only possible when the vorticity is zero.

4. SOLUTIONS OF THE FIELD EQUATIONS WITH A GIVEN EQUATION OF STATE

In this section, ρ is the rest-mass density of the fluid. For a barotropic fluid one has $p = p(\rho)$, but very little is known about physically reasonable functional forms of p at the high densities and temperatures in the initial stages of a "hot big bang" cosmology. However, over a limited range of ρ a "γ law" is applicable:

$$p \propto \rho^{\gamma}, \quad \gamma = \text{const}. \tag{4.1}$$

For an adiabatic process the relativistic first law of thermodynamics[5] is

$$\frac{d\mu}{\mu + p} = \frac{d\rho}{\rho} \tag{4.2}$$

or

$$p\frac{d\mu}{dp} - \frac{\mu}{\gamma} = \frac{p}{\gamma},$$

from which

$$\mu = \begin{cases} cp^{1/\gamma} + p/(\gamma - 1), & \gamma \neq 1, \\ cp + p\log p, & \gamma = 1, \quad c = \text{const}. \end{cases} \tag{4.3}$$

Tooper[6] has considered the case $\mu = \rho + p/(\gamma - 1)$, and others[7] have considered the case $c = 0$. For simplicity, this latter equation of state is used:

$$p = (\gamma - 1)\mu. \tag{4.4}$$

If the system is to be consistent with causality and mechanically stable, then[7] $1 \leq \gamma \leq 2$. We now reconsider those cases for which an equation of state (4.4) can be used, and which are not dust models, i.e., $\gamma > 1$.

---

[4] R. P. Kerr, J. Math. & Mech. **12**, 33 (1963).

[5] L. D. Landau, and E. M. Lifshitz, *Fluid Mechanics* (Pergamon Press, Inc., New York, 1959), p. 499.

[6] R. F. Tooper, Astrophys. J. **142**, 1541 (1965).

[7] H. Bondi, Proc. Roy. Soc. (London) **A282**, 303 (1964); B. K. Harrison, K. S. Thorne, M. Wakano, and J. A. Wheeler, *Gravitation Theory and Gravitational Collapse* (The University of Chicago Press, Chicago, 1965), Chap. 9.

Case Ib: The solution is

$$\begin{aligned} &r = \Lambda, \\ &p = p_0/F^2(x^1), \quad p_0 \text{ const}, \\ &\omega = \omega_0/F(x^1); \quad {\omega_0}^2 = p_0/2(\gamma - 1), \\ &\tau = \tau_0/F^2(x^1), \quad \tau_0 = (3\gamma - 2)p_0/2(\gamma - 1), \\ &x^1 = \int (cF^4 - \Lambda F^2 + p_0)^{-\frac{1}{2}}\, dF, \quad c \text{ const}, \end{aligned} \tag{4.5}$$

plus equations determining ϵ, E, H.

Case Ic:

$$\begin{aligned} &\omega = \omega_0/Y^2(x^1), \quad \omega_0 \text{ const}, \\ &r = K/Y^2(x^1), \quad K \text{ const}, \\ &\tau = cY^{-4(3\gamma-2)/(5\gamma-4)} + 2{\omega_0}^2/Y^4, \quad c \text{ const}, \\ &\mu = 2(\Lambda - cY^{-4(3\gamma-2)/(5\gamma-4)})/(3\gamma - 2), \\ &x^1 = \int \Big[KY^2 + \Big(\frac{c}{3\gamma - 2}\Big) Y^{8(\gamma-1)/(5\gamma-4)} \\ &\qquad - {\omega_0}^2 - \Big(\frac{\gamma}{3\gamma - 2}\Big) Y^4 \Big]^{-\frac{1}{2}} dY, \end{aligned} \tag{4.6}$$

plus equations determining ϵ, E, H. There is an explicit solution if $\gamma = \frac{4}{3}$, $c = 0$.

Case Id: No fluid solutions have been found, even with $\epsilon E = 0$. For dust solutions see Ref. 1.

There are solutions for a vacuum electromagnetic field ($p = \mu = 0$) which are equivalent to the corresponding solutions for Taub space, one of which has already been found by Brill[8] (see Case IIIc).

Case IIa: Homogeneous solutions have been found for $\Lambda = \tau = r = 0$. They are:

1. $\alpha = \beta$; isotropic case

$$\begin{aligned} X(x^0) = Y(x^0) &= (x^0)^{\frac{2}{3}}, \\ \mu = p/(\gamma - 1) &= \frac{4}{3\gamma^2 (x^0)^2}. \end{aligned} \tag{4.7}$$

This is the obvious generalization of the Einstein–de Sitter universe.

2. $\alpha \neq \beta$, $\gamma \neq 2$.

$$\begin{aligned} x^0 &= \int (Y^{3(2-\gamma)/2} + c)^{(\gamma-1)/(2-\gamma)} Y^{\frac{1}{2}}\, dY, \\ &\qquad c = \text{const}, \quad c \neq 0, \\ X &= (Y^{3(2-\gamma)/2} + c)^{1/(2-\gamma)} Y^{-\frac{1}{2}}, \\ p/(\gamma - 1) &= \mu = 3/(XY^2)^{\gamma}. \end{aligned} \tag{4.8}$$

The solution for $\gamma = \frac{4}{3}$ (which can be obtained explicitly) has been obtained by Doroshkevich[9] and Thorne.[10] As $x^0 \to \infty$, $X \to Y$, and the solution tends asymptotically to the isotropic case. For a full discussion of these solutions and their asymptotic properties, see Doroshkevich,[9] Thorne,[10] and a further paper by the present authors (to appear).

3. $\alpha \neq \beta$, $\gamma = 2$.

$$\begin{aligned} &X(x^0) = (x^0)^{1/(1+2\lambda)} \quad \lambda \text{ const}, \quad \lambda > 0, \quad \text{or } \lambda < -2, \\ &Y(x^0) = X^{\lambda}, \\ &\mu = \frac{(2 + \lambda)}{(1 + 2\lambda)^2 (x^0)^2}. \end{aligned} \tag{4.9}$$

Kantowski[11] has obtained the solutions for $K = \pm 1$ with $\gamma = \frac{4}{3}, 2$. They are:

1. $\gamma = \frac{4}{3}$.

$$\begin{aligned} X &= \beta^2(x^0)/\nu(x^0), \\ Y &= \nu^2/\beta, \\ \mu &= a^2/3\nu^4, \\ \nu &= (c\beta + \beta^2 - K\beta^4/a^2)^{\frac{1}{2}}, \\ x^0 &= \frac{3}{a} \int \nu\, d\beta, \quad a, c \text{ const}, \quad a > 1. \end{aligned} \tag{4.10}$$

2. $\gamma = 2$, $K = 1$.

$$\begin{aligned} x^0 &= a^2 c \int [X \cosh (a \log X)]^{-2}\, dX, \\ Y &= ac[X \cosh (a \log X)]^{-1}, \\ \mu &= (X/ac)^2 \cosh^4 (a \log X)(1 - a^{-2}). \end{aligned} \tag{4.11}$$

For $K = -1$, replace "cosh" by "sinh". a, c are constants, $a > 1$.

It is of interest to consider a homogeneous electromagnetic field by itself, $p = \mu = 0$. There is also a Class III one detailed later. The nonstationary model in this class is

$$\begin{aligned} &E = E_0/Y^2(x^0), \quad E_0, H_0 \text{ const}, \\ &H = H_0/Y^2(x^0), \\ &\tau = \tau_0/Y^4(x^0), \quad \tau_0 = \tfrac{1}{2}(E_0^2 + H_0^2), \\ &x^0 = \int (\tfrac{1}{3}\Lambda Y^4 - KY^2 + cY - \tau_0)^{-\frac{1}{2}} Y dY, \quad c = \text{const}, \\ &X(x^0) = (\tfrac{1}{3}\Lambda Y^2 - K + c/Y - \tau_0/Y^2)^{\frac{1}{2}}. \end{aligned} \tag{4.12}$$

The integration can be performed analytically if $c = 0$, or $\Lambda = 0$, and Λ, K, c, τ_0 are given. These solutions are in fact the nonstationary regions of the

[8] D. R. Brill, Phys. Rev. **133**, B845 (1964).

[9] A. S. Doroshkevich, Astrophysics, **1**, 138 (1965).

[10] K. S. Thorne, Astrophys. J. **148**, 51 (1967).

[11] R. Kantowski, Ph.D. thesis, Texas University (1966).

Reissner–Nordstrom type solutions.[12] We obtain the Robinson–Bertotti electromagnetic solution[13] when $\alpha = \beta = 0$.

Case IIb: The Friedmann solutions ($\epsilon E = 0$) are well known.[14] No solutions have been found with $\epsilon E \neq 0$, except in the static case. For this case, a solution is only possible for $\Lambda > 0$, $K < 0$, $p > 0$:

$$\begin{aligned}
&\tau = \tau_0 / Y^{4(3\gamma-2)/(5\gamma-4)}, \quad \tau_0 \text{ const},\\
&r = K/Y^2,\\
&p = \frac{2\Lambda(\gamma-1)}{(3\gamma-2)} - 2\left(\frac{\gamma-1}{3\gamma-2}\right)\tau_0 Y^{-4(3\gamma-2)/(5\gamma-4)},\\
&x^1 = \int \{-K - [\gamma\Lambda/(3\gamma-2)]Y^2 \\
&\qquad - [\gamma/(3\gamma-2)]\tau_0 Y^{-2\gamma/(5\gamma-4)}\}^{-\frac{1}{2}}\, dY,\\
&\qquad\qquad \text{plus equations for } E, H.
\end{aligned} \tag{4.13}$$

An exact integration is only possible in the dust case. No solution is possible with this equation of state in the shear case $\alpha = 0$, $\beta \neq 0$.

Case IIc: This is the most general case, which in the isotropic case has been extensively studied in the context of stellar evolution (e.g., Ref. 15). Most work has been numerical, and so a simple but unrealistic barotropic equation of state is no advantage. Authors who have determined an analytical solution have usually imposed some further conditions, e.g., extra conditions between the unknown functions, and then have deduced the equation of state, which is usually quite complicated. Our method of integration requires the specification of α, ϵ, $\dot{u}$ as functions of x^1 at some initial x^0, and p as a function of x^0, for some fixed x^1. These functions can be quite arbitrary. Thus the integration of this case is not possible until these extra conditions have been imposed.

If $p = \mu = 0$, we obtain the Reissner–Nordstrom family of solutions,[12] and if in addition $\tau = 0$, the "Schwarzschildian" solutions.[16]

Case IIIb: No analytic solutions have been found for a perfect fluid, or even for dust, see Ref. 1. For a vacuum field one obtains the Taub–NUT spaces. Brill[8] has found the solution for a sourceless electromagnetic field, $p = \mu = 0 = \Lambda$, with $K = 1$:

$$\begin{aligned}
&Y = (Y_0^2 + \theta^2)^{\frac{1}{2}},\\
&X = \left(\frac{X_0 + c_0\theta}{Y_0^2 + \theta^2} - 4Y_0^2\right)^{\frac{1}{2}}, \quad X_0, Y_0, c_0 \text{ const},\\
&r = Y^{-2},\\
&k = XY^{-2},\\
&\tau = \tau_0/Y^4, \quad \tau_0 \text{ const},\\
&\theta(x^0) = \int^{x^0} (X/2Y_0)\, dx^0,\\
&E = 2(2\tau_0)^{\frac{1}{2}} Y_0\theta/Y^4,\\
&H = 2(2\tau_0)^{\frac{1}{2}}(Y^2 - 2\theta^2)/Y^4.
\end{aligned} \tag{4.14}$$

This solution does not show geometrical singularities on the timelike surface on which $X = 0$, which separates the closed universe and the outer asymptotically flat one. The corresponding solution for the NUT part of the space (Case Ic) is

$$\begin{aligned}
&Y = (Y_0 + \theta^2)^{\frac{1}{2}},\\
&F = \left(\frac{F_0 + c_0\theta}{Y_0^2 + \theta^2} - 4Y_0^2\right)^{-\frac{1}{2}}, \quad C_0, F_0, Y_0 \text{ const},\\
&r = Y^{-2},\\
&\omega = 1/FY^2,\\
&\theta(x^1) = \int^{x^1} (FY_0)^{-1}\, dx^1,\\
&\tau = \tau_0/Y^4, \quad \tau_0 \text{ const},\\
&E = 2(2\tau_0)^{\frac{1}{2}} Y_0\theta/Y^4,\\
&H = (2\tau_0)^{\frac{1}{2}}(Y^2 - 2\theta^2)/Y^4.
\end{aligned} \tag{4.15}$$

To summarize, we have found all the solutions for which an equation of state is predetermined in Sec. 3. In Sec. 4 we have tried to complete the classification by integrating the equations after giving an equation of state. We have found all the Class I solutions for a fluid except for the general Case Id. For this case the dust solutions have been found previously (see Ref. 1), and a vacuum electromagnetic field solution was essentially obtained by Brill.[8]

Class II solutions are incomplete. There are Case IIa solutions for a fluid, and for a vacuum electromagnetic field, but a solution for both has not been found. There is no static or shear ($\alpha = 0$) case.

Case IIb solutions for a barotropic fluid are well known if $\epsilon E = 0$. If $\epsilon E \neq 0$, we have found the static solution; there is no shear ($\alpha = 0$) solution, and the nonstationary case remains unsolved.

Case IIc: Little progress has been made, even though much study has been devoted to this case by many workers, although much of the work has been done for a nonbarotropic equation of state.

[12] B. Carter (unpublished).
[13] B. Bertotti, Phys. Rev. **116**, 1331 (1959); I. Robinson, Bull. Acad. Pol. Sci. **7**, 351 (1959).
[14] G. E. Tauber, J. Math. Phys. **8**, 118 (1967).
[15] G. C. McVittie, Astrophys. J. **143**, 682 (1966); R. F. Tooper, Astrophys. J. **140**, 434 (1964); **143**, 465 (1966), H. A. Buchdahl, *ibid.* **140**, 1512 (1964); **147**, 310 (1967); and Ref. 7.
[16] M. Cahen and R. Debever, Compt. Rend. **260**, 815 (1965).

Case III: All the fluid solutions have been found, except for Case IIIc, but the dust solutions have not been given for this case yet. Brill[8] has obtained a vacuum electromagnetic field solution for Case IIIb.

5. DISCUSSION OF THE GENERAL CASE

If the extra terms q, π in the energy–momentum tensor are nonzero, then the system is no longer in thermal equilibrium. The mean velocity u^a is essentially defined as the normalized flux of some quantity, e.g., number of particles, momentum, energy, entropy, and so several definitions are possible. Only in thermal equilibrium do these definitions agree. This difficulty can be avoided in several situations. For example, one could consider a perfect fluid through which passes an energy flux (e.g., of neutrinos) which does not interact with the fluid, a problem which has been discussed by several authors, e.g., Ref. 17. To consider this as a one-fluid problem, it is necessary to introduce a π term representing anisotropic pressures as well as a q term, but π can be written in terms of the other quantities ($\pi = 3q$) and the problem is well defined once an equation of state for q is given. Equation (2.7) now acts as a restriction on the geometry.

Even where viscous effects are important, it may be possible under favorable conditions to use simple methods,[18] but in general it will be necessary to use the general relativistic Boltzmann equation. T_{ab} must be written

$$T^{ab}(x) = \int F(\mathbf{x}, \mathbf{p}) p^a p^b d^4p, \tag{5.1}$$

where $F(\mathbf{x}, \mathbf{p})$ is the particle distribution function in 8-dimensional phase space. The previously assumed L.R.S. of T_{ab} gives no information as to the possible symmetry of $F(\mathbf{x}, \mathbf{p})$. The concept of local dynamical symmetry (L.D.S.), i.e., L.R.S. on the 4-manifold of phase space spanned by the momenta, was introduced by Tauber and Weinberg,[19] and recently extended by Ehlers, Geren, and Sachs.[20] They showed that if the particles have nonzero mass and if g is 3-dimensional, then L.D.S. implies that space–time is stationary or a shearfree Robertson–Walker type. If the particles do have zero rest mass, then an extra assumption $\omega = 0$ or $\dot{u} = 0$ is required.

Suppose L.R.S. holds in the 4-manifold of space–time spanned by the $F(\mathbf{x}, \mathbf{p})$, and the system is in near equilibrium:

$$F(\mathbf{x}, \mathbf{p}) = F_0(\mathbf{x}, \mathbf{p}) + \epsilon F_1(\mathbf{x}, \mathbf{p}), \tag{5.2}$$

where $F_0(\mathbf{x}, \mathbf{p})$ is an equilibrium distribution and ϵ is small. Then the alternative definitions of u^a are nearly equivalent and a suitable 4-velocity can be defined. This gives a model to which (2.7) is applicable. Then, q, π are small quantities and $\omega k \sim \epsilon$. There are three disjoint and exhaustive classes of solutions.

Class I: $\omega \sim 1$, $k \sim \epsilon$. Then $a, \dot{u} \sim 1$; $\alpha, \beta \sim \epsilon$. The solutions are nearly stationary.

Class II: $\omega = k = 0$. The metric of Theorem 3 can be introduced.

Class III: $\omega \sim \epsilon$; $k \sim 1$; $\alpha, \beta, \sim 1$; $a, \dot{u}, \partial_1 k, \partial_1 \alpha, \partial_1 \beta, \partial_1 \tau \sim \epsilon$. The solutions are nearly homogeneous.

In general it will be extremely difficult to integrate the equations exactly; there are in fact no known analytic solutions when dissipative processes are present. Vaidya[17] has obtained an analytical solution for the case where an energy flux is present. He solved the equations by imposing functional restraints on the metric-tensor components and then deducing the equations of state and energy generation. Although this method often leads to solutions which are neither simple nor physically very meaningful, it may well be the only easy way to obtain solutions. Part of the difficulty lies in obtaining simple but plausible forms for the energy-transport and viscosity coefficients. Some results have already been obtained,[21] but the study of relativistic nonequilibrium statistical mechanics is still in its infancy.

ACKNOWLEDGMENTS

The authors are indebted to Professor C. W. Misner for several useful discussions and to Dr. D. W. Sciama for his continuous advice and encouragement.

APPENDIX A: THE JACOBI IDENTITIES, MAXWELL EQUATIONS, AND FIELD EQUATIONS

The e^0 axis has been taken along u^a, and the e_1 axis along the axis of symmetry. L.R.S. has been invoked to set

$$\gamma^0_{02} = \gamma^0_{03} = \gamma^0_{31} = \gamma^0_{32} = \gamma^1_{02} = \gamma^1_{03} = \gamma^3_{01} = \gamma^2_{01} = \gamma^1_{21} = \gamma^1_{31} = \gamma^2_{13} = 0,$$

$$\gamma^0_{10} = \dot{u}, \gamma^1_{01} = -\alpha, \gamma^2_{02} = \gamma^3_{03} = -\beta,$$

$$\gamma^2_{12} = \gamma^3_{13} = a, \gamma^1_{32} = k, \gamma^3_{21} = m, \gamma^3_{23} = s.$$

[17] C. W. Misner, Phys. Rev. **137**, B1360 (1965); C. W. Misner and D. H. Sharp, Phys. Letters **15**, 279 (1965); P. C. Vaidya, Astrophys. J. **144**, 943 (1966); R. W. Lindquist, Ann. Phys. (N.Y.) **37**, 487 (1966).

[18] C. W. Misner, Nature **210**, 40 (1967).

[19] G. E. Tauber and J. W. Weinberg, Phys. Rev. **122**, 1342 (1961).

[20] J. Ehlers, P. Geren, and R. K. Sachs (unpublished).

[21] N. A. Chernikov, Acta Phys. Polon. **23**, 629 (1963).

Then the Jacobi identities are:

$$\binom{0}{123}\quad \partial_1\omega = \omega(2a + \dot{u}),$$

$$\binom{0}{023}\quad -2\partial_0\omega = 4\omega\beta + k\dot{u},$$

$$\binom{1}{123}\quad -\partial_1 k = 2\omega\alpha - 2ak,$$

$$\binom{3}{123}\quad \partial_1 s = as,$$

$$\binom{2}{012}\quad -\partial_1\beta - \partial_0 a = \beta\dot{u} + \alpha a - \omega m + e_2\cdot\dot{e}_3 m,$$

$$\binom{3}{023}\quad \partial_0 s = -\beta s,$$

$$\binom{1}{023}\quad -\partial_0 k = k(2\beta - \alpha),$$

$$\binom{2}{013}\quad \partial_1(e_2\cdot\dot{e}_3 - \omega) = -\dot{u}(e_2\cdot\dot{e}_3 - \omega),$$

$$\binom{3}{012}\quad \partial_1(e_2\cdot\dot{e}_3 - \omega) - \sigma m = -\dot{u}(e_2\cdot\dot{e}_3 - \omega) + \alpha m.$$

The Maxwell equations $F^{ab}{}_{;b} = \epsilon u^a$, $(\eta^{ab}{}_{cd}F^{cd})_{;b} = 0$ reduce to

$$\partial_1 E = 2aE - 2\omega H + \epsilon,$$
$$\partial_1 H = 2aH + 2\omega E,$$
$$\partial_0 E = -2\beta E - kH,$$
$$\partial_0 H = -2\beta H + kE.$$

The field equations are:

$$(00)\quad -\Lambda + \tfrac{3}{2}p + \tfrac{1}{2}\mu + \tau = \partial_1\dot{u} - \partial_0(\alpha + 2\beta) - \alpha^2 - 2\beta^2 + 2\omega^2 + \dot{u}^2 - 2a\dot{u},$$

$$(11)\quad \Lambda - \tfrac{1}{2}p + \tfrac{1}{2}\mu - \tau + 2\pi = \partial_0\alpha + \partial_1(2a - \dot{u}) + \alpha^2 + 2\alpha\beta - \dot{u}^2 - 2a^2 + \tfrac{1}{2}(k^2 - m^2),$$

$$(22)\quad \Lambda - \tfrac{1}{2}p + \tfrac{1}{2}\mu + \tau - \pi = \partial_0\beta + \partial_1 a + \alpha\beta + 2\beta^2 + \partial_2 s - s^2 + 2\omega e_2\cdot\dot{e}_3 - \tfrac{1}{2}(m^2 + k^2) + km - 2a^2 + a\dot{u},$$

$$(33)\quad \Lambda - \tfrac{1}{2}p + \tfrac{1}{2}\mu + \tau - \pi = \partial_0\beta + \partial_1 a + \alpha\beta + 2\beta^2 + \partial_2 s - s^2 + 2e_2\cdot\dot{e}_3\omega + \tfrac{1}{2}(m^2 - k^2) - 2a^2 + a\dot{u},$$

$$(01)\quad -q = -2\partial_1\beta + 2a(\beta - \alpha) - \omega k,$$

$$(23)\quad \tfrac{1}{2}\partial_1(m - k) = \alpha\omega - \tfrac{1}{2}m\dot{u} + a(m - k).$$

APPENDIX B: PROOFS OF THEOREMS 1 AND 3

It is convenient to prove Theorems 1 and 3 together. A coordinate transformation

$$x^{0'} = x^0,\quad x^{\gamma'} = x^{\gamma'}(x^i)$$

is used to set $e_0{}^\gamma = \partial_0(x^\gamma) = 0$. The coordinates are now comoving with only x^0 varying along the world lines. A transformation $x^{0'} = x^{0'}(x^i)$, $x^{\nu'} = x^\nu$, can be used to set $e_0{}^0 = \partial_0(x^0) = F(x^i)$, where F is for the moment arbitrary. Then $g_{00} = -1/F^2$. Choose F in a convenient way on a 3-surface, $x^1 = \text{const}$, and propagate F by requiring that

$$\partial_1(\log F) = -\dot{u}. \tag{B1}$$

Then F is equivalent to the index function introduced by Synge[22]; the symmetry conditions introduced here have implied that the fluid is holonomic. (See also Ref. 23.)

The remaining coordinate freedom is $x^{0'} = x^0 + f(x^\nu)$, $x^{\nu'} = x^{\nu'}(x^\sigma)$. The vectors e_2, e_3 are arbitrary by a rotation

$$e_2' = e_2\cos\theta + e_3\sin\theta,$$
$$e_3' = -e_2\sin\theta + e_3\cos\theta,$$

where

$$\theta = \theta(x^i).$$

$\theta_{,0}$ can be fixed by the requirement $e_2\cdot\dot{e}_3 = \omega$ at all points, and $\theta_{,1}$ by $\gamma^3{}_{12} = -m = 0$ on $x^0 = c^0$. Then $\gamma^3{}_{02} = \gamma^2{}_{03} = \omega$, and from $\binom{3}{012}$, $\partial_0 m = -\alpha m$ so that $m = 0$ everywhere.

In an initial hypersurface $x^0 = c^0$, relabel the points

$$x^{0'} = x^0 = c^{0'}, x^{1'} = x^1,\quad x^{2'} = x^{2'}(x^\sigma),\quad x^{3'} = x^{3'}(x^\sigma)$$

so that $e_1{}^2 = \partial_1(x^2) = 0$, $e_1{}^3 = \partial_1(x^3) = 0$ in $x^0 = c^0$. From (2.1),

$$[e_0, e_1]x^2 = \partial_0 e_1{}^2 = \dot{u}\partial_0 x^2 - \alpha\partial_1 x^2 = -\alpha e_1{}^2.$$

Therefore $e_1{}^2 = 0$ everywhere; similarly $e_1{}^3 = 0$ everywhere.

Now make a transformation $x^{0'} = x^0 + f(x^\nu)$, $x^{\nu'} = x^\nu$ to set $e_1{}^0 = \partial_1 x^0 = 0$ in some surface $x^0 = \text{const}$. From (2.1),

$$[e_0, e_1]x^0 = \partial_0 e_1{}^0 - \partial_1 F = F\dot{u} - \alpha e_1{}^0.$$

From (B1), $\partial_0 e_1{}^0 = -\alpha e_1{}^0$, and so $e_1{}^0 = 0$ everywhere.

The remaining coordinate freedom is

$$x^{0'} = x^0 + f(x^2, x^3),\quad x^{1'} = x^{1'}(x^1, x^2, x^3),$$
$$x^{2'} = x^{2'}(x^2, x^3),\quad x^{3'} = x^{3'}(x^2, x^3).$$

[22] J. L. Synge, Proc. London Math. Soc. **43**, 376 (1937).
[23] A. Lichnerowicz, *Theories Relativistes de la Gravitation et de l'Electromagnetisme* (Masson & Cie., Paris, 1955), p. 92.

Since $\gamma^3{}_{01} = \gamma^3{}_{12} = \gamma^3{}_{20} = 0$, e_3 is hypersurface orthogonal; choose these surfaces to be $x^3 = \text{const}$. Then $\partial_2 x^3 = e_2{}^3 = 0$. This determines x^3 up to $x^{3'} = x^{3'}(x^3)$. Make a transformation

$$x^{0'} = x^0 + f(x^2, x^3), \quad x^{\nu'} = x^\nu,$$

to set $e_2{}^0 = 0$ in a 2-surface $x^0 = c^0, x^1 = c^1$. From (2.1),

$$[e_0, e_2]x^0 = \partial_0 e_2{}^0 = -\beta e_2{}^0,$$
$$[e_1, e_2]x^0 = \partial_1 e_2{}^0 = a e_2{}^0,$$
$$\text{and so } e_2{}^0 = 0 \text{ everywhere.}$$

As $\gamma^2{}_{03} = \gamma^2{}_{31} = \gamma^2{}_{01} = 0$, e_2 is hypersurface orthogonal. If these surfaces are chosen to be the surfaces $x^2 = \text{const}$, then $e_3{}^2 = 0$. Since $\gamma^1{}_{02} = 0$, x^1 can be chosen so that $e_2{}^1 = \partial_2 x^1 = 0$. We can specify $\theta = \theta(x^2, x^3)$ on an arbitrary 2-surface C: $x^0 = c^0$, $x^1 = c^1$, and this can be done so that $\gamma^2{}_{32} = 0$ on C and $\partial_3 s = 0$ on a line $x^2 = c^2$ in C. From $\binom{2}{032}$ and (31),

$$\partial_0 \gamma^2{}_{32} = -\beta \gamma^2{}_{32}, \quad \partial_1 \gamma^2{}_{32} = a \gamma^2{}_{32},$$

and so $\gamma^2{}_{32} = 0$ everywhere.

Similarly, from $\binom{3}{023}$, $\binom{3}{123}$, (22), and the commutation relations, $\partial_3 s = 0$. It can be shown that F can be chosen as a function only of x^0, x^1. From (2.1),

$$[e_0, e_3]x^i = \partial_0 e_3{}^i = -\beta e_3{}^i, \quad \text{where } i = 0 \text{ or } 3,$$
$$[e_1, e_3]x^i = \partial_1 e_3{}^i = a e_3{}^i, \quad \text{and so } e_3{}^0 = y(x^2, x^3) e_3{}^3.$$

If and only if $\omega = 0$, u^a is hypersurface orthogonal and we can set $e_3{}^0 = y = 0$.

This completes the outline of the proof of Theorem 1. For fuller details see Ref. 1. The only remaining freedom is the specification of θ and s at one point.

To prove Theorem 3 it is convenient to consider the three cases separately.

Case I: $\gamma^1{}_{23} = \gamma^1{}_{30} = \gamma^1{}_{02} = 0$ imply e_1 is hypersurface orthogonal. This implies that we can set $e_3{}^1 = 0$. x^1 is free by a transformation $x^{1'} = x^{1'}(x^1)$. $\gamma^1{}_{01} = \gamma^1{}_{21} = \gamma^1{}_{31} = 0$ imply $\partial_0 e_1{}^1 = \partial_2 e_1{}^1 = \partial_3 e_1{}^1 = 0$, and $e_1{}^1 = e_1{}^1(x^1)$. The commutation relations (2.1) applied to x^2 give $e_2{}^2 = B(x^1) f_2{}^2(x^2)$, and the freedom of x^2 can be used to set $f_2{}^2 = 1$. Similarly $e_3{}^3 = B(x^1) f(x^2) f_3{}^3(x^3)$, and we can set $f_3{}^3 = 1$. From $\binom{0}{123}$, $\partial_1 \omega = (2a + \dot{u})\omega$. Writing

$$\partial_1 \quad \text{as} \quad ', \quad \omega'/\omega + F'/F = 2a,$$

and so

$$\omega(x^1) = cB^2/F,$$

where c is constant. Also

$$[e_2, e_3]x^0 = \partial_2 e_3{}^0 = -2\omega F + s e_3{}^0,$$

and so

$$\partial y(x^2, x^3)/\partial x^2 = -2c/f(x^2). \tag{B2}$$

Case II: Since $\omega = k = 0$, we can set $y = e_3{}^1 = 0$; then $x^{0'} = x^0 + \text{const}$, $x^{1'} = x^{1'}(x^1)$ is the remaining freedom of choice of x^0, x^1.

The commutation relations applied to x^1, x^2, x^3 in turn give

$$e_1{}^1 = A(x^0, x^1),$$
$$e_2{}^2 = B(x^0, x^1) f_2{}^2(x^2), \quad \text{and we can set } f_2{}^2 = 1,$$
$$e_3{}^3 = B(x^0, x^1) f(x^2) f_3{}^3(x^3), \quad \text{and we can set } f_3{}^3 = 1.$$

Case III: Since $\omega = 0$, we can set $y = e_3{}^1 = 0$. Since $\dot{u} = 0$, we can put $F = 1$. The commutation relations give the same results as in Case II, but A and B are independent of x^1. From $\binom{1}{023}$, $\binom{1}{123}$, $\partial_0 k = k(\alpha - 2\beta)$, $\partial_1 k = 0$. Therefore $k = 2CB^2(x^0)/A(x^0)$, C a constant. The remaining commutation relations applied to x^1 give

$$e_3{}^1 = B(x^0) f(x^2) h(x^2, x^3),$$

where

$$\partial h(x^2, x^3)/\partial x^2 = -2C/f(x^2). \tag{B3}$$

The remaining commutation relations applied to x^3 give

$$s = B(x^0, x^1) \frac{\partial}{\partial x^2} (\log f(x^2))$$

for all three cases. From (22), (33), $\partial_2 s - s^2$ is a function of x^0, x^1 only, and so $\partial_2 s - s^2 = B^2(x^0, x^1)K$, where K is a constant, and

$$\frac{d^2 t(x^2)}{(dx^2)^2} + K t(x^2) = 0,$$

where $t(x^2) = 1/f(x^2)$. K is arbitrary and determines the form of $t(x^2)$ and $f(x^2)$. Specifying the constants c, C then gives $h(x^2, x^3)$, $y(x^2, x^3)$ from (B2) and (B3). (See Ref. 1.)

Now all the $e_a{}^i$ are known, g_{ij} can be determined and has the form (2.8), where $X = A^{-1}$, $Y = B^{-1}$.

With this coordinate system, the rotation coefficients can be written in terms of the metric components as follows:

$$\alpha = F \frac{\partial}{\partial x^0} (\log X),$$

$$\beta = F \frac{\partial}{\partial x^0} (\log Y),$$

$$a = -\frac{1}{X} \frac{\partial}{\partial x^1} (\log Y),$$

$$\dot{u} = -\frac{1}{X} \frac{\partial}{\partial x^1} (\log F),$$

$$s = K/Y^2, \quad K \text{ const.}$$

The other coefficients ω, k can be determined from the Jacobi identities.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 7 JULY 1968

Relativistic Effects in Atomic Fine Structure. II

LLOYD ARMSTRONG, JR.
Westinghouse Research Laboratories, Pittsburgh, Pennsylvania

(Received 16 October 1967)

The methods of second quantization are used to investigate matrix elements of the Breit equation in mixed configurations. In the special case of a single configuration, the results are shown to be identical to those obtained in the previous paper of this series using equivalent operators. Application of second-quantization methods to calculations involving relativistic atomic states is shown to be extremely straightforward and to provide a simplicity and clarity unattainable using the equivalent-operator technique.

I. INTRODUCTION

In a previous paper,[1] we obtained equivalent operators for the Breit interaction. These equivalent operators were defined such that their matrix elements evaluated with nonrelativistic wavefunctions were equal to matrix elements of the Breit operator evaluated using relativistic wavefunctions. The derivation of these equivalent operators involved evaluating the Breit operator between relativistic wavefunctions $|l^2SL)$ and expanding this result in a sum over nonrelativistic matrix elements of coupled double tensors[2] $W^{(k_1K_1)k}W^{(k_2K_2)k}$. Expansion coefficients were obtained by utilizing orthogonality and summation rules for $6j$ and $9j$ symbols.

In this paper we wish to present an alternative derivation of these equivalent operators and to extend the previous results by obtaining operators suitable for use in mixed configurations. The present derivation is based upon the creation and destruction operators of second quantization.[3] Use of this method simplifies the calculation immensely and provides valuable insight into the meaning and origin of equivalent operators.

II. DERIVATION OF THE OPERATORS

In the second-quantization formalism, one-electron operators $F = \sum_i f_i$ are written

$$F = \sum q_\alpha^+ \langle\alpha| f |\beta\rangle q_\beta, \tag{1}$$

where the sum is over all states α and β, and q_α^+ is a creation operator whose effect on the vacuum is to create a state $|\alpha\rangle$. The destruction operator q_α, when operating to the left on the vacuum state, produces a state $\langle\alpha|$. In the same manner, two-electron operators $G = \sum_{i<j} g_{ij}$ are written

$$G = \tfrac{1}{2}\sum q_\alpha^+ q_\beta^+ \langle\alpha_1\beta_2| g_{12} |\gamma_1\delta_2\rangle q_\delta q_\gamma, \tag{2}$$

where the subscripted numbers in the matrix element refer to electron number, and the sum is over all α, β, γ, and δ. A product of creation operators acting on the vacuum produces a determinental product state

$$q_\alpha^+ q_\beta^+ \cdots q_\phi^+ |0\rangle = \{\alpha\beta\cdots\phi\}. \tag{3}$$

An arbitrary wavefunction Ψ can therefore be represented by some properly normalized sum over products of creation operators, e.g.,

$$|l^2SLM_SM_L\rangle = \sum (2)^{-\frac{1}{2}} \begin{pmatrix} l & l & L \\ M_{l1} & M_{l2} & -M_L \end{pmatrix} \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & S \\ M_{s1} & M_{s2} & -M_S \end{pmatrix} \times (-1)^{M_L+M_S}[L, S]^{\frac{1}{2}} q_\alpha^+ q_\beta^+ |0\rangle, \tag{4}$$

where $\alpha = (\frac{1}{2}lM_{s1}M_{l1})$ and $\beta = (\frac{1}{2}lM_{s2}M_{l2})$.

In the present case, the terms f and g_{12} appearing in (1) and (2) will be terms in the Breit Hamiltonian; the states α, β, γ, and δ should therefore be relativistic electron states. As given in REFS, these relativistic states have the form

$$|ljm) = \begin{pmatrix} F/r & |ljm\rangle \\ iG/r & |\bar{l}jm\rangle \end{pmatrix}, \tag{5}$$

where $\bar{l} = l \pm 1$ as $j = l \pm \frac{1}{2}$, and $|ljm\rangle$ is an ordinary nonrelativistic state having the quantum numbers indicated. In (5), and throughout the paper, we indicate relativistic states by a rounded ket, e.g., $|ljm)$, and nonrelativistic states by an angular ket, e.g., $|ljm\rangle$. The wavefunction given by (5) is an eigenfunction of J^2; by analogy with the discussion given by Lawson and Macfarlane[4] for the nonrelativistic case, we can therefore say that the creation operator for such a wavefunction, $q^+(ljm)$, must transform like a component of a spherical tensor of rank j. Also transforming like a component of a spherical tensor of rank j is the operator $\tilde{q}(ljm) = (-1)^{j-m}q(lj-m)$.

[1] Lloyd Armstrong, Jr., J. Math. Phys. 7, 1891 (1966), to be referred to hereafter as REFS.

[2] Brian R. Judd, *Operator Techniques in Atomic Spectroscopy* (McGraw-Hill Book Company, Inc., New York, 1963).

[3] Brian R. Judd, *Second Quantization and Atomic Spectroscopy* (The Johns Hopkins Press, Baltimore, 1967).

[4] R. D. Lawson and M. H. Macfarlane, Nucl. Phys. 66, 80 (1965).

As was shown in REFS, the Hamiltonian terms of interest are given by

$$H_\alpha + H_\beta + H_\gamma + H_\delta, \tag{6}$$

where

$$H_\alpha = \sum_i H_\alpha^i, \quad H_\alpha^i = -Ze^2/r_i,$$

$$H_\beta = \sum_{i<j} e^2/r_{ij},$$

$$H_\gamma = \tfrac{1}{2}e^2 \sum_{i<j} (\boldsymbol{\alpha}_i \cdot \boldsymbol{\alpha}_j)/r_{ij},$$

and

$$H_\delta = -\tfrac{1}{2}e^2 \sum_{i<j} (\boldsymbol{\alpha}_i \cdot \mathbf{r}_{ij})(\boldsymbol{\alpha}_j \cdot \mathbf{r}_{ij})/r_{ij}^3.$$

The expression for H_α in the second-quantization formalism is easily obtained using (1):

$$H_\alpha = -Ze^2 \sum q^+(l_1 j_1 m_1)\,(l_1 j_1 m_1|\, 1/r\, |l_2 j_2 m_2) \times q(l_2 j_2 m_2), \tag{7}$$

where the sum is over all l_1, l_2, j_1, j_2, m_1, and m_2. In REFS, the matrix element appearing in (7) was found to be

$$(l_1 j_1 m_1|\, \frac{1}{r}\, |l_2 j_2 m_2) = \delta(l_1, l_2)\delta(j_1, j_2)\delta(m_1, m_2) \times \int \frac{F_j^2 + G_j^2}{r}\, dr. \tag{8}$$

Using this and the transformation properties of the operators, (7) becomes

$$H_\alpha = +Ze^2 \sum [j]^{\frac{1}{2}} \int \frac{F_j^2 + G_j^2}{r}\, dr\, \delta(l_1, l_2) \times \delta(j_1, j_2)[\mathbf{q}^+(l_1 j_1)\mathbf{q}(l_1 j_1)]^0, \tag{9}$$

where the term in brackets indicates that the two tensors of rank j are coupled to 0.

The two-particle forms of H_β, H_γ, and H_δ were shown in REFS to be of the general form $g_{12} = \sum_k g_1^k \cdot g_2^k$. Using this form of the two-body interaction in (2) shows that the second-quantization form of H_β, H_γ, H_δ is given by

$$\begin{aligned} &\tfrac{1}{2} \sum q^+(l_1 j_1 m_1) q^+(l_2 j_2 m_2)([l_1 j_1 m_1]_1 [l_2 j_2 m_2]_2 \\ &\quad \times |g_1^k \cdot g_2^k|\, [l_3 j_3 m_3]_1 [l_4 j_4 m_4]_2) q(l_4 j_4 m_4) q(l_3 j_3 m_3) \\ &= \tfrac{1}{2} \sum [k]^{-1} (l_1 j_1 \| \, g_1^k \, \| l_3 j_3)(l_2 j_2 \| \, g_2^k \, \| l_4 j_4) \\ &\quad \times \Big\{ [\mathbf{q}^+(l_1 j_1)\mathbf{q}(l_3 j_3)]^k \cdot [\mathbf{q}^+(l_2 j_2)\mathbf{q}(l_4 j_4)]^k \\ &\quad - [\mathbf{q}^+(l_1 j_1)\mathbf{q}(l_4 j_4)]^0 \left(\frac{[k]^2}{[j_1]}\right)^{\frac{1}{2}} (-1)^{j_1 + j_3} \delta(l_2 l_3)\delta(j_2 j_3) \Big\}, \end{aligned} \tag{10}$$

where the sum is over k and all values of l and j. The second term in brackets in (10) results from the use of fermion anticommutation relations in changing the position of $q(l_3 j_3 m_3)$ relative to $q^+(l_2 j_2 m_2)$.

The second-quantization form of operators, given by (1) and (2), is most useful if the states α, β, γ, and δ are of the same type as the single-particle states which compose the wavefunctions that will be used to evaluate matrix elements. For example, if the wavefunctions are composed of single-particle states coupled together through jj coupling, the operators should be of the form $q^+(ljm)$, etc. In REFS, the wavefunctions used are of the form $|l^N SLJM)$; implicit in the derivation given is the definition that a single-particle wavefunction $|\frac{1}{2} l m_s m_l)$ is given by

$$|\tfrac{1}{2} l m_s m_l) = \sum_j [j]^{\frac{1}{2}} (-1)^{l - \frac{1}{2} - m} \begin{pmatrix} \frac{1}{2} & l & j \\ m_s & m_l & -m \end{pmatrix} |ljm). \tag{11}$$

This definition amounts to ignoring the coupling of the small components of the wavefunction, and the small component of the left-hand side of (11) contains a combination of states $|\frac{1}{2} l + 1 m_s m_{l+1})$ and $|\frac{1}{2} l - 1 m_s m_{l-1})$. The many-particle wavefunction $|l^N SLJM)$ of REFS is created by coupling the relativistic single-particle functions given by Eq. (11) as if they were nonrelativistic wavefunctions $|\frac{1}{2} l m_s m_l\rangle$ being coupled to $|l^N SLJM\rangle$. Again, this amounts to ignoring the coupling of the small components of the wavefunctions. For the more general case, we also define in the same manner the relativistic wavefunction $|l_1^N S_1 L_1, l_2^M S_2 L_2 \cdots ; SL)$, corresponding to the nonrelativistic function $|l_1^N S_1 L_1, l_2^M S_2 L_2 \cdots ; SL\rangle$.

Equation (11) can be used to define the operator $q^+(\frac{1}{2} l m_s m_l)$, which creates the state on the left of (11), in terms of $q^+(ljm)$. Taking adjoints of both sides leads to an identical expression for $\tilde{q}(\frac{1}{2} l m_s m_l)$ in terms of $\tilde{q}(ljm)$. In order to compare (9) and (10) with the equivalent operators obtained in REFS, therefore, we must write the operators $q(ljm)$ in terms of operators $q(\frac{1}{2} l m_s m_l)$. This is easily done using familiar recoupling coefficients. In place of (9) we obtain

$$\begin{aligned} H_\alpha &= Ze^2 \sum [j][k]^{\frac{1}{2}} (-1)^{l + \frac{1}{2} + k + j} \begin{Bmatrix} \frac{1}{2} & \frac{1}{2} & k \\ l & l & j \end{Bmatrix} \\ &\quad \times \int \frac{F_j^2 + G_j^2}{r}\, dr [\mathbf{q}^+(\tfrac{1}{2} l)\mathbf{q}(\tfrac{1}{2} l)]^{(kk)0} \\ &= Ze^2 \sum [j][k]^{\frac{1}{2}} (-1)^{l - \frac{1}{2} + k + j} \begin{Bmatrix} \frac{1}{2} & \frac{1}{2} & k \\ l & l & j \end{Bmatrix} \\ &\quad \times \int \frac{F_j^2 + G_j^2}{r}\, \mathbf{R}^{(kk)0}(l, l). \end{aligned} \tag{12}$$

The sum in (12) is over j, k, and l, and the expression $[\mathbf{q}^+(\frac{1}{2} l)\mathbf{q}(\frac{1}{2} l)]^{(kk)0}$ implies that the "large components"

of the operators are coupled to a total spin **k** and total orbital angular momentum **k**, which are then coupled to a total momentum of 0. In the second expression on the right of (12), these coupled operators have been replaced by $-\mathbf{R}^{(kk)0}(l,l)$; using an argument similar to that used by Judd[3] in the nonrelativistic case, $R^{(K_1K_2)K}$ can be shown to be equivalent to a sum over all electrons of a single-particle operator

$$\mathbf{R}^{(K_1K_2)K}(l_1l_2) = \sum_{i=1}^{N} \mathbf{r}_i^{(K_1K_2)K}(l_1l_2), \tag{13}$$

where

$$(\tfrac{1}{2}l_\alpha \| r^{(K_1K_2)}(l_1l_2) \| \tfrac{1}{2}l_\beta) = [K_1, K_2]^{\frac{1}{2}}\delta(l_\alpha, l_1)\delta(l_2, l_\beta). \tag{14}$$

Because of Eq. (14) and the manner in which the relativistic many-particle functions are constructed, $R^{(K_1K_2)K}(l_1l_2)$, evaluated between relativistic wavefunctions $|l_1^N S_1L_1, l_2^M S_2L_2 \cdots ; SLJ)$, will produce exactly the same result as will Feneuille's[5] $W^{(K_1K_2)K}(l_1l_2)$, evaluated between nonrelativistic functions

$$|l_1^N S_1L_1, l_2^M S_2L_2 \cdots ; SLJ\rangle.$$

In the case of $l_1 = l_2$, $W^{(K_1K_2)K}(l_1l_2)$ becomes equivalent to $W^{(K_1K_2)K}$ of Judd[2]; the results of the present work can therefore be easily compared with the results of REFS by letting $l_1 = l_2 = l_3 = l_4$ and replacing $R^{(K_1K_2)K}(l_1l_2)$ by $W^{(K_1K_2)K}$. Thus we see that the operator in Eq. (12) is exactly equivalent to Eqs. (10) and (12) of REFS.

In the same manner, the right-hand side of Eq. (10) becomes

$$\begin{aligned}
&\tfrac{1}{2}\sum [k]^{-1}(l_1j_1\| g_1^k \|l_3j_3)(l_2j_2\| g_2^k \|l_4j_4) \\
&\quad\times \left[\begin{Bmatrix} \frac{1}{2} & \frac{1}{2} & k_1 \\ l_1 & l_3 & K_1 \\ j_1 & j_3 & k \end{Bmatrix} \begin{Bmatrix} \frac{1}{2} & \frac{1}{2} & k_2 \\ l_2 & l_4 & K_2 \\ j_2 & j_4 & k \end{Bmatrix} \right. \\
&\quad\times [k_1, K_1, k_2, K_2, j_1, j_2, j_3, j_4]^{\frac{1}{2}} \\
&\quad\times \mathbf{R}^{(k_1K_1)k}(l_1l_3)\cdot\mathbf{R}^{(k_2K_2)k}(l_2l_4) \\
&\quad - (-1)^{k_1+l_1+j_3+\frac{1}{2}}[k][k_1]^{\frac{1}{2}} \begin{Bmatrix} \frac{1}{2} & \frac{1}{2} & k_1 \\ l_1 & l_4 & j_1 \end{Bmatrix} \\
&\quad\times \left. \delta(l_2l_3)\delta(j_2j_3)\mathbf{R}^{(k_1k_1)0}(l_1l_4) \right],
\end{aligned} \tag{15}$$

where the sum is over k_1, K_1, k_2, K_2, k, and all l and j values. Because of (13), the $\mathbf{R}^{(k_1K_1)k}\cdot\mathbf{R}^{(k_2K_2)k}$ term in (15) can be written $\sum_{i,j}\mathbf{r}_i^{(k_1K_1)k}\cdot\mathbf{r}_j^{(k_2K_2)k}$; however, by considering a matrix element of $\mathbf{r}_i^{(K_1K_1)K}\cdot\mathbf{r}_i^{(K_2K_2)K}$, one can show in a straightforward manner that the second term in (15) cancels out the contribution to the first term from the part of the sum over i and j in which $i = j$. Equation (15) then can be written

$$\begin{aligned}
&\tfrac{1}{2}\sum [k]^{-1}[k_1, K_1, k_2, K_2, j_1, j_2, j_3, j_4]^{\frac{1}{2}} \\
&\times \begin{Bmatrix} \frac{1}{2} & \frac{1}{2} & k_1 \\ l_1 & l_3 & K_1 \\ j_1 & j_3 & k \end{Bmatrix} \begin{Bmatrix} \frac{1}{2} & \frac{1}{2} & k_2 \\ l_2 & l_4 & K_2 \\ j_2 & j_4 & k \end{Bmatrix} (l_1j_1\| g_1^k \|l_3j_3) \\
&\times (l_2j_2\| g_2^k \|l_4j_4) \sum_{i\neq j} \mathbf{r}_i^{(k_1K_1)k}(l_1, l_3)\cdot\mathbf{r}_j^{(k_2K_2)k}(l_2, l_4),
\end{aligned} \tag{16}$$

where the sum is over k_1, K_1, k_2, K_2, k, and all l and j values.

Turning our attention now to the reduced matrix elements appearing in (16), for H_β we find that

$$\begin{aligned}
&(l_1j_1\| g_1^k \|l_3j_3)(l_2j_2\| g_2^k \|l_4j_4) \\
&= e^2\iint (l_1j_1\| h_1^k \|l_3j_3)(l_2j_2\| h_2^k \|l_4j_4) \frac{r_<^k}{r_>^{k+1}}\, dr_1\, dr_2,
\end{aligned} \tag{17}$$

with

$$\begin{aligned}
(l_1j_1\| h^k \|l_3j_3) &= (-1)^{j_1-\frac{1}{2}} \begin{pmatrix} j_1 & k & j_3 \\ -\frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix} \\
&\quad\times [j_1, j_3]^{\frac{1}{2}}(\Delta(l_1, l_3, k)F_1F_2 + \Delta(\bar{l}_1, \bar{l}_3, k)G_1G_2),
\end{aligned}$$

where $\Delta(l_1, l_3, k)$ is one if l_1, l_3, and k can form a triangle and their sum is even, and zero otherwise. For H_γ we find

$$\begin{aligned}
&(l_1j_1\| g_1^k \|l_3j_3)(l_2j_2\| g_2^k \|l_4j_4) = -\frac{e^2}{2} \\
&\quad\times \sum_\beta \iint (l_1j_1\| (\alpha_1C_1^\beta)^k \|l_3j_3)(l_2j_2\| (\alpha_2C_2^\beta)^k \|l_4j_4) \\
&\quad\times (-1)^{1+k+\beta}\frac{r_<^\beta}{r_>^{\beta+1}}\, dr_1\, dr_2,
\end{aligned} \tag{18}$$

with

$$\begin{aligned}
(l_1j_1\| (\alpha C^\beta)^k \|l_3j_3) &= i[k, j_1, j_3]^{\frac{1}{2}} \\
&\quad\times \left\{ \sqrt{2}\,(-1)^{l_1+\beta+1} \begin{pmatrix} 1 & \beta & k \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} j_1 & j_3 & k \\ -\frac{1}{2} & -\frac{1}{2} & 1 \end{pmatrix} \right. \\
&\quad\times [\Delta(l_1, \bar{l}_3, \beta)F_1G_3 + \Delta(\bar{l}_1, l_3, \beta)G_1F_3] \\
&\quad + (-1)^{j_3+\frac{1}{2}} \begin{pmatrix} 1 & \beta & k \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} j_1 & j_3 & k \\ -\frac{1}{2} & +\frac{1}{2} & 0 \end{pmatrix} \\
&\quad\times \left. [\Delta(l_1\bar{l}_3\beta)F_1G_3 - \Delta(\bar{l}_1, l_3, \beta)G_1F_3] \right\}.
\end{aligned}$$

To obtain H_δ, we must note that Eq. (25) of REFS for an expansion of $(r_{12}r_{12})^2/r_{12}^3$ is not correct for mixed configurations. For the more general case, we

[5] S. Feneuille, J. Phys. Radium **28**, 61 (1967).

must write

$$\frac{(r_{12}r_{12})^2}{r_{12}^3} = \sum_\beta (-1)^\beta \frac{r_<^\beta}{r_>^{\beta+1}} \times \left\{ (\mathbf{C}_1^\beta \mathbf{C}_2^\beta)^2 \left[\frac{(8\beta)(\beta+1)(2\beta+1)}{(15)(2\beta-1)(2\beta+3)} \right]^{\frac{1}{2}} - (\mathbf{C}_<^{\beta-2}\mathbf{C}_>^\beta)^2 \left[\frac{\beta(\beta-1)(2\beta-3)(2\beta+1)}{5(2\beta-1)} \right]^{\frac{1}{2}} + (\mathbf{C}_<^\beta \mathbf{C}_>^{\beta+2})^2 \left[\frac{(\beta)(\beta+2)(2\beta+1)(2\beta+5)}{5(2\beta+3)} \right]^{\frac{1}{2}} \right\}, \tag{19}$$

where $\mathbf{C}_<$ implies either $\mathbf{C}_1$ or $\mathbf{C}_2$, directly as $r_<$ implies r_1 or r_2, etc. This can be written more compactly as

$$\sum_{\beta,\varphi,\delta} (-1)^\beta \frac{r_<^\beta}{r_>^{\beta+1}} (\mathbf{C}_<^\varphi \mathbf{C}_>^\delta)^2 F(\beta\varphi\delta), \tag{20}$$

where the sum over φ is limited to $\varphi = \beta - 2, \beta$ and the sum over δ to $\delta = \beta, \beta + 2$; $F(\beta\varphi\delta)$ is the term multiplying $(C^\varphi C^\delta)^2$ in Eq. (19). Using (20), we find that for H_δ

$$\begin{aligned}(l_1 j_1 \| g_1^k \| l_3 j_3)(l_2 j_2 \| g_2^k \| l_4 j_4) &= -\frac{e^2}{6} \sum_\beta \iint (l_1 j_1 \| (\alpha_1 C_1^\beta)^k \| l_3 j_3) \\ &\times (l_2 j_2 \| (\alpha_2 C_2^\beta)^k \| l_4 j_4)(-1)^{1+k+\beta} \frac{r_<^\beta}{r_>^{\beta+1}} dr_1\, dr_2 \\ &+ \frac{5e^2}{2} \sum_{\beta\varphi\delta} F(\beta\varphi\delta) \begin{Bmatrix} 1 & 1 & 2 \\ \varphi & \delta & k \end{Bmatrix} \int_0^\infty dr_2 \\ &\times \left[\int_0^{r_2} dr_1 \frac{r_1^\beta}{r_2^{\beta+1}} (l_1 j_1 \| (\alpha_1 C_1^\varphi)^k \| l_3 j_3) \right. \\ &\times (l_2 j_2 \| (\alpha_2 C_2^\delta)^k \| l_4 j_4) \\ &+ \left. \int_{r_2}^\infty dr_1 \frac{r_2^\beta}{r_1^{\beta+1}} (l_1 j_1 \| (\alpha C_1^\delta)^k \| l_3 j_3) \right. \\ &\times \left. (l_2 j_2 \| (\alpha_2 C_2^\varphi)^k \| l_4 j_4) \right]. \end{aligned} \tag{21}$$

We note from Eq. (17) that Eq. (15) of REFS should contain $(-1)^{j_1-\frac{1}{2}}$ rather than $(-1)^{j_3-\frac{1}{2}}$; the proper phase was used in all of the calculations. Equation (18) indicates that Eq. (21) of REFS is also in error, with the first term properly having a phase $(-1)^{l+1+\beta}$ rather than $(-1)^{l+1}$, and the expression G_3F_1 properly being G_1F_3. Finally,[6] the right-hand side of Eq. (29) of REFS should be multiplied by a factor of 2 and $r_<^\beta/r_>^{\beta+1}$ replaced by $r_1^\beta/r_2^{\beta+1}$, with the integral over r_1 having the limits 0 and r_2.

III. DISCUSSION

Using the second-quantization formalism, we have obtained expressions for the operators of the Breit equation which can be used to evaluate matrix elements between relativistic states of the form $|l_1{}^N S_1 L_1; l_2{}^M S_2 L_2; \cdots; SLJM)$. The angular dependence of these expressions is contained in the operators $R^{(K_1K_2)K}(l_1 l_2)$; evaluation of these operators in the relativistic scheme is exactly equivalent mathematically to the evaluation of the usual double tensors[2] in the nonrelativistic scheme. The equivalent operators of REFS can therefore be viewed as the result of manipulating the Breit operator into such a form that the powerful tensor techniques of Racah[2,7] can be fruitfully used in the relativistic scheme.

The present method simplifies immensely the calculation of REFS and, because of this increased simplicity, the results of REFS become more transparent. For example, the products of $9j$ symbols appearing in the results of REFS are shown by the present derivation to be merely transformation coefficients from jj to LS coupling.

---

[6] I should like to thank Dr. S. Feneuille for bringing this error to my attention.

[7] G. Racah, Phys. Rev. **62**, 438 (1942); **63**, 367 (1943); **76**, 1352 (1949).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 7 JULY 1968

Error Bounds in Spectroscopy and Nonequilibrium Statistical Mechanics

R. G. GORDON*
Department of Chemistry, Harvard University, Cambridge, Massachusetts

(Received 24 October 1967)

Upper and lower bounds are derived for the response of a system initially in equilibrium to a weak damped-harmonic perturbation. The only information required in the construction of these error bounds are equilibrium properties of the unperturbed system, in particular, the equilibrium moments of the spectral density corresponding to the perturbation are used. The error bounds are shown to be the most precise possible, given only this equilibrium information. As an example, error bounds are obtained for the shape of a nuclear-resonance spectrum due to dipolar broadening in solids, and good agreement is obtained with nuclear-magnetic resonance experiments on CaF_2. The error bounds are applicable to many other areas of spectroscopy and nonequilibrium statistical mechanics.

I. INTRODUCTION

A problem of central importance in spectroscopy and nonequilibrium statistical mechanics is the calculation of spectral densities for time-dependent fluctuations. For, according to the general fluctuation-dissipation theorem, these spectral densities determine the linear dissipative response of a system to various perturbations.[1] For example, spectral density corresponding to the fluctuation of the dipole moment determines the absorption coefficient for radiation. The spectrum of the density fluctuations determines the scattering cross section for light and neutrons.[2] The heat conductivity depends upon the spectral density of the heat-current fluctuations, and similar relations hold for other transport coefficients.[3,4]

In this paper we shall define precisely the extent to which such linear responses are determined by equilibrium properties of the unperturbed system. In particular, the equilibrium properties which we consider are the frequency moments of various spectral densities. Assuming that a finite number of these equilibrium moments have been calculated, we determine rigorous upper and lower bounds to the linear response of the system to a perturbation of finite duration in time. We consider in detail the response to a perturbation with a damped-harmonic time dependence. This amounts to studying the spectral density "broadened" by convolution with a Lorentzian line shape of any given width. The restriction to a damped-harmonic perturbation is not essential, and similar methods can be used to obtain error bounds to other forms of time-dependent perturbations.

These upper and lower bounds on the response are shown to be the most precise that are possible, using only the given number of equilibrium moments. In addition, the spectral density, which has the specified moments and *attains* the upper bound to the response, is shown to be unique, and this "maximal" spectral density is constructed explicitly from the moments. Similarly, the spectral density which attains the minimum response is shown to be unique, and its construction is given.

These methods are used to study the response of a set of interacting spins on an infinite rigid lattice to a damped-harmonic applied magnetic field. In other words, we calculate the upper and lower error bounds for the shape of a nuclear-magnetic resonance (NMR) spectrum of a dipole-coupled solid, using the moments of the spectral density calculated by Van Vleck[5] for this case. The calculated spectrum is broadened by a "theoretical resolution function" of Lorentzian shape, and any width one chooses. Naturally, the broader the chosen "theoretical resolution function," the narrower the error bounds become, but the more detail is lost by the broadening. The more moments used, the more precise the error bounds become for any fixed resolution. In any case, one has the assurance that the best possible use is being made of the known moments. These calculations are compared with NMR experiments on CaF_2, and agreement is obtained within the combined error bounds of theory and experiment.

Moments have been calculated for other spectral densities, including absorption of radiation,[6] neutron scattering,[7] and Raman light scattering.[8] The present method of moment inversion is being applied to these

* Sloan Foundation Fellow.

[1] R. Kubo, Can. J. Phys. **34**, 1274 (1956); J. Phys. Soc. Japan **12**, 570 (1957).

[2] L. van Hove, Phys. Rev. **95**, 249 (1954).

[3] M. S. Green, J. Chem. Phys. **20**, 1281 (1952); **22**, 398 (1954); H. Mori, Phys. Rev. **112**, 1829 (1958).

[4] For recent reviews of this field, see R. W. Zwanzig, Ann. Rev. Phys. Chem. **16**, 67 (1965); R. G. Gordon, Advan. Magnetic Resonance **3**, 1 (1967).

[5] J. H. Van Vleck, Phys. Rev. **74**, 1168 (1948).

[6] R. G. Gordon, J. Chem. Phys. **41**, 1819 (1964).

[7] G. Placzek, Phys. Rev. **86**, 377 (1952).

[8] R. G. Gordon, J. Chem. Phys. **40**, 1973 (1964).

and other cases. These investigations define rigorously and precisely the extent to which equilibrium properties of a system determine its nonequilibrium behavior.

II. MATHEMATICAL FORMULATION

We consider a system described by a Hamiltonian H_0 subject to a perturbation H_1 of the form

$$H_1 = H_{\text{int}} \cos \omega_0 t \exp(-\epsilon |t|), \tag{1}$$

where H_{int} depends only on the variables in the system. The time-dependent amplitude of the perturbation is chosen here to represent a damped-harmonic perturbing field (for example, an applied electric or magnetic field). The linear response $\bar{I}(\omega_0, \epsilon)$ of the system to this pulse has the general form given by perturbation theory

$$\bar{I}(\omega_0, \epsilon) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega_0 t} e^{-\epsilon|t|} C(t)\, dt, \tag{2}$$

where

$$C(t) \equiv \langle H_{\text{int}}(0) H_{\text{int}}(t) \rangle \tag{3}$$

is the correlation function of the interaction Hamiltonian, averaged over the unperturbed equilibrium ensemble of the system Hamiltonian H_0. The time-dependence in $H_{\text{int}}(t)$ is that generated by H_0 alone, so that $C(t)$ represents the time-dependence of the equilibrium fluctuations of H_{int}.

The physical meaning attached to this response is different in different physical situations. In an absorption experiment it is related to probability of absorption, and in a scattering experiment it is related to the probability or cross section for scattering. Because of this probability interpretation, it is clear that $\bar{I}(\omega_0, \epsilon)$ is a positive (or at least nonnegative) quantity

$$\bar{I}(\omega_0, \epsilon) \geq 0. \tag{4}$$

Mathematically, the positive-definiteness of the response is clear from the alternate expression for time-dependent response in terms of the absolute square of a quantum-mechanical matrix element.[4] The thermodynamic interpretation of this positivity is that the entropy production[9] due to external perturbations is positive. The positive nature of the response is an important fact to be used in the construction of our error bounds.

The convolution theorem for Fourier transforms allows us to write the product in Eq. (2) as the convolution

$$\bar{I}(\omega_0, \epsilon) = \frac{\epsilon}{\pi} \int_{-\infty}^{\infty} \frac{I(\omega)\, d\omega}{(\omega_0 - \omega)^2 + \epsilon^2}, \tag{5}$$

where

$$I(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} C(t) e^{-i\omega t}\, dt. \tag{6}$$

We note that

$$I(\omega) = \lim_{\epsilon \to 0^+} \bar{I}(\omega_0, \epsilon); \tag{7}$$

i.e., $I(\omega)$ is the idealized limiting response to an infinite undamped perturbation. The function $I(\omega)$ is also nonnegative, since it is a special case of $\bar{I}(\omega_0, \epsilon)$.

The relevant equilibrium properties for discussing a linear response are the Taylor expansion coefficients for the correlation function $C(t)$:

$$C(t) = \sum_{n=0}^{\infty} (it)^n \frac{\mu_n}{n!}, \tag{8}$$

where

$$\mu_n \equiv (-i)^n \frac{d^n C(t)}{dt^n}\bigg|_{t=0} = (-i)^n \left\langle H_{\text{int}}(0) \frac{d^n H_{\text{int}}(t)}{dt^n} \right\rangle\bigg|_{t=0}. \tag{9}$$

Now, inverting the Fourier transform, Eq. (6), gives

$$C(t) = \int_{-\infty}^{\infty} e^{i\omega t} I(\omega)\, d\omega. \tag{10}$$

Thus carrying out the differentiations indicated in Eq. (9) gives the well-known result,

$$\mu_n = \int_{-\infty}^{\infty} \omega^n I(\omega)\, d\omega, \tag{11}$$

that the coefficients of the time expansion of a correlation function are frequency moments of the corresponding spectral density. These moments are equilibrium properties of the unperturbed system [through Eq. (9)], and with a finite amount of effort one may expect to calculate a finite number, say $2M$, of them.

We can now state our problem: Given the equilibrium moments $\mu_0, \mu_1, \cdots, \mu_{2M}$ of a spectral density $I(\omega)$, find the minimum and maximum possible values of the response $\bar{I}(\omega_0, \epsilon)$, for each frequency ω_0 and pulse width ϵ^{-1}.

III. ERROR BOUNDS

In the last section we posed the problem of finding error bounds for the linear response of a system to a particular pulsed perturbation. Now we may state the essential mathematical features of the problem:

Given the $2M + 1$ real numbers $\mu_0, \mu_1, \cdots, \mu_{2M}$, find the minimum and maximum possible values of the linear functional $\bar{I}(\omega_0, \epsilon)$ defined by

$$\bar{I}(\omega_0, \epsilon) \equiv \frac{\epsilon}{\pi} \int_{-\infty}^{\infty} \frac{I(\omega)\, d\omega}{(\omega - \omega_0)^2 + \epsilon^2}, \tag{12}$$

[9] S. R. de Groot and P. Mazur, *Non-Equilibrium Thermodynamics* (North-Holland Publ. Co., Amsterdam, 1962), Ch. VIII.

where ω_0 and ϵ are fixed real numbers, and $I(\omega)$ ranges over the class of all nonnegative distributions, subject to the $2M+1$ constraints

$$\mu_n = \int_{-\infty}^{\infty} \omega^n I(\omega)\, d\omega; \quad n = 0, 1, \cdots, 2M. \tag{13}$$

The chief difficulty in actually constructing such maximum and minimum values is that an infinite number of distinct distributions $I(\omega)$ can have the specified moments. Therefore we clearly cannot construct each such distribution separately, and find its value of $\bar{I}(\omega_0, \epsilon)$. Rather, we must construct a general test, involving only a finite number of steps, of whether *any* $I(\omega)$ exists, which has the proper moments, *and* whose broadened distribution $\bar{I}(\omega_0, \epsilon)$ takes some prescribed value (say I_0). Thus we restate the problem:

Given the real numbers $\mu_0, \mu_1, \mu_2 \cdots \mu_{2M}, \omega_0$, and the positive numbers ϵ and I_0, does there exist any distribution $I(\omega)$ such that

$$\mu_n = \int_{-\infty}^{\infty} \omega^n I(\omega)\, d\omega, \quad \text{for} \quad n = 0, 1, \cdots, 2M, \tag{14a}$$

$$I(\omega) \geq 0, \quad \text{for all real } \omega, \tag{14b}$$

and

$$I_0 = \frac{\epsilon}{\pi} \int_{-\infty}^{\infty} \frac{I(\omega)\, d\omega}{(\omega - \omega_0)^2 + \omega^2}. \tag{14c}$$

Simply matching the moments [condition (a) above] can always be satisfied by an infinite number of distributions $I(\omega)$.[10] Requiring, in addition, that the distribution be nonnegative [condition (b)] greatly restricts the possibilities. A necessary and sufficient condition for the existence of some distribution $I(\omega)$, satisfying both conditions (a) and (b), is that all the determinants[11]

$$\Delta_n \equiv \begin{vmatrix} \mu_0 & \mu_1 & \mu_2 & \cdots & \mu_n \\ \mu_1 & \mu_2 & \mu_3 & \cdots & \mu_{n+1} \\ & & \cdot & & \\ & & \cdot & & \\ & & \cdot & & \\ \mu_n & \mu_{n+1} & \cdots & \cdots & \mu_{2n} \end{vmatrix} \geq 0 \tag{15}$$

be nonnegative for $n = 0, 1, 2 \cdots 2M$. In practice, this test is difficult to apply because of the large amount of arithmetic required to evaluate these determinants. We have been able to replace this test for the existence of a nonnegative solution by an equivalent one involving only $\sim M^2$ arithmetic operations. This new method has the additional advantage that also testing for condition (c) requires only a small additional effort. The method is suggested by writing the broadened response $\bar{I}$ in the form

$$\bar{I}(\omega_0, \epsilon) = -\frac{1}{\pi} \operatorname{Im} \int_{-\infty}^{\infty} \frac{I(\omega)\, d\omega}{-\omega_0 + i\epsilon + \omega}, \tag{16}$$

which is exactly equivalent to Eq. (12), but rewritten in complex form. This suggests that we consider the complex function

$$\mathfrak{I}(z) \equiv \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{I(\omega)\, d\omega}{(z + \omega)}, \tag{17}$$

since, if we set $z = -\omega_0 + i\epsilon$, the desired response is

$$\bar{I}(\omega_0, \epsilon) = -\operatorname{Im} [\mathfrak{I}(-\omega_0 + i\epsilon)]. \tag{18}$$

With these definitions, the response to an undamped perturbation is

$$I(\omega) = -\lim_{\epsilon \to 0^+} \operatorname{Im} [\mathfrak{I}(-\omega_0 + i\epsilon)]. \tag{19}$$

The integral $\mathfrak{I}(z)$ can be expressed in terms of its associated continued fraction[12]

$$A(z) = \cfrac{\alpha_1}{z + \alpha_2 - \cfrac{\alpha_2\alpha_3}{z + \alpha_3 + \alpha_4 - \cfrac{\alpha_4\alpha_5}{z + \alpha_5 + \alpha_6 - \cdots}}}, \tag{20}$$

where the expansion coefficients α_n can be determined uniquely from the condition that $\mathfrak{I}(z)$ and $A(z)$ have the same formal expansions in $(1/z)$:

$$\mathfrak{I}(z) = \frac{1}{\pi z} \int_{-\infty}^{\infty} I(\omega)\, d\omega - \frac{1}{\pi z^2} \int_{-\infty}^{\infty} \omega I(\omega)\, d\omega + \frac{1}{\pi z^3} \int_{-\infty}^{\infty} \omega^2 I(\omega)\, d\omega \cdots = \frac{1}{\pi} \sum_{n=0}^{\infty} (-1)^n \mu_n / z^{n+1}. \tag{21}$$

The α_n can be constructed from the μ_n using determinantal expressions,[12] or much more conveniently by the use of a simple product–difference recursion relation.[13] Only $\sim M^2$ arithmetic operations are required to construct the first M terms in the expansion. [It can be shown that if the fraction $A(z)$ converges as the number of terms increases, it converges to $\mathfrak{I}(z)$. However, we shall *not* need to assume convergence in order to obtain our error bounds.]

The first useful feature of the associated continued fraction is that by inspection we can test for condition (b): The given moments $\mu_0, \mu_1, \cdots, \mu_{2M}$ could have been derived from some nonnegative distribution $I(\omega)$ if and only if $\alpha_{2n}\alpha_{2n+1} \geq 0$, for $n = 0, 1, \cdots, M$.[14]

[10] J. A. Shohat and J. D. Tamarkin, *The Problem of Moments*, Mathematical Survey No. 1 (American Mathematical Society, Providence, R.I., 1950), Theorem 3.11, p. 103.

[11] Reference 10, Theorem 1.2, p. 5.

[12] Reference 10, p. 31.

[13] R. G. Gordon, J. Math. Phys. **9**, 655 (1968).

[14] Reference 10, Theorem 2.3, p. 31.

The associated continued fraction can also be used to test for condition (c), i.e., to find out whether a given number I_0 is a possible value for $\bar{I}(\omega_0, \epsilon)$ for some $I(\omega)$ which also satisfies conditions (a) and (b). To carry out this test, we form polynomials $P_n(z)$ and $Q_n(z)$ according to the pair of recursion relations ($n \geq 2$)

$$P_{n+1}(z) = (z + \alpha_{2n-1} + \alpha_{2n})P_n(z) - \alpha_{2n}\alpha_{2n+1}P_{n-1}(z), \tag{22a}$$

$$Q_{n+1}(z) = (z + \alpha_{2n-1} + \alpha_{2n})Q_n(z) - \alpha_{2n}\alpha_{2n+1}Q_{n-1}(z), \tag{22b}$$

starting from the initial values

$$P_1(z) = 0, \tag{23a}$$

$$P_2(z) = \alpha_1, \tag{23b}$$

$$Q_1(z) = 1, \tag{24a}$$

$$Q_2(z) = z + \alpha_2. \tag{24b}$$

These polynomials are closely related to the value of the associated fraction $A(z)$ truncated after M terms, which has the value

$$A_M(z) = P_M(z)/Q_M(z). \tag{25}$$

Following Hamburger,[15] we consider a more general approximant to $A(z)$,

$$C_{M+1}(z, \tau) = \cfrac{\alpha_1}{z + \alpha_2 - \cfrac{\alpha_2\alpha_3}{z + \alpha_3 + \alpha_4 - \cfrac{\alpha_4\alpha_5}{\ddots \cfrac{-\alpha_{2M-2}\alpha_{2M-1}}{z + \alpha_{2M-1} + \alpha_{2M} - \tau}}}}, \tag{26}$$

where τ is any real number. In terms of the polynomials $P_n(z)$ and $Q_n(z)$, we have

$$C_{M+1}(z, \tau) = \frac{P_{M+1}(z) - \tau P_M(z)}{Q_{M+1}(z) - \tau Q_M(z)}. \tag{27}$$

As special cases of this we have

$$C_{M+1}(z; \pm\infty) = A_M(z), \tag{28}$$

and

$$C_{M+1}(z; 0) = A_{M+1}(z). \tag{29}$$

As τ ranges over the real axis from $-\infty$ to ∞, $C_{M+1}(z; \tau)$ sweeps around a circle in the complex plane, which circle we denote $C_{M+1}(z)$. Evidently, from Eqs. (28) and (29), the circle $C_{M+1}(z)$ passes through the points $A_M(z)$ and $A_{M+1}(z)$. Now we can apply a theorem[15,16] which tells us three things:

(1) To each point on the circle $C_{M+1}(z)$ there exists one and only one distribution $I(\omega)$, which satisfies conditions (a) and (b), and for which

$$\bar{I}(\omega_0, \epsilon) = -\text{Im } C_{M+1}(z; \tau). \tag{30}$$

(2) To each point within the circle there exist an infinite number of such distributions $I(\omega)$.

(3) To each point outside the circle there exist no such distributions. Thus, the allowed range of $\bar{I}(\omega_0, \epsilon)$, as $I(\omega)$ ranges over all possible nonnegative distributions $I(\omega)$, which have the given $2M$ moments, is

$$\min_\tau [-\text{Im } C_{M+1}(z; \tau)] \leq \bar{I} \leq \max_\tau [-\text{Im } C_{M+1}(z; \tau)], \tag{31}$$

since the solutions on the circumference of the circle enclose those in the interior, and there are no solutions outside the circle.

The minimization and maximization problems of the functional $\bar{I}(\omega_0, \epsilon)$ over the space of functions $I(\omega)$ have now been reduced to the minimization and maximization of a single function $-\text{Im } C_{M+1}(z; \tau)$ of a single real variable τ. Geometrically, this problem amounts to no more than finding the top and bottom of the circle. From Eqs. (27) and (31) one finds readily that

$$\max \{\bar{I}_M(\omega_0, \epsilon)\} = \frac{[\text{Re } P_{M+1}(z) \text{ Re } Q_M(z) - \text{Re } Q_{M+1}(z) \text{ Re } P_M(z)]}{[\text{Im } Q_{M+1}(z) \text{ Re } Q_M(z) - \text{Re } Q_{M+1}(z) \text{ Im } Q_M(z)]}, \tag{32}$$

and

$$\min \{\bar{I}_M(\omega_0, \epsilon)\} = \frac{[\text{Im } P_{M+1}(z) \text{ Im } Q_M(z) - \text{Im } Q_{M+1}(z) \text{ Im } P_M(z)]}{[\text{Im } Q_{M+1}(z) \text{ Re } Q_M(z) - \text{Re } Q_{M+1}(z) \text{ Im } Q_M(z)]}, \tag{33}$$

where $z = -\omega_0 + i\epsilon$. A typical set of circles is plotted in Fig. 1.

From the method of construction, it is obvious that there are an infinite number of distributions $I(\omega)$ for which $\bar{I}(\omega_0, \epsilon)$ attains any value I_0 that we may choose in the range:

$$\min \{\bar{I}(\omega_0, \epsilon)\} < I_0 < \max \{\bar{I}(\omega_0, \epsilon)\}. \tag{34}$$

[15] H. Hamburger, Math. Ann. **81**, 235 (1920); **82**, 120, 168 (1921).
[16] Reference 10, Theorem 2.7, p. 48.

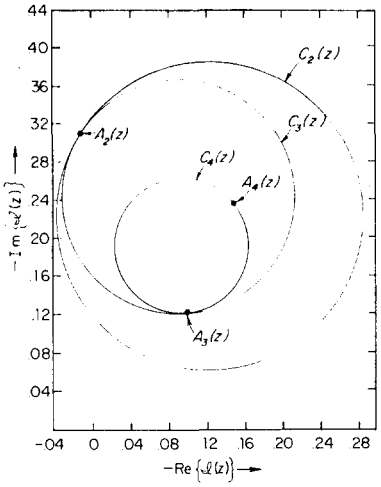


FIG. 1. An example of the circles $C_M(z)$ in the complex plane.

Thus the error bounds which we have obtained are the most precise, which we could hope to obtain for the response $\bar{I}(\omega_0, \epsilon)$, given only the equilibrium moments $\mu_0, \mu_1, \cdots, \mu_{2M}$.

The distributions $I(\omega)$ for which $\bar{I}(\omega_0, \epsilon)$ actually *attains* its maximum and minimum values, respectively, must be *unique* since they each correspond to a single point on the circumference of the circle $C_{M+1}(z)$ (the bottom and top of the circle, respectively). It is not necessary to actually construct these distributions, since we have already constructed the error bounds, Eqs. (32) and (33), without solving explicitly for the distributions. Nevertheless, it is interesting to construct these distributions. We will call the distribution, which maximizes the response $\bar{I}(\omega_0, \epsilon)$, the maximal distribution, and write it as $I_{\max}(\omega_0, \epsilon)$. Similarly, we define the minimal distribution $I_{\min}(\omega_0, \epsilon)$ to be that distribution, which minimizes $\bar{I}(\omega_0, \epsilon)$, subject to the constraints

$$I_{\max}(\omega_0, \epsilon) \geq 0, \tag{35a}$$

$$I_{\min}(\omega_0, \epsilon) \geq 0, \tag{35b}$$

and

$$\mu_n = \int_{-\infty}^{\infty} \omega^n I_{\max}(\omega, \epsilon)\, d\omega, \tag{36a}$$

$$\mu_n = \int_{-\infty}^{\infty} \omega^n I_{\min}(\omega, \epsilon)\, d\omega, \tag{36b}$$

for $n = 0, 1, 2, \cdots, 2M$.

The first information we need for these constructions are the values of τ which correspond to the top and bottom of the circle $C_{M+1}(z)$. These are

$$\tau_{\min} = \mathrm{Im}\, Q_{n+1}(z)/\mathrm{Im}\, Q_n(z), \tag{37}$$

$$\tau_{\max} = \mathrm{Re}\, Q_{n+1}(z)/\mathrm{Re}\, Q_n(z). \tag{38}$$

For these values of τ, $\mathfrak{J}(z)$ takes the values given by Eq. (26), when values of τ are substituted. If we can rewrite the resulting fraction as a sum of partial fractions,

$$\mathfrak{J}(z) \equiv \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{I(\omega)\, d\omega}{z+\omega} = \sum_{j=1}^{M} \frac{\rho_M(j)}{z+\xi_{jM}}. \tag{39}$$

Then we shall have constructed the maximal and minimal solutions as sums of δ functions, with weights $\rho_M(j)$, at positions $-\xi_{jM}$. These weights and positions are most easily calculated in terms of the eigenvectors and the (real) eigenvalues of a certain real symmetric tridiagonal matrix[13] of order M. Thus the ξ_{jM} are real, and the weights ρ are readily shown to be positive, as they must be if the maximal and minimal solutions are to be nonnegative. Thus the maximal and minimal solutions are uniquely given as a finite set of δ functions. Of course, there will in general be a different maximal and minimal solution for each ω_0 and ϵ one may choose.

Of particular interest is the maximal solution in which we let the width ϵ of the broadening function tend to zero

$$\lim_{\epsilon \to 0^+} I_{\max}(\omega_0, \epsilon) = I_{\max}(\omega_0, 0). \tag{40}$$

This is the solution which concentrates the greatest possible distribution at ω_0, consistent with the $2M$ known moments of the distribution. This special case of our results is essentially equivalent to the well-known Tchebycheff inequality.[17]

The methods developed here may be extended to give error bounds for the response to other forms of perturbation, such as a Gaussian or rectangular broadening function. However, the results will be qualitatively similar to those given here for Lorentzian broadening, so the more complicated constructions required for other forms of perturbation need not be elaborated here.

IV. APPLICATION TO NUCLEAR-MAGNETIC RESONANCE

As an example we consider the application of the results of the last section to calculating the line shape in the nuclear-magnetic resonance spectra of solids. For nuclei of spin one-half, the dominant interaction between the spins is just the magnetic-dipole–dipole interaction.[18] This interaction is suitable for our purposes, since the dipolar spin Hamiltonian may be estimated accurately from the atomic geometry of the crystal. To a very good approximation the nuclei in an ionic crystal may be assumed to be fixed at their

[17] P. Tchebycheff, J. Math. Pures Appl. (2) **19**, 157 (1874); Ref. 10, Lemma 2.13, p. 43.

[18] See, for example, A. Abragam, *The Principles of Nuclear Magnetism* (Oxford University Press, London, 1961), Ch. IV.

lattice positions, and the dipolar spin Hamiltonian calculated accordingly.

Van Vleck[5] has calculated the first four moments of the NMR spectrum for a dipolar interaction and has specialized the results to the case of the CaF_2 crystal. These theoretical moments are in good agreement with measurements[19] of the moments of the NMR spectrum.

We have adopted these four theoretical moments,[5,19] and also the sixth and eighth moments estimated by Abragam[18] from an analytic fit to the experimental curve. From these eight moments, we have constructed the upper and lower bounds, Eqs. (32) and (33), for several different resolutions, ϵ. A typical set of these upper and lower bounds are plotted as a function of frequency ω_0 in Fig. 2, for fixed $\epsilon = 1$. Units of frequency are $\mu_2^{\frac{1}{2}}$. Experimental curves[19] for the nuclear resonance of CaF_2, with the external field direction along the [111] crystal axis, were appropriately broadened by convolution with a Lorentzian curve

$$L(\omega) = \frac{\epsilon/\pi}{\epsilon^2 + (\omega - \omega_0)^2},$$

with width $\epsilon = 1$. These experimental points are plotted in Fig. 2, with an assumed experimental uncertainty of $\pm 10\%$ of the peak height. The agreement between theory and experiment is satisfactory. The comparison is significant, since there are no unknown parameters, and no theoretical approximations other than the choice of Hamiltonian for the system. The theory includes a rigorous bound for its possible error, and the theoretical error can be reduced by evaluating more terms in a well-defined sequence of approximations.

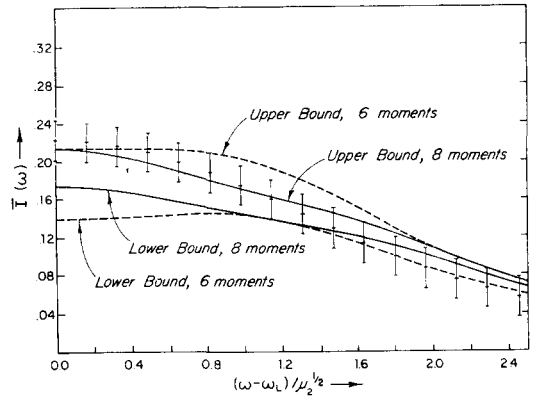


FIG. 2. Comparison of theory and experiment: shape of the nuclear magnetic resonance spectrum of CaF_2 crystal. See text for full explanation.

ACKNOWLEDGMENT

This work was supported in part by the National Science Foundation.

[19] C. R. Bruce, Phys. Rev. **107**, 43 (1957).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 7 JULY 1968

Construction of an Invariance from a Conservation Law, and Vice Versa, in Classical Mechanics

PHILIP PEARLE

Department of Physics, Case Western Reserve University, Cleveland, Ohio

(Received 13 June 1967)

Special classes of problems (each class is called a mechanics) are discussed for which it is either always possible to construct a conservation law when one is given an invariance of the equations of motion, or else it is always possible to construct an invariance when one is given a conservation law. Of chief interest is the introduction of a set of mechanics obeying the latter alternative in the form: (A) Given any constant of the motion, one can always construct a linear combination of its first derivatives (with fixed constant coefficients) which will be an invariance generator. Hamiltonian mechanics is, roughly speaking, "derived" as a consequence of (A) and the additional postulate: (B) One can always find a function from which, when one constructs the linear combination of its first derivatives, one obtains the motion generator. Another interesting mechanics satisfying (A) [but not (B)] is discussed. This is a class of harmonic-oscillator problems, called oscillator mechanics. It is shown that there is a class of problems belonging to both Hamiltonian mechanics and oscillator mechanics. This class of problems has an algebra that is the same as the algebra of quantum mechanics.

I. INTRODUCTION

For any Hamiltonian mechanics problem, there is a useful rule which permits one to construct an invariance (a one-parameter group of transformations which leave the equations of motion invariant) from any constant of the motion. This invariance will also be a group of canonical transformations. (For a familiar example, if a component of momentum is

lattice positions, and the dipolar spin Hamiltonian calculated accordingly.

Van Vleck[5] has calculated the first four moments of the NMR spectrum for a dipolar interaction and has specialized the results to the case of the CaF_2 crystal. These theoretical moments are in good agreement with measurements[19] of the moments of the NMR spectrum.

We have adopted these four theoretical moments,[5,19] and also the sixth and eighth moments estimated by Abragam[18] from an analytic fit to the experimental curve. From these eight moments, we have constructed the upper and lower bounds, Eqs. (32) and (33), for several different resolutions, ϵ. A typical set of these upper and lower bounds are plotted as a function of frequency ω_0 in Fig. 2, for fixed $\epsilon = 1$. Units of frequency are $\mu_2^{\frac{1}{2}}$. Experimental curves[19] for the nuclear resonance of CaF_2, with the external field direction along the [111] crystal axis, were appropriately broadened by convolution with a Lorentzian curve

$$L(\omega) = \frac{\epsilon/\pi}{\epsilon^2 + (\omega - \omega_0)^2},$$

with width $\epsilon = 1$. These experimental points are plotted in Fig. 2, with an assumed experimental uncertainty of $\pm 10\%$ of the peak height. The agreement between theory and experiment is satisfactory. The comparison is significant, since there are no unknown parameters, and no theoretical approximations other than the choice of Hamiltonian for the system. The theory includes a rigorous bound for its possible error, and the theoretical error can be reduced by evaluating more terms in a well-defined sequence of approximations.

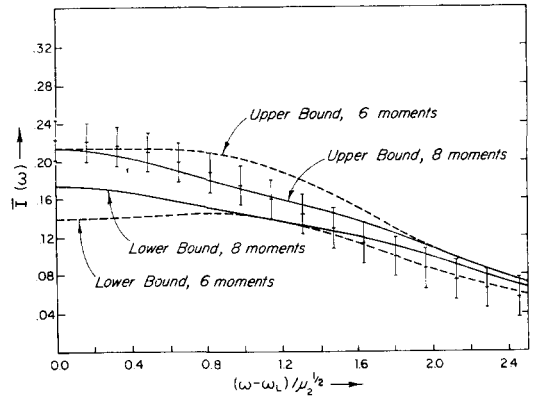


FIG. 2. Comparison of theory and experiment: shape of the nuclear magnetic resonance spectrum of CaF_2 crystal. See text for full explanation.

ACKNOWLEDGMENT

This work was supported in part by the National Science Foundation.

[19] C. R. Bruce, Phys. Rev. **107**, 43 (1957).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 7 JULY 1968

Construction of an Invariance from a Conservation Law, and Vice Versa, in Classical Mechanics

PHILIP PEARLE

Department of Physics, Case Western Reserve University, Cleveland, Ohio

(Received 13 June 1967)

Special classes of problems (each class is called a mechanics) are discussed for which it is either always possible to construct a conservation law when one is given an invariance of the equations of motion, or else it is always possible to construct an invariance when one is given a conservation law. Of chief interest is the introduction of a set of mechanics obeying the latter alternative in the form: (A) Given any constant of the motion, one can always construct a linear combination of its first derivatives (with fixed constant coefficients) which will be an invariance generator. Hamiltonian mechanics is, roughly speaking, "derived" as a consequence of (A) and the additional postulate: (B) One can always find a function from which, when one constructs the linear combination of its first derivatives, one obtains the motion generator. Another interesting mechanics satisfying (A) [but not (B)] is discussed. This is a class of harmonic-oscillator problems, called oscillator mechanics. It is shown that there is a class of problems belonging to both Hamiltonian mechanics and oscillator mechanics. This class of problems has an algebra that is the same as the algebra of quantum mechanics.

I. INTRODUCTION

For any Hamiltonian mechanics problem, there is a useful rule which permits one to construct an invariance (a one-parameter group of transformations which leave the equations of motion invariant) from any constant of the motion. This invariance will also be a group of canonical transformations. (For a familiar example, if a component of momentum is

conserved, one constructs the corresponding canonical group of translations which leaves the Hamiltonian, and thus the equations of motion, invariant.) Conversely, if one has an invariance, one can construct a constant of the motion from the group generator, provided these transformations are canonical. However, if the invariance does not consist of canonical transformations, then one cannot construct the constant of the motion.

For an arbitrary mechanics problem, there exists no rule enabling one to construct an invariance from a conservation law, nor a rule enabling one to construct a conservation law from an invariance. If there were such rules, then two different sets of equations of motion which have a constant of motion in common would necessarily have an invariance in common (and vice versa), but this is not the case. In an arbitrary mechanics problem there is generally no connection between a *single* constant of the motion and a *single* invariance.[1]

The interesting questions arise as to whether there are classes of problems other than Hamiltonian mechanics for which any conservation law begets an invariance, and whether there are classes of problems for which any invariance begets a conservation law (which is not the case with Hamiltonian mechanics). We shall point out that such classes of problems (we will call each class a "mechanics") do exist. For example, a class of harmonic-oscillator problems called "oscillator mechanics" will be discussed for which there are three different ways of constructing three (possibly different) constants of the motion from any given invariance generator, as well as a rule for constructing an invariance generator from any constant of the motion.

One may also inquire as to whether Hamiltonian mechanics is uniquely characterized by its relationship between a conservation law and an invariance. We shall, roughly speaking, "derive" Hamiltonian mechanics from the following two requirements:

(A) For each problem in the mechanics, one can always construct an invariance from a constant of the motion: The invariance generator will be a linear combination (with certain fixed constant coefficients) of the first derivatives of the conserved quantity.

(B) A particular function can always be found which will enable one to construct the *motion generator* in this way.

Other mechanics, notably the oscillator mechanics mentioned above, satisfy property (A) but not property (B).

As an interesting example, we shall investigate in detail the class of problems which is common to both oscillator mechanics and Hamiltonian mechanics. This class of problems, called "in-common" mechanics, is especially rich in its connections between invariances and conservation laws. In addition, we will show that it has an algebra that is the same as the algebra of quantum ("matrix") mechanics, a fact first noticed by Strocchi.[2]

The content of the various sections will now be outlined. Section II contains a brief summary, without proofs, of the needed relationships between equations of motion, invariance generators, and conserved quantities.[3] In Sec. III a few examples are given of mechanics for which constants of the motion (or constants) can always be constructed from any invariance generator. A presentation of oscillator mechanics follows in Sec. IV.

Section V begins our discussion of mechanics in which a conservation law determines an invariance. Restrictions on the class of problems satisfying requirement (A) (stated above) are derived. In Sec. VI it is shown that requirements (A) and (B) define Hamiltonian mechanics. Section VII points out that oscillator mechanics satisfies requirement (A). Lastly, Sec. VIII contains the discussion of in-common mechanics.

II. NECESSARY RELATIONSHIPS

In this section we shall briefly review the relationships which will be needed in the following sections: the proofs can be found elsewhere.[3]

The equations of motion for a system of n degrees of freedom are

$$dx^i/dt = b^i(x^1, \cdots, x^n), \quad i = 1, \cdots, n. \tag{1}$$

One can find $n-1$ independent constants of the motion $\Phi^\alpha(x^1, \cdots, x^n)$ $(\alpha = 1, \cdots, n-1)$, which do not depend upon the time variable, and which therefore satisfy the partial differential equation

$$d\Phi^\alpha/dt = \mathbf{b} \cdot \nabla \Phi^\alpha = 0 \tag{2}$$

$(\mathbf{b} \cdot \nabla \equiv \sum_{i=1}^n b^i \partial/\partial x^i)$. In addition, we introduce another constant of the motion which depends linearly on time, $T(x^1, \cdots, x^n) - t$. The function T must satisfy

$$\mathbf{b} \cdot \nabla T = 1. \tag{3}$$

The n functions Φ^α, T are independent, but they are not unique.

[1] E. P. Wigner, Progr. Theoret. Phys. (Kyoto) **11**, 437 (1954); R. Houtappel, H. Van Dam, and E. Wigner, Rev. Mod. Phys. **37**, 595 1965).

[2] F. Strocchi, Rev. Mod. Phys. **38**, 36 (1966). This connection between a classical mechanical system and the algebra of quantum mechanics was also found independently by the author; he is indebted to Professor W. Tobocman for bringing Strocchi's paper to his attention. The approach presented here is quite different from that of Strocchi.

[3] L. P. Eisenhart, *Continuous Groups of Transformations* (Dover Publications, Inc., New York, 1961); see especially Sec. 25.

If a one-parameter set of transformations

$$\bar{x}^i = X^i(x^1, \cdots, x^n; a) \tag{4}$$

("a" is the parameter) is also a one-parameter group of transformations, then Eq. (4) can also be written, in an obvious notation, as

$$\bar{x}^i = e^{a\mathbf{g}\cdot\nabla} x^i, \tag{5}$$

where $g^i(x^1, \cdots, x^n) \equiv \partial X^i/\partial a|_{a=0}$. The linear operator $\mathbf{g}\cdot\nabla$ is called the group generator.

If the equations of motion (1) are expressed in terms of the variables $\bar{x}^i$, they will have the form $d\bar{x}^i/dt = \bar{b}^i(\bar{x}^i, \cdots, \bar{x}^n)$. If the functions $\bar{b}^i$ satisfy the relationship $\bar{b}^i(\bar{x}^i, \cdots, \bar{x}^n) = b^i(\bar{x}^i, \cdots, \bar{x}^n)$ for all a, then the equations of motion are invariant under the one-parameter group of transformations. It can be shown that the necessary and sufficient condition for $\mathbf{g}\cdot\nabla$ to generate an invariance of the equations of motion characterized by the vector $\mathbf{b}$ is

$$[\mathbf{b}\cdot\nabla, \mathbf{g}\cdot\nabla] = 0. \tag{6}$$

(This is the commutator bracket.) Equation (6) is equivalent to the set of n first-order partial differential equations for the components g^j:

$$\mathbf{b}\cdot\nabla g^j = \mathbf{g}\cdot\nabla b^j. \tag{7}$$

The most general invariance generator satisfying Eq. (6) [or Eq. (7)] can be constructed as follows. First we pick any $n-1$ independent constants of the motion Φ^α, and a function T. We invert the n equations $\phi^\alpha = \Phi^\alpha(x)$, $\phi^n = T(x)$ to obtain $x^i = X^i(\phi^1, \cdots, \phi^n)$. It can be shown that each of the following transformations are invariances:

$$\bar{x}^i = X^i(\phi^1 + a, \phi^2, \cdots, \phi^n), \tag{8_1}$$

$$\bar{x}^i = X^i(\phi^1, \phi^2 + a, \cdots, \phi^n), \tag{8_2}$$

$$\vdots$$

$$\bar{x}^i = X^i(\phi^1, \phi^2, \cdots, \phi^n + a). \tag{8_n}$$

The invariance generator for the αth invariance is

$$\mathbf{g}_\alpha\cdot\nabla \equiv \sum_{i=1}^{n} (\partial x^i/\partial\phi^\alpha)\partial/\partial x^i = \partial/\partial\phi^\alpha, \tag{9}$$

In particular, the nth invariance generator constructed in this fashion is also the motion generator, for

$$g_n^i = \partial x^i/\partial\phi^n = b^i. \tag{10}$$

These n invariance generators all commute:

$$[\mathbf{g}_\alpha\cdot\nabla, \mathbf{g}_\beta\cdot\nabla] = 0 \quad (\alpha, \beta = 1, \cdots, n). \tag{11}$$

Finally, it can be shown that the most general invariance generator $\mathbf{g}\cdot\nabla$ is a linear combination of these n invariance generators:

$$\mathbf{g}\cdot\nabla = \sum_{\alpha=1}^{n} h^\alpha(\Phi^1(x), \cdots, \Phi^{n-1}(x))\mathbf{g}_\alpha\cdot\nabla, \tag{12}$$

where the h^α are n arbitrary functions of the $n-1$ constants of the motion.

III. MECHANICS WHERE AN INVARIANCE DETERMINES A CONSERVATION LAW

In this section, we shall present a few examples of mechanics in which constants of the motion can be constructed from invariances. We may demand of such a mechanics that it fulfill these rough requirements:

1. The connection between an invariance and a conservation law must be "simple" in some sense.
2. The connection must hold for all invariances in any problem belonging to the mechanics.
3. The mechanics must hold for an infinite number of problems for each value of n, and for an infinity of different values of n.

For our first example, let us look for a mechanics in which the existence of any invariance generator $\mathbf{g}\cdot\nabla$ implies that $\sum_{i=1}^n g^i(x)$ is a constant of the motion. By summing Eq. (7) over the index j, we find that

$$\mathbf{b}\cdot\nabla\sum_{j=1}^{n} g^j = \mathbf{g}\cdot\nabla\sum_j b^j. \tag{13}$$

Since for this mechanics $\sum_j g^j$ is a constant of the motion, the left-hand side of this equation vanishes. By our second requirement for a mechanics, the right-hand side must then vanish for *any* generator. This leads to the *necessary* condition that $\sum_j b^j(x)$ must be a constant. By putting $\sum_j b^j(x) = \text{const}$ back into Eq. (13), we see that this is a *sufficient* condition for any $\sum_j g^j$ to be a constant of the motion *or a constant.* [In fact, we know that at least one invariance generator yields a constant, since $\sum_j b^j(x)$ is a constant.]

In a similar manner, one may look for mechanics in which

$$(\alpha)\ \mathbf{g}\cdot\mathbf{x} \quad \text{or} \quad (\beta)\ \mathbf{g}\cdot\mathbf{b} \quad \text{or} \quad (\gamma)\ \nabla\cdot\mathbf{g} \tag{14}$$

are constants of the motion. By applying suitable operations to Eqs. (7), one finds the following *necessary* conditions for each of these three different mechanics:

$$(\alpha)\ \mathbf{b}\cdot\mathbf{x} = \text{const},$$

$$(\beta)\ (\mathbf{b}\cdot\nabla)\mathbf{b} = -\nabla\frac{(\mathbf{b}\cdot\mathbf{b})}{2}, \tag{15}$$

$$(\gamma)\ \nabla\cdot\mathbf{b} = \text{const}.$$

By putting Eqs. (15) back into Eq. (7), one finds that these are *sufficient* conditions for the quantities in Eqs. (14) to be constants of the motion, or constants.

Each **b** vector corresponding to a problem in any one of the above three mechanics satisfies a restriction [e.g., Eq. (15)]. The restriction nevertheless admits an infinity of problems, so our third requirement for a mechanics is satisfied. The conserved quantity constructed out of an invariance generator **g** [e.g., Eqs. (14)] is linear in the components of **g**, and we may regard this as fulfilling the first requirement of simplicity. The second requirement is fulfilled because a quantity whose time derivative vanishes is constructed out of *any* invariance generator. However, some of these quantities are constants, and while, strictly speaking, a constant is a conserved quantity, it is not a very useful conserved quantity.

In the next section we will describe a mechanics for which it is always possible to construct a conserved quantity, which is *not* a constant, out of any invariance generator.

IV. OSCILLATOR MECHANICS

We consider a class of harmonic-oscillator problems for which the equations of motion are

$$dx^i/dt = \sum_{j=1}^{n} x^j B_{ji}, \tag{16}$$

where the matrix elements B_{ji} are constants and the matrix B is an antisymmetric nonsingular matrix ($B_{ji} = -B_{ij}$, $\det B \neq 0$). We will find this mechanics to be of special interest in the next sections. It is not a merely academic mechanics, for it will later be pointed out that a subset of this mechanics has an algebra which is identical to the algebra of quantum mechanics. In this section, we will show that this mechanics is remarkably rich in connections between an invariance generator $\mathbf{g}\cdot\nabla$ and conservation laws: $\mathbf{g}\cdot\mathbf{x}$, $\mathbf{g}\cdot\mathbf{b}$, and $\mathbf{g}\cdot\mathbf{g}$ are each constants of the motion or constants! Furthermore, we will see that they cannot all be constants, so that a suitable function of all three [such as $(\mathbf{g}\cdot\mathbf{x})^2 + (\mathbf{g}\cdot\mathbf{b})^2 + (\mathbf{g}\cdot\mathbf{g})^2$] can be constructed which will always be a constant of the motion.

We begin by constructing Φ^α, T, and $\mathbf{g}\cdot\nabla$ for the equations of motion (16). Constants of the motion must satisfy Eq. (2):

$$\sum_{i,j} x^j B_{ji} \frac{\partial}{\partial x^i}\Phi = 0$$

and they can immediately be found, in terms of the eigenvectors of the matrix B. It is well known that a nonsingular antisymmetric matrix must be of even order ($n/2$ is therefore an integer), and must possess purely imaginary nonzero eigenvalues $\pm i\omega_\alpha$ ($\alpha = 1, \cdots, n/2$). We shall define $\omega_{-\alpha} \equiv -\omega_\alpha$ so that we may number the eigenvalues from $\alpha = -n/2$ to $\alpha = +n/2$. The eigenvector $\mathbf{e}_\alpha$ corresponding to the eigenvalue $i\omega_\alpha$ satisfies these equations[4]:

$$B\mathbf{e}_\alpha = i\omega_\alpha \mathbf{e}_\alpha, \quad \mathbf{e}_\alpha B = -i\omega_\alpha \mathbf{e}_\alpha,$$
$$\mathbf{e}_\alpha^* = \mathbf{e}_{-\alpha}, \quad \mathbf{e}_\alpha^* \cdot \mathbf{e}_\beta = \delta_{\alpha\beta}.$$

B_{ij}, written in terms of its eigenvectors and eigenvalues, looks like this:

$$B_{ij} = \sum_{\alpha=-n/2}^{n/2} i\omega_\alpha e_\alpha^i e_\alpha^{*j}.$$

It can readily be demonstrated that

$$\Phi^{\alpha\beta} \equiv \frac{(\mathbf{x}\cdot\mathbf{e}_\alpha)^{\omega_\alpha}}{(\mathbf{x}\cdot\mathbf{e}_\beta)^{\omega_\beta}}, \quad \beta \neq \alpha = \frac{-n}{2}, \cdots, \frac{+n}{2} \tag{17}$$

are all (mostly complex) constants of the motion, and there are exactly $n-1$ independent functions among these. A particularly interesting set of $n/2$ real constants of the motion are $(\Phi^{-\alpha,\alpha})^{1/\omega_\alpha}$, because they are real quadratic forms:

$$(\Phi^{-\alpha,\alpha})^{1/\omega_\alpha} = (\mathbf{x}\cdot\mathbf{e}_\alpha)(\mathbf{e}_\alpha^*\cdot\mathbf{x}) \quad \left(\alpha = 1, \cdots, \frac{n}{2}\right). \tag{18}$$

$n/2 - 1$ further real constants of the motion can be constructed from the $\Phi^{\alpha\beta}$, e.g., $\Phi^{\alpha\beta} + \Phi^{\alpha\beta *}$, $i(\Phi^{\alpha\beta} - \Phi^{\alpha\beta *})$, or $\Phi^{\alpha\beta}\Phi^{\alpha\beta *}$.

Functions satisfying $\mathbf{b}\cdot\nabla\, T = 1$ are

$$T^\alpha = \frac{1}{i\omega_\alpha}\ln \mathbf{e}_\alpha^*\cdot\mathbf{x}, \quad \text{or} \quad \frac{T^\alpha + T^{-\alpha}}{2} = \frac{1}{2i\omega_\alpha}\ln\frac{\mathbf{e}_\alpha^*\cdot\mathbf{x}}{\mathbf{e}_\alpha\cdot\mathbf{x}},$$

the latter set being real.

A complete commuting set of (complex) invariance generators are

$$\mathbf{g}_\alpha\cdot\nabla \equiv (\mathbf{x}\cdot\mathbf{e}_\alpha)(\mathbf{e}_\alpha^*\cdot\nabla) \quad \left(\alpha = \frac{-n}{2}, \cdots, \frac{n}{2}\right), \tag{19}$$

as may be verified by direct computation. An arbitrary real generator may therefore be written as [Eq. (12)]

$$\mathbf{g}\cdot\nabla = \sum_{\alpha=-n/2}^{n/2} h^\alpha(\Phi^{\beta\gamma})\mathbf{g}_\alpha\cdot\nabla, \quad h^{*\alpha} = h^{-\alpha},$$

of which a particular case is $\mathbf{b}\cdot\nabla$:

$$\mathbf{b}\cdot\nabla = \sum_{\alpha=-n/2}^{n/2} i\omega_\alpha \mathbf{g}_\alpha\cdot\nabla. \tag{20}$$

Now we are in a position to see that $\mathbf{g}\cdot\mathbf{x}$, $\mathbf{g}\cdot\mathbf{b}$, $\mathbf{g}\cdot\mathbf{g}$ are all functions of constants of the motion. $\mathbf{g}_\alpha\cdot\mathbf{x} = \mathbf{g}_{-\alpha}\cdot\mathbf{x}$ are the constants of the motion displayed in Eq. (18), so therefore [using $\mathbf{g}_\alpha\cdot\mathbf{g}_\beta = (\mathbf{g}_\alpha\cdot\mathbf{x})\delta_{\alpha,-\beta}$]

$$\mathbf{g}\cdot\mathbf{x} = \sum_{\alpha=0}^{n/2}(h^{\alpha^*} + h^\alpha)(\mathbf{g}_\alpha\cdot\mathbf{x}), \tag{21a}$$

$$\mathbf{g}\cdot\mathbf{b} = \sum_{\alpha=0}^{n/2} i\omega_\alpha(h^{\alpha^*} - h^\alpha)(\mathbf{g}_\alpha\cdot\mathbf{x}), \tag{21b}$$

$$\mathbf{g}\cdot\mathbf{g} = 2\sum_{\alpha=0}^{n/2}|h^\alpha|^2(\mathbf{g}_\alpha\cdot\mathbf{x}) \tag{21c}$$

[4] * refers to complex conjugation; $\delta_{\alpha\beta}$ is the Kronecker delta function; the ith component of the column vector $B\mathbf{e}_\alpha$ is $\sum_{j=1}^n B_{ij}e_\alpha^j$.

are all functions of constants of the motion, which means that they are either constants of the motion or constants.

It might be suspected that not *all* of the equations (21) can be constants, regardless of what functions h^α are chosen. This suspicion is correct, as is proven in Appendix A. Therefore, for this mechanics, we may state the rule that, from any invariance generator $\mathbf{g}\cdot\nabla$, the following quantity may be constructed:

$$(\mathbf{g}\cdot\mathbf{x})^2 + (\mathbf{g}\cdot\mathbf{b})^2 + (\mathbf{g}\cdot\mathbf{g})^2,$$

and it will always be a constant of the motion. (We do not choose $\mathbf{g}\cdot\mathbf{x} + \mathbf{g}\cdot\mathbf{b} + \mathbf{g}\cdot\mathbf{g}$ because, for example, we may set each $h = -1$; this results in $\mathbf{g}\cdot\mathbf{b} = 0$, $\mathbf{g}\cdot\mathbf{x} = -\mathbf{g}\cdot\mathbf{g}$, and the sum will vanish identically.)

V. MECHANICS WHERE A CONSERVATION LAW DETERMINES AN INVARIANCE

One can imagine many different rules for constructing a generator $\mathbf{g}\cdot\nabla$ out of a known constant of the motion Φ, e.g., $g^i \equiv \partial(\sum_j \partial\Phi/\partial x^j)/\partial x^i$ or $g \equiv \int_0^{x^i} \Phi(x^1, \cdots, x^i, \cdots, x^n)dx^i$ or $g^i \equiv \partial^2\Phi/\partial\mathbf{x}^{i2}$. However, we will focus our attention, in the remainder of this paper, on what appears to be the "simplest" available rule. We shall look for mechanics in which an invariance generator $\mathbf{g}\cdot\nabla$ may be constructed from any constant of the motion Φ according to the rule[5]

$$g^i\frac{\partial}{\partial x^i} = M^{ik}\frac{\partial\Phi}{\partial x^k}\frac{\partial}{\partial x^i}, \tag{22}$$

where the M^{ik} are to be constants, fixed for all of the problems involving n degrees of freedom in a particular mechanics. Because of the restriction to this class of problems, our three requirements for a mechanics become:

(1) The connection between invariance and conservation law is to be given by Eq. (22).

(2) For any problem, any constant of the motion $\Phi(x)$ as well as any function $T(x)$ must generate an invariance.[6]

(3) (a) The $n \times n$ matrix M must be nonsingular.

(b) The matrix M should be characterized by a property that is independent of n, and defined for an infinity of values of n.

Requirements (2) and (3a) insure that we will be able to construct a complete set of generators from a functionally independent set of constants of the motion. For if we construct n generators $\mathbf{g}_\alpha\cdot\nabla$ from n independent functions $\Phi^1, \cdots, \Phi^{n-1}$, $T \equiv \Phi^n$, the matrix Θ whose elements are $\Theta_{k\alpha} \equiv \partial\Phi^\alpha/\partial x^k$ will be nonsingular. Then the matrix G whose elements are $G_{i\alpha} \equiv g^i_\alpha$ will also be nonsingular (since $\det G = \det M \cdot \det\Theta \neq 0$) so the $\mathbf{g}_\alpha$ will be linearly independent.

Requirement (3b) enables us to say that problems with different values of n belong to the same mechanics.

To see what constraints Eq. (22) imposes, we substitute this expression for g^i into Eqs. (7) and obtain after some manipulation

$$\frac{\partial}{\partial x^k}\left[M^{jk}b^i\frac{\partial\Phi}{\partial x^i}\right] = M^{jk}\frac{\partial b^i}{\partial x^k}\frac{\partial\Phi}{\partial x^i} + \frac{\partial b^j}{\partial x^i}M^{ik}\frac{\partial\Phi}{\partial x^k}. \tag{23}$$

The left-hand side of Eq. (23) vanishes, since $\Phi(x)$ satisfies $\mathbf{b}\cdot\nabla\Phi(x) = \text{const}$. Designating by B the matrix whose elements are $B_{ki} \equiv \partial b^i/\partial x^k$, and rewriting Eq. (23) n times with n independent functions Φ^α, we obtain succinctly

$$MB\Theta + \tilde{B}M\Theta = 0.$$

Since Θ is a nonsingular matrix, we may multiply this equation on the right by Θ^{-1}, arriving at

$$MB + \tilde{B}M = 0. \tag{24}$$

Equation (24) is a necessary restriction that the $\mathbf{b}$ vector of a problem must satisfy if it is to belong to the mechanics characterized by a known matrix M [that is, if it is to be always possible to construct an invariance generator from any solution of $\mathbf{b}\cdot\nabla\Phi = \text{const}$, according to the rule given in Eq. (22)]. It is also a sufficient restriction, as can be seen by reversing the order of the above arguments.

Equation (24) may also be regarded as a restriction imposed on a matrix M by a known class of $\mathbf{b}$ vectors.

Before investigating specific mechanics which satisfy Eq. (24), we shall derive another relationship between B and M. This is a necessary consequence of Eq. (24), which may be written as

$$\frac{\partial b^j}{\partial x^i} = -M^{-1}{}_{ir}\frac{\partial b^r}{\partial x^s}M^{sj}. \tag{25}$$

We can differentiate Eq. (25) with respect to x^k and utilize the symmetrical relationship $\partial^2 b^j/\partial x^k\partial x^i = \partial^2 b^j/\partial x^i\partial x^k$ to obtain

$$\frac{\partial^2 b^j}{\partial x^k\partial x^i} = -M^{-1}{}_{ir}\frac{\partial^2 b^r}{\partial x^k\partial x^s}M^{sj} = -M^{-1}{}_{kr}\frac{\partial^2 b^r}{\partial x^i\partial x^s}M^{sj}.$$

Upon postmultiplication by $M^{-1}{}_{jl}$, this equation becomes

$$\frac{\partial}{\partial x^l}\left(M^{-1}{}_{ir}\frac{\partial b^r}{\partial x^k} - M^{-1}{}_{kr}\frac{\partial b^r}{\partial x^i}\right) = 0. \tag{26}$$

Equation (26) says that $M^{-1}\tilde{B} - B\tilde{M}^{-1}$ must be a

[5] We are using the convention that repeated Latin indices are to be summed over.

[6] This means that any solution of $\mathbf{b}\cdot\nabla\Phi = \text{const}$ will produce an invariance generator. If the constant is nonzero, $\Phi/\text{const} \equiv T$ will satisfy $\mathbf{b}\cdot\nabla T = 1$. Then, if T produces an invariance generator [according to the rule in Eq. (22)], const $\times\, T = \Phi$ will also produce this invariance generator (multiplied by the constant).

constant matrix. After premultiplying this matrix by M and postmultiplying by M, we arrive at a necessary condition:

$$MB - \tilde{B}\tilde{M} = C, \tag{27}$$

where C, a constant matrix, must obviously be antisymmetric.

Equation (27) may be added to the transpose of Eq. (24) to yield the interesting relation

$$(M + \tilde{M})B = C. \tag{28}$$

Equation (28) can be a severe restriction on the functional dependence of $\mathbf{b}$ upon x, for it says that the symmetric part of the *constant* matrix M, multiplied by the matrix B *which depends upon* x, must equal some *constant* (antisymmetric) matrix. However, it is no restriction at all if the matrix M is antisymmetric, since then Eq. (28) is identically satisfied by any B (with $C = -\tilde{C} = 0$). It will be shown in the next section that if M is antisymmetric, the associated mechanics is equivalent (to within a linear transformation of coordinates) to Hamiltonian mechanics. Following this, in Sec. VII, we will consider other mechanics where M is not antisymmetric.

VI. HAMILTONIAN-EQUIVALENT MECHANICS

Suppose we have a class of $\mathbf{b}$ vectors forming a mechanics, so that Eq. (24) is satisfied, and M is known. Although an invariance generator may always be constructed [Eq. (22)] from any solution of $\mathbf{b} \cdot \nabla\Phi = \text{const}$, one will often be unable to construct the *particular* invariance generator $\mathbf{b} \cdot \nabla$. This leads us to ask: What are the mechanics in which a function $H(x)$ can always be found, enabling us to construct the invariance generator $\mathbf{b} \cdot \nabla$ itself according to the rule

$$b^i \frac{\partial}{\partial x^i} = M^{ik} \frac{\partial H(x)}{\partial x^k} \frac{\partial}{\partial x^i}. \tag{29}$$

The answer will be that, for such a mechanics, it is necessary and sufficient that M be antisymmetric. When M is antisymmetric, $\mathbf{b} \cdot \nabla H(x) = 0$ [by Eq. (29)], so that $H(x)$ is in fact a constant of the motion. It will be seen that all of these mechanics are "equivalent" to Hamiltonian mechanics in the following sense.

When we say that two mechanics are equivalent, we shall mean that any problem in one mechanics can be transformed into a problem in the other mechanics by a linear transformation

$$x'^r = A^{rs}x^s + c^r$$

(A^{rs}, c^r are constants). Under this transformation, various quantities in the "unprimed" mechanics are related to the same quantities in the "primed" mechanics by

$$b'^r = A^{rs}b^s, \quad g'^r = A^{rs}g^s,$$
$$B' = \tilde{A}^{-1}B\tilde{A}, \quad M' = AM\tilde{A}. \tag{30}$$

We will now show that requirements 1, 2, and 3, together with the additional requirement of Eq. (29), define Hamiltonian-equivalent mechanics. Since $\mathbf{b}$ is given by Eq. (29), the matrix elements of B are

$$B_{ki} \equiv \frac{\partial b^i}{\partial x^k} = M^{ij} \frac{\partial^2 H}{\partial x^k \partial x^i},$$

and Eq. (24) can be written as

$$M^{kl} \frac{\partial^2 H}{\partial x^l \partial x^j} \tilde{M}^{ji} + M^{kl} \frac{\partial^2 H}{\partial x^l \partial x^j} M^{ji} = 0. \tag{31}$$

Since M is nonsingular, and if at least one problem in the class has a nonvanishing Hessian determinant $|\partial^2 H/\partial x^l \partial x^j| \neq 0$, we may premultiply Eq. (31) by M^{-1} and by the inverse of the matrix whose elements are $\partial^2 H/\partial x^l \partial x^j$ to obtain

$$M + \tilde{M} = 0. \tag{32}$$

Thus it is necessary that M be antisymmetric.

We can also prove that in a mechanics where M is antisymmetric, $\mathbf{b}$ can always be constructed from some scalar function $H(x)$, according to the rule given in Eq. (29). We may set $b^j(x) \equiv M^{js}h^s(x)$ [this defines the n functions $h^s(x)$] without loss of generality, because M is a nonsingular matrix. Equation (24) (which is satisfied since $\mathbf{b}$ is in the mechanics) tells us that

$$M^{ir} \frac{\partial h^s}{\partial x^r} \tilde{M}^{sj} + M^{ir} \frac{\partial h^r}{\partial x^s} M^{sj} = 0. \tag{33}$$

After setting $\tilde{M}^{sj} = -M^{sj}$, we may premultiply and postmultiply Eq. (33) by M^{-1}, which yield

$$\frac{\partial h^s}{\partial x^r} = \frac{\partial h^r}{\partial x^s}. \tag{34}$$

Equations (34) are the well-known conditions which are necessary and sufficient for $\mathbf{h}(x) = \nabla H(x)$ [$H(x)$ is some scalar function], and our proof is complete.

To show that all mechanics for which M is antisymmetric are equivalent to Hamiltonian mechanics, we remark that any nonsingular antisymmetric matrix must be of even order and may be brought by a congruence transformation to the block form:

$$J \equiv AM\tilde{A} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \tag{35}$$

(where 1 is the $n/2 \times n/2$ identity matrix). *But Eq. (29)*

with M replaced by the matrix J defines Hamiltonian mechanics.[7] To see this, we identify $x^1, \cdots, x^{n/2}$ with $q^1, \cdots, q^{n/2}$ and we identify $x^{n/2+1}, \cdots, x^n$ with $p^1, \cdots, p^{n/2}$. Then, for example, using Eq. (29), we see that

$$\frac{dF(x)}{dt} = \mathbf{b} \cdot \nabla F = \frac{\partial F(x)}{\partial x^i} J^{ij} \frac{\partial H(x)}{\partial x^j} \equiv \{F(x), H(x)\}, \tag{36}$$

where $\{F, H\}$ is the familiar Poisson bracket expression. The condition [Eq. (6)] that $\mathbf{g} \cdot \nabla$ [Eq. (22)] must generate an invariance becomes

$$0 = [\mathbf{g} \cdot \nabla, \mathbf{b} \cdot \nabla] f(x) = \frac{\partial}{\partial x^i}\left(\frac{\partial \Phi}{\partial x^r} J^{rs} \frac{\partial H}{\partial x^s}\right) J^{ij} \frac{\partial f(x)}{\partial x^j}$$
$$\equiv \{\{\Phi, H\}, f(x)\}, \tag{37}$$

i.e., $\mathbf{g} \cdot \nabla$ constructed from Φ will be an invariance generator in Hamiltonian mechanics, if $\{\Phi, H\} =$ const, as is well known.

Thus we have shown that either the antisymmetry of M or the condition that $\mathbf{b} \cdot \nabla$ be constructible from a scalar function [according to the rule in Eq. (29)] are each necessary and sufficient for a mechanics to be Hamiltonian-equivalent.

VII. MECHANICS WHERE M IS SYMMETRIC

Equation (28) provides the strongest possible restriction when M is symmetric (we shall not discuss the general case of mechanics for which M is neither symmetric nor antisymmetric). Any mechanics where M is symmetric is equivalent to a mechanics where M is a diagonal matrix with ± 1 as diagonal matrix elements (any nonsingular symmetric matrix can be brought to such a form by a congruence transformation). When M is such a diagonal matrix, Eq. (28) can be solved for B (using $\tilde{M} = M$, $M^2 = 1$):

$$B = MC/2. \tag{38}$$

This expression for B identically satisfies Eq. (24).

Since B must be a constant matrix, it follows that $b^i(x)$ can only depend linearly upon x. These problems will therefore have solutions $x^i(t)$ which can oscillate sinusoidally or exponentially increase or decrease with time.

The mechanics where $M = 1$ is particularly interesting. An invariance generator $\mathbf{g} \cdot \nabla$ is very simply constructed from a conserved quantity Φ: $\mathbf{g} = \nabla\Phi$. From Eq. (38), we see that B must be an antisymmetric matrix. Of course, no function $H(x)$ can be found for which $\mathbf{b} = \nabla H(x)$, since this is not a Hamiltonian-equivalent mechanics.

Finally we focus our attention on the mechanics where $M = 1$ and B is nonsingular (and therefore of even order, since B is antisymmetric). This is just the oscillator mechanics which we discussed in Sec. IV! These problems have only oscillatory solutions, essentially because the roots of a nonsingular antisymmetric matrix like B are nonvanishing pure imaginary numbers. Now we see that, for any problem in this mechanics, not only can constants of the motion always be constructed from any invariance generator (since at least one of $\mathbf{g} \cdot \mathbf{x}$, $\mathbf{g} \cdot \mathbf{b}$, $\mathbf{g} \cdot \mathbf{g}$ is a constant of the motion), but also an invariance generator can always be constructed from any constant of the motion (by the rule $\mathbf{g} = \nabla\Phi$). Certainly this oscillator mechanics is rich in its connections between invariances and conservation laws.

VIII. "IN COMMON" MECHANICS

For our final investigation, we shall inquire as to whether Hamiltonian mechanics and the oscillator mechanics possess any problems in common. We shall find that such an "in common" mechanics (as we shall call it) is possible—and much of its algebra is the algebra of quantum ("matrix") mechanics (in a complex $n/2$ dimensional space and in the "Schrödinger picture").

The problems of in common mechanics must be harmonic-oscillator problems, for which $b^i = J^{ik}\partial H(x)/\partial x^k$ is a linear function in x. $H(x)$ must therefore be a quadratic form in x:

$$H(x) \equiv \tfrac{1}{2} x^r H_{rs} x^s;$$
$$H_{rs} = H_{sr}, \quad \frac{\partial H_{rs}}{\partial x^i} = 0 \quad (\text{all } i, r, s). \tag{39}$$

From the requirement that the problems lie in Hamiltonian mechanics, we see that

$$b^i = J^{ik} H_{kj} x^j, \quad B_{ji} \equiv \frac{\partial b^i}{\partial x^j} = -H_{jk} J^{ki}. \tag{40}$$

From the requirement that the problems lie in oscillator mechanics, we see that B must be antisymmetric: $HJ + \widetilde{(HJ)} = 0$. (Here we have written H to denote the constant matrix whose elements are H_{jk}.) Using the symmetry of H and the antisymmetry of J, this restriction upon H becomes

$$[J, H] = 0. \tag{41}$$

Equation (41) is the necessary (and sufficient)

[7] C. Kilmister, *Hamiltonian Mechanics* (John Wiley & Sons, Inc., New York, 1964), Chap. 4. The author was fortunate enough to be able to read a set of notes by E. Wichmann from a course in mechanics which was given at Berkeley in 1959, and it is Wichmann's notation that is followed here.

condition for a problem to belong to both Hamiltonian and oscillator mechanics. If we write the symmetric matrix H in block form

$$H \equiv \begin{bmatrix} H_{11} & \tilde{H}_{21} \\ H_{21} & H_{22} \end{bmatrix}, \quad \tilde{H}_{11} = H_{11}, \quad \tilde{H}_{22} = H_{22}, \quad (42)$$

(the H_{11}, H_{21}, H_{22} are $n/2 \times n/2$ matrices,) we find that satisfaction of Eq. (41) requires $H_{11} = H_{22}$ and $\tilde{H}_{21} = -H_{21}$. We will rename the symmetric matrix H_{11}, calling it H^R, and rename the antisymmetric matrix H_{21}, calling it H^I (the reason for this notation will appear shortly). Then the symmetric matrix H and the antisymmetric matrix B can be written as

$$H = \begin{bmatrix} H^R & -H^I \\ H^I & H^R \end{bmatrix}; \quad B = -\begin{bmatrix} H^I & H^R \\ -H^R & H^I \end{bmatrix}; \quad \tilde{H}^R = H^R, \quad \tilde{H}^I = -H^I, \quad (43)$$

while the equations of motion appear in matrix form as

$$\frac{d}{dt}\begin{bmatrix} x^1 \\ \cdot \\ \cdot \\ \cdot \\ x^n \end{bmatrix} = \begin{bmatrix} H^I, & H^R \\ & \\ & \\ -H^R, & H^I \end{bmatrix}\begin{bmatrix} x^1 \\ \cdot \\ \cdot \\ \cdot \\ x^n \end{bmatrix}. \quad (44)$$

The constants of the motion and the invariance generators for any oscillator-mechanics problem were given in Sec. IV, and naturally these results hold for in-common mechanics problems as well. However, because of the peculiarly simple form of the matrix B [Eq. (43)], we may rewrite these quantities in terms of the eigenfunctions and eigenvalues of the $n/2$ dimensional matrix $\bar{H} \equiv H^R + iH^I$. Because $\bar{H}$ is Hermitian, it has $n/2$ real eigenvalues ω_α and eigenvectors ψ_α satisfying

$$\bar{H}\psi_\alpha = \omega_\alpha\psi_\alpha, \quad \psi_\alpha^*\bar{H} = \omega_\alpha\psi_\alpha^*, \quad \psi_\alpha^* \cdot \psi_\beta = \delta_{\alpha\beta} \quad \left(\alpha, \beta = 1, \cdots, \frac{n}{2}\right).$$

It is readily verified that the eigenvectors $\mathbf{e}_\alpha$ of B, which satisfy $B\mathbf{e}_\alpha = i\omega_\alpha\mathbf{e}_\alpha$, can be written in block form as

$$\sqrt{2}\,\mathbf{e}_\alpha = \begin{bmatrix} \psi_\alpha \\ -i\psi_\alpha \end{bmatrix}, \quad \sqrt{2}\,e_{-\alpha} = \begin{bmatrix} \psi_\alpha^* \\ i\psi_\alpha^* \end{bmatrix} \quad i = 1, \cdots, \frac{n}{2}.$$

Let us introduce the $(n/2)$-dimensional complex vector $\mathbf{z}$ and the complex operator $\bar{\nabla}$ by

$$z^k \equiv \frac{1}{\sqrt{2}}(x^k + ix^{n/2+k}),$$

$$\bar{\nabla}^k \equiv \frac{\partial}{\partial x^k} - i\frac{\partial}{\partial x^{n/2+k}} = \sqrt{2}\frac{\partial}{\partial z^k}.$$

Using this notation, for example, the $n/2$ constants of the motion which can be written as quadratic forms are

$$(\mathbf{x} \cdot \mathbf{e}_\alpha)(\mathbf{e}_\alpha^* \cdot \mathbf{x}) \equiv (\mathbf{z}^* \cdot \psi_\alpha)(\psi_\alpha^* \cdot \mathbf{z}), \quad \left(\alpha = 1, \cdots, \frac{n}{2}\right), \quad (45)$$

while a complete commuting set of n (complex) invariance generators is

$$\mathbf{g}_\alpha \cdot \nabla \equiv (\mathbf{z}^* \cdot \psi_\alpha)(\psi_\alpha^* \cdot \bar{\nabla}^*) \quad \text{and} \quad (\mathbf{g}_\alpha \cdot \nabla)^*, \quad \left(\alpha = 1, \cdots, \frac{n}{2}\right). \quad (46)$$

In this notation, the expression for $dF(x)/dt = \mathbf{b} \cdot \nabla F = \{F(x), H(x)\}$ becomes

$$\frac{dF(x)}{dt} = -i\sum_{\alpha=1}^{n/2}\omega_\alpha[(\bar{\nabla}F \cdot \psi_\alpha)(\psi_\alpha^* \cdot \mathbf{z}) - (\mathbf{z}^* \cdot \psi_\alpha)(\psi_\alpha^* \cdot \bar{\nabla}^* F)] \quad (47)$$

[using the expression of Eq. (20) for $\mathbf{b} \cdot \nabla$, and Eq. (46)]. By utilizing the expansion for $\bar{H}$ in terms of its eigenvectors

$$\bar{H}_{jk} = \sum_{\alpha=1}^{n/2}\omega_\alpha\psi_\alpha^j\psi_\alpha^{*k}, \quad (i, k = 1, \cdots, n/2),$$

Eq. (47) may also be written as

$$i\frac{dF(x)}{dt} = \sum_{j,k=1}^{n/2}\left[\frac{\partial F}{\partial z^j}\bar{H}_{jk}z^k - z^{j*}\bar{H}_{jk}\frac{\partial F}{\partial z^{k*}}\right]. \quad (48)$$

Because in-common mechanics problems belong to both Hamiltonian mechanics and oscillator mechanics, we can construct two different invariance generators from a single constant of the motion. For example, the generators constructed from the conserved quantities in Eq. (45), according to the oscillator rule $\mathbf{g} \cdot \nabla = \nabla\Phi \cdot \nabla$, are $(\mathbf{g}_\alpha \cdot \nabla)^* + (\mathbf{g}_\alpha \cdot \nabla)$; the generators constructed from the same functions according to the Hamiltonian rule

$$\mathbf{g} \cdot \nabla = J^{ij}(\partial\Phi/\partial x^j)\partial/\partial x^i$$

are

$$i[(\mathbf{g}_\alpha \cdot \nabla)^* - (\mathbf{g}_\alpha \cdot \nabla)].$$

In order to see how the algebra of in-common mechanics is related to the algebra of matrix mechanics, we consider Eq. (48) with $F(x) = z^j$:

$$i\frac{dz^j}{dt} = \sum_{k=1}^{n/2}\bar{H}_{jk}z^k, \quad \left(j = 1, \cdots, \frac{n}{2}\right). \quad (49)$$

Thus Hamilton's equation of motion in classical mechanics now resembles Schrödinger's equation in $(n/2)$-dimensional matrix mechanics, where $\bar{H}$ takes the place of the $n/2 \times n/2$ Hamiltonian matrix divided

by $\hbar$ and the z^k take the place of the $n/2$ components of the state vector ψ.

Next let $F(x)$ be chosen as a quadratic form $\frac{1}{2}x^iF_{ij}x^j$ ($F_{ij} = F_{ji}$, $\partial F^{ij}/\partial x^k = 0$), and suppose that $[F, J] = 0$ (F is the matrix whose elements are F_{ij}). Following an argument identical to the one we used to determine the structure of the matrix H [Eqs. (42) to (43)], we find that we can write F in terms of $(n/2)$-dimensional matrices F^R and F^I:

$$F = \begin{bmatrix} F^R & -F^I \\ F^I & F^R \end{bmatrix}, \quad \tilde{F}^R = F^R, \quad \tilde{F}^I = -F^I. \tag{50}$$

We may define a Hermitian matrix $\bar{F} \equiv F^R + iF^I$, and we may write the quadratic form as

$$\sum_{j,k=1}^{n/2} z^{j*}\bar{F}_{jk}z^k.$$

When this special function of x is inserted into Eq. (48), we obtain

$$i\frac{d}{dt}\sum_{j,k=1}^{n/2} z^{j*}\bar{F}_{jk}z^k = \sum_{j,k=1}^{n/2} z^{j*}[\bar{F}, \bar{H}]_{jk}z^k, \tag{51}$$

where $[\bar{F}, \bar{H}]_{jk}$ are the elements of the commutator of the matrices $\bar{F}$ and $\bar{H}$. Equation (51) is to be compared with the expression in matrix mechanics for the time derivative of a matrix element.

The fact that a classical Hamiltonian system of oscillators exists whose algebra is similar to that of matrix mechanics, was first noted by Strocchi.[2] Matrix mechanics not only differs from in-common mechanics in the interpretation of such quantities as the variables z^k, it also differs in that the quantities and manipulations necessary to matrix mechanics are more restricted than those possible with in common mechanics. For example, only linear sums of $n/2$ of the constants of the motion [those in Eq. (45)] can be written in the form of the quadratic expression that that appears in Eq. (51). One simply does not need to work with the other $n/2 - 1$ constants of the motion which are not quadratic forms when doing matrix mechanics.

As a last demonstration of the similarity of the algebras of in-common mechanics and matrix mechanics, we shall consider a "transformation theory" of in common mechanics. That is, we shall look for all transformations of variables from x to $\bar{x}$, such that when the equations of motion of any in-common problem are expressed in the new variables $\bar{x}$, the problem still remains in in-common mechanics. It will be shown that these transformations are unitary transformations, when they are regarded as transforming the complex variables z into $\bar{z}$.

If a problem is to remain in Hamiltonian mechanics after a transformation of variables, then it is well known that $\bar{x}^i = X^i(x)$ must be a canonical transformation. The condition that a transformation be canonical is that the new variables $\bar{x}$ obey the Poisson-bracket relations

$$J^{rs} \equiv \frac{\partial \bar{x}^r}{\partial x^i} J^{ij} \frac{\partial \bar{x}^s}{\partial x^j} \equiv \{\bar{x}^r, \bar{x}^s\}. \tag{52}$$

If a problem is to remain in oscillator mechanics after a transformation of variables, the transformation $\bar{x}^i = X^i(x)$ must take the equations of motion $dx^i/dt = x^jB_{ji}$ into equations of the form $d\bar{x}^i/dt = \bar{x}^j\bar{B}_{ji}$ (where B, $\bar{B}$ are real, nonsingular, antisymmetric, constant $n \times n$ matrices). As is shown in Appendix B, an arbitrary orthogonal transformation

$$\bar{x}^i = U_{ij}x^j \tag{53}$$

($U\tilde{U} = 1$, $U^*_{ij} = U_{ij}$, $\partial U^{ij}/\partial x^k = 0$) can be considered as the most general transformation having this desired property. Upon inserting Eq. (53) into Eq. (52), we obtain the relation $J = UJ\tilde{U}$. Therefore U is restricted by the following two conditions:

$$[U, J] = 0, \quad \tilde{U}U = 1. \tag{54}$$

The vanishing of the commutator implies, as we have seen, that U can be written in block form as

$$U = \begin{bmatrix} U^R & -U^I \\ U^I & U^R \end{bmatrix}, \tag{55}$$

where U^R, U^I are arbitrary real $(n/2)$-dimensional matrices. The condition that U is an orthogonal matrix requires that

$$U^RU^R + U^IU^I = 1, \quad U^IU^R = U^RU^I, \quad U^RU^I = U^IU^R. \tag{56}$$

The transformation in Eq. (53), satisfying Eqs. (55) and (56), is canonical in both Hamiltonian mechanics and oscillator mechanics (here a "canonical transformation" for any mechanics is taken to mean a transformation which takes any problem of the mechanics into another problem of the mechanics). To see that the transformation is canonical for in-common mechanics also, it is convenient to introduce the $n/2 \times n/2$ complex matrix $\bar{U} \equiv U^R + iU^I$. Using Eq. (56), we find that

$$\bar{U}^+\bar{U} = (\tilde{U}^R - i\tilde{U}^I)(U^R + iU^I) = 1,$$

so that $\bar{U}$ is a unitary matrix. Writing $\bar{z}^j \equiv \bar{x}^j + i\bar{x}^{n/2+j}$ ($j = 1, \cdots, n/2$), Eq. (53) becomes

$$\bar{z}^j = \sum_{j=1}^{n/2} \bar{U}_{jk}z^k. \tag{57}$$

As is well known, the equations of motion (49) do not change their form under the unitary transformation (57), so Eq. (53) is also a canonical transformation for in-common mechanics.

APPENDIX A

In Sec. IV, it was shown that for the oscillator mechanics, three functions of any invariance generator vector $\mathbf{g}$ must be either constants or constants of the motion. In the notation introduced there, these three functions are

$$\mathbf{g}\cdot\mathbf{x} = \sum_{\alpha=0}^{n/2}(h^{\alpha^*} + h^{\alpha})(\mathbf{g}_\alpha\cdot\mathbf{x}), \tag{21a}$$

$$\mathbf{g}\cdot\mathbf{b} = \sum_{\alpha=0}^{n/2} i\omega_\alpha(h^{\alpha^*} - h^{\alpha})(\mathbf{g}_\alpha\cdot\mathbf{x}), \tag{21b}$$

$$\mathbf{g}\cdot\mathbf{g} = 2\sum_{\alpha=0}^{n/2}|h^{\alpha}|^2(\mathbf{g}_\alpha\cdot\mathbf{x}). \tag{21c}$$

In this appendix, we will prove that at least one of expressions (21a–c) must be a constant of the motion.

In order to show that h^α cannot be chosen to make Eqs. (21a–c) all constants, we will suppose that $\mathbf{g}\cdot\mathbf{x} = c_1$, $\mathbf{g}\cdot\mathbf{b} = c_2$, $\mathbf{g}\cdot\mathbf{g} = c_3$ (c_1, c_2, c_3 are finite constants) and obtain a contradiction. First we observe that $\mathbf{g}\cdot\mathbf{g} > 0$ because it is the sum of positive terms (remember that $\mathbf{g}_\alpha\cdot\mathbf{x} = |\mathbf{e}_\alpha\cdot\mathbf{x}|^2$), which can only vanish identically if all $h^\alpha = 0$ (i.e., if $\mathbf{g}\equiv 0$, which is unadmissable). We will see that $c_3 = 0$ is nevertheless a necessary consequence of assuming that Eqs. (21) are all constants, and thus we will obtain the contradiction.

Let us choose $\mathbf{x}$ so that, for one particular β, $\mathbf{g}_\beta\cdot\mathbf{x} = \epsilon$ (ϵ is any small positive number), while for $\alpha\neq\beta$, the $\mathbf{g}_\alpha\cdot\mathbf{x}$ are arbitrary. We can see that as $\epsilon\to 0$, $|h^\beta|^2$ cannot become infinite more rapidly than $1/\epsilon$, or else c_3 would be infinite. This means that both $(h^{\beta^*} + h^\beta)(\mathbf{g}_\beta\cdot\mathbf{x})$ and $(h^{\beta^*} - h^\beta)(\mathbf{g}_\beta\cdot\mathbf{x})$ must vanish as $\epsilon\to 0$, at least as fast as $(\epsilon)^{\frac{1}{2}}$. Furthermore, unless $|h^\beta|^2$ goes to infinity exactly as $1/\epsilon$, $|h^\beta|^2(\mathbf{g}_\beta\cdot\mathbf{x})$ will also vanish as $\epsilon\to 0$. Let us now let $\epsilon = 0$, which removes a single term from each of the sums (21a), (21b), while the term $|h^\beta|^2\mathbf{g}_\beta\cdot\mathbf{x}$ will only not vanish from the sum (21c) if $|h^\beta|^2\propto 1/\epsilon$ as $\epsilon\to 0$.

Continuing in this fashion, we may further restrict $\mathbf{x}$ so that for some particular γ, $\mathbf{g}_\gamma\cdot\mathbf{x} = \epsilon'$ (ϵ' is any small positive number) and $\mathbf{g}_\beta\cdot\mathbf{x} = 0$, while for $\alpha\neq\beta,\gamma$, the $\mathbf{g}_\alpha\cdot\mathbf{x}$ are still arbitrary. Once more we may conclude that when $\epsilon' = 0$, $(h^{\gamma^*} + h^\gamma)(\mathbf{g}_\gamma\cdot\mathbf{x})$ and $i\omega_\gamma(h^{\gamma^*} - h^\gamma)(\mathbf{g}_\gamma\cdot\mathbf{x})$ must vanish from the sums (21a), (21b), while $|h^\gamma|^2(\mathbf{g}_\gamma\cdot\mathbf{x})$ will only not vanish if $|h^\gamma|^2\propto 1/\epsilon'$ as $\epsilon'\to 0$.

Repeating this argument $n/2 - 1$ times, we eventually end up with only one term in each of the sums (21a), (21b):

$$c_1 = (\bar{h}^{\alpha^*} + \bar{h}^{\alpha})z, \tag{A1a}$$

$$c_2 = i\omega_\alpha(\bar{h}^{\alpha^*} - \bar{h}^{\alpha})z, \tag{A1b}$$

$$c_3 = 2|\bar{h}^{\alpha}|^2 z + (\text{terms which are positive or zero}). \tag{A1c}$$

Here we have set $\mathbf{g}_\alpha\cdot\mathbf{x} = z$ (where α is the remaining $n/2$th index) and $\bar{h}^\alpha(z)$ is the value of h^α when all of its arguments except z vanish or become infinite. [The arguments of h^α are the $\Phi^{\alpha\beta}$'s, all of which (except $\Phi^{\alpha,-\alpha} = \mathbf{g}_\alpha\cdot\mathbf{x}\equiv z$) vanish or become infinite as the various ϵ's are consecutively set equal to zero.]

Finally, repeating the argument for the index α, we see that when $z = 0$ the last terms in the sums (A1a), (A1b) vanish, so we conclude that $c_1 = c_2 = 0$. But this means that the right-hand sides of equations (A1a), (A1b) vanish, even when $z\neq 0$, and this in turn implies that $\bar{h}^\alpha(z)$ must vanish for all values of z. Therefore $\bar{h}^\alpha(z) = 0$ for all z, including $z = 0$ [since $\bar{h}^\alpha(z)$ is a continuous function of z].

This last result is what we have striven to obtain. Since each h^α is a finite-valued continuous function of its arguments, no matter how $\mathbf{x}$ approaches zero, the limit of h^α when $\mathbf{x}\to 0$ is the same as the limit of $\bar{h}^\alpha(z)$ as $z\to 0$, i.e., $h^\alpha\to 0$ for all α. But, according to Eq. (21c), this means that $c_3 = 0$, and we have our contradiction.

APPENDIX B

We wish to prove that, as stated in Sec. VIII, the most general transformation $\bar{x}^i = X^i(x)$ which takes any problem of oscillator mechanics into another problem in oscillator mechanics (i.e., a "canonical" transformation for oscillator mechanics) can be considered to be an orthogonal transformation $\bar{x}^i = U_{ij}x^j$ ($\tilde{U}U = 1$, $U^* = U$, $\partial U/\partial x^i = 0$).

Before the coordinate transformation is performed, the equations of motion are $dx^i/dt = x^r B^{(\lambda)}_{ri}$ (the superscript λ is present to label the different antisymmetric nonsingular matrices B). After the canonical coordinate transformation, the equations of motion must have the form $d\bar{x}^i/dt = \bar{x}^r\bar{B}^{(\lambda)}_{ri}$, where $\bar{B}^{(\lambda)}$ must also be antisymmetric and nonsingular. It is easy to see that x^rB_{ri} transforms like a contravariant vector under any coordinate transformation, so that $B^{(\lambda)}$ and $\bar{B}^{(\lambda)}$ are related in this way:

$$\bar{x}^s\bar{B}^{(\lambda)}_{sj} = x^rB^{(\lambda)}_{ri}\frac{\partial\bar{x}^j}{\partial x^i}. \tag{B1}$$

Now we observe that because Eq. (B1) holds for all *nonsingular* antisymmetric matrices $B^{(\lambda)}$, it must hold

for *singular* antisymmetric matrices as well. This is because any singular antisymmetric matrix $B^{(\lambda')}$ can be written as linear sum of selected nonsingular antisymmetric matrices: $B^{(\lambda')} = \sum_\lambda B^{(\lambda)}$. By summing Eq. (B1) over this selected set of matrices, we see that Eq. (B1) must hold for any antisymmetric matrix ($\bar{B}^{(\lambda')}$ may also turn out to be a singular matrix).

A special set of singular matrices $B^{(\mu\nu)}$ may be chosen, whose matrix elements are

$$B_{ij}^{(\mu\nu)} = \delta_{\mu i}\delta_{\nu j} - \delta_{\nu i}\delta_{\mu j} \quad \mu, \nu = 1, \cdots, n, \quad \mu \neq \nu. \tag{B2}$$

Upon inserting Eq. (B2) into Eq. (B1), we obtain

$$\left(x^\mu \frac{\partial}{\partial x^\nu} - x^\nu \frac{\partial}{\partial x^\mu}\right)\bar{x}^j = \bar{x}^s \bar{B}_{sj}^{(\mu\nu)}, \tag{B3}$$

where $\bar{B}^{(\mu\nu)}$ is some antisymmetric matrix corresponding to $B^{(\mu\nu)}$.

Equation (B3) states that the n functions $\bar{x}^j(x)$ are basis functions for the n-dimensional real representation of the n-dimensional rotation group,[8] and the matrices $\bar{B}^{\mu\nu}$ are matrix generators for this representation. The matrices given in Eq. (B2) are also matrix generators for this same representation, and both $B^{(\mu\nu)}$ and $\bar{B}^{(\mu\nu)}$ generate representations whose matrices are real orthogonal. It follows that these representations are related by a real orthogonal transformation:

$$B^{(\mu\nu)} = \tilde{U}\bar{B}^{(\mu\nu)}U, \quad (\tilde{U}U = 1). \tag{B4}$$

Now that the matrix U has been introduced, we may use it to define a new set of variables x' which will temporarily prove useful in helping us solve Eqs. (B3) for $\bar{x}$:

$$\bar{x}^s \equiv U_{sr}x'^r. \tag{B5}$$

Upon inserting Eq. (B5) into Eq. (B3) [and using Eqs. (B4), (B2)], we arrive at

$$\left(x^\mu \frac{\partial}{\partial x^\nu} - x^\nu \frac{\partial}{\partial x^\mu}\right)x'^j = x'^s(\tilde{U}\bar{B}^{(\mu\nu)}U)_{sj} = x'^\mu\delta_{\nu j} - x'^\nu\delta_{\mu j}. \tag{B6}$$

Equations (B6) constitute a set of partial differential equations for the functions x'^j, which we will now proceed to solve. If neither μ nor ν equals j, then the right-hand side of Eq. (B6) will vanish. Of the equations

$$\left(x^\mu \frac{\partial}{\partial x^\nu} - x^\nu \frac{\partial}{\partial x^\mu}\right)x'^j = 0 \quad (\mu, \nu \neq j), \tag{B7}$$

we may select $n - 2$ equations for which μ is fixed at some value, and ν takes on the values from 1 to n, excluding j and μ. Any other equation of the set (B7) can be written as a linear sum (with coefficients which are functions of x) of these $n - 2$ equations. These $n - 2$ equations form a complete set[9] and we therefore conclude from the theory of sets of first-order partial differential equations that their most general solution is an arbitrary function of any two independent solutions. Two such solutions are $x'^j = x^j$ and $x'^j = \sum_{i=1}^n x^{i2} \equiv x^2$, so the most general solution of Eqs. (B7) is

$$x'^j = f^j(x^j, x^2), \tag{B8}$$

where the f^j are any arbitrary functions.

Returning to the set of Eqs. (B6), we now set $\mu = j$:

$$\left(x^j \frac{\partial}{\partial x^\nu} - x^\nu \frac{\partial}{\partial x^j}\right)f^j(x^j, x^2) = -f^\nu(x^\nu, x^2), \quad \nu \neq j. \tag{B9}$$

When we perform the differentiation, Eq. (B9) becomes

$$x^\nu f_1^j(x^j, x^2) = f^\nu(x^\nu, x^2), \tag{B10}$$

where f_1^j is the derivative of f^j with respect to its first argument. Equation (B10) can only be satisfied if f_1^j is independent of x^j since f on the right-hand side cannot be written as a function of x^j. Upon writing $f_1^j \equiv f(x^2)$, we see that the solution of Eq. (B10) [and therefore Eqs. (B6)] is

$$x'^j = f^j = x^j f(x^2). \tag{B11}$$

Therefore the most general canonical transformation for oscillator mechanics is [from Eqs. (B5), (B11)]

$$\bar{x}^s = f(x^2)U_{sr}x^r. \tag{B12}$$

Under the transformation (B12), the equations of motion $dx^i/dt = x^r B_{ri}$ become $d\bar{x}^i/dt = \bar{x}^r(UB\tilde{U})_{ri}$ [using $df(x^2)/dt = 2(df/dx^2)(dx^i/dt)x^i = 2(df/dx^2) \times x^r B_{ri}x^i = 0$], a result which is independent of which $f(x^2)$ function is chosen. If we are only interested in transformations which produce different equations of motion, we may, without loss of generality in this respect, set $f(x^2) = 1$. Thus we have shown that the most general canonical transformation in oscillator mechanics can be considered to be an orthogonal transformation $\bar{x}^s = U_{sr}x^r$.

[8] F. Murnaghan, *The Unitary and Rotation Groups* (Spartan Books, Washington D.C., 1962); M. Hammermesh, *Group Theory* (Addison-Wesley Publishing Company, Inc., London, 1961).

[9] G. Duff, *Partial Differential Equation* (The University of Toronto Press, Toronto, 1956).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 7 JULY 1968

Symmetric Representation for Three-Body Problems. II. Motion in Space*

R. C. WHITTEN† AND F. T. SMITH
Stanford Research Institute, Menlo Park, California

(Received 31 July 1967)

The symmetric representation of three bodies is extended to the case of motion in space. It is found that the generalized angular-momentum eigenfunctions are linear combinations of products of elements of the "kinematic-rotation" group with elements of the spatial-rotation group. Recurrence relations among the expansion coefficients are derived; values are tabulated for $\lambda \leq 4$. Finally, coupling coefficients for the eigenfunctions are derived.

I. INTRODUCTION

The problem of the symmetric or "democratic" representation of the quantum mechanics of three bodies has been the subject of rather intensive investigation for the past several years.[1–5] The various approaches have used group-theoretical methods,[2,3] as well as more direct ones, wherein the eigenvalue differential equation for the generalized angular momentum (GAM) is solved.[4,5] In each case, the solution is expressed as a linear combination of products of elements of the rotation group (expressed as functions of a set of Euler angles that specify the orientation of the plane of the particles) with functions of a pair of "kinematic" angles Θ and Φ. In the papers published thus far these functions appear to be complicated quantities, with no obvious properties that may be invoked in order to simplify the mathematical manipulations. Actually, they were shown in an earlier paper[1] to transform like elements of the rotation group $SU(2)$ under so-called "kinematic rotations," and this important property is exploited in the present one. Hence one would expect the GAM wavefunction to be given by linear combinations of products of the group elements corresponding to spatial rotations and to kinematic rotations; we show that this is indeed the case.

Before doing so, however, it is very worthwhile to review briefly the kinematic coordinates presented in Paper I. As in most treatments of the mechanics of three bodies, their relative positions are described by two vectors, one $\boldsymbol{\xi}^1$ (suitably normalized), which joins particles 1 and 2, the other $\boldsymbol{\xi}^2$ (also suitably normalized) which extends from the center of mass of the (1, 2) pair to particle 3. The reader is referred to Paper I for specific formulas in terms of masses and position vectors in an arbitrary coordinate system. In terms of the kinematic angles, we can write (in the principal axis system)

$$\begin{aligned} \xi_1^1 &= \rho \cos\Theta \cos\Phi, \\ \xi_2^1 &= -\rho \sin\Theta \sin\Phi, \\ \xi_1^2 &= \rho \cos\Theta \sin\Phi, \\ \xi_2^2 &= \rho \sin\Theta \cos\Phi, \end{aligned} \qquad \rho = |\boldsymbol{\xi}|, \tag{1}$$

which are, in fact, definitions of ρ, Θ, and Φ. These definitions were not presented explicitly in the earlier paper, but they are easily derived from the expressions for the area of the figures enclosed by the three bodies, the moments and products of inertia of the figure, etc. For example, the area is

$$A = \tfrac{1}{2}(\xi_1^1\xi_2^2 - \xi_2^1\xi_1^2) = \frac{\rho^2}{4}\sin 2\Theta, \tag{2}$$

and the product of inertia is

$$E = \tfrac{1}{2}(\xi_1^1\xi_2^1 + \xi_1^2\xi_2^2) = 0. \tag{3}$$

Equation (3) proves that the coordinate axes specified are indeed the principal axes of the figure. We shall call this the *body-fixed system*. In Paper I the system was generalized by permitting rotations about an axis perpendicular to the plane of the particles. The GAM eigenfunctions of the system were found to be elements of the rotation group

$$\psi = D^{\lambda/2}_{m/2,-\sigma/2}\left(2\gamma, \frac{\pi}{2} - 2\Theta, -2\Phi\right), \tag{4}$$

where γ specifies the spatial rotation, λ is the GAM quantum number, m is the angular-momentum quantum number, and σ is another quantum number which specifies the configuration of the particles. In the present paper, we further generalize this formula to include rotations in space.

* This work was supported by Stanford Research Institute as part of its program of internal research and development.

† Present address: Ames Research Center (National Aeronautics and Space Administration), Moffett Field, California.

[1] F. T. Smith, J. Math. Phys. **3**, 735 (1962), referred to as Paper I.

[2] A. J. Dragt, J. Math. Phys. **6**, 533 (1965).

[3] E. Chacón and M. Moshinsky, Rev. Mex. Fis. **14**, 119 (1965).

[4] W. Zickendraht, Ann. Phys. (N.Y.) **35**, 18 (1965).

[5] V. V. Pustovalov and Yu. A. Simonov, Zh. Eksp. Teor. Fiz. **51**, 345 (1966) [Sov. Phys.—JETP **24**, 230 (1967)].

II. MODIFIED COORDINATE SYSTEM

To permit the use of the algebra of irreducible tensors, it is necessary to transform from the Cartesian system of Eq. (1) to the more convenient spherical representation on the unit sphere ($\rho = 1$). We will also perform a "kinematic rotation" by replacing Θ by $\Theta - (\pi/4)$:

$$\begin{aligned} \bar{\xi}^1_{+1} &= -(1/\sqrt{2})(\xi^1_1 + i\xi^1_2), \quad \xi^1_{-1} = (1/\sqrt{2})(\xi^1_1 - i\xi^1_2), \\ \bar{\xi}^2_{+1} &= -(1/\sqrt{2})(\xi^2_1 + i\xi^2_2), \quad \xi^2_{-1} = (1/\sqrt{2})(\xi^2_1 - i\xi^2_2); \end{aligned} \tag{5}$$

and

$$\begin{aligned} \bar{\xi}^{+1}_{+1} &= -(\xi^1_{+1} + i\xi^2_{+1}) = e^{i\Phi}\cos\Theta, \\ \bar{\xi}^{+1}_{-1} &= -(\xi^1_{-1} + i\xi^2_{-1}) = -e^{i\Phi}\sin\Theta, \\ \bar{\xi}^{-1}_{+1} &= (\xi^1_{+1} - i\xi^2_{+1}) = -e^{-i\Phi}\sin\Theta, \\ \bar{\xi}^{-1}_{-1} &= (\xi^1_{-1} - i\xi^2_{-1}) = e^{-i\Phi}\cos\Theta. \end{aligned} \tag{6}$$

This can be written in the general form

$$\bar{\xi}^\sigma_K = (-1)^{(\frac{1}{2}+\sigma/2)(\frac{1}{2}-K/2)}\,\mathfrak{D}^{\frac{1}{2}}_{\sigma/2,K/2}(2\Phi, 2\Theta, 0), \tag{7}$$

where $\mathfrak{D}$ is used hereafter for elements of the kinematic rotation group and D represents spatial rotations. Since the $\boldsymbol{\xi}^\sigma$ are spatial (but not kinematic!) vectors, it is a simple matter to rotate them to any orientation specified by the Euler angles φ, θ, and γ:

$$\begin{aligned} \xi^\sigma_m &= \sum_{K=\pm 1} D^1_{mK}(\varphi, \theta, \gamma)\,\bar{\xi}^\sigma_K \\ &= \sum_{K=\pm 1} (-1)^{(\frac{1}{2}+\sigma/2)(\frac{1}{2}-K/2)}\, D^1_{mK}(\varphi, \theta, \gamma) \\ &\quad \times \mathfrak{D}^{\frac{1}{2}}_{\sigma/2,K/2}(2\Phi, 2\Theta, 0). \end{aligned} \tag{8a}$$

Note that the angle of rotation γ in the plane of the particles is contained in D^1 rather than in $\mathfrak{D}^{\frac{1}{2}}$. There is nothing unique about this; we could just as well have written

$$\begin{aligned} \xi^\sigma_m = \sum_K (-1)^{(\frac{1}{2}+\sigma/2)(\frac{1}{2}-K/2)}\, D^1_{mK}(\varphi, \theta, 0) \\ \times \mathfrak{D}^{\frac{1}{2}}_{\sigma/2,K/2}(2\Phi, 2\Theta, 2\gamma), \end{aligned} \tag{8b}$$

or we could have put part of γ in D^1 and part in $\mathfrak{D}^{\frac{1}{2}}$. This separation being immaterial, we shall hereafter omit the arguments of the spatial- and kinematic-rotation matrices to simplify the notation.

Previously we mentioned that the ξ^σ_m do not transform like vectors under kinematic rotations. This is obvious from Eq. (7), which suggests that they may be proportional to the components of a (single-valued) spinor; the single valuedness is a consequence of the appearance of 2Φ and 2Θ as the arguments of $\mathfrak{D}^{\frac{1}{2}}$. The spinor elements are ξ^{+1}_{+1}, ξ^{+1}_{-1}, $-\xi^{-1}_{+1}$, and ξ^{-1}_{-1}. It is evident that suitable bilinear combinations of the spinor elements will transform as vectors under kinematic rotations and as scalars under spatial rotations. Examples of these are

$$\mathfrak{D}^1_{00} = \cos 2\Theta = (\xi^{+1}_{+1}\xi^{-1}_{-1} - \xi^{+1}_{-1}\xi^{-1}_{+1}), \tag{9}$$

which is equal to four times the area of the triangle enclosed by the particle, and the spatial invariants

$$\begin{aligned} \boldsymbol{\xi}^{+1}\cdot\boldsymbol{\xi}^{+1} &= (1/\sqrt{2})\mathfrak{D}^1_{10}, \\ \boldsymbol{\xi}^{-1}\cdot\boldsymbol{\xi}^{-1} &= -(1/\sqrt{2})\mathfrak{D}^1_{-10}. \end{aligned} \tag{10}$$

The apparent simplicity of the foregoing analysis of the $\boldsymbol{\xi}$ in terms of elements of the kinematic-rotation group is somewhat deceptive. The interval of Θ on which the $\psi^{\lambda\sigma j}$ are orthogonal differs from that on which the $\mathfrak{D}^{\lambda/2}_{\sigma/2,K/2}(2\Phi, 2\Theta, 0)$ are orthogonal. As we shall see later, this introduces some complications because the usual orthogonality relations for elements of $SU(2)$ do not apply. The density function for invariant integration is

$$\sin 4\Theta \sin\theta;$$

the ranges of integration are

$$\begin{aligned} 0 &\le \varphi \le 2\pi, \\ 0 &\le \theta \le 2\pi, \\ 0 &\le \gamma < 2\pi, \\ 0 &\le \Theta \le \pi/4, \\ 0 &\le \Phi \le 2\pi. \end{aligned}$$

We are now almost to the point where the wavefunctions of the system can be derived. In fact, the ξ^σ_m are the wavefunctions corresponding to $\lambda = 1$ and j (orbital angular momentum—OAM) $= 1$ (see Table II of Ref. 2). Before proceeding in this direction, however, we digress slightly by reviewing the properties of the various irreducible tensors that can be constructed from the GAM operator.

III. TENSOR OPERATORS

In paper I the GAM tensor $\boldsymbol{\Lambda}$ was defined (in atomic units) as

$$\Lambda^{ij}_{kl} = -i\left(\xi^i_k \frac{\partial}{\partial \xi^j_l} - \xi^j_l \frac{\partial}{\partial \xi^i_k}\right). \tag{11}$$

Like any tensor of rank two, symmetric and antisymmetric parts can be extracted from $\boldsymbol{\Lambda}$:

$$\begin{aligned} \Sigma_{ij} &= \Lambda^{12}_{ij} - \Lambda^{21}_{ij} \quad \text{(symmetric)}, \\ L_{ij} &= \Lambda^{11}_{ij} + \Lambda^{22}_{ij} \quad \text{(antisymmetric)}, \\ Y_{ij} &= \Lambda^{11}_{ij} - \Lambda^{22}_{ij} \quad \text{(symmetric)}, \\ A_{ij} &= \Lambda^{12}_{ij} + \Lambda^{21}_{ij} \quad \text{(antisymmetric)}, \end{aligned} \tag{12}$$

of which $L_{ij} \equiv L_k$ (ijk cyclic) is the kth component of the orbital angular momentum. A more useful representation for these tensors can be obtained by

transforming to the spherical (irreducible tensor[2]) representation:

$$\begin{aligned}
\Sigma^2_{\pm 2} &= \tfrac{1}{4}(\Sigma_{11} - \Sigma_{22} \pm 2i\Sigma_{12}),\\
\Sigma^2_{\pm 1} &= \mp(1/\sqrt{2})(\Sigma_{13} \pm i\Sigma_{23}),\\
\Sigma^2_0 &= (1/\sqrt{6})(3\Sigma_{33} - 2\Sigma^0),\\
\Sigma^0 &= \tfrac{1}{2}(\Sigma_{11} + \Sigma_{22} + \Sigma_{33}),\\
L_{\pm 1} &= \pm(1/\sqrt{2})(L_1 \pm iL_2),\\
L_0 &= L_3,\\
\Xi^{\pm 1}_m &= \mp(1/\sqrt{2})(Y_m \pm iA_m),\\
Y_{\pm 1} &= \mp(1/\sqrt{2})(Y_1 \pm iY_2),\\
Y_0 &= Y_3,\\
A_{\pm 1} &= \mp(1/\sqrt{2})(A_1 \pm iA_1),\\
A_0 &= A_0.
\end{aligned} \tag{13}$$

From these irreducible operators it is possible to construct the Lie algebras which underlie the groups $SU(3)$ and $SU(4)$. The study of $SU(3)$ has been particularly popular in recent years because of its applications in elementary-particle theory. Biedenharn and Baird[6] have published quite exhaustive investigations of the properties of the general unitary unimodular group $SU(n)$ in which they showed how to construct all of the invariants. According to their results, the three-body problem can, in principle, be treated entirely within $SU(3)$, although the invariants are not particularly useful. The invariant Casimir operators are the GAM, $\mathbf{\Lambda}^2$, and the operator

$$\Sigma_d = \mathbf{L} \cdot \mathbf{\Sigma} \cdot \mathbf{L}, \tag{14}$$

which can be shown to be the determinant of the $\mathbf{\Sigma}$ tensor; the remainder of the mutually commuting set consists of the zeroth component of OAM, L_0, and the zeroth component of $\mathbf{\Sigma}$, Σ^2_0, and another bilinear operator constructed from $\Sigma^2_{\pm 2}$ and L_0,

$$M^2 = L_0^2 + \tfrac{1}{2}(\Sigma^2_2\Sigma^2_{-2} + \Sigma^2_{-2}\Sigma^2_2). \tag{15}$$

Unfortunately, M^2 is not a very useful constant of the motion in the three-body problem. Rather, it is to be preferred that the OAM operator L^2 be used because it has real physical significance in the problem. However, if M^2 is replaced by L^2, the zeroth component of $\mathbf{\Sigma}$ is no longer a member of the mutually computing set and we immediately step outside the true representations of $SU(3)$ when we use it. In order completely to describe the 3-body motion it is necessary to find another mutually commuting dynamical variable. If we return to the set of operators [which, by the way, include linear combinations of all the generators of $SU(4)$], we immediately find that Σ^0 fills our requirement. It lies outside $SU(3)$, however, and so it is necessary to go into $SU(4)$ to obtain a satisfactory treatment. The importance of $SU(4)$ to the system which includes Σ^0 is evident from the identity

$$[\Sigma^0, \Xi^{\pm 1}_m] = \pm 2\Xi^{\pm 1}_m, \tag{16}$$

in which the Ξ^{μ}_m play the role of ladder operators which raise or lower the eigenvalue of Σ^0. The Ξ^{μ}_m are the additional generators required to extend the group from $SU(3)$ to $SU(4)$. Of course, it must be realized that, by demanding that L^2 be one of the commuting operators, the three-body eigenfunctions are *not* representations of $SU(4)$. They correspond to a projection of the representation into a subspace in which OAM is conserved. The general $SU(4)$ representation contains a mixture of the $SU(2)$ representations which correspond to kinematic and spatial rotations. Further investigations of the properties of true $SU(4)$ representations and their relations to the two $SU(2)$ groups mentioned above will be the topic of a future paper.

The eigenvalues of the operators Λ^2, Σ^0, L^2, and L_0 are well known[2] and are only summarized here:

$$\begin{aligned}
&\Lambda^2\psi = \lambda(\lambda + 4)\psi; \quad (\lambda \text{ integer}),\\
&\Sigma^0\psi = \sigma\psi; \quad \lambda \geq \sigma; \quad (\sigma = 0, \pm 2, \pm 4, \cdots)\\
&\qquad\qquad \text{or} \quad (\sigma = \pm 1, \pm 3, \pm 5, \cdots),\\
&L^2\psi = l(l + 1)\psi; \quad \lambda \geq l \quad (l \text{ integer}),\\
&L_0\psi = m\psi; \quad l \geq |m|.
\end{aligned} \tag{17}$$

The eigenvalues of Σ_d, called ω hereafter, are not so well known. In fact, the eigenfunctions obtained later will not be explicitly labeled according to ω, although we devise a procedure for computing it. The important point to bear in mind is that the eigenfunctions are sometimes degenerate with respect to the dynamical variables Λ^2, L^2, L_0, and Σ^0. Σ_d simply breaks this degeneracy; any method which affects the breaking of the degeneracy is satisfactory for our purposes. The approach developed in Sec. IV yields multiple solutions which correspond to *linear combinations* of the ω. Thus, the eigenfunctions of Λ^2

[6] L. C. Biedenharn, J. Math. Phys. **4**, 436 (1963); G. E. Baird and L. C. Biedenharn, J. Math. Phys. **4**, 1449 (1963); **5**, 1723, 1730 (1964).

obtained in this paper are not strictly eigenfunctions of Σ_d. They are only linear combinations of the latter, but this is immaterial to our development.

Before going on to construct the GAM eigenfunctions, it is instructive to look further at the algebraic form of some of the generators of $SU(4)$ in terms of generators of the two $SU(2)$ groups:

$$\Sigma_m^2 = \frac{\sqrt{2}}{\mathfrak{D}_{00}^1}\Big[-\mathfrak{D}_{00}^1\Big(D_{m2}^2\mathcal{K}_{+1} - D_{m-2}^2\mathcal{K}_{-1} + \frac{1}{\sqrt{3}}D_{m0}^2\mathcal{K}_0\Big) + \frac{1}{\sqrt{2}}(\mathfrak{D}_{01}^1 D_{m1}^2 L_{+1} - \mathfrak{D}_{0-1}^1 D_{m-1}^2 L_{-1}) + (D_{m-1}^2 L_{+1} + D_{m+1}^2 L_1)\Big], \quad (18)$$

$$\Xi_m^{+1} = \frac{1}{\mathfrak{D}_{00}^1}[-\sqrt{2}\,\mathfrak{D}_{00}^1 D_{m0}^1 \mathfrak{L}_{+1} + \mathfrak{D}_{10}^1(D_{m-1}^1 L_{+1} - D_{m1}^1 L_{-1}) - D_{1-1}^1 D_{m-1}^1 L_{-1} + D_{01}^1 D_{m1}^1 L_{+1}], \quad (19)$$

where

$$\mathcal{K}_{\pm 1} = \frac{2}{2\sqrt{2}}\Big[-\frac{\partial}{\partial\Theta} \pm i\csc 2\Theta\Big(\frac{\partial}{\partial\Theta} - \cos 2\Theta\frac{\partial}{\partial\gamma}\Big)\Big], \quad (20)$$

$$\mathcal{K}_0 = -\frac{i}{2}\frac{\partial}{\partial\Phi},$$

$$\mathfrak{L}_{\pm 1} = \frac{1}{2\sqrt{2}}\Big[-\frac{\partial}{\partial\Theta} \pm i\csc 2\Theta\Big(\frac{\partial}{\partial\gamma} - \cos 2\Theta\frac{\partial}{\partial\Phi}\Big)\Big] \quad (21)$$

are, respectively, the ladder operators with respect to k and σ for the $SU(2)$ rotation group elements $\mathfrak{D}_{\sigma/2,k/2}^{\lambda/2}$; $L_{\pm 1}$ are the usual generators of D_{mk}^j (ladder operators in the "body-fixed" coordinate system). The generators (23) and (24) could be employed to construct the eigenfunctions, although we shall employ a different route for this purpose.

We conclude this section by writing down an expression for the Casimir operator Λ^2 in terms of the angle variables. It is quite similar in form to that given by Zickendraht[4] with the difference that we exploit the appearance of the appropriate $SU(2)$ generators and group elements:

$$\Lambda^2 = -\tfrac{1}{2}(\mathcal{K}_{+1}\mathcal{K}_{-1} + \mathcal{K}_{-1}\mathcal{K}_{+1}) + \mathcal{K}_0^2 - \frac{1}{\sqrt{2}\,\mathfrak{D}_{00}^1}\Big[\mathfrak{D}_{0-1}^1\mathcal{K}_{+1} - \mathfrak{D}_{01}^1\mathcal{K}_{-1} + \frac{1}{\sqrt{2}}(L^2 - L_0^2)\Big] - \frac{1}{\sqrt{2}\,(\mathfrak{D}_{00}^1)^2}(\mathfrak{D}_{0-1}^1 L_{+1}^2 - \mathfrak{D}_{01}^1 L_{-1}^2), \quad (22)$$

where $L_0 \equiv -i(\partial/\partial\gamma)$. The forms of Eqs. (19), (20), and (22) immediately suggest the ansatz

$$\psi_\lambda = \sum_{\lambda'=\sigma}^{\lambda}\sum_k \alpha_k^{\lambda'\sigma j\omega}\mathfrak{D}_{\sigma/2,k/2}^{\lambda'} D_{mk}^j \quad (23)$$

as a trial solution. It is apparent that the parity of $j + \lambda$ can be either odd or even and the two never mix within the same eigenfunction. It should be noted that we do not demand a single representation for the kinematic-rotation group elements. Such a condition would be much too stringent: $\psi^{\lambda\sigma jn\omega}$ would, in general, be an eigenfunction of Λ^2 for motion in a plane, but not for motion in three dimensions.

This is an important and useful representation because it permits the use of the well-understood properties of the unitary unimodular group $SU(2)$. In practical calculations, the α's can be carried as algebraic quantities and evaluated only in the final result. The properties of this expansion are much more transparent than those obtained heretofore.[2-5]

IV. THE GAM EIGENFUNCTIONS

One can employ the forms of $\psi_{\frac{1}{2},\sigma,l,m} \equiv \xi_m^\sigma$ given in Eq. (8a) to build up the eigenfunctions for larger values of λ, σ, and j by writing irreducible tensor products. This would be quite a simple matter if the ξ transformed as spinors under kinematic rotations; unfortunately, they do not because of the appearance of the factor $(-1)^{(1+\sigma)(1-K)/4}$ in each term. It will be apparent in the following that it is this factor which leads to the summation over λ' in Eq. (23) as well as to certain restrictions on the OAM j. The simplest case to treat is the one where $\sigma = \lambda$. Then we start with

$$\xi_m^1 = \sum_{K=\pm 1}(-1)^{(1-K)/2}\mathfrak{D}_{\frac{1}{2},K/2}^{\frac{1}{2}} D_{mk}^1. \quad (8c)$$

It is very simple now to construct $\psi_{2,2,jm}$ by compounding two of the ξ_m^1:

$$\psi_{22jm} = \sum_{m'}(1, m'1, m - m' \mid 11jm)\xi_{m'}^1\xi_{m-m'}^1 = \sum_{KK'}(1, K - K', 1K' \mid 11jK) \times (\tfrac{1}{2}, (K - K')/2, \tfrac{1}{2}, K/2 \mid \tfrac{1}{2}, \tfrac{1}{2}, 1, K/2) \times (-1)^{(1-K)/2}\mathfrak{D}_{1K/2}^1 D_{mk}^j, \quad (24)$$

where we have used the Clebsch–Gordan series[7] to compound the rotational matrix elements. By direct

[7] A. R. Edmonds, *Angular Momentum in Quantum Mechanics* (Princeton University Press, Princeton, N.J., 1957); the phases of all Clebsch–Gordan (vector-coupling) coefficients are as given in this work.

comparison with Eq. (23), it is obvious that

$$\alpha_K^{2,2,j} = \sum_{K'} (1, K-K', 1, K' \mid 11jK) \times (\tfrac{1}{2}, (K-K')/2, \tfrac{1}{2}, K'/2 \mid \tfrac{1}{2}, \tfrac{1}{2}, 1K/2) \times (-1)^{1-K/2}. \quad (25)$$

Since the ξ can be coupled in only one way, Σ_d has only one eigenvalue. One can then go on to construct $\psi_{\lambda\lambda jm}$ for larger values of λ by use of the following recurrence relation, which is easily derived:

$$\alpha_K^{\lambda\lambda j\omega} = \sum_{K'} (-1)^{(1-K')/2} (j', K-K', 1, K' \mid j'1jK) \times (\lambda - \tfrac{1}{2}, (K-K')/2, \tfrac{1}{2}, K'/2 \mid \lambda - \tfrac{1}{2}, \tfrac{1}{2}, \lambda/2, K/2) \times \alpha_{K-K'}^{\lambda-1,\lambda-1,j'\omega}. \quad (26)$$

It can be shown that there is only one coupling scheme for $\psi_{\lambda\lambda jm\omega}$ so that Σ_d has a single eigenvalue for all cases where $\sigma = \lambda$. Thus we can omit ω and write $\psi_{\lambda\lambda jm}$.

The general case where $\sigma < \lambda$ is more difficult to construct. As a particularly simple case of $\sigma < \lambda$, let us construct ψ for $j = \lambda$, starting with $\lambda = 2$, $\sigma = 0$, and $m = 2$:

$$\begin{aligned}\psi_{2022} &= \sum_{m'} (1, 2-m', 1, m' \mid 1122) \xi_{m'}^1 \xi_{2-m}^{-1} = \xi_1^1 \xi_1^{-1} \\ &= \frac{1}{\sqrt{2}} \sum_{KK'} \sum_{\lambda=0,2} (\tfrac{1}{2}, (K-K')/2, \tfrac{1}{2}, K'/2 \mid \tfrac{1}{2}, \tfrac{1}{2}, \lambda/2, K/2) \times (1, K-K', 1, K' \mid 112K) \times (-1)^{(1-K')/2} \mathfrak{D}_{0K/2}^{\lambda/2} D_{2K}^2 \\ &= (1/\sqrt{2})[\mathfrak{D}_{01}^1 D_{22}^2 - (1/\sqrt{3}) D_{20}^2 - \mathfrak{D}_{0-1}^1 D_{2-2}^2]. \quad (27)\end{aligned}$$

Eigenfunctions[8] corresponding to $m = 0, -2$ are obtained by use of the appropriate ladder operators which leads merely to replacement of D_{2K}^2 by D_{mk}^2. One can now proceed to construct any $\psi_{\lambda\sigma jm}$ $(j = \lambda)$ by compounding the ξ in a similar manner:

$$\begin{aligned}\psi_{\lambda\sigma\lambda\lambda} &= (\xi_{+1}^{+1})^{(\lambda+\sigma)/2} (\xi_{-1}^{-1})^{(\lambda-\sigma)/2} \\ &= \sum_K \sum_{\lambda''=\sigma}^{\lambda} \alpha_K^{\lambda''\sigma\lambda} \mathfrak{D}_{\sigma/2,K/2}^{\lambda''/2} D_{\lambda K}^{\lambda}, \quad (28)\end{aligned}$$

where

$$\begin{aligned}\alpha_K^{\lambda,\tau,\lambda+1}(\lambda+1) = \sum_{\lambda''=\sigma-\sigma'}^{\lambda} \sum_{k'=\pm1} \alpha_{K-K'}^{\lambda'',\sigma-\sigma',\lambda}(\lambda) \alpha_{K'}^{1\sigma'1}(1) \\ \times (\lambda''/2, (\sigma-\sigma')/2, \tfrac{1}{2}, \sigma'/2 \mid \lambda''/2, \tfrac{1}{2}, \lambda'/2, \sigma/2) \\ \times (\lambda''/2, (K-K')/2, \tfrac{1}{2}, K'/2 \mid \lambda''/2, \tfrac{1}{2}, \lambda'/2, K/2) \\ \times (\lambda, K-K', 1, K \mid \lambda, 1, \lambda+1, K) \quad (29)\end{aligned}$$

[8] This is identical to Dragt's Eq. (6.79) if the explicit values of $\mathfrak{D}_{0,\pm1}^1$ are substituted into our Eq. (27).

is the recurrence relation for the α's and the quantum number ω has been omitted because $\psi_{\lambda\sigma\lambda\lambda}$ can be constructed in only one way. The index of α appearing in parenthesis is necessary to show the value of the GAM quantum number λ to which the α belongs; when $\sigma = \lambda$, there is no ambiguity and we omit the parenthetical index.

If we try to construct $\psi_{\lambda\lambda jj}$ for the case in which $j + K$ (or $j + \lambda$) is odd, we find that it vanishes. This can be shown by constructing ψ_{2211} by the method outlined above. The vanishing of the GAM eigenfunctions for larger values of $\sigma \pm \lambda$ with $\lambda + j$ odd can be proved by induction.

Eigenfunctions corresponding to $\lambda > \sigma$, $\lambda > j$ for $\lambda + j$ either even or odd can be constructed in a similar manner. However, they are, in general, nonorthogonal and the Schmidt orthogonalization procedure must be resorted to. The general procedure is to begin by building up a wavefunction which is orthogonal to those of lower quantum numbers insofar as the orbital angular momentum is concerned:

$$\varphi = \sum_m (j_1, m, j_2, -m \mid j_1, j_2, j, 0) \times \psi_{\lambda_1\sigma_1 j_1 m\omega_1} \psi_{\lambda_2,\sigma_2,j_2,-m,\omega_2}, \quad (30)$$

which is really a linear combination of the orthogonal functions

$$\psi_{\lambda_1+\lambda_2,\sigma_1+\sigma_2,j,0,\omega}\ \psi_{\lambda_1+\lambda_2-1,\sigma_1+\sigma_2,j,0,\omega'}, \cdots, \psi_{|\lambda_1-\lambda_2|,\sigma_1+\sigma_2,j,0,\omega^{(n)}},$$

where the ω's refer as usual to distinguishable solutions for given $(\lambda\sigma jm)$. Hence, we can write

$$\begin{aligned}\chi_{\lambda_1+\lambda_2,\sigma_1+\sigma_2,j,0,\omega} &= \sum_m (j_1, m, j_2, -m \mid j_1 j_2 j0)\, \psi_{\lambda_1\sigma_1 j_1 m\omega_1}\, \psi_{\lambda_2\sigma_2 j_2 - m\omega_2} \\ &\quad - \sum_{\omega'} \sum_{\lambda'=|\sigma_1+\sigma_2|}^{\lambda_1+\lambda_2-2} \beta^{\lambda',|\sigma_1+\sigma_2|,j\omega_1\omega_2\omega'}\, \psi_{\lambda',\sigma_1+\sigma_2,j,0,\omega'}, \quad (31)\end{aligned}$$

where the χ's are not yet orthogonal with respect to ω. However, they are expandable in the form given in Eq. (28) and the corresponding coefficients $\gamma_K^{\lambda'\sigma j\omega}(\lambda)$ are related to the α's of lower order by the recurrence relation

$$\begin{aligned}\gamma_K^{\lambda'\sigma j\omega}(\lambda) &= \sum_{\lambda_1'=|\sigma|} \sum_{\lambda_2'=|\sigma-\sigma_1|} \sum_{K'=\pm1} (j_1, K-K', j_2, K' \mid j_1 j_2 jK) \\ &\quad \times \left(\frac{\lambda_1'}{2}, \frac{K-K'}{2}, \frac{\lambda_2'}{2}, \frac{K'}{2} \,\middle|\, \frac{\lambda_1'}{2}\frac{\lambda_2'}{2}\frac{\lambda'}{2}\frac{K_2}{2}\right) \\ &\quad \times \left(\frac{\lambda_1'}{2}, \frac{\sigma_1}{2}, \frac{\lambda_2'}{2}, \frac{\sigma-\sigma_1}{2} \,\middle|\, \frac{\lambda_1'}{2}\frac{\lambda_2'}{2}\frac{\lambda'}{2}\frac{\sigma}{2}\right) \\ &\quad \times \alpha_{K-K'}^{\lambda_1'\sigma_1 j_1\omega}(\lambda_1)\, \alpha_{K'}^{\lambda_2',\sigma-\sigma_1,j_2\omega_2}(\lambda_2) \\ &\quad - \sum_{\omega'} \sum_{\lambda''=|\sigma|}^{\lambda-2} \beta^{\lambda''|\sigma|j\omega_1\omega_2}\, \alpha_K^{\lambda'\sigma j\omega'}(\lambda''), \quad (32)\end{aligned}$$

where

$$\beta^{\lambda''\sigma j\omega_1\omega_2\omega}$$

$$= \left\{ \sum_{\lambda''=|\sigma|}^{\lambda} \sum_{\lambda'=|\sigma|}^{\lambda} \sum_{K} \alpha_K^{\lambda''\sigma j\omega}(\lambda'')\, \alpha_K^{\lambda'\sigma j\omega}(\lambda'')\, [\lambda''', \lambda'; \sigma, K] \right\}^{-1}$$

$$\times \sum_{\lambda'=|\sigma|}^{\lambda''} \sum_{\lambda_1'=|\sigma_1|}^{\lambda_1} \sum_{\lambda_2'=|\sigma-\sigma_1|}^{\lambda_2} \sum_{\lambda''=|\sigma|}^{\lambda''} \sum_{K} \sum_{K'} \alpha_{K-K'}^{\lambda_1'\sigma_1 j_1\omega_1}(\lambda_1)$$

$$\times \alpha_{K'}^{\lambda_2'\sigma-\sigma_1 j\omega_1}(\lambda_2)\, \alpha_K^{\lambda'''\sigma j\omega}(\lambda'')[\lambda', \lambda'''; \sigma, K]$$

$$\times (j_1 K - K' j_2 K' \,|\, j_1 j_2 j K)$$

$$\times \left(\frac{\lambda_1'}{2}\frac{K-K'}{2}\frac{\lambda_2'}{2}\frac{K'}{2} \,\middle|\, \frac{\lambda_1'}{2}\frac{\lambda_2'}{2}\frac{\lambda'}{2}\frac{K}{2} \right)$$

$$\times \left(\frac{\lambda_1'}{2}\frac{\sigma_1}{2}\frac{\lambda_2'}{\sigma}\frac{\sigma-\sigma_1}{2} \,\middle|\, \frac{\lambda_1'}{2}\frac{\lambda_2'}{2}\frac{\lambda'}{2}\frac{\sigma}{2} \right), \tag{33}$$

if $\lambda \geq j$, and zero otherwise; also $\lambda_1 + \lambda_2 = \lambda$ and $K(K')$ runs from $-j$ to $j(-j_2$ to $j_2)$ if $\lambda + j(\lambda_2 + j_2)$ are even, from $-j+1$ to $j-1$ $(-j_2+1$ to $j_2-1)$ if $\lambda + j(\lambda_2 + j_2)$ are odd. The symbol $[\lambda', \lambda'''; \sigma, K]$, which represents the "overlap" of the $\mathfrak{D}^{\lambda/2}_{\sigma/2,K/2}$, is defined as

$$[\lambda', \lambda'''; \sigma, K] \equiv (\mathfrak{D}^{\lambda'/2}_{\sigma/2,K/2}, \mathfrak{D}^{\lambda'''/2}_{\sigma/2,K/2})$$

when the domain of integration and the density function are as given in Sec. II.

The possibility of employing different combinations of λ_1 and λ_2 gives rise to the existence of multiple solutions in certain cases. The number of possible solutions for various combinations of λ, σ, and j has been obtained by Zickendraht[3] and is listed in Appendix A for easy reference. Note the difference between our λ and σ and Zickendraht's λ and μ. A systematic procedure for finding the multiple solutions has not yet been developed. Our approach is based partially on the realization that the sequence in which the ξ_m^μ are coupled in building up the wavefunctions determines the eigenvalue ω.

It now remains to orthogonalize the $\chi_{\lambda\sigma jm}$ with respect to ω by means of the Schmidt procedure. First, we select one of the χ's as an eigenfunction corresponding to the first solution of the set; that is,

$$\psi_{\lambda\sigma jm\omega_1} = \chi_{\lambda\sigma jm(1)}$$

or

$$\alpha_K^{\lambda''\sigma j\omega_1}(\lambda) = \gamma_K^{\lambda''\sigma j(1)}(\lambda). \tag{34}$$

Then we compute $\psi^{\lambda\sigma jm\omega_2}$ such that it is orthogonal to $\psi^{\lambda\sigma jm\omega_1}$ with respect to ω:

$$\psi_{\lambda\sigma j\omega_2} = \chi_{\lambda\sigma jm(2)} - \psi_{\lambda\sigma jm\omega_1}(\chi_{\lambda\sigma jm(2)}, \psi_{\lambda\sigma jm\omega_1}) \times (\psi_{\lambda\sigma jm\omega_1}, \psi_{\lambda\sigma jm\omega_1})^{-1}. \tag{35}$$

The process is repeated until all unique solutions corresponding to given $(\lambda\sigma j)$ have been obtained after which they are normalized (see Appendix A for an exhaustive listing of the number of unique solutions for each case). The α's for the ith solution are then

$$\alpha_K^{\lambda'\sigma j\omega_i}(\lambda) = \gamma_K^{\lambda'\sigma j(i)}(\lambda) - \sum_{l=1}^{i-1} \left[\alpha_K^{\lambda'\sigma j\omega_l}(\lambda) \sum_{\lambda''\lambda'''} \sum_{K'} \alpha_{K'}^{\lambda''\sigma j\omega_l}(\lambda)\, \gamma_{K'}^{\lambda''\sigma j(i)}(\lambda) \times [\lambda'', \lambda'''; \sigma, K] \right] \left[\sum_{\lambda''\lambda'''} \sum_{K'} \alpha_{K'}^{\lambda''\sigma j\omega_l}(\lambda) \alpha_{K'}^{\lambda'''\sigma j\omega_l}(\lambda) \times [\lambda'', \lambda'''; \sigma, K] \right]^{-1}. \tag{36}$$

The method of building-up followed by orthogonalization with respect to λ and ω was employed to construct the eigenfunctions listed in Appendix B.

An alternative approach is to construct the $\psi^{\lambda,\pm\lambda,j,m}$ by the building-up process discussed previously. The ladder operators, given by Eq. (19) or, equivalently, by

$$\Xi_m^\mu = 2(-1)^{(1+\mu)/2} \sum_{m'} (-1)^{m'} \times (1, m', 1, m - m' \,|\, 111m) \frac{\partial}{\partial \xi_{m-m'}^{-\mu}}, \tag{37}$$

are used to raise or lower, as desired, the quantum number σ. Since Ξ does not commute with L^2, a linear combination of states of OAM $l-1$ and $l+1$ result from the operation of Ξ_m^μ on the eigenfunction; the one corresponding to the desired eigenvalue of L^2 is then selected. The process of building-up followed by orthogonalization appears to be the simpler of the two approaches and has thus been employed to compute the eigenfunctions listed in Appendix B.

There is yet another way to attack the problem, and this has its origin in Eq. (22) for Λ^2 and the ansatz (23). This is very similar in spirit to Zickendraht's approach, but is distinguished from it by leading to a recurrence relation among the expansion coefficients for general $(\lambda\sigma j\omega)$. After some straightforward but

tedious algebraic manipulation, one obtains the result:

$$\begin{aligned}
&\frac{2}{3}\sum_{\lambda'=\sigma}^{\lambda}\alpha_k^{\lambda'\sigma j\omega}(\lambda)\,[(\lambda-\lambda')(\lambda+\lambda'+2)+2\lambda]\\
&\quad\times\left(\frac{\lambda'}{2},\frac{\sigma}{2},2,0\,\middle|\,\frac{\lambda'}{2},2,\frac{\lambda''}{2},\frac{\sigma}{2}\right)\\
&\quad\times\left(\frac{\lambda'}{2},\frac{k}{2},2,0\,\middle|\,\frac{\lambda'}{2},2,\frac{\lambda''}{2},\frac{k}{2}\right)\\
&\quad+\tfrac{1}{3}\alpha_k^{\lambda''\sigma j\omega}(\lambda)\,[(\lambda-\lambda'')(\lambda+\lambda''+2)+2\lambda]\\
&\quad+2\sum_{\lambda'=\sigma}^{\lambda}\alpha_k^{\lambda'\sigma j\omega}(\lambda)\sqrt{\tfrac{1}{3}}\,\lambda'(\lambda'+2)\\
&\quad\times\left[\left(\frac{\lambda'}{2},\frac{k-2}{2},1,1\,\middle|\,\frac{\lambda'}{2},1,\frac{\lambda'}{2},\frac{k}{2}\right)\right.\\
&\quad\times\left(\frac{\lambda'}{2},\frac{k-2}{2},2,1\,\middle|\,\frac{\lambda'}{2},2,\frac{\lambda''}{2},\frac{k}{2}\right)\\
&\quad-\left(\frac{\lambda'}{2},\frac{k+2}{2},1,-1\,\middle|\,\frac{\lambda'}{2},1,\frac{\lambda'}{2},\frac{k}{2}\right)\\
&\quad\left.\times\left(\frac{\lambda'}{2},\frac{k+2}{2},2,1\,\middle|\,\frac{\lambda'}{2},2,\frac{\lambda''}{2},\frac{k}{2}\right)\right]\\
&\quad\times\left(\frac{\lambda'}{2},\frac{\sigma}{2},2,0\,\middle|\,\frac{\lambda'}{2},2,\frac{\lambda''}{2},\frac{\sigma}{2}\right)-2\alpha_k^{\lambda''\sigma j\omega}(\lambda)\\
&\quad\times[j(j+1)-k^2]+2\sqrt{2}\,j(j+1)\sum_{\lambda'}\left[\alpha_{k-2}^{\lambda'\sigma j\omega}(\lambda)\right.\\
&\quad\times(j,k,1,-1\,|\,j,1,j,k-1)\\
&\quad\times(j,k-1,1,-1\,|\,j,1,j,k-2)\\
&\quad\times\left(\frac{\lambda'}{2},\frac{k-2}{2},1,1\,\middle|\,\frac{\lambda'}{2},1,\frac{\lambda''}{2},\frac{k}{2}\right)-\alpha_{k+2}^{\lambda'\sigma j\omega}(\lambda)\\
&\quad\times(j,k,1,1\,|\,j,1,j,k+1)\\
&\quad\times(j,k+1,1,1\,|\,j,1,j,k+2)\\
&\quad\left.\times\left(\frac{\lambda'}{2},\frac{k+2}{2},1,-1\,\middle|\,\frac{\lambda'}{2},1,\frac{\lambda''}{2},\frac{k}{2}\right)\right]\\
&\quad\times\left(\frac{\lambda'}{2},\frac{\sigma}{2},1,0\,\middle|\,\frac{\lambda'}{2},1,\frac{\lambda''}{2},\frac{\sigma}{2}\right)=0, \qquad (38)
\end{aligned}$$

which can, in principle, be employed to compute all the expansion coefficients for any set of quantum numbers. Although it would be extremely tedious in practice, it can be used to verify that the eigenfunctions obtained by taking tensor products do satisfy the eigenvalue equation $\Lambda^2\psi=\lambda(\lambda+4)\psi$. This was done for a number of the lower-order eigenfunctions listed in Appendix B.

We conclude this section by showing how to evaluate the eigenvalues of Σ_d. The scalar product $\mathbf{L}\cdot\boldsymbol{\Sigma}\cdot\mathbf{L}=\Sigma_d$ is first expanded as

$$\Sigma_d=\sum_{\mu\nu}(-1)^{\mu}(1,\nu,2,\mu-\nu\,|\,121\mu)\,L_{-\mu}\Sigma^2_{\mu-\nu} \qquad (39)$$

and allowed to operate on the function $\psi^{\lambda\sigma j\omega}$. The result,

$$\begin{aligned}
\Sigma_d\psi^{\lambda\sigma jm\omega}=&\sum_{\substack{k\\ \mu\nu}}\alpha_k^{\lambda'\sigma j\omega}(\lambda)\,(-)^{\mu+\nu}\\
&\times\{(1,\mu+\nu,2,-\nu\,|\,121,\mu)\\
&\times(j,m+\mu+\nu,1,-\mu-\nu\,|\,j1jm)\\
&\times(j,m+\nu,1,\mu\,|\,j1j,m+\mu+\nu)\\
&\times j(j+1)+(1,\nu,2,\mu-\nu\,|\,121\mu)\\
&\times(2,-\nu,1,\mu\,|\,212,\mu-\nu)\\
&\times(j,m+\nu,1,-\nu\,|\,j1jm)\\
&\times[j(j+1)]^{\frac{1}{2}}\}\sum_{-\mu}^{2}D^j_{m+\nu,k}\,\mathfrak{D}^{\lambda'/2}_{\sigma/2,k/2}, \qquad (40)
\end{aligned}$$

can be simplified by substituting the following form of $\boldsymbol{\Sigma}^2$ [see Eq. (18)]:

$$\Sigma^2_{-\mu}=\sum_r O_r D^2_{-\mu,r}, \qquad (41)$$

where O_r is an appropriate operator. Then

$$\begin{aligned}
\Sigma_d\psi^{\lambda\sigma jm\omega}=&-\frac{[(2j-1)(2j+3)j(j+1)]^{\frac{1}{2}}}{10}\\
&\times\sum_{k\lambda' j'}\alpha_k^{\lambda'\sigma j\omega}(\lambda)\,Q^{\lambda' j'}_{\sigma mk}, \qquad (42)
\end{aligned}$$

where

$$Q^{\lambda' j}_{\sigma mk}=\sum_r(j,k,2,r\,|\,j,2,j,k+r)\,O_r\,D^j_{m,k+r}\,\mathfrak{D}^{\lambda 1/2}_{\sigma/2,k/2}. \qquad (43)$$

Next we let

$$P^j\sum_0^2\psi^{\lambda\sigma jm\omega}=(j,m,2,0\,|\,j2jm)\sum_{k\lambda'}\alpha_k^{\lambda'\sigma j\omega}(\lambda)\,Q^{\lambda' j}_{\sigma mk}. \qquad (44)$$

Comparison with Eq. (42) shows that

$$\Sigma_d \psi^{\lambda\sigma j m\omega} = \frac{j(j+1)(2j-1)(2j+3)}{j(j+1)-m^2}\frac{1}{\sqrt{10}} \times P^j \sum_0^2 \psi^{\lambda\sigma j m\omega}, \quad (45)$$

which leads immediately to the eigenvalues ω by application of the form of Σ_0^2 given by Eq. (18) to the eigenfunctions. Because the eigenvalues of Σ_d are not of great importance for present purposes, this program has not been carried out for the eigenfunctions listed in Appendix B. This aspect of the problem will be more fully developed in a future paper.

V. ADDITION OF GAM

It is apparent that Eq. (31) is very close to yielding an addition theorem for GAM. We complete the derivation by first rewriting Eq. (36) as

$$\chi_{\lambda\sigma j m\omega_r} = \psi_{\lambda\sigma j m\omega_r} + \sum_{s=1}^{r-1} (\chi_{\lambda\sigma j m\omega_r}, \psi_{\lambda\sigma j m\omega_s}),$$

$$\psi_{\lambda\sigma j m\omega_s} = \sum_{s=1}^{r} R_{rs}\, \psi_{\lambda\sigma j m\omega_s}, \quad (46)$$

where

$$R_{rs} = \delta_{rs} + (\chi_{\lambda\sigma j m\omega_r}, \psi_{\lambda\sigma j m\omega_s}), \quad (s < r), \quad (47)$$

and the ψ's are assumed to be normalized. We now substitute the above expression for χ into Eq. (31), multiply by both sides of the equation by $(j_1, m_1', j_2, m - m_2' \,|\, j_1 j_2 j, m)$, and sum over m_1. The result is [using the orthogonality properties of the $SU(2)$ Clebsch–Gordan coefficients]

$$\psi_{\lambda_1\sigma_1 j_1 m_1\omega}\psi_{\lambda_2,\sigma-\sigma_1,j_2,m-m_1,\omega_2} = \sum_{js}(j_1, m_1, j_2, m - m_1 \,|\, j_1 j_2 j m) \times \left(R_{rs}\,\psi^{\lambda\sigma j m\omega_s} + \sum_{\lambda'} \beta^{\lambda'\sigma j m\omega_1\omega_2\omega_s}\,\psi^{\lambda'\sigma j m\omega_s}\right)$$

$$= \sum_{\lambda js}(j_1, m_1, j_2, m - m_1 \,|\, j_1, j_2, j, m) \times \begin{pmatrix} \lambda_1\omega_1; & \lambda_2\omega_2; & \lambda\omega \\ j_1\sigma_1; & j_2\sigma_2; & j\sigma \end{pmatrix} \psi^{\lambda\sigma j m\omega_s}, \quad (48)$$

where the subscript r refers to the combination ω_1, ω_2.

Equation (48) is not, of course, a prescription for computing the GAM coupling coefficients

$$\begin{pmatrix} \lambda_1\omega_1; & \lambda_2\omega_2; & \lambda\omega \\ j_1\sigma_1; & j_2\sigma_2; & j\sigma \end{pmatrix},$$

because we have not specified a method of finding the matrix elements P_{rs}; Eq. (47) is only an implicit relationship. Equation (48) does, however, demonstrate the existence of the coupling coefficients. It also suggests a way of computing them. Suppose we expand the GAM eigenfunction on both sides as in Eq. (23), and then take the scalar product with the function $\psi^{\lambda'\sigma' j' m'\omega'}$. The result is

$$\begin{pmatrix} \lambda_1\omega_1; & \lambda_2\omega_2; & \lambda\omega \\ j_1\sigma_1; & j_2\sigma_2; & j\sigma \end{pmatrix} = \frac{8\pi^2}{2j+1}\delta_{\sigma_2,\sigma-\sigma_1}\sum_{\lambda'}\sum_{\lambda''}\sum_{\lambda_1'}\sum_{\lambda_2'}\sum_{K,K'} \times \alpha_K^{\lambda'\sigma j\omega}(\lambda)\alpha_{K'}^{\lambda_1'\sigma_1 j_1\omega_1}(\lambda_1)\alpha_{K-K'}^{\lambda_2'\sigma-\sigma_1 j_2\omega_2}(\lambda_2) \times (j_1 K_1, j_2 K - K_1 \,|\, j_1 j_2 j K) \times \left(\frac{\lambda_1'}{2}, \frac{K'}{2}, \frac{\lambda_2'}{2}, \frac{K-K'}{2} \,\middle|\, \frac{\lambda_1'}{2}\frac{\lambda_2'}{2}\frac{\lambda'}{2}\frac{K}{2}\right) \times \left(\frac{\lambda_1'}{2}, \frac{\sigma_1}{2}, \frac{\lambda_2'}{2}, \frac{\sigma-\sigma_1}{2} \,\middle|\, \frac{\lambda_1'}{2}, \frac{\lambda_2'}{2}, \frac{\lambda'}{2}\frac{\sigma}{2}\right) \times [\lambda', \lambda''; \sigma, K], \quad (49)$$

from which all the coupling coefficients can, in principle, be computed.

We close this section by using the coupling coefficients to show the manner in which the eigenvalues ω combine. If we let Σ_d operate on both sides of Eq. (48), we obtain

$$(\omega_1 + \omega_2)\,\psi_{\lambda_1\sigma_1 j_1 m_1\omega_1}\,\psi_{\lambda_2\sigma-\sigma_1,j_2,m-m_1\omega_2} = \sum_{\lambda j\omega}(j_1, m_1 j_2, m - m_1 \,|\, j_1 j_2 j m) \times \begin{pmatrix} \lambda_1\omega_1; & \lambda_2\omega_2; & \lambda\omega \\ j_1\sigma_1; & j_2\sigma_2; & j\sigma \end{pmatrix} \omega\,\psi_{\lambda\sigma j m\omega}. \quad (50)$$

Employing the orthogonality of the coupling coefficients, one further obtains

$$\omega = \sum_{\sigma_1\omega_1\omega_2} \begin{pmatrix} \lambda_1\omega_1; & \lambda_2\omega_2; & \lambda\omega \\ j_1\sigma_1; & j_2\sigma_2; & j\sigma \end{pmatrix}(\omega_1 + \omega_2). \quad (51)$$

Additional properties of the GAM operators, eigenfunctions and coupling coefficients will be explored in future papers.

ACKNOWLEDGMENT

The authors are grateful to Dr. F. M. Chilton for helpful discussions.

APPENDIX A: NUMBER OF POSSIBLE SOLUTIONS FOR VARIOUS VALUES OF $\lambda\sigma j^4$

The number n of possible solutions for positive and negative σ are added together. For $\sigma = 0$, the number of solutions must be divided by 2.

Case 1: j odd, λ odd.

$\lvert\sigma\rvert$	n
λ	2
$\lambda - 2$	2
$\lambda - 4$	4
$\lambda - 6$	4
$\lambda - 8$	6
$\lambda - 10$	6
.	
.	
.	

$$n_{\max} = j + 1 \quad \text{for} \quad j \leq \frac{\lambda + 1}{2}$$

$$n_{\max} = \lambda - j + 2 \quad \text{for} \quad j \geq \frac{\lambda + 1}{2}.$$

Case 2: j even, λ even.

$\lvert\sigma\rvert$	n
λ	2
$\lambda - 4$	4
$\lambda - 8$	6
$\lambda - 12$	8
.	
.	
.	

$$n_{\max} = j + 2 \quad \text{for} \quad j \leq \lambda/2$$

$$n_{\max} = \lambda - 1 + 2 \quad \text{for} \quad j \geq \lambda/2.$$

$\lvert\sigma\rvert$	n
$\lambda - 2$	2
$\lambda - 6$	4
$\lambda - 10$	6
.	
.	
.	

$$n_{\max} = j \quad \text{for} \quad j \leq \lambda/2 + 1$$

$$n_{\max} = \lambda - j + 2 \quad \text{for} \quad j \geq \lambda/2 + 1.$$

Case 3: j odd, λ even.

$\lvert\sigma\rvert$	n
$\lambda - 2$	2
$\lambda - 6$	4
$\lambda - 10$	6
.	
.	
.	

$$n_{\max} = j + 1 \quad \text{for} \quad j \leq \frac{\lambda}{2}$$

$$n_{\max} = \lambda - j + 1 \quad \text{for} \quad j \geq \frac{\lambda}{2}$$

$\lvert\sigma\rvert$	n
$\lambda - 4$	2
$\lambda - 8$	4
.	
.	
.	

$$n_{\max} = j - 1 \quad \text{for} \quad j \leq \frac{\lambda}{2} + 1$$

$$n_{\max} = \lambda - j + 1 \quad \text{for} \quad j \geq \frac{\lambda}{2} + 1.$$

Case 4: j even, λ odd.

$\lvert\sigma\rvert$	n
$\lambda - 2$	2
$\lambda - 4$	2
$\lambda - 6$	4
$\lambda - 8$	4
$\lambda - 10$	6
$\lambda - 12$	6
.	
.	
.	

$$n_{\max} = j \quad \text{for} \quad j \leq \frac{\lambda + 1}{2}$$

$$n_{\max} = \lambda - j + 1 \quad \text{for} \quad j \geq \frac{\lambda + 1}{2}.$$

APPENDIX B: GAM-EIGENFUNCTION EXPANSION COEFFICIENTS

Unnormalized expansion coefficients of the GAM eigenfunctions for $\lambda \leq 4$ are presented in the following listing. The quantum number ω is suppressed unless the eigenfunctions are otherwise degenerate. Otherwise, only a solution number in parentheses, [i.e., (2) as in $[\alpha_K^{\lambda^\tau \sigma j(2)}(\lambda)]$, rather than the value of ω itself] is given. The other index (λ) in parentheses is also suppressed since no ambiguity can result.

$\lambda\sigma j$	
000	$\alpha_0^{000} = 1$
111	$\alpha_1^{111} = 1;\ \alpha_{-1}^{111} = -1$
1–11	$\alpha_1^{1-11} = 1;\ \alpha_{-1}^{1-11} = 1$
222	$\alpha_2^{222} = 1;\ \alpha^{222} = -1/\sqrt{3};\ \alpha_{-2}^{222} = 1$
2–22	$\alpha_2^{2-22} = 1;\ \alpha_0^{2-22} = 1/\sqrt{3};\ \alpha_{-2}^{2-22} = 1$
220	$\alpha_0^{220} = 1$
2–20	$\alpha_0^{2-20} = 1$
202	$\alpha_2^{202} = 1;\ \alpha_0^{002} = 1/\sqrt{3};\ \alpha_{-2}^{202} = -1$
201	$\alpha_0^{201} = 1$
333	$\alpha_3^{333} = 1;\ \alpha_1^{333} = -1/\sqrt{5};\ \alpha_{-1}^{333} = 1/\sqrt{5};\ \alpha_{-3}^{333} = -1$
3–33	$\alpha_3^{3-33} = 1;\ \alpha_1^{3-33} = 1/\sqrt{5};\ \alpha_{-1}^{3-33} = 1/\sqrt{5};\ \alpha_{-3}^{3-33} = 1$
331	$\alpha_1^{331} = -1;\ \alpha_{-1}^{331} = 1$
3–31	$\alpha_1^{3-31} = 1;\ \alpha_1^{3-31} = 1$
313	$\alpha_3^{313} = 1;\ \alpha_1^{313} = -1/\sqrt{45};\ \alpha_{-1}^{313} = -1/\sqrt{45};\ \alpha_{-3}^{313} = 1$ $\alpha_1^{113} = 4/\sqrt{45};\ \alpha_{-1}^{113} = -4/\sqrt{45}$
3–13	$\alpha_3^{3-13} = 1;\ \alpha_1^{3-13} = 1/\sqrt{45};\ \alpha_{-1}^{3-13} = -1/\sqrt{45};\ \alpha_{-3}^{313} = -1$ $\alpha_1^{1-13} = 4/\sqrt{45};\ \alpha_{-1}^{1-13} = 4/\sqrt{45}$
312	$\alpha_1^{312} = 1;\ \alpha_{-1}^{312} = -1;\ \alpha_1^{112} = \tfrac{1}{2};\ \alpha_{-1}^{112} = \tfrac{1}{2}$
3–12	$\alpha_1^{3-12} = -1;\ \alpha_{-1}^{3-12} = 1;\ \alpha_1^{1-12} = -\tfrac{1}{2};\ \alpha_{-1}^{1-12} = \tfrac{1}{2}$
311	$\alpha_1^{311} = 1;\ \alpha_{-1}^{311} = 1$
3–11	$\alpha_1^{3-11} = 1;\ \alpha_{-1}^{3-11} = -1$
444	$\alpha_4^{444} = 1;\ \alpha_2^{444} = -1/\sqrt{7};\ \alpha_0^{444} = (3/35)^{\frac{1}{2}};\ \alpha_{-2}^{444} = -1/\sqrt{7};\ \alpha_{-4}^{444} = 1$
4–44	$\alpha_4^{4-44} = 1;\ \alpha_2^{4-44} = 1/\sqrt{7};\ \alpha_0^{4-44} = (3/35)^{\frac{1}{2}};\ \alpha_{-2}^{4-44} = 1/\sqrt{7};\ \alpha_{-4}^{4-44} = 1$
442	$\alpha_2^{444} = 1;\ \alpha_0^{444} = -\tfrac{2}{3};\ \alpha_{-2}^{444} = 1$
4–42	$\alpha_2^{4-42} = 1;\ \alpha_0^{4-44} = \tfrac{2}{3};\ \alpha_{-2}^{4-44} = 1$
440	$\alpha_0^{440} = 1$
4–40	$\alpha_0^{4-40} = 1$
424	$\alpha_4^{424} = 1;\ \alpha_2^{424} = -\tfrac{1}{2}\sqrt{7};\ \alpha_{-2}^{424} = \tfrac{1}{2}\sqrt{7};\ \alpha_{-4}^{424} = -1$ $\alpha_2^{224} = \tfrac{3}{2}\sqrt{7};\ \alpha_0^{224} = -\tfrac{3}{2}\sqrt{35};\ \alpha_{-2}^{224} = \tfrac{3}{2}\sqrt{7}$
4–24	$\alpha_4^{4-24} = 1;\ \alpha_2^{4-24} = \tfrac{1}{2}\sqrt{7};\ \alpha_{-2}^{4-24} = -\tfrac{1}{2}\sqrt{7};\ \alpha_{-4}^{4-24} = -1$ $\alpha_2^{2-24} = \tfrac{3}{2}\sqrt{7};\ \alpha_0^{2-24} = \tfrac{3}{2}\sqrt{35};\ \alpha_{-2}^{2-24} = \tfrac{3}{2}\sqrt{7}$
423	$\alpha_2^{423} = 1;\ \alpha_0^{423} = -2/\sqrt{5};\ \alpha_{-2}^{423} = 1;\ \alpha_2^{223} = 1;\ \alpha_{-2}^{223} = -1$
4–23	$\alpha_2^{4-23} = 1;\ \alpha_0^{4-23} = 2/\sqrt{5};\ \alpha_{-2}^{4-23} = 1;\ \alpha_2^{2-23} = -1;\ \alpha_{-2}^{2-23} = 1$
422	$\alpha_2^{422} = 1;\ \alpha_{-2}^{422} = -1;\ \alpha_2^{222} = \tfrac{1}{7};\ \alpha_0^{222} = \tfrac{2\sqrt{3}}{7};\ \alpha_{-2}^{222} = \tfrac{1}{7}$
4–22	$\alpha_2^{4-22} = -1;\ \alpha_{-2}^{4-22} = 1;\ \alpha_2^{2-22} = \tfrac{1}{7};\ \alpha_0^{2-22} = -\tfrac{2\sqrt{3}}{7};$ $\alpha_{-2}^{2-22} = \tfrac{1}{7}$
421	$\alpha_0^{421} = -1;\ \alpha_0^{221} = \sqrt{3}$
4–21	$\alpha_0^{4-21} = 1;\ \alpha_0^{2-21} = \sqrt{3}$
404	$\alpha_4^{404} = 1;\ \alpha_0^{404} = -1/\sqrt{105};\ \alpha_{-4}^{404} = 1;\ \alpha_2^{204} = \sqrt{\tfrac{3}{7}};\ \alpha_{-2}^{204} = -\sqrt{\tfrac{3}{7}}$ $\alpha_0^{004} = 4/\sqrt{105}$
403	$\alpha_2^{403} = 1;\ \alpha_{-2}^{403} = -1;\ \alpha_0^{203} = 3/\sqrt{5}$

$\lambda\sigma j$

402(1) $\alpha_2^{402} = 1$; $\alpha_0^{402} = -\frac{2}{7}$; $\alpha_{-2}^{402} = 1$; $\alpha_2^{202} = -\dfrac{3}{35\sqrt{3}}$

$\alpha_{-2}^{202} = \dfrac{3}{35\sqrt{3}}$; $\alpha_0^{002} = -\frac{6}{35}$

(2) $\alpha_2^{402} = 1$; $\alpha_0^{402} = -\dfrac{892}{35\cdot 71}\alpha_{-2}^{402} = 1$; $\alpha_2^{202} = \frac{26}{35}\cdot\frac{334}{355}\sqrt{3}$

$\alpha_{-2}^{202} = -\frac{26}{35}\frac{334}{355}\sqrt{3}$; $\alpha_0^{002} = \dfrac{46}{35\cdot 71}$

400 $\alpha_0^{400} = 1$.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 7 JULY 1968

Nearest-Neighbor Spacing Distribution for an Ensemble of Random Matrices with a Nonrandom Bias*

J. F. McDONALD† AND L. D. FAVRO**
Department of Physics, Wayne State University, Detroit, Michigan

(Received 15 June 1967)

The nearest-neighbor spacing distribution for the biased Gaussian distribution $\exp[-\gamma \operatorname{Tr}(H-H_0)^2]$ is calculated in the limits of large and small γ. The results for the small-γ limit are found to approach those for an unbiased Gaussian distribution (i.e., H_0 proportional to the unit matrix). To order γ^2 the mth-order spacing distribution is expressed in terms of the mth-order spacing distribution for the unbiased Gaussian ensemble. The results for the large-γ limit are found to depend strongly on H_0. That is, the nearest-neighbor spacing distribution reflects the structure of H_0. Thus, a biased distribution offers a possible description for the experimental deviations from the Wigner surmise and apparent multiple-peak structure exhibited by some nuclear-spacing data.

1. INTRODUCTION

There has been much success in predicting the statistical properties of the highly excited energy levels of complex systems using ensembles of random matrices.[1] In particular, the nearest-neighbor spacing distribution for levels with the same symmetry quantum numbers in many heavy nuclei is found to be fitted within the experimental accuracy by the so-called Wigner surmise

$$p(x) = \frac{\pi}{2} x \exp(-\pi x^2/4), \tag{1}$$

where $x = S/\bar{S}$ with S the spacing and $\bar{S}$ the average spacing.[2] The calculation of the nearest-neighbor spacing distribution from various ensembles of random matrices yield numerical results which agree with (1) in the region of major probability ($S < 3\bar{S}$) to within the experimental error.[3]

It is also found, however, that there are experimental deviations from (1). In particular, certain nuclear spacing data appears to have a multiple-peak structure.[4] Also, deviations from (1) occur in cases where some observables are only approximately conserved.[5] We shall be concerned only with the first type of deviation in this paper. The second problem will be discussed in a second paper.

The usual approach in the statistical theory of energy levels assumes that nothing is known about the Hamiltonian of the system under consideration with the possible exception of its symmetry properties. In what will be done here, it is assumed that more information is known. In particular, it is assumed that we are given the eigenvalues λ_n of some Hamiltonian H_0, where H_0 is some approximation to the actual Hamiltonian of the system being considered and which reflects the underlying structure of the energy spectrum. The lack of complete knowledge again leads to the use of a statistical approach exactly as before, with the exception that the information contained in H_0 must be included. That is, the matrix-element distribution must be biased by H_0.

The deviations of the ensemble from H_0 could be of any magnitude in general. When the deviations are small, one expects to see the structure of H_0 very strongly. On the other hand, when the deviations are large, one expects the structure to be washed out and complete randomness to return to the ensemble.

The general procedure which will be followed is the same as that for unbiased ensembles, with the exception that a biased matrix-element distribution will be used. Since the general case seems to be intractable, we will consider only certain limiting cases and restrict our discussion primarily to the nearest-neighbor spacing distribution for a biased Gaussian ensemble.

2. DISTRIBUTION

In the following we shall consider only the orthogonal ensemble.[6] Thus we need consider only real symmetric matrices. This corresponds to systems

* Some of this material is based on portions of a thesis presented by one of the authors (J. F. M.) to Wayne State University in partial fulfillment of the requirements for the Ph.D. degree.

† Present address: Department of Mathematics, University of Windsor, Windsor, Ontario.

** Supported in part by a Faculty Research Fellowship, Wayne State University.

[1] Most of the pertinent papers up through 1963, as well as an excellent introductory review of this subject, can be found in C. E. Porter, *Statistical Theories of Spectra: Fluctuations* (Academic Press Inc., New York, 1965).

[2] E. P. Wigner, "Conference on Neutron Physics by Time of Flight, Gatlinburg, Tennessee, November 1956"; see Oak Ridge Natl. Lab. Report ORNL-2309 (1957), p. 67.

[3] See, for example, M. L. Mehta, Nucl. Phys. **18**, 395 (1960); M. Gaudin, Nucl. Phys. **25**, 447 (1961); H. S. Leff, J. Math. Phys. **5**, 763 (1964).

[4] See, for example, J. B. Garg *et al.*, Phys. Rev. **134**, B985 (1964).

[5] N. Rosenzweig and C. E. Porter, Phys. Rev. **120**, 1698 (1960).

[6] F. J. Dyson, J. Math. Phys. **3**, 140 (1962).

which are time-reversal invariant and either rotationally invariant or have integral spin. In addition, we will restrict our attention to the biased analog of the Gaussian ensemble[7]

$$P(H, H_0, \gamma)\, dH = \eta\{\exp[-\gamma \operatorname{Tr}(H - H_0)^2]\}\, dH, \tag{2}$$

where

$$\eta = (\gamma/\pi)^{N/2}(2\gamma/\pi)^{N(N-1)/4} \tag{3}$$

and

$$dH = \prod_{i=1}^{N} dH_{ii} \prod_{j>k} dH_{jk}. \tag{4}$$

The ensemble given by (2) has been investigated for the case when H_0 is proportional to the identity matrix and also for the general case in which H_0 is two-dimensional.[8]

The parameter γ enables us to vary the distribution from the completely nonrandom H_0 (i.e., γ large) to the completely random case (i.e., γ small).[9] It should be noted that this distribution is representationally invariant if H_0 has the same transformation properties as H.

A formal expression for the joint eigenvalue distribution corresponding to (2) can be obtained in the usual manner.[10] The result can be written as

$$P(E, \gamma) = \eta\left(\prod_{i>j} |E_i - E_j|\right)\int h(\phi)\, d\phi \times \exp\left[-\gamma \sum_i \left(E_i^2 + \lambda_i^2 - 2\sum_j A_{ij}^2 \lambda_i E_j\right)\right], \tag{5}$$

where A_{ij} is the matrix which connects the representations in which H and H_0 are diagonal, and $h(\phi)$ is proportional to the weight function for the rotation parameter, ϕ_j. Note that, unlike the unbiased distribution (i.e., the case in which all of the λ_n are degenerate), the integration over the ϕ_j is in general not trivial. For an arbitrary γ and H_0 this integration seems to be intractable. Thus, for a general H_0 we will resort to considering only the limiting cases of large and small γ.

3. DEFINITION OF THE LARGE-γ LIMIT

It is of primary concern to determine how γ restricts the position and spacings of the levels in the ensemble relative to the position and spacings of the levels of H_0. Thus, we would like to calculate quantities such as

$$\langle E_n\rangle, \quad \langle E_n - E_{n-1}\rangle, \quad \langle (E_n - \langle E_n\rangle)^2\rangle$$

and

$$\langle [(E_n - E_{n-1}) - (\langle E_n\rangle - \langle E_{n-1}\rangle)]^2\rangle,$$

where it is assumed that the E_n are ordered in the same way as their subscripts, and where the brackets $\langle\ \rangle$ indicate the average over the ensemble, i.e.,

$$\langle f(H)\rangle \equiv \int_{\text{over ensemble}} f(H) P(H, H_0, \gamma)\, dH. \tag{6}$$

In general, these quantities cannot be calculated from Eq. (6). However, for the large-γ limit we can calculate an expansion for them in inverse powers of γ by expanding the particular $f(H)$ of interest using perturbation theory. That is, if γ is large enough, $H - H_0$ can be treated as a random perturbation [of order $1/(\gamma)^{\frac{1}{2}}$] on H_0. The conditions on γ for the convergence of the expansions obtained in this manner constitute the definition of the large-γ limit.

The convergence of such perturbation expansions, of course, depends strongly on the structure of H_0. It is found that such expansions are meaningful in general only if most of the spacings in the H_0 spectrum fall into two classes. The λ_n are degenerate [i.e., $\gamma(\lambda_n - \lambda_{n-1})^2 \ll 1$] or they are far apart [i.e., $\gamma(\lambda_n - \lambda_{n-1})^2 \gg 1$].

If the perturbation calculations are carried out to order $1/\gamma$, three effects are noted. First, neglecting for now the repulsion between originally degenerate levels, the repulsion between levels which were originally far apart causes the entire spectrum to expand. The mean position of each degenerate level $\bar{E}_n$ is shifted out (relative to the center of the spectrum) from the corresponding λ_n. If Δ is the average spacing of the H_0 spectrum, the levels near the center are shifted an amount of order $1/\gamma\Delta$, while those near the ends of the spectrum are shifted an amount of order $(\ln N)/\gamma\Delta$. However, even though the mean position of levels near the ends of the spectrum are a large distance from the corresponding λ_n, the distance between two adjacent mean positions differs from $\lambda_n - \lambda_{n-1}$ only by an amount of order $1/\gamma\Delta$.

The second effect is due to the repulsion of originally degenerate levels. To first order in perturbation theory this effect may be obtained for an m-fold degenerate level by considering an m-dimensional ordinary Gaussian ensemble with the origin shifted to the corresponding $\bar{E}_n$. That is, the originally degenerate levels have an average spacing of order[11] $(m\gamma)^{-\frac{1}{2}}$ and are centered on the corresponding $\bar{E}_n$. Finally, as long as $[m/(\gamma\Delta^2)]^{\frac{1}{2}} \ll 1$ for all the degenerate levels, the probability of two levels which were originally a large distance apart being close together is small.

[7] The ensemble given by (2) was shown to be the solution to a Brownian motion model for random matrices, where H_0 corresponds to the matrix at $t = 0$; see F. J. Dyson, J. Math. Phys. 3, 1191 (1962).

[8] N. Rosenzweig, Nuovo Cimento 38, 1047 (1965).

[9] The large- and small-γ limits correspond to the small- and large-t limits for Dyson's Brownian motion model.

[10] See, for example, C. E. Porter and N. Rosenzweig, Suomalaisen Tiedeakat. Toimituksia, Ser. AVI, No. 44 (1960).

[11] See, for example, the introductory review of Ref. 1.

These perturbation methods, when applied to the eigenvectors of H, also indicate that the distribution of parameters corresponding to rotations in nondegenerate planes are sharply peaked.[12] The peaks correspond to values for which the E_n and λ_n axes in that plane are parallel or antiparallel. Thus, there are a series of peaks in the integrand of Eq. (5). Each peak corresponds to a particular association of the E_n with the λ_n and a particular alignment of the E_n and λ_n axes for that association. To obtain the *ordered* joint eigenvalue distribution, only one such peak need be considered.

4. CALCULATION OF THE JOINT EIGENVALUE AND NEAREST-NEIGHBOR SPACING DISTRIBUTION IN THE LARGE-γ LIMIT

In the large-γ limit we have a set of large parameters, namely, $\gamma(\lambda_n - \lambda_m)^2 \gg 1$, where $\lambda_n \neq \lambda_m$. We wish to obtain asymptotic expansions for the integrals involved in the joint eigenvalue distribution and the nearest-neighbor spacing distribution in terms of these parameters. The method we will use to obtain the asymptotic expansion is due to Laplace.[13]

We shall calculate the *particular ordered distribution* for which E_n corresponds to λ_n in the limit $\gamma \to \infty$. From the results of the perturbation calculations of the previous section we see that there is a peak in the angular distribution of (5), which corresponds to a parallel alignment of the two sets of axes (i.e., $A_{ij} = \delta_{ij}$ for $\lambda_i \neq \lambda_j$).[14] We take as our rotation parameters ϵ_{ij}, $i > j$, where the ϵ_{ij} are defined such that, to first order, $A_{ii} = 1$ and $A_{ij} = -A_{ji} = \epsilon_{ij}$, $i > j$. From the orthogonality conditions on the A_{ij} it then follows that to second order

$$A_{ii} = 1 - 2\sum_{i>j} \epsilon_{ij}^2, \tag{7}$$

$$A_{ij} = \epsilon_{ij} + B_{ij}, \qquad i > j, \tag{8}$$

$$A_{ij} = -\epsilon_{ji} + C_{ij}, \quad i < j, \tag{9}$$

where B_{ij} and C_{ij} are quadratic in the ϵ_{ij}.

We now expand $\mathrm{Tr}\,(H - H_0)^2$ in terms of the ϵ_{ij} (note that we formally expand in terms of *all* of the ϵ_{ij}). The ϵ_{ij} corresponding to rotations in degenerate planes are in general not small, but they are not explicitly contained in $\mathrm{Tr}\,(H - H_0)^2$ anyway. To obtain a zeroth-order approximation to the integral, one need only retain terms up to quadratic. To second order in the ϵ_{ij}

$$\mathrm{Tr}\,(H - H_0)^2 \simeq \sum_i (E_i - \lambda_i)^2 + 2\sum_{i>j} (\lambda_i - \lambda_j)(E_i - E_j)\epsilon_{ij}^2. \tag{10}$$

Next the weight function in the ϵ_{ij} space must be expanded in terms of those ϵ_{ij} which appear explicitly in (10) (i.e., $\lambda_i \neq \lambda_j$). Since we only want the zeroth-order term in the expansion of the integral, we need only the zeroth-order term in the expansion of the volume element. Also, since the zeroth-order term depends only on those ϵ_{ij} not contained in (10) (i.e., $\lambda_i = \lambda_j$), integration over these parameters will yield only a multiplicative constant. Thus, we need not explicitly determine the term. However, we must show that it is not identically zero. This can be done by evaluating the volume element at the point $\epsilon_{ij} = 0$ for all $i > j$ using the methods of Hurwitz.[15]

Finally, one can extend the limits on the ϵ_{ij} contained in (10) to plus and minus infinity,[16] and carry out the integrations. The result of this integration is the *ordered* joint eigenvalue distribution

$$P(E, \gamma) \sim C\left[\prod_{\lambda_i > \lambda_j} (E_i - E_j)^{\frac{1}{2}}\right]\left[\prod_{\substack{\lambda_k = \lambda_l \\ k>l}} |E_k - E_l|\right] \times \exp\left(-\gamma \sum_i (E_i - \lambda_i)^2\right)[1 + O(1/\gamma\Delta^2)], \tag{11}$$

where C is a constant. Note that the levels corresponding to a particular degenerate λ have not been ordered.

To obtain the nearest-neighbor spacing distribution we again use Laplace's method. Thus, expanding in powers of $E_i - \bar{E}_i$, we obtain

$$P(E, \gamma) \sim C \prod_{\substack{\lambda_k = \lambda_l \\ k>l}} [|E_k - E_l|] \times \exp\left(-\gamma \sum_i (E_i - \bar{E}_i)^2\right)[1 + O(\gamma\Delta^2)], \tag{12}$$

where the $\bar{E}_n$ correspond to the values of E_n for which

$$P(E, \gamma) \Big/ \left[C \prod_{\substack{\lambda_k = \lambda_l \\ k>l}} |E_k - E_l|\right]$$

is a maximum. Note that $\bar{E}_i = \bar{E}_j$ if $\lambda_i = \lambda_j$. Also, it follows from the perturbation calculations of the previous section that $\bar{E}_i - \bar{E}_j = \lambda_i - \lambda_j + O(1/\gamma\Delta)$ if

[12] If the perturbation becomes too large (i.e., if the condition $\gamma\Delta^2 \gg 1$ is violated), the peaks broaden and the approximation which will be developed in the next section will fail. As $\gamma\Delta^2$ decreases, the eigenvectors of H become, more and more, mixtures of eigenvectors of H_0 corresponding to different nondegenerate λ_n, and $H - H_0$ is no longer a small perturbation.

[13] N. G. de Gruijn, *Asymptotic Methods in Analysis* (North-Holland Publishing Company, Amsterdam, 1961), p. 60.

[14] There are other peaks in the distribution. However, they correspond at most to having one or more of the corresponding axes antiparallel instead of parallel. All such peaks are equivalent, and we need consider only one.

[15] A. Hurwitz, *Collection of Modern Mathematical Classics* (McGraw-Hill Book Co., Inc., New York, 1960), p. 186.

[16] The corrections due to the limits are of order $\exp[-\gamma(\lambda_i - \lambda_j)^2]$, $\lambda_i \neq \lambda_j$.

$\lambda_i \neq \lambda_j$. The limits on the $(E_i - \bar{E}_i)$ can be extended to plus and minus infinity. The corrections due to the limits are of order $e^{-\gamma(\lambda_i-\lambda_j)^2}$, $\lambda_i \neq \lambda_j$. Thus, to this approximation, the E_n corresponding to different degenerate levels are statistically independent. Further, the joint distribution for those levels corresponding to a particular degenerate level (say m-fold degenerate) is just the result one obtains for an unbiased Gaussian ensemble of m-dimensional matrices (with the origin of the energy scale shifted to the corresponding $\bar{E}_n$).

For definiteness, let us assume that H_0 is such that it contains degeneracies from one to m, and such that the large spacings are given by a set of numbers Δ_β, $\beta = 1, \cdots, n$. To the order of our approximation two general types of contribution to the spacing distribution must be considered. The first type arises from the degenerate levels and can be written as

$$p_{\mathrm{I}}(s, \gamma) \sim (N-1)^{-1} \sum_{l=2}^{m} (l-1)\alpha_l P_l(S, \gamma), \quad (13)$$

where P_l is the result (normalized to unity) obtained from an l-dimensional unbiased Gaussian ensemble, and where α_l is the number of times an l-fold degenerate level occurs in the H_0 spectrum.

The other general type of contribution arises from defining the spacing as $S = |E_i - E_j|$, where $\lambda_i \neq \lambda_j$, and where λ_i is a member of a p-fold degenerate level and λ_j is a member of a nearest-neighbor q-fold degenerate level. This contribution is calculated by considering only these $p + q$ levels, since the integration over all the other levels just yields a multiplicative constant. This contribution will be denoted as $P_{pq}(S, \gamma, \Delta_\beta)$, where Δ_β is the original spacing between these p-fold and q-fold degenerate levels. The total contribution from such configurations can be written as

$$p_{\mathrm{II}}(S, \gamma) \sim (N-1)^{-1} \sum_{p \geq q, \beta} \alpha_{pq\beta} P_{pq}(S, \gamma, \Delta_\beta), \quad (14)$$

where $\alpha_{pq\beta}$ is the total number of times that such a configuration occurs in the H_0 spectrum. The total spacing distribution $p(S, \gamma)$ is

$$p(S, \gamma) = p_{\mathrm{I}}(S, \gamma) + p_{\mathrm{II}}(S, \gamma). \quad (15)$$

No attempt was made to do calculations with an H_0 based on a particular physical model. Instead, various simple (in terms of the level structure of H_0) models were assumed and the calculations carried out to illustrate the effects of various features (degeneracies, etc.) of H_0 on the final distribution. The results for two such models are given here.

First consider a model in which H_0 is three-dimensional (i.e., $N = 3$) and has a doubly degenerate level a distance Δ_1 from the other level. In the limit $\gamma\Delta_1^2 \gg 1$ the above approximation can be used. For this H_0 there is only one Δ_β, namely, Δ_1. The only nonzero α's are $\alpha_2 = \alpha_{211} = 1$. Thus,

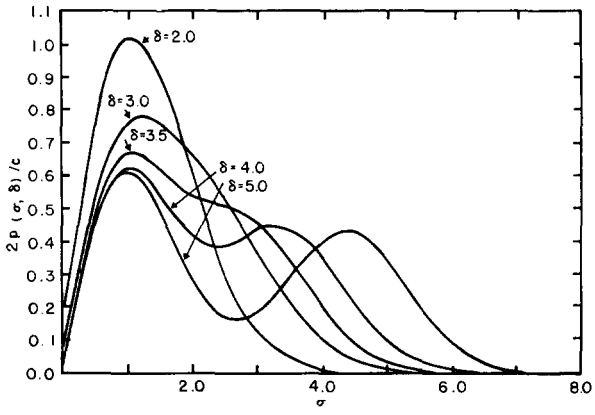


FIG. 1. Plots of $2p(\sigma, \delta)/C$ for a three-dimensional H_0 with a doubly degenerate level at various values of the original spacing δ.

$$p(S, \gamma) \sim \tfrac{1}{2}[P_2(S, \gamma) + P_{12}(S, \gamma, \Delta_1)]. \quad (16)$$

It is easily shown that

$$P_2(\sigma) = C\sigma e^{-\sigma^2/2} \quad (17)$$

and

$$P_{12}(\sigma, \delta) = C\left[\frac{\sqrt{6}}{12} e^{-(\sigma-\delta)^2/6} + \frac{(\delta-\sigma)}{4}\left\{1 + \Phi\left(\frac{\delta-\sigma}{6}\right)\right\}\right] e^{-(\delta-\sigma)^2/2}, \quad (18)$$

where C is a constant, and where the energy scale has been changed so that $\sigma = \gamma^{\frac{1}{2}}S$, $\delta = \gamma^{\frac{1}{2}}\Delta_1$ (σ and δ are dimensionless).

In Fig. 1 we have plotted $2p(\sigma, \delta)/C$ for various values of δ. It is seen that for $\delta = 2$ there is only one peak; however, as δ increases, a second peak emerges. The first peak results from p_{I} and the second peak from p_{II}. It should be noted that for a given value of δ the second peak occurs at a value of σ less than δ because the repulsion between the degenerate levels tends, in the lowest order, to decrease the spacing.

A second, and more physically realistic, model for H_0 is a $2N$-dimensional Hamiltonian with N equally spaced (spacing Δ_1) doubly degenerate levels. In the limit $\gamma\Delta_1^2 \gg 1$ our large-γ approximation can be used. For this model $\alpha_2 = N$, $\alpha_{221} = N - 1$, and all other α's vanish. Hence,

$$p(S, \gamma) \sim \frac{N}{2N-1} P_2(S, \gamma) + \frac{N-1}{2N-1} P_{22}(S, \gamma, \Delta_1). \quad (19)$$

It is easily shown that

$$P_{22}(\sigma, \delta) = C\Big[\tau\Big(\frac{\sigma-\delta}{2}\Big) + \frac{1}{3\sqrt{\pi}}\Big\{e^{-(\sigma-\delta)^2} - (\sqrt{\pi}/\sqrt{3})(\sigma-\delta) \times \Big[1 - \Phi\Big(\frac{\sigma-\delta}{\sqrt{3}}\Big)\Big]\Big\}e^{-2(\sigma-\delta)^2/3}, \quad (20)$$

where

$$\tau(x) = \int_0^\infty e^{-2(t+x)^2} e^{-2x^2}(2t^2-1)\Phi(t)\,dt. \quad (21)$$

In Fig. 2 we have made plots of $2p(\sigma, \delta)/C$ (in the limit $N \to \infty$) for various values of δ. The integral $\tau[(\sigma-\delta)/2]$ was calculated numerically on the IBM 7074 at Wayne State University. The curves are very similar to those obtained for the three-dimensional case. The main difference is that here the second peak emerges at a larger value of δ and the distance between peaks is smaller here for a given δ than it was there. This is a result of the fact that the levels of H_0 on both sides of a spacing δ are now doubly degenerate. Thus the repulsion between each degenerate pair contributed to a decrease in the value of σ at which the second peak occurs.

The question of where (i.e., for what values of σ) the above approximations are valid is difficult to answer. Ideally we would like to calculate higher terms in the expansion to see where and how fast the asymptotic expansion appears to be converging. In general, it is very difficult to obtain such terms in the expansion of the integral over the rotation parameters by using the ϵ_{ij} as parameters.[17]

There are, however, two reasons why one might expect the approximation to be good in the region of major probability. First, for small values of N the validity of the approximation can easily be tested. For example, for $N = 2$ the calculation can be carried out exactly,[8] and the results can be compared with the approximate result if $\gamma(\lambda_1 - \lambda_2)^2 \gg 1$. For $N = 3$ we were not able to carry out the calculation exactly. However, the higher terms in the asymptotic expansion are easily obtained. In both of these cases it appears that the error is negligible in the region of major probability if $\gamma\Delta_0^2 > 4$. It is not obvious that such results will be valid for large N. However, the perturbation calculations of the previous section yield

[17] For a particular H_0 it may be possible to introduce a more convenient set of parameters. For example, for the $2N$-dimensional H_0 discussed above, a convenient set may be obtained by performing rotations in each of the degenerate planes and then defining additional rotations in such a way as to obtain a completely general rotation. For a discussion of these parameters, see J. F. McDonald, "A Statistical Theory of Energy Levels of Complex Quantum Systems: A Biased Distribution," Ph.D. thesis, Wayne State University (1967).

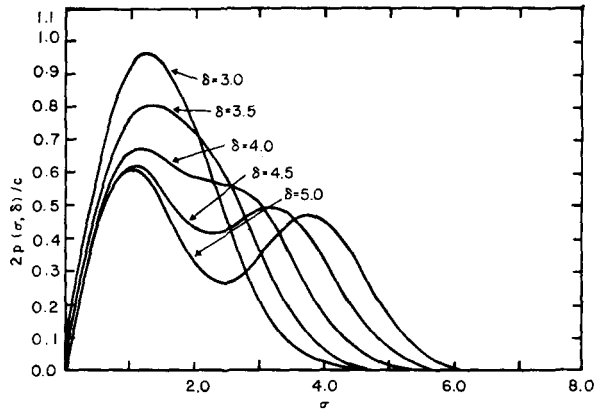


FIG. 2. Plots of $2p(\sigma, \delta)/C$ for a $2N$-dimensional H_0 with equally spaced doubly degenerate levels at various values of the original spacing δ, in the limit $N \to \infty$.

expansions for $\bar{S}$ and $\bar{S}^2$ in powers of $1/\gamma\Delta_0^2$ which converge rapidly for $(1/\gamma\Delta_0^2) \gg 1$ for arbitrary N. Hopefully, this would indicate that our expansion for the spacing distribution should also converge rapidly under these conditions in the region of major probability.

5. SMALL-γ LIMIT

As previously indicated, we expect the small-γ limit to approach the completely random case. That is, we expect H_0 to be, in some sense, a nonrandom perturbation on the random matrix H. Since the presence of the A_{ij} is the only thing which causes a deviation from the random case, it seems appropriate to expand that part of the distribution function which involves the A_{ij}. Thus, we take

$$\eta \exp\,[-\gamma\,\mathrm{Tr}\,(H - H_0)^2] = \eta \exp\Big(-\gamma \sum_i (E_i^2 + \lambda_i^2)\Big)\Big[1 + 2\gamma \sum_{i,j} A_{ij}^2 \lambda_i E_j + 2\gamma^2 \sum_{\substack{i,j \\ k,l}} A_{ij}^2 A_{kl}^2 \lambda_i \lambda_k E_j E_l + \cdots\Big]. \quad (22)$$

Using the methods developed by Ullah,[18] one can explicitly calculate the averages over the rotation parameters and obtain the joint eigenvalue distribution to the first few powers of γ. If we retain terms only to order γ^2, the joint eigenvalue distribution can be written

$$p(E, \gamma) \simeq \eta \exp\,[-\gamma(\mathrm{Tr}\,H^2 + \mathrm{Tr}\,H_0^2)] \times \Big\{1 + \frac{4\gamma^2\,\mathrm{Tr}\,H_0^2}{N(N-1)(N-2)} \times [N\,\mathrm{Tr}\,H^2 + (\mathrm{Tr}\,H)^2]\Big\}, \quad (23)$$

[18] N. Ullah, Nucl. Phys. **58**, 65 (1964).

where for convenience we have chosen the origin of the energy scale such that $\mathrm{Tr}\, H_0 = 0$.

The nth-order (i.e., with n levels between the levels under consideration) spacing distribution $P^{(n)}(S, \gamma)$, which one obtains from (23), can be expressed in terms of the corresponding distribution $p_0^{(n)}(S, \gamma)$, which one obtains from the ordinary Gaussian ensemble as

$$P^{(n)}(S, \gamma) \simeq \exp(-\gamma\, \mathrm{Tr}\, H_0^2)\left[p_0^{(n)}(S, \gamma) + \frac{4\gamma^2\, \mathrm{Tr}\, H_0^2}{(N-1)(N+2)} \times \left\{ \frac{1}{2\gamma}\, p_0^{(n)}(S, \gamma) - \eta \frac{\partial}{\partial \gamma} (p_0^{(n)}(S, \gamma)/\eta) \right\} \right]. \quad (24)$$

Unfortunately, explicit analytic expressions for the $p_0^{(n)}$ do not exist.

However, as was previously noted, the Wigner surmise is a good approximation for $p_0^{(0)}$ if S is not too large. Thus, to obtain some idea of the conditions on γ for this approximation to be valid (this constitutes the definition of the small-γ limit), we take

$$p_0^{(0)}(S, \gamma) \simeq 2N\gamma S \exp(-\gamma N S^2). \quad (25)$$

It then follows that

$$p^{(0)}(S, \gamma) \simeq [1 + 4\gamma^2 S^2\, \mathrm{Tr}\, H_0^2)/N] p_0^{(0)}(S, \gamma). \quad (26)$$

Note that a factor $e^{\lambda\, \mathrm{Tr} H_0^2} = 1 + \gamma\, \mathrm{Tr}\, H_0^2 + \cdots$ has been absorbed into the normalization. Thus, it appears that the condition which must be satisfied for our approximation to be valid is

$$[4\gamma^2 S^2\, \mathrm{Tr}\, H_0^2] \ll 1. \quad (27)$$

Further, since[19] $\bar{S}^2 \simeq \pi/4\gamma N$ (note that, if $\bar{S}$ is to be independent of N, $\gamma \propto 1/N$), this can be rewritten as

$$(\pi\gamma\, \mathrm{Tr}\, H_0^2)/N^2 \ll 1. \quad (28)$$

As an approximation we can take $\mathrm{Tr}\, H_0^2 \simeq N^2\Delta^2/3$, where Δ is some mean spacing of H_0. Then the condition becomes

$$N\gamma \ll 1/\Delta^2, \quad (29)$$

which indicates that the small-γ limit should be valid if the mean spacing $\bar{S}$ is large compared to Δ.

[19] See, for example, the introductory review of Ref. 1.

6. SUMMARY AND CONCLUSIONS

In this paper we have investigated the effects of biasing a Gaussian ensemble of matrices with a nonrandom matrix H_0. In particular, we considered the extremes when the influence of H_0 is strong and weak. The results for the limit when the influence is weak were found to approach the corresponding results for an unbiased Gaussian ensemble as was anticipated.

However, the results for the limit when the influence is strong were quite different. They were found to be strongly dependent on H_0 as expected. Thus, it appears that such a biased ensemble might very well be used to explain the deviations of experimental data from the Wigner surmise and the apparent structure exhibited by some nearest-neighbor spacing-distribution data.

At worst, the biased ensemble can be used to eliminate certain effects as the cause of structure in the data. For example, if one calculates H_0 for a particular nucleus using the shell model, and then calculates the nearest-neighbor spacing distribution only to find that the theoretical and experimental distributions disagree, one can eliminate shell-model structure as the cause of structure in the distribution.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 7 JULY 1968

Mean Entropy of States in Quantum-Statistical Mechanics

OSCAR E. LANFORD III
University of California, Berkeley, California
AND
DEREK W. ROBINSON
CERN, Geneva

(Received 6 September 1967)

The equilibrium states for an infinite system of quantum mechanics may be represented by states over suitably chosen C^* algebras. We consider the problem of associating an entropy with these states and finding its properties, such as positivity, subadditivity, etc. For the states of a quantum-spin system, a mean entropy is defined and it is shown that this entropy is affine and upper semicontinuous.

1. INTRODUCTION

In the algebraic theory of statistical mechanics the class of possible equilibrium states is defined as the subclass K of states ρ, over the C^* algebra $\mathcal{A}$ of local observables, which satisfy certain subsidiary conditions of a physical origin. Firstly, it is assumed that the theory is invariant under a symmetry group G (the translation group R^ν, for example), and the states $\rho \in K$ considered are taken to be G invariant. Secondly, as one wishes to describe only systems with a finite number of particles in each finite subsystem, extra conditions must be introduced. The consequence of these latter "finite mean density" conditions can be described as follows: If $\Lambda \subset R^\nu$ is an open set of compact closure and $\mathcal{A}(\Lambda) \subset \mathcal{A}$ is the corresponding subalgebra of strictly local observables, then a state $\rho \in K$ must be such that its restriction to each $\mathcal{A}(\Lambda)$ is described by a density matrix ρ_Λ acting on a Hilbert space $\mathcal{H}_\Lambda$. As a direct result of this last property we may introduce, for each $\rho \in K$, a family of entropies $S(\rho_\Lambda)$ by the definition

$$S(\rho_\Lambda) = -\mathrm{Tr}_{\mathcal{H}_\Lambda}(\rho_\Lambda \log \rho_\Lambda).$$

Consequently, we may study properties of $S(\rho_\Lambda)$, attempt to introduce for each $\rho \in K$ an entropy per unit volume $S(\rho)$, and, subsequently, analyse the linearity and continuity properties, etc., of $S(\rho)$.

The program outlined above was recently completed by Ruelle, in collaboration with one of the present authors (D. W. R.), in the framework of classical statistical mechanics.[1] The purpose of the present paper is to attempt the same program for quantum-statistical mechanics. In this latter setting many difficulties arise due to noncommutativity, and our results are complete only for quantum-spin systems. In the general case many interesting problems remain open.

[1] D. W. Robinson and D. Ruelle, Commun. Math. Phys. **5**, 288 (1967).

2. GENERAL FORMULATION

We want to investigate both continuous infinite quantum-statistical systems and lattice systems. Thus, we consider a C^* algebra $\mathcal{A}$ and a collection $\{\mathcal{A}(\Lambda)\}$ of C^* subalgebras of $\mathcal{A}$, where Λ runs over:

(i) the bounded open sets in R^ν for continuous systems;

(ii) the finite subsets of Z^ν for lattice systems.

We suppose that these subalgebras satisfy the following axioms:

(1) $\mathcal{A}(\Lambda_1) \subset \mathcal{A}(\Lambda_2)$ if $\Lambda_1 \subset \Lambda_2$.

(2) For each Λ, $\mathcal{A}(\Lambda)$ is isomorphic to $\mathfrak{L}(\mathcal{H}_\Lambda)$ for some Hilbert space $\mathcal{H}_\Lambda$. We will usually identify $\mathcal{A}(\Lambda)$ with $\mathfrak{L}(\mathcal{H}_\Lambda)$, although this is not strictly compatible with axiom (1).

(3) $\mathcal{A}$ is the norm closure of $\cup_\Lambda \mathcal{A}(\Lambda)$.

(4) $\mathcal{A}(\Lambda_1 \cup \Lambda_2)$ is generated by $\mathcal{A}(\Lambda_1) \cup \mathcal{A}(\Lambda_2)$ in the weak operator topology on $\mathfrak{L}(\mathcal{H}_{\Lambda_1 \cup \Lambda_2})$.

(5) Let G denote the group of translations, i.e., $G = Z^\nu$ for lattice systems and $G = R^\nu$ for continuous systems. Then G acts on $\mathcal{A}$ by automorphisms τ_x in such a way that $\tau_x(\mathcal{A}(\Lambda)) = \mathcal{A}(\Lambda + x)$ for all regions Λ and translations x.

Finally, we need a condition expressing the independence of observables belonging to disjoint regions. This condition may take one of two forms, depending on whether we are considering bosons or fermions:

(6) Either

(a) If Λ_1 and Λ_2 are any two disjoint regions, then $\mathcal{A}(\Lambda_1)$ commutes with $\mathcal{A}(\Lambda_2)$; or

(b) Each $\mathcal{A}(\Lambda)$ is generated by a set of creation and annihilation operators satisfying the canonical anticommutation relations, and, if Λ_1 and Λ_2 are disjoint regions, the creation and annihilation operators for Λ_1 anticommute with those for Λ_2.

These axioms describe several kinds of physical systems:

(1) Ordinary continuous quantum systems, either bosons or fermions.

(2) Quantum lattice systems, again either bosons or fermions, with finitely many creation and annihilation operators associated with each lattice site. For fermion lattice systems, $\mathcal{H}_\Lambda$ is finite-dimensional for each finite set Λ, but for boson systems this is, of course, not true.

(3) Quantum-spin systems. In this case, $\mathcal{H}_\Lambda$ is finite-dimensional for each bounded region Λ, but the different unit rays in $\mathcal{H}_{\{x\}}$, where x is a lattice point, are interpreted as describing different polarization states of a particle localized at x rather than varying occupation numbers for that point. We will assume that such systems satisfy axiom (6a).

The statistical-mechanical states of $\mathcal{A}$ are those which, when restricted to an $\mathcal{A}(\Lambda)$, are given by a density matrix. In other words, such a state ρ defines, for each region Λ, a positive operator ρ_Λ on $\mathcal{H}_\Lambda$, with $\mathrm{Tr}_{\mathcal{H}_\Lambda}(\rho_\Lambda) = 1$, such that

$$\rho(A) = \mathrm{Tr}_{\mathcal{H}_\Lambda}(\rho_\Lambda A),$$

if $A \in \mathcal{A}(\Lambda) = \mathfrak{L}(\mathcal{H}_\Lambda)$. This statement imposes no restriction on ρ if $\mathcal{H}_\Lambda$ is finite-dimensional; otherwise, it corresponds to the requirement that there be only finitely many particles in the region Λ.

Every statistical-mechanical state ρ defines a family $\{\rho_\Lambda\}$ of density matrices. Conversely, the assignment of a density matrix to each bounded region defines a statistical-mechanical state on $\mathcal{A}$, provided that the assignment satisfies the obvious compatibility condition that, if $\Lambda_1 \subset \Lambda_2$ and if $A \in \mathcal{A}(\Lambda_1)$, then

$$\mathrm{Tr}_{\mathcal{H}_{\Lambda_1}}(\rho_{\Lambda_1} A) = \mathrm{Tr}_{\mathcal{H}_{\Lambda_2}}(\rho_{\Lambda_2} A).$$

We can reformulate the compatibility condition as follows: If $\Lambda_1 \subset \Lambda_2$, then $\mathcal{A}(\Lambda_1)$ is a type I factor contained in $\mathcal{A}(\Lambda_2) = \mathfrak{L}(\mathcal{H}_{\Lambda_2})$. Hence, we may factorize

$$\mathcal{H}_{\Lambda_2} = \mathcal{H}_{\Lambda_1} \otimes \mathcal{H}'$$

in such a way that an operator A in $\mathcal{A}(\Lambda_1) = \mathfrak{L}(\mathcal{H}_{\Lambda_1})$ is identified with the operator $A \otimes 1$ on $\mathcal{H}_{\Lambda_1} \otimes \mathcal{H}'$. [The space $\mathcal{H}'$ may be identified with $\mathcal{H}_{\Lambda_2 - \Lambda_1}$, but operators in $\mathcal{A}(\Lambda_2 - \Lambda_1)$ do not factorize as nicely as those in $\mathcal{A}(\Lambda_1)$ unless algebras for disjoint regions commute. See below.] The compatibility condition may now be formulated as

$$\rho_{\Lambda_1} = \mathrm{Tr}_{\mathcal{H}'}(\rho_{\Lambda_2}),$$

where $\mathrm{Tr}_{\mathcal{H}'}$ means the partial trace with respect to $\mathcal{H}'$, i.e., if $\{\varphi_i\}$ is an orthonormal basis for $\mathcal{H}_{\Lambda_1}$ and $\{\psi_i\}$ is an orthonormal basis for $\mathcal{H}'$, then

$$(\rho_{\Lambda_1}\varphi_i, \varphi_k) = \sum_{j=1}^{\infty} [\rho_{\Lambda_2}(\varphi_i \otimes \psi_j), \varphi_k \otimes \psi_j].$$

The condition that a state be translation-invariant may easily be formulated in terms of the corresponding system of density matrices. For any region Λ and any translation x, τ_x is an isomorphism of $\mathcal{A}(\Lambda)$ onto $\mathcal{A}(\Lambda + x)$. Since $\mathcal{A}(\Lambda)$ is identified with $\mathfrak{L}(\mathcal{H}_\Lambda)$ and $\mathcal{A}(\Lambda + x)$ with $\mathfrak{L}(\mathcal{H}_{\Lambda+x})$, there is a unitary operator $U_{\Lambda,x}$ from $\mathcal{H}_\Lambda$ to $\mathcal{H}_{\Lambda+x}$ which induces this isomorphism, and $U_{\Lambda,x}$ is determined up to a multiplicative constant. Then the state defined by the system $\{\rho_\Lambda\}$ of density matrices is translation-invariant if and only if

$$\rho_{\Lambda+x} = U_{\Lambda,x}\rho_\Lambda U_{\Lambda,x}^{-1},$$

for all regions Λ and translations x.

We now want to make a more careful analysis of the relation of $\rho_{\Lambda_1 \cup \Lambda_2}$ to ρ_{Λ_1} and ρ_{Λ_2} when Λ_1 and Λ_2 are disjoint regions. We have already remarked that the inclusion of $\mathcal{A}(\Lambda_1)$ in $\mathcal{A}(\Lambda_1 \cup \Lambda_2)$ gives a factorization of $\mathcal{H}_{\Lambda_1 \cup \Lambda_2}$ as $\mathcal{H}_{\Lambda_1} \otimes \mathcal{H}'$, where operators in $\mathcal{A}(\Lambda_1)$ go over into operators of the form $A \otimes 1$. If we are considering a boson system or a spin system, then the commutant of $\mathcal{A}(\Lambda_1)$ in $\mathcal{A}(\Lambda_1 \cup \Lambda_2)$ is precisely $\mathcal{A}(\Lambda_2)$. In this case, there is an essentially unique way to identify $\mathcal{H}'$ with $\mathcal{H}_{\Lambda_2}$, and operators in $\mathcal{A}(\Lambda_2)$ take the form $1 \otimes A$ on $\mathcal{H}_{\Lambda_1} \otimes \mathcal{H}_{\Lambda_2}$. Hence we have

$$\rho_{\Lambda_1} = \mathrm{Tr}_{\mathcal{H}_{\Lambda_2}}(\rho_{\Lambda_1 \cup \Lambda_2}), \quad \rho_{\Lambda_2} = \mathrm{Tr}_{\mathcal{H}_{\Lambda_1}}(\rho_{\Lambda_1 \cup \Lambda_2}).$$

For fermion systems, although $\mathcal{H}'$ has the same dimension as $\mathcal{H}_{\Lambda_2}$, there is no unique sensible way to identify $\mathcal{H}'$ with $\mathcal{H}_{\Lambda_2}$. Nevertheless, by using the special structure of fermion systems, we can construct a useful identification of $\mathcal{H}'$ with $\mathcal{H}_{\Lambda_2}$. This we do as follows: Let N_1 and N_2 denote the number operators for the regions Λ_1 and Λ_2, respectively. Then a simple calculation with the anticommutation relations shows that the commutant of $\mathcal{A}(\Lambda_1)$ in $\mathcal{A}(\Lambda_1 \cup \Lambda_2)$ is precisely $(-1)^{N_1N_2}\mathcal{A}(\Lambda_2)(-1)^{N_1N_2}$. Therefore, we can identify $\mathcal{H}'$ with $\mathcal{H}_{\Lambda_2}$ in such a way that if A is in $\mathcal{A}(\Lambda_2) = \mathfrak{L}(\mathcal{H}_{\Lambda_2})$, then A goes over into

$$(-1)^{N_1N_2}(1 \otimes A)(-1)^{N_1N_2},$$

in $\mathfrak{L}(\mathcal{H}_{\Lambda_1} \otimes \mathcal{H}_{\Lambda_2})$. With this identification we have

$$\rho_{\Lambda_1} = \mathrm{Tr}_{\mathcal{H}_{\Lambda_2}}(\rho_{\Lambda_1 \cup \Lambda_2}),$$
$$\rho_{\Lambda_2} = \mathrm{Tr}_{\mathcal{H}_{\Lambda 1}}[(-1)^{N_1N_2}\rho_{\Lambda_1 \cup \Lambda_2}(-1)^{N_1N_2}].$$

The second of these equations can be simplified if we assume that ρ is an even state of $\mathcal{A}$. By definition, an element A of some $\mathcal{A}(\Lambda)$ is *odd* if

$$(-1)^{N(\Lambda)}A(-1)^{N(\Lambda)} = -A,$$

where $N(\Lambda)$ is the number operator for the region Λ. A state ρ of $\mathcal{A}$ is *even* if ρ vanishes on every odd element of every $\mathcal{A}(\Lambda)$. If ρ is now an even state and A is an element of $\mathcal{A}(\Lambda_2)$, then

$$\rho(A) = \mathrm{Tr}_{\mathcal{H}_{\Lambda_1} \otimes \mathcal{H}_{\Lambda_2}}[\rho_{\Lambda_1 \cup \Lambda_2}(-1)^{N_1N_2}(1 \otimes A)(-1)^{N_1N_2}]$$
$$= \mathrm{Tr}_{\mathcal{H}_{\Lambda_1} \otimes \mathcal{H}_{\Lambda_2}}[\rho_{\Lambda_1 \cup \Lambda_2}(1 \otimes A)].$$

To prove this equation, note that we can write A as the sum of an even part and an odd part, that the odd part contributes nothing to $\rho(A)$, and that the even part commutes with $(-1)^{N_1 N_2}$. Collecting these results we have:

Proposition 1: Let ρ be a statistical-mechanical state of the C^* algebra $\mathcal{A}$, and let $\{\rho_\Lambda\}$ be the corresponding system of density matrices. If $\mathcal{A}$ is the algebra for a fermion system, we further assume that ρ is even. Then, if Λ_1 and Λ_2 are disjoint regions, we may identify $\mathcal{H}_{\Lambda_1 \cup \Lambda_2}$ with $\mathcal{H}_{\Lambda_1} \otimes \mathcal{H}_{\Lambda_2}$ in such a way that

$$\rho_{\Lambda_1} = \mathrm{Tr}_{\mathcal{H}_{\Lambda_2}}(\rho_{\Lambda_1 \cup \Lambda_2}) \quad \rho_{\Lambda_2} = \mathrm{Tr}_{\mathcal{H}_{\Lambda_1}}(\rho_{\Lambda_1 \cup \Lambda_2}).$$

Note that if we are dealing with a fermion system, then translation invariance implies that the state ρ is even. To show this[2] let A be an odd element of some $\mathcal{A}(\Lambda)$ and let x be a translation large enough so that $\Lambda + nx$ does not intersect Λ for $n = 1, 2, 3, \cdots$ Let

$$A_N = \frac{1}{N} \sum_{n=0}^{N-1} \tau_{nx} A.$$

Now

$$\begin{aligned} \|A_N\|^2 = \|A_N^* A_N\| &\leq \|\{A_N^*, A_N\}\| \\ &\leq \frac{1}{N^2} \sum_{n,m=0}^{N-1} \|\{(\tau_{nx}A)^*, \tau_{mx}A\}\| \\ &\leq \frac{2}{N} \|A\|^2, \end{aligned}$$

where the last inequality is a consequence of

$$\{(\tau_{nx}A)^*, \tau_{mx}A\} = 0, \quad \text{for} \quad n \neq m.$$

However, by translation invariance,

$$\rho(A) = \rho(A_N) = \lim_{N \to \infty} \rho(A_N).$$

But as $\lim_{N \to \infty} \|A_N\| = 0$, then $\lim_{N \to \infty} \rho(A_N) = 0$ and thus $\rho(A) = 0$.

Given a statistical-mechanical state ρ and a region Λ, we can define the entropy of the region Λ as follows:

$S(\rho_\Lambda) = +\infty$, if $\rho_\Lambda \log \rho_\Lambda$ is not of trace class on $\mathcal{H}_\Lambda$

$= -\mathrm{Tr}_{\mathcal{H}_\Lambda}(\rho_\Lambda \log \rho_\Lambda)$, otherwise.

In defining the operator $\rho_\Lambda \log \rho_\Lambda$ we use the usual convention $x \log x = 0$, for $x = 0$.

3. BASIC INEQUALITIES FOR THE ENTROPY

Lemma 1[3]*:* If A and B are positive, self-adjoint, trace-class operators on a Hilbert space $\mathcal{H}$, then

$$\mathrm{Tr}_{\mathcal{H}}[A \log A - A \log B - A + B] \geq 0.$$

Proof: Let $\psi_i(\varphi_i)$ be a complete orthonormal set of eigenfunctions of $A(B)$ and let $a_i(b_i)$ be the corresponding eigenvalues. Let $U = (U_{ij})$ be a unitary mapping defined by

$$\psi_i = \sum_j U_{ij} \varphi_j.$$

Now

$$\begin{aligned} (\psi_i | A \log A - A \log B | \psi_i) &= a_i \left\{ \log a_i - \sum_j |U_{ij}|^2 \log b_j \right\} \\ &\geq a_i \left\{ \log a_i - \log \sum_j |U_{ij}|^2 b_j \right\} \\ &\geq a_i - \sum_j |U_{ij}|^2 b_j \\ &= (\psi_i | A - B | \psi_i), \end{aligned}$$

where, to obtain the first inequality, we have used the concavity of the logarithm and, to obtain the second inequality, we have used

$$\log x \geq 1 - 1/x \quad (x > 0).$$

The result follows by summation.

Lemma 2: If A and B be positive, self-adjoint operators on a Hilbert space $\mathcal{H}$, then, for $1 \geq \alpha \geq 0$,

$$[\alpha A + (1-\alpha)B] \log [\alpha A + (1-\alpha)B] \leq \alpha A \log A + (1-\alpha) B \log B,$$

and furthermore

$$A \geq B \geq 0 \quad \text{implies} \quad \log A \geq \log B.$$

The statements of the lemma are special consequences of the theory of convex and monotone operator functions initially developed by Löwner.[4] For further results, the reader may consult Ref. 5. The details of the application of the general theory to the case at hand are worked out in Refs. 6 and 7. Moreover, we do not need the full force of the first inequality of the lemma, but only the inequality obtained by taking its trace; this latter inequality can be proved without use of the general theory of convex operator functions.[7,8]

We remark that Lemma 1 may be deduced from the first statement of Lemma 2. We preferred, however, to give the simple straightforward proof reproduced above.

---

[2] This proof was independently discovered by R. T. Powers.

[3] This lemma, together with its proof, was communicated to one of us (D. W. R.) by Professor R. Jost, who attributed it to O. Klein. If $\mathrm{Tr}(A) = \mathrm{Tr}(B) = 1$, this lemma is a particular case of Theorem 1 of H. Umegaki, Kodai Math. Sem. Rep. **14**, 59 (1962).

[4] C. Löwner, Math. Z. **38**, 177 (1934).

[5] J. Bendat and S. Sherman, Trans. Am. Math. Soc. **79**, 58 (1955).

[6] M. Nakamura and H. Umegaki, Proc. Japan Acad. **37**, 149 (1961).

[7] C. Davis, Proc. Japan Acad. **37**, 533 (1961).

[8] I. E. Segal, J. Math. & Mech. **9**, 623 (1960).

The preceding lemmas may now be used to deduce the following results for the quantum entropy, specializations of which appear in Refs. 9 and 10.

Theorem 1: Let ρ be a statistical-mechanical state of the C^* algebra $\mathcal{A}$, and let $\{\rho_\Lambda\}$ be the corresponding system of density matrices. If $\mathcal{A}$ is the algebra for a fermion system, we assume further that ρ is an even state. Then the associated entropy $S(\rho_\Lambda)$ is a positive set function, i.e.,

$$S(\rho_\Lambda) \geq 0,$$

which is subadditive, i.e.,

$$S(\rho_{\Lambda_1 \cup \Lambda_2}) \leq S(\rho_{\Lambda_1}) + S(\rho_{\Lambda_2}),$$

if

$$\Lambda_1 \cap \Lambda_2 = \phi.$$

Further, if $\{\rho_\Lambda^{(1)}\}$ and $\{\rho_\Lambda^{(2)}\}$ are two families of density matrices and $1 \geq \alpha \geq 0$, then

$$\begin{aligned}\alpha S(\rho_\Lambda^{(1)}) + (1-\alpha) S(\rho_\Lambda^{(2)}) &\leq S[\alpha\rho_\Lambda^{(1)} + (1-\alpha)\rho_\Lambda^{(2)}] \\ &\leq \alpha S(\rho_\Lambda^{(1)}) + (1-\alpha) S(\rho_\Lambda^{(2)}) \\ &\quad - \alpha \log \alpha - (1-\alpha)\log(1-\alpha).\end{aligned}$$

Proof: The positivity of $S(\rho_\Lambda)$ is an immediate consequence of the normalization of ρ_Λ and the fact that

$$-\lambda \log \lambda \geq 0, \quad 1 \geq \lambda \geq 0.$$

The subadditivity property follows from Lemma 1, Proposition 1, and the identification $\mathcal{H} = \mathcal{H}_{\Lambda_1} \otimes \mathcal{H}_{\Lambda_2}$, $A = \rho_{\Lambda_1 \cup \Lambda_2}$, and $B = \rho_{\Lambda_1} \otimes \rho_{\Lambda_2}$. The final inequality is a consequence of Lemma 2. The lower bound is immediately obtained from the first statement of that lemma, while the upper bound is obtained from the second statement as follows: We have

$$\alpha\rho_\Lambda^{(1)} + (1-\alpha)\rho_\Lambda^{(2)} \geq \alpha\rho_\Lambda^{(1)} \geq 0,$$

and hence

$$\log[\alpha\rho_\Lambda^{(1)} + (1-\alpha)\rho_\Lambda^{(2)}] \geq \log \alpha\rho_\Lambda^{(1)}.$$

Thus

$$\begin{aligned}-\alpha \operatorname{Tr}[\rho_\Lambda^{(1)} \log(\alpha\rho_\Lambda^{(1)} + (1-\alpha)\rho_\Lambda^{(2)})] &\leq -\alpha \operatorname{Tr}[\rho_\Lambda^{(1)} \log \alpha\rho_\Lambda^{(1)}] \\ &= -\alpha \operatorname{Tr}[\rho_\Lambda^{(1)} \log \rho_\Lambda^{(1)}] - \alpha \log \alpha.\end{aligned}$$

Similarly,

$$\begin{aligned}&-(1-\alpha)\operatorname{Tr}[\rho_\Lambda^{(2)} \log(\alpha\rho_\Lambda^{(1)} + (1-\alpha)\rho_\Lambda^{(2)})] \\ &\quad \leq -(1-\alpha)\operatorname{Tr}[\rho_\Lambda^{(2)} \log(1-\alpha)\rho_\Lambda^{(2)}] \\ &\quad = -(1-\alpha)\operatorname{Tr}[\rho_\Lambda^{(2)} \log \rho_\Lambda^{(2)}] - (1-\alpha)\log(1-\alpha).\end{aligned}$$

Adding the last two inequalities yields the desired result.

4. MEAN ENTROPY—THE QUANTUM LATTICE SYSTEM

The next desirable aim would be to define an entropy per unit volume, or mean entropy, by establishing, under suitable restrictions, the existence of $S(\rho_\Lambda)/V(\Lambda)$ in the limit of Λ increasing such that the volume $V(\Lambda) \to \infty$. Unfortunately, we are at present able to do this only for a quantum lattice system, and even then it is not possible to establish the existence of the limit in the most desirable generality. A possible means of improving our results is discussed in the concluding section.

Let us define for $a = (a_1, \cdots, a_\nu) \in Z^\nu$ and $a_1 > 0, \cdots, a_\nu > 0$ the parallelepiped $\Lambda(a)$ by

$$\Lambda(a) = \{x \in Z^\nu;\ 0 < x_i \leq a_i \quad \text{for} \quad i = 1, 2, \cdots, \nu\}.$$

The measure (volume) $V[\Lambda(a)]$ of $\Lambda(a)$ is then given by $\prod_{i=1}^{\nu} a_i$.

Theorem 2: If the family $\rho = \{\rho_\Lambda\}$ of density matrices of a quantum lattice system satisfies the conditions of normalization, compatibility, and translation invariance, then the mean entropy

$$S(\rho) = \lim_{a_1, \ldots, a_\nu \to \infty} \frac{S(\rho_{\Lambda(a)})}{V(\Lambda(a))}$$

exists, and in fact

$$S(\rho) = \inf_{a_1, \ldots, a_\nu} \frac{S(\rho_{\Lambda(a)})}{V(\Lambda(a))}.$$

Further, $S(\rho)$ is an affine function. Explicitly, if $\rho^{(1)} = \{\rho_\Lambda^{(1)}\}$ and $\rho^{(2)} = \{\rho_\Lambda^{(2)}\}$ are two appropriate families of density matrices and $1 \geq \alpha \geq 0$, then

$$S[\alpha\rho^{(1)} + (1-\alpha)\rho^{(2)}] = \alpha S(\rho^{(1)}) + (1-\alpha) S(\rho^{(2)}).$$

Proof: By translation invariance, $S[\Lambda(a)]$ is a function of $a_1, \cdots, a_\nu$ only. Moreover, if we are dealing with a fermion system, translation invariance implies that the state ρ is even (see Sec. 2). But if we introduce a function $S(a_1, \cdots, a_\nu)$ through the definition

$$S(a_1, \cdots, a_\nu) = S[\Lambda(a)],$$

the subadditivity of $S(\Lambda)$ implies that $S(a_1, \cdots, a_\nu)$ is subadditive in each argument a_i $(1 \leq i \leq \nu)$ separately, i.e.,

$$\begin{aligned}S(a_1, \cdots, a_i^{(1)} + a_{i\nu}^{(2)}, \cdots, a_\nu) &= S(a_1, \cdots, a_i^{(1)}, \cdots a_\nu) \\ &\quad + S(a_i, \cdots, a_i^{(2)}, \cdots, a_\nu).\end{aligned}$$

[9] M. Delbrück and G. Molière, Abhandl. Preuss. Akad. P. 1 (1937).

[10] R. Jost, Helv. Phys. Acta **20**, 491 (1947).

A standard argument [cf. Ref. 11, Lemma A1] establishes the existence of

$$S(\rho) = \lim_{a_1,\ldots,a_\nu \to \infty} \frac{S(a_1,\cdots,a_\nu)}{a_1 a_2,\cdots,a_\nu} = \inf_{a_1,\ldots,a\nu} \frac{S(a_1,\cdots,a_\nu)}{a_1 a_2,\cdots,a_\nu}.$$

The affine property of $S(\rho)$ follows from the last statement of Theorem 1, if one takes $\Lambda = \Lambda(a)$, divides by $V[\Lambda(a)]$, and takes the appropriate limit.

5. PROPERTIES OF THE MEAN ENTROPY

For fermion lattice systems and spin systems we can exploit the finite dimensionality of the $\mathfrak{K}_\Lambda$'s to prove some additional properties of the mean entropy.

Theorem 3: Let $\mathcal{A}$ be the C^* algebra for a fermion lattice system or a spin lattice system. If x is a lattice point, let N denote the dimension of $\mathfrak{K}_{\{x\}}$. Equip the set of states of $\mathcal{A}$ with the weak * topology. Then

(1) For any invariant state ρ of $\mathcal{A}$, $0 \leq S(\rho) \leq \text{Log } N$.

(2) The mean entropy is an upper semicontinuous function on the set of invariant states of $\mathcal{A}$. If F is any closed subset of the set of invariant states of $\mathcal{A}$, then the restriction of the mean entropy to F attains its maximum.

(3) If ρ is an invariant state of $\mathcal{A}$, and if μ_ρ is the unique probability measure with barycenter ρ concentrated on the extremal invariant states of $\mathcal{A}$, then

$$S(\rho) = \int d\mu_\rho(\rho') S(\rho').$$

In physical language, statement 3 says that if ρ is an average of pure phases, then the mean entropy of ρ is the same average of the entropies of the pure phases making up ρ. For the existence and uniqueness of the measure μ_ρ, see Ref. 12, Theorem 2, or Ref. 13, Theorem 3. We remark that, under the hypotheses of this theorem, $\mathcal{A}$ is separable so there are no technical difficulties about the sense in which the measure is concentrated on the extremal invariant states.

Proof: For any finite set Λ of lattice points, the dimension of $\mathfrak{K}_\Lambda$ is $N^{V(\Lambda)}$. Now

$$S(\rho_\Lambda) = -\sum_{i=1}^{N^{V(\Lambda)}} \lambda_i \log \lambda_i,$$

where the λ_i are the eigenvalues of ρ_Λ. By elementary estimates, if $\mu_1,\cdots,\mu_n$ are positive real numbers with $\sum_i \mu_i = 1$, then

$$-\sum_{i=1}^{n} \mu_i \log \mu_i \leq \log n.$$

Hence,

$$S(\rho_\Lambda) \leq \log (N^{V(\Lambda)}) = V(\Lambda) \log N.$$

Dividing by $V(\Lambda)$ and taking the limit of infinite volume gives

$$S(\rho) \leq \log N.$$

Since $S(\rho)$ is nonnegative by definition, we have proved part (1) of Theorem 3.

To prove part (2), observe first that the ρ_Λ's are continuous functions of ρ and that the eigenvalues of ρ_Λ vary continuously with ρ_Λ by perturbation theory. Since $-\lambda \log \lambda$ is a continuous function of λ,

$$S(\rho_\Lambda) = -\sum_i \lambda_i \log \lambda_i$$

is a continuous function of ρ. But

$$S(\rho) = \inf_\Lambda \left\{\frac{S(\rho_\Lambda)}{V(\Lambda)}\right\},$$

where the infimum is to be taken over all rectangles. Thus, $S(\rho)$ is the infimum of a family of continuous functions and is therefore upper semicontinuous. In particular, if F is any closed set of invariant states on $\mathcal{A}$, then the restriction of S to F takes on its maximum, since any upper semicontinuous function on a compact set takes on its maximum.

Furthermore, since $S(\rho)$ is both affine and upper semicontinuous, it respects barycentric decompositions. More precisely, if μ is any probability measure on the set of invariant states of $\mathcal{A}$, and if the barycenter of μ is ρ, then

$$S(\rho) = \int d\mu(\rho') S(\rho').$$

(This follows from Lemma 10 of Ref. 14.) In particular, if μ_ρ is the unique decomposition of ρ into extremal invariant states,[15] then the above equation holds. This proves part (3) and completes the proof of the theorem.

6. CONCLUSION

While we have been able to obtain most of the desired results concerning the entropy of a quantum spin system, the position is less satisfactory in other cases. The main gap in the development is the failure to establish the existence of the mean entropy $S(\rho)$ under general circumstances. In classical statistical

[11] D. Ruelle, J. Math. Phys. **6**, 201 (1965).
[12] D. Kastler and D. W. Robinson, Commun. Math. Phys. **3**, 51 (1966).
[13] O. E. Lanford and D. Ruelle, J. Math. Phys. **8**, 1460 (1967).
[14] G. Choquet and P. A. Mayer, Ann. Inst. Fourier **13**, 139 (1963).
[15] Note that the uniqueness proofs given in Refs. 12 and 13 for such decompositions are valid even for Fermi systems. In the Fermi case, $\mathcal{A}$ is $R\nu$ (or $Z\nu$) Abelian, in the sense of Ref. 13, as an argument similar to that appearing after Proposition 1 readily establishes.

mechanics[1] these existence problems were solved by showing that the entropy satisfied a condition of strong subadditivity. One could believe, and even support one's belief by heuristic physical arguments, that the same condition holds for the quantum entropy.

Conjecture: The quantum entropy $S(\rho_\Lambda)$ satisfies the inequality

$$S(\rho_{\Lambda_1\cup\Lambda_2}) + S(\rho_{\Lambda_1\cap\Lambda_2}) \leq S(\rho_{\Lambda_1}) + S(\rho_{\Lambda_2}).$$

A satisfactory discussion of the existence of the mean entropy would ensue, if this conjecture were proved. There would, however, still exist a problem in establishing a barycentric decomposition of the mean entropy in the general case because, although it would clearly be an affine function, it could not be expected to be upper semicontinuous.

We have not discussed in any detail the physical relevance of the mean entropy which we have introduced but postpone this to a forthcoming publication.[16]

ACKNOWLEDGMENTS

The authors are grateful to Professor D. Ruelle for many helpful and stimulating conversations. One of us (D. W. R.) would like to thank Professor R. Jost for a number of informative discussions. Finally, we would like to thank Professor P. Porcelli for his hospitality at the University of Louisiana during the Baton Rouge Symposium on C^* algebras where the present collaboration began.

[16] D. W. Robinson, Commun. Math. Phys. **6**, 151 (1967).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 7 JULY 1968

Parameter Differentiation of Exponential Operators and the Baker–Campbell–Hausdorff Formula

M. LUTZKY
U.S. Naval Ordnance Laboratory, White Oak, Silver Spring, Maryland

(Received 8 November 1967)

Parameter differentiation of exponential operators is used to derive a method for obtaining the Baker–Campbell–Hausdorff coefficients in a more explicit form than is available from the standard Hausdorff recursion formula. In passing, a derivation of the Hausdorff recursion formula is given which is simpler than the proof usually presented.

I. INTRODUCTION

The problem of solving the equation $e^z = e^x e^y$ for z, where x and y are noncommuting operators, frequently arises in physics. For instance, Weiss and Maradudin[1] discussed this problem in a study of x-ray scattering in crystals, and Snider[2] encountered it in the course of an investigation involving a linearization of the Boltzmann equation. The classical solution, given by Hausdorff,[3] involves an expansion of z into an infinite series of terms homogeneous in y, the successive terms being found from a recursion formula which utilizes the Hausdorff polarization operator. The recursion formula is known as the Baker–Campbell–Hausdorff formula (BCH), and has been discussed recently by several authors.[4] Unfortunately, the BCH formula is somewhat difficult to use for the computation of higher-order terms. In this paper, we use the method of parameter differentiation of exponential operators[5] to derive in a rather simple way a more explicit form for the BCH coefficients. In passing, we also show how the method may be used to provide a derivation of the BCH recursion formula that is somewhat simpler than the proof usually given.

II. PRELIMINARY DEFINITIONS AND FORMULAS

Our formulas will be considerably simplified by the use of the curly commutator bracket, recursively defined by

$$\{y, x^0\} = y, \tag{1}$$

$$\{y, x^{k+1}\} = [\{y, x^k\}, x]. \tag{2}$$

If $f(x)$ has a power-series expansion

$$f(x) = \sum_{k=0}^{\infty} a_k x^k,$$

[1] G. H. Weiss and A. A. Maradudin, J. Math. Phys. **3**, 771 (1962).
[2] R. F. Snider, J. Math. Phys. **5**, 1586 (1964).
[3] F. Hausdorff, Ber. Verhandl. Saechs. Akad. Wiss. Leipzig, Math.-Naturw. Kl. **58**, 19 (1906).
[4] W. Magnus, Commun. Pure Appl. Math. **7**, 649 (1954); J. Wei, J. Math. Phys. **4**, 1337 (1963); and Ref. 1.
[5] R. M. Wilcox, J. Math. Phys. **8**, 962 (1967).

mechanics[1] these existence problems were solved by showing that the entropy satisfied a condition of strong subadditivity. One could believe, and even support one's belief by heuristic physical arguments, that the same condition holds for the quantum entropy.

Conjecture: The quantum entropy $S(\rho_\Lambda)$ satisfies the inequality

$$S(\rho_{\Lambda_1\cup\Lambda_2}) + S(\rho_{\Lambda_1\cap\Lambda_2}) \leq S(\rho_{\Lambda_1}) + S(\rho_{\Lambda_2}).$$

A satisfactory discussion of the existence of the mean entropy would ensue, if this conjecture were proved. There would, however, still exist a problem in establishing a barycentric decomposition of the mean entropy in the general case because, although it would clearly be an affine function, it could not be expected to be upper semicontinuous.

We have not discussed in any detail the physical relevance of the mean entropy which we have introduced but postpone this to a forthcoming publication.[16]

ACKNOWLEDGMENTS

The authors are grateful to Professor D. Ruelle for many helpful and stimulating conversations. One of us (D. W. R.) would like to thank Professor R. Jost for a number of informative discussions. Finally, we would like to thank Professor P. Porcelli for his hospitality at the University of Louisiana during the Baton Rouge Symposium on C^* algebras where the present collaboration began.

[16] D. W. Robinson, Commun. Math. Phys. **6**, 151 (1967).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 7 JULY 1968

Parameter Differentiation of Exponential Operators and the Baker–Campbell–Hausdorff Formula

M. LUTZKY
U.S. Naval Ordnance Laboratory, White Oak, Silver Spring, Maryland

(Received 8 November 1967)

Parameter differentiation of exponential operators is used to derive a method for obtaining the Baker–Campbell–Hausdorff coefficients in a more explicit form than is available from the standard Hausdorff recursion formula. In passing, a derivation of the Hausdorff recursion formula is given which is simpler than the proof usually presented.

I. INTRODUCTION

The problem of solving the equation $e^z = e^x e^y$ for z, where x and y are noncommuting operators, frequently arises in physics. For instance, Weiss and Maradudin[1] discussed this problem in a study of x-ray scattering in crystals, and Snider[2] encountered it in the course of an investigation involving a linearization of the Boltzmann equation. The classical solution, given by Hausdorff,[3] involves an expansion of z into an infinite series of terms homogeneous in y, the successive terms being found from a recursion formula which utilizes the Hausdorff polarization operator. The recursion formula is known as the Baker–Campbell–Hausdorff formula (BCH), and has been discussed recently by several authors.[4] Unfortunately, the BCH formula is somewhat difficult to use for the computation of higher-order terms. In this paper, we use the method of parameter differentiation of exponential operators[5] to derive in a rather simple way a more explicit form for the BCH coefficients. In passing, we also show how the method may be used to provide a derivation of the BCH recursion formula that is somewhat simpler than the proof usually given.

II. PRELIMINARY DEFINITIONS AND FORMULAS

Our formulas will be considerably simplified by the use of the curly commutator bracket, recursively defined by

$$\{y, x^0\} = y, \tag{1}$$

$$\{y, x^{k+1}\} = [\{y, x^k\}, x]. \tag{2}$$

If $f(x)$ has a power-series expansion

$$f(x) = \sum_{k=0}^{\infty} a_k x^k,$$

[1] G. H. Weiss and A. A. Maradudin, J. Math. Phys. **3**, 771 (1962).
[2] R. F. Snider, J. Math. Phys. **5**, 1586 (1964).
[3] F. Hausdorff, Ber. Verhandl. Saechs. Akad. Wiss. Leipzig, Math.-Naturw. Kl. **58**, 19 (1906).
[4] W. Magnus, Commun. Pure Appl. Math. **7**, 649 (1954); J. Wei, J. Math. Phys. **4**, 1337 (1963); and Ref. 1.
[5] R. M. Wilcox, J. Math. Phys. **8**, 962 (1967).

then we further *define*

$$\{y,f(x)\} = \sum_{k=0}^{\infty} a_k \{y, x^k\}. \tag{3}$$

It is easy to show that

$$\{\{y,f(x)\},g(x) = \{y,f(x)g(x)\}, \tag{4}$$

and if $f(x)g(x) = 1$, it is clear that

$$\{y,f(x)\} = w \tag{5a}$$

implies

$$y = \{w, g(x)\}. \tag{5b}$$

We shall also require a method of computing $d^l e^z/dt^l$, where t is a scalar variable, z is an operator, and $z = z(t)$. The simplest approach is first to derive an expression for the operator $V(s)$ which satisfies the equation (s is a scalar variable)[6]

$$e^{s(A+B)} = e^{sA}V(s), \tag{6}$$

with the initial condition $V(0) = I$.

Multiplying (6) on the left by e^{-sA} and differentiating with respect to s, we obtain the following differential equation for $V(s)$:

$$\frac{dV}{ds} = (e^{-sA}Be^{sA})V(s). \tag{7}$$

Solving (7) by iteration, we obtain the infinite series

$$V(s) = I + \int_{\xi=0}^{\xi=s} Q(\xi)\,d\xi + \int_{\xi_2=0}^{\xi_2=s}\int_{\xi_1=0}^{\xi_1=\xi_2} Q(\xi_2)Q(\xi_1)\,d\xi_1\,d\xi_2 + ,\cdots, \text{ etc.}, \tag{8}$$

where

$$Q(\xi) = e^{-\xi A}Be^{\xi A}. \tag{8'}$$

We may now compute the derivatives of e^z; thus

$$\frac{de^z}{dt} = \lim_{h\to 0}\left\{\exp\left[z(t) + h\frac{dz}{dt}\right] - \exp[z(t)]\right\}\Big/h. \tag{9}$$

If, in (6) and (8), we set $s = 1$, $A = z(t)$, $B = h(dz/dt)$, and substitute into (9), we obtain

$$\frac{de^z}{dt} = \lim_{h\to 0} e^z[V(1) - I]/h = e^z\int_0^1 e^{-\xi z}\frac{dz}{dt}e^{\xi z}\,d\xi. \tag{10}$$

Using the well-known formula[7]

$$e^{-x}ye^x = y + [y,x] + \frac{1}{2!}\{y,x^2\} + \cdots = \{y,e^x\}, \tag{11}$$

Eq. (10) may be put in the form

$$\frac{de^z}{dt} = e^z\left\{\frac{dz}{dt}, \frac{e^z - 1}{z}\right\}. \tag{12}$$

Equations (10) and (12) have appeared in the literature in various forms; a survey is given by Wilcox.[8] The proof given here[9] seems to be the most simple and direct, and is easily extended to yield higher derivatives. Thus, for the second derivative, we may put

$$\begin{aligned}\frac{d^2e^z}{dt^2} &= \lim_{h\to 0}[\exp(z(t+h)) + \exp(z(t-h)) - 2\exp(z)]/h^2 \\ &= \lim_{h\to 0}\left[\exp\left(z + \frac{dz}{dt}h + \frac{d^2z}{dt^2}\frac{h^2}{2!}\right) + \exp\left(z - \frac{dz}{dt}h + \frac{d^2z}{dt^2}\frac{h^2}{2!}\right) - 2\exp(z)\right]\Big/h^2.\end{aligned} \tag{13}$$

Setting $s = 1$, $A = z(t)$,

$$B = \pm\frac{dz}{dt}h + \frac{d^2z}{dt^2}\frac{h^2}{2!}$$

in (6) and (8), and substituting the results into (13), we find, letting $h \to 0$,

$$\frac{d^2e^z}{dt^2} = e^z\left\{\frac{d^2z}{dt^2}, \frac{e^z-1}{z}\right\} + 2e^z\int_{\xi=0}^{\xi=1}\left\{\frac{dz}{dt}, e^{\xi z}\right\}\left\{\frac{dz}{dt}, \frac{e^{\xi z}-1}{z}\right\}d\xi.$$

We obtain a slightly different form of this equation, more convenient for our purposes, by making use of the following integration by parts:

$$\int_{\xi=0}^{\xi=1}(UW + WU)\,d\xi = \left\{\frac{dz}{dt}, \frac{e^z-1}{z}\right\}^2,$$

where

$$U = \left\{\frac{dz}{dt}, e^{\xi z}\right\}$$

and

$$W = \left\{\frac{dz}{dt}, \frac{e^{\xi z}-1}{z}\right\}.$$

This result enables us to obtain the formula

$$\frac{d^2e^z}{dt^2} = e^z\left\{\frac{d^2z}{dt^2}, \frac{e^z-1}{z}\right\} + e^z\left\{\frac{dz}{dt}, \frac{e^z-1}{z}\right\}^2 + e^z\int_{\xi=0}^{\xi=1}(UW - WU)\,d\xi, \tag{14}$$

with U and W defined as above.

[6] R. Karplus and J. Schwinger, Phys. Rev. **73**, 1020 (1948); see also R. P. Feynman, Phys. Rev. **84**, 108 (1951).

[7] See, for instance, W. H. Louisell, *Radiation and Noise in Quantum Electronics* (McGraw-Hill Book Co., Inc., New York, 1964), p. 101.

[8] Reference 5.

[9] G. Hinds (private communication).

III. THE BCH COEFFICIENTS

We are now in a position to derive expressions for the BCH coefficients. We define $z(t)$ by the expression

$$e^z = e^x e^{ty}. \tag{15}$$

Differentiating with respect to t, using (12), we obtain

$$e^z y = e^z \left\{ \frac{dz}{dt}, \frac{e^z - 1}{z} \right\}. \tag{16}$$

Cancelling e^z and solving for dz/dt by means of (5), we obtain the following differential equation for $z(t)$:

$$\frac{dz}{dt} = \left\{ y, \frac{z}{e^z - 1} \right\}. \tag{17}$$

We assume an expression for $z(t)$ of the form

$$z(t) = \sum_{l=0}^{\infty} z_l t^l, \tag{18}$$

where the z_l are independent of t. It is clear from (15) that z_l is of degree l in y. Furthermore,

$$z_0 = z(0) = x \tag{19}$$

and

$$z_1 = \left. \frac{dz}{dt} \right|_{t=0} = \left\{ y, \frac{x}{e^x - 1} \right\}. \tag{20}$$

The classic Hausdorff recursion formula for z_l is obtainable from (17) and (18) (see Appendix); however, we prefer to proceed in a manner analogous to the derivation of z_1. That is, to obtain z_2 we specialize Eq. (14) by assuming the definition of $z(t)$ given by (15). When $z(t)$ is so defined, we may put

$$\frac{d^2 e^z}{dt^2} = e^z y^2 \tag{21}$$

and

$$\left\{ \frac{dz}{dt}, \frac{e^z - 1}{z} \right\}^2 = \left\{ \left\{ y, \frac{z}{e^z - 1} \right\}, \frac{e^z - 1}{z} \right\}^2 = y^2, \tag{22}$$

where we have used (17) and (4) to obtain (22).

Using (21) and (22), Eq. (14) becomes

$$\left\{ \frac{d^2 z}{dt^2}, \frac{e^z - 1}{z} \right\} = -\int_{\xi=0}^{\xi=1} [U, W]\, d\xi.$$

This equation may be solved for d^2z/dt^2 by means of (5); the result is

$$\frac{d^2 z}{dt^2} = -\left\{ \int_{\xi=0}^{\xi=1} [U, W]\, d\xi, \frac{z}{e^z - 1} \right\}. \tag{23}$$

Furthermore, since

$$z_2 = \frac{1}{2} \left. \frac{d^2 z}{dt^2} \right|_{t=0},$$

we obtain, letting $t \to 0$ and using (19) and the first equality of (20),

$$z_2 = -\frac{1}{2} \left\{ J, \frac{x}{e^x - 1} \right\}, \tag{24}$$

where

$$J = \int_{\xi=0}^{\xi=1} \left[\{z_1, e^{\xi x}\}, \left\{ z_1, \frac{e^{\xi x} - 1}{x} \right\} \right] d\xi. \tag{25}$$

These equations yield z_2 as a function of z_1 and x; we may obtain z_2 as a function of y and x by replacing z_1 with the right-hand side of (20). The result is then

$$z_2 = -\frac{1}{2} \left\{ K, \frac{x}{e^x - 1} \right\}, \tag{26}$$

where

$$K = \int_{\xi=0}^{\xi=1} \left[\left\{ y, \frac{x e^{\xi x}}{e^x - 1} \right\}, \left\{ y, \frac{e^{\xi x} - 1}{e^x - 1} \right\} \right] d\xi. \tag{27}$$

In a similar manner, z_l can be found for any l, by first calculating $d^l e^z/dt^l$, solving the resulting equation for $d^l z/dt^l$, and letting $t \to 0$. If we then set $t = 1$ in Eq. (18), we see that $z(1) = \sum_{k=0}^{\infty} z_k$ provides an expression for the operator z in the equation $e^z = e^x e^y$.

Snider[10] has obtained an expression for z_2 in terms of the commutator "super operator," which, however, seems to be more complicated than the form given here. In particular, it is somewhat difficult to expand his expression into a usable series, whereas (26) may be expanded in a straightforward way, as we now show.

IV. EXPANSION IN SERIES

To expand z_1, we note the well-known series

$$\frac{x}{e^x - 1} = \sum_{k=0}^{\infty} \frac{B_k}{k!} x^k, \tag{28}$$

where the B_k are the Bernoulli numbers. The first few numbers are $B_0 = 1$, $B_1 = -\frac{1}{2}$, $B_2 = \frac{1}{6}$, $B_4 = -\frac{1}{30}$, $B_6 = \frac{1}{42}$ and $B_{2k+1} = 0$ for $k > 0$. Thus, for z_1 we obtain the well-known expression

$$z_1 = y - \tfrac{1}{2}[y, x] + \tfrac{1}{12}\{y, x^2\} - \tfrac{1}{720}\{y, x^4\} + \cdots.$$

The series (28) must also be used in the expansion of z_2, as can be seen from (26). However, for the evaluation of z_2 we must, in addition, expand the integral (27). To accomplish this, we need the expressions

$$\frac{x e^{\xi x}}{e^x - 1} = \sum_{k=0}^{\infty} \varphi_k(\xi) x^k \tag{29}$$

and

$$\frac{e^{\xi x} - 1}{e^x - 1} = \sum_{k=0}^{\infty} \Psi_k(\xi) x^k. \tag{30}$$

[10] Reference 2, Appendix B.

Then K in Eq. (27) may be written

$$K = \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} [\{y, x^k\}, \{y, x^l\}] \int_{\xi=0}^{\xi=1} \varphi_k(\xi) \Psi_l(\xi)\, d\xi. \quad (31)$$

The functions $\varphi_k(\xi)$ appearing in (29) are closely related to the Bernoulli polynomials, and the following properties are easily proved:

$$\varphi_k(\xi) = \frac{1}{k!} \sum_{m=0}^{k} \binom{k}{m} B_m \xi^{k-m},$$

$$\frac{d}{d\xi} \varphi_{k+1}(\xi) = \varphi_k(\xi),$$

$$\varphi_k(0) = \frac{B_k}{k!},$$

$$\varphi_k(1) = \varphi_k(0) \quad \text{for} \quad k \neq 1,$$

$$\varphi_0(\xi) = 1,$$

$$\int_{\xi=0}^{\xi=1} \varphi_k(\xi)\, d\xi = \begin{cases} 0 & \text{for} \quad k > 0, \\ 1 & \text{for} \quad k = 0. \end{cases}$$

For the functions $\Psi_k(\xi)$ appearing in (30), we have

$$\Psi_k(\xi) = \varphi_{k+1}(\xi) - \varphi_{k+1}(0).$$

Also

$$\int_{\xi=0}^{\xi=1} \varphi_k(\xi) \Psi_l(\xi)\, d\xi = \begin{cases} \varphi_{l+1}(0) & \text{for} \quad k = 0, \quad l > 0, \\ \varphi_{k+1}(0) & \text{for} \quad l = 0, \quad k > 0. \end{cases}$$

Using these properties, we may derive the expansions

$$K = 2 \sum_{k=1}^{\infty} \{y, x^k, y\} \frac{B_{k+1}}{(k+1)!} + \sum_{l=1}^{\infty} \sum_{k=1}^{\infty} \Delta_{kl} I_{kl}, \quad (32)$$

and, from (26) and (32),

$$z_2 = -\sum_{l=1}^{\infty} \sum_{k=0}^{\infty} \{y, x^l, y, x^k\} \frac{B_{l+1}}{(l+1)!} \frac{B_k}{k!} - \frac{1}{2} \sum_{k=1}^{\infty} \sum_{l=1}^{\infty} \sum_{j=0}^{\infty} I_{kl} \{\Delta_{kl}, x^j\} \frac{B_j}{j!}, \quad (33)$$

where we have used the abbreviations

$$\{y, x^k, y\} = [\{y, x^k\}, y],$$

$$\{y, x^l, y, x^k\} = \{\{y, x^l, y\}, x^k\},$$

$$\Delta_{kl} = [\{y, x^k\}, \{y, x^l\}],$$

$$I_{kl} = \int_{\xi=0}^{\xi=1} \varphi_k(\xi) \varphi_{l+1}(\xi)\, d\xi.$$

In evaluating terms of these series, we may note that

$$\Delta_{kl} = -\Delta_{lk} \quad \text{and} \quad I_{kl} = 0, \quad \text{when} \quad k = l + 2.$$

The method described in this paper may be used to calculate any z_l, either in closed form as an integral [such as Eqs. (26) and (27)] or as an infinite series whose general term is known [as in (33)]. The calculations clearly become increasingly laborious as l increases, but it seems doubtful that explicit formulas exist which are any simpler than the ones obtainable in this manner.

APPENDIX

We now show that the classical Hausdorff recursion formula is easily derivable from the differential equation (17). We define the Hausdorff polarization operator uD_x, which is a linear operator acting only on other linear operators, by the formula

$$(uD_x)x^n = \sum_{k=0}^{k=n-1} x^k u x^{n-k-1}. \quad (A1)$$

The quantity u is itself an operator, which may be a function of the operator x. Equation (A1) may be used to define the action of the polarization operator on any function of x which is expandable in a power series in x.

We assume that the operators x, y, and z satisfy Eq. (15). If $f(y, z)$ is expandable in a power series in y and z, it is clear from the definition of the polarization operator that

$$\frac{d}{dt} f(y, z) = \left(\frac{dz}{dt} D_z\right) f(y, z). \quad (A2)$$

(Remember that z is a function of t, and y is not.) In particular, dz/dt is defined as a function of y and z by (17), so that we may write

$$\frac{d^2z}{dt^2} = \left(\frac{dz}{dt} D_z\right) \frac{dz}{dt} = \left(\left\{y, \frac{z}{e^z - 1}\right\} D_z\right) \frac{dz}{dt},$$

and in general we have, by the same reasoning,

$$\frac{d^n z}{dt^n} = \left(\left\{y, \frac{z}{e^z - 1}\right\} D_z\right) \frac{d^{n-1} z}{dt^{n-1}}. \quad (A3)$$

If we assume for $z(t)$ an expansion of the form (18), and let $t \to 0$ in (A3), we obtain, using (19),

$$n z_n = \left(\left\{y, \frac{x}{e^x - 1}\right\} D_x\right) z_{n-1}, \quad (A4)$$

which is precisely the Hausdorff recursion formula. Although in appearance this recursion is quite compact, it is in practice very cumbersome to use in obtaining series such as (33), and it does not lead at all to explicit integral expressions such as (26).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 7 JULY 1968

Remarks on the Burgers Equation

M. ROSENBLATT
University of California, San Diego, California

(Received 21 November 1967)

Periodic and aperiodic solutions of the Burgers equation $u_t + uu_x = \mu u_{xx}$, $\mu > 0$, are studied in this paper. A harmonic analysis of the solutions is carried out and the form of the spectrum is estimated for large time. Corresponding estimates of energy decay are also made. In Burgers' work on this equation, the case in which $\mu \downarrow 0$ with t fixed, and one then lets $t \to \infty$, is studied. In our investigation, a fixed value of $\mu > 0$ is taken and then one lets $t \to \infty$. A similar analysis is also carried out for an irrotational solution of a similar 3-dimensional system of equations. For large time and moderate wavenumbers there is, to the first order, a drift of spectral mass from low wavenumbers to higher wavenumbers. Comments are also made on the asymptotic distribution of a class of random solutions.

INTRODUCTION

Periodic and aperiodic solutions of the Burgers equation are studied in this paper. A harmonic analysis of the solutions is carried out and the form of the spectrum is estimated asymptotically for large time. Corresponding estimates of energy decay are also made. A similar analysis is also carried out for an irrotational solution of a similar 3-dimensional system of equations. For large time and moderate wavenumbers there is, to the first order, a drift of spectral mass from low wavenumbers to higher wavenumbers. Comments are also made on the asymptotic distribution of a class of random solutions of the equations.

1. THE BURGERS EQUATION

Burgers introduced the equation

$$u_t + uu_x = \mu u_{xx}, \quad \mu > 0, \tag{1.1}$$

as a model equation for $t > 0$, $-\infty < x < \infty$, in order to get some insight into turbulence.[1] Hopf[2] and Cole[3] showed that the study of this equation could be reduced to that of the heat equation. Specifically, if $u_0(x) = u(0, x)$ is continuous and

$$\int_0^x u_0(\xi)\, dx = o(x^2) \tag{1.2}$$

as $|x| \to \infty$, then

$$u(x, t) = \frac{\int_{-\infty}^{\infty} \frac{x - y}{t} \exp\left[-\frac{1}{2\mu} F(x, y, t)\right] dy}{\int_{-\infty}^{\infty} \exp\left[-\frac{1}{2\mu} F(x, y, t)\right] dy}, \tag{1.3}$$

with

$$F(x, y, t) = \frac{(x - y)^2}{2t} + \int_0^y u_0(y)\, dy,$$

is the unique solution of the initial-value problem. Letting

$$\varphi_0(y) = \exp\left[-\frac{1}{2\mu}\int_0^y u_0(\eta)\, d\eta\right],$$

$$\varphi(x, t) = (4\pi\mu t)^{-\frac{1}{2}} \int_{-\infty}^{\infty} \varphi_0(y) \exp\left[-\frac{(x - y)^2}{4\mu t}\right] dy, \tag{1.4}$$

it is clear that $\varphi(x, t)$ is a nonnegative solution of the heat equation

$$\varphi_t = \mu\varphi_{xx} \tag{1.5}$$

and that

$$u(x, t) = -2\mu \frac{\partial}{\partial x} \log \varphi(x, t). \tag{1.6}$$

Because of the continuity of u_0,

$$\lim_{\substack{t \to 0 \\ y \to x}} u(t, y) = u_0(x). \tag{1.7}$$

It is sometimes convenient to consider solutions of Burgers equation on the circle (or equivalently periodic solutions of the equation). One then has a corresponding result on existence and uniqueness of the solution without condition (1.2). A standard argument indicates that the solutions of (1.1), of period 2π with continuous initial data $u_0(x)$ and

$$\int_0^{2\pi} u_0(x)\, dx = 0, \tag{1.8}$$

are given in terms of solutions $\varphi(x, t)$ of the heat equation with period 2π,

$$\varphi(x, t) = \sum y_j \exp(-j^2\mu t) \exp(ijx), \tag{1.9}$$

with initial condition

$$\varphi(x, 0) = \varphi_0(x) = \sum y_j \exp(ijx) \tag{1.10}$$

given by (1.4) in terms of $u_0(x)$. Again, φ is positive so that $y_j = \bar{y}_{-j}$ and $y_0 > 0$. Then

$$u(x, t) = -2\mu \frac{\partial}{\partial x} \log \varphi(x, t).$$

[1] G. K. Batchelor, *The Theory of Homogeneous Turbulence* (Cambridge University Press, 1953).
[2] E. Hopf, Commun. Pure Appl. Math. 3, 201 (1950).
[3] J. D. Cole, Quart. Appl. Math. 9, 226 (1951).

It will sometimes be convenient to look at random solutions of the Burgers equation that are stationary in x. We look in particular at some stationary solutions obtained from solutions of the heat equation that are stationary in x. If φ is a solution of the heat equation on the circle $[0, 2\pi)$ (or, equivalently, a periodic solution with period 2π) it takes the form (1.9). Stationarity up to second-order moments implies that the coefficients y_j in representation (1.10) are orthogonal

$$Ey_j\bar{y}_k = \delta_{j-k}a_j \geq 0, \quad \sum a_j < \infty.$$

Stationarity up to second-order moments for solutions of the heat equation on the full real line implies that one can write

$$\varphi(x, t) = \int_{-\infty}^{\infty} e^{ix\lambda}e^{-\mu\lambda^2 t}\, dY(\lambda), \tag{1.11}$$

where $Y(\lambda)$ is a process with orthogonal increments

$$E\, dY(\lambda)\, d\bar{Y}(\mu) = \delta_{\lambda-\mu}\, dF(\lambda), \tag{1.12}$$

with F a bounded nondecreasing function. The sequence $\{a_j\}$ and function $F(\lambda)$ correspond to the second-order spectral decomposition of the process in the periodic and general case.

2. PERIODIC SOLUTIONS

Consider a periodic solution of the Burgers equation with period 2π and spatial mean zero; that is, condition (1.8) is satisfied. For continuous initial data, it has already been noted that the solution can be given in terms of a positive solution $\varphi(x, t)$ of the heat equation (1.5) with $y_0 > 0$. The continuity of initial data implies that

$$\sum_j |y_j|^2 < \infty. \tag{2.1}$$

In view of (1.6), the solution $u(x, t)$ of the Burgers equation can be written,

$$u(x, t) = -2\mu \frac{\sum_j y_j(ij) \exp(-j^2\mu t) \exp(ijx)}{\sum_j y_j \exp(-j^2\mu t) \exp(ijx)} = \sum_k Z_k(t) \exp(ikx). \tag{2.2}$$

If we formally look at the Burgers equation in the Fourier domain,

$$Z_k'(t) + \sum_j Z_{k-j}(t)Z_j(t)ij = -\mu k^2 Z_k(t) \tag{2.3}$$

is obtained for the Fourier coefficients $Z_k(t)$ of $u(x, t)$. This equation is certainly satisfied for sufficiently large t. We shall obtain estimates of $Z_k(t)$ with error term valid for large t and moderate k from the representation (2.2). These estimates will be used to show that, in the range of large t and moderate k, there is a unidirectional transfer of spectral mass from wavenumbers $1, 2, \cdots, k$ to $k + 1$ to the first order. This is consistent with some of the heuristic arguments employed in the discussion of nonlinear phenomena like turbulence (see Refs. 1–3).

Make the strong assumption

$$\sum_{j\neq 0} |y_j| < y_0, \tag{2.4}$$

or the stronger assumption (for a later computation)

$$\sum_{j\neq 0} |y_j|\, j^2 < y_0, \tag{2.4'}$$

so that the expansion

$$u(x, t) = 2\mu \frac{\psi_x}{y_0 - \psi} = 2\mu \sum_{j=0}^{\infty} y_0^{-j-1} \frac{\partial}{\partial x} \psi^{j+1} \Big/ (j + 1) \tag{2.5}$$

is valid, where

$$\varphi(x, t) = y_0 - \psi. \tag{2.6}$$

Even if (2.4) or (2.4') are not satisfied, they will be satisfied at some later time t by $y_j \exp(-j^2\mu t)$ so that the qualitative features of our conclusions are valid for any periodic solution satisfying (1.8).

Now

$$Z_k(t) = 2\mu ik \sum_{j=0}^{\infty} \frac{1}{j+1} \sum_{\substack{s_1+\cdots+s_{j+1}=k \\ s_i\neq 0}} y_{s_1}\cdots y_{s_{j+1}} \times \exp[-(s_1^2 + \cdots + s_{j+1}^2)\mu t] y_0^{-j-1} \tag{2.7}$$

and

$$Z_k'(t) = 2\mu^2 ik \sum_{j=0}^{\infty} \frac{1}{(j+1)} \times \sum_{\substack{s_1+\cdots+s_{j+1}=k \\ s_i\neq 0}} y_{s_1}\cdots y_{s_{j+1}}(-s_1^2 - \cdots - s_{j+1}^2) \times \exp[-(s_1^2 + \cdots + s_{j+1}^2)\mu t] y_0^{-j-1}. \tag{2.8}$$

Our aim is now to give the principal terms for $Z_k(t)$, $Z_k'(t)$ together with bounds for the error terms. First consider $Z_k(t)$ with $k \geq 1$ [$Z_{-k}(t) = \bar{Z}_k(t)$ since the solutions are real]. The principal term for large t is

$$2\mu i(y_1/y_0)^k \exp(-k\mu t) \tag{2.9}$$

corresponding to the term in the inner summation of (2.7) with $j + 1 = k$ and $s_1 = s_2 = \cdots = s_k = 1$. All the other terms of the inner summation corresponding

to $j+1=k$ are in total bounded by

$$2\mu k\Big[\sum_{s\neq 0}|y_s/y_0|\exp(-s^2\mu t)\Big]^{k-1} \times\Big[\sum_{s\neq 0,1}|y_s/y_0|\exp(-s^2\mu t)\Big]. \quad (2.10)$$

The total of the terms of the inner summation corresponding to fixed $j+1$ with $j+1>k$ is bounded by

$$2\mu k\Big[\sum_{s\neq 0}|y_s/y_0|\exp(-s^2\mu t)\Big]^{j+1} \quad (2.11)$$

Under assumption (2.4), the total of the terms of the inner summation corresponding to fixed $j+1$ with $j+1<k$ is bounded by

$$2\mu k\exp\{-M_{j+1}\mu t\}, \quad (2.12)$$

where

$$M_{j+1}=\min_{\substack{s_1+\cdots+s_{j+1}=k\\ s_i\neq 0}}\{s_1^2+\cdots+s_{j+1}^2\}. \quad (2.13)$$

However, $M_{j+1}\geq k+2\{k-(j+1)\}$ if $j+1<k$. Thus, for large t,

$$|Z_k(t)-2\mu i(y_1/y_0)^k e^{-k\mu t}|<\mu kC\exp[-(k+1)\mu t], \quad (2.14)$$

where C is an absolute constant. Under assumption (2.4′) a similar argument implies that $Z_k'(t)$ is to the first order

$$-2\mu^2 ik(y_1/y_0)^k e^{-k\mu t} \quad (2.15)$$

and that

$$|Z_k'(t)+2\mu^2 ik(y_1/y_0)e^{-k\mu t}|<\mu^2 kC\exp[-(k+1)\mu t], \quad (2.16)$$

with C an absolute constant. Let us now look at Eq. (2.3) with the estimates given by (2.14) and (2.16). Notice that

$$\Big|\sum_{\substack{j<0\\ \text{or } j>k}} Z_{k-j}(t)Z_j(t)ij\Big|\leq Ck^2e^{-\mu t}e^{-(k+1)\mu t}, \quad (2.17)$$

where C is an absolute constant. This implies that

$$Z'(t)+\sum_{j=1}^{k-1}Z_{k-j}(t)Z_j(t)ij+\mu k^2Z_k(t) = O\{(k^3e^{-\mu t})\exp[-(k+1)\mu t]\}, \quad (2.18)$$

as $t\to\infty$ where the constant, corresponding to the O order of magnitude estimate, is independent of k. At the very least, if $k^3\exp(-\mu t)=o(1)$ as $t\to\infty$, the error estimate on the right of Eq. (2.18) is negligible compared to the individual terms on the left of the equation since each of those terms is precisely of order of magnitude $\exp(-k\mu t)$. It follows that we have a drift of spectral mass from lower wavenumbers $1\leq j<k$ to k through the convolution term

$$\sum_{j=1}^{k-1}Z_{k-j}(t)Z_j(t)ij, \quad (2.19)$$

as given in a number of heuristic discussions of turbulence. This is enough to counterbalance the dissipative term $-\mu k^2Z_k(t)$ to such an extent that $Z_k(t)$ is of order of magnitude $\exp(-k\mu t)$ [see (2.14)] rather than $\exp(-k^2\mu t)$, as it would be if the convolution transfer term did not enter. Notice that all three terms of Eq. (2.3) are of order $\exp(-k\mu t)$.

Cole[4] has recently obtained estimates related to (2.14).

3. ENERGY ESTIMATES

We first consider a real deterministic solution of the Burgers equation on the real line. *Let u be a solution with u_t, u_x bounded and $u\in L$,*

$$\lim_{|x|\to\infty}u(x)=0. \quad (3.1)$$

The assumptions on u imply that $u\in L^2$. On multiplying the Burgers equation by u and integrating with respect to x, the following is obtained:

$$\begin{aligned}\int_{-N}^{N}[uu_t+u^2u_x]\,dx &= \frac{1}{2}\frac{d}{dt}\int_{-N}^{N}u^2\,dx+\tfrac{1}{3}[u^3(N)-u^3(-N)]\\ &= \mu\int_{-N}^{N}uu_{xx}\,dx\\ &= \mu uu_x\Big|_{-N}^{N}-\mu\int_{-N}^{N}u_x^2\,dx. \end{aligned} \quad (3.2)$$

The relation

$$\frac{1}{2}\frac{d}{dt}\int u^2\,dx=-\mu\int u_x^2\,dx \quad (3.3)$$

is obtained in the limit as $N\to\infty$ and describes the rate of decay of energy.

A similar argument can be used to give a corresponding result in the case of a random solution of the Burgers equation stationary in x. It will be more than enough to *assume that moments of u and its partial derivatives u_t, u_x,* and u_{xx} *exist up to fourth order.* Multiply the Burgers equation by u and take expectations to obtain

$$E\{uu_t+u_xu^2\}=\mu E(uu_{xx}). \quad (3.4)$$

The simple equalities

$$Eu_xu^2=\frac{1}{3}\frac{\partial}{\partial x}Eu^3=0$$

[4] J. D. Cole (unpublished).

and

$$\frac{\partial}{\partial x} Euu_x = Eu_x^2 + Euu_{xx} = 0$$

imply that

$$\frac{1}{2}\frac{d}{dt} Eu^2 = -\mu Eu_x^2. \tag{3.5}$$

We shall get estimates of the energy and the rate of decay of energy when t is large for a class of solutions of the Burgers equation in both the deterministic solution of the Burgers equation of the form $u = -2\mu\varphi_x(c + \varphi)^{-1}$, where $\max_x |\varphi| < c$ and

$$\varphi(x, t) = \int \exp(ix\lambda - \mu t\lambda^2) h(\lambda)\, d\lambda, \tag{3.6}$$

where $h \in L$ and is bounded. Now

$$\int |u(x, t)|^2\, dx = 4\mu^2 \int \left| \frac{\varphi_x}{c + \varphi} \right|^2 dx$$

and, since $|\varphi| \to 0$ as $t \to \infty$,

$$\int |u(x, t)|^2\, dx \cong 4\mu^2 c^{-2} \int |\varphi_x|^2\, dx, \tag{3.7}$$

as $t \to \infty$. The energy content of the solution u of the Burgers equation given above is the same as that of the solution $2\mu c^{-1}\varphi_x$ of the heat equation as $t \to \infty$. The rate of decay of total energy is governed by $\int |u_x|^2\, dx$ and

$$u_x(x, t) = -2\mu[(c + \varphi)\varphi_{xx} - \varphi_x^2](c + \varphi)^{-2}. \tag{3.8}$$

Thus

$$\int |u_x(x, t)|^2\, dx \cong 4\mu^2 c^{-4} \int |(c + \varphi)\varphi_{xx} - \varphi_x^2|^2\, dx,$$

as $t \to \infty$. Since

$$\int |\varphi_{xx}|^2 |c + \varphi|^2\, dx \cong c^2 \int |\varphi_{xx}|^2\, dx$$

as $t \to \infty$ and

$$\int (c + \varphi)\varphi_{xx}\varphi_x^2\, dx = \int \varphi\varphi_{xx}\varphi_x^2\, dx = -\frac{1}{3}\int \varphi_x^4\, dx,$$

it follows that

$$\int |u_x|^2\, dx \cong 4\mu^2 c^{-2} \int |\varphi_{xx}|^2\, dx + \tfrac{2}{3}\mu^2 c^{-4} \int |\varphi_x|^4\, dx, \tag{3.9}$$

as $t \to \infty$. Since

$$\sup_x |\varphi_x(x, t)| \leq \int \exp(-\mu t\lambda^2) |\lambda| |h(\lambda)|\, d\lambda = O[(\mu t)^{-1}]$$

as $t \to \infty$, then

$$\int |\varphi_x(t, x)|^4\, dx \leq C(\mu t)^{-2} \int \lambda^2 \exp(-2\mu t\lambda^2) |h(\lambda)|^2\, d\lambda$$

as $t \to \infty$ with C a constant. If $h(\lambda) \cong k\lambda^\alpha$, $\alpha \geq 0$, and $k \neq 0$, as $\lambda \downarrow 0$, we shall have the second term on the right of relation (3.9) of smaller order of magnitude than the first as $t \to \infty$, so that

$$\int |u_x|^2\, dx \cong 4\mu^2 c^{-2} \int |\varphi_{xx}|^2\, dx \tag{3.10}$$

as $t \to \infty$. The asymptotic rate of decay of energy will then be the same as for the solution $2\mu c^{-1}\varphi_x$ of the heat equation. Notice that our discussion of total energy and its rate of decay for large t depends on the low-frequency range of λ in the neighborhood of zero.

Similar conclusions can be obtained for a random solution $u = -2\mu\varphi_x(c + \varphi)^{-1}$, where

$$\varphi(x, t) = \int \exp(ix\lambda - \mu t\lambda^2)\, dZ(\lambda), \tag{3.11}$$

with $\varphi(x, 0)$ a bounded process stationary in x, of absolute value less than c and mean zero. Here $Z(\cdot)$ is the random spectral function corresponding to $\varphi(x, 0)$. We shall assume that $\varphi(x, 0)$ has second-order spectral density $f(\lambda)$ and fourth-order cumulant spectral density $g(\lambda_1, \lambda_2, \lambda_3)$. These will be well defined and continuous if the second-order moments and fourth-order cumulants are integrable as functions of the one and three linearly independent differences of spatial arguments respectively (see Ref. 5 for a discussion of higher-order spectra). As before, $|\varphi| \to 0$ as $t \to \infty$, so that

$$Eu^2 \cong 4\mu^2 c^{-2} E\varphi_x^2 = 4\mu^2 c^{-2} \int \lambda^2 \exp(-2\mu t\lambda^2) f(\lambda)\, d\lambda \tag{3.12}$$

as $t \to \infty$. The rate of dissipation of energy is given by (3.5), with u_x given by (3.8). Then

$$Eu_x^2 \cong 4\mu^2 c^{-4} E\,|(c + \varphi)\varphi_{xx} - \varphi_x^2|^2,$$

with

$$E\,|(c + \varphi)\varphi_{xx} - \varphi_x^2|^2 \cong c^2 E\varphi_{xx}^2 + \tfrac{5}{3} E\varphi_x^4.$$

Now, $E\varphi_x^4 = c_4(\varphi_x) + 3\sigma^4(\varphi_x)$, where c_4 is the fourth cumulant of φ_x and $\sigma^2(\varphi_x)$ is the variance of φ_x. Thus

$$\begin{aligned} E\varphi_x^4 = 3\left[\int \lambda^2 \exp(-2\mu t\lambda^2) f(\lambda)\, d\lambda\right]^2 \\ + \int \lambda_1 \lambda_2 \lambda_3 (-\lambda_1 - \lambda_2 - \lambda_3) \\ \times \exp\{-\mu t[\lambda_1^2 + \lambda_2^2 + \lambda_3^2 + (\lambda_1 + \lambda_2 + \lambda_3)^2]\} \\ \times g(\lambda_1, \lambda_2, \lambda_3)\, d\lambda_1\, d\lambda_2\, d\lambda_3. \end{aligned} \tag{3.13}$$

[5] Ya. G. Sinai, Theory Probability Appl. (USSR) **8**, 429 (1963).

Assuming that f and g are bounded and that f is not zero in the neighborhood of zero,

$$E\varphi_{xx}^2 \sim (\mu t)^{-\frac{5}{2}},$$

while

$$E\varphi_x^4 = O[(\mu t)^{-3}],$$

as $t \to \infty$. In that case the principal contribution to the decay of energy would be expected from the term $E\varphi_{xx}^2$ when t is large.

4. A SOLUTION ON THE REAL LINE

In this section we shall again look at a solution of the Burgers equation of the form $u = 2\mu\Psi_x(c - \Psi)^{-1}$ with $\max_x |\Psi| < c$ and Ψ given by (3.6). The function h is assumed to be bounded, real, and continuous with $h(0) = c > 0$. Our object is to get an estimate of the Fourier transform $f(\lambda, t)$ (with respect to x) of $u(x, t)$ at a fixed $\lambda \neq 0$ as $t \to \infty$. This is a limited parallel to the computation carried out in Sec. 2 for a periodic solution. Notice that as t gets large this will fall outside of both the energy-containing and energy-dissipative range for which $\lambda \sim (\mu t)^{-\frac{1}{2}}$ as $t \to \infty$ by arguments like those given in the section on energy estimates.

Now

$$u(x, t) = 2\mu \sum_{k=0}^{\infty} \frac{y_0^{-k-1} \dfrac{\partial}{\partial x} \Psi^{k+1}}{(k+1)}. \tag{4.1}$$

Consider the Fourier transform of Ψ^k $(k > 1)$

$$(2\pi)^{-(k-1)} \int \exp\left\{-\mu t\left[\left(\lambda - \sum_1^{k-1} \lambda_i\right)^2 t + \sum_1^{k-1} \lambda_i^2\right]\right\} \times h\left[\lambda - \sum_1^{k-1} \lambda_i\right] \prod_1^{k-1} h(\lambda_i)\, d\lambda_1 \cdots d\lambda_{k-1}. \tag{4.2}$$

Set

$$\lambda_i = \frac{\lambda}{k} + (2\mu t)^{-\frac{1}{2}} z_i.$$

Integral (4.2) can then be rewritten as

$$(2\pi)^{-(k-1)} \exp(-\mu t\lambda^2/k)(2\mu t)^{-(k-1)/2} \times \int \exp\left\{-\frac{1}{2}\left[\left(\sum_1^{k-1} z_i\right)^2 + \sum_1^{k-1} z_i^2\right]\right\} h\left[\frac{\lambda}{k} - (2\mu t)^{-\frac{1}{2}} \sum_1^{k-1} z_i\right] \times \prod_1^{k-1} h\left(\frac{\lambda}{k} + (2\mu t)^{-\frac{1}{2}} z_i\right) dz_1 \cdots dz_{k-1}. \tag{4.3}$$

This suggests the estimate

$$\exp(-\mu t\lambda^2/k)\left\{h\left(\frac{\lambda}{k}\right)\right\}^k (4\pi\mu t)^{-(k-1)/2} k^{-\frac{1}{2}} \tag{4.4}$$

for integral (4.3) to the first order and therefore

$$2\mu i\lambda \sum_{k=1}^{\infty} \exp(-\mu t\lambda^2/k)\left[y_0^{-1} h\left(\frac{\lambda}{k}\right)\right]^k (4\pi\mu t)^{-(k-1)/2} k^{-\frac{3}{2}} \tag{4.5}$$

as an estimate of $f(\lambda, t)$. For $\lambda \sim (\mu t)^{-1}$, it is clear that only the term corresponding to $k = 1$ is of interest as $t \to \infty$. Consider a fixed $\lambda > 0$ and assume that

$$h(\lambda) = h_0 + \lambda h_1 + o(\lambda), \quad \lambda > 0, \tag{4.6}$$

in the neighborhood of $\lambda = 0$. The following approximation is obtained by taking into account those terms in the summation (4.5) for values of k in the neighborhood of $k_{\max}$, which maximizes the typical term

$$f(\lambda, t) \sim 2\mu i\lambda \exp\left(-\lambda\left\{\mu t\left[\tfrac{1}{2} \log\left(\frac{4\pi\mu t y_0^2}{h_0^2}\right)\right]\right\}^{\frac{1}{2}} + O[\log(\mu t)]\right). \tag{4.7}$$

5. AN OVERSIMPLIFIED VERSION OF THE NAVIER-STOKES EQUATIONS

A three-dimensional example like the Burgers equation can be obtained by looking at the Navier-Stokes system of equations without the continuity equation and the pressure terms, and by looking for solutions that are gradients. We wish to consider Cauchy's problem for the equations

$$\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla)\mathbf{v} = \mu\Delta\mathbf{v}, \quad \mu > 0, \tag{5.1}$$

where $\mathbf{v} = (v_1, v_2, v_3)$ is a 3-vector depending on 3 spatial variables and time, and the initial vector $\mathbf{v}_0 = \mathbf{v}(x, 0)$ satisfies curl $\mathbf{v}_0 = 0$. This system is briefly mentioned at the end of Cole's paper[3] but is studied at some length in the paper of Kuznetsov and Rozhdestvenskii.[6] The fact that curl $\mathbf{v}_0 = 0$ implies that a solution with this initial specification is irrotational for all time. By introducing a scalar function φ satisfying

$$\mathbf{v} = -2\mu\nabla \log \varphi, \tag{5.2}$$

the initial-value problem for Eq. (5.1) is reduced to that of

$$\frac{\partial \varphi}{\partial t} = \mu\Delta\varphi. \tag{5.3}$$

If we set $\varphi(x, t) = y_0 - \Psi$, with $\max |\varphi| < y_0$ as before,

$$-2\mu \log \varphi - 2\mu \log y_0 = 2\mu \sum_{j=0}^{\infty} y_0^{-j-1} \Psi^{j+1}/(j+1). \tag{5.4}$$

[6] N. N. Kuznetsov and B. L. Rozhdestvenskii, Z. Vycisl. Mat. Fiz. 1, 217 (1961).

An argument like that given in Sec. 4 now suggests approximating the Fourier transform of Ψ^k $(k \geq 1)$ by

$$\exp(-\mu t\boldsymbol{\lambda}^2/k)\left\{h\left(\frac{\boldsymbol{\lambda}}{k}\right)\right\}^k(4\pi\mu t)^{-\frac{3}{2}(k-1)}k^{-\frac{3}{2}} \quad (5.5)$$

if Ψ is given by

$$\Psi(\mathbf{x}, t) = \int \exp(i\mathbf{x}\cdot\boldsymbol{\lambda} - \mu t\boldsymbol{\lambda}^2)h(\boldsymbol{\lambda})\, d\boldsymbol{\lambda} \quad (5.6)$$

where $\mathbf{x}$ and $\boldsymbol{\lambda}$ are 3-vectors, and h is bounded, real, and continuous. One can then proceed further and obtain estimates somewhat like (4.7) under further assumptions on h analogous to those made in Sec. 4. Instead of proceeding in this direction, we shall consider periodic solutions of system (5.1) in the manner of Sec. 2.

Let $\mathbf{v}$ be an irrotational solution of Eq. (5.1) with period 2π in each spatial variable and

$$\int_0^{2\pi} v^{(1)}(\mathbf{x}, t)\, dx_1 = \int_0^{2\pi} v^{(2)}(\mathbf{x}, t)\, dx_2 = \int_0^{2\pi} v^{(3)}(\mathbf{x}, t)\, dx_3 = 0. \quad (5.7)$$

Then the solution $\mathbf{v}$ is given by (5.2) in terms of a solution

$$\varphi = y_0 - \Psi = y_0 - \sum_{\mathbf{k}\neq\mathbf{0}} y_{\mathbf{k}} \exp(-\mathbf{k}^2\mu t)\exp(i\mathbf{k}\cdot\mathbf{x}) \quad (5.8)$$

of the heat equation. Notice that $\mathbf{k}$ is a wavenumber 3-vector of integers and $\mathbf{0}$ is the zero 3-vector. Assume, as in Sec. 2, that

$$\sum_{\mathbf{k}} |y_{\mathbf{k}}|\, \mathbf{k}^2 < y_0, \quad (5.9)$$

We can then get estimates for the vectors $\mathbf{z}_{\mathbf{k}}(t)$ and $\mathbf{z}'_{\mathbf{k}}(t)$, for large t, where the $\mathbf{z}_{\mathbf{k}}(t)$ are the Fourier coefficients in the harmonic analysis of $\mathbf{v}$:

$$\mathbf{v}(\mathbf{x}, t) = \sum \mathbf{z}_{\mathbf{k}}(t)e^{i\mathbf{k}\cdot\mathbf{x}} \quad (5.10)$$

Then

$$\mathbf{z}_{\mathbf{k}}(t) = 2\mu i\mathbf{k}\sum_{j=0}^{\infty}\frac{1}{(j+1)} \times \sum_{\substack{s_1+\cdots+s_{j+1}=\mathbf{k}\\ s_i\neq 0}} e^{-(s_1^2+\cdots+s_{j+1}^2)\mu t}y_{s_1}\cdots y_{s_{j+1}}y_0^{-j-1}. \quad (5.11)$$

For large t, it is clear that the first-order contribution to $\mathbf{z}_{\mathbf{k}}(t)$ in (5.11) will come from vectors $\mathbf{s} \neq 0$ with the following properties. The component s_i of $\mathbf{s}$ will be zero if $k_i = 0$, while $s_i = 0, 1$ $(s_i = 0, -1)$ if $k_i > 0$ $(k_i < 0)$. Call this set of vectors $U(\mathbf{k})$. Thus, if all the components of $\mathbf{k}$ are positive, $U(\mathbf{k})$ consists of the vectors (1, 0, 0), (0, 1, 0), (0, 0, 1), (1, 1, 0), (1, 0, 1), (0, 1, 1), and (1, 1, 1). If $\mathbf{k}$ has its first component negative and the others zero, $U(\mathbf{k})$ contains only (−1, 0, 0). Let us set

$$A(\mathbf{k}) = \sum_{j\geq 0}\frac{1}{j+1}\sum_{\substack{s_1+\cdots+s_{j+1}=\mathbf{k}\\ s_i\in U(\mathbf{k})}} y_{s_1}\cdots y_{s_{j+1}}y_0^{-j-1}. \quad (5.12)$$

Then, for fixed $\mathbf{k}$,

$$\mathbf{z}_{\mathbf{k}}(t) = 2\mu i\mathbf{k}A(\mathbf{k})\exp\left[-\sum_{i=1}^{3}|k_i|\,\mu t\right][1 + o(e^{-\mu t})] \quad (5.13)$$

and

$$\mathbf{z}'_{\mathbf{k}}(t) = -2\mu^2 i\mathbf{k}A(\mathbf{k})\left(\sum_{i=1}^{3}|k_i|\right) \times \exp\left(-\sum_{i=1}^{3}|k_i|\,\mu t\right)[1 + o(e^{-\mu t})], \quad (5.14)$$

as $t \to \infty$. Notice that the analog of the system (2.3) is

$$\mathbf{z}'_{\mathbf{k}}(t) + \sum_{\mathbf{k}'} i[\mathbf{k}'\cdot\mathbf{z}_{\mathbf{k}-\mathbf{k}'}(t)]\mathbf{z}_{\mathbf{k}'}(t) = -\mu\,|\mathbf{k}|^2\,\mathbf{z}_{\mathbf{k}}(t). \quad (5.15)$$

This is certainly satisfied by (5.11), for t sufficiently large. We shall just briefly remark on the analog of Eq. (2.18) for wavevectors $\mathbf{k}$ all of whose components are nonnegative. Given any two wavevectors $\mathbf{k}$ and $\mathbf{k}'$, $\mathbf{k} \geq \mathbf{k}'$ means that $\mathbf{k} - \mathbf{k}'$ is a vector all of whose components are nonnegative. The estimates (5.14) and (5.15) with error terms imply that

$$\mathbf{z}'_{\mathbf{k}}(t) + \sum_{\mathbf{k}\geq\mathbf{k}'\geq 0} i[\mathbf{k}'\cdot\mathbf{z}_{\mathbf{k}-\mathbf{k}'}(t)]\mathbf{z}_{\mathbf{k}'}(t) + \mu\,|\mathbf{k}|^2\,\mathbf{z}_{\mathbf{k}}(t) = O\left\{\exp\left[-\left(1+\sum_{i=1}^{3}|k_i|\right)\mu t\right]\right\}, \quad (5.16)$$

if $\mathbf{k} \geq \mathbf{0}$, $\mathbf{k} \neq \mathbf{0}$ with $\mathbf{k}$ fixed as $t \to \infty$. The error estimate is small compared to each of the terms on the left of (5.16) which are of order of magnitude $\exp\{-\sum_{i=1}^{3}|k_i|\,\mu t\}$. One has a drift of spectral mass from wavevectors $\mathbf{k}'$, $0 \leq \mathbf{k}' \leq \mathbf{k}$ to $\mathbf{k}$ through the truncated convolution term as before. If $\mathbf{k} \leq 0$, the convolution term would be summed over $\mathbf{k}'$, $\mathbf{k} \leq \mathbf{k}' \leq 0$, with a similar modification for $\mathbf{k}$ belonging to some other octant.

6. ASYMPTOTIC DISTRIBUTION OF A CLASS OF RANDOM SOLUTIONS.

Let us now consider random solutions of the Burgers equation that are stationary in x. Our object is to get some information on the asymptotic distribution of the solution as $t \to \infty$. The case of periodic solutions satisfying condition (1.8) is rather uninteresting, since it is clear from (2.2) that

$$u(x, t) \cong -2\mu(y_1/y_0)e^{ix}e^{-\mu t} \quad (6.1)$$

as $t \to \infty$. The asymptotic distribution of $u(x, t)$ (assuming that $y_0 > \delta > 0$ with probability one) is determined by the joint distribution of y_0 and y_1.

Consider now a random solution of the Burgers equation given by

$$u = 2\mu\Psi_x(c - \Psi)^{-1}, \tag{6.2}$$

with c a positive constant and Ψ a strictly stationary solution of the heat equation (1.5) in x with

$$\max_x |\Psi_0(x)| < c, \quad E\Psi_0(x) \equiv 0, \tag{6.3}$$

where $\Psi_0(x) = \Psi(x, 0)$. Notice that (6.3) implies that $\max |\Psi(x, t)| < c$ for all $t > 0$. The stationarity of Ψ_0 in x implies that $u_0(x) = u(x, 0)$, $u(x, t)$, and $\Psi(x, t)$ are strictly stationary. Let the second-order spectral distribution function of Ψ_0 be $F(\lambda)$,

$$E\Psi_0(y + x)\Psi_0(y) = \int e^{ix\lambda}\, dF(\lambda). \tag{6.4}$$

Since $u_0(x) = 2\mu\Psi_{0x}(c - \Psi_{0x})^{-1}$, it is clear that at the very least we ought to have

$$\int \lambda^2\, dF(\lambda) < \infty, \tag{6.5}$$

in order that $u_0(x)$ be well defined at least in mean square. Let us assume that Ψ_0 is strongly (or uniformly) mixing in the following sense.[7,8] Let A, B be two events with A determined by conditions on the random variables $\Psi_0(x)$ with $x \leq y$ and B determined by the random variables $\Psi_0(x)$ with $x \geq y + d$, $d > 0$. The process Ψ_0 is said to be strongly mixing if

$$|P(AB) - P(A)P(B)| < \alpha(d), \tag{6.6}$$

where $\alpha(d)$ is a function of d that tends to zero as $d \to \infty$ and that otherwise is independent of A, B, and y. This condition of strong mixing is a strong form of asymptotic independence as the events A and B are determined by random variables that are further and further removed spatially. Conditions sufficient for strong mixing are discussed in Refs. 7 and 8. Notice that strong mixing for Ψ_0 implies that u_0 is strongly mixing. Since

$$\Psi(x, t) = \int_{-\infty}^{\infty} (4\pi\mu t)^{-\frac{1}{2}} \exp\left\{-\frac{(x - y)^2}{4\mu t}\right\}\Psi_0(y)\, dy,$$

it is clear that $\max |\Psi(x, t)| \to 0$ as $t \to \infty$ and that

$$u(x, t) \cong 2\mu c^{-1}\Psi_x, \tag{6.7}$$

as $t \to \infty$. Now:

$$\Psi_x(x, t) = -\int (4\pi\mu t)^{-\frac{1}{2}} \exp\left\{-\frac{(x - y)^2}{4\mu t}\right\}\Psi_{0y}(y)\, dy. \tag{6.8}$$

Equation (6.8) indicates that Ψ_x is obtained from Ψ_{0x} by narrow band pass filtering about frequency zero and one knows that Ψ_{0x} is strongly mixing (since Ψ_0 is strongly mixing). Under appropriate conditions on second and fourth order moments (see Ref. 8), one would expect Ψ_x to be asymptotically normally distributed (when properly normalized) as $t \to \infty$. The assumption that Ψ_0 is strongly mixing implies that F is absolutely continuous with spectral density $f(\lambda)$. But then Ψ_{0x} has spectral density $\lambda^2 f(\lambda)$. One of the conditions required for asymptotic normality of Ψ_x is that the variance of $(4\pi\mu t)^{\frac{1}{2}}\Psi_x$ diverge as $t \to \infty$. But this variance

$$\sigma^2[(4\pi\mu t)^{\frac{1}{2}}\Psi_x] = 4\pi\mu t \int \lambda^2 e^{-2\mu t\lambda^2} f(\lambda)\, d\lambda = O(1). \tag{6.9}$$

We do not have divergence of (6.9) as $t \to \infty$ and so $u(x, t)$ as given by (6.7), is typically not asymptotically normal. The following objection might be brought against our argument. Even though u as given by (1.6) is stationary in x, the function φ might not be and it is this class of random solutions one ought to look at. Nonetheless, we shall see that random solutions of (5.1) stationary in $\mathbf{x}$ that are obtained from corresponding stationary solutions of the heat equation (5.3) by (5.2) are typically asymptotically normal under analogous conditions. This could be thought of as a difference between the one- and three-dimensional equations.

Let us consider a random solution of Eq. (5.1) given by

$$\mathbf{v} = 2\mu(c - \Psi)^{-1}\nabla\Psi \tag{6.10}$$

with c a positive constant and Ψ a strictly stationary solution of (5.1) in $\mathbf{x}$ with

$$\max_{\mathbf{x}} |\Psi_0(\mathbf{x})| < c, \quad E\Psi_0(\mathbf{x}) \equiv 0, \tag{6.11}$$

where $\Psi_0(\mathbf{x}) = \Psi(\mathbf{x}, 0)$. Let F be the second-order spectral distribution function of Ψ_0

$$E\Psi_0(\mathbf{y} + \mathbf{x})\Psi_0(\mathbf{y}) = \int \exp(i\mathbf{x}\cdot\boldsymbol{\lambda})\, dF(\boldsymbol{\lambda}). \tag{6.12}$$

One would have to require that

$$\int |\boldsymbol{\lambda}|^2\, dF(\boldsymbol{\lambda}) < \infty, \tag{6.13}$$

if $\mathbf{v}$ is to be well defined in mean square. The following definition is a possible generalization of the concept of strong mixing for a process $\Psi_0(\mathbf{x})$ with parameter $\mathbf{x}$ a point in 3-space. Consider any two events A, B determined by the random variables $\{\Psi_0(\mathbf{x}), \mathbf{x} \in I\}$ and $\{\Psi_0(\mathbf{x}'), \mathbf{x}' \in I'\}$, respectively. Let C_1, C_2 be the convex hulls of the set of points $\{\mathbf{x}, \mathbf{x} \in I\}$ and $\{\mathbf{x}', \mathbf{x}' \in I'\}$, respectively. Set d equal to the distance between C_1, C_2. We shall say that $\Psi_0(\mathbf{x})$ is strongly mixing if condition (6.6) is satisfied for all events A, B, where $\alpha(d)$ is a function of d that tends to zero as $d \to \infty$. If Ψ_0 is strongly mixing, then F is absolutely

[7] A. N. Kolmogorov and Y. A. Rozanov, Theory Probability Appl. (USSR) 5, 204 (1960).
[8] M. Rosenblatt, Quart. Appl. Math. 18, 387 (1961).

continuous with spectral density $f(\boldsymbol{\lambda})$. As $t \to \infty$ one can show that

$$\mathbf{v}(\mathbf{x}, t) \cong 2\mu c^{-1} \nabla \Psi \tag{6.14}$$

and

$$(4\pi\mu t)^{\frac{3}{2}} \nabla \Psi(\mathbf{x}, t) = -\int_{-\infty}^{\infty} \exp\left\{-\frac{|\mathbf{x}-\mathbf{y}|^2}{4\mu t}\right\} \nabla \Psi_0(\mathbf{y})\, d\mathbf{y}. \tag{6.15}$$

The spectral density of $(4\pi\mu t)^{\frac{3}{2}} \Psi_{x_i}(\mathbf{x}, t)$ is

$$\sigma^2[(4\pi\mu t)^{\frac{3}{2}} \Psi_{x_i}(\mathbf{x}, t)] = (4\pi\mu t)^3 \int \lambda_i^2 \exp(-2\mu t |\boldsymbol{\lambda}|^2) f(\boldsymbol{\lambda})\, d\boldsymbol{\lambda} \to \infty, \tag{6.16}$$

as $t \to \infty$ if $f(\boldsymbol{\lambda})$ is continuous and bounded away from zero at $\boldsymbol{\lambda} = 0$. Actually, less stringent conditions on $f(\boldsymbol{\lambda})$ at zero would suffice for (6.16). Again, it is clear from (6.15) that we are narrow band pass filtering about $\boldsymbol{\lambda} = 0$. But now because of (6.16), broad conditions on fourth-order moments assure asymptotic normality of (6.10).

ACKNOWLEDGMENTS

I should like to thank Professor Burgers, Professor Saffman, and Professor Whitham for their helpful comments.