JOURNAL OF
MATHEMATICAL
PHYSICS

VOLUME 6, NUMBER 5 MAY 1965

Relativistic Formulation of the Lifetime Matrix in the Potential Theory of Collision

TRAN TRONG GIEN
Department of Physics, Ohio University, Athens, Ohio
(Received 23 June 1964)

The lifetime matrices for collision processes as described by the Klein–Gordon and Dirac equations are formulated. The results obtained are exactly the same as in the case of a collision described by the Schrödinger equation. This gives us a consistent operator theory for the collision lifetime in the case of potential scattering.

INTRODUCTION

THROUGH a series of papers published by Smith,[1-5] the formulation of the lifetime matrix for collision processes and its applications were first developed and introduced into quantum mechanics. In these papers, Smith focuses his attention on the interaction described by a Schrödinger equation.

As is well known, most collision processes in nuclear and atomic physics are described by the Schrödinger equation. The question is now raised: is it possible to formulate a lifetime matrix for a collision process at higher energy, where within the potential theory of scattering, the Dirac and Klein–Gordon equations are used to describe the dynamics of these collisions?

This paper is devoted to the formulation of the lifetime matrix for interactions by Dirac or Klein–Gordon particles and proves that with an equivalent proposed form for the Q matrix, the relation between the Q and S matrices is the same as in the case of a Schrödinger particle.

To illustrate matters more simply, we study first the one-dimensional elastic scattering by Dirac particles, where no spin flip is assumed. It is shown that although the Dirac wavefunction is more complicated due to the presence of a spin part, the result still comes out nicely at the final step with the usual simple relation between the Q matrix and the phase shift.

Next, the case of elastic scattering in space with the central force field will be studied. It is well known that for a Dirac equation with a central force field, there are two different phase shifts for a state with a fixed value of j. These phase shifts are classified by opposite values of the operator κ and can be associated with those of the Dirac states which are well defined not only with a value of j, but also with a space parity. This property of the Dirac states implies that diagonal elements of the Q matrix should be formulated with states which have a well defined value of κ. Exactly the same relation between the elements of the Q and S matrices is found in this case.

The lifetime matrix for inelastic scattering by a Dirac particle is also formulated and again the relationship between the S and Q matrices is proved.

For the Klein–Gordon equation, in order to simplify matters, we omit the investigation of inelastic scattering and only consider elastic scattering. The results of this investigation show that the Q matrix formulation for the Klein–Gordon particles

[1] F. T. Smith, Phys. Rev. **118**, 349 (1960).
[2] F. T. Smith, J. Chem. Phys. **36**, 248 (1962).
[3] F. T. Smith, J. Chem. Phys. **38**, 1304 (1963).
[4] F. T. Smith, Phys. Rev. **130**, 394 (1963).
[5] F. T. Smith, Phys. Rev. **131**, 2803 (1963).

is still possible and the relation between the Q and S matrices is still valid.

I. DIRAC PARTICLE

A. One-Dimensional Elastic Scattering

The one-dimensional Dirac equation with the potential $V(x)$ is

$$[i\alpha_x\, d/dx - \beta m - V(x)]\psi = E\psi. \tag{1}$$

This equation has the asymptotic solution,

$$_\infty\psi = A(u\, e^{-ikx} - e^{i\eta}\, \chi e^{ikx}), \tag{2}$$

where ue^{-ikx}, χe^{ikx} are respectively incoming and outgoing parts, u and χ, the corresponding spinors, η is the phase shift, and A is a normalization constant. The natural units, $\hbar = c = 1$, are used here. It should be noted that the representations of the spinors u and χ are similar, except k in u should be replaced by $-k$ in χ.

The definition of the lifetime Q for the collision by a Dirac particle must be consistent with the non-relativistic definition as given by Smith.[1] Q should still be defined as the ratio of the average excess number of particles in the region of x from 0 to R to the inward or outward flux across a boundary at $x = R$, in the limit as R tends to ∞. This excess number of particles is the difference between the number of particles in the central region with the interaction present and the number of particles in the central region when the interaction is absent. The average over R of this quantity is denoted by $\langle I(R)\rangle$.

The probability density and the flux through a boundary at large are

$$\rho = \psi^\dagger(x)\psi(x),$$

and

$$\mathfrak{F} = -{}_\infty\psi^\dagger(x)\alpha_x\, {}_\infty\psi(x).$$

Therefore, the explicit form of the delay time is defined as

$$Q = \lim_{R\to\infty}\left[\int_0^R (\psi^\dagger(x)\psi(x) - \langle {}_\infty\psi^\dagger\, {}_\infty\psi\rangle)\, dx\right]_{\mathrm{Av}} \div |\mathrm{flux}|, \tag{3}$$

where

$$\langle {}_\infty\psi^\dagger\, {}_\infty\psi\rangle = \lim_{L\to\infty}\frac{1}{L}\int_0^L {}_\infty\psi^\dagger\, {}_\infty\psi\, dx. \tag{4}$$

As in the case of a scattering by a Schrödinger particle, it is possible to derive from the Dirac equation (1) the equation

$$\int_0^R \psi^\dagger\psi\, dx = i\left({}_\infty\psi^\dagger\alpha_x\frac{\partial\, {}_\infty\psi}{\partial E}\right)^R. \tag{5}$$

This equation is equivalent to Eq. (19) of Ref. 1. It is also possible to show that

$$\langle {}_\infty\psi^\dagger\, {}_\infty\psi\rangle = 2A^*Au^\dagger u. \tag{6}$$

The inward flux is equal to

$$\mathfrak{F}_+ = -A^*Au^\dagger\alpha_x u. \tag{7}$$

Explicit calculation of Q gives the following result,

$$Q = \lim_{R\to\infty}\left[\frac{\partial\eta}{\partial E} + 2\,\frac{u^\dagger\alpha_x\,\partial u/\partial E}{u^\dagger\alpha_x u}\sin(\eta + 2kR)\right]_{\mathrm{Av}},$$

or,

$$Q = \partial\eta/\partial E. \tag{8}$$

Thus, one sees that even for the scattering of a Dirac particle, one arrives at the result derived by Wigner.[6]

With $S = e^{i\eta}$, Q is linked to S by the relation

$$Q = -i(\partial S/\partial E)S^\dagger. \tag{9}$$

B. Elastic Scattering in Space

The Dirac equation for the central force is

$$[-\boldsymbol{\alpha}\cdot\mathbf{k} - \beta m + V(r)]\psi = E\psi \tag{10}$$

Defining the new operators[7]:

(a) the radial velocity operator

$$\alpha_r = (\boldsymbol{\alpha}\cdot\mathbf{r})/r, \tag{11}$$

(b) the radial momentum operator

$$k_r = (\mathbf{r}\cdot\mathbf{k} - i)/r, \tag{12}$$

(c) and the operator κ, where

$$\kappa = \rho(\boldsymbol{\sigma}'\cdot\mathbf{L} + 1), \tag{13}$$

the Dirac equation is written as

$$[-\alpha_r k_r - i\alpha_r\beta\kappa/r - \beta m + V(r)]\psi = E\psi. \tag{14}$$

The eigenvalues of κ are: ±1, ±2, ±3, $\pm4,\cdots$, $\pm(j + \frac{1}{2}),\ \cdots$.

By using a representation in which H and κ are diagonal, the radial part of the Dirac state

$$\psi_r = \frac{1}{r}\begin{pmatrix}F(r)\\ G(r)\end{pmatrix}$$

can be automatically separated.

The radial part of the Dirac equation splits into two parts

---

[6] E. Wigner, Phys. Rev. 98, 145 (1955).

[7] L. I. Schiff, *Quantum Mechanics* (McGraw-Hill Book Company, Inc., New York, 1955), p. 334.

$$d^2g/dr^2 + \left(k^2 - \kappa(\kappa+1)r^{-2} + V^2 - 2EV - \frac{\alpha'\kappa}{\alpha r} + \frac{1\alpha''}{2\alpha} - \frac{3\alpha'^2}{4\alpha^2}\right)g = 0 \quad (15a)$$

and

$$d^2f/dr^2 + \left(k^2 - \kappa(\kappa-1)r^{-2} + V^2 - 2EV + \frac{\beta'\kappa}{\beta r} + \frac{1\beta''}{2\beta} - \frac{3\beta'^2}{4\beta^2}\right)f = 0, \quad (15b)$$

where

$$g = G/\alpha^{\frac{1}{2}}, \quad f = F/\beta^{\frac{1}{2}}, \quad \alpha' = d\alpha/dr, \quad \beta' = d\beta/dr,$$
$$\alpha'' = d^2\alpha/dr^2, \quad \beta'' = d^2\beta/dr^2,$$

where $\alpha = E + m - V(r)$; $\beta = E - m - V(r)$.

It is also well known that to each value j of the total angular momentum J, correspond two opposite values of κ, $\kappa = \pm(j + \frac{1}{2})$. These two Dirac states behave asymptotically as states of opposite space parity. These states of opposite parity have different phase shifts. We call them successively η_κ^{I} [for the case $\kappa = (j + \frac{1}{2})$] and $\eta_\kappa^{\mathrm{II}}$ [for the case $\kappa = -(j + \frac{1}{2})$]. It should be noted that the great and small parts of a j state with fixed value of κ have a common phase shift.[8]

At large r, if the potential $V(r)$ falls off faster than $1/r^2$, these Dirac states can be written explicitly as

$${}_\infty\psi_r^{\mathrm{I}} = \frac{A}{r}\begin{bmatrix} \beta^{\frac{1}{2}}I_{|\kappa|-1}(kr) - e^{i\eta_\kappa^{\mathrm{I}}}\beta^{\frac{1}{2}}I^*_{|\kappa|-1}(kr) \\ \alpha^{\frac{1}{2}}I_{|\kappa|}(kr) - e^{i\eta_\kappa^{\mathrm{I}}}\alpha^{\frac{1}{2}}I^*_{|\kappa|}(kr) \end{bmatrix}, \quad (16a)$$

when $k = j + \frac{1}{2}$, and

$${}_\infty\psi_r^{\mathrm{II}} = \frac{A}{r}\begin{bmatrix} \beta^{\frac{1}{2}}I_{|\kappa|}(kr) - e^{i\eta_\kappa^{\mathrm{II}}}\beta^{\frac{1}{2}}I^*_{|\kappa|}(kr) \\ \alpha^{\frac{1}{2}}I_{|\kappa|-1}(kr) - e^{i\eta_\kappa^{\mathrm{II}}}\alpha^{\frac{1}{2}}I^*_{|\kappa|-1}(kr) \end{bmatrix}, \quad (16b)$$

when $k = -(j + \frac{1}{2})$.

The fact that there are two different phase shifts for Dirac states defined with opposite values of κ implies that these collision states have different delay times. Diagonal elements of the Q matrix should be therefore defined with a state, not only having fixed value of j, but also a definite space parity.

The definition of the matrix element $Q_{\kappa\kappa}$ is then

$$Q_{\kappa\kappa} = \lim_{R\to\infty}\left[\int^{r<R}(\psi_\kappa^\dagger\psi_\kappa - \langle{}_\infty\psi_\kappa^\dagger\,{}_\infty\psi_\kappa\rangle)\,d\tau\right]_{\mathrm{Av}} \div |\mathrm{flux}|, \quad (17)$$

where

$$\langle{}_\infty\psi_\kappa^\dagger\,{}_\infty\psi_\kappa\rangle = \frac{1}{V}\int^{r<R}{}_\infty\psi_\kappa^\dagger\,{}_\infty\psi_\kappa\,d\tau. \quad (18)$$

V is the volume of the sphere of radius R. It should be remarked that the average was taken over a sphere of finite radius R instead of a sphere of infinite radius, because $\langle{}_\infty\psi_\kappa^\dagger\,{}_\infty\psi_\kappa\rangle$ falls off as fast as $1/R^2$. The limit as $R \to \infty$ should be taken only when the integration over $d\tau$ has been performed.

It is still possible to show that the matrix element $Q_{\kappa\kappa}$ is equal to $\partial\eta_\kappa/\partial E$, where η_κ is the corresponding phase shift. In fact, from (14), the following relation can be derived

$$\int^{r<R}\psi_\kappa^\dagger\psi_\kappa\,d\tau = i\left(\mathfrak{R}_{\kappa,r}^\dagger\alpha_r\frac{\partial\mathfrak{R}_{\kappa,r}}{\partial E}\right)^R, \quad (19)$$

where

$$\mathfrak{R}_{\kappa,r} = r\psi_{\kappa,r}. \quad (20)$$

An equivalent expression is (the flux is assumed normalized to unity)

$$\int^{r<R}\psi_\kappa^\dagger\psi_\kappa\,d\tau = \left(F_\kappa^*\frac{\partial G_\kappa}{\partial E} - G_\kappa^*\frac{\partial F_\kappa}{\partial E}\right)^R = \partial\eta_\kappa/\partial E + 2R/v + \cdots, \quad (21)$$

where the notation $+ \cdots$ stands for the oscillating bounded terms which vanish as the average over R is taken. This equation is equivalent to Eq. (28) of Ref. 1.

The flux through the surface of a sphere of large radius R can be computed easily and is equal to

$$\mathfrak{F}_+ = -2k\,|A|^2. \quad (22)$$

The normalization to unit flux would mean choosing $|A|^2 = 1/2k$.

It is easy to show that the leading term of $\int^{r<R}\langle{}_\infty\psi_\kappa^\dagger{}_\infty\psi_\kappa\rangle d\tau$ in the limit as $R \to \infty$ is $2R/v$, where v is the ordinary velocity of the particle.

Thus,

$$Q_{\kappa\kappa} = \partial\eta_\kappa/\partial E = iS_\kappa\,\partial S_\kappa^\dagger/\partial E, \quad (23)$$

where

$$S_\kappa = e^{i\eta_\kappa}.$$

If the sign of κ is not known, one must superpose states of opposite values of κ and definite j in order to calculate the lifetime. One can write

$$\psi_j = a_{j1}\psi_\kappa^{\mathrm{I}} + a_{j2}\psi_\kappa^{\mathrm{II}}. \quad (24)$$

The lifetime Q_{jj} of a state ψ_j is related to the lifetimes $Q_{\kappa\kappa}^{\mathrm{I}}$ and $Q_{\kappa\kappa}^{\mathrm{II}}$ through the relation

$$Q_{jj} = |a_{j1}|^2\,Q_{\kappa\kappa}^{\mathrm{I}} + |a_{j2}|^2\,Q_{\kappa\kappa}^{\mathrm{II}}, \quad (25)$$

[8] These results can be easily derived.

where $|a_{i1}|^2$ and $|a_{i2}|^2$ are the usual probabilities. They are normalized such that

$$|a_{i1}|^2 + |a_{i2}|^2 = 1.$$

C. Inelastic Collision

The case where there are only two compound particles in each reaction channel will be considered. The reaction could be either of the type

$$A + B \rightarrow A^* + B^*,$$

or

$$A + B \rightarrow C + D.$$

The system of particles is represented by a state ψ_i, the single label j represents a set of quantum numbers such as the total angular momentum, the eigenvalue of κ, etc., as well as the total internal energy of the compound particles, which in the incoming channel of the state ψ_i is E_i.

The asymptotic behavior of the system corresponding to the incoming channel j can be described by

$$\phi_j = \mathcal{G}_j(\theta, \varphi, r)\omega_j(\mathrm{s})/r, \tag{26}$$

where $\omega_j(\mathrm{s})$ is the product of the wavefunctions describing the internal behavior of the colliding compound particles in the incoming channel j. The total internal energy is E_j with $(E - E_j)^2 = k_j^2 + \mu^2$, μ is the reduced mass of the particles, and $\mathcal{G}_j(\phi, \varphi, r)/r$ describes the relative motion of the colliding compound particles and is assumed to be a solution of a free Dirac equation. The complete state function of the system is

$${}_\infty\psi_i = \phi_i - \sum_k S_{ik}\phi_k^{(+)}. \tag{27}$$

The notation $\phi_k^{(+)}$ has been used to indicate the asymptotic outgoing states

$$\phi_k^{(+)} = \mathcal{G}_i^{(+)}(\theta, \varphi, r)\omega_i(\mathrm{s})/r. \tag{28}$$

The external region of the configuration space corresponding to a channel α is characterized by the vanishing of a certain term V_α of the Hamiltonian H.[9] If there is no exchange of particles in the reaction, V_α is assumed the same before and after the reaction occurs. The reaction occurs in one unique channel α in the restricted sense. In this case, $H - V_\alpha = H_\alpha$ will be the Hamiltonian of a set of eigenstates ${}_\alpha\psi_\alpha$ which are products of the free wavefunction describing the relative motion and the wavefunctions describing the internal states of the compound particles.

[9] H. Ekstein, Phys. Rev. **101**, 880 (1956).

In this section, only the reactions of this type are considered, since the Q matrix of a reaction with exchange of particles can be derived from the former by a natural extension.[1]

The element of the Q matrix is defined as

$$Q_{ij} = \lim_{R\to\infty}\left[\int^{|r|}\int^{|s|<R} \psi_i^\dagger\psi_i\, d\tau_r\, d\tau_s - R\sigma_{ij}\right]_{\mathrm{Av}}, \tag{29}$$

where

$$\sigma_{ij} = \lim_{R\to\infty}\frac{1}{R}\int^{|r|}\int^{|s|<R} {}_\infty\psi_j^\dagger\, {}_\infty\psi_i\, d\tau_r\, d\tau_s. \tag{30}$$

To simplify matters, we examine the problem for a simple system which depends only on two coordinates: x, the coordinate of the relative motion which must be always greater or equal to zero and y, the coordinate of internal motion. The extension to the more general case is straightforward.

In this case the incoming and outgoing wavefunctions at large x are,

$$\phi_i = A_i\, e^{-ik_i x}u_i\omega_i(y), \tag{31a}$$

and

$$\phi_k^{(+)} = A_k\, e^{ik_k x}\chi_k\omega_k(y). \tag{31b}$$

Within a unique channel, as defined above, the states ω_i form a complete set of eigenfunctions. They are orthogonal and normalized to unity,

$$\int_{-\infty}^{+\infty}\omega_i^*(y)\omega_j(y)\, dy = \delta_{ij}. \tag{32}$$

The following relation can be found

$$\int_0^R dx\int_{-\infty}^{+\infty} dy\psi_j^\dagger\psi_i = \int_{-\infty}^{+\infty} dy\left(i\,{}_\infty\psi_j^\dagger\alpha_x\frac{\partial\, {}_\infty\psi_i}{\partial E}\right)_{x=R}, \tag{33}$$

which is equivalent to equation (41) of Ref. 1.

The right-hand side of Eq. (33) can be calculated explicitly by using (31a), (31b) for ϕ_i and $\phi_k^{(+)}$. The following result is obtained,

$$\begin{aligned}\int_0^R dx\int_{-\infty}^{+\infty} dy\psi_j^\dagger\psi_i &= \frac{R}{v_i}\delta_{ij} + R\sum_k S_{jk}^*\frac{1}{v_k}S_{ik} - \sum_k S_{jk}^*\frac{\partial S_{ik}}{\partial E}\\ &\quad - \frac{1}{2i\mu}\frac{1}{v_j^2}S_{ji}^*\, e^{-2ik_jR} - \frac{1}{2i\mu}\frac{1}{v_i^2}S_{ji}\, e^{2ik_jR}.\end{aligned} \tag{34}$$

It is also possible to show

$$R\sigma_{ij} = \frac{R}{v_i}\delta_{ij} + \sum_k S_{jk}^*S_{ik}\frac{R}{v_k}. \tag{35}$$

Thus,

$$Q_{ij} = -i \sum_k S^*_{ik} \frac{\partial S_{ik}}{\partial E}, \tag{36}$$

or in matrix form,

$$\mathbf{Q} = -i(\partial \mathbf{S}/\partial E)S^\dagger = i\mathbf{S}\,\partial \mathbf{S}^\dagger/\partial E. \tag{37}$$

Integrating this last equation with suitable boundary conditions, we get

$$\mathbf{S} = 1 - i \int_E^{+\infty} \mathbf{Q}(E')\mathbf{S}(E')\,dE'. \tag{38}$$

It should be remarked that the Dirac equation describes particles of spin $\frac{1}{2}$ and since we have used it to describe the relative motion of the compound particles, the combined system in the center of mass must have total internal angular momentum equal to $\frac{1}{2}$. A reaction in which each channel has compound particles, one with total internal angular momentum equal to zero and the other to $\frac{1}{2}$ can be considered as an example for this.

II. KLEIN-GORDON PARTICLE

Only potentials with certain forms are acceptable for the Klein–Gordon equation.[10] Because of this limitation, the study of the inelastic collisions with the Klein–Gordon equation is obscure. For simplicity, we only investigate the formulation of the lifetime matrix for elastic collisions by Klein–Gordon particles.

A. One-Dimensional Elastic Scattering

The one-dimensional Klein–Gordon equation with the potential $V(x)$ is,

$$-d^2\psi/dx^2 + m^2\psi = [E - V(x)]^2\psi. \tag{39}$$

We consider the case where $V(x)$ is a component of a covariant 4-vector. The coordinate axes are chosen such that the three other components of this 4-vector are zero.

The asymptotic solution of the Klein–Gordon equation is

$$_\infty\psi = A(e^{-ikx} - e^{i\eta}e^{ikx}), \tag{40}$$

where η is the phase shift of the interaction and $k = (E^2 - m^2)^{\frac{1}{2}}$.

The lifetime Q is defined as

$$Q = \lim_{R\to\infty}\left[\int_0^R (\rho - \langle\rho_0\rangle)\,dx\right]_{\text{Av}} \div |\text{flux}|, \tag{41}$$

where ρ and ρ_0 are the charge densities of the Klein–Gordon equation with and without the interaction being present.

In this case,

$$\rho = [E - V(x)]\psi^*\psi/m, \tag{42a}$$

and,

$$\rho_0 = E_\infty\psi^*_\infty\psi/m. \tag{42b}$$

The flux of the Klein–Gordon equation can be calculated by using the familiar expression

$$\mathfrak{F} = \frac{1}{2mi}\left(\psi^*\frac{d\psi}{dx} - \psi\frac{d\psi^*}{dx}\right). \tag{43}$$

Thus, the inward flux through a boundary at large x is

$$\mathfrak{F}_+ = -k\,|A|^2/m. \tag{44}$$

If $|\mathfrak{F}|$ is normalized to 1, $|A|^2 = m/k$.

From the Klein–Gordon equation, it is possible to derive the following equation which is equivalent to Eq. (19) of Ref. 1:

$$\int_0^R \rho\,dx = \frac{1}{2m}\left[\frac{\partial\psi}{\partial E}\frac{\partial\psi^*}{\partial x} - \psi^*\frac{\partial^2\psi}{\partial x\,\partial E}\right]^R. \tag{45}$$

Explicit calculation of the right-hand side gives

$$\int_0^R \rho\,dx = \frac{2E}{k}R + \frac{\partial\eta}{\partial E} + \frac{1}{k}\frac{\partial k}{\partial E}\sin(\eta + 2kR). \tag{46}$$

It can be shown that

$$\langle\rho_0\rangle \equiv \lim_{L\to\infty}\frac{1}{L}\int_0^L \frac{E_\infty\psi^*_\infty\psi}{m}\,dx = \frac{2E}{k}. \tag{47}$$

Thus, we arrive at the same relation between Q and the delay time $\partial\eta/\partial E$,

$$Q = \partial\eta/\partial E. \tag{48}$$

Thus,

$$Q = -i(\partial S/\partial E)S^\dagger, \tag{49}$$

where

$$S = e^{i\eta}.$$

A slight difference in the definition of the lifetime for the Klein–Gordon case should be pointed out. The lifetime Q was defined with the charge densities ρ and ρ_0. This would mean with an appropriate potential $V(x)$, ρ could be negative in a certain range of x.

B. Elastic Scattering in Three Dimensions

Here, the potential is assumed to have arisen from a central force such that $V = V(r)$.

The Klein–Gordon equation is written as

$$(-\nabla^2 + m^2)\psi(\mathbf{r}) = [E - V(r)]^2\psi(\mathbf{r}). \tag{50}$$

[10] Reference 7, p. 321.

The radial wave equation is

$$\frac{d^2\phi_l}{dr^2} + \left(k^2 - \frac{l(l+1)}{r^2}\right)\phi_l + (V^2 - 2EV)\phi_l = 0, \tag{51}$$

where $k^2 = E^2 - m^2$.

If the potential $V(r)$ falls off faster than $1/r^2$, the asymptotic solution of the Klein–Gordon equation is

$$_\infty\psi_l(r, \theta, \varphi) = {}_\infty\phi(r)Y_{lm}(\theta, \varphi)/r, \tag{52}$$

with

$$_\infty\phi_l = A[I_l(kr) - e^{i\eta_l}I_l^*(kr)]. \tag{53}$$

$I_l(kr)$ and $I_l^*(kr)$ are the usual incoming and outgoing radial waves.

The elements of the Q matrix are to be associated with lifetimes of the collision systems which have well-defined angular momentum.

Thus,

$$Q_{ll} = \lim_{R\to\infty}\left[\int^{r<R}\left(\psi_l^*\frac{E-V}{m}\psi_l - \left\langle\frac{{}_\infty\psi_l^*E_\infty\psi_l}{m}\right\rangle\right)d\tau\right]_{\text{Av}} \div |\text{flux}|, \tag{54}$$

where

$$\left\langle\frac{{}_\infty\psi_l^*E_\infty\psi_l}{m}\right\rangle = \frac{1}{V}\int^{r<R}\frac{{}_\infty\psi_l^*E_\infty\psi_l}{m}d\tau = \frac{1}{V}\int_0^R\phi_l^*\frac{E}{m}\phi_l\,dr. \tag{55}$$

The inward flux through the surface of a large sphere of radius R is calculated in the usual way and we get

$$\mathfrak{F}_+ = -k\,|A|^2/m. \tag{56}$$

$|\mathfrak{F}|$ is normalized to unit flux, thus $|A|^2 = m/k$. In a similar manner, the following relation is found

$$\int^{r<R}\frac{\psi_l^*(E-V)\psi_l}{m} = \frac{1}{2m}\left(\frac{\partial\phi_l^*}{\partial r}\frac{\partial\phi_l}{\partial E} - \phi_l^*\frac{\partial^2\phi_l}{\partial E\,\partial r}\right)^R = 2R/v + \partial\eta_l/\partial E + \cdots. \tag{57}$$

Also,

$$\int^{r<R}\left\langle\frac{{}_\infty\psi_l^*E_\infty\psi_l}{m}\right\rangle d\tau = \frac{2}{v}R + \cdots. \tag{58}$$

Thus,

$$Q_{ll} = \partial\eta_l/\partial E = -iS_l^*\,\partial S_l/\partial E, \tag{59}$$

where

$$S_l = e^{i\eta_l}.$$

Thus, the same relation between the Q and S matrices is found,

$$\mathbf{Q} = -i(\partial\mathbf{S}/\partial E)\mathbf{S}^\dagger, \tag{60}$$

where **S** is a diagonal matrix with the diagonal elements $S_l = e^{i\eta_l}$.

III. CONCLUSION

Together with the results obtained by Smith in his first paper,[1] this work shows that the lifetime matrix can be formulated for all three dynamical equations (Dirac, Klein–Gordon, and Schrödinger) which are the available equations in the potential theory of scattering. The Dirac and Klein–Gordon equations are used in the case where relativistic effect must be considered.

We have also shown that most of the applications of the lifetime matrix as given in Ref. 1 can be derived in a similar fashion for the Klein–Gordon or Dirac equations although the computation in general needs more effort.

It should be pointed out that the relation between the Q and S matrices were only proved for potentials which fall off faster than $1/r^2$. For the scattering by a Coulomb potential, certainly, a new development is needed.

Since the relation between the Q and S matrices can be proved with all three dynamical equations within the potential theory of scattering, it is suggested that the relation between the Q and S matrices be accepted as a definition of one in terms of the other and vice versa in the theories where no potential exists. Since the theories for the S matrix have been well developed and established, the properties of Q could be known through those of S.

ACKNOWLEDGMENTS

The author wishes to thank Dr. Lawrence Gallaher of the Georgia Institute of Technology for his suggestions and stimulating discussions, Dr. Felix T. Smith of the Stanford Research Institute for his encouraging correspondences, his corrections, and suggestions, and Dr. Richard Koshel who helped in revising the manuscript.

A Set of Harmonic Functions for the Group $SU(3)$

MIRZA A. B. BÉG*
Institute for Advanced Study, Princeton, New Jersey, and The Rockefeller Institute, New York, New York†
AND
HENRI RUEGG
Institute for Advanced Study, Princeton, New Jersey, and CERN, Geneva, Switzerland†
(Received 26 August 1964)

We construct a set of harmonic functions carrying all the irreducible representations of the group $SU(3)$. Some features of these functions are discussed in detail.

1. INTRODUCTION

An important problem in the representation theory of linear Lie groups is the construction of basis vectors which transform according to irreducible representations of the group. For $GL(n)$ the problem can be solved completely by constructing tensors of assigned symmetry, in an n-dimensional (in general complex) vector space.[1] By appropriate specialization of these tensors we can construct basis vectors for any of the classical subgroups of $GL(n)$, for example $SU(n)$. The actual construction of these tensors involves standard manipulations with the algebra of the symmetric group; the process, however, becomes rapidly messy as we consider representations described by tensors of rank higher than 3.

An alternate procedure for constructing basis vectors, which bypasses the tensor analysis, is afforded by the theory of harmonic functions.[2] For our purpose, these functions may be defined as follows: Let $\mathfrak{M}$ be a differentiable manifold on which we can implement, as mappings, the transformations of the group $\mathcal{G}$. Furthermore, let $\mathfrak{M}$ be such that for every pair of points on $\mathfrak{M}$, there exists an element of $\mathcal{G}$ which takes one into the other. Now define a metric on the manifold

$$ds^2 = g_{ij}\, dx^i\, dx^j$$

and the invariant Laplace–Beltrami operator[3]

$$\Delta_{\mathfrak{M}} = \frac{1}{g^{\frac{1}{2}}} \frac{\partial}{\partial x^i}\left(g^{\frac{1}{2}} g^{ij} \frac{\partial}{\partial x^j}\right),$$

$$g \equiv \det(g_{ij}), \qquad g^{ij} = (g^{-1})_{ij}.$$

We shall refer to the eigenfunctions of $\Delta_{\mathfrak{M}}$ as the harmonic functions of $\mathcal{G}$ on $\mathfrak{M}$. By requiring these functions to be simultaneous eigenfunctions of all available commuting operators viz. (1) commuting generators of $\mathcal{G}$; (2) invariant operators of $\mathcal{G}$; (3) invariant operators of the subgroups of $\mathcal{G}$, we can construct basis vectors out of the harmonic functions.[4] While it is obvious from the start that these vectors carry representations of $\mathcal{G}$, it can be demonstrated without much difficulty that the representations they carry are, in fact, irreducible. It is important to note, however, that manifolds may exist, which satisfy the conditions stated earlier, and which lead to some but not all the irreducible representations of $\mathcal{G}$.

The above remarks can be illustrated by considering the group $SU(2)$.[5] The two-dimensional sphere is an invariant manifold and the eigenfunctions are the usual spherical harmonics $Y_l^m(\theta, \phi)$. These functions are simultaneous eigenfunctions of the two commuting operators and carry $(2l+1)$-dimensional irreducible representations of $SU(2)$. Since l is an integer, we obtain in this way only one-half of all the representations; to obtain all the representations the manifold must be enlarged. The connection with tensor analysis is immediate; the Y_l^m are nothing but the components of irreducible Cartesian tensors constructed out of the coordinates of a single point on the 2-sphere.

The purpose of the present paper is to construct a set of harmonic functions which carry all the irreducible representations of the group $SU(3)$. This group has assumed a position of paramount importance in the theory of strongly interacting elementary

* Work supported by the National Science Foundation at the Institute for Advanced Study and supported, in part, by the U. S. Atomic Energy Commission during the author's stay at Brookhaven, Summer 1964.
† Present address.
[1] See, e.g., H. Weyl, *The Classical Groups* (Princeton University Press, Princeton, New Jersey, 1946), Chap. IV.
[2] The mathematical literature on this subject is very extensive. We shall presume no foreknowledge on the part of the reader beyond some simple facts about the ordinary spherical harmonics. For a modern discussion, and a fairly complete list of references, see F. A. Berezin and I. M. Gel'fand, Amer. Math. Soc. Translations **21**, 193 (1962); F. A. Berezin, *ibid.*, p. 239 *et seq.*
[3] See, e.g., O. Veblen, *Invariants of Quadratic Differential Forms* (Cambridge Tracts in Mathematics and Mathematical Physics, No. 24, 1952).
[4] H. Weyl, Ann. Math. **35**, 486 (1934).
[5] For $SU(2)$, our notation follows that of A. R. Edmonds, *Angular Momentum in Quantum Mechanics* (Princeton University Press, Princeton, New Jersey, 1957).

particles[6] and it seems desirable to study this group in all its aspects. The physicist who is also an expert in group theory may find our considerations an interesting variant on customary procedures and be amused by the very simple nature of the harmonic functions which, to the best of our knowledge, have not been recorded in the literature. For the neophyte, we hope that the paper is also of some didactic value.

2. REPRESENTATIONS OF $SU(2)$

In order to establish the notation, and illustrate the procedure in a familiar context, we construct functions carrying all the representations of $SU(2)$. There is no new result in this section.

Let z_1, z_2 be two complex variables providing a point field for implementing the transformations of $SU(2)$. $[|z_1|^2 + |z_2|^2 = 1]$. Introduce spherical coordinates

$$z_1 = e^{i\phi_1}\cos\theta; \qquad z_2 = e^{i\phi_2}\sin\theta \tag{2.1}$$

and let $0 \leq \theta \leq \pi/2$, $0 \leq \phi_1 \leq 2\pi$, $0 \leq \phi_2 \leq 2\pi$. The complex vector (z_1, z_2) sweeps over all the points of a 3-sphere, thus the defining point field is congruent to the group manifold itself. This is a very special feature of $SU(2)$ and does not occur for the higher groups whose dimensionality is larger than the dimensionality of the defining representation.

The invariant metric is given by

$$\begin{aligned} ds^2 &= |dz_1|^2 + |dz_2|^2 \\ &= d\theta^2 + \cos^2\theta\, d\phi_1^2 + \sin^2\theta\, d\phi_2^2 \end{aligned} \tag{2.2}$$

and the Laplace–Beltrami operator is

$$\begin{aligned} \Delta_3 = \frac{1}{\sin\theta\cos\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\cos\theta\frac{\partial}{\partial\theta}\right) \\ + \frac{1}{\cos^2\theta}\frac{\partial^2}{\partial\phi_1^2} + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi_2^2}. \end{aligned} \tag{2.3}$$

The eigenvalues of Δ_3 are $-n(n+2)$ where n is a positive integer,[7] the eigenfunctions are solutions of

$$[\Delta_3 + n(n+2)]Y = 0. \tag{2.4}$$

In Eq. (2.4) the variables separate completely and the general solution can be constructed out of elementary solutions in the usual manner. The elementary solutions, regular over the point field, are given by

$$Y^n_{m_1m_2} = e^{i(m_1\phi_1 + m_2\phi_2)}\, d^{\frac{1}{2}n}_{\frac{1}{2}(m_1+m_2),\frac{1}{2}(m_1-m_2)}(2\theta) \tag{2.5}$$

where the d functions are defined in terms of Jacobi polynomials, exactly as in Edmonds.[5] The functions $Y^n_{m_1m_2}$ constitute a complete orthogonal set with the density function $g^{\frac{1}{2}} = \sin\theta\cos\theta$.

In order to identify the separation constants, n, m_1, m_2, we need the Lie algebra in terms of differential operators involving θ, ϕ_1, ϕ_2. The requisite expressions are

$$H_1 = \frac{1}{2i}\left(\frac{\partial}{\partial\phi_1} - \frac{\partial}{\partial\phi_2}\right), \tag{2.6}$$

$$E_+ = \frac{1}{2\sqrt{2}}e^{i(\phi_1-\phi_2)}\left[\frac{\partial}{\partial\theta} + \frac{\tan\theta}{i}\frac{\partial}{\partial\phi_1} + \frac{\cot\theta}{i}\frac{\partial}{\partial\phi_2}\right], \tag{2.7}$$

$$E_- = E_+^\dagger, \tag{2.8}$$

with the Casimir operator[8] given by

$$\begin{aligned} I^2 &\equiv H_1^2 + \{E_+, E_-\} \\ &= -\frac{1}{4}\left[\frac{\partial^2}{\partial\theta^2} + (\cot\theta - \tan\theta)\frac{\partial}{\partial\theta}\right. \\ &\qquad \left. + \frac{1}{\cos^2\theta}\frac{\partial^2}{\partial\phi_1^2} + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi_2^2}\right]. \end{aligned} \tag{2.9}$$

Hence

$$I^2 Y^n_{m_1m_2} = I(I+1)Y^n_{m_1m_2}, \quad I = \tfrac{1}{2}n, \tag{2.10}$$

$$H_1 Y^n_{m_1m_2} = I_3 Y^n_{m_1m_2}, \quad I_3 = \tfrac{1}{2}(m_1 - m_2). \tag{2.11}$$

Note that we get a set of harmonic functions for every allowed value of $\frac{1}{2}(m_1 + m_2)$; $2I + 1$ in all. The occurrence of this extra index is of course due to the fact that we have been working on the full group manifold and have thus picked up not only the harmonic functions but also the representation matrices.

3. REPRESENTATIONS OF $SU(3)$

A. Preliminary Considerations

Let (z_1, z_2, z_3) be three complex variables subject to the constraint

$$|z_1|^2 + |z_2|^2 + |z_3|^2 = 1. \tag{3.1}$$

The complex vector z_i now sweeps out the points of a 5-sphere. Clearly the 5-sphere is an invariant manifold for $SU(3)$, indeed it is an invariant manifold for $SO(6)$ which is homomorphic to $SU(4)$ and contains $SU(3)$ as a subgroup. By working on this manifold we do not expect to pick up all the irreducible representations of $SU(4)$; we shall show, however, that one can pick up all the irreducible representations of $SU(3)$.

[6] M. Gell-Mann, Phys. Rev. **125**, 1067 (1962); Y. Ne'eman, Nucl. Phys. **26**, 222 (1961); A. Salam and J. C. Ward, Nuovo Cimento **20**, 419 (1961).

[7] The eigenvalues of the Laplace–Beltrami operator on a ν sphere are $-n(n+\nu-1)$, where n is a positive integer.

[8] By Casimir operator we shall mean any operator of second or higher degree, constructed from the Lie algebra and commuting with all the infinitesimal generators.

The transformations of $SU(3)$ are implemented as mappings of the form $(z_1, z_2, z_3) \to (z_1', z_2', z_3')$. We can, in fact, demonstrate quite easily that to every pair of points $P \equiv (z_1, z_2, z_3)$, $Q \equiv (z_1', z_2', z_3')$, there always exists a transformation of $SU(3)$ that will take P to Q (or Q to P). Let us denote by R the point (1, 0, 0) and let u_1, u_2, u_3 be a triad of complex numbers such that

$$u_1^* z_1 + u_2^* z_2 + u_3^* z_3 = 0, \tag{3.2}$$

$$|u_1|^2 + |u_2|^2 + |u_3|^2 = 1. \tag{3.3}$$

Then the matrix

$$M(P \to R) = \begin{Vmatrix} z_1^* & z_2^* & z_3^* \\ u_1^* & u_2^* & u_3^* \\ w_1 & w_2 & w_3 \end{Vmatrix}, \tag{3.4}$$

where $w_i = \sum_{jk} \epsilon_{ijk} z_j u_k$, is unitary unimodular and takes P to R. Similarly, we can construct a matrix which takes Q to R, and hence the matrices which take P to Q,

$$M(P \to Q) = M^\dagger(Q \to R) M(P \to R), \tag{3.5}$$

and Q to P,

$$M(Q \to P) = M^\dagger(P \to Q). \tag{3.6}$$

Thus the 5-sphere is a suitable manifold for studying the irreducible representations of $SU(3)$. It will be convenient to parametrize the manifold in spherical coordinates

$$\left.\begin{aligned} z_1 &= e^{i\phi_1} \cos\theta \\ z_2 &= e^{i\phi_2} \sin\theta \cos\xi \\ z_3 &= e^{i\phi_3} \sin\theta \sin\xi \end{aligned}\right\}, \tag{3.7}$$

$$0 \le \phi_i \le 2\pi; \quad 0 \le \theta \le \frac{\pi}{2}; \quad 0 \le \xi \le \frac{\pi}{2}. \tag{3.8}$$

The invariant metric is then given by

$$\begin{aligned} ds^2 &= |dz_1|^2 + |dz_2|^2 + |dz_3|^2 \\ &= d\theta^2 + \cos^2\theta\, d\phi_1^2 \\ &\quad + \sin^2\theta (d\xi^2 + \cos^2\xi\, d\phi_2^2 + \sin^2\xi\, d\phi_3^2). \end{aligned} \tag{3.9}$$

Note the embedding of the group space of $SU(2)$ inside this manifold.

B. The Lie Algebra

Before we proceed further it will be convenient to have at hand the generators of infinitesimal transformations of $SU(3)$. The simplest procedure is to write down the generators of $SU(2)$ in complex Cartesian coordinates; the analogous construction for $SU(3)$ is then quite transparent. We shall label the generators and their linear combinations in accordance with Behrends *et al.*[9]; the numerical factors etc. are chosen such that the commutation relations agree with those of this reference.

The eight independent generators, with six of them combined into raising and lowering operators, can be exhibited as

$$\begin{aligned} H_1 &= (z_2 \partial_2 - z_3 \partial_3 - z_2^* \partial_2^* + z_3^* \partial_3^*)/2\sqrt{3} \\ &= \frac{1}{2i\sqrt{3}} \left(\frac{\partial}{\partial\phi_2} - \frac{\partial}{\partial\phi_3} \right), \end{aligned} \tag{3.10}$$

$$\begin{aligned} H_2 &= (-2z_1 \partial_1 + z_2 \partial_2 + z_3 \partial_3 \\ &\qquad + 2z_1^* \partial_1^* - z_2^* \partial_2^* - z_3^* \partial_3^*)/6 \\ &= \frac{1}{6i} \left(-2 \frac{\partial}{\partial\phi_1} + \frac{\partial}{\partial\phi_2} + \frac{\partial}{\partial\phi_3} \right), \end{aligned} \tag{3.11}$$

$$\begin{aligned} E_{+1} &= (z_2 \partial_3 - z_3^* \partial_2^*)/6^{\frac{1}{2}} \\ &= \frac{1}{2\sqrt{6}} e^{i(\phi_3 - \phi_2)} \left[\frac{\partial}{\partial\xi} + \frac{\tan\xi}{i} \frac{\partial}{\partial\phi_2} + \frac{\cot\xi}{i} \frac{\partial}{\partial\phi_3} \right], \end{aligned} \tag{3.12}$$

$$\begin{aligned} E_{+2} &= (z_2 \partial_1 - z_1^* \partial_2^*)/6^{\frac{1}{2}} \\ &= \frac{1}{2\sqrt{6}} e^{i(\phi_2 - \phi_1)} \left[-\cos\xi \frac{\partial}{\partial\theta} + \cot\theta \sin\xi \frac{\partial}{\partial\xi} \right. \\ &\qquad \left. + \frac{\cot\theta}{i\cos\xi} \frac{\partial}{\partial\phi_2} + \frac{\tan\theta\cos\xi}{i} \frac{\partial}{\partial\phi_1} \right], \end{aligned} \tag{3.13}$$

$$\begin{aligned} E_{+3} &= (z_3 \partial_1 - z_1^* \partial_3^*)/6^{\frac{1}{2}} \\ &= \frac{1}{2\sqrt{6}} e^{i(\phi_3 - \phi_1)} \left[-\sin\xi \frac{\partial}{\partial\theta} - \cot\theta \cos\xi \frac{\partial}{\partial\xi} \right. \\ &\qquad \left. + \frac{\cot\theta}{i\sin\xi} \frac{\partial}{\partial\phi_3} + \frac{\tan\theta\sin\xi}{i} \frac{\partial}{\partial\phi_1} \right], \end{aligned} \tag{3.14}$$

$$E_{-\alpha} = E_\alpha^\dagger, \qquad \alpha = 1, 2, 3. \tag{3.15}$$

(Note: $\partial_i \equiv \partial/\partial z_i$, $*$ means complex conjugate, $\dagger$ means Hermitian conjugate.)

From the generators one can construct the quadratic and cubic Casimir operators for $SU(3)$. For the quadratic operator we have

$$\begin{aligned} C_2 &= H_1^2 + H_2^2 + \sum_\alpha \{E_\alpha, E_{-\alpha}\} \\ &= -\frac{1}{12} \left[\frac{\partial^2}{\partial\theta^2} + (3\cot\theta - \tan\theta) \frac{\partial}{\partial\theta} + \frac{1}{\sin^2\theta} \frac{\partial^2}{\partial\xi^2} \right. \\ &\quad + \frac{1}{\sin^2\theta} (\cot\xi - \tan\xi) \frac{\partial}{\partial\xi} + \frac{1}{\cos^2\theta} \frac{\partial^2}{\partial\phi_1^2} \\ &\quad + \frac{1}{\sin^2\theta\cos^2\xi} \frac{\partial^2}{\partial\phi_2^2} + \frac{1}{\sin^2\theta\sin^2\xi} \frac{\partial^2}{\partial\phi_3^2} \\ &\quad \left. + \frac{1}{3} \left(\frac{\partial}{\partial\phi_1} + \frac{\partial}{\partial\phi_2} + \frac{\partial}{\partial\phi_3} \right)^2 \right]. \end{aligned} \tag{3.16}$$

[9] R. E. Behrends, J. Dreitlein, C. Fronsdal, and B. W. Lee, Rev. Mod. Phys. 34, 1 (1962).

Notice the existence of a linear operator

$$C_1 = \frac{1}{i}\left(\frac{\partial}{\partial\phi_1} + \frac{\partial}{\partial\phi_2} + \frac{\partial}{\partial\phi_3}\right) \tag{3.17}$$

which commutes with all the generators of $SU(3)$ and thus lies in the center. We leave it as an exercise to the reader to show that the cubic Casimir operator can be written as

$$C_3 = C_1[27C_2 - C_1^2 + 9]/162. \tag{3.18}$$

Finally we must choose an embedding of the Lie algebra of $SU(2)$ inside that of $SU(3)$. Our choice is dictated by the requirement that $SU(2)$ and $U(1)$ appear as mutually commuting subgroups of $SU(3)$.[10] This requirement is manifestly satisfied if we take (H_1, E_{+1}, E_{-1}) to constitute the Lie algebra of $SU(2)$ and H_2 as the infinitesimal generator of $U(1)$. The Casimir operator of this $SU(2)$ is given by

$$I^2 = -\frac{1}{4}\left[\frac{\partial^2}{\partial\xi^2} + (\cot\xi - \tan\xi)\frac{\partial}{\partial\xi} + \frac{1}{\cos^2\xi}\frac{\partial^2}{\partial\phi_2^2} + \frac{1}{\sin^2\xi}\frac{\partial^2}{\partial\phi_3^2}\right]. \tag{3.19}$$

C. Harmonic Functions of $SU(3)$ on the 5-Sphere

Corresponding to the metric Eq. (3.9) we obtain the Laplace–Beltrami operator

$$\Delta_5 = \frac{1}{\sin^3\theta\cos\theta}\frac{\partial}{\partial\theta}\left(\sin^3\theta\cos\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\cos^2\theta}\frac{\partial^2}{\partial\phi_1^2} + \frac{1}{\sin^2\theta}\left[\frac{1}{\sin\xi\cos\xi}\frac{\partial}{\partial\xi}\left(\sin\xi\cos\xi\frac{\partial}{\partial\xi}\right) + \frac{1}{\cos^2\xi}\frac{\partial^2}{\partial\phi_2^2} + \frac{1}{\sin^2\xi}\frac{\partial^2}{\partial\phi_3^2}\right]$$
$$= \tfrac{1}{3}C_1^2 - 12C_2. \tag{3.20}$$

The eigenvalues of Δ_5 are[7] $-n(n+4)$, where n is a positive integer; the equation to be solved is therefore

$$\Delta_5 Y + n(n+4)Y = 0. \tag{3.21}$$

Since the variables separate completely in Eq. (3.21) we can again construct the general solution out of elementary solutions. The dependence of the elementary regular solutions on ξ, ϕ_2, ϕ_3 is in fact already known to us from our discussion of $SU(2)$; the dependence on ϕ_1 is trivial. Separating out these variables we obtain

$$Y_n^{m_1m_2m_3I} = \Theta_n^{m_1,I}(\theta) \times d^I_{\frac{1}{2}(m_2+m_3),\frac{1}{2}(m_2-m_3)}(2\xi)e^{i(m_1\phi_1+m_2\phi_2+m_3\phi_3)}, \tag{3.22}$$

where $\Theta_n^{m_1,I}$ satisfies the differential equation

$$\frac{1}{\sin^3\theta\cos\theta}\frac{d}{d\theta}\left(\sin^3\theta\cos\theta\frac{d\Theta}{d\theta}\right) + \left[n(n+4) - \frac{m_1^2}{\cos^2\theta} - \frac{4I(I+1)}{\sin^2\theta}\right]\Theta = 0. \tag{3.23}$$

If we choose $\cos 2\theta$ as the independent variable in this equation, it reduces to a standard Papperitz equation[11] soluble in terms of hypergeometric functions. It is more convenient, however, to put $\Theta \sin\theta = \Theta_1$. The regular solution for Θ_1 then turns out to be simply another d function.[5] Carrying out the necessary reduction we obtain

$$Y_n^{m_1m_2m_3I} = \frac{e^{i(m_1\phi_1+m_2\phi_2+m_3\phi_3)}}{\sin\theta} \times d^{\frac{1}{2}(n+1)}_{\frac{1}{2}(m_1+2I+1),\frac{1}{2}(m_1-2I-1)}(2\theta) \times d^I_{\frac{1}{2}(m_2+m_3),\frac{1}{2}(m_2-m_3)}(2\xi). \tag{3.24}$$

The separation constants which occur in these functions can be readily identified in terms of the parameters customarily used[9] for characterizing the irreducible representations of $SU(3)$ and for labeling the states in a given representation.

The irreducible representations are defined through the transformation properties of symmetric traceless tensors with p contravariant and q covariant indices. If such a tensor is constructed out of a single complex vector (z_1, z_2, z_3), its components are homogeneous polynomials of degree $p + q$. Hence

$$n = p + q. \tag{3.25}$$

Now consider the transformation of these polynomials under the simultaneous phase transformations $\phi_i \to \phi_i + \epsilon$. Clearly the polynomials get multiplied by a common phase factor $e^{i\epsilon(p-q)}$. The eigenvalues of the infinitesimal generator are therefore $p - q$. This infinitesimal generator is none other than the central operator C_1 introduced earlier; hence the identification

$$m_1 + m_2 + m_3 = p - q. \tag{3.26}$$

The reader may check these identifications by calculating the eigenvalues of, say, the quadratic Casimir operator, which are $\frac{1}{3}(p^2+q^2+pq+3p+3q)$.

The states within a representation can be labeled through the eigenvalues of I^2, $I_3 = \sqrt{3}H_1$, $Y = 2H_2$,

[10] There are two algebraically distinct embeddings of the Lie algebra of $SU(2)$ in $SU(3)$. The embedding other than the one adopted here corresponds globally to $SO(3)$ and is discussed in detail by A. Dragt, "Classification of Three Particle States According to $SU(3)$" (IAS Preprint).

[11] See, e.g., P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill Book Company, Inc., New York, 1953), pp. 539.

i.e., the total isotopic spin, its projection along the 3-axis, and the hypercharge, respectively. Note that the physical hypercharge could have been defined as some other multiple of H_2 rather than twice H_2. The factor of 2 is chosen in accordance with the eightfold way of Ref. 6.

With these identifications we can write the harmonic functions as

$$\psi^{(p,q)}_{I,I_3,Y} = \frac{1}{\sin\theta} d^{\frac{1}{6}(p+q+1)}_{\frac{1}{6}(p-q-3Y+6I+3),\frac{1}{6}(p-q-3Y-6I-3)}(2\theta) \times d^{I}_{\frac{1}{3}(p-q)+\frac{1}{2}Y,I_3}(2\xi) \times e^{\frac{1}{2}i(p-q)(\phi_1+\phi_2+\phi_3)} e^{iI_3(\phi_2-\phi_3)} \times e^{\frac{1}{2}i(Y)(-2\phi_1+\phi_2+\phi_3)}. \quad (3.27)$$

The functions $\psi^{(p,q)}_{I,I_3,Y}$ constitute a complete orthogonal set over the interval

$$0 \le \theta \le \tfrac{1}{2}\pi, \quad 0 \le \xi \le \tfrac{1}{2}\pi, \quad 0 \le \phi_i \le 2\pi \quad (i = 1, 2, 3) \quad (3.28)$$

with the density function

$$g^{\frac{1}{2}} = \cos\theta \sin^3\theta \cos\xi \sin\xi. \quad (3.29)$$

Furthermore, these functions are regular (and thus single valued) for all nonnegative integer values of p and q. Thus they are carriers of all the irreducible representations of $SU(3)$.

It is important to note that while the $\psi^{(p,q)}_{I,I_3,Y}$ carry all the irreducible representations of $SU(3)$, they also carry *some* representations of $U(3)$. These representations of $U(3)$ are such that the extra conserved quantum number is precisely $p - q$, as is obvious from the fact that the operator C_1 can be regarded as a generator of the nine-dimensional group $U(3)$. The emergence of these representations of $U(3)$ is due to the fact that the harmonic functions were constructed out of the coordinates of a single triplet z^i and its complex conjugate $z_i \sim (z^i)^*$ (antitriplet). A simple example will suffice to illustrate the point: Let z^i, u^i be two distinct triplets and let us compound them to form the basis vectors of the adjoint representation of $SU(3)$. There are three *distinct* constructions

$$\psi_0 = u^i z_i - \tfrac{1}{3}\delta^i_i (u^\dagger z), \quad (3.30)$$

$$\psi_{+1} = u^i \epsilon_{jkl} u^k z^l, \quad (3.31)$$

$$\psi_{-1} = u_i \epsilon^{ikl} u_k z_l. \quad (3.32)$$

Under phase transformations $z^i \to e^{i\phi} z^i$, $u^i \to e^{i\phi} u^i$, $\psi_0 \to \psi_0$, $\psi_{+1} \to e^{i3\phi}\psi_{+1}$, $\psi_{-1} \to e^{-i3\phi}\psi_{-1}$. By restricting ourselves to a single triplet, we can realize ψ_0 but not ψ_{+1} or ψ_{-1}, i.e., we pick up an octet which transforms irreducibly under $U(3)$ and is left invariant by those transformations of $U(3)$ which do not belong to $SU(3)$.

4. CONCLUDING REMARKS

We have found a set of harmonic functions capable of carrying all the irreducible representations of $SU(3)$. These functions constitute a complete orthogonal set on the 5-sphere. By doing some "surgery" on the 5-sphere we could have restricted ourselves to representations of the quotient group $SU(3)/Z_3$ which may well be the group of interest in physics, rather than the full $SU(3)$. This "surgery" consists of defining the manifold through

$$0 \le \theta \le \tfrac{1}{2}\pi, \quad 0 \le \xi \le \tfrac{1}{2}\pi, \quad 0 \le \phi_1 + \phi_2 + \phi_3 \le 2\pi, \quad 0 \le \phi_2 - \phi_3 \le 4\pi, \quad 0 \le -2\phi_1 + \phi_2 + \phi_3 \le 4\pi. \quad (4.1)$$

The representations are then characterized by

$$p - q = 0 \bmod 3. \quad (4.2)$$

A shortcoming of these harmonic functions is the emergence of very specific representations of $U(3)$. If all physical states carry these representations, invariance under $U(3)$ would result in the additive conservation of $p - q$; while this is all right for meson systems, it is certainly wrong for baryon systems. One can get rid of these representations altogether by constructing *bilocal* harmonics of arbitrary parentage

$$\Psi^{(p,q)}_{I,I_3,Y}(z,u) = \sum_{\substack{(p_1,q_1)\\(p_2,q_2)\\ \nu}} P^{(p,q)\nu}_{(p_1,q_1),(p_2,q_2)} \times \sum_{\substack{I'I_3'Y'\\ I''I_3''Y''}} \psi^{(p_1,q_1)}_{I',I_3',Y'}(z)\psi^{(p_2,q_2)}_{I'',I_3'',Y''}(u) \times \begin{bmatrix} (p_1,q_1) & (p_2,q_2) & (p,q)_\nu \\ (I'I_3'Y') & (I''I_3''Y'') & (II_3Y) \end{bmatrix}, \quad (4.3)$$

where the last symbol in the rhs denotes a Clebsch–Gordan coefficient[12] for $SU(3)$ (the index ν reflects the fact that the group is not simply reducible) and the P's are arbitrary parentage coefficients. The ψ's are eigenfunctions of commuting operators constructed out of the infinitesimal generators, these generators may be taken to be a linear superposition of the infinitesimal generators at z and u respectively.[13] Note that the $U(3)$ operator $C_1 = C_1(z) +$

[12] M. Moshinsky, Rev. Mod. Phys. **34**, 813 (1962); J. J. deSwart, *ibid.* **35**, 916 (1963).

[13] Equation (3.18) must now be read as

$$C_3 = J[27C_2 - J^2 + 9]/162,$$

where J is *some* operator whose eigenvalue is $p - q$.

$C_1(u)$ is no longer diagonal. We can, of course, adjust the parentage so that C_1 is diagonal; the harmonics would then carry *all* the irreducible representations of $U(3)$ as well.

Finally we remark that the most general sets of harmonics would be obtained by working on the group manifold itself. Here we are faced with a problem of mathematical tractability that we have not yet resolved.

ACKNOWLEDGMENTS

We are indebted to Professor T. Regge for help and advice during the course of this work and to Professor J. R. Oppenheimer for hospitality at the Institute for Advanced Study. One of us (M.A.B. Bég) is indebted to Professor G. C. Wick for the hospitality extended to him at Brookhaven.

It is a pleasure to thank Dr. A. Dragt and Professor F. Gürsey for enjoyable discussions.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 6, NUMBER 5 MAY 1965

Operators that Lower or Raise the Irreducible Vector Spaces of U_{n-1} Contained in an Irreducible Vector Space of U_n*

JOHN G. NAGEL† AND MARCOS MOSHINSKY
Instituto de Física, Universidad de México, México, D. F.
(Received 21 May 1964)

We define operators that lower or raise the irreducible vector spaces of a semisimple subgroup of a semisimple Lie group contained in an irreducible vector space of the group. We determine the lowering and raising operators for the canonical subgroup U_{n-1} of the unitary group U_n. With the help of these operators, which are polynomial functions of the generators of U_n, and the corresponding operators for the subgroups in the canonical chain $U_n \supset U_{n-1} \supset \cdots \supset U_2 \supset U_1$ we can obtain, in this chain, the full set of normalized basis vectors of an irreducible vector space of U_n from any given normalized basis vector of the vector space. In particular we can obtain, using only the lowering operators, the set of basis vectors from the basis vector of highest weight of the vector space. This result is of importance in applications to many-body problems and in the determination of the Wigner coefficients of U_n. In future papers we plan to determine the lowering and raising operators for the orthogonal and symplectic groups.

I. INTRODUCTION AND SUMMARY

THE authors have been interested in developing group theoretical techniques which could be used for the analysis of a wide class of many-body problems.[1,2] They have also been concerned with the determination of the Wigner coefficients of the unitary groups,[3,4] coefficients which are presently of interest both in many-body problems and in elementary-particle physics.[3] One of the points of these programs was to determine the full set of basis vectors of an irreducible vector space of a semisimple Lie group. This has led to the introduction of the group theoretical concept of operators that lower or raise the irreducible vector spaces of a semisimple subgroup of a semisimple Lie group contained in an irreducible vector space of the group. These lowering (raising) operators shall be polynomial functions of the generators of the group that, when acting on a basis vector of an irreducible vector space of the group which is of given weight with respect to the subgroup, lower (raise) the weight. Furthermore they shall, when acting on the basis vector of highest weight of an irreducible vector space of the subgroup contained in the vector space of the group, transform it into the basis vector of highest weight of a lowered (raised) irreducible vector space of the subgroup.

The lowering operators in particular enable us to obtain, from the basis vector of highest weight of the highest irreducible vector space of the subgroup contained in an irreducible vector space of the group, the basis vectors of highest weight of all the other irreducible vector spaces of the subgroup contained in the vector spaces of the group.

One can choose the weight generators of a semisimple Lie group in such a way that the basis vector

* Work supported by Comisión Nacional de Energía Nuclear (México).
† Exchange Professor of the Organization of American States. Permanent Address: Blegdamsvej 17, Copenhagen, Denmark.
[1] M. Moshinsky, "Group Theory and the Many Body Problem," in *Physics of Many Particle System*, edited by E. Meeron (Gordon and Breach, New York, 1965), see also M. Moshinsky, Phys. Letters **6**, 305 (1963).
[2] M. Moshinsky and J. G. Nagel, Phys. Letters **5**, 173 (1963).
[3] M. Moshinsky, Rev. Mod. Phys. **34**, 813 (1962).
[4] M. Moshinsky, J. Math. Phys. **4**, 1128 (1963).

$C_1(u)$ is no longer diagonal. We can, of course, adjust the parentage so that C_1 is diagonal; the harmonics would then carry *all* the irreducible representations of $U(3)$ as well.

Finally we remark that the most general sets of harmonics would be obtained by working on the group manifold itself. Here we are faced with a problem of mathematical tractability that we have not yet resolved.

ACKNOWLEDGMENTS

We are indebted to Professor T. Regge for help and advice during the course of this work and to Professor J. R. Oppenheimer for hospitality at the Institute for Advanced Study. One of us (M.A.B. Bég) is indebted to Professor G. C. Wick for the hospitality extended to him at Brookhaven.

It is a pleasure to thank Dr. A. Dragt and Professor F. Gürsey for enjoyable discussions.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 6, NUMBER 5 MAY 1965

Operators that Lower or Raise the Irreducible Vector Spaces of U_{n-1} Contained in an Irreducible Vector Space of U_n*

JOHN G. NAGEL† AND MARCOS MOSHINSKY
Instituto de Física, Universidad de México, México, D. F.
(Received 21 May 1964)

We define operators that lower or raise the irreducible vector spaces of a semisimple subgroup of a semisimple Lie group contained in an irreducible vector space of the group. We determine the lowering and raising operators for the canonical subgroup U_{n-1} of the unitary group U_n. With the help of these operators, which are polynomial functions of the generators of U_n, and the corresponding operators for the subgroups in the canonical chain $U_n \supset U_{n-1} \supset \cdots \supset U_2 \supset U_1$ we can obtain, in this chain, the full set of normalized basis vectors of an irreducible vector space of U_n from any given normalized basis vector of the vector space. In particular we can obtain, using only the lowering operators, the set of basis vectors from the basis vector of highest weight of the vector space. This result is of importance in applications to many-body problems and in the determination of the Wigner coefficients of U_n. In future papers we plan to determine the lowering and raising operators for the orthogonal and symplectic groups.

I. INTRODUCTION AND SUMMARY

THE authors have been interested in developing group theoretical techniques which could be used for the analysis of a wide class of many-body problems.[1,2] They have also been concerned with the determination of the Wigner coefficients of the unitary groups,[3,4] coefficients which are presently of interest both in many-body problems and in elementary-particle physics.[3] One of the points of these programs was to determine the full set of basis vectors of an irreducible vector space of a semisimple Lie group. This has led to the introduction of the group theoretical concept of operators that lower or raise the irreducible vector spaces of a semisimple subgroup of a semisimple Lie group contained in an irreducible vector space of the group. These lowering (raising) operators shall be polynomial functions of the generators of the group that, when acting on a basis vector of an irreducible vector space of the group which is of given weight with respect to the subgroup, lower (raise) the weight. Furthermore they shall, when acting on the basis vector of highest weight of an irreducible vector space of the subgroup contained in the vector space of the group, transform it into the basis vector of highest weight of a lowered (raised) irreducible vector space of the subgroup.

The lowering operators in particular enable us to obtain, from the basis vector of highest weight of the highest irreducible vector space of the subgroup contained in an irreducible vector space of the group, the basis vectors of highest weight of all the other irreducible vector spaces of the subgroup contained in the vector spaces of the group.

One can choose the weight generators of a semisimple Lie group in such a way that the basis vector

* Work supported by Comisión Nacional de Energía Nuclear (México).
† Exchange Professor of the Organization of American States. Permanent Address: Blegdamsvej 17, Copenhagen, Denmark.
[1] M. Moshinsky, "Group Theory and the Many Body Problem," in *Physics of Many Particle System*, edited by E. Meeron (Gordon and Breach, New York, 1965), see also M. Moshinsky, Phys. Letters **6**, 305 (1963).
[2] M. Moshinsky and J. G. Nagel, Phys. Letters **5**, 173 (1963).
[3] M. Moshinsky, Rev. Mod. Phys. **34**, 813 (1962).
[4] M. Moshinsky, J. Math. Phys. **4**, 1128 (1963).

of highest weight of an irreducible vector space of the group is also the basis vector of highest weight of the highest irreducible vector space of a given subgroup of the group. So we can then, considering a chain of subgroups of the group, repeat the preceding analysis for each link of the chain and thus from the basis vector of highest weight of an irreducible vector space of the group obtain the full set of basis vectors of the vector space.

We now particularize our discussion to the unitary group U_n of n dimensions and its canonical subgroup U_{n-1}. In Sec. II we determine the explicit expressions (2.27) for the lowering and raising operators in terms of the generators of U_n. In Sec. III we derive recursion relations expressing the lowering and raising operators of U_n in terms of those of unitary groups of lower dimensions. In Sec. IV we discuss the matrix elements of the commutators of the lowering and raising operators with respect to the basis vectors of an irreducible vector space of U_n. These matrix elements are used in Sec. V in the derivation of the normalization coefficients of the lowering and raising operators. The normalization coefficients, which are given in Eqs. (5.13), are fundamental if we want to obtain a normalized basis vector when applying a lowering or a raising operator to a given normalized basis vector. In Eq. (5.8) we obtain any arbitrary normalized basis vector of an irreducible vector space of U_n in the $U_n \supset U_{n-1} \supset \cdots \supset U_2 \supset U_1$ chain by applying normalized lowering operators of these groups to the normalized basis vector of highest weight of the vector space.

The procedure followed here for the determination of the lowering and raising operators of the unitary groups is applicable to the determination of the lowering and raising operators of the orthogonal and symplectic groups. In future papers we plan to determine these operators.

II. LOWERING AND RAISING OPERATORS OF U_n

1. The Unitary Group U_n

Before applying the definition of the lowering and raising operators to the unitary group U_n of n dimensions, we first write down a few of the well-known properties of U_n.

The generators of U_n, which are denoted by $\mathcal{C}_\mu^{\mu'}$ where $1 \leq \mu, \mu' \leq n$, have the Hermiticity properties

$$\mathcal{C}_\mu^{\mu'\dagger} = \mathcal{C}_{\mu'}^{\mu}, \tag{2.1}$$

and fulfill the commutation relations

$$[\mathcal{C}_\mu^{\mu'}, \mathcal{C}_{\mu''}^{\mu'''}] = \delta_{\mu''}^{\mu'}\mathcal{C}_\mu^{\mu'''} - \delta_\mu^{\mu'''}\mathcal{C}_{\mu''}^{\mu'}. \tag{2.2}$$

From these commutation relations one obtains as special cases

$$[\mathcal{C}_\mu^{\mu}, \mathcal{C}_{\mu'}^{\mu'}] = 0,$$

and

$$[\mathcal{C}_\mu^{\mu}, \mathcal{C}_{\mu'}^{\mu''}] = (\delta_{\mu'}^{\mu} - \delta_\mu^{\mu''})\mathcal{C}_{\mu'}^{\mu''},$$

from which one sees that the generators can be divided into the three sets $\mathcal{C}_{\mu'}^{\mu}$, $\mu < \mu'$, $\mathcal{C}_\mu^\mu$ and $\mathcal{C}_\mu^{\mu'}$ $\mu < \mu'$ of lowering, weight, and raising generators respectively.[5,4]

An irreducible vector space of U_n can be characterized by $[h_{\mu n}]$ $1 \leq \mu \leq n$ which is the highest of the weights of the basis vectors of the vector space.[5,4] The basis vector of highest weight is unique even though U_n is not semisimple, due to the homomorphism of $U_1 \times SU_n$ with U_n, where SU_n is semisimple.[5,4] The $h_{\mu n}$ which are integers fulfill[6]

$$h_{1n} \geq h_{2n} \geq \cdots \geq h_{n-1n} \geq h_{nn}. \tag{2.3a}$$

We define a canonical subgroup of a group as a subgroup for which equivalent irreducible vector spaces of the subgroup contained in an arbitrary irreducible vector space of the group do not appear more than once.[7] For U_n a canonical subgroup is $U_{n-1} + 1$

$$\left(U_n\right) \supset \left(\begin{array}{c|c} U_{n-1} & \\ \hline & 1 \end{array}\right)$$

whose generators are $\mathcal{C}_\mu^{\mu'}$ $1 \leq \mu, \mu' < n$ and where $[h_{\mu n-1}]$ $1 \leq \mu < n$ is the highest of the weights of the basis vectors of an irreducible vector space of $U_{n-1} + 1$ and characterizes the vector space. The $h_{\mu\, n-1}$, besides satisfying

$$h_{1\,n-1} \geq h_{2\,n-1} \geq \cdots \geq h_{n-2\,n-1} \geq h_{n-1\,n-1}, \tag{2.3b}$$

also fulfill[6]

$$h_{\mu n} \geq h_{\mu\, n-1} \geq h_{\mu+1\, n}, \tag{2.4}$$

such that an irreducible vector space $[h_{\mu n-1}]$ fulfilling the inequalities (2.4) appears once and only once in the irreducible vector space $[h_{\mu n}]$.

A chain of canonical subgroups of U_n, whose irreducible vector spaces completely characterize the basis vectors of an irreducible vector space of U_n, is then the chain[7,4,8]

$$\left(U_n\right) \supset \left(\begin{array}{c|c} U_{n-1} & \\ \hline & 1 \end{array}\right) \supset \cdots \supset \left(\begin{array}{c|c} U_2 & \\ \hline & 1 \\ \hline & \ddots \\ \hline & 1 \end{array}\right) \supset \left(\begin{array}{c|c} U_1 & \\ \hline & 1 \\ \hline & 1 \\ \hline & \ddots \\ \hline & 1 \end{array}\right)$$

$$U_\nu \supset U_{\nu-1} + 1, \qquad 1 < \nu \leq n.$$

[5] G. Racah, lecture notes on Group Theory and Spectroscopy (CERN 61-8 1961).

[6] H. Weyl, *The Theory of Groups and Quantum Mechanics* (Dover Publications, Inc., New York), p. 391.

[7] G. E. Baird and L. C. Biedenharn, J. Math. Phys. **4**, 1449 (1963).

[8] I. M. Gelfand and M. L. Zetlin, Doklady Akad. Nauk USSR **71**, 825 (1950).

The $\mathcal{C}_\mu^{\mu'}$ $1 \leq \mu, \mu' \leq \nu \leq n$ are the generators of U_ν and $[h_{\mu\nu}]$ $1 \leq \mu \leq \nu \leq n$ characterizes an irreducible vector space of U_ν. The $h_{\mu\nu}$ fulfill relations similar to Eqs. (2.3) and (2.4) and so we can then, using the notation of Gelfand and Zetlin,[8] completely characterize the normalized basis vectors belonging to an irreducible vector space $[h_{\mu n}]$ of U_n by

$$|h_{\mu\nu}\rangle \equiv \left| \begin{array}{ccccccccc} h_{1n} & & \cdots & & \cdots & & & & h_{nn} \\ & h_{1\,n-1} & \cdots & & & & h_{n-1\,n-1} & & \\ & & \cdots & & & & & & \\ & & & h_{12} & & h_{22} & & & \\ & & & & h_{11} & & & & \end{array} \right\rangle, \qquad 1 \leq \mu \leq \nu \leq n.$$

We now find the eigenvalues of the unitary invariant of the group U_μ

$$N_\mu \equiv \sum_{\lambda=1}^{\mu} \mathcal{C}_\lambda^\lambda, \qquad 1 \leq \mu \leq n.$$

The eigenvalue is the same for all the basis vectors of an irreducible vector space of U_μ and can thus be found by acting on the basis vector of highest weight and so one has

$$N_\mu \left| \begin{array}{ccccccccc} h_{1n} & & \cdots & & \cdots & & & & h_{nn} \\ & & \cdots & & \cdots & & & & \\ & h_{1\mu} & \cdots & & h_{\mu\mu} & & & & \\ & & \cdots & & & & & & \\ & & & h_{11} & & & & & \end{array} \right\rangle = \sum_{\lambda=1}^{\mu} h_{\lambda\mu} \left| \begin{array}{ccccccccc} h_{1n} & & \cdots & & \cdots & & & & h_{nn} \\ & & \cdots & & \cdots & & & & \\ & h_{1\mu} & \cdots & & h_{\mu\mu} & & & & \\ & & \cdots & & & & & & \\ & & & h_{11} & & & & & \end{array} \right\rangle.$$

Hence

$$\mathcal{C}_\mu^\mu |h_{\lambda\nu}\rangle = (N_\mu - N_{\mu-1}) |h_{\lambda\nu}\rangle \equiv w_\mu |h_{\lambda\nu}\rangle, \tag{2.5}$$

where the μth component w_μ of the weight of $|h_{\lambda\nu}\rangle$ is given by

$$w_\mu = \sum_{\lambda=1}^{\mu} h_{\lambda\mu} - \sum_{\lambda=1}^{\mu-1} h_{\lambda\mu-1}. \tag{2.6}$$

Defining

$$\left|\begin{matrix} h \\ q \end{matrix}\right\rangle \equiv \left|\begin{matrix} h_1 & \cdots & h_n \\ q_1 & \cdots & q_{n-1} \end{matrix}\right\rangle \equiv \left| \begin{array}{ccccccccc} h_1 & & \cdots & & \cdots & & & & h_n \\ & q_1 & \cdots & & & & q_{n-1} & & \\ & & \cdots & & q_1 & \cdots & & q_{n-2} & \\ & & & & \cdots & & & & \\ & & & & & q_1 & & q_2 & \\ & & & & & & q_1 & & \end{array} \right\rangle,$$

one obtains from Eq. (2.6)

$$\mathcal{C}_\mu^\mu \left|\begin{matrix} h \\ q \end{matrix}\right\rangle = q_\mu \left|\begin{matrix} h \\ q \end{matrix}\right\rangle, \qquad 1 \leq \mu < n, \tag{2.7}$$

i.e. $|{}^h_q\rangle$ is the basis vector of highest weight of the irreducible vector space $[q_1 \cdots q_{n-1}]$ of U_{n-1}, so one has, using the raising generators of U_{n-1},

$$\mathcal{C}_\mu^{\mu'} \left|\begin{matrix} h \\ q \end{matrix}\right\rangle = 0, \qquad 1 \leq \mu < \mu' < n. \tag{2.8}$$

The basis vector

$$\left|\begin{matrix} h_1 & \cdots & h_n \\ h_1 & \cdots & h_{n-1} \end{matrix}\right\rangle \equiv \left| \begin{array}{ccccccccc} h_1 & & \cdots & & \cdots & & & & h_n \\ & h_1 & \cdots & & & & h_{n-1} & & \\ & & \cdots & & & & & & \\ & & & h_1 & & h_2 & & & \\ & & & & h_1 & & & & \end{array} \right\rangle,$$

is, as seen from the inequalities (2.4), the basis vector of highest weight of the highest irreducible vector space $[h_1 \cdots h_{n-1}]$ of U_{n-1} contained in the irreducible vector space $[h_1 \cdots h_n]$ of U_n. From Eqs. (2.7) and (2.6) one obtains

$$\mathcal{C}_\mu^\mu \left|\begin{matrix} h_1 & \cdots & h_n \\ h_1 & \cdots & h_{n-1} \end{matrix}\right\rangle = h_\mu \left|\begin{matrix} h_1 & \cdots & h_n \\ h_1 & \cdots & h_{n-1} \end{matrix}\right\rangle, \qquad 1 \leq \mu \leq n,$$

so one sees that it is also the basis vector of highest weight of the irreducible vector space $[h_1 \cdots h_n]$ of U_n, and hence one has, using the raising generators of U_n,

$$\mathcal{C}_\mu^{\mu'} \left|\begin{matrix} h_1 & \cdots & h_n \\ h_1 & \cdots & h_{n-1} \end{matrix}\right\rangle = 0, \qquad 1 \leq \mu < \mu' \leq n.$$

2. Definition of the Lowering and Raising Operators

We now define the lowering operator L_n^m of $U_n \supset U_{n-1} + 1$ as a polynomial function of the generators of U_n with the following two properties. The first is that L_n^m, when acting on a basis vector $|h_{\lambda\nu}\rangle$ of an irreducible vector space $[h_{\lambda n}]$ of U_n, which is of weight w_μ $1 \leq \mu < n$ given by Eq. (2.6) with respect to U_{n-1}, lowers the mth component of the weight by 1; i.e., L_n^m gives, when acting on $|h_{\lambda\nu}\rangle$, an arbitrary linear combination of basis vectors $|h'_{\lambda\nu}\rangle$ of weight w'_μ where

$$w'_\mu = w_\mu - \delta_\mu^m, \qquad 1 \leq \mu < n, \tag{2.9'}$$

so that from Eq. (2.6)

$$\sum_{\lambda=1}^{\nu} h'_{\lambda\nu} = \sum_{\lambda=1}^{\nu} h_{\lambda\nu} - \begin{cases} 0 & 1 \leq \nu < m \\ 1 & m \leq \nu < n, \end{cases} \tag{2.10'}$$

and besides,

$$h'_{\lambda n} = h_{\lambda n}, \qquad 1 \leq \lambda \leq n.$$

From Eq. (2.9') it then follows

$$[\mathcal{C}_\mu^\mu, L_n^m] |h_{\lambda\nu}\rangle = (\mathcal{C}_\mu^\mu L_n^m - L_n^m \mathcal{C}_\mu^\mu) |h_{\lambda\nu}\rangle$$
$$= (w'_\mu - w_\mu) L_n^m |h_{\lambda\nu}\rangle = -\delta_\mu^m L_n^m |h_{\lambda\nu}\rangle,$$

which is valid for all $|h_{\lambda\nu}\rangle$ so one obtains

$$[\mathfrak{C}_\mu^\mu, L_n^m] = -\delta_\mu^m L_n^m, \quad 1 \leq \mu, m < n. \tag{2.11'}$$

The second property of L_n^m is that, when acting on $|{}^h_q\rangle$, the irreducible vector space $[q_1 \cdots q_{n-1}]$ of U_{n-1} shall be lowered to $[q_1' \cdots q_{n-1}']$ such that the lowered weight w_μ' given by Eq. (2.9') becomes the highest of the weights of the basis vectors of the vector space $[q_1' \cdots q_{n-1}']$, i.e., L_n^m shall fulfill

$$L_n^m \left| \begin{matrix} h_1 \cdots\cdots\cdots h_n \\ q_1 \cdots q_m \cdots q_{n-1} \end{matrix} \right\rangle \propto \left| \begin{matrix} h_1 \cdots\cdots\cdots\cdots h_n \\ q_1 \cdots q_m - 1 \cdots q_{n-1} \end{matrix} \right\rangle, \quad 1 \leq m < n. \tag{2.12'}$$

Hence

$$[\mathfrak{C}_\mu^{\mu'}, L_n^m] \left| {}^h_q \right\rangle = (\mathfrak{C}_\mu^{\mu'} L_n^m - L_n^m \mathfrak{C}_\mu^{\mu'}) \left| {}^h_q \right\rangle = 0, \quad 1 \leq \mu < \mu' < n, \tag{2.13'}$$

which however does not imply that the commutator is zero, as the $|{}^h_q\rangle$ do not constitute a complete set of basis vectors.

From the commutation relations (2.2) and the Jacobi identity one can, by induction, easily prove that if L_n^m fulfills

$$[\mathfrak{C}_\lambda^{\lambda+1}, L_n^m] \left| {}^h_q \right\rangle = 0, \quad 1 \leq \lambda < n-1, \tag{2.14'}$$

then it also fulfills Eq. (2.13').

From Eq. (2.10') one sees that $L_n^m |h_{\lambda\nu}\rangle$ is not completely determined by the conditions (2.9') and (2.12') except (apart from an arbitrary factor) when $|h_{\lambda\nu}\rangle$ is of the form $|{}^h_q\rangle$. Clearly therefore Eqs. (2.11') and (2.13') are not sufficient to determine L_n^m uniquely. It is important though to notice that, in order to generate the full set of basis vectors of an irreducible vector space of U_n from its basis vector of highest weight, one only needs one possible solution for the lowering operators. We therefore settle for a particular solution of Eqs. (2.11') and (2.13').

The raising operator R_m^n is now defined in a completely similar way, i.e., as its first property R_m^n gives, when acting on $|h_{\lambda\nu}\rangle$ of weight w_μ $1 \leq \mu < n$ with respect to U_{n-1}, an arbitrary linear combination of basis vectors $|h'_{\lambda\nu}\rangle$ of weight w_μ' where

$$w_\mu' = w_\mu + \delta_m^\mu, \quad 1 \leq \mu < n. \tag{2.9''}$$

Hence

$$\sum_{\lambda=1}^{\nu} h'_{\lambda\nu} = \sum_{\lambda=1}^{\nu} h_{\lambda\nu} + \begin{cases} 0 & 1 \leq \nu < m \\ 1 & m \leq \nu < n, \end{cases} \tag{2.10''}$$

$$h'_{\lambda n} = h_{\lambda n}, \quad 1 \leq \lambda < n,$$

and

$$[\mathfrak{C}_\mu^\mu, R_m^n] = +\delta_m^\mu R_m^n \quad 1 \leq \mu, m < n. \tag{2.11''}$$

The second property of R_m^n is that

$$R_m^n \left| \begin{matrix} h_1 \cdots\cdots\cdots h_n \\ q_1 \cdots q_m \cdots q_{n-1} \end{matrix} \right\rangle \propto \left| \begin{matrix} h_1 \cdots\cdots\cdots\cdots h_n \\ q_1 \cdots q_m + 1 \cdots q_{n-1} \end{matrix} \right\rangle, \quad 1 \leq m < n. \tag{2.12''}$$

Hence

$$[\mathfrak{C}_\mu^{\mu'}, R_m^n] \left| {}^h_q \right\rangle = 0, \quad 1 \leq \mu < \mu' < n, \tag{2.13''}$$

which is fulfilled if

$$[\mathfrak{C}_\lambda^{\lambda+1}, R_m^n] \left| {}^h_q \right\rangle = 0, \quad 1 \leq \lambda < n-1. \tag{2.14''}$$

In the same way as was indicated for the lowering operators one sees that Eqs. (2.11'') and (2.13'') do not determine R_m^n uniquely, so again we shall settle for a particular solution of these equations. With the help of L_n^m and R_m^n any other lowering or raising operator can be formed as products of these.

3. Derivation of the Lowering and Raising Operators

We now derive explicit formulas for the lowering and raising operators satisfying Eqs. (2.11) and (2.14), starting with the lowering operators.

The most general polynomial function L_n^m of the generators which satisfies Eq. (2.11') is a linear combination of products of the form $\mathfrak{C}_{\mu_2}^{\mu_1'} \mathfrak{C}_{\mu_4}^{\mu_3'} \cdots \mathfrak{C}_{\mu_{q-2}}^{\mu_{q-3}'} \mathfrak{C}_{\mu_q}^{\mu_{q-1}'}$ which, as seen from the commutation relations (2.2), have all the indices $\mu \neq m, n$ paired, while the index m appears once more above than below and hence the index n once more below than above. Using the commutation relations we can then write L_n^m as

$$L_n^m = \sum_{p=0}^{n-2} \sum_{\mu_p, \mu_{p-1}, \cdots, \mu_2, \mu_1 = 1}^{n-1} \mathfrak{C}_{\mu_1}^m \mathfrak{C}_{\mu_2}^{\mu_1} \cdots \mathfrak{C}_{\mu_p}^{\mu_{p-1}} \mathfrak{C}_n^{\mu_p} \times \mathfrak{D}_{m\mu_1\mu_2\cdots\mu_{p-1}\mu_p n}, \quad 1 \leq m < n, \quad m \neq \mu_i \neq n$$

$$\text{for } 1 \leq i \leq p, \text{ and } \mu_i \neq \mu_j \text{ for } i \neq j, \tag{2.15}$$

where the term corresponding to $p = 0$ is, by definition, the term not containing any μ_i, i.e., the term $\mathfrak{C}_n^m \mathfrak{D}_{mn}$, and where $\mathfrak{D}_{m\mu_1\cdots\mu_p n}$ is a function of closed cycles of the generators, i.e.,

$$\mathfrak{D}_{m\mu_1\cdots\mu_p n} = \mathfrak{D}_{m\mu_1\cdots\mu_p n}(\mathfrak{C}_{\lambda_2}^{\lambda_1} \mathfrak{C}_{\lambda_3}^{\lambda_2} \cdots \mathfrak{C}_{\lambda_1}^{\lambda_q}), \quad q \geq 1, \quad 1 \leq \lambda_i \leq n. \tag{2.16}$$

If in a given term of Eq. (2.15) the product $\mathfrak{C}_{\mu_2}^{\mu_1} \mathfrak{C}_{\mu_3}^{\mu_2} \cdots \mathfrak{C}_{\mu_p}^{\mu_{p-1}} \mathfrak{C}_n^{\mu_p}$ contains a raising generator of U_{n-1}, i.e., $\mu_1 < m$ or $\mu_{i+1} < \mu_i$, it could, using the

commutation relations (2.2), be moved to the right, and so when acting on $|{}^{h}_{q}\rangle$ in Eq. (2.12′) it would, according to Eq. (2.8), give zero. Therefore, if we have a solution of Eqs. (2.11′) and (2.12′) of the form (2.15), we can construct another solution of the form

$$L_n^m = \sum_{p=0}^{n-m-1} \sum_{\mu_p > \mu_{p-1} > \cdots > \mu_2 > \mu_1 = m+1}^{n-1} \mathcal{C}_{\mu_1}^{m} \mathcal{C}_{\mu_2}^{\mu_1} \cdots \mathcal{C}_{\mu_p}^{\mu_{p-1}} \mathcal{C}_{n}^{\mu_p} \times \mathfrak{D}_{m\mu_1\mu_2\cdots\mu_{p-1}\mu_p n}, \qquad 1 \leq m < n, \tag{2.17}$$

in which the product does not contain any raising generators of U_{n-1}. This reflects the nonuniqueness of the lowering operators mentioned in the previous subsection and for the sake of simplicity we restrict ourselves to lowering operators of the form (2.17), which then are equivalent to Eq. (2.15) with respect to Eq. (2.12′), but however not with respect to Eq. (2.9′).

We now need to apply the condition (2.14′) to L_n^m of the form (2.17) with $\mathfrak{D}_{m\mu_1\cdots\mu_p n}$ of the form (2.16) in order to obtain the final formula for L_n^m. Let us however first define some useful operators

$$\mathcal{E}_{\mu\mu'} \equiv \mathcal{C}_{\mu}^{\mu} - \mathcal{C}_{\mu'}^{\mu'} + \mu' - \mu, \tag{2.18}$$

which, as seen from Eqs. (2.18), (2.1), and (2.2), have the properties

$$\mathcal{E}_{\mu'\mu} = -\mathcal{E}_{\mu\mu'}, \tag{2.19}$$

$$\mathcal{E}_{\mu\mu''} + \mathcal{E}_{\mu''\mu'} = \mathcal{E}_{\mu\mu'}, \tag{2.20}$$

$$\mathcal{E}_{\mu\mu'}^{\dagger} = \mathcal{E}_{\mu\mu'}, \tag{2.21}$$

and

$$[\mathcal{E}_{\mu\mu'}, \mathcal{C}_{\mu''}^{\mu'''}] = \{(\delta_{\mu'}^{\mu} - \delta_{\mu}^{\mu'''}) - (\delta_{\mu''}^{\mu'} - \delta_{\mu}^{\mu'''})\}\mathcal{C}_{\mu''}^{\mu'''}. \tag{2.22}$$

From Eqs. (2.18) and (2.5) one obtains the eigenvalues of the $\mathcal{E}_{\mu\mu'}$

$$\mathcal{E}_{\mu\mu'} |h_{\lambda\nu}\rangle = (w_\mu - w_{\mu'} + \mu' - \mu) |h_{\lambda\nu}\rangle, \qquad 1 \leq \mu, \mu' \leq n, \tag{2.23}$$

where w_μ and $w_{\mu'}$ are given by Eq. (2.6). In the special case when $|h_{\lambda\nu}\rangle$ is of the form $|{}^{h}_{q}\rangle$ and $1 \leq \mu, \mu' < n$ one obtains from Eqs. (2.18) and (2.7)

$$\mathcal{E}_{\mu\mu'} \left|\begin{matrix} h \\ q \end{matrix}\right\rangle = q_{\mu\mu'} \left|\begin{matrix} h \\ q \end{matrix}\right\rangle, \qquad 1 \leq \mu, \mu' < n, \tag{2.24}$$

where

$$q_{\mu\mu'} \equiv q_\mu - q_{\mu'} + \mu' - \mu,$$

from which follows

$$q_{\mu'\mu} = -q_{\mu\mu'}, \tag{2.25}$$

and

$$q_{\mu\mu''} + q_{\mu''\mu'} = q_{\mu\mu'},$$

and besides, from Eq. (2.3b)

$$q_{\mu\mu'} \gtreqless 0, \qquad \mu' \gtreqless \mu. \tag{2.26}$$

From Eq. (2.23) we see that with respect to the basis vectors $|h_{\lambda\nu}\rangle$ the operator $\mathcal{E}_{\mu\mu'}$ is represented by a diagonal matrix some of whose diagonal elements can be zero, i.e., the matrix is not nonsingular. However for the particular basis vectors $|{}^{h}_{q}\rangle$ the eigenvalues $q_{\mu\mu'}$ of $\mathcal{E}_{\mu\mu'}$ for $1 \leq \mu, \mu' < n$, $\mu \neq \mu'$ are, according to Eq. (2.26), always different from zero.

In Appendix A it is then shown [see Eq. (A1)] that

$$L_n^m = \left(\sum_{p=0}^{n-m-1} \sum_{\mu_p > \mu_{p-1} > \cdots > \mu_2 > \mu_1 = m+1}^{n-1} \mathcal{C}_{\mu_1}^{m} \mathcal{C}_{\mu_2}^{\mu_1} \cdots \mathcal{C}_{\mu_p}^{\mu_{p-1}} \mathcal{C}_{n}^{\mu_p} \times \prod_{i=1}^{p} \mathcal{E}_{m\mu_i}^{-1} \right) \prod_{\mu=m+1}^{n-1} \mathcal{E}_{m\mu}, \qquad 1 \leq m < n, \tag{2.27a′}$$

which is homogeneous of degree $n - m$ in the operators $\mathcal{C}_{\mu'}^{\mu}$ and $\mathcal{E}_{m\mu}$, indeed satisfies Eq. (2.14′), so we now have a simple solution for the lowering operators. In Eq. (2.27a′) we have for $m = n - 1$ used the definition [see Eq. (A2)]

$$\prod_{\mu=n}^{n-1} \mathcal{E}_{m\mu} \equiv 1, \tag{2.28′}$$

and we from here on define in a similar way all products that have the same range for the dummy index.

One sees that L_n^m written in the form

$$L_n^m = \prod_{\mu=m+1}^{n-1} \mathcal{E}_{m\mu} \sum_{p=0}^{n-m-1} \sum_{\mu_p > \mu_{p-1} > \cdots > \mu_2 > \mu_1 = m+1}^{n-1} \left(\prod_{i=1}^{p} \mathcal{E}_{m\mu_i}^{-1} \right) \times \mathcal{C}_{n}^{\mu_p} \mathcal{C}_{\mu_p}^{\mu_{p-1}} \cdots \mathcal{C}_{\mu_2}^{\mu_1} \mathcal{C}_{\mu_1}^{m}, \qquad 1 \leq m < n, \tag{2.27b′}$$

satisfies Eq. (2.11′). It can also be proven, in a way similar to the one which was used in the appendix for L_n^m given by Eq. (2.27a′), that L_n^m given by Eq. (2.27b′) fulfills Eq. (2.14′) and thus also is a lowering operator. In Sec. III we show that this operator is, in fact, equal to the one given by Eq. (2.27a′), which also is the reason for not distinguishing typografically between them. Equation (2.27b′) can now be considered as a useful alternative way of writing our lowering operator.

Note that the appearance of the inverse of the operators $\mathcal{E}_{\mu\mu'}$ in Eqs. (2.27a′) and (2.27b′) is just a matter of notation, as each factor in $\prod_{i=1}^{p} \mathcal{E}_{m\mu_i}^{-1}$ has a counterpart in $\prod_{\mu=m+1}^{n-1} \mathcal{E}_{m\mu}$ that cancels it.

The derivation of the raising operators follows the same lines as that of the lowering operators and one obtains

$$R_m^n = \Bigg(\sum_{p=0}^{m-1} \sum_{\mu_p > \mu_{p-1} > \cdots > \mu_2 > \mu_1 = 1}^{m-1} \mathcal{C}_m^{\mu_p} \mathcal{C}_{\mu_p}^{\mu_{p-1}} \cdots \mathcal{C}_{\mu_2}^{\mu_1} \mathcal{C}_{\mu_1}^{n}$$

$$\times \prod_{i=1}^{p} \mathcal{E}_{m\mu_i}^{-1} \Bigg) \prod_{\mu=1}^{m-1} \mathcal{E}_{m\mu}, \qquad 1 \leq m < n, \qquad (2.27a'')$$

or

$$R_m^n = \prod_{\mu=1}^{m-1} \mathcal{E}_{m\mu} \sum_{p=0}^{m-1} \sum_{\mu_p > \mu_{p-1} > \cdots > \mu_2 > \mu_1 = 1}^{m-1} \Bigg(\prod_{i=1}^{p} \mathcal{E}_{m\mu_i}^{-1} \Bigg)$$

$$\times \mathcal{C}_{\mu_1}^{n} \mathcal{C}_{\mu_2}^{\mu_1} \cdots \mathcal{C}_{\mu_p}^{\mu_{p-1}} \mathcal{C}_{m}^{\mu_p}, \qquad 1 \leq m < n, \qquad (2.27b'')$$

which are homogeneous of degree m in the operators $\mathcal{C}_{\mu'}^{\mu}$ and $\mathcal{E}_{m\mu}$, and where for $m = 1$

$$\prod_{\mu=1}^{0} \mathcal{E}_{m\mu} \equiv 1. \qquad (2.28'')$$

From here on we define in a similar way to (2.28'') all products that have the same range for the dummy index.

As both our lowering and raising operators are defined, according to Eq. (2.12), by their effect on the basis vectors of *highest* weight with respect to U_{n-1}, we cannot expect that the explicit expressions for them, (2.27') and (2.27''), should be symmetrical, e.g., apart from the factors $\mathcal{E}_{\mu\mu'}$, the L_n^m are expressed exclusively in terms of lowering generators, while the R_m^n are expressed in terms of lowering generators and a single raising generator $\mathcal{C}_{\mu_1}^{n}$. This lack of symmetry is also reflected by the fact that

$$L_n^{m\dagger} \not\propto R_m^n, \qquad (2.29)$$

which can be seen from Eqs. (2.27) by using Eqs. (2.1) and (2.21).

Special Cases

Using the general formulas (2.27) for the lowering and raising operators for $n = 2$, 3, and 4 we obtain the explicit expressions

U_2:

$$L_2^1 = \mathcal{C}_2^1, \qquad R_1^2 = \mathcal{C}_1^2.$$

U_3:

$$L_3^1 = \mathcal{C}_3^1 \mathcal{E}_{12} + \mathcal{C}_2^1 \mathcal{C}_3^2 = \mathcal{E}_{12} \mathcal{C}_3^1 + \mathcal{C}_3^2 \mathcal{C}_2^1,$$

$$L_3^2 = \mathcal{C}_3^2;$$

$$R_1^3 = \mathcal{C}_1^3,$$

$$R_2^3 = \mathcal{C}_2^3 \mathcal{E}_{21} + \mathcal{C}_2^1 \mathcal{C}_1^3 = \mathcal{E}_{21} \mathcal{C}_2^3 + \mathcal{C}_1^3 \mathcal{C}_2^1.$$

U_4:

$$L_4^1 = \mathcal{C}_4^1 \mathcal{E}_{12} \mathcal{E}_{13} + \mathcal{C}_2^1 \mathcal{C}_4^2 \mathcal{E}_{13} + \mathcal{C}_3^1 \mathcal{C}_4^3 \mathcal{E}_{12} + \mathcal{C}_2^1 \mathcal{C}_3^2 \mathcal{C}_4^3$$

$$= \mathcal{E}_{12} \mathcal{E}_{13} \mathcal{C}_4^1 + \mathcal{E}_{13} \mathcal{C}_4^2 \mathcal{C}_2^1 + \mathcal{E}_{12} \mathcal{C}_4^3 \mathcal{C}_3^1 + \mathcal{C}_4^3 \mathcal{C}_3^2 \mathcal{C}_2^1,$$

$$L_4^2 = \mathcal{C}_4^2 \mathcal{E}_{23} + \mathcal{C}_3^2 \mathcal{C}_4^3 = \mathcal{E}_{23} \mathcal{C}_4^2 + \mathcal{C}_4^3 \mathcal{C}_3^2,$$

$$L_4^3 = \mathcal{C}_4^3;$$

$$R_1^4 = \mathcal{C}_1^4,$$

$$R_2^4 = \mathcal{C}_2^4 \mathcal{E}_{21} + \mathcal{C}_2^1 \mathcal{C}_1^4 = \mathcal{E}_{21} \mathcal{C}_2^4 + \mathcal{C}_1^4 \mathcal{C}_2^1,$$

$$R_3^4 = \mathcal{C}_3^4 \mathcal{E}_{31} \mathcal{E}_{32} + \mathcal{C}_3^1 \mathcal{C}_1^4 \mathcal{E}_{32} + \mathcal{C}_3^2 \mathcal{C}_2^4 \mathcal{E}_{31} + \mathcal{C}_3^2 \mathcal{C}_2^1 \mathcal{C}_1^4$$

$$= \mathcal{E}_{31} \mathcal{E}_{32} \mathcal{C}_3^4 + \mathcal{E}_{32} \mathcal{C}_1^4 \mathcal{C}_3^1 + \mathcal{E}_{31} \mathcal{C}_2^4 \mathcal{C}_3^2 + \mathcal{C}_1^4 \mathcal{C}_2^1 \mathcal{C}_3^2.$$

III. RECURSION RELATIONS

We develop here some useful recursion relations for the lowering and raising operators. Performing the summation over μ_p in the formulas (2.27') for L_n^m in two steps, first summing up to $n' - 1$ where $m \leq n' < n$ and then summing from n' up to $n - 1$, one readily obtains from Eq. (2.27a')

$$L_n^m = \sum_{\mu=n'}^{n-1} L_\mu^m \mathcal{C}_n^\mu \prod_{\lambda=\mu+1}^{n-1} \mathcal{E}_{m\lambda} + [\mathcal{C}_n^{n'}, L_{n'}^m] \prod_{\lambda=n'}^{n-1} \mathcal{E}_{m\lambda},$$

$$1 \leq m \leq n' < n, \qquad (3.1a')$$

and from Eq. (2.27b')

$$L_n^m = \sum_{\mu=n'}^{n-1} \Bigg(\prod_{\lambda=\mu+1}^{n-1} \mathcal{E}_{m\lambda} \Bigg) \mathcal{C}_n^\mu \mathcal{C}_\mu^m + \prod_{\lambda=n'}^{n-1} \mathcal{E}_{m\lambda} [\mathcal{C}_n^{n'}, L_{n'}^m],$$

$$1 \leq m \leq n' < n, \qquad (3.1b')$$

where in both cases

$$L_m^m \equiv 1. \qquad (3.2')$$

Equations (3.1') are seen to be recursion relations expressing the lowering operators of U_n in terms of those of unitary subgroups of U_n. From Eqs. (3.1a') and (3.1b') one obtains in both cases for $n' = n - 1$

$$L_n^m = \mathcal{C}_n^{n-1} L_{n-1}^m \mathcal{E}_{m\,n-1} - \mathcal{E}_{m\,n-1} L_{n-1}^m \mathcal{C}_n^{n-1},$$

$$1 \leq m < n - 1. \qquad (3.3')$$

Since from both Eq. (2.27a') and (2.27b') we get the same expression

$$L_{m+1}^m = \mathcal{C}_{m+1}^m,$$

it follows from Eq. (3.3') by induction that L_n^m given by Eq. (2.27a') is equal to L_n^m given by Eq. (2.27b'), which proves our statement in Sec. II3. As another special case of Eqs. (3.1') one obtains for $n' = m$

$$L_n^m = \sum_{\mu=m}^{n-1} L_\mu^m \mathcal{C}_n^\mu \prod_{\lambda=\mu+1}^{n-1} \mathcal{E}_{m\lambda}, \qquad 1 \leq m < n, \qquad (3.4a')$$

and

$$L_n^m = \sum_{\mu=m}^{n-1} \Bigg(\prod_{\lambda=\mu+1}^{n-1} \mathcal{E}_{m\lambda} \Bigg) \mathcal{C}_n^\mu L_\mu^m, \quad 1 \leq m < n. \qquad (3.4b')$$

From Eqs. (2.27″) for R^n_m one obtains in a similar way the recursion relations

$$R^n_m = \sum_{\mu=1}^{n'} R^\mu_m \mathcal{C}^n_\mu \prod_{\lambda=1}^{\mu-1} \mathcal{E}_{m\lambda} + [R^{n'}_m, \mathcal{C}^n_{n'}] \prod_{\lambda=1}^{n'} \mathcal{E}_{m\lambda}, \quad 1 \le n' \le m, \tag{3.1a″}$$

or

$$R^n_m = \sum_{\mu=1}^{n'} \left(\prod_{\lambda=1}^{\mu-1} \mathcal{E}_{m\lambda} \right) \mathcal{C}^n_\mu R^\mu_m + \prod_{\lambda=1}^{n'} \mathcal{E}_{m\lambda} [R^{n'}_m, \mathcal{C}^n_{n'}], \quad 1 \le n' \le m, \tag{3.1b″}$$

where

$$R^{n'}_m = \left(\sum_{p=0}^{m-1} \sum_{\mu_p > \mu_{p-1} > \cdots > \mu_2 > \mu_1 = n'+1}^{m-1} \mathcal{C}^{\mu_p}_m \mathcal{C}^{\mu_{p-1}}_{\mu_p} \cdots \mathcal{C}^{\mu_1}_{\mu_2} \mathcal{C}^{n'}_{\mu_1} \times \prod_{i=1}^{p} \mathcal{E}^{-1}_{m\mu_i} \right) \prod_{\mu=n'+1}^{m-1} \mathcal{E}_{m\mu}, \quad 1 \le n' < m \tag{3.5a}$$

or

$$R^{n'}_m = \prod_{\mu=n'+1}^{m-1} \mathcal{E}_{m\mu} \sum_{p=0}^{m-1} \sum_{\mu_p > \mu_{p-1} > \cdots > \mu_2 > \mu_1 = n'+1}^{m-1} \left(\prod_{i=1}^{p} \mathcal{E}^{-1}_{m\mu_i} \right) \times \mathcal{C}^{\mu_p}_m \mathcal{C}^{\mu_{p-1}}_{\mu_p} \cdots \mathcal{C}^{\mu_1}_{\mu_2} \mathcal{C}^{n'}_{\mu_1}, \quad 1 \le n' < m, \tag{3.5b}$$

and

$$R^m_m \equiv 1. \tag{3.2″}$$

From Eqs. (3.1″) one obtains for $n' = 1$

$$R^n_m = \mathcal{E}_{m1} R^1_m \mathcal{C}^n_1 - \mathcal{C}^n_1 R^1_m \mathcal{E}_{m1}, \quad 1 < m < n, \tag{3.3″}$$

and for $n' = m$

$$R^n_m = \sum_{\mu=1}^{m} R^\mu_m \mathcal{C}^n_\mu \prod_{\lambda=1}^{\mu-1} \mathcal{E}_{m\lambda}, \quad 1 \le m < n, \tag{3.4a″}$$

and

$$R^n_m = \sum_{\mu=1}^{m} \left(\prod_{\lambda=1}^{\mu-1} \mathcal{E}_{m\lambda} \right) \mathcal{C}^n_\mu R^\mu_m, \quad 1 \le m < n. \tag{3.4b″}$$

Notice that L^m_μ $m < \mu < n$, which appears in the recursion relations for L^m_n, is a lowering operator of U_μ, while R^μ_m $1 \le \mu < m$ is not a raising operator of U_μ and in fact is, apart from the $\mathcal{E}_{\mu\mu'}$, expressed exclusively in terms of lowering generators. Yet we denote these operators by the same letter as that of the raising operators R^μ_m $m < \mu$ of U_μ as there is no possibility of confusion, since for the operators defined in Eqs. (3.5), $1 \le \mu < m$ while for the raising operators of U_μ, $m < \mu$. This again reflects the lack of symmetry between the lowering and the raising operators.

IV. MATRIX ELEMENTS OF THE COMMUTATORS OF THE LOWERING AND RAISING OPERATORS

In this section, we derive some properties of the lowering and raising operators which are of importance when we want to obtain a given normalized basis vector from another given normalized basis vector.

From Eq. (2.1) one sees that the lowering generators in Eq. (2.27a′) when acting on $\langle {h \atop q} |$ act as raising generators, and hence from Eq. (2.8) give zero if they belong to U_{n-1}; so we have, also using the definition (3.2′),

$$\left\langle {h \atop q} \right| L^m_{n'} \equiv \left(L^{m\dagger}_{n'} \left| {h \atop q} \right\rangle \right)^\dagger = \delta^m_{n'} \left\langle {h \atop q} \right|, \quad 1 \le m \le n' < n. \tag{4.1′}$$

Now using Eqs. (3.4′), (4.1′), (2.22), and (2.24) we obtain

$$\left\langle {h \atop q} \right| L^m_n = \left\langle {h \atop q} \right| \sum_{\mu=m}^{n-1} L^m_\mu \mathcal{C}^\mu_n \prod_{\lambda=\mu+1}^{n-1} \mathcal{E}_{m\lambda} = \left\langle {h \atop q} \right| \mathcal{C}^m_n \prod_{\lambda=m+1}^{n-1} \mathcal{E}_{m\lambda} = \prod_{\mu=m+1}^{n-1} (q_{m\mu} + 1) \left\langle {h \atop q} \right| \mathcal{C}^m_n, \quad 1 \le m < n. \tag{4.2′}$$

We may note that even though, by definition [see Eq. (2.12′)],

$$L^m_n \left| {h_1 \cdots \cdots h_n \atop q_1 \cdots q_m \cdots q_{n-1}} \right\rangle \propto \left| {h_1 \cdots \cdots \cdots h_n \atop q_1 \cdots q_m - 1 \cdots q_{n-1}} \right\rangle, \tag{4.3}$$

we have that

$$\left\langle {h_1 \cdots \cdots h_n \atop q_1 \cdots q_m \cdots q_{n-1}} \right| L^m_n \propto \left\langle {h_1 \cdots \cdots \cdots h_n \atop q_1 \cdots q_m + 1 \cdots q_{n-1}} \right|, \tag{4.4′}$$

the reason for this apparent paradox being that the $|{h \atop q}\rangle$ do not constitute a complete set of basis vectors so that $\langle {h \atop q}| \mathcal{C}^m_n$ contains basis vectors which are not of the form $\langle {h \atop q}|$. Equation (4.4′) also reflects Eq. (2.29).

Using Eq. (4.2′) twice we obtain

$$\left\langle {h \atop q} \right| L^m_n L^{m'}_n = \prod_{\mu=m+1}^{n-1} (q_{m\mu} + 1) \left\langle {h \atop q} \right| \mathcal{C}^m_n L^{m'}_n = \prod_{\mu=m+1}^{n-1} (q_{m\mu} + 1) \left\langle {h \atop q} \right| L^{m'}_n \mathcal{C}^m_n = \prod_{\mu=m+1}^{n-1} (q_{m\mu} + 1) \times \prod_{\mu=m'+1}^{n-1} (q_{m'\mu} + 1) \left\langle {h \atop q} \right| \mathcal{C}^{m'}_n \mathcal{C}^m_n, \quad 1 \le m < m' < n. \tag{4.5a}$$

In a similar but slightly more complicated way, using Eq. (4.2′) and the explicit formula (2.27a′) for L^m_n, one also obtains

$$\left\langle {h \atop q} \right| L^{m'}_n L^m_n = \prod_{\mu=m+1}^{n-1} (q_{m\mu} + 1) \prod_{\mu=m'+1}^{n-1} (q_{m'\mu} + 1) \times \left\langle {h \atop q} \right| \mathcal{C}^m_n \mathcal{C}^{m'}_n, \quad 1 \le m < m' < n. \tag{4.5b}$$

We now obtain from Eqs. (4.5a) and (4.5b), using the abbreviated notation

$$|q_m\ q_{m'}\rangle \equiv \left|\begin{matrix} h_1 & \cdots & \cdots & \cdots & h_n \\ q_1 & \cdots & q_m & \cdots & q_{m'} \cdots q_{n-1}\end{matrix}\right\rangle, \qquad 1 \le m < m' < n,$$

that

$$\langle q_m - 1\ q_{m'} - 1|\ [L_n^m, L_n^{m'}]\ |q_m\ q_{m'}\rangle = 0. \tag{4.6'}$$

For the operators defined in Eq. (3.5) we have correspondingly

$$\left\langle \begin{matrix} h \\ q \end{matrix}\right| R_m^{n'} = \delta_m^{n'} \left\langle \begin{matrix} h \\ q \end{matrix}\right|, \qquad 0 < n' \le m < n, \tag{4.1''}$$

from which follows for the raising operators

$$\left\langle \begin{matrix} h \\ q \end{matrix}\right| R_m^n = \prod_{\mu=1}^{m-1} (q_{m\mu} - 1) \left\langle \begin{matrix} h \\ q \end{matrix}\right| \mathfrak{C}_m^n, \tag{4.2''}$$

$$\left\langle \begin{matrix} h_1 & \cdots & \cdots & h_n \\ q_1 & \cdots & q_m & \cdots q_{n-1}\end{matrix}\right| R_m^n \propto \left\langle \begin{matrix} h_1 & \cdots & \cdots & \cdots & h_n \\ q_1 & \cdots & q_m - 1 & \cdots & q_{n-1}\end{matrix}\right|, \tag{4.4''}$$

and

$$\langle q_m + 1\ q_{m'} + 1|\ [R_m^n, R_{m'}^n]\ |q_m\ q_{m'}\rangle = 0, \qquad 1 \le m < m' < n. \tag{4.6''}$$

Let us now consider the matrix elements of the commutator of a lowering with a raising operator. Using Eqs. (4.2''), (3.4a'), and (4.1'), we obtain

$$\left\langle \begin{matrix} h \\ q \end{matrix}\right| R_m^n L_n^{m'} = \prod_{\mu=1}^{m-1} (q_{m\mu} - 1) \left\langle \begin{matrix} h \\ q \end{matrix}\right| \mathfrak{C}_m^n \sum_{\mu=m'}^{n-1} L_\mu^{m'} \mathfrak{C}_n^\mu \prod_{\lambda=\mu+1}^{n-1} \varepsilon_{m'\lambda}$$

$$= \prod_{\mu=1}^{m-1} (q_{m\mu} - 1) \prod_{\mu=m'+1}^{n-1} (q_{m'\mu} + 1) \left\langle \begin{matrix} h \\ q \end{matrix}\right| \mathfrak{C}_m^n \mathfrak{C}_n^{m'}, \qquad 1 \le m < m' < n, \tag{4.7'}$$

and, using Eqs. (4.2'), (3.4a''), and (4.1''), we obtain similarly

$$\left\langle \begin{matrix} h \\ q \end{matrix}\right| L_n^{m'} R_m^n = \prod_{\mu=1}^{m-1} (q_{m\mu} - 1) \prod_{\mu=m'+1}^{n-1} (q_{m'\mu} + 1) \times \left\langle \begin{matrix} h \\ q \end{matrix}\right| \mathfrak{C}_n^{m'} \mathfrak{C}_m^n, \qquad 1 \le m < m' < n. \tag{4.7''}$$

We now obtain from Eqs. (4.7'), (4.7''), (2.2), and (2.8),

$$\langle q_m + 1\ q_{m'} - 1|\ [R_m^n, L_n^{m'}]\ |q_m\ q_{m'}\rangle$$

$$= \prod_{\mu=1}^{m-1} q_{m\mu} \prod_{\mu=m'+1}^{n-1} q_{m'\mu} \times \langle q_m + 1\ q_{m'} - 1|\ [\mathfrak{C}_m^n, \mathfrak{C}_n^{m'}]\ |q_m\ q_{m'}\rangle$$

$$= \prod_{\mu=1}^{m-1} q_{m\mu} \prod_{\mu=m'+1}^{n-1} q_{m'\mu} \times \langle q_m + 1\ q_{m'} - 1|\ \mathfrak{C}_m^{m'}\ |q_m\ q_{m'}\rangle = 0, \qquad 1 \le m < m' < n. \tag{4.8a}$$

In a similar but slightly more complicated way, using the explicit formulas (2.27) for L_n^m and $R_{m'}^n$, one obtains

$$\langle q_m - 1\ q_{m'} + 1|\ [R_{m'}^n, L_n^m]\ |q_m\ q_{m'}\rangle = 0, \qquad 1 \le m < m' < n. \tag{4.8b}$$

Since from Eq. (4.3)

$$[L_n^m, L_n^{m'}]\ |q_m\ q_{m'}\rangle \propto |q_m - 1\ q_{m'} - 1\rangle, \qquad 1 \le m < m' < n,$$

it follows from Eq. (4.6') that

$$[L_n^m, L_n^{m'}] \left|\begin{matrix} h \\ q \end{matrix}\right\rangle = 0, \qquad 1 \le m < m' < n. \tag{4.9'}$$

In a similar way one obtains from Eqs. (4.6''), (4.8a), and (4.8b)

$$[R_m^n, R_{m'}^n] \left|\begin{matrix} h \\ q \end{matrix}\right\rangle = 0, \quad 1 \le m < m' < n, \tag{4.9''}$$

$$[R_m^n, L_n^{m'}] \left|\begin{matrix} h \\ q \end{matrix}\right\rangle = 0, \quad 1 \le m < m' < n, \tag{4.10a}$$

and

$$[R_{m'}^n, L_n^m] \left|\begin{matrix} h \\ q \end{matrix}\right\rangle = 0, \quad 1 \le m < m' < n. \tag{4.10b}$$

In Sec. V we show however that, in general [see (5.15)],

$$[R_m^n, L_n^m] \left|\begin{matrix} h \\ q \end{matrix}\right\rangle \neq 0, \qquad 1 \le m < n. \tag{4.11}$$

It might be mentioned that a stronger property than Eqs. (4.9) seems to hold, namely

$$[L_n^m, L_n^{m'}] = [R_m^n, R_{m'}^n] = 0, \qquad 1 \le m < m' < n,$$

but a general proof of this has as yet not been found.

V. NORMALIZATION COEFFICIENTS

In the previous sections we were only interested in lowering and raising operators that would take us from one basis vector to another one, without requiring that the last basis vector should be normalized if the first one was normalized. In the applications it is of course very convenient to deal with normalized basis vectors and so we derive in this section the normalization coefficients of the lowering and raising operators when acting on $\left|\begin{matrix} h \\ q \end{matrix}\right\rangle$.

The normalization coefficients, and hence also the

normalized lowering operators, are functions of $h_1, \cdots, h_n$ and $q_1, \cdots, q_{n-1}$ and are denoted by

$$N\begin{pmatrix} q_1 \cdots q_m \cdots q_{n-1} \\ h_1 \cdots\cdots\cdots h_n \\ q_1 \cdots q_m - 1 \cdots q_{n-1} \end{pmatrix} \equiv N^{q_1 \cdots q_m \cdots q_{n-1}}_{q_1 \cdots q_m - 1 \cdots q_{n-1}} \equiv N^{q_m}_{q_m - 1}, \qquad 1 \leq m < n,$$

and

$$\mathfrak{L}\begin{pmatrix} q_1 \cdots q_m \cdots q_{n-1} \\ h_1 \cdots\cdots\cdots h_n \\ q_1 \cdots q_m - 1 \cdots q_{n-1} \end{pmatrix} \equiv \mathfrak{L}^{q_1 \cdots q_m \cdots q_{n-1}}_{q_1 \cdots q_m - 1 \cdots q_{n-1}} \equiv \mathfrak{L}^{q_m}_{q_m - 1}, \qquad 1 \leq m < n,$$

respectively, where we use the abbreviated notation when there is no possibility of confusion. The normalized lowering operators can then be written as

$$\mathfrak{L}^{q_m}_{q_m - 1} = (N^{q_m}_{q_m - 1})^{-1} L^m_n, \tag{5.1'}$$

and fulfill by definition

$$\left| \begin{matrix} h_1 \cdots\cdots\cdots h_n \\ q_1 \cdots q_m - 1 \cdots q_{n-1} \end{matrix} \right\rangle \equiv \mathfrak{L}^{q_m}_{q_m - 1} \left| \begin{matrix} h_1 \cdots\cdots\cdots h_n \\ q_1 \cdots q_m \cdots q_{n-1} \end{matrix} \right\rangle = (N^{q_m}_{q_m - 1})^{-1} L^m_n \left| \begin{matrix} h_1 \cdots\cdots\cdots h_n \\ q_1 \cdots q_m \cdots q_{n-1} \end{matrix} \right\rangle. \tag{5.2'}$$

From Eqs. (5.2') and (4.2') we then have for the normalization coefficients

$$N^{q_m}_{q_m - 1} = \left\langle \begin{matrix} h_1 \cdots\cdots\cdots h_n \\ q_1 \cdots q_m - 1 \cdots q_{n-1} \end{matrix} \right| L^m_n \left| \begin{matrix} h_1 \cdots\cdots\cdots h_n \\ q_1 \cdots q_m \cdots q_{n-1} \end{matrix} \right\rangle = \prod_{\mu=m+1}^{n-1} q_{m\mu} \left\langle \begin{matrix} h_1 \cdots\cdots\cdots h_n \\ q_1 \cdots q_m - 1 \cdots q_{n-1} \end{matrix} \right| \times \mathfrak{C}^m_n \left| \begin{matrix} h_1 \cdots\cdots\cdots h_n \\ q_1 \cdots q_m \cdots q_{n-1} \end{matrix} \right\rangle. \tag{5.3'}$$

For the normalization coefficients of the raising operators and the normalized raising operators we have similarly

$$N\begin{pmatrix} q_1 \cdots q_m \cdots q_{n-1} \\ h_1 \cdots\cdots\cdots h_n \\ q_1 \cdots q_m + 1 \cdots q_{n-1} \end{pmatrix} \equiv N^{q_1 \cdots q_m \cdots q_{n-1}}_{q_1 \cdots q_m + 1 \cdots q_{n-1}} \equiv N^{q_m}_{q_m + 1}, \qquad 1 \leq m < n,$$

and

$$\mathfrak{R}\begin{pmatrix} q_1 \cdots q_m \cdots q_{n-1} \\ h_1 \cdots\cdots\cdots h_n \\ q_1 \cdots q_m + 1 \cdots q_{n-1} \end{pmatrix} \equiv \mathfrak{R}^{q_1 \cdots q_m \cdots q_{n-1}}_{q_1 \cdots q_m + 1 \cdots q_{n-1}} \equiv \mathfrak{R}^{q_m}_{q_m + 1}, \qquad 1 \leq m < n.$$

For the normalized raising operators we then have

$$\mathfrak{R}^{q_m}_{q_m + 1} = (N^{q_m}_{q_m + 1})^{-1} \mathfrak{R}^n_m, \tag{5.1''}$$

fulfilling

$$\left| \begin{matrix} h_1 \cdots\cdots\cdots h_n \\ q_1 \cdots q_m + 1 \cdots q_{n-1} \end{matrix} \right\rangle \equiv \mathfrak{R}^{q_m}_{q_m + 1} \left| \begin{matrix} h_1 \cdots\cdots\cdots h_n \\ q_1 \cdots q_m \cdots q_{n-1} \end{matrix} \right\rangle = (N^{q_m}_{q_m + 1})^{-1} R^n_m \left| \begin{matrix} h_1 \cdots\cdots\cdots h_n \\ q_1 \cdots q_m \cdots q_{n-1} \end{matrix} \right\rangle, \tag{5.2''}$$

and for the normalization coefficients we have

$$N^{q_m}_{q_m + 1} = \left\langle \begin{matrix} h_1 \cdots\cdots\cdots h_n \\ q_1 \cdots q_m + 1 \cdots q_{n-1} \end{matrix} \right| \mathfrak{R}^n_m \left| \begin{matrix} h_1 \cdots\cdots\cdots h_n \\ q_1 \cdots q_m \cdots q_{n-1} \end{matrix} \right\rangle = \prod_{\mu=1}^{m-1} q_{m\mu} \left\langle \begin{matrix} h_1 \cdots\cdots\cdots h_n \\ q_1 \cdots q_m + 1 \cdots q_{n-1} \end{matrix} \right| \times \mathfrak{C}^n_m \left| \begin{matrix} h_1 \cdots\cdots\cdots h_n \\ q_1 \cdots q_m \cdots q_{n-1} \end{matrix} \right\rangle. \tag{5.3''}$$

From Eqs. (5.3') and (5.3'') we obtain, using Eqs. (2.1), (2.25), and (2.26), the symmetry relation of the normalization coefficients of the raising operators with those of the lowering operators

$$N^{q_m}_{q_m + 1} = \left(\prod_{\mu=1}^{m-1} q_{m\mu} \Big/ \prod_{\mu=m+1}^{n-1} (q_{m\mu} + 1) \right) N^{q_m + 1}_{q_m} = (-)^{m-1} \left(\prod_{\mu=1}^{m-1} q_{\mu m} \Big/ \prod_{\mu=m+1}^{n-1} (q_{m\mu} + 1) \right) N^{q_m + 1}_{q_m}. \tag{5.4}$$

We are now able to obtain with the help of the operators L^m_n and R^n_m any normalized basis vector $|^h_{q'}\rangle$ from any given normalized basis vector $|^h_q\rangle$. Denoting the general normalization coefficients by

$$N\begin{pmatrix} q_1 \cdots q_{n-1} \\ h_1 \cdots h_n \\ q'_1 \cdots q'_{n-1} \end{pmatrix} \equiv N^{q_1 \cdots q_{n-1}}_{q_1' \cdots q_{n-1}'},$$

we have

$$\left| \begin{matrix} h_1 \cdots h_n \\ q'_1 \cdots q'_{n-1} \end{matrix} \right\rangle = (N^{q_1 \cdots q_{n-1}}_{q_1' \cdots q_{n-1}'})^{-1} (L^1_n)^{q_1 - q_1'} \cdots \times (R^n_{n-1})^{q_{n-1}' - q_{n-1}} \left| \begin{matrix} h_1 \cdots h_n \\ q_1 \cdots q_{n-1} \end{matrix} \right\rangle, \tag{5.5}$$

where we have considered an example with $q'_1 < q_1, \cdots, q'_{n-1} > q_{n-1}$. In general, when $q'_m \leq q_m$,

one has $(L_n^m)^{q_m - q_m'}$ and when $q_m' \geq q_m$ one has $(R_m^n)^{q_m' - q_m - q_m}$ in the product of the operators acting on $|{}^{h}_{q}\rangle$ in Eq. (5.5).

Due to Eqs. (4.9) and (4.10) the general normalization coefficients in Eq. (5.5) is independent of the order of the lowering and raising operators, and so if we consider some value q_m'', where due to Eq. (4.11) either $q_m' \leq q_m'' \leq q_m$ or $q_m' \geq q_m'' \geq q_m$, $1 \leq m < n$, we readily obtain for the normalization coefficients

$$N^{q_1 \cdots q_{n-1}}_{q_1' \cdots q_{n-1}'} = N^{q_1 \cdots q_{n-1}}_{q_1'' \cdots q_{n-1}''} N^{q_1'' \cdots q_{n-1}''}_{q_1' \cdots q_{n-1}'}. \tag{5.6}$$

From Eq. (5.5) follows specially

$$\begin{vmatrix} h_1 & \cdots & h_n \\ q_1 & \cdots & q_{n-1} \end{vmatrix} \rangle = (N^{h_1 \cdots h_n}_{q_1 \cdots q_{n-1}})^{-1} \times \prod_{\mu=1}^{n-1} (L_n^\mu)^{h_\mu - q_\mu} \begin{vmatrix} h_1 & \cdots & h_n \\ h_1 & \cdots & h_{n-1} \end{vmatrix} \rangle, \tag{5.7}$$

so we can now obtain any given normalized basis vector $|h_{\mu\nu}\rangle$ of an irreducible vector space $[h_{\mu n}]$ of U_n from its normalized basis vector of highest weight

$$\begin{vmatrix} h_{1n} & \cdots & h_{nn} \\ h_{1n} & \cdots & h_{n-1n} \end{vmatrix} \rangle$$

as

$$\begin{vmatrix} h_{1n} & \cdots & h_{nn} \\ & h_{1\,n-1} \cdots h_{n-1\,n-1} & \\ & \cdots & \\ & h_{12}\ h_{22} & \\ & h_{11} & \end{vmatrix} \rangle$$

$$= \left[\prod_{\nu=2}^{n} N \begin{pmatrix} h_{1\nu} & \cdots & h_{\nu-1\,\nu} \\ h_{1\nu} & \cdots & h_{\nu\nu} \\ h_{1\nu-1} & \cdots & h_{\nu-1\,\nu-1} \end{pmatrix} \right]^{-1} (L_2^1)^{h_{12} - h_{11}} \times \prod_{\mu=1}^{2} (L_3^\mu)^{h_{\mu 2} - h_{\mu 3}} \cdots \prod_{\mu=1}^{n-2} (L_{n-1}^\mu)^{h_{\mu n-1} - h_{\mu n}} \times \prod_{\mu=1}^{n-1} (L_n^\mu)^{h_{\mu n} - h_{\mu n-1}} \begin{vmatrix} h_{1n} & \cdots & h_{nn} \\ & h_{1n} \cdots h_{n-1n} & \\ & \cdots & \\ & h_{1n}\ h_{2n} & \\ & h_{1n} & \end{vmatrix} \rangle, \tag{5.8}$$

where it should be noted that the lowering operators belonging to distinct subgroups of U_n are to be maintained in the order given above.

The explicit formulas for the normalization coefficients can be obtained in two different ways. A self-contained, though somewhat lengthy, derivation has been obtained by the authors,[9] but is not given in this paper. We give here a simpler, though not self-contained, derivation using the formulas for the matrix elements of the generators of U_n obtained by Gelfand and Zetlin[8] and later rederived by other authors.[7,9]

With the phase convention

$$\langle h_{\mu\nu}' | \mathfrak{C}_n^{n-1} | h_{\mu\nu} \rangle \geq 0, \tag{5.9}$$

the matrix elements of $\mathfrak{C}_n^m$ are given by[8,7,9]

$$\left\langle \begin{vmatrix} h_{1n} & \cdots & h_{nn} \\ & h_{1n-1} \cdots h_{l_{n-1}n-1} - 1 \cdots h_{n-1\,n-1} & \\ & \cdots & \\ & h_{1m} \cdots h_{l_{mm}} - 1 \cdots h_{mm} & \\ & h_{1m-1} \cdots h_{m-1\,m-1} & \\ & \cdots & \\ & h_{11} & \end{vmatrix} \right.$$

$$\left. \times \mathfrak{C}_n^m \begin{vmatrix} h_{1n} & \cdots & h_{nn} \\ & h_{1n-1} \cdots h_{l_{n-1}n-1} \cdots h_{n-1\,n-1} & \\ & \cdots & \\ & h_{1m} \cdots h_{l_{mm}} \cdots h_{mm} & \\ & h_{1m-1} \cdots h_{m-1\,m-1} & \\ & \cdots & \\ & h_{11} & \end{vmatrix} \right\rangle$$

$$= \prod_{\lambda=m+1}^{n-1} S(l_{\lambda-1} - l_\lambda)[h_{l_\lambda \lambda, l_{\lambda-1}\lambda-1}(h_{l_\lambda\lambda, l_{\lambda-1}\lambda-1} + 1)]^{-\frac{1}{2}} \times \prod_{\lambda=m+1}^{n} \left\{ - \frac{\prod_{\kappa=1}^{\lambda-2} h_{l_{\lambda-1}\lambda-1, \kappa\lambda-2}}{\prod_{\substack{\kappa=1 \\ \kappa \neq l_{\lambda-1}}}^{\lambda-1} h_{l_{\lambda-1}\lambda-1, \kappa\lambda-1}} \right.$$

$$\left. \times \frac{\prod_{\kappa=1}^{\lambda} (h_{l_{\lambda-1}\lambda-1, \kappa\lambda} - 1)}{\prod_{\substack{\kappa=1 \\ \kappa \neq l_{\lambda-1}}}^{\lambda-1} (h_{l_{\lambda-1}\lambda-1, \kappa\lambda-1} - 1)} \right\}^{\frac{1}{2}}, \quad 1 \leq l_\lambda \leq \lambda, \tag{5.10}$$

where

$$h_{\mu\nu, \mu'\nu'} \equiv h_{\mu\nu} - h_{\mu'\nu'} + \mu' - \mu,$$

and

$$S(x) \equiv \begin{cases} 1 \\ -1 \end{cases} \text{ for } x \gtrless 0.$$

From Eqs. (5.3′) and (5.10) one then obtains for normalization coefficients of the lowering operators

[9] J. G. Nagel and M. Moshinsky, Rev. Mexicana Fis. (to be published).

$$N\begin{bmatrix} q_1 \cdots q_m \cdots q_{n-1} \\ h_1 \cdots\cdots\cdots\cdots h_n \\ q_1 \cdots q_m - 1 \cdots q_{n-1} \end{bmatrix}$$

$$= \Big[\Big(\prod_{\mu=m+1}^{n-1} q_{m\mu} \Big/ \prod_{\mu=1}^{m-1} (q_{\mu m}+1)\Big) \times \prod_{\mu=1}^{m} (h_\mu - q_m + m - \mu + 1) \times \prod_{\mu=m+1}^{n} (q_m - h_\mu + \mu - m - 1)\Big]^{\frac{1}{2}}$$

$$= \Big[-\Big(\prod_{\mu=m+1}^{n-1} q_{m\mu} \Big/ \prod_{\mu=1}^{m-1} (q_{m\mu}-1)\Big) \times \prod_{\mu=1}^{n} (q_m - h_\mu + \mu - m - 1)\Big]^{\frac{1}{2}}. \quad (5.11')$$

From Eq. (5.11′) one now sees that the phase convention for the normalization coefficients of the lowering operators equivalent to the phase convention (5.9) is

$$N\begin{bmatrix} q_1 \cdots q_m \cdots q_{n-1} \\ h_1 \cdots\cdots\cdots\cdots h_n \\ q_1 \cdots q_m - 1 \cdots q_{n-1} \end{bmatrix} > 0. \quad (5.12)$$

From the symmetry relation (5.4) one sees however, that this phase convention does not imply that the normalization coefficients of the raising operators are also all positive.

Using Eq. (5.6) successively and Eq. (5.11′) one now obtains the general normalization coefficients of the lowering operators as

$$N\begin{bmatrix} q_1 \cdots q_{n-1} \\ h_1 \cdots h_n \\ q_1' \cdots q_{n-1}' \end{bmatrix} = \prod_{m=1}^{n-1} N\begin{bmatrix} q_1' \cdots q_{m-1}' q_m q_{m+1} \cdots q_{n-1} \\ h_1 \cdots\cdots\cdots\cdots\cdots h_n \\ q_1' \cdots q_{m-1}' q_m' q_{m+1} \cdots q_{n-1} \end{bmatrix}$$

$$= \prod_{m=1}^{n-1} \prod_{q_m''=q_m'+1}^{q_m} N\begin{bmatrix} q_1' \cdots q_{m-1}' \; q_m'' \; q_{m+1} \cdots q_{n-1} \\ h_1 \cdots\cdots\cdots\cdots\cdots\cdots h_n \\ q_1' \cdots q_{m-1}' q_m'' - 1 q_{m+1} \cdots q_{n-1} \end{bmatrix}$$

$$= \Big[\prod_{\mu>\lambda=1}^{n-1} \frac{(q_\lambda - q_\mu + \mu - \lambda)!}{(q_\lambda' - q_\mu' + \mu - \lambda)!} \prod_{\mu\geq\lambda=1}^{n-1} \frac{(h_\lambda - q_\mu' + \mu - \lambda)!}{(h_\lambda - q_\mu + \mu - \lambda)!} \times \prod_{\mu>\lambda=1}^{n} \frac{(q_\lambda - h_\mu + \mu - \lambda - 1)!}{(q_\lambda' - h_\mu + \mu - \lambda - 1)!}\Big]^{\frac{1}{2}}, \quad q_\mu' \leq q_\mu, \quad (5.13')$$

from which one, as a special case, obtains the normalization coefficients appearing in Eq. (5.7),

$$N\begin{bmatrix} h_1 \cdots h_{n-1} \\ h_1 \cdots h_n \\ q_1 \cdots q_{n-1} \end{bmatrix} = \Big[\prod_{\mu\geq\lambda=1}^{n-1} \frac{(h_\lambda - q_r + \mu - \lambda)!}{(q_\lambda - q_\mu + \mu - \lambda)!} \times \prod_{\mu>\lambda=1}^{n} \frac{(h_\lambda - h_\mu + \mu - \lambda - 1)!}{(q_\lambda - h_\mu + \mu - \lambda - 1)!}\Big]^{\frac{1}{2}}. \quad (5.14)$$

For the normalization coefficients of the raising operators one obtains from the symmetry relation (5.4) and Eq. (5.11′)

$$N\begin{bmatrix} q_1 \cdots q_m \cdots q_{n-1} \\ h_1 \cdots\cdots\cdots\cdots h_n \\ q_1 \cdots q_m + 1 \cdots q_{n-1} \end{bmatrix}$$

$$= (-)^{m-1}\Big[\Big(\prod_{\mu=1}^{m-1} q_{\mu m} \Big/ \prod_{\mu=m+1}^{n-1} (q_{m\mu}+1)\Big) \times \prod_{\mu=1}^{m} (h_\mu - q_m + m - \mu) \times \prod_{\mu=m+1}^{n} (q_m - h_\mu + \mu - m)\Big]^{\frac{1}{2}}$$

$$= (-)^{m-1}\Big[-\Big(\prod_{\mu=1}^{m-1} q_{m\mu} \Big/ \prod_{\mu=m+1}^{n-1} (q_{m\mu}+1)\Big) \times \prod_{\mu=1}^{n} (q_m - h_\mu + \mu - m)\Big]^{\frac{1}{2}}, \quad (5.11'')$$

from which one then obtains in the same way as for the lowering operators

$$N\begin{bmatrix} q_1 \cdots q_{n-1} \\ h_1 \cdots h_n \\ q_1' \cdots q_{n-1}' \end{bmatrix} = (-)^{\sum_{\mu=1}^{n-1} (\mu-1)(q_\mu' - q_\mu)}$$

$$\times \Big[\prod_{\mu>\lambda=1}^{n-1} \frac{(q_\lambda - q_\mu + \mu - \lambda)!}{(q_\lambda' - q_\mu' + \mu - \lambda)!} \prod_{\mu\geq\lambda=1}^{n-1} \frac{(h_\lambda - q_\mu + \mu - \lambda)!}{(h_\lambda - q_\mu' + \mu - \lambda)!} \times \prod_{\mu>\lambda=1}^{n} \frac{(q_\lambda' - h_\mu + \mu - \lambda - 1)!}{(q_\lambda - h_\mu + \mu - \lambda - 1)!}\Big]^{\frac{1}{2}}, \quad q_\mu' \geq q_\mu. \quad (5.13'')$$

From Eqs. (5.11′) and (5.11″) one obtains

$$N_{q_m}^{q_{m-1}} N_{q_{m-1}}^{q_m} = (-)^{m-1} \prod_{\mu=1}^{m} (h_\mu - q_m + m - \mu + 1) \times \prod_{\mu=m+1}^{n} (q_m - h_\mu + \mu - m - 1)$$

$$= -\prod_{\mu=1}^{n} (q_m - h_\mu + \mu - m - 1).$$

It then follows that

$$N^{q_m+1}_{q_m} N^{q_m}_{q_{m+1}} = N^{q_m}_{q_{m+1}} N^{q_m+1}_{q_m}$$
$$= (-)^{m-1} \prod_{\mu=1}^{m} (h_\mu - q_m + m - \mu)$$
$$\times \prod_{\mu=m+1}^{n} (q_m - h_\mu + \mu - m)$$
$$= -\prod_{\mu=1}^{n} (q_m - h_\mu + \mu - m),$$

so one sees that in general

$$N^{q_m-1}_{q_m} N^{q_m}_{q_{m-1}} \neq N^{q_m+1}_{q_m} N^{q_m}_{q_{m+1}}, \qquad (5.15)$$

which proves Eq. (4.11).

Special Cases

As special cases for $n = 2$ and 3 one obtains from Eqs. (5.11) the normalized lowering and raising operators in Eqs. (5.1), and from Eq. (5.14) the normalized basis vector $|h_{\mu\nu}\rangle$ in Eq. (5.8),

U_2:

$$\mathfrak{L}\begin{pmatrix} & q_1 & \\ h_1 & & h_2 \\ & q_1 - 1 & \end{pmatrix} = [(h_1 - q_1 + 1)(q_1 - h_2)]^{-\frac{1}{2}} L_2^1,$$

$$\mathfrak{R}\begin{pmatrix} & q_1 & \\ h_1 & & h_2 \\ & q_1 + 1 & \end{pmatrix} = [(h_1 - q_1)(q_1 - h_2 + 1)]^{-\frac{1}{2}} R_1^2;$$

$$\left|\begin{matrix} h_1 & h_2 \\ & q_1 \end{matrix}\right\rangle = \left[\frac{(q_1 - h_2)!}{(h_1 - q_1)!\,(h_1 - h_2)!}\right]^{\frac{1}{2}} (L_2^1)^{h_1 - q_1} \left|\begin{matrix} h_1 & h_2 \\ & h_1 \end{matrix}\right\rangle.$$

U_3:

$$\mathfrak{L}\begin{pmatrix} & q_1\ q_2 & \\ h_1 & h_2 & h_3 \\ & q_1 - 1\ q_2 & \end{pmatrix} = [(q_1 - q_2 + 1)(h_1 - q_1 + 1)(q_1 - h_2)(q_1 - h_3 + 1)]^{-\frac{1}{2}} L_3^1,$$

$$\mathfrak{L}\begin{pmatrix} & q_1\ q_2 & \\ h_1 & h_2 & h_3 \\ & q_1\ q_1 - 1 & \end{pmatrix} = \left[\frac{(q_1 - q_2 + 2)}{(h_1 - q_2 + 2)(h_2 - q_2 + 1)(q_2 - h_3)}\right]^{\frac{1}{2}} L_3^2;$$

$$\mathfrak{R}\begin{pmatrix} & q_1\ q_2 & \\ h_1 & h_2 & h_3 \\ & q_1 + 1\ q_2 & \end{pmatrix} = \left[\frac{(q_1 - q_2 + 2)}{(h_1 - q_1)(q_1 - h_2 + 1)(q_1 - h_3 + 2)}\right]^{\frac{1}{2}} R_1^3,$$

$$\mathfrak{R}\begin{pmatrix} & q_1\ q_2 & \\ h_1 & h_2 & h_3 \\ & q_1\ q_2 + 1 & \end{pmatrix} = -[(q_1 - q_2 + 1)(h_1 - q_2 + 1)(h_2 - q_2)(q_2 - h_3 + 1)]^{-\frac{1}{2}} R_2^3;$$

$$\left|\begin{matrix} h_1 & h_2 & h_3 \\ & q_1\ q_2 & \\ & r_1 & \end{matrix}\right\rangle = \left[\frac{(r_1 - q_2)!\,(q_1 - q_2 + 1)(q_1 - h_2)!\,(q_1 - h_3 + 1)!\,(q_2 - h_3)!}{(q_1 - r_1)!\,(h_1 - q_1)!\,(h_1 - q_2 + 1)!\,(h_2 - q_2)!\,(h_1 - h_2)!\,(h_1 - h_3 + 1)!\,(h_2 - h_3)!}\right]^{\frac{1}{2}}$$
$$\times (L_2^1)^{q_1 - r_1} (L_3^1)^{h_1 - q_1} (L_3^2)^{h_2 - q_2} \left|\begin{matrix} h_1 & h_2 & h_3 \\ & h_1 & h_2 \\ & & h_1 \end{matrix}\right\rangle.$$

APPENDIX A: PROOF OF THE FORMULA FOR THE LOWERING OPERATORS

We prove here that L_n^m written in the form

$$L_n^m = \left(\sum_{p=0}^{n-m-1} \sum_{\mu_p > \mu_{p-1} > \cdots > \mu_2 > \mu_1 = m+1}^{n-1} \mathfrak{C}_{\mu_1}^m \mathfrak{C}_{\mu_2}^{\mu_1} \cdots \mathfrak{C}_{\mu_p}^{\mu_{p-1}} \mathfrak{C}_n^{\mu_p} \prod_{i=1}^{p} \mathfrak{E}_{m\mu_i}^{-1}\right) \prod_{\mu=m+1}^{n-1} \mathfrak{E}_{m\mu}, \qquad 1 \leq m < n, \qquad (A1)$$

where

$$\prod_{\mu=n}^{n-1} \mathfrak{E}_{m\mu} \equiv 1, \qquad (A2)$$

satisfies Eq. (2.14′)

$$[\mathfrak{C}_\lambda^{\lambda+1}, L_n^m] \left|\begin{matrix} h \\ q \end{matrix}\right\rangle = 0, \qquad 1 \leq \lambda < n - 1. \qquad (A3)$$

We shall treat the three cases $\lambda > m$, $\lambda = m$ and $\lambda < m$ separately.

$$m < \lambda < n - 1.$$

In calculating the commutator of $\mathcal{C}_\lambda^{\lambda+1}$ with L_n^m given by (A1) we consider first the commutations of $\mathcal{C}_\lambda^{\lambda+1}$ with the products of the $\mathcal{E}_{m\mu}$ and then with those terms in L_n^m, where the products $\mathcal{C}_{\mu_1}^m \mathcal{C}_{\mu_2}^{\mu_1} \cdots \mathcal{C}_{\mu_p}^{\mu_{p-1}} \mathcal{C}_n^{\mu_p}$ either contain $\lambda + 1$ or λ, or contain both λ and $\lambda + 1$, or do not contain neither λ nor $\lambda + 1$. Using the commutation relations (2.2) and (2.22) and Eqs. (2.8), (2.19), (2.20), and (2.18), one obtains

$$[\mathcal{C}_\lambda^{\lambda+1}, L_n^m] \left|\begin{matrix} h \\ q \end{matrix}\right\rangle = \sum_{p=0}^{n-m-1} \sum_{\mu_p > \cdots > \mu_1 = m+1}^{n-1} \mathcal{C}_{\mu_1}^m \cdots \mathcal{C}_n^{\mu_p} \left[\mathcal{C}_\lambda^{\lambda+1}, \prod_{i=1}^{p} \mathcal{E}_{m\mu_i}^{-1} \prod_{\kappa=m+1}^{n-1} \mathcal{E}_{m\kappa} \right] \left|\begin{matrix} h \\ q \end{matrix}\right\rangle$$

$$+ \sum_{i=0}^{\lambda-m-1} \sum_{p=i}^{n-\lambda-2+i} \sum_{\mu_p > \cdots > \mu_{i+1} = \lambda+2}^{n-1} \sum_{\mu_i > \cdots > \mu_1 = m+1}^{\lambda-1} \Bigg([\mathcal{C}_\lambda^{\lambda+1}, \mathcal{C}_{\mu_1}^m \cdots \mathcal{C}_{\lambda+1}^{\mu_i} \mathcal{C}_{\mu_{i+1}}^{\lambda+1} \cdots \mathcal{C}_n^{\mu_p}] \prod_{i=1}^{p} \mathcal{E}_{m\mu_i}^{-1} \prod_{\substack{\kappa=m+1 \\ \kappa \neq \lambda+1}}^{n-1} \mathcal{E}_{m\kappa}$$

$$+ [\mathcal{C}_\lambda^{\lambda+1}, \mathcal{C}_{\mu_1}^m \cdots \mathcal{C}_\lambda^{\mu_i} \mathcal{C}_{\mu_{i+1}}^{\lambda} \cdots \mathcal{C}_n^{\mu_p}] \prod_{i=1}^{p} \mathcal{E}_{m\mu_i}^{-1} \prod_{\substack{\kappa=m+1 \\ \kappa \neq \lambda}}^{n-1} \mathcal{E}_{m\kappa}$$

$$+ [\mathcal{C}_\lambda^{\lambda+1}, \mathcal{C}_{\mu_1}^m \cdots \mathcal{C}_\lambda^{\mu_i} \mathcal{C}_{\lambda+1}^{\lambda+1} \mathcal{C}_{\mu_{i+1}}^{\lambda+1} \cdots \mathcal{C}_n^{\mu_p}] \prod_{i=1}^{p} \mathcal{E}_{m\mu_i}^{-1} \prod_{\substack{\kappa=m+1 \\ \kappa \neq \lambda, \lambda+1}}^{n-1} \mathcal{E}_{m\kappa} \Bigg) \left|\begin{matrix} h \\ q \end{matrix}\right\rangle$$

$$= \Bigg\{ \sum_{i=0}^{\lambda-m-1} \sum_{p=i}^{n-\lambda-2+i} \sum_{\mu_p > \cdots > \mu_{i+1} = \lambda+2}^{n-1} \sum_{\mu_i > \cdots > \mu_1 = m+1}^{\lambda-1} (\mathcal{C}_{\mu_1}^m \cdots \mathcal{C}_\lambda^{\mu_i} \mathcal{C}_{\mu_{i+1}}^{\lambda+1} \cdots \mathcal{C}_n^{\mu_p} (\mathcal{E}_{m\lambda} - \mathcal{E}_{m\lambda+1})$$

$$+ \mathcal{C}_{\mu_1}^m \cdots \mathcal{C}_\lambda^{\mu_i} (\mathcal{C}_\lambda^\lambda - \mathcal{C}_{\lambda+1}^{\lambda+1}) \mathcal{C}_{\mu_{i+1}}^{\lambda+1} \cdots \mathcal{C}_n^{\mu_p}) \prod_{i=1}^{p} \mathcal{E}_{m\mu_i}^{-1} \Bigg\} \prod_{\substack{\kappa=m+1 \\ \kappa \neq \lambda, \lambda+1}}^{n-1} \mathcal{E}_{m\kappa} \left|\begin{matrix} h \\ q \end{matrix}\right\rangle$$

$$= \Bigg\{ \sum_{i=0}^{\lambda-m-1} \sum_{p=i}^{n-\lambda-2+i} \sum_{\mu_p > \cdots > \mu_{i+1} = \lambda+2}^{n-1} \sum_{\mu_i > \cdots > \mu_1 = m+1}^{\lambda-1} \mathcal{C}_{\mu_1}^m \cdots \mathcal{C}_\lambda^{\mu_i} \mathcal{C}_{\mu_{i+1}}^{\lambda+1} \cdots \mathcal{C}_n^{\mu_p} \prod_{i=1}^{p} \mathcal{E}_{m\mu_i}^{-1} \Bigg\}$$

$$\times \prod_{\substack{\kappa=m+1 \\ \kappa \neq \lambda, \lambda+1}}^{n-1} \mathcal{E}_{m\kappa} (\mathcal{E}_{\lambda+1\lambda} + \mathcal{C}_\lambda^\lambda - \mathcal{C}_{\lambda+1}^{\lambda+1} + 1) \left|\begin{matrix} h \\ q \end{matrix}\right\rangle = 0.$$

$$m = \lambda < n - 1$$

In this case one obtains in a similar way

$$[\mathcal{C}_m^{m+1}, L_n^m] \left|\begin{matrix} h \\ q \end{matrix}\right\rangle = \sum_{p=0}^{n-m-1} \sum_{\mu_p > \cdots > \mu_1 = m+1}^{n-1} \mathcal{C}_{\mu_1}^m \cdots \mathcal{C}_n^{\mu_p} \left[\mathcal{C}_m^{m+1}, \prod_{i=1}^{p} \mathcal{E}_{m\mu_i}^{-1} \prod_{\mu=m+1}^{n-1} \mathcal{E}_{m\mu} \right] \left|\begin{matrix} h \\ q \end{matrix}\right\rangle$$

$$+ \sum_{p=0}^{n-m-2} \sum_{\mu_p > \cdots > \mu_1 = m+2}^{n-1} \Bigg([\mathcal{C}_m^{m+1}, \mathcal{C}_{\mu_1}^m \cdots \mathcal{C}_n^{\mu_p}] \prod_{i=1}^{p} \mathcal{E}_{m\mu_i}^{-1} \prod_{\mu=m+1}^{n-1} \mathcal{E}_{m\mu}$$

$$+ [\mathcal{C}_m^{m+1}, \mathcal{C}_{m+1}^m \mathcal{C}_{\mu_1}^{m+1} \cdots \mathcal{C}_n^{\mu_p}] \prod_{i=1}^{p} \mathcal{E}_{m\mu_i}^{-1} \prod_{\mu=m+2}^{n-1} \mathcal{E}_{m\mu} \Bigg) \left|\begin{matrix} h \\ q \end{matrix}\right\rangle$$

$$= \Bigg\{ \sum_{p=0}^{n-m-2} \sum_{\mu_p > \cdots > \mu_1 = m+2}^{n-1} (-\mathcal{C}_{\mu_1}^{m+1} \cdots \mathcal{C}_n^{\mu_p} \mathcal{E}_{mm+1} + (\mathcal{C}_m^m - \mathcal{C}_{m+1}^{m+1}) \mathcal{C}_{\mu_1}^{m+1} \cdots \mathcal{C}_n^{\mu_p}) \prod_{i=1}^{p} \mathcal{E}_{m\mu_i}^{-1} \Bigg\} \prod_{\mu=m+2}^{n-1} \mathcal{E}_{m\mu} \left|\begin{matrix} h \\ q \end{matrix}\right\rangle$$

$$= \Bigg\{ \sum_{p=0}^{n-m-2} \sum_{\mu_p > \cdots > \mu_1 = m+2}^{n-1} \mathcal{C}_{\mu_1}^{m+1} \cdots \mathcal{C}_n^{\mu_p} \prod_{i=1}^{p} \mathcal{E}_{m\mu_i}^{-1} \Bigg\} \prod_{\mu=m+2}^{n-1} \mathcal{E}_{m\mu} (-\mathcal{E}_{mm+1} + \mathcal{C}_m^m - \mathcal{C}_{m+1}^{m+1} + 1) \left|\begin{matrix} h \\ q \end{matrix}\right\rangle = 0.$$

$$1 \leq \lambda < n$$

In this case we have

$$[\mathcal{C}_\lambda^{\lambda+1}, L_n^m] \left|\begin{matrix} h \\ q \end{matrix}\right\rangle = \sum_{p=0}^{n-m-1} \sum_{\mu_p > \cdots > \mu_1 = m+1}^{n-1} \mathcal{C}_{\mu_1}^m \cdots \mathcal{C}_n^{\mu_p} \left[\mathcal{C}_\lambda^{\lambda+1}, \prod_{i=1}^{p} \mathcal{E}_{m\mu_i}^{-1} \prod_{\mu=m+1}^{n-1} \mathcal{E}_{m\mu} \right] \left|\begin{matrix} h \\ q \end{matrix}\right\rangle = 0,$$

so we have now finally proven that L_n^m of the form (A1) satisfies Eq. (A3).

Symbolic Calculus of the Wiener Process and Wiener–Hermite Functionals*

TSUTOMU IMAMURA†
Boston University, Boston, Massachusetts
WILLIAM C. MEECHAM
University of Minnesota, Minneapolis, Minnesota
AND
ARMAND SIEGEL
Boston University, Boston, Massachusetts
(Received 13 December 1963)

A new definition is given for the "ideal random function" (derivative of the Wiener function), which separates out infinite factors by fullest exploitation of the possibilities of the Dirac delta function. By allowing all integrals to be written formally as sums, this facilitates the definition and manipulation of the Wiener–Hermite functionals, especially for vector random processes of multiple argument. Expansion of a random function in Wiener–Hermite functionals is discussed. An expression is derived for the expectation value of the product of any number of Wiener–Hermite functionals; this is all that is needed in principle to obtain full statistical information from the Wiener–Hermite functional expansion of a random function. The method is illustrated by the calculation of the first correction to the flatness factor (measure of Gaussianity) of a nearly-Gaussian random function.

I. INTRODUCTION

IN certain physical problems, one may wish to analyze a random function f in terms of its deviation from some Gaussian function. Toward this end, Cameron and Martin[1] and Wiener[2] have given two formulations of the theory of Wiener–Hermite functionals; when f is expanded in series of such functionals, the first term is Gaussian, and the higher terms are in a sense successively further and further removed from Gaussianity.

Such an expansion is distinct in principle from the well-known Gram–Charlier or Edgworth series. In the latter, the probability distribution function is expanded in a series whose first term is a Gaussian function; in the present case, the distributed random variable, i.e., the argument of the distribution function, is expanded instead. This latter procedure might be expected to be useful when the mathematical conditions which define the problem (e.g., a differential equation) apply to the distributed variable rather than to the distribution function. If an expansion of the random function is then used, the conditions expressed by the differential equation are transformed into conditions on the "coefficients" (actually integral kernels) of the functional expansion.

As an application of such an expression, one may have in mind a situation in which the argument R is of multiple nature, e.g., a vector; moreover, the function f may itself be a vector, which is equivalent to a scalar function with a discrete argument (for the vector index of f) incorporated into R. The present paper is in fact intended for such a purpose, since it arose from the necessity of furnishing a mathematical background for certain papers on turbulence theory[3] by the present authors. The formulas needed in this case are considerably more complicated in detail than those of the single argument developed in Refs. 1 and 2, even though similar in principle. Hence some simplifying scheme is desirable. In this paper we put forth first, by way of preparation, such a simplifying scheme: A new "symbolic calculus" of the ideal random function (the derivative of the Wiener function). With its aid we then define a simple set of Hermite functionals, and develop a method of evaluating expectation values of products of such functionals. The work is done in terms of vector, or multiple, argument R, except for an introductory discussion of the case of a single, scalar argument. We conclude with an illustrative problem, the computation of the "flatness factor" (deviation from Gaussianity) due to the second and third terms of an Hermite functional expansion.

* Work of T. I. (in full) and A. S. (in part) supported by U. S. Air Force Office of Scientific Research; of W. C. M. (in part) by the Office of Naval Research.
† Present address: Kwansei Gakuin University, Nishinomiya, Japan.
[1] R. H. Cameron and W. T. Martin, Ann. Math. 48, 385 (1947).
[2] N. Wiener, *Nonlinear Problems in Random Theory* (Tech. Press, Cambridge, Massachusetts and John Wiley & Sons, Inc., New York, 1958).
[3] W. C. Meecham and A. Siegel, Phys. Fluids 7, 1178 (1964); A. Siegel, T. Imamura, and W. C. Meecham, J. Math. Phys. 6, 707 (1964). A preliminary report by Siegel, Imamura, and Meecham appeared in Phys. Fluids 6, 1519 (1963).

II. DEFINITION OF THE IDEAL RANDOM FUNCTION

A. Single Argument

As the basis for constructing our series, we use the "ideal random function" $a(x)$ of a scalar variable x. This improper function may be defined in various ways: (a) as the derivative of Wiener's random function[4-6] $X(x)$, or (b) by the moment equations

$$\langle a(x)\rangle = 0, \tag{2.1}$$

$$\langle a(x_1)a(x_2)\rangle = \delta(x_1 - x_2), \tag{2.2}$$

plus further moment equations expressing the condition that a be Gaussian.[7] We propose here still another definition:

$$a(x) = \xi(x)[\delta(0)]^{\frac{1}{2}}, \tag{2.3}$$

where $\delta(0)$ is the Dirac delta function of zero argument, and $\xi(x)$ is in turn defined as follows: for any fixed x it is a Gaussian random variable having zero mean and unit variance, and it is independent of $\xi(x')$ whenever $x' \neq x$. This completes the definition of $a(x)$. The extraction of the factor $[\delta(0)]^{\frac{1}{2}}$ has made it possible to express the properties of $a(x)$ in terms of those of the random variable $\xi(x)$, which is a harmless Gaussian random variable, i.e., $\langle\xi(x)\rangle = 0$, $\langle\xi^2(x)\rangle = 1$, $\langle\xi^n(x)\rangle = 0$ (n odd), $\langle\xi^m(x)\rangle = (m-1)!!$ (m even).

The equivalence of this to the above definitions (a) and (b) may be shown with the use of the identity

$$\delta(x_1 - x_2) = \delta_{x_1x_2}\delta(0), \tag{2.4}$$

where $\delta_{x_1x_2}$ is the *Kronecker* delta, equal to 1 when $x_1 = x_2$, to zero otherwise. This identity is proved in the Appendix [Eq. (A6)].

The equivalence to Definition (b) follows with the use of Eq. (2.4) from

$$\langle\xi(x_1)\xi(x_2)\rangle = \delta_{x_1x_2}, \tag{2.5}$$

which is a direct consequence of the definition of $\xi(x)$; Eqs. (2.1) and (2.2) and all the higher moment equations entailed by the Gaussianity of $a(t)$ are easily derived in this way.

The equivalence to Wiener's definition (a) is demonstrated by the fact that the integral of $a(x)$,

$$X(x) = \int_0^x a(x')\,dx' \tag{2.6}$$

has all the properties of the Wiener function: It is Gaussian because it is the sum of the Gaussian random variables $a(x)$ or $\xi(x)$; it has zero mean, if we interchange the order of integration with the operation of taking the mean,

$$\langle X(x)\rangle = \int_0^x \langle a(x')\rangle\,dx' = 0; \tag{2.7}$$

it has variance x:

$$\langle X^2(x)\rangle = \int_0^x\int_0^x \langle a(x')a(x'')\rangle\,dx'\,dx''$$

$$= \int_0^x\int_0^x \delta(x' - x'')\,dx'\,dx'' = x; \tag{2.8}$$

and its increment $X(x_2) - X(x_1)$ over an interval (x_1, x_2) is independent of the increment over any interval not overlapping (x_1, x_2), since the variables $a(x)$ summed in

$$X(x_2) - X(x_1) = \int_{x_1}^{x_2} a(x)\,dx \tag{2.9}$$

are independent of those which contribute to the increment over any nonoverlapping interval.

The conceptual usefulness of the function ξ lies in that with its aid the formalism of the ideal random function of a continuous variable is reduced to that of a function of a discrete variable, equivalent to a set of independent discrete Gaussian variables; the latter being a much more elementary concept than that of the Wiener function. By making the formalisms of discrete and continuous variables equivalent, it does for the Wiener process something quite analogous to what the Dirac δ-function method does for quantum mechanics in breaking down the formal distinction between the function-space transformation theory of operators having continuous and those having discrete eigenvalue specta.

In order to bring out the full utility of this scheme, it is desirable to make even more explicit than is usually done the way in which the δ-function calculus depends on the interpretation of integrals as sums. This is done in the Appendix.

Besides reducing a mathematically sophisticated concept formally to a more elementary one, the reduction of integrals to sums is extremely useful expositionally where the Wiener process is a vector rather than a scalar function. In such a case it may be said to depend on a discrete index as well as on a continuous variable, and in the rigorous theory the resulting combination of sums *and* integrals in

[4] N. Wiener, J. Math. and Phys. **2**, 131 (1923).
[5] N. Wiener, Acta Math. 55, 117 (1930), Sec. 13.
[6] R. E. A. C. Paley and N. Wiener, *Fourier Transforms in the Complex Domain* (American Mathematical Society, New York, 1934), Chap. 9.
[7] M. C. Wang and G. E. Uhlenbeck, Rev. Mod. Phys. **17**, 323 (1945), Sec. 9a.

many common formulas becomes very complicated in appearance. Our formalism allows us to write such sums-and-integrals simply as sums, and to combine discrete and continuous variables into one multiple variable. This results in considerable simplification of notation, relatively, in a field where the formalism tends to become complicated. We shall see how this works out in Sec. IIIB.

B. Multiple Argument, Including Discrete Indices

If the function to be expanded, f, of argument x is a tensor of rank d, it will have discrete-valued indices $\alpha^{(1)}, \alpha^{(2)}, \cdots, \alpha^{(d)}$ to denote its components. We assume that x, in turn, is a multiple or vector with c components, each component ranging over a continuum of values. If f is thought of as a vector in function space, the components of x are a set of c continuous-valued indices on the same footing as the $\alpha^{( )}$, and we denote them by $x^{(1)}, x^{(2)}, \cdots, x^{(c)}$.[8] (N. B.: Vectors or quantities having multiple components will enter the discussion in so many interrelated ways that it would be profitless and even confusing to distinguish them by typographical devices, such as boldface type. It has already been stated that in this section x stands for a c-dimensional vector, or a c-tuple. In other places it will always be made clear whether x stands for a single or multiple variable.)

If α stands for the d-tuple $\alpha^{(1)}, \cdots, \alpha^{(d)}$, and x (when the context is appropriate) for the c-tuple $x^{(1)}, \cdots, x^{(c)}$, the ideal random function of α and x is

$$a(\alpha, x), \tag{2.10}$$

which must satisfy the moment equations

$$\begin{aligned} \langle a(\alpha, x)\rangle &= 0, \\ \langle a(\alpha_1, x_1)a(\alpha_2, x_2)\rangle &= \delta_{\alpha_1\alpha_2}\delta(x_2 - x_1), \end{aligned} \tag{2.11}$$

where the δ functions are multiple δ functions, in keeping with the multiple nature of their arguments; plus further moment equations, as in the case of a single argument, to ensure Gaussianity. To set up a transformation analogous to (2.3) in this case, we extract a factor $\delta^{\frac{1}{2}}(0)$ for each continuous index:

$$a(\alpha, x) = \xi(\alpha, x)\delta^{c/2}(0), \tag{2.12}$$

in which $\delta(0)$ is the one-dimensional Dirac δ function of argument zero [with zero argument there is no convenient way of indicating multiple character in a Dirac delta function, so we write $\delta(0)$ only to denote a single, or one-dimensional, "Dirac infinity"]. The variables α and x are now on the same footing, at least as far as the properties of $\xi(\alpha, x)$ are concerned; thus we define the $(c + d)$-tuple

$$R = (\alpha, x)$$

and put

$$a(\alpha, x) = a(R) = \xi(R)\delta^{c/2}(0). \tag{2.13}$$

$a(R)$ is now defined as a random variable by defining $\xi(R)$, and this in turn is done by exact analogy to the simple case: $\xi(R)$ is for any given R a Gaussian random variable of zero mean and unit variance, independent of $\xi(R')$ whenever $R' \neq R$.

This definition implies

$$\begin{aligned} \langle \xi(R)\rangle &= 0, \\ \langle \xi(R_1)\xi(R_2)\rangle &= \delta_{R_1R_2} \equiv \delta_{R_1{}^{(1)}R_2{}^{(1)}} \cdots \delta_{R_1{}^{(c+d)}R_2{}^{(c+d)}}, \end{aligned} \tag{2.14}$$

etc.; or, in words, that to each point of the space of R there is associated an independent Gaussian variable of zero mean and unit variance.

The generalized Wiener function for a multiple argument can be constructed from this ideal random function. It has to be defined as a function of sets of points of the space $\{R\}$; thus the values it takes on will be, strictly speaking, analogous not to $X(x)$ but to

$$\int_{(s)} dX(x') = \int_{(s)} a(x')\, dx', \tag{2.15}$$

where x is a single variable and s is a measurable set of points x [$X(x)$ is of course the special case of (2.15) for which s is the interval $(0, x)$]. Therefore we discuss sets S of points of $\{R\}$. For practical purposes these must be constructed additively from subsets of points x (now again x stands for a multiple variable) each having some fixed α value, i.e., there is a range S_d of α values comprised in S and for each α, S has a subset $S_c(\alpha)$ of x values, so that S is built up as the sum

$$S = \sum_{\alpha \in S_d} S_c(\alpha). \tag{2.16}$$

$S_c(\alpha)$ is assumed measurable, and the measure of S is

$$M(S) = \sum_{\alpha \in S_d} M[S_c(\alpha)]. \tag{2.17}$$

The generalized Wiener function, which we denote by X, is defined by

$$X(S) = \sum_{R \in S} a(R) = \sum_{\alpha \in S_d} \int_{x \in S_c(\alpha)} a(\alpha, x)\, dx. \tag{2.18}$$

As a sum of Gaussian random variables of zero mean, this will be itself Gaussian and have zero mean. Its variance can be computed by the usual

[8] d stands for the number of discrete-valued indices, c for the number of continuous-valued indices.

method, with details adapted to the multiple nature of the argument:

$$\begin{aligned}\langle X^2(S)\rangle &= \sum_{\alpha\in S_d}\int_{x\in S_e(\alpha)} dx \\ &\quad\times \sum_{\alpha'\in S_d}\delta_{\alpha\alpha'}\int_{x'\in S_e(\alpha')}\delta(x-x')\,dx' \\ &= \sum_{\alpha\in S_d}\int_{x\in S_e(\alpha)} dx\int_{x'\in S_e(\alpha)}\delta(x-x')\,dx' \\ &= \sum_{\alpha\in S_d}\int_{x\in S_e(\alpha)} dx \\ &= \sum_{\alpha\in S_d} M[S_e(\alpha)],\end{aligned} \tag{2.19}$$

agreeing with (2.16).

III. WIENER-HERMITE FUNCTIONALS

A. Review and Reformulation of Hermite Functions of Denumerably Many Variables

Our formulation of the Wiener–Hermite functionals will be based on the point of view that the space of functions f is a vector space, with each value of R (or each set of values of its components) playing the role of an index. We can then construct these functionals entirely along the lines of the formalism of ordinary Hermite functions of multidimensional argument.

We therefore now give, as background, the definition of an appropriate set of Hermite functions of a denumerable set of variables $\{\xi_i\}$. The range of values of i need not be specified; summations will always be assumed to be carried over all variables ξ_i.

Using as weighting function $\exp[-(\sum \xi_i^2)/2]$, one possible set of Hermite polynomials forming a complete orthogonal set over $(-\infty, \infty)$ in all the ξ's would consist of functions[9] of the form

$$e^{+\frac{1}{2}\sum\xi_i^2}\left[\prod_i\left(-\frac{\partial}{\partial\xi_i}\right)^{n_i}\right]e^{-\frac{1}{2}\sum\xi_i^2}, \tag{3.1}$$

where the set $\cdots n_i \cdots$ runs over all possible sets of nonnegative integers.[10]

The above way of writing the Hermite functions is not convenient for defining the functionals, or any Hermite functions of an infinite number of variables, since it would explicitly require the specification of an infinite number of n_i—an infinite number of which are zero anyway, for functions or functionals of finite order. It is better to specify the variables with respect to which differentiation occurs, repeating them to take care of multiple differentiation. Thus we put

$$H^{(n)}(i_1, i_2, \cdots, i_n) = e^{\frac{1}{2}\sum\xi_i^2}\prod_{k=1}^{n}\left(-\frac{\partial}{\partial\xi_{i_k}}\right)e^{-\frac{1}{2}\sum\xi_i^2}. \tag{3.2}$$

The complete set is obtained by letting n range from zero to infinity, and letting the $i_1, i_2, \cdots, i_n$ for any given n range over all possible sets of n i's, repeated i's being allowed.

It is more or less obvious that this set is the same as the previous one, and we omit the proof. One essential remark should be made concerning the notation of (3.2): the quantities $i_1, i_2, \cdots, i_n$ in the parenthesis of $H^{(n)}(i_1, i_2, \cdots, i_n)$ are not the arguments of $H^{(n)}$ in the usual sense of Hermite functions, rather the ξ's are the arguments in this sense; but in the usual sense of the theory of *random* functions $H^{(n)}$ is a random function of the $i_1, i_2, \cdots, i_n$. It will be noted that the superscript n is not strictly necessary, since the order of the polynomial is indicated by the number of indices i, but this redundant index is often a help in the readability of formulas.

B. Explicit Definition of the Functionals

When it is realized that for each different value of i we have in the above a completely independent variable ξ_i, the formal result obtained can immediately be taken over for functionals. We simply replace i by R, writing $\xi(R)$, and realize that again for each different R, i.e., for each distinct set of values of $R^{(1)}, \cdots, R^{(c+d)}$, we have an independent variable $\xi(R)$. (We use subscripts to denote different values of R, viz.: R_k; the subscript is to be distinguished from the superscript, which labels *components* of R.) The Wiener–Hermite functionals are then defined by

$$\begin{aligned}H^n(R_1, R_2, \cdots, R_n) &= \delta^{nc/2}(0)\,e^{\frac{1}{2}\sum_R\xi^2(R)} \\ &\quad\times\prod_{k=1}^{n}\left(-\frac{\partial}{\partial\xi(R_k)}\right)e^{-\frac{1}{2}\sum_R\xi^2(R)},\end{aligned} \tag{3.3}$$

in complete analogy to the discrete polynomials (3.2), except for the factor $\delta^{nc/2}(0)$, which are explained later. The remainder of this section will

[9] If the range of values of i is infinite, $\sum \xi_i^2$ in the weighting function will usually be divergent, hence (3.1) and (3.2) will not then exist in a strict sense. However, these expressions are used only in a formal sense. The final physical quantity of any calculation in a theory like this is an expectation value, and expectation values whenever reasonably defined are convergent quantities.

[10] A. Erdélyi *et al.*, *Higher Transcendental Functions* (McGraw-Hill Book Company, Inc., New York, 1953), Vol. 2, Chap. 12. In this reference a more general definition is used, in which the weighting function is the exponential function of minus an arbitrary positive-definite quadratic form in the ξ's. We have specialized to the unit matrix as the matrix of the quadratic form.

be devoted to discussion of this formula and to expressing it in various alternative ways.

(1) The summation over R in the exponent is a simplifying convention which takes advantage of the reduction of integration to summation made possible by our symbolic calculus. It may more familiarly be interpreted as an integral as far as the continuous components of R are concerned; the transformation is straightforward:

$$\sum_R \xi^2(R) = \sum_\alpha \sum_x \xi^2(\alpha, x) = \sum_\alpha \sum_x a^2(\alpha, x)\ \delta^{-c}(0)$$
$$= \sum_\alpha \int a^2(\alpha, x)\, dx, \tag{3.4}$$

where the integral is over the entire range of points x. The replacement of $\delta^{-c}(0)$ by dx is based on Eq. (A5) of the Appendix.

The integral in (3.4) diverges, of course, with probability one[4]; but see Footnote (9).

(2) The factor $\delta^{nc/2}(0)$ in the definition of $H^{(n)}$ ensures Dirac δ-function normalization for the continuous components of R (see Sec. IVC, second paragraph). With it, the derivatives with respect to $\xi(R_k)$ can be written as functional derivatives as far as the continuous components of R_k are concerned: With the use of (2.13), we have

$$\delta^{nc/2}(0) \prod_{k=1}^{n} \left(-\frac{\partial}{\partial \xi(R_k)}\right) = \prod_{k=1}^{n} \left(-\frac{1}{\delta^{-c}(0)} \frac{\partial}{\partial a(R_k)}\right)$$
$$= \prod_{k=1}^{n} \left(-\frac{1}{dx} \frac{\partial}{\partial a(R_k)}\right), \tag{3.5}$$

where $dx = dx^{(1)}\, dx^{(2)} \cdots dx^{(c)}$. When R has only continuous components, the factors of the product (3.5) are functional derivatives, and may be written in familiar form

$$\frac{1}{dx} \frac{\partial}{\partial a(R_k)} = \frac{\delta}{\delta a(R_k)}. \tag{3.6}$$

When R has both discrete and continuous components, the factors of (3.5) are hybrids, being ordinary partial derivatives relative to the discrete components of R_k and functional derivatives relative to the continuous components; and it would then seem best to leave $H^{(n)}$ in the form (3.3).

(3) The definition (3.2) for the Hermite functions, as distinguished from (3.1), exhibits their tensor character. Similarly the form (3.3) exhibits the functionals as tensors in function space: Let $\xi(\cdot)$ denote the vector whose components are the $\xi(R)$, R being allowed to range over all of its possible values. That is, if $e(R)$ is the unit vector in function space for argument R,

$$\xi(\cdot) = \sum_R e(R)\xi(R). \tag{3.7}$$

Denote the gradient in this function space by

$$\frac{\partial}{\partial \xi(\cdot)} = \sum_R e(R) \frac{\partial}{\partial \xi(R)}. \tag{3.8}$$

Then the Wiener–Hermite functionals (3.3) are the components of the n-adic

$$\mathsf{H}^{(n)} = \delta^{nc/2}(0)\, e^{\frac{1}{2}\xi(\cdot)\cdot\xi(\cdot)}[-\partial/\partial\xi(\cdot)]^n\, e^{-\frac{1}{2}\xi(\cdot)\cdot\xi(\cdot)}, \tag{3.9}$$

where the dot between the $\xi(\cdot)$ denotes scalar product in function space and the nth power denotes the direct product of n vector factors $\partial/\partial\xi(\cdot)$. The Wiener–Hermite functional is obtained from the tensor as

$$H^{(n)}(R_1, R_2, \cdots, R_n) = e(R_1)e(R_2) \cdots e(R_n)\cdot \mathsf{H}^{(n)}, \tag{3.10}$$

in which $e(R_1)\ e(R_2) \cdots e(R_n)$ is a direct product, and the dot denotes an n-fold scalar product. Since (3.3) is independent of the order of differentiation, the order of the unit vectors in (3.10) may be changed without affecting the result.

IV. EXPECTATION VALUES

A. Outline of Method

The expansion of a random function f in Wiener–Hermite functionals is given by

$$f(R) = \sum_{R_1} K^{(1)}(R; R_1)H^{(1)}(R_1)$$
$$+ \sum_{R_1, R_2} K^{(2)}(R; R_1, R_2)H^{(2)}(R_1, R_2) + \cdots$$
$$+ \sum_{R_1, R_2, \cdots, R_n} K^{(n)}(R; R_1, R_2, \cdots, R_n)$$
$$\times H^{(n)}(R_1, R_2, \cdots, R_n) + \cdots. \tag{4.1}$$

[It has been assumed that $f(R)$ has zero mean for all R, which removes any $H^{(0)}$ term and does not lead to any significant loss of generality.] Here it must be borne in mind that the kernels K are ordinary functions of their arguments, while the H's are random functions.

The statistical properties of f will be determined by its moments, i.e., by expectation values of the form

$$\langle f(R')f(R'') \cdots f(R^{(N)})\rangle. \tag{4.2}$$

The computation of such an expression can by commuting the $\langle\ \rangle$ operation with the sums in (4.1), be reduced to the evaluation of integrals of products of the K's, since the expectation values of products of H's are invariably combinations of Dirac and Kronecker δ's. Hence we now address ourselves to

the problem of the evaluation of the expectation values of products of the H's.

From the Gaussian distribution of ξ we have

$$\langle H^{(n_1)}H^{(n_2)}\cdots H^{(n_l)}\rangle$$
$$= \int H^{(n_1)}H^{(n_2)}\cdots H^{(n_l)}\, e^{-\frac{1}{2}\sum\xi^2(R)} \prod_R \frac{d\xi(R)}{(2\pi)^{\frac{1}{2}}}$$
$$= \int e^{-\frac{1}{2}\sum\xi^2(R)}H^{(n_1)}H^{(n_2)}\cdots H^{(n_l)}e^{-\frac{1}{2}\sum\xi^2(R)} \prod_R \frac{d\xi(R)}{(2\pi)^{\frac{1}{2}}} \tag{4.3}$$

(in this and other equations for expectation values, arguments are to be thought of as assigned to the $H^{(n)}$, but are omitted for the sake of brevity). The motive for the last form of writing[11] is that exp $[-\frac{1}{4}\sum \xi^2(R)]$ is the Wiener–Hermite function-functional (not polynomial-functional) of order zero, and the H's can be written very simply in terms of "creation" and "destruction" operators which raise or lower the order of Wiener–Hermite function-functionals. [A new degree of verbal fine-structure has been introduced here. In mathematical physics, the distinction between Hermite polynomials and Hermite functions is that the latter contain as an extra factor the square root of the weighting function associated with the former, so that the inner product for the Hermite functions has weighting function unity. However, we have earlier defined Hermite "functionals" as generalizations of Hermite "polynomials" in which the number of variables becomes nondenumerable, i.e., the functionals have in their inner product the (generalized) weighting function of Hermite polynomials rather than unity. If the functional generalization of an Hermite polynomial is a Wiener–Hermite functional, what shall we call that of an Hermite function? We could have avoided this difficulty by calling the functional generalization of the polynomial a "polynomialal", but this seemed unbearable on grounds of esthetics and euphony, and people might have thought we were only stuttering. Therefore we decided as follows: Where the distinction between weighting functions needs to be made clear, to speak of the functionals with weighting function exp $(-\sum \xi^2(R)/2)$ as "polynomial-functionals," and to refer to the others as "function-functionals"; and where the distinction is not required by the context, to drop the modifier.] That is, writing

$$h^{(m)}(R_1, R_2, \cdots, R_m) = e^{-\frac{1}{2}\sum\xi^2(R)}H^{(m)}(R_1, R_2, \cdots, R_m) \tag{4.4}$$

for the Wiener–Hermite function-functional, (4.3) will become a power of 2π times the inner product of $h^{(0)}$ (on the left) with a linear combination of h's of various orders. Only the $h^{(0)}$ term on the right gives a nonvanishing contribution, and the problem therefore may be said to be that of finding its numerical coefficient, since

$$(2\pi)^{-\frac{1}{2}(n_1+\cdots+n_l)}(h^{(0)}, h^{(0)}) = \int e^{-\frac{1}{2}\sum\xi^2(R)} \prod_R \frac{d\xi(R)}{(2\pi)^{\frac{1}{2}}} = 1; \tag{4.5}$$

and the evaluation of this coefficient can be simplified considerably by the creation–destruction-operator device.

B. Wiener–Hermite Functions and Functionals in Terms of Creation and Destruction Operators

The starting point for expressing $H^{(n)}(R_1, R_2, \cdots, R_n)$ in terms of creation and destruction operators is the following expression[12] for the Hermite *polynomial* of a single argument y:

$$H_m(y) = \sum_{k=0}^{m} \binom{m}{k} C^{m-k}\, D^k \tag{4.6}$$

where

$$C = \tfrac{1}{2}y - d/dy, \qquad D = \tfrac{1}{2}y + d/dy \tag{4.7}$$

are, respectively, creation and destruction operators on the Hermite *functions* h_n, i.e., given

$$h_0 = e^{-y^2/4}, \tag{4.8}$$

then

$$Ch_n = h_{n+1}, \qquad Dh_n = nh_{n-1}. \tag{4.8'}$$

Warning: In the context of discussions of Hermite functions of a single variable $H_m(y)$, y *is* the argument of H_m (cf. Sec. IIIA). Our notation for Hermite functions of a single variable therefore is not consistent with that for multiple variables. The distinction is clear if it is realized that in the former the H is labeled by a subscript, and in the later by a superscript.

It will be observed that the expression for $H_m(y)$ in (4.6) would be the expansion of $(C + D)^m$ according to the binomial theorem if all D operators were moved to the right and C operators to the left in

[11] Equation (4.3) may be interpreted as follows: The *probabilistic* expectation value of the *functional* $H^{(n_1)}\cdots H^{(n_l)}$ is equal to the *quantum-mechanical* expectation value of the *operator functional* $H^{(n_1)}\cdots H^{(n_l)}$ of the operators $\xi(R_1)\cdots$ in the "vacuum state" (multidimensional harmonic oscillator ground state), whose wavefunction (before normalization) is exp $(-\frac{1}{4}\sum \xi^2(R))$.

[12] A. Siegel, "A New Expansion of the Differential Operator for the Time Development of Fluctuation Distributions," in *Rarefied Gas Dynamics*, edited by L. Talbot (Academic Press Inc., New York, 1961).

the individual terms of the expansion. We accordingly write[13]

$$H_m(y) = :(C + D)^m:, \tag{4.9}$$

where the colons : : denote an *ordered product* in which the C and D factors of the operator inside are rearranged so that a C operator never operates before a D operator.

The generalization of (4.9) to the functionals is readily done as follows: Since differentiations commute, all derivatives for equal values of R in (3.3) may be grouped together. Suppose there are n_1 differentiations $\partial/\partial R'$; n_2 differentiations $\partial/\partial R''$, $\cdots$; n_s differentiations $\partial/\partial R^s$, $\cdots$. Then (3.3) may be transformed into

$$\begin{aligned}&\delta^{nc/2}(0)\, e^{\frac{1}{2}\xi^2(R')}[-\partial/\partial\xi(R')]^{n_1} e^{-\frac{1}{2}\xi^2(R')} \\ &\quad\times e^{\frac{1}{2}\xi^2(R'')}[-\partial/\partial\xi(R'')]^{n_2} e^{-\frac{1}{2}\xi^2(R'')}\cdots \\ &\quad\times e^{\frac{1}{2}\xi^2(R^s)}[-\partial/\partial\xi(R^s)]^{n_s} e^{-\frac{1}{2}\xi^2(R^s)}\cdots .\end{aligned} \tag{4.10}$$

[This is equivalent to the previously discarded way of writing, namely (3.1), which happens to be useful for this particular application.] But this is a product of Hermite functions of single argument:

$$\begin{aligned}&H^{(n)}(R_1, R_2, \cdots, R_n) \\ &= \delta^{nc/2}(0) H_{n_1}[\xi(R')] H_{n_2}[\xi(R'')]\cdots H_{n_s}[\xi(R^s)]\cdots ,\end{aligned} \tag{4.11}$$

which may be written, according to (4.9), as

$$\begin{aligned}H^{(n)}(R_1, R_2, \cdots, R_n) = \delta^{nc/2}(0) &:[C(R') + D(R')]^{n_1}: \\ &\times :[C(R'') + D(R'')]^{n_2}:\cdots \\ &\times :[C(R^s) + D(R^s)]^{n_s}:\cdots ,\end{aligned} \tag{4.12}$$

where $C(R)$, $D(R) \equiv \xi(R)/2 \mp d/d\xi(R)$ But. R', R'', $\cdots$ are all unequal to one another, hence the C's and D's in the *different* ordered products commute with one another. Such ordering in this expression as is due to the order in which R', R'' etc. appear is therefore immaterial and the over-all product may without changing its value be ordered with all C operators on the left and D operators on the right. Denoting the ordering of all C, D operators regardless of their R tags by the same symbol : :, means that

$$\begin{aligned}&H^{(n)}(R_1, R_2, \cdots, R_n) = \delta^{nc/2}(0) :[C(R') + D(R')]^{n_1} \\ &\times [C(R'') + D(R'')]^{n_2}\cdots [C(R^s) + D(R^s)]^{n_s}\cdots :.\end{aligned} \tag{4.13}$$

[13] G. C. Wick, Phys. Rev. **80**, 268 (1950), introduced the notation of Eq. (4.9).

But now, again due to the commutation of C and D operators belonging to different R values, the sorting out of equal R values which led to (4.10) may be reversed and the $C + D$ factors restored to their original order as in (3.3). We thus obtain the result

$$\begin{aligned}&H^{(n)}(R_1, R_2, \cdots, R_n) \\ &\quad = \delta^{nc/2}(0) :\prod_{k=1}^{n}[C(R_k) + D(R_k)]:.\end{aligned} \tag{4.14}$$

It is worth noticing that this is the simplest imaginable generalization of the one-discrete-variable formula (4.9). [Note that it reduces correctly to (4.9) if the R_k are discrete-valued and all equal.] To go from (4.9) to (4.14) all one does is the following: (a) Attach a factor $\delta^{n/2}$ for each continuous variable, (b) insert between the : : a factor $C(R_k) + D(R_k)$ for each variable R_k instead of m factors $C + D$. But simple as the result may be, it does not appear to be self-evident.

C. Calculation of Expectation Values

As stated in Sec. IVA, we are seeking formulas for the expectation values of products of Hermite functionals. We derive these by a method borrowed from the quantum theory of fields, but make the exposition self-contained and not dependent on quantum-field-theoretical vocabulary. The result, which we prove, is

$$\begin{aligned}&\langle H^{(n_1)}(R_1, \cdots, R_{n_1}) H^{(n_2)}(R_{n_1+1}, \cdots, R_{n_1+n_2})\cdots \\ &\times H^{(n_m)}(R_{n_1+n_2+\cdots+n_{m-1}+1}, \cdots, R_{n_1+n_2+\cdots+n_m})\rangle \\ &= \begin{cases}[\delta(0)]^{\frac{1}{2}(n_1+\cdots n_m)C} \sum\limits_{\substack{\text{distinct}\\ \text{exogamous}\\ \text{pairings}}} \prod\limits_{\substack{\text{exogamous}\\ \text{pairs}}} \delta_{R_iR_j} \\ \qquad \text{if} \quad n_1 + \cdots + n_m \quad \text{is even}, \\ 0 \qquad \text{if} \quad n_1 + \cdots + n_m \quad \text{is odd},\end{cases}\end{aligned} \tag{4.15}$$

where the symbols are to be interpreted as follows:

$$\prod_{\substack{\text{exogamous}\\ \text{pairs}}} \delta_{R_iR_j} \tag{4.16}$$

is a product in which each R index from R_1 to $R_{n_1\cdots n_m}$ appears just once as a subscript of the Kronecker delta multiplicand, subject to the restriction that each pair coupled in a delta function be (using a term borrowed from anthropology) "exogamous" in the sense that the two spouses in it come from different H functions. In the summation, such a product appears just once for each distinct way of arranging all the R variables in exogamous pairs (mere interchange of arguments in

a pair is not considered to give a distinct way of pairing).

It will be noted that the power of $\delta(0)$ in (4.15) is just such as to make a Dirac δ function out of each Kronecker δ that has continuous indices when the multidimensional Kronecker δ is written out in terms of unidimensional ones.

We shall see in the derivation that the result (4.15) can be given a simple heuristic interpretation.

We now give the proof, breaking it down into five steps.

1. Expectation of ξ^n

As is well known, for a Gaussian ξ of zero mean,

$$\begin{aligned} \langle \xi^n \rangle &= (n-1)!! \qquad (n \text{ even}) \\ &= 0 \qquad\qquad (n \text{ odd}). \end{aligned} \tag{4.17}$$

We wish to emphasize, for the sake of our proof, that this quantity is equal for any n (even *or* odd) to the number of ways of arranging n points (or objects of any kind; but points lend themselves better to our later interpretation) in pairs such that every point is paired with just one other point.

2. Expectation of a Product of ξ's of Different Arguments

Since ξ's of different arguments are independent, if we have k points $R_1, R_2, \cdots, R_k$ we may write

$$\langle \xi(R_1)\xi(R_2)\cdots\xi(R_k)\rangle = \langle \xi^a\rangle\langle \xi^b\rangle\cdots, \tag{4.18}$$

if $a + b + \cdots = k$ and the $R_1, R_2, \cdots, R_k$ are equal in groups containing a, b, $\cdots$ members respectively and unequal outside these groups. From this and the pair interpretation of the last paragraph,

$$\langle \xi(R_1)\xi(R_2)\cdots\xi(R_k)\rangle = \left\{\begin{array}{l}\text{number of ways of arranging the } k \text{ } R\text{'s} \\ \text{so that each } R \text{ is paired with just one} \\ \text{equal } R.\end{array}\right\} \tag{4.19}$$

This number may be evaluated as follows: Define an "indicator" function of $R_1, R_2, \cdots, R_k$, which is equal to unity when every R is paired with just one equal R, and zero otherwise. Such an indicator function is equal to

$$\prod_{\text{pairs}} \delta_{R_i R_j}, \qquad (k \text{ even}) \tag{4.20}$$

the product over pairs being defined such that every two R's which are paired in the given arrangement appear in the same δ function; and to zero if k is odd. The number of arrangements in (4.19) is then the sum of the indicator function over distinct pairings, i.e.,

$$\langle \xi(R_1)\xi(R_2)\cdots\xi(R_k)\rangle = \begin{cases} \sum_{\substack{\text{distinct}\\ \text{pairings}}} \prod_{\text{pairs}} \delta_{R_i R_j} & \text{if } k \text{ is even}, \\ 0 & \text{if } k \text{ is odd}. \end{cases} \tag{4.21}$$

3. Expectation of a Product of $H^{(n)}$'s (Preliminary Discussion)

The value of

$$\begin{aligned} &[\delta(0)]^{-(n_1+\cdots+n_m)C} \\ &\quad\times \langle H^{(n_1)}(R_1\cdots R_{n_1})H^{(n_2)}(R_{n_1+1},\cdots R_{n_1+n_2})\cdots \\ &\quad\times H^{(n_m)}(R_{n_1+\cdots+n_{m-1}+1},\cdots R_{n_1+\cdots+n_m})\rangle \\ &\quad= \langle :\xi(R_1)\cdots\xi(R_{n_1}):\,:\xi(R_{n_1+1})\cdots\xi(R_{n_1+n_2}):\cdots \\ &\quad\times :\xi(R_{n_1+\cdots+n_{m-1}+1})\cdots\xi(R_{n_1+\cdots+n_m}):\rangle \end{aligned} \tag{4.22}$$

differs from that of

$$\langle \xi(R_1)\cdots\xi(R_{n_1})\xi(R_{n_1+1})\cdots\xi(R_{n_1+\cdots+n_m})\rangle, \tag{4.23}$$

i.e., from an expectation of the same form as that evaluated in the previous paragraph, only in that every C, D-operator product obtained when the former is multiplied out has, in the latter, the C operators within each : : moved to the left of all D operators in the same : :. To see what difference this makes, we need a lemma on products of C and D operators, proved in the next paragraph.

4. Expectation of a Product of C and D Operators of Like Index

Let Q be a product of C and D operators operating on the same variable, in any order. In order to have nonzero expectation value, the numbers of C and D operators must be equal, and the number of C operators to the right of any point in the product must be not less than the number of D operators to the right of the same point.[14] Assuming these conditions, the value of $\langle Q\rangle$ may be obtained as follows: Commute each and every C operator in Q to the left of all D operators. By commuting a C operator to the left of all D operators we mean replacing it by the operator obtained by transferring it to the left of all D operators plus a sum constructed as follows: Strike out of Q the C in question and one of the D operators to its left; the sum is then to consist of one term so obtained

[14] To see this, and also to interpret some of the ensuing arguments, the quantum-mechanical picture of Footnote 11 is helpful. One then regards the expectation value of Q as a *quantum* expectation value of the operator in the "vacuum state" h_0, i.e., as an inner product

$$\langle Q\rangle = (2\pi)^{-\frac{1}{2}}(h_0, Qh_0).$$

Thus Q must transform h_0 into a multiple of itself, i.e., have as many creation as destruction operators, in order to have nonvanishing $\langle Q\rangle$; etc.

for each D operator to the left of the C operator concerned. This is justified by the identity

$$D^m C = CD^m + D^{m-1}[D, C] + D^{m-2}[D, C]D + \cdots + [D, C]D^{m-1}, \quad (4.24)$$

combined with (a) the fact that the expectation value of the first term vanishes, $\langle CD^m \rangle = 0$ [cf. Footnote 14: $(h_0, CD^m h_0) = (Dh_0, D^m h_0) = (0, D^m h_0)$, if we note that $D = C^+$]; (b) the fact that $[D, C] = 1$.

When this process has been carried out for all C operators, it will be seen that $\langle Q \rangle$ is equal to the expectation value of a sum of unit operators, one for each way in which each C operator in Q can be paired with a D operator on its left. Since $\langle 1 \rangle = 1$, $\langle Q \rangle$ = Numbers of ways of arranging the C and D operators of Q in pairs, each pair being such that its C operator lies to the right of its D operator in the original order of Q. Figure 1 illustrates a permissible mode of pairing.

5. Expectation of a Product of $H^{(n)}$'s (Conclusion)

Consider the result of the previous section relative to that of Sec. IVC3. According to IVC4, a pairing of a C and a D operator of like index can contribute to the expectation value of a product of operators of the same index if and only if the C operator lies to the right of the D operator. But in IVC3 we saw that all C operators within a : : lie to the *left* of all D operators within the same : :. Hence (4.22) differs from (4.23) only in the omission from the former of all contributions from the pairing of R's which lie in the same : :. If such R's are called *endogamous* (as antonym of *exogamous*), the theorem (4.15) follows.

V. CALCULATION OF THE FIRST APPROXIMATION TO THE FLATNESS FACTOR FROM THE WIENER-HERMITE FUNCTIONAL EXPANSION

The flatness factor of the random variable $f(R)$ (obtained from the random function f by assigning the value R to its argument) is defined as

$$\Phi[f(R)] \equiv \langle f^4(R) \rangle / \langle f^2(R) \rangle^2. \quad (5.1)$$

Since $\Phi = 3$ for a Gaussian random variable, $\Phi - 3$ is a measure of the deviation of $f(R)$ from Gaussianity,[15] and will be essentially determined by the higher terms of the series (4.1). (Absence of higher terms is a sufficient, but not necessary, condition for Gaussianity; the vanishing of $\Phi - 3$ is a necessary condition, but not sufficient.)

[15] As a single random variable. We are not now discussing the joint distribution, or joint Gaussianity, of two or more random variables $f(R_1), f(R_2), \cdots$.

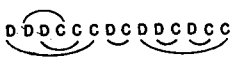


FIG. 1. A permissible mode of pairing of C and D operators.

Assuming that successive terms of the functional expansion make successively smaller contributions to physical magnitudes, we proceed to evaluate the numerator and denominator of Φ in second approximation. With the notation

$$(K^{(l)}, H^{(l)}) \equiv \sum_{R_1, R_2, \cdots R_l} K^{(l)}(R; R_1, R_2, \cdots, R_l) \times H^{(l)}(R_1, R_2, \cdots, R_l), \quad (5.2)$$

Eq. (4.1) becomes

$$f(R) = (K^{(1)}, H^{(1)}) + (K^{(2)}, H^{(2)}) + (K^{(3)}, H^{(3)}) + \cdots; \quad (5.3)$$

then

$$\langle f^2(R) \rangle = \langle (K^{(1)}, H^{(1)})^2 + 2(K^{(1)}, H^{(1)})(K^{(2)}, H^{(2)}) + (K^{(2)}, H^{(2)})^2 + 2(K^{(1)}, H^{(1)})(K^{(3)}, H^{(3)}) + \cdots \rangle \quad (5.4)$$

and

$$\langle f^4(R) \rangle = \langle (K^{(1)}, H^{(1)})^4 + 4(K^{(1)}, H^{(1)})^3 (K^{(2)}, H^{(2)}) + 4(K^{(1)}, H^{(1)})^3 (K^{(3)}, H^{(3)}) + 6(K^{(1)}, H^{(1)})^2 (K^{(2)}, H^{(2)})^2 + \cdots \rangle. \quad (5.5)$$

The next step is to distribute the expectation operation over the several sums. Then we commute $\langle\ \rangle$ and $\sum$ operations. As for the step following this, we would have for example from the first term in the last form of writing of $\langle f^2(R) \rangle$, at this stage,

$$\sum_{R_1, R_2} K^{(1)}(R_1) K^{(2)}(R_2) \langle H^{(1)}(R_1) H^{(1)}(R_2) \rangle. \quad (5.6)$$

Thus the problem has now been reduced to the evaluation of expectation values of products of H's. Actually the only such products that must be included are the following:

$$\langle H^{(1)}(R_1) H^{(1)}(R_2) \rangle = \delta(R_1 - R_2) \quad (5.7)$$

$$\langle H^{(2)}(R_1, R_2) H^{(2)}(R_3, R_4) \rangle = \delta(R_1 - R_3)\delta(R_2 - R_4) + \delta(R_1 - R_4)\delta(R_2 - R_3) \quad (5.8)$$

$$\langle H^{(1)}(R_1) H^{(1)}(R_2) H^{(1)}(R_3) H^{(1)}(R_4) \rangle = \delta(R_1 - R_2)\delta(R_3 - R_4) + \delta(R_1 - R_3)\delta(R_2 - R_4) + \delta(R_1 - R_4)\delta(R_2 - R_3), \quad (5.9)$$

$$\begin{aligned}\langle H^{(1)}(R_1)H^{(1)}(R_2)H^{(2)}(R_3,R_4)H^{(2)}(R_5,R_6)\\ = \delta(R_1-R_2)\delta(R_3-R_5)\delta(R_4-R_6)\\ + \delta(R_1-R_2)\delta(R_3-R_6)\delta(R_4-R_5)\\ + \delta(R_1-R_3)\delta(R_2-R_5)\delta(R_4-R_6)\\ + \delta(R_1-R_3)\delta(R_2-R_6)\delta(R_4-R_5)\\ + \delta(R_1-R_5)\delta(R_2-R_3)\delta(R_4-R_6)\\ + \delta(R_1-R_5)\delta(R_2-R_4)\delta(R_3-R_6)\\ + \delta(R_1-R_4)\delta(R_2-R_6)\delta(R_3-R_5)\\ + \delta(R_1-R_4)\delta(R_2-R_5)\delta(R_3-R_6)\\ + \delta(R_1-R_6)\delta(R_2-R_3)\delta(R_4-R_5)\\ + \delta(R_1-R_6)\delta(R_2-R_4)\delta(R_3-R_5) \qquad (5.10)\end{aligned}$$

$$\begin{aligned}\langle H^{(1)}(R_1)H^{(1)}(R_2)H^{(1)}(R_3)H^{(3)}(R_1,R_2,R_3)\rangle\\ = \delta(R_1-R_4)\delta(R_2-R_5)\delta(R_3-R_6)\\ + \delta(R_1-R_4)\delta(R_2-R_6)\delta(R_3-R_5)\\ + \delta(R_1-R_5)\delta(R_2-R_4)\delta(R_3-R_6)\\ + \delta(R_1-R_5)\delta(R_2-R_6)\delta(R_3-R_4)\\ + \delta(R_1-R_6)\delta(R_2-R_5)\delta(R_3-R_4)\\ + \delta(R_1-R_6)\delta(R_2-R_4)\delta(R_3-R_5). \qquad (5.11)\end{aligned}$$

The values given have been obtained from Eq. (4.15). All others arising from terms which appear explicitly in Eqs. (5.4) and (5.5) vanish, either because the total number of variables appearing in the H's is odd, making it impossible to pair every R, or because, although pairing of all R's is possible, there are no possible exogamous pairings [the sole term of this kind is the fourth term in (5.4)].

We now assume, with no actual loss of generality (because of the invariance of the Wiener–Hermite functionals to permutations of their indices), that every kernel $K^{(l)}(R; R_1, R_2, \cdots, R_l)$ is invariant with respect to all permutations of the indices $R_1, R_2, \cdots, R_l$. This has the following effect: (a) All products of δ's within any one of (5.8), (5.9), and (5.11) are mutually equivalent, and (b) the first four such products in (5.10) are mutually equivalent, as are the remaining six.

We then find

$$\begin{aligned}\langle f^2(R)\rangle &= (K^{(1)}\cdot K^{(1)}) + 2(K^{(2)} : K^{(2)}) + \cdots\\ \langle f^4(R)\rangle &= 3(K^{(1)}\cdot K^{(1)})^2 + 12(K^{(1)}\cdot K^{(1)})(K^{(2)} : K^{(2)})\\ &\quad + 24(K^{(1)}K^{(1)}K^{(1)} \vdots K^{(3)})\\ &\quad + 48((K^{(1)}\cdot K^{(2)})\cdot(K^{(1)}\cdot K^{(2)})) + \cdots, \qquad (5.12)\end{aligned}$$

in which a dot between two terms in parentheses indicates summation over any pair of disposable indices (i.e., indices $R_1, R_2, \cdots$, but never R itself, which have not yet been summed out), taking one from each side of the dot; e.g.,

$$(K^{(1)}\cdot K^{(2)}) = \int K^{(1)}(R; R_1)K^{(2)}(R; R_1, R_2)\, dR_1, \qquad (5.13)$$

leaving a "disposable" index R_2.

Suppose now that the $K^{(l)}$ are of order of magnitude regularly decreasing with increasing l, as they would be if $K^{(l)}$ were proportional to the lth power of a small constant. Then to second approximation,

$$\begin{aligned}\langle f^2(R)\rangle^2 &\cong (K^{(1)}\cdot K^{(1)})^2 + 4(K^{(1)}\cdot K^{(1)})(K^{(2)} : K^{(2)})\\ \langle f^4(R)\rangle &\cong 3\langle f^2(R)\rangle^2 + 24(K^{(1)^3} \vdots K^{(3)})\\ &\quad + 48((K^{(1)}\cdot K^{(2)})\cdot(K^{(1)}\cdot K^{(2)})) \qquad (5.14)\end{aligned}$$

and

$$\begin{aligned}\Phi[f(R)] \cong 3 &+ 24[(K^{(1)^3} \vdots K^{(3)})\\ &+ 2((K^{(1)}\cdot K^{(2)})\cdot(K^{(1)}\cdot K^{(2)}))]/(K^{(1)}\cdot K^{(1)})^2 \qquad (5.15)\end{aligned}$$

The second term on the right-hand side is the correction to Gaussianity in first approximation.

VI. RELATION TO MORE RIGOROUS THEORY

A. Expression of Functionals in Terms of the Wiener Function

We discuss here how the preceding formalism can be related to the mathematically rigorous formulation of the Wiener function. The basis of the rigorous theory is the function $X(x)$, with *postulated* properties identical with those *derived* by us in the paragraph containing Eqs. (2.7)–(2.9), which function is shown by Wiener in Refs. 4–6 to be well defined.

A linear functional of X is written by Wiener as a Stieltjes integral

$$\int f(x)\, dX(x). \qquad (6.1)$$

(Even this does not exist in the strict mathematical sense, but is rigorously defined in terms of the quantity obtained through integration by parts:

$$-\int X(x)\, df(x), \qquad (6.2)$$

provided $f(x)$ is such that this last integral does exist.) A linear functional as written in our formalism,

$$\begin{aligned}\int f(x)a(x)\, dx &= \int f(x)H_1[a(x)]\, dx\\ &= \int f(x)H^{(1)}(x)\, dx,\end{aligned} \qquad (6.3)$$

derives its rigorous meaning from the replacement of $a(x)dx$ by $dX(x)$, which puts it into the form (6.1). The evaluation of expectation values of products of such linear functionals as (6.3), which we perform here by exchanging the order of the operations of integration and taking the expectation value, is justified by the fact that the results agree with those obtained by Wiener.[5]

The use of the Wiener–Hermite functionals defined in this paper is really only an extension of the use of the ideal random function $a(x)$ in place of $X(x)$. *Linear* functionals in terms of our Wiener–Hermite functionals are of the form (6.3) and have therefore already been justified. Higher-order functionals may be considered term by term, and each higher term can be reduced to something which can readily be interpreted in terms of such integrals as (6.1). As an example, take the second-order Wiener–Hermite functional. Since

$$H^{(2)}(x_1, x_2) = a(x_1)a(x_2) - \delta(x_1 - x_2), \tag{6.4}$$

it can be given a meaning by integrating over x_1 and x_2:

$$\int_0^{x_2}\int_0^{x_1} H^{(2)}(x_1', x_2')\, dx_1'\, dx_2' = X(x_1)X(x_2) - M(x, x), \tag{6.5}$$

where $M(x_1, x_2)$ means the smaller of x_1 and x_2. In this way $H^{(2)}(x_1, x_2)$ is reduced to the product of Wiener functions and the harmless $M(x_1, x_2)$, a nonrandom variable. The second-order term in a Wiener–Hermite functional expansion is thus interpretable as a Stieltjes double integral, e.g.,

$$\iint f(x_1, x_2)\, dX(x_1)\, dX(x_2) - \iint f(x_1, x_2)\, dM(x_1, x_2). \tag{6.6}$$

Expectation values of products of functions of this type, involving any number of products of differentials of $X(x)$'s, do not differ in principle from expectation values of products of linear functionals, which have already been discussed.

B. Relation to Functionals of Cameron and Martin

Our functionals are in form a special case of those of Cameron and Martin.[1] The nature of the relationship can be seen most readily if we take the simple case in which R is a single continuous variable, which we shall call Q at this point, running from $-\infty$ to $+\infty$. Let the functions $\alpha(\omega, Q)$, with ω running over a discrete spectrum of values, form a complete orthonormal basis in the space of functions of Q, satisfying the unitarity properties

$$\sum_\omega \alpha^*(\omega, Q)\alpha(\omega, Q') = \delta(Q - Q') \tag{6.7}$$

$$\int \alpha(\omega, Q)\alpha^*(\omega', Q)\, dQ = \delta_{\omega\omega'}. \tag{6.8}$$

Then the random functions

$$A(\omega) = \int \alpha(\omega, Q)\, dX(Q) \tag{6.9}$$

are a set of independent-multi-Gaussian variables of mean zero and variance one. The functionals of Cameron and Martin can be expressed in terms of the $A(\omega)$ in the form used by us in Eq. (3.3):

$$H^{(n)}(\omega_1, \omega_2, \cdots, \omega_n) = e^{\frac{1}{2}\sum_\omega A^2(\omega)} \prod_{k=1}^{n}\left(-\frac{\partial}{\partial A(\omega_k)}\right) e^{-\frac{1}{2}\sum_\omega A^2(\omega)}. \tag{6.10}$$

The sum over ω in the exponents is of purely formal significance, since it diverges almost everywhere in Wiener space (i.e., with probability one). Our functionals in the case of one-dimensional continuous variable R are obtained by replacing ω by R and putting

$$\alpha(R, Q) = \delta^{\frac{1}{2}}(R - Q), \tag{6.11}$$

[the latter of which does satisfy (6.7) and (6.8), as may be seen by formal manipulation]. In this case one obtains

$$A(R) = \int \delta^{\frac{1}{2}}(R - Q)\, dX(Q) = \delta^{\frac{1}{2}}(0)\, dX(R) = \delta^{\frac{1}{2}}(0)[dX(R)/dR]\, dR = \delta^{-\frac{1}{2}}(0)a(R) = \xi(R), \tag{6.12}$$

where $a(R)$ and $\xi(R)$ are the same as in Eq. (2.3). Substitution of $\xi(R)$ for $A(\omega)$ in (6.10) yields the $H^{(n)}$ of (3.3), apart from the normalizing factor $\delta^{nc/2}(0)$.

The Cameron and Martin functionals are Hermite functions of a denumerable set of independent linear functionals of the Wiener function. In our case the linear functionals are specialized to be proportional to the values of the derivative of the Wiener function itself at the various values of its argument. Since the argument is continuum-valued, the results are improper functionals. But the use of a spatial or spatiotemporal argument makes for much greater clarity of interpretation when these functionals are used in the solution of differential equations in space and time.

APPENDIX: INTEGRALS AS SUMS IN THE δ-FUNCTION CALCULUS

The basic definition of the δ function $\delta(y - x)$ is that, for a certain class of functions $f(x)$,

$$\int f(x)\, \delta(y - x)\, dx = f(y) \tag{A1}$$

if the region of integration of x includes the point $x = y$. When the requirement (A1) for valid use of the δ function is satisfied, this same integral (A1) may be interpreted formally as a sum, along the following lines: (1) Consider the integral sign to stand for the operation of summation over all points in the region of integration. In a symbolic nonrigorous calculus such as this the nondenumerability of the set of points of summation should not be considered a hindrance to such an interpretation; but the reader may, if necessary, overcome this difficulty in the usual way by likening the integral to a finite Riemann sum with exceedingly fine subdivision of the region of integration. (1) In the summand $f(x)\delta(y - x)dx$, consider the differential dx to be a constant magnitude independent of x. (In the Riemann-sum interpretation, this means that the region of integration is divided into *equal* intervals, their length in the symbolic notation then being denoted by dx, which is then the "infinitesimal" of old-fashioned treatments of calculus.) Then it may be taken outside the integral (= sumation) sign

$$\int f(x)\ \delta(y - x)\ dx = dx \int f(x)\ \delta(y - x). \tag{A2}$$

(3) Put

$$\delta(y - x) = \delta_{yx}/dx, \tag{A3}$$

where δ_{yx} is the *Kronecker* delta, equal to 1 when $y = x$, to zero otherwise. (In the Riemann-sum formulation, we would have to say that δ_{yx} is equal to 1 whenever x and y are in the same interval of the Riemann subdivision, to zero otherwise.)

Substituting (A3) into (A2),

$$\begin{aligned} \int f(x)\ \delta(y - x)\ dx &= \int f(x)\ \frac{\delta_{yx}}{dx}\ dx \\ &= \int f(x)\ \delta_{yx} = f(y), \end{aligned} \tag{A4}$$

agreeing with (A1). Since it gives the same result as (A1), this point of view must be considered valid wherever (A1) is valid. It is seen that it can easily be taken over to functions of more than one variable, and to multiple integration.

It should be pointed out that, putting $y = x$ in (A3) we obtain the useful identity

$$\delta(0) = 1/dx, \tag{A5}$$

expressing an inverse relationship between the magnitudes of the "infinity" $\delta(0)$ and the "infinitesimal" dx. Substituting into Eq. (A3),

$$\delta(y - x) = \delta_{yx}\delta(0). \tag{A6}$$

Wiener–Hermite Expansion in Model Turbulence in the Late Decay Stage*

ARMAND SIEGEL AND TSUTOMU IMAMURA†
Boston University, Boston, Massachusetts
AND
WILLIAM C. MEECHAM
University of Minnesota, Minneapolis, Minnesota
(Received 12 June 1964)

The Wiener–Hermite functional expansion, which is the expansion of a random function about a Gaussian function, is here substituted into the Burgers one-dimensional model equation of turbulence. The result is a hierarchy of equations which (along with initial conditions) determine the kernel functions which play the role of expansion coefficients in the series. Initial conditions are postulated, based on physical reasoning, criteria of simplicity, and the assumption that the series is to represent the late decay stage (in which the Gaussian correction is small and also decreasing with time). These are shown to justify an iterative solution to the equations. The first correction to the Gaussian approximation is calculated. This is then tested by evaluating the correction to the flatness factor, which for an exactly Gaussian function has the value 3, but which has been found by experiment (in real three-dimensional fluids, of course) to have a value which deviates from the Gaussian value increasingly rapidly with the order of the derivative. We utilize this effect as a test of the inherent ability of the Wiener–Hermite expansion to bring to realization the physical properties implicit in the Navier–Stokes or Burgers equations. The various contributions to the flatness-factor deviation, when computed, do show a potential capability of providing a theoretical basis for the effect.

I. INTRODUCTION

IN the statistical theory of turbulence[1,2] the velocity field $\mathbf{u}(\mathbf{x}, t)$ as a function of position x and time t, which satisfies the Navier–Stokes equation

$$\partial\mathbf{u}/\partial t + \mathbf{u}\cdot\nabla\mathbf{u} = -\nabla p/\rho + \nu\nabla^2\mathbf{u} \qquad (1.1)$$

(ρ = density, p = pressure, and ν = kinematic viscosity) is regarded as a random variable. One of the best-established experimental facts in this field is that in the early stages of time decay of the turbulent motion, the velocity field at a point is Gaussianly distributed to within experimental error, and the joint distribution of the velocities at two different points is joint-Gaussian except when the points are close together, in which case strong deviations from Gaussianity occur.[3] Moreover, theory indicates that in the late stages of decay the joint distribution tends to purely Gaussian form.[4] These observations form the motivation of an expansion, originally proposed by Wiener[5] and later revived in modified form by two of the present authors,[6] of the velocity field in a series of functionals of the "ideal random function" or derivative of the Wiener random function. We call such a series the "Wiener–Hermite series." Its first-order term is exactly Gaussian, and the higher-order terms contribute successive corrections to the Gaussian form. The experimental data on the approximate Gaussianity of the velocity field lead to the expectation that this series would be rapidly convergent, hence a manageable and accurate mathematical method for turbulence calculations should be obtainable by discarding all but its first few terms. (It has been pointed out to us[7] that the Wiener–Hermite series has the important advantage that, since it is an expansion of the random function and not of its probability distribution, all quantities computed from it are, by construction, implicitly derived from a *positive-definite* probability distribution.)

To our knowledge, no further development of this method (other than a preliminary report by ourselves[8] of the work to be described more fully here) has been published to date (see, however, Footnote 11). The main reasons for this are undoubtedly (a) that the mathematical properties of such expansions first needed to be developed, and (b) that

* Work of A. S. (in part) and T. I. (in full) supported by U. S. Air Force Office of Scientific Research; of W. C. M. (in part) by the Office of Naval Research.
† Present address: Kwansei Gakuin University, Nishinomiya, Japan.
[1] G. I. Taylor, Proc. Roy. Soc. A151, 424 (1935) and further references which may be found in Ref. 2.
[2] G. K. Batchelor, *The Theory of Homogeneous Turbulence* (Cambridge University Press, New York, 1956).
[3] Experimental references on the distribution of $\mathbf{u}$, with interpretations, will be found in Ref. 2, Chap. VIII.
[4] Reference 2, Sec. 5.4.
[5] N. Wiener, *Fifth International Congress for Applied Mechanics* (John Wiley & Sons, Inc., New York, 1939).
[6] A. Siegel and W. C. Meecham, Bull. Am. Phys. Soc. Ser. II, 4, 197 (1959).
[7] By R. H. Kraichnan.
[8] A. Siegel, T. Imamura, and W. C. Meecham, "Wiener-Hermite Functional Expansion in Turbulence with the Burgers Model," Physics of Fluids 6, 1519 (1963).

the structure of the expansion, though not complicated, was so unfamiliar[9] that no physical interpretation was available to guide the intuition in handling it. The authors of this paper are publishing at this time a series of publications meant as pilot studies in the application of this method, which we consider of great potential value in the study of nonlinear stochastic equations, to turbulence theory. The first of these papers,[10] to be referred to here as IMS, is an exposition purely of the mathematical methods and is adapted to a range of physical applications more general than turbulence theory.

The present paper and another which appears elsewhere[11] are applications of the method in the regions of small and large Reynolds numbers, respectively. In keeping with the spirit of a pilot study, meant mainly to develop insight into the general structure of the method, it will be drastically simplified in three principal ways:

(a) We use the model equation of Burgers,[12]

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} - \nu(\partial^2 u/\partial x^2) = 0, \qquad (1.2)$$

(where u is a scalar function, x a scalar position variable, t is the time, and ν is a constant) instead of the Navier–Stokes equation. Besides making both the spatial variable and the field itself one instead of three dimensional, this equation eliminates the pressure term of the Navier–Stokes equation; yet the work of Moomaw[13] indicates that, with respect to statistical properties, the Burgers model behaves quite satisfactorily, and should therefore be a good vehicle for our study.

(b) We attempt a description only of the late decay stage of the random field. "Decay" means that the forces which have produced the motion of the fluid have been turned off; hence the mean velocity is decreasing due to dissipation. This means that we do not attempt to describe the way in which the universal features of the turbulence field arise from a wide variety of forces generating the flow. "Late" means that the mean velocity has decayed sufficiently to make the governing equations quasi-linear, i.e., nonlinear terms are treated as nominally small corrections although their behavior is of all-important interest to us as a test of the validity of the expansion. Model turbulence under conditions where the nonlinear terms are of crucial importance has been treated elsewhere by two of the present authors (Ref. 11).

(c) In solving our equations for the time development of the field, the initial conditions which define the solutions cannot be obtained entirely from physical or experimental considerations, but must be chosen partly according to criteria of simplicity and certain requirements of convergence.

In addition to the above special assumptions, we restrict ourselves to the case of spatial homogeneity.

II. EQUATIONS OF MOTION AND SOLUTION PROCEDURE

A. The Equations of Motion

We take the Wiener–Hermite series in the form

$$u(x, t) = \int K^{(1)}(x - x_1)H^{(1)}(x_1)\,dx_1 + \iint K^{(2)}(x - x_1, x - x_2)H^{(2)}(x_1, x_2)\,dx_1\,dx_2 + \cdots . \qquad (2.1)$$

The $K^{(i)}$, which are implicitly functions of t = time, are ordinary (nonrandom) functions. The absence of a $K^{(0)}$ term implies that the mean value of u vanishes, i.e., that we are moving with the mean velocity of the fluid. We assume the $K^{(i)}$ to be invariant with respect to all permutations of their spatial arguments, since no generality is lost thereby with respect to the moments of the distribution of u, which determine its probability distribution. The integrations are from $-\infty$ to $+\infty$. The $H^{(i)}$ are the Wiener–Hermite functionals defined in IMS. This series converges in the mean-square sense[14] for $u(x, t)$ continuous in x, provided the use of a continuous argument for the constructed functionals is justified in terms of the given problem (cf. IMS, Sec. VI.2). Its first term has a Gaussian distribution. The higher terms are *statistically orthogonal* to the first term. [Two random variables, A and B, are "statistically orthogonal" if they are uncorrelated, $\langle AB\rangle = 0$. The use of the former term puts emphasis on the equivalent role of the expectation value as an inner product, since if A and B are functions of a random variable x with probability density

[9] It is not unfamiliar in electrical engineering, of course, where it has already been in use for some years; for references, see N. Wiener, *Nonlinear Problems in Random Theory* (Technology Press, Cambridge, Massachusetts, and John Wiley & Sons, Inc., New York, 1958).

[10] T. Imamura, W. C. Meecham, and A. Siegel, J. Math. Phys. **6**, 695 (1965).

[11] W. C. Meecham and A. Siegel, Phys. Fluids **7**, 1178 (1964).

[12] J. M. Burgers, Verhandel. Koninkl. Ned. Akad. Wetenschap. Afdel. Natuurk. (Sec. I) **17**, No. 2 (1939); Proc. Acad. Sci. Amsterdam **53**, 247 (1950).

[13] D. W. Moomaw, *A Study of Burgers' Model Equation with Application to Statistical Theories of Turbulence*, Ph.D. Thesis, University of Michigan, Ann Arbor (1962).

[14] R. H. Cameron and W. T. Martin, Ann. Math. **48**, 385 (1947).

$P(x)$, $\langle AB\rangle = \int A(x)B(x)\cdot P(x)\,dx$. Note that statistically orthogonal variables need not be differently distributed; in fact, they can be identically distributed, yet completely uncorrelated.] This is the closest one can come, at least in a general description, to regarding them as the locus of the deviation of the distribution from Gaussianity; in fact, we shall see [v. the expression for the flatness factor, Eq. (3.2)] that the deviation from Gaussianity does not come purely from terms higher than the first, but involves interactions between the first term and the higher terms.

The remainder of this section consists of a summary of the method whereby a set of differential equations for the $K^{(i)}$ can be derived from the Burgers equation; details are relegated to the Appendix. Our actual applications require only the equations for the first three nonvanishing kernels, in Fourier-transformed form. The reader primarily interested in what we have done concretely to solve the Burgers equation statistically may then prefer to skim through the summarized derivation until he reaches the equations for the said three kernels, Eqs. (2.10)–(2.12).

Since the integrals in the above equation are convolutions, we can use the conventional notation whereby if f, g are functions of i variables,

$$f * g \equiv \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x - x_1, \cdots, x - x_i) \times g(x_1, \cdots, x_i)\,dx_1 \cdots dx_i. \tag{2.2}$$

Then Eq. (2.1) becomes

$$u = \sum_{i=1}^{\infty} K^{(i)} * H^{(i)}. \tag{2.3}$$

The dependence of the $K^{(i)}$ on x and $x_1, \cdots, x_i$ only through the differences $x - x_1, \cdots$ implies an assumption of statistical homogeneity of u in x, corresponding to the common assumption of homogeneity in turbulence.

An infinite set of equations for the time variation of the kernels $K^{(i)}$ can be derived by substituting the expansion for u into the Burgers equation (2) and equating the kernels of the mutually orthogonal terms of the resulting expression to zero. First introduce the notation u' for the linear part of the operation in Eq. (1.2),

$$u'(x, t) \equiv [\partial/\partial t - \nu(\partial^2/\partial x^2)]u(x, t), \tag{2.4}$$

whereby (2) becomes

$$u' + \tfrac{1}{2}(\partial/\partial x)u^2 = 0. \tag{2.5}$$

Substituting (2.3) into (2.5),

$$\sum_i K^{(i)\prime} * H^{(i)} + \tfrac{1}{2}\sum_{i,j}\left(\frac{\partial}{\partial x}[K^{(i)}K^{(j)}]\right) * (H^{(i)}H^{(j)}) = 0. \tag{2.6}$$

Multiplying throughout by $H^{(l)} = H^{(l)}(x_1^{(L)}, x_2^{(L)}, \cdots x_l^{(L)})$ and taking the expectation value, we obtain

$$\sum_i K^{(i)\prime} * \langle H^{(l)}H^{(i)}\rangle + \tfrac{1}{2}\sum_{i,j}\left(\frac{\partial}{\partial x}[K^{(i)}K^{(j)}]\right) * \langle H^{(l)}H^{(i)}H^{(j)}\rangle = 0 \qquad (l = 0, 1, 2, \cdots). \tag{2.7}$$

The convolutions in this expression are to be taken only between functions of like superscripts, thus there is no convolution over the arguments of $H^{(l)}$.

A formula for the expectation value of products of Wiener–Hermite functionals is given in IMS, Eq. (4.15). The convolution integrals can then be carried out. This work is done in detail in the Appendix, where a general result, Eq. (A1) or (A12), is derived. We work primarily with the Fourier transform of this equation. Let us write the Fourier transform of a function $f(x)$ (which may depend on other variables not here expressed) by simply replacing x by the Fourier-conjugate variable κ, viz.:

$$f(\kappa_1, \kappa_2, \cdots) \equiv \int e^{i(\kappa_1 x_1 + \kappa_2 x_2 + \cdots)} f(x_1, x_2, \cdots)\,dx_1\,dx_2 \cdots. \tag{2.8}$$

Then it is shown in the Appendix that the equation of motion for the Fourier transform $K^{(l)}(\boldsymbol{\kappa}^{(L)})$ of the kernel $K^{(l)}(\mathbf{x}^{(L)})$ is [Eq. (A16)]

$$\left[\frac{\partial}{\partial t} + \nu\left(\sum_{\lambda=1}^{l}\kappa_\lambda\right)^2\right]K^{(l)}(\kappa^{(l)}) - \frac{i}{2l!}\left(\sum_{\lambda=1}^{l}\kappa_\lambda\right)\sum_{i,j}^{*} C(l, i, j) \times \sum_{\text{app}}(K^{(i)}(-\kappa_1^{(L)}, \cdots -\kappa_{(i-j+l)/2}; \kappa^{(I+J-L)/2}), K^{(j)}(\kappa_{1+(i-j+l)/2}^{(L)}, \cdots \kappa_l^{(L)}; \kappa^{(I+J-L)/2})) = 0, \qquad l = 0, 1, 2, \cdots. \tag{2.9}$$

The summations implied by $\sum_{i,j}^{*}$ and $\sum_{\text{app}}$ are defined in detail in the Appendix. The inner product is defined by

$$(f(\mathfrak{n}, \xi), g(\zeta, \xi)) \equiv (2\pi)^{-\alpha}\int [f(\mathfrak{n}, \xi)]^* g(\zeta, \xi)\,d\xi$$

where the asterisk denotes complex conjugate, and α is the number of components of ξ.

The first four equations are as follows:

$l = 0$: In this case all terms vanish identically, either because of the hypothesis $K^{(0)} = 0$, or because the assumption of spatial homogeneity—the dependence of the $K^{(i)}$ of the series in its primary form, Eq. (2.1), on x only through the $x - x_1$, etc.—makes inner products in which all variables are integrated out, as must be the case with $l = 0$, constant in x. The latter reason is the one that holds for the nonlinear term of this equation, and is evidently necessary for the consistency of the equations of motion with $K^{(0)} = 0$.

For $l = 1$, 2, and 3, we have:

$$[\partial/\partial t + \nu\kappa_1^2]K^{(1)}(\kappa_1) = 2i\kappa_1(K^{(2)}\cdot K^{(1)}) + \cdots, \quad (2.10)$$

$$[\partial/\partial t + \nu(\kappa_1 + \kappa_2)^2]K^{(2)}(\kappa_1, \kappa_2) = (i/2)(\kappa_1 + \kappa_2)K^{(1)}(\kappa_1)K^{(1)}(\kappa_2) + i(\kappa_1 + \kappa_2)[3(K^{(3)}\cdot K^{(1)} + 2(K^{(2)}\cdot K^{(2)})] + \cdots, \quad (2.11)$$

$$[\partial/\partial t + \nu(\kappa_1 + \kappa_2 + \kappa_3)^2]K^{(3)}(\kappa_1, \kappa_2, \kappa_3) = (i/3)(\kappa_1 + \kappa_2 + \kappa_3)PK^{(1)}(\kappa_1)K^{(2)}(\kappa_2, \kappa_3) + i(\kappa_1 + \kappa_2 + \kappa_3)[4(K^{(1)}\cdot K^{(4)}) + 2P(K^{(2)}\cdot K^{(3)})] + \cdots. \quad (2.12)$$

These are all the equations we need for the present work. Nonlinear terms have been put on the right-hand side. Each of these equations contains an infinite number of terms, but those in which the sum of the superscripts of the K's in the inner product exceeds the superscript of the linear term by more than two have been omitted. The operator P in the $K^{(3)}$ equations sums the function following it over all three cyclic permutations of the variables κ_1, κ_2, and κ_3.

Inner products with a *dot* are to be integrated over only one variable:

$$(K^{(i)}(\kappa', -\kappa_1, -\kappa_2, \cdots, -\kappa_{i-1}) \cdot K^{(j)}(\kappa', \kappa_i, \kappa_{i+1}, \cdots, \kappa_{i+j-2})) = (2\pi)^{-1}\int [K^{(i)}(\kappa', -\kappa_1, \cdots, -\kappa_{i-1})]^* \times K^{(j)}(\kappa', \kappa_i, \cdots, \kappa_{i+j-2})\, d\kappa'. \quad (2.13)$$

which is evidently a function of $i + j - 2$ variables. In the inner products in the equations of motion the integrated variables and the disposable variables (those which are not integrated over) are not explicitly indicated. But there can be no ambiguity if one uses the above explicit formula for the inner product and inserts for $\kappa_1, \cdots \kappa_{i+j-2}$ the variables that appear on the left-hand side of the equation of motion concerned.

B. Initial Conditions

1. Form of the Initial Conditions

We seek solutions to the three equations of motion (2.10)–(2.12) which have the property of rapid convergence of the Wiener–Hermite series for $u(x, t)$. The convergence (more strictly, decreasingness of successive terms) depends quite strongly on the initial conditions. In this and the next section, we investigate the problem of the choice of initial conditions.

If the Wiener–Hermite series for u converges rapidly, one would also expect the right-hand sides of the equations of motion to fall off rapidly with increasing values of the superscripts involved. We make this also a condition on the $K^{(i)}$. As a simple way to formalize this condition, we suppose the terms on the right-hand side to decrease with the number of integrations in the inner product; another suitable number, which increases in a way necessarily correlated with the number of integrations, is, of course, the sum of the superscripts involved, but we shall see that it is the number of integrations that more directly causes the decrease.

The natural measure of smallness of these terms is the value of t, measured from the beginning of decay or any time thereafter, since the linearization and supposed increase of Gaussianity both develop with increase of time during the decay period. We shall see that a decrease of the inner products, as negative powers of t which are increasingly negative according to the number of integrations, can be realized as a rather natural outcome of the exponential form of the solutions of the equations of motion.

If the above is true, the equations of motion become in first approximation

$$\frac{\partial}{\partial t}K^{(1)}(\kappa) + \nu\kappa^2 K^{(1)}(\kappa) = 0, \quad (2.14)$$

$$\frac{\partial}{\partial t}K^{(2)}(\kappa_1, \kappa_2) + \nu(\kappa_1 + \kappa_2)^2 K^{(2)}(\kappa_1, \kappa_2) - \frac{i}{2}(\kappa_1 + \kappa_2)K^{(1)}(\kappa_1)K^{(1)}(\kappa_2) = 0. \quad (2.15)$$

$$\frac{\partial}{\partial t}K^{(3)}(\kappa_1, \kappa_2, \kappa_3) + \nu(\kappa_1 + \kappa_2 + \kappa_3)^2 K^{(3)}(\kappa_1, \kappa_2, \kappa_3) - \frac{i}{3}(\kappa_1 + \kappa_2 + \kappa_3)[K^{(1)}(\kappa_1)K^{(2)}(\kappa_2, \kappa_3) + K^{(1)}(\kappa_2)K^{(2)}(\kappa_3, \kappa_1) + K^{(1)}(\kappa_3)K^{(2)}(\kappa_1, \kappa_2)] = 0. \quad (2.16)$$

In our paper dealing with the regime of large Reynolds number (Ref. 11), the term $2i\kappa_1(K^{(2)}\cdot K^{(1)})$ on the right-hand side of Eq. (2.14) is retained. This is because of the emphasis in that paper on results involving the nature of the energy spectrum function; it is shown there that energy conservation, in an approximation based on the kernels $K^{(1)}$ and $K^{(2)}$, requires the retention of this term. However, the emphasis in this paper is more directly on the kernels themselves, and the term in question, being of higher order than those retained, is seen to result only in a negligible correction.

It will be a necessary condition on our solutions to these Eqs. (2.14)–(2.16) that, when they are substituted into the neglected terms, the values of the latter do indeed approach zero with increasing t relative to the terms retained.

The equations can be solved successively, since the inhomogeneous part of each equation depends only on the solution of the preceding one. In fact, we have here the basis for an iterative scheme which can, at least formally, lead to exact solutions, but we shall solve only to lowest order. The solutions will depend on the initial values of the kernels. We designate the initial values by the symbols $K_0^{(1)}$, $K_0^{(2)}$, and $K_0^{(3)}$.

Straightforwardly solving the successive equations in terms of arbitrary unspecified forms of the initial values of the functions allows us to characterize the forms of the right-hand sides more exactly, leading to the general form (for all three equations)

$$\left[\frac{\partial}{\partial t}+\nu\left(\sum_{\lambda=1}^{i}\kappa_\lambda\right)^2\right]K^{(i)}(\boldsymbol{\kappa},t) = \sum_r f_r^{(i)}(\boldsymbol{\kappa})\exp\left[-\nu g_r^{(i)}(\boldsymbol{\kappa})t-h_r^{(i)}(\kappa)\right],\quad i=1,2,3. \tag{2.17}$$

The right-hand side is the nonlinear term of the preceding equations, more exactly specified. The functions h_r depend on initial conditions, the g_r on the linear operators of the preceding equations, and the f_r on both. Writing the general result down at the outset in this way allows us to investigate the problem of initial conditions. These, as we have said previously, are to be chosen with an eye to physical reasonableness and so as to satisfy the requirements of decreasingness of the successive terms of the expansion of u and of the nonlinear terms of the equations of motion.

The complete solution of (2.17) is

$$K^{(i)} = K^{(i)}_{\text{comp}} + K^{(i)}_{\text{ptc}},$$

which is the sum of the complementary function

$$K^{(i)}_{\text{comp}}(\boldsymbol{\kappa},t) = K_0^{(i)}(\boldsymbol{\kappa})\exp\left[-\nu\left(\sum\kappa_\lambda\right)^2 t\right], \tag{2.18}$$

and the particular integral

$$K^{(i)}_{\text{ptc}}(\boldsymbol{\kappa},t) = \sum_r \exp\left[-h_r^{(i)}(\boldsymbol{\kappa})\right] \times\int_0^t \exp\left[-\nu\left(\sum\kappa_\lambda\right)^2(t-t')\right]f_r^{(i)}(\boldsymbol{\kappa}) \times\exp\left[-\nu g_r^{(i)}(\boldsymbol{\kappa})t'\right]dt' = \sum_r \frac{f_r^{(i)}(\boldsymbol{\kappa})\exp\left[-h_r^{(i)}(\boldsymbol{\kappa})\right]}{\nu\left[\left(\sum\kappa\right)^2-g_r^{(i)}(\boldsymbol{\kappa})\right]} \times\left\{\exp\left[-\nu g_r^{(i)}(\boldsymbol{\kappa})t\right]-\exp\left[-\nu\left(\sum\kappa\right)^2 t\right]\right\}. \tag{2.19}$$

The Gaussian form of the exponential function in the solution of the homogeneous equation, Eq. (2.18), with variance proportional to t^{-1}, along with the fact that the equation for $K^{(1)}$ is homogeneous, means that the time dependence of all the inhomogeneous terms in the equations is of the same Gaussian type (i.e., having variance proportional to t^{-1}), and therefore that all solutions are also linear combinations of functions of this type. The assumption that we are working in the decay stage therefore means that the initial values of the solutions contain multivariate Gaussian functions of the κ variables in every term. Moreover, with increasing t, any factor in the solution which converges less rapidly in $\boldsymbol{\kappa}$ than the Gaussian is dominated within the region of $\boldsymbol{\kappa}$ where the Gaussian is appreciable, by the lowest terms of its Taylor expansion.[4] Since we are working only in the region of late decay, we therefore assume a homogeneous combination of powers of the $\boldsymbol{\kappa}$ variables as the coefficient of the Gaussian term in the initial value of each K.

The above assumptions determine the initial value of $K^{(1)}$ as

$$K_0^{(1)}(\kappa) = A(i\kappa)^m e^{-l^2\kappa^2}, \tag{2.20}$$

where A is real; the factor i^m is inserted in order to make $K^{(1)}(x)$ a real function. The constant l is in the lowest approximation half the correlation length of $u(x)$ at $t=0$. [To see that $2l$ is the correlation length, we note that the expansion of the correlation function in terms of the K's is

$$\langle u(x)u(x+r)\rangle = (2\pi)^{-1}\int e^{-i\kappa r}\left|K^{(1)}(\kappa)\right|^2 d\kappa + 2(2\pi)^{-2} \times\int e^{-i(\kappa_1+\kappa_2)r}\left|K^{(2)}(\kappa_1,\kappa_2)\right|^2 d\kappa_1\, d\kappa_2 + 6(2\pi)^{-3} \times\int e^{-i(\kappa_1+\kappa_2+\kappa_3)r}\left|K^{(3)}(\kappa_1,\kappa_2,\kappa_3)\right|^2 d\kappa_1\, d\kappa_2\, d\kappa_3 + \cdots. \tag{2.21}$$

If this may be approximated by the first term, Eq. (2.20) gives

$$\langle u(x)u(x+r)\rangle \propto e^{-r^2/8l^2}, \quad \text{Q.E.D.}]$$

Thus our calculations are valid only at a stage late enough to be characterized in lowest approximation by a single correlation length. If the time origin is late enough in the decay period, l becomes proportional to $(\nu T)^{\frac{1}{2}}$, where T is of the order of the time from the onset of the single-correlation-length régime to the chosen origin [just as we shall find with the initial condition just given that the correlation length of the time-dependent solution is $(l^2 + \nu t)^{\frac{1}{2}} \to \nu t$ for $t \gg l^2/\nu$; this condition implies that t is large enough so that the large eddies have decayed.]

Now consider the second-order kernel. We have chosen as initial condition

$$K_0^{(2)}(\kappa_1, \kappa_2) = B[i(\kappa_1 + \kappa_2)]^n e^{-l^2(\kappa_1^2+\kappa_2^2)}. \quad (2.22)$$

In this case, the remarks preceding Eq. (2.21) do not determine a unique initial form for the function. They merely require that the polynomial factor be homogeneous in the κ_i and that the exponent be a quadratic form in the κ_i. The particular form above represents a simple and somewhat restricted choice, arrived at as follows: The form of the homogeneous operator for $K^{(2)}$ suggests that $(\kappa_1 + \kappa_2)$ replaces κ when we go from $K^{(1)}$ to $K^{(2)}$; this leads to the polynomial factor $(\kappa_1 + \kappa_2)^n$. The quadratic form in the exponent, on the other hand, cannot simply be $(\kappa_1 + \kappa_2)^2$ since this would not allow integrals over κ_1 and κ_2 of $K_0^{(2)}$ or its powers to converge [see Eq. (2.25)]; such convergence requires that the exponent contain the sum of the squares of two independent functions of κ_1 and κ_2 with positive coefficients (which are squared correlation lengths). The choice made is, again, simple but quite restrictive. In particular, it involves only a single convergence length, which has been made equal to that of the $K^{(1)}$ kernel.

For the third-order kernel we have chosen, in strict analogy to $K^{(2)}$,

$$K_0^{(3)}(\kappa_1, \kappa_2, \kappa_3) = C[i(\kappa_1 + \kappa_2 + \kappa_3)]^p e^{-l^2(\kappa_1^2+\kappa_2^2+\kappa_3^2)}. \quad (2.23)$$

With the initial condition (2.21) the solution of (2.14) is

$$K^{(1)}(\kappa; t) = A(i\kappa)^m e^{-(l^2+\nu t)\kappa^2}. \quad (2.24)$$

With the initial condition (2.22), the complementary function of Eq. (2.15) is

$$K_{\text{comp}}^{(2)}(\kappa_1, \kappa_2; t) = B[i(\kappa_1 + \kappa_2)]^n e^{-l^2(\kappa_1^2+\kappa_2^2)-\nu t(\kappa_1+\kappa_2)^2}, \quad (2.25)$$

showing the necessity for incorporating convergence factors with respect to two independent variables into the initial condition. From (2.23), we have

$$K_{\text{comp}}^{(3)}(\kappa_1, \kappa_2, \kappa_3; t) = C[i(\kappa_1 + \kappa_2 + \kappa_3)]^p e^{-l^2(\kappa_1^2+\kappa_2^2+\kappa_3^2)-\nu t(\kappa_1+\kappa_2+\kappa_3)^2}. \quad (2.26)$$

2. Choice of m and n

The choice of the exponents m and n in the initial values of the kernels is determined by the decay and the convergence of the series representing the mean square velocity in terms of the $K^{(i)}$. We have

$$\langle u^2 \rangle = (K^{(1)}, K^{(1)}) + 2(K^{(2)}, K^{(2)}) + 6(K^{(3)}, K^{(3)}) + \cdots; \quad (2.27)$$

where we put a comma in the inner product to indicate integration over *all* variables [compare with Eq. (2.13)]:

$$(K^{(m)}, K^{(m)}) = \int K^2(x_1, x_2, \cdots x_m)\, dx_1 \cdots dx_m$$

$$= (2\pi)^{-m} \int |K(\kappa_1, \kappa_2, \cdots \kappa_m)|^2\, d\kappa_1 \cdots d\kappa_m. \quad (2.28)$$

We can obtain a necessary condition for convergence by considering the contributions of $K^{(1)}(\kappa; t)$ and of $K_{\text{comp}}^{(2)}(\kappa_1, \kappa_2; t)$ to $\langle u^2 \rangle$ individually and comparatively. With (2.24) and (2.25) we find

$$(K^{(1)}, K^{(1)}) = (\text{const})\, A^2(\nu t + l^2)^{-(m+\frac{1}{2})}, \quad (2.29)$$

$$(K_{\text{comp}}^{(2)}, K_{\text{comp}}^{(2)}) = (\text{const})\, B^2 l^{-1}(\nu t + l^2/2)^{-(n+\frac{1}{2})}. \quad (2.30)$$

If $\langle u^2 \rangle$ is to decay, both these terms must decay individually, and we have

$$m > -\tfrac{1}{2}, \qquad n > -\tfrac{1}{2}. \quad (2.31)$$

If $(K^{(2)}, K^{(2)})$ is not to dominate $(K^{(1)}, K^{(1)})$ in the course of time, we must have

$$n \geqq m. \quad (2.32)$$

Another condition can be obtained by considering the particular solution of (2.15). Direct inversion of the operator $\partial/\partial t + \nu(\kappa_1 + \kappa_2)^2$ gives

$$K_{\text{pto}}^{(2)} = -(iA^2/4\nu)(\kappa_1\kappa_2)^{m-1}(\kappa_1 + \kappa_2)\, e^{-l^2(\kappa_1^2+\kappa_2^2)} \times [e^{-\nu t(\kappa_1^2+\kappa_2^2)} - e^{-\nu t(\kappa_1+\kappa_2)^2}]. \quad (2.33)$$

Changing to variables $\xi_1 = (\nu t)^{\frac{1}{2}}\kappa_1$, etc., we have

$$(K_{\text{pto}}^{(2)}, K_{\text{pto}}^{(2)}) = \text{const}\, (\nu t)^{-2m} \iint (\xi_1\xi_2)^{2(m-1)}(\xi_1 + \xi_2)^2 \times e^{-(2l^2/\nu t)(\xi_1^2+\xi_2^2)}[e^{-(\xi_1^2+\xi_2^2)} - e^{-(\xi_1+\xi_2)^2}]^2\, d\xi_1\, d\xi_2. \quad (2.34)$$

Consider now the smallest integer value of m permitted by (2.31) (we do not consider the more complicated possibility of noninteger values of m and n, since it is possible to solve the convergence problem without them), namely zero. The integral

in (2.34) does converge in this case (the integrand can be seen to be nonsingular, since the bracketed term approaches $\xi_1\xi_2$ as $\xi_1\xi_2$ approaches zero), but it approaches a constant as t approaches infinity. Thus it contributes a nondecaying value to $\langle u^2\rangle$, not cancellable by any other contribution, which is not permissible.

Our final choice for m, which is justified by later calculations, is the next integer, $m = 1$. We also take $n = 1$, which is consistent with (2.32). It might seem as if this will fail to provide increasing Gaussianity with time, since it makes $(K^{(2)}_{\text{comp}}, K^{(2)}_{\text{comp}})$ which is the dominant part of $(K^{(2)}, K^{(2)})$, decay at the same rate as $(K^{(1)}, K^{(1)})$ asymptotically in time. However, the vanishing of $(K^{(2)}, K^{(2)})$ relative to $(K^{(1)}, K^{(1)})$ is not a necessary condition for Gaussianity. We adopt as criterion of Gaussianity the value 3 for the flatness factor (to be discussed in more detail in Sec. III), admittedly not a sufficient condition in the precise sense, but one which we consider sufficient in terms of the precision required in this context. And it will turn out that the choice of initial conditions made here does lead to a flatness factor that approaches 3 asymptotically.

The role of $K^{(3)}$ can be considered in a similar fashion. However, we simply postulate for the exponent in K_0^3 that $p = 1$, which is analogous to the choice $n = 1$ for $K^{(2)}$. For future reference, we write down here the particular solution for $K^{(3)}$:

$$K^{(3)}_{\text{ptc}} = -(i/6\nu)(\kappa_1 + \kappa_2 + \kappa_3)\, e^{-l^2(\kappa_1^2+\kappa_2^2+\kappa_3^2)}$$
$$\times [A(B + A^2/4\nu)P e^{-\nu t(\kappa_1^2+[\kappa_2+\kappa_3]^2)} - A(A^2/4\nu + 3B)$$
$$\times e^{-\nu t(\kappa_1+\kappa_2+\kappa_3)^2} - (A^3/2\nu)\, e^{-\nu t(\kappa_1^2+\kappa_2^2+\kappa_3^2)}]. \quad (2.35)$$

As with $K^{(2)}_{\text{ptc}}$, this is obtained by direct inversion of the linear operator of the equation of motion. P is a symmetrizing operator as in Eq. (2.12).

C. Self-Consistency of the Approximation Scheme

The self-consistency of our iteration method of constructing solutions requires that the terms neglected in the equations of motion be small when evaluated through substitution of the zero-order solutions into these equations. We do not have a completely general proof of this as yet, but we show in this section that at least the leading omitted terms are small, and are so for a reason which indicates that the terms beyond them diminish successively more and more strongly.

The order of decrease is, as mentioned before, postulated to be determined by the number of integrations or, secondarily, by the sum of the superscripts, in a term. By this criterion, the largest neglected term in the equation for $K^{(1)}(\kappa)$ is $2i\kappa(K^{(1)}(\kappa')\cdot K^{(2)}(\kappa', \kappa))$. Thus we consider

$$\kappa(K^{(1)}(\kappa')\cdot K^{(2)}(\kappa', \kappa))$$
$$= \text{const}\,\kappa \int \kappa'(\kappa + \kappa')\, e^{-(\nu t + l^2)(2\kappa'^2+\kappa^2)}$$
$$\times [-1+ae^{-2\nu t(\kappa\kappa')}]\, d\kappa' < \text{const}\, e^{-\frac{3}{2}\nu t\kappa^2}[\text{const}\, t^{-\frac{3}{2}}\kappa e^{-\frac{1}{2}\nu t\kappa^2}$$
$$+ \text{const}\, t^{-\frac{3}{2}}\kappa + \text{const}\, t^{-1}\kappa^2 + \text{const}\, t^{-\frac{1}{2}}\kappa^3], \quad (2.36)$$

where

$$a = 1 + 4\nu B/A^2. \quad (2.37)$$

The $<$ sign in (2.36) is obtained by discarding l^2 in the result of the integration.

The form of $(K^{(1)}\cdot K^{(2)})$ as displayed in (2.36) is such that the desired negligibility compared to $K^{(1)}$ cannot possibly be of a pointwise nature. This is because of the more slowly-decreasing (with κ) exponential in its three last terms, which makes them larger than $K^{(1)}$ for large κ. [It follows that our representation of the energy spectrum converges the more poorly, the larger the wavenumber.] However, the observable results being treated here are not the individual values of the $K^{(i)}$, but integrals of products of them. Hence, the criterion of magnitude of a term is more accurately taken as the integral of its square—essentially its average behavior over the domain of κ. This recognizes that the important region of κ, due to the Gaussian factor, is that in which its values are of order $(\nu t)^{-\frac{1}{2}}$. Since κ may be measured in units $(\nu t)^{-\frac{1}{2}}$, it may be inferred from (2.36) by simple dimensional considerations that the integral of the square of this nominally leading term of the omitted part of the $K^{(1)}$ equation behaves, for large t, homogeneously as $t^{-9/2}$. This is to be compared with a value of $t^{-7/2}$ for the second term in Eq. (2.14), and is seen to provide the necessary "convergence" for the omitted term according to the assumed criterion.

As a check on this result, one may easily find the first iterative correction to $K^{(1)}$ by substituting the solution of (2.14) into the first omitted term, treat it as an inhomogeneous term in the differential equation, and evaluate its contribution to $K^{(1)}$ by applying the inverse of the operator on the left-hand side of (2.14) to it. When the result is combined with the zeroth iterate, the added contribution to the integrated square behaves as $t^{-\frac{1}{2}}$ times the integrated square of the zeroth iterate alone.

The results of the last two paragraphs may be summed up by saying that the first omitted term in the equation of motion of $K^{(1)}$ behaves as $t^{-\frac{1}{2}}$ times the term retained, and when treated as an

inhomogeneous term in the equation gives rise to a contribution to the solution which behaves as $t^{-\frac{1}{2}}$ times the zeroth iterate. This conclusion would also follow if each κ factor in any expression were treated as implicitly proportional to $t^{-\frac{1}{2}}$, as would follow by dimensional reasoning from the form of the exponent of the Gaussian.

We now turn to the omitted terms of the $K^{(2)}$ equation, and find the integrated-square estimate for their magnitude. Formally, the largest of these terms are, omitting irrelevant constants,

$$(\kappa_1 + \kappa_2)(K^{(2)}(-\kappa_1, \kappa')\cdot K^{(2)}(\kappa_2, \kappa')) \equiv F(\kappa_1, \kappa_2) \quad (2.38)$$

and

$$(\kappa_1 + \kappa_2)(K^{(1)}(\kappa')\cdot K^{(3)}(\kappa', \kappa_1, \kappa_2)) \equiv G(\kappa_1, \kappa_2). \quad (2.39)$$

For reasons which will appear, it is necessary to tie our considerations to the restricted applications of this paper, which consist of the calculation of the inner products $(K^{(i)}, K^{(i)})$, where $i = 1, 2, 3$; $((K^{(1)}\cdot K^{(2)})\cdot(K^{(1)}\cdot K^{(2)}))$; and $(K^{(1)}K^{(1)}K^{(1)}, K^{(3)})$.

The complete solution of Eq. (2.15) is, for large t,

$$K^{(2)}(\kappa, \kappa') \sim i(A^2/4\nu)(\kappa + \kappa')[-e^{-(\nu t + l^2)(\kappa^2 + \kappa'^2)} + ae^{-(\nu t + l^2)(\kappa^2 + \kappa'^2) - 2\nu t\kappa\kappa'}, \quad (2.40)$$

where a is the same as in (2.34). Considering now

$$(\kappa_1 + \kappa_2)^2 K^{(2)}(\kappa_1, \kappa_2) \equiv H(\kappa_1, \kappa_2), \quad (2.41)$$

which appears in the second term of Eq. (2.15), the two terms in brackets of (2.40) contribute different values to the integrated square $\int H^2(\kappa_1, \kappa_2)d\kappa_1\, d\kappa_2$: The first gives t^{-4}, while the second gives $t^{-7/2}$. This is because the first term of (2.40) is Gaussian with variance proportional to t^{-1} with respect to *both* of any *two* independent linear combinations of κ_1 and κ_2, while the second is so only with respect to $\kappa_1 + \kappa_2$, while being Gaussian with variance proportional only to l^2 with respect to $\kappa_1 - \kappa_2$. Thus the retained terms in Eq. (2.15) decrease as $t^{-7/2}$ in the integrated-square sense. The largest terms of $F(\kappa_1, \kappa_2)$ and $G(\kappa_1, \kappa_2)$ in this sense turn out to decrease as $t^{-9/2}$, hence their decreasingness is satisfactory as long as the integrated square is an appropriate measure. This is the case for the mean-square value of $u(x, t)$, to which $K^{(2)}$ contributes as $(K^{(2)}, K^{(2)})$. On the other hand, $K^{(2)}$ contributes to the flatness factor through $((K^{(1)}\cdot K^{(2)})\cdot(K^{(1)}\cdot K^{(2)}))$. In this inner product the presence of the $K^{(1)}$ factors provides Gaussian variances proportional to t^{-1} for both variables, hence the mean magnitudes of correction terms should be estimated as if such variances were present in all integrals. From this point of view we find a controlling magnitude of t^{-3} for the retained term $H(\kappa_1, \kappa_2)$, and the same is true for the third term of (2.15); while F and G give t^{-4}. Hence this criterion again gives satisfactory decreasingness.

The second term in Eq. (2.40) has a mean square which behaves, itself, as $t^{-\frac{1}{2}}$, which is larger than that of the first term. Thus it may appear that for order-of-magnitude consistency, the first term should be dropped. However, we have just seen that the two terms contribute in the same order in t to the flatness factor, hence both must be retained.

A similar treatment of the omitted term $2iP(\kappa_1 + \kappa_2 + \kappa_3)(K^{(2)}\cdot K^{(3)})$ of the $K^{(3)}$ equation leads to the same quantitative results. We have nothing to say as yet about $(K^{(1)}\cdot K^{(4)})$; in the present treatment we consider the first three kernels as a closed system to the required approximation, assuming that further work will bear this out.

We are now in a position to indicate the essential meaning of our result. The method given shows an underlying feature which not only accounts for its success, but indicates that it should continue to give satisfactory results when applied to the higher neglected terms. This feature is that, dimensionally, each integration over a κ variable behaves as a factor $t^{-\frac{1}{2}}$, provided this variable has a variance proportional to t^{-1} in the Gaussian factor (which always turns out to be the case). Thus the integrated square of such an integral has a factor t^{-1} due to such an integration. The higher neglected terms all have two, three, ... integrations over κ variables. If these behave like those we have already considered, they will indeed be asymptotically small in t according to the successively increasing sums of their indices, which are necessarily correlated with the number of κ integrations in the respective terms. This is the reasoning which we promised to justify, at the end of the first paragraph of Sec. IIB.1.

III. CALCULATION OF THE FLATNESS-FACTOR DEVIATION

A. The Flatness Factor as a Test of the Validity of the Expansion

1. General Discusison

A very striking experimental manifestation of the non-Gaussian nature of the velocity field in real turbulent flow is the rapid increase of the flatness factor of the nth derivative of the velocity field with n. The flatness factor of a random variable x which has zero mean is defined as the ratio of its fourth moment to the square of its second moment:

$$\Phi[x] = \langle x^4\rangle/\langle x^2\rangle^2. \quad (3.1)$$

(The square brackets in $\Phi[x]$ indicate that Φ, which is of course not a function of x, is a functional of x if x is regarded as a function of its parameter of distribution.)

For a Gaussian x, $\Phi[x]$ is exactly 3. Although having this value is not sufficient to make a distribution Gaussian, since there are an infinite number of higher moments which must in a Gaussian distribution obey analogous but mutually independent conditions for their relations to the second moment, this condition is a sufficient one if the distribution is approximated through its first four moments.

The Wiener–Hermite series is by its very form capable of giving separate expression, through its higher terms, to the deviations of a random function from Gaussianity (in the next section we take note of certain restrictions on this capability). In our hierarchy of equations the higher terms are determined by the Gaussian term. More completely stated, the following is true: The higher terms are determined by: (1) *The form of the lowest term,* which is arbitrary within wide limits. $K^{(1)}$ might conceivably be almost any function, it need not necessarily be κ times a Gaussian, although this is what we have assumed. (2) *The initial conditions on the kernels,* whose influence is never wiped out, no matter how large t becomes. This is seen in our calculations of the flatness factor, in which the particular solution for any kernel contributes to the same order as the complementary function. Our "initial conditions," of course, are applied *at a time origin not earlier than the onset of late decay.* They therefore embody any characteristics the flow pattern may have acquired during the creation of the turbulence by the applied forces, which have been turned off well before our initial instant. If the universal equilibrium hypothesis is true, our initial conditions are determined by the form of this univeral equilibrium, provided it has had time to set in by the time the late decay (quasilinear) stage begins. It is interesting that this influence of the form of the flow, universal or not, at the time of onset of late decay, persists indefinitely thereafter, through the constants a and b [see Eq. (3.29)]. (3) *The form of the Burgers equation.*

What influence do the above factors have on our results? The hope on which this work is based is that *the flatness factors we compute do strongly reflect the influence of the form of the Burgers equation, and that the Burgers equation in turn is for this purpose a good facsimile of the Navier–Stokes equation and thence of experimental reality.* The built-in presuppositions expressed in (1) and (2) above are potential sources of artifacts which might interfere with this program. As for (1), we have already supplied the justification of this standard assumption (see Footnote 4). With regard to (2), we have already given the arguments for the naturalness of the general form of the initial conditions used, in Sec. IIB.2. There remains, however, within this general form, the aforementioned arbitrariness of a and b. The fixing of these constants must wait on the determination of the form of the Wiener–Hermite kernels in the universal equilibrium stage, a problem to which we address ourselves elsewhere (Ref. 11, and further work in progress).

Having stated these reservations, we proceed on the assumption that our methods do provide an interesting test of the ability of the Wiener–Hermite expansion to reproduce the experimental effect of the rapid increase of the flatness factor of the nth derivative of the velocity field with n. From Eq. (5.15) of IMS, the flatness factor of the nth derivative of $u(x, t)$ is to first approximation

$$\Phi[u^{(n)}] \cong 3 + 24[(K^{(1)(n)}K^{(1)(n)}K^{(1)(n)}, K^{(3)}) + 2((K^{(1)(n)} \cdot K^{(2)(n)}) \cdot (K^{(1)(n)} \cdot K^{(2)(n)}))] \times (K^{(1)(n)}, K^{(1)(n)})^{-2} \quad (3.2)$$

where

$$K^{(i)(n)} \equiv (\partial/\partial x)K^{(i)}(x - x_1, \cdots x - x_i; t). \quad (3.3)$$

It should perhaps be mentioned that the flatness factor we are considering is that of the distribution of the function $u^{(n)}(x, t)$ at a single point x, or, in the language of statistics, that of the "marginal distribution" of $u^{(n)}(x, t)$. Thus the question of the Gaussianity of the *joint* distribution, as such, of the values of $u^{(n)}(x, t)$ for different points x, is not directly involved. However, the result is very strongly dependent on this joint distribution indirectly, particularly on the joint distribution for two points in the neighborhood of one another, since the derivative is essentially the difference between the values of $u^{(n)}(x, t)$ at two infinitesimally separated points.

2. Effect of Transformations of the Probability Space on the Flatness Factor

The present section concerns a matter of a rather special nature, but one of sufficient importance to warrant its interpolation at this point. Namely: a Gaussian random function does not always have a Wiener–Hermite expansion whose terms vanish beyond the second order.

We can prove this, at least up to a point, as follows.

It will be recalled that $H^{(1)}(x)$ represents a vector in a certain function space. Let us represent this vector for the sake of a more descriptive notation by the symbol ξ. The infinite-dimensional space of ξ is called its sample space (Wiener space, or "differential space"). The fact that it is Gaussianly distributed may be expressed by the statement that its probability density may be "visualized" as a mass density having a univariate Gaussian distribution along any straight line in the probability space (this means that we are also using the fact that ξ_i and ξ_j are independent if $i \neq j$). In fact, we may say more, because all components of ξ are identically distributed and have zero means; thus the probability density is also hyperspherically symmetrical about the origin. Now, ξ itself is represented by a Wiener–Hermite series containing only one term,

$$\xi = \int \delta(x - x')H^{(1)}(x')\,dx' = H^{(1)}(x). \qquad (3.4)$$

We can construct a function of ξ which is exactly Gaussian but, at least formally, has an expansion not at all restricted to the form of a linear functional of $H^{(1)}(x)$, as follows: Let the ξ that lie in any hyperspherical surface centered at the origin be subjected to a rigid rotation with the surface, but with the amount of rotation to vary (continuously, if we wish) with distance from the origin. Denote the mapping so obtained by $\mathbf{f}(\xi)$, which is a function of ξ. Since the density at all points of a hyperspherical surface is constant, $\mathbf{f}(\xi)$ has an exactly Gaussian distribution. But due to the varying amount of rotation with distance from the origin, $\mathbf{f}$ is not at all a linear function of ξ, and its Wiener–Hermite expansion will contain terms of all orders.

The above is, so far as we now know, not a proof but only a conjecture, for the reason that we have not taken into account the well-known effect by which the probability in Wiener space is found to be infinitely concentrated in the neighborhood of a single spherical surface. If only one value of the radius matters, the above transformation might, in effect, be no different from a rigid rotation of the entire space by the value of the rotation at the surface of concentration of the probability. In this event, the transformation would be linear, and $\mathbf{f}(\xi)$ representable in the form $\int K(x, x')H^{(1)}(x')\,dx'$.

If the conjecture *were* true, what consequences would it have for our work? No matter how close to multivariate Gaussian a function might be, its series might converge very slowly. Thus there is no guarantee that our criterion of decreasingness is satisfied by actual solutions of the Burgers equation unless they are somehow transformed, even when they are close to Gaussian. If convergence is poor, the second term on the right-hand side of Eq. (3.2) may be large even for a marginally Gaussian variable. On the other hand, if convergence is rapid, the random function *is* approximately multivariate Gaussian, and the second term on the right-hand side of (3.2) is a reliable index of the degree of deviation from Gaussianity.

The above discussion gives some information about the extent to which our work, from a fundamental mathematical point of view, tests the reliability of the Wiener–Hermite series as applied to flow equations of the type of the Navier–Stokes equation. By looking for a rapidly converging series, we may be restricting ourselves to a special class of nearly-Gaussian solutions.

B. Calculation of the Flatness Factor Deviations

In this section, we outline the various steps in evaluating the integrals necessary for finding the flatness factor. The inner products are all calculated in the wavenumber representation.

From Eqs. (A11) and (A13), $\partial/\partial x$ is represented by $-i(\kappa_1 + \kappa_2 + \cdots)$ when operating on functions of wavenumber variables. We are interested only in the asymptotic form of the flatness factor, for which $\nu t \gg l^2$. We can then write down the asymptotic formulas

$$K^{(1)(n)} \sim (-1)^n i^{n+1} A\kappa^{n+1} e^{-t\kappa^2}, \qquad (3.5)$$

$$K^{(2)(n)} \sim (-1)^n i^{n+1}(A^2/4\nu)(\kappa_1 + \kappa_2)^{n+1} \times [-e^{-\nu t(\kappa_1^2+\kappa_2^2)} + a\, e^{-\nu t(\kappa_1+\kappa_2)^2}], \qquad (3.6)$$

$$K^{(3)(n)} \sim (-1)^n i^{n+1}(A^3/24\nu^2)(\kappa_1 + \kappa_2 + \kappa_3)^{n+1} \times [2\, e^{-\nu t(\kappa_1^2+\kappa_2^2+\kappa_3^2)} - aP\, e^{-\nu t(\kappa_1^2+[\kappa_2+\kappa_3]^2)} + b\, e^{-\nu t(\kappa_1+\kappa_2+\kappa_3)^2}]. \qquad (3.7)$$

In these equations, a is as before [Eq. (2.41)], while b is given by

$$b = 1 + 12\nu B/A^2 + 24\nu^2 C/A^3. \qquad (3.8)$$

The inner products in the flatness factor are simply variations of a single type of integral: the integrand is the product of a polynomial in several variables with an exponential function of a positive-definite quadratic form in the same variables. The methods used to evaluate these integrals are described in some detail in the Appendix, Sec. 2. Here we merely state the results.

One obtains immediately

$$(K^{(1)(n)}, K^{(1)(n)}) \sim 2^{-n-\frac{3}{2}}(n + \tfrac{1}{2})!\, A^2(\nu t)^{-n-\frac{3}{2}}. \qquad (3.9)$$

The other inner products involved in the flatness factor contain three terms each. We put

$$I_1^{(n)} = \int [\kappa_1\kappa_2(\kappa_1+\kappa_3)(\kappa_2+\kappa_3)]^{n+1} \times e^{-2\nu t(\kappa_1^2+\kappa_2^2+\kappa_3^2)}\, d\kappa_1\, d\kappa_2\, d\kappa_3, \tag{3.10}$$

$$I_2^{(n)} = \int [\kappa_1\kappa_2(\kappa_1+\kappa_3)(\kappa_2+\kappa_3)]^{n+1} \times e^{-2\nu t(\kappa_1^2+\kappa_2^2+\kappa_3^2+\kappa_3\kappa_3)}\, d\kappa_1\, d\kappa_2\, d\kappa_3, \tag{3.11}$$

$$I_3^{(n)} = \int [\kappa_1\kappa_2(\kappa_1+\kappa_3)(\kappa_2+\kappa_3)]^{n+1} \times e^{-2\nu t(\kappa_1^2+\kappa_2^2+\kappa_3^2+\kappa_1\kappa_3+\kappa_2\kappa_3)}\, d\kappa_1\, d\kappa_2\, d\kappa_3. \tag{3.12}$$

Then

$$((K^{(1)(n)}\cdot K^{(2)(n)})\cdot(K^{(1)(n)}\cdot K^{(2)(n)})) \sim (A^6/16\nu^2)[I_1^{(n)} - 2aI_2^{(n)} + a^2I_3^{(n)}]. \tag{3.13}$$

Also we define

$$I_4^{(n)} = \int [\kappa_1\kappa_2\kappa_3(\kappa_1+\kappa_2+\kappa_3)]^{n+1} \times e^{-2\nu t(\kappa_1^2+\kappa_2^2+\kappa_3^2)}\, d\kappa_1\, d\kappa_2\, d\kappa_3, \tag{3.14}$$

$$I_5^{(n)} = \int [\kappa_1\kappa_2\kappa_3(\kappa_1+\kappa_2+\kappa_3)]^{n+1} \times e^{-\nu t[2\kappa_1^2+\kappa_2^2+\kappa_3^2+(\kappa_2+\kappa_3)^2]}\, d\kappa_1\, d\kappa_2\, d\kappa_3, \tag{3.15}$$

$$I_6^{(n)} = \int [\kappa_1\kappa_2\kappa_3(\kappa_1+\kappa_2+\kappa_3)]^{n+1} \times e^{-\nu t[\kappa_1^2+\kappa_2^2+\kappa_3^2+(\kappa_1+\kappa_2+\kappa_3)^2]}\, d\kappa_1\, d\kappa_2\, d\kappa_3, \tag{3.16}$$

in terms of which we have

$$(K^{(1)(n)}K^{(1)(n)}K^{(1)(n)}, K^{(3)(n)}) \sim (-1)^n(A^6/24\nu^2)[-2I_4^{(n)} + 3aI_5^{(n)} - bI_6^{(n)}]. \tag{3.17}$$

The results for the $I^{(n)}$ are as follows (summations are over positive integer values, subject to such further restrictions as may be indicated in individual cases):

$$I_1^{(n)} = (2\nu t)^{-2n-7/2} \sum_{g_1,g_2=0}^{[(n+1)/2]} \binom{n+1}{2g_1}\binom{n+1}{2g_2} \times (n+\tfrac{1}{2}-g_1)!\,(n+\tfrac{1}{2}-g_2)!\,(g_1+g_2-\tfrac{1}{2})!, \tag{3.18}$$

in which $[(n+1)/2]$ stands for the largest integer not greater than $(n+1)/2$.

$$I_2^{(n)} = 2^{-n-\frac{1}{2}}3^{-n/2-1}(\nu t)^{-2n-7/2} \sum_{c,d_1,d_2=0}{}' \binom{n+1}{c}\binom{n+1}{d_1,\,d_2} \times (-1)^{d_1}3^{-\frac{1}{2}(c+d_2)}(\tfrac{1}{2}[2n+1-d_1-d_2])! \times (\tfrac{1}{2}[n-c+d_1])!\,(\tfrac{1}{2}[n+c+d_2])!, \tag{3.19}$$

in which $\binom{n+1}{d_1,\,d_2}$ stands for the trinomial coefficient $(n+1)!/d_1!d_2!(n+1-d_1-d_2)!$, and the prime on the summation sign indicates the restrictions

$$n \text{ even } \begin{cases} c \text{ even} : d_1, d_2 \text{ odd}, \\ c \text{ odd} : d_1, d_2 \text{ even}, \end{cases} \tag{3.20a}$$

$$n \text{ odd} : c, d_1, d_2 \text{ all odd or all even}. \tag{3.20b}$$

$$I_3^{(n)} = 2^{-2n-3}(\nu t)^{-2n-7/2} \sum_{c,d=0}^{n+1} \binom{n+1}{c}\binom{n+1}{d}(-2)^{-(c+d)} \times (n-c+\tfrac{1}{2})!\,(n-d+\tfrac{1}{2})!\,(c+d-\tfrac{1}{2})!. \tag{3.21}$$

$$I_4^{(n)} = (2\nu t)^{-2n-7/2} \sum_{c_2,c_3}{}' \binom{n+1}{c_2,\,c_3}\left(\frac{[2n+1-c_2-c_3]}{2}\right)! \times \left(\frac{[n+c_2]}{2}\right)!\left(\frac{[n+c_3]}{2}\right)!, \tag{3.22}$$

in which the prime on the summation sign stands for the restrictions

$$n \text{ even} : c_2, c_3 \text{ both odd}, \tag{3.23a}$$

$$n \text{ odd} : c_2, c_3 \text{ both even}, \tag{3.23b}$$

$$c_2 + c_3 \leq n+1. \tag{3.23c}$$

$$I_5^{(n)} = 2^{-2n-5/2}3^{-1/2}(\nu t)^{-2n-7/2} \times \sum_{c,d=0}^{n+1}{}' \binom{n+1}{c}\binom{n+1}{d}(-1)^{n-d+1}2^c3^{-c/2-d} \times \left(\frac{[2n-c+1]}{2}\right)!\left(\frac{[2n-2d+1]}{2}\right)!\left(\frac{[c+2d-1]}{2}\right)!, \tag{3.24}$$

where the prime on the summation sign indicates that c is to be summed over even values only.

$$I_6^{(n)} = 2^{-3n-4}(\nu t)^{-2n-7/2} \sum_{r,s=0}^{n+1} \binom{n+1}{r}\binom{n+1}{s}(-1)^{r+s}2^{r-s} \times (n+\tfrac{1}{2}-r+s)!\,(n+\tfrac{1}{2}-s)!\,(r-\tfrac{1}{2})!. \tag{3.25}$$

The quantity of interest to us is the deviation of the flatness factor from the Gaussian value of 3. We refer to this quantity for brevity as the flatness-factor deviation, and denote its value for the nth derivative by the symbol ϕ_n. From Eq. (3.2) we have

$$\phi_n = 24[(K^{(1)(n)}K^{(1)(n)}K^{(1)(n)}, K^{(3)}) + 2((K^{(1)(n)}\cdot K^{(2)(n)}), (K^{(1)(n)}\cdot K^{(2)(n)}))] \times (K^{(1)(n)}, K^{(1)(n)})^{-2}. \tag{3.26}$$

Defining

$$\bar{I}_i^{(n)} = I_i^{(n)}/I_0^{(n)}, \tag{3.27}$$

where

$$I_0^{(n)} = (K^{(1)(n)}\cdot K^{(1)(n)})^2/A^4, \tag{3.28}$$

we have

$$\phi_n = \pi^{\frac{1}{2}} A^2 \nu^{-2} (\nu t)^{-\frac{1}{2}} (\phi_{n0} + a\phi_{n1} + a^2\phi_{n2} + b\phi_{n3}), \quad (3.29)$$

where

$$\phi_{n0} = (\nu t/\pi)^{\frac{1}{2}} [3\bar{I}_1^{(n)} - 2(-1)^n \bar{I}_4^{(n)}], \quad (3.30)$$

$$\phi_{n1} = (\nu t/\pi)^{\frac{1}{2}} [-6\bar{I}_2^{(n)} + 3(-1)^n \bar{I}_5^{(n)}], \quad (3.31)$$

$$\phi_{n2} = (\nu t/\pi)^{\frac{1}{2}} \cdot 3\bar{I}_3^{(n)}, \quad (3.32)$$

$$\phi_{n3} = (\nu t/\pi)^{\frac{1}{2}} (-1)^{n+1} I_6^{(n)}. \quad (3.33)$$

The quantities ϕ_{n0}, ϕ_{n1}, ϕ_{n2}, ϕ_{n3} have been defined to be time independent, and the entire asymptotic time dependence of ϕ_n is exhibited in Eq. (3.29) through the explicit factor $t^{-\frac{1}{2}}$ which appears there. This rate of decay of the non-Gaussian part agrees with that found by Moomaw,[13] although he finds a different dependence on the Reynolds number. Equation (3.29) also explicitly reveals the dependence of the flatness-factor deviation on initial conditions, since these are expressed entirely through the coefficients a and b. The values of $\phi_{n0} \cdots \phi_{n3}$ for n equal to 0, 1, 2, and 3 are given in Table I (some errors in these results as presented in Ref. 7 have been corrected here). It will be seen that in the main they have a tendency to increase in magnitude with n, but that this tendency is not entirely a consistent one. The behavior of ϕ_n itself with increasing n is evidently dependent on initial conditions, and it will not increase with n under all circumstances.

One obvious test that can be made is to ascertain the behavior of ϕ_n in the simple case where $u(x, t)$ consists initially entirely of a Gaussian term: Then $a = b = 1$; the values of ϕ_n, with constant factors and the $t^{-\frac{1}{2}}$ dependence removed, are shown in Table II. The increase with n is extremely rapid. To what extent do our results mimic the characteristics of real turbulence? According to the data of Batchelor and Townsend,[15] in real turbulence ϕ_0 is slightly negative, while ϕ_1, ϕ_2, and ϕ_3 are strongly positive and increasing in the order given. Since in the case $a = b = 1$ of our calculations ϕ_0 is a small number arising from the linear combination

TABLE I. Values of the ϕ_{nj}.

n	ϕ_{n0}	ϕ_{n1}	ϕ_{n2}	ϕ_{n3}
0	2.12	−4.90	2.25	0.75
1	5.66	−9.45	3.18	2.28
2	6.70	−8.94	2.97	4.46
3	14.44	−12.8	3.12	10.0

[15] G. K. Batchelor and A. A. Townsend, Proc. Roy. Soc. London A199, 238 (1949); Ref. 2 of this paper, p. 174.

TABLE II. Values of ϕ_n for $a = b = 1$.

n	$(\nu^2/A^2)(\nu t/\pi)^{\frac{1}{2}}\phi_n$
0	0.22
1	1.67
2	5.19
3	14.8

of numbers which are large relative to itself, only a small change in the coefficients—say, a slight increase in a, or a slight decrease in b—would make ϕ_0 slightly negative but leave the others positive and strongly increasing with n. It will also be noted that with the result we have obtained, the conditions for *positive* flatness-factor deviation (apart from the slightly negative $n = 0$ case) and for *increase with n*, two features found in the experiments, display a strong mutual consistency.

Since the comparison of our results with experiment is on a speculative basis as yet, we should sum up the limitations of our accomplishment in the present calculation: (1) We have not provided an intuitive basis for the phenomenon reproduced. (2) We have not investigated alternate choices for the algebraic factors in the initial forms of the $K^{(i)}$. (3) We do not yet know how to evaluate the constants a and b, or their counterparts in a more general theory. (4) We have oversimplified the matter of inferring the properties of the Burgers model from experimental results in real turbulence. Each of the above limitations suggests an avenue of future investigation, which we intend to explore.

ACKNOWLEDGMENTS

The early thinking which underlies our work was done during the year 1957–1958, during which the first- and third-named authors were at the University of Michigan. In particular, the first-named author (A.S.) was on leave at that time as a Guggenheim Fellow, and wishes to thank the John Simon Guggenheim Memorial Foundation for its support, and the Physics Department of the University of Michigan for its hospitality during this period.

We wish to thank Norman Goldman of the Computing Center of Boston University for some of the numerical calculations leading to the results in Table III.

APPENDIX

1. Equations of Motion of the Wiener-Hermite Kernels

The expectation values in Eq. (2.7) are given in terms of products of δ functions in IMS Eq. (4.15).

We prove that this allows us to transform it into

$$l!\,K^{(l)\prime}(\mathbf{x}-\mathbf{x}^{(L)}) + \frac{1}{2}\frac{\partial}{\partial x}\sum_{i,j}{}^{*}\, C(l,i,j) \times \sum_{\text{app}} \int K^{(i)}(x-x_1^{(L)},\cdots,x-x_{(i-j+l)/2}^{(L)}; \times \mathbf{x}-\mathbf{x}^{(I+J-L)/2})K^{(j)}(x-x_{1+(i-j+l)/2}^{(L)},\cdots x_l^{(L)}; \times \mathbf{x}-\mathbf{x}^{(I+J-L)/2})\,d\mathbf{x}^{(I+J-L)/2}=0, \quad l=0,1,2,\cdots. \tag{A1}$$

The terms in this expression are defined as follows: $\mathbf{x}-\mathbf{x}^{(L)}$ symbolizes a set of l scalar variables, $x-x_1^{(L)},\cdots,x-x_l^{(L)}$; $\mathbf{x}-\mathbf{x}^{(I+J-L)/2}$ an analogously constructed set of $(i+j-l)/2$ variables $x-x_1^{(I+J-L)/2},\cdots x-x_{(i+j-l)/2}^{(I+J-L)/2}$. $\sum_{\text{app}}$ is the sum over all ways of *apportioning* the individual scalar variables of $\mathbf{x}-\mathbf{x}^{(L)}$ over two functions $K^{(i)}$ and $K^{(j)}$, independently of order or position within the two functions. $C(l,i,j)$ is the number of distinct arrangements, for a given apportionment, obtainable by permuting the variables within the individual functions. The asterisk on the sum over i and j signifies that the sum is restricted to values of i and j that satisfy

$$i+j\geq l\geq |i-j| \qquad p(i+j)=p(l), \tag{A2}$$

where p is the parity of the indicated arguments.

Equation (A1) is proved as follows: The functional forms of the expectation values of products of H's, such as appear in Eq. (2.7), are given in Eq. (4.15) of IMS, which in the case of continuous variables only reads

$$\langle H^{(\alpha)}(R_1,R_2,\cdots R_\alpha)H^{(\beta)}(R_{\alpha+1},\cdots R_{\alpha+\beta})\cdots H^{(\zeta)} \times(\cdots R_{\alpha+\beta+\cdots+\zeta})\rangle = \sum_{\substack{\text{distinct}\\ \text{exogamous}\\ \text{pairings}}} \prod_{\substack{\text{exogamous}\\ \text{pairs }(i,j)}} \delta(R_i-R_j), \tag{A3}$$

the product over "exogamous pairs (i,j)" being such that every R on the left side of the equation appears just once in a δ function, and the two R's in any δ function come from different H's. The sum is over distinct ways of pairing R's in this exogamous fashion; two pairings are distinct if they cannot be made identical by a reordering of R's which does not divorce a pair. The expectation value *vanishes* if no such exogamous pairings exist.

In the case of the expectation value of the product of two H's in Eq. (2.7) it is easy to write down a simplified form for (2.8). Let us write $x_1^{(I)}, x_2^{(I)}, \cdots, x_i^{(I)}$ for the several variables of $H^{(i)}$, with a parallel notation for those of $H^{(j)}$, etc. These variables will also be denoted, collectively, by $x^{(I)}$, etc. It is evident by inspection that

$$\langle H^{(l)}(\mathbf{x}^{(L)})H^{(i)}(\mathbf{x}^{(I)})\rangle = 0 \qquad (l\neq i)$$
$$= \sum_{\text{perm }(\mathbf{x}^{(L)})} \delta(\mathbf{x}^{(L)}-\mathbf{x}^{(I)}) \qquad (l=i), \tag{A4}$$

"perm $\mathbf{x}^{(L)}$" meaning permutations of the individual variables of $x^{(L)}$. Hence, in view of the symmetry of the $K^{(i)}$, the first linear term of (2.7) is

$$l!\,K^{(l)\prime}(\mathbf{x}-\mathbf{x}^{(L)}), \tag{A5}$$

where $\mathbf{x}-\mathbf{x}^{(L)}$ stands for the set of variables $x-x_1^{(L)}, x-x_2^{(L)},\cdots,x-x_l^{(L)}$.

In dealing with the triple product of H's, no simple expression is possible because of notational difficulties, and it seems more efficient to reason directly from (A3) to the expression, analogous to (A5), which is obtained for the nonlinear term after integrating out the δ functions. In every exogamous pairing generated by the triple-product expectation value, l of the I and J variables are paired with L variables. The integration over these, due to the δ function, makes them equal to their L-variable mates. The remaining $(i+j-l)/2$ links of the pairing are between I and J variables; integrating over, say, the J variables that are linked with I variables, the δ function simply makes them equal to their I-variable mates; the integration over the latter then gives an inner product of the two kernels with respect to these variables, which, since there are $(i+j-l)/2$ of them, we denote collectively by $\mathbf{x}^{(I+J-L)/2}$.

As a result of the integrations, then, the number of L variables in $K^{(i)}$ is i minus $(i+j-l)/2$, the number of I variables paired with J variables, or $(i-j+l)/2$. Similarly, the number of L variables in $K^{(j)}$ is $(j-i+l)/2$.

It follows from the above that one contribution to the nonlinear term of (2.7) is

$$\frac{1}{2}\frac{\partial}{\partial x}\int K^{(i)}(x-x_1^{(L)},\cdots x-x_{(i-j+l)/2}^{(L)};\mathbf{x}-\mathbf{x}^{(I+J-L)/2}) \times K^{(j)}(x-x_{1+(i-j+l)/2},\cdots x-x_l^{(L)}; \times \mathbf{x}-\mathbf{x}^{(I+J-L)/2})\,d\mathbf{x}^{(I+J-L)/2}, \tag{A6}$$

provided i and j are values such that this pairing is possible. The entire nonlinear term will consist of a sum over such possible values of i and j, further summed, for i and j given, over permutations of the variables which give other permissible pairings.

One condition on i and j is that $(i-j+l)/2$, $(j-i+l)/2$, and $(i+j-l)/2$ all be integers in order for the above pairing to be possible. For this it is sufficient that

$$p(i+j)=p(l), \tag{A7}$$

where p is the parity of the indicated argument. Moreover, these three quantities must be nonnegative, for which it is necessary and sufficient that $i + j \geqq l \geqq |i - j|$. This completes the conditions on the sum over i and j.

The set of pairings of the variables may be classified according to the L variables that are assigned to $K^{(i)}$ and to $K^{(j)}$. We call such an assignment, apart from the positions the L variables occupy in the functions, an "apportionment." Clearly, different apportionments give different results, except when i and j are equal, and there will have to be a sum over apportionments, $\sum_{\text{app}}$. For a given apportionment, any further permutations of the variables obtained by other variations of the pairing does not change the result, because of the symmetry of the kernels, since such permutations only vary the positions of the variables within a kernel; thus their effect is to multiply the given term by a constant $C(l, i, j)$. We do not bother to evaluate this constant in general because we are concerned here only with a few sets of small values of l, i and j, for which its value is readily obtainable by inspection.

The final result obtained for the nonlinear term is then

$$\begin{aligned}&\frac{1}{2}\frac{\partial}{\partial x}\sum_{i,j} C(l, i, j)\\ &\times \sum_{\text{app}} \int K^{(i)}(x - x_1^{(L)}, \cdots, x - x_{(i-j+l)/2}^{(L)};\\ &\times \mathbf{x} - \mathbf{x}^{(I+J-L)/2}) K^{(j)}(x - x_{1+(i-j+l)/2}^{(L)}, \cdots x - x_l^{(L)};\\ &\times \mathbf{x} - \mathbf{x}^{(I+J-L)/2})\, d\mathbf{x}^{(I+J-L)/2} \qquad \text{(A9)}\end{aligned}$$

the sum over i and j being such that

$$i + j \geq l \geq |i - j| \qquad p(i + j) = p(l). \qquad \text{(A10)}$$

This completes the proof of Eq. (A1).

Since x appears in Eq. (A1) everywhere in a difference $\mathbf{x} - \mathbf{x}^{(E)}$, it is actually superfluous and can be eliminated: The integration variables $\mathbf{x} - \mathbf{x}^{(I+J-L)/2}$ may be replaced by $\mathbf{x}^{(I+J-L)/2}$, while leaving the integration element unchanged, without affecting the value of the integral. Also the variables $\mathbf{x} - \mathbf{x}^{(L)}$ may be replaced by $\mathbf{x}^{(L)}$; under this transformation $\partial/\partial x$ transforms as follows:

$$\begin{aligned}\frac{\partial}{\partial x} f(\mathbf{x} - \mathbf{x}^{(L)}) &= \sum_{\lambda=1}^{l} \frac{\partial}{\partial(x - x_\lambda)} f(\mathbf{x} - \mathbf{x}^{(L)})\\ &\to \sum_{\lambda=1}^{l} \frac{\partial}{\partial x_\lambda} f(\mathbf{x}^{(L)}).\end{aligned} \qquad \text{(A11)}$$

The equations of motion then become

$$\begin{aligned}&\left[\frac{\partial}{\partial t} - \nu\left(\sum_{\lambda=1}^{l} \frac{\partial}{\partial x_\lambda^{(L)}}\right)^2\right]\left[K^{(l)}(\mathbf{x}^{(L)}) + \frac{1}{2l!}\left(\sum_{\lambda=1}^{l} \frac{\partial}{\partial x_\lambda^{(L)}}\right)\right.\\ &\times \sum_{i,j}{}^{*}\, C(l, i, j) \sum_{\text{app}} \int K^{(i)}(x_1^{(L)}, \cdots, x_{(i-j+l)/2}^{(L)};\\ &\times \mathbf{x}^{(I+J-L)/2}) K^{(j)}(x_{1+(i-j+l)/2}^{(L)}, \cdots x_l^{(L)};\\ &\times \mathbf{x}^{(I+J-L)/2})\, d\mathbf{x}^{(I+J-L)/2} = 0. \qquad \text{(A12)}\end{aligned}$$

We prefer to work with these equations in their Fourier-transformed versions. With later needs in mind, we introduce here a notation and formalism for this purpose, which is somewhat more general than our present purposes require. The Fourier conjugate variable of any x is denoted by κ, with super- and subscripts to match the x; the κ's will be referred to as "wavenumber variables." The Fourier transformation of a function with respect to any of its variables is expressed by simply replacing the variables in question by their conjugates. If $f(\mathbf{x}^{(A)}, \mathbf{x}^{(B)})$ has Fourier transform $f(\boldsymbol{\kappa}^{(A)}, \mathbf{x}^{(B)})$ with respect to its A variables, we use the Fourier transformation in the form

$$\begin{aligned}f(\boldsymbol{\kappa}^{(A)}, \mathbf{x}^{(B)}) &= \int e^{i\mathbf{x}^{(A)}\cdot\mathbf{x}^{(A)}} f(\mathbf{x}^{(A)}, \mathbf{x}^{(B)})\, d\mathbf{x}^{(A)};\\ f(\mathbf{x}^{(A)}, \mathbf{x}^{(B)}) &= (2\pi)^{-\alpha} \int e^{-i\mathbf{x}^{(A)}\cdot\mathbf{x}^{(A)}} f(\boldsymbol{\kappa}^{(A)}, \mathbf{x}^{(B)})\, d\mathbf{x}^{(A)}.\end{aligned} \qquad \text{(A13)}$$

We also require the Parseval theorem for the integral of a product of two functions over some but not necessarily all of their variables, namely

$$\begin{aligned}\int [f(\mathbf{y}, \mathbf{x})]^* g(\mathbf{z}, \mathbf{x})\, d\mathbf{x} &= (2\pi)^{-(\beta+\gamma)} \int e^{-i(\mathbf{y}\cdot\mathbf{n}+\mathbf{z}\cdot\boldsymbol{\zeta})}\\ &\quad \times (f(-\mathbf{n}, \boldsymbol{\xi})\cdot g(\boldsymbol{\zeta}, \boldsymbol{\xi}))\, d\mathbf{n}\, d\boldsymbol{\zeta}\end{aligned} \qquad \text{(A14)}$$

(asterisk denotes complex conjugate of the function within the brackets), where $\boldsymbol{\xi}$, $\mathbf{n}$, $\boldsymbol{\zeta}$ are Fourier-conjugate variables to $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{z}$, and β, γ are the numbers of components of $\mathbf{n}$, $\boldsymbol{\zeta}$, respectively. The inner product is defined by

$$\begin{aligned}(f(\mathbf{n}, \boldsymbol{\xi})\cdot g(\boldsymbol{\zeta}, \boldsymbol{\xi})) &\equiv (2\pi)^{-\alpha} \int [f(\mathbf{n}, \boldsymbol{\xi})]^* g(\boldsymbol{\zeta}, \boldsymbol{\xi})\, d\boldsymbol{\xi}\\ &= \int [f(\mathbf{n}, \mathbf{x})]^* g(\boldsymbol{\zeta}, \mathbf{x})\, d\mathbf{x},\end{aligned} \qquad \text{(A15)}$$

where α is the number of components of $\boldsymbol{\xi}$, in general, but the presence of the minus sign in front of $\mathbf{n}$ in the inner product of Eq. (A14) is to be especially noted.

With these definitions we obtain

$$\left[\frac{\partial}{\partial t}+\nu\left(\sum_{\lambda=1}^{l}\kappa_\lambda\right)^2\right]K^{(l)}(\kappa^{(L)})-\frac{i}{2l!}\left(\sum_{\lambda=1}^{l}\kappa_\lambda\right)$$
$$\times\sum_{i,j}{}^{*}\,C(l,i,j)\sum_{\text{app}}(K^{(i)}(-\kappa_1^{(L)}\cdots-\kappa_{(i-j+l)/2}^{(L)};$$
$$\kappa^{(I+J-L)/2})\cdot K^{(j)}(\kappa_{1+(i-j+l)/2},\cdots,\kappa_l^{(L)};$$
$$\kappa^{(I+J-L)/2}))=0,\qquad l=0,1,2,\cdots. \tag{A16}$$

2. Evaluation of Integrals Contributing to the Flatness Factor

The integrations of (3.10)–(3.12) and (3.14)–(3.16) leading to the expressions (3.18)–(3.25) were done in a variety of ways, which are briefly described here.

$I_1^{(n)}$, $I_2^{(n)}$, $I_4^{(n)}$, $I_5^{(n)}$ were done as follows: First, diagonalize the quadratic form in the exponent, if not already diagonal, by the appropriate orthogonal transformation of variables. Next, write the integral in terms of a sum of integrals over the positive ranges $(0,\infty)$ only, by changing the signs of integration variables where necessary. Then transform to variables x_i which are proportional to the squares of the variables in the diagonalized exponent obtained in the first step, with proportionality constants so chosen as to make the exponent equal to $-(x_1+x_2+x_3)$. The coefficient of the exponential in the integrand is then a product of multinomials in the x_i, and these are then expanded by the multinomial theorem. The terms resulting from the latter expansion are then evaluated by the formula

$$\int_0^\infty x^k e^{-x}\,dx=k!. \tag{A17}$$

Only odd half integer values of k occur.

In the cases of $I_3^{(n)}$ and $I_6^{(n)}$, the individual multinomial factors were rather laborious, and for the sake of simplicity, the diagonalization was done by completing the square in the exponential (giving a nonorthogonal transformation) instead of by a principal-axis transformation.

The net transformations which turn the exponential into $-(x_1+x_2+x_3)$ are as follows:

$$I_2^{(n)},\ I_4^{(n)}:\ (\nu t)^{\frac12}\kappa_1=2^{-\frac12}x_1^{\frac12},$$
$$(\nu t)^{\frac12}\kappa_2=2^{-\frac12}x_2^{\frac12}+2^{-\frac12}3^{-\frac12}x_3^{\frac12}, \tag{A18}$$
$$(\nu t)^{\frac12}\kappa_3=-2^{-\frac12}x_2^{\frac12}+2^{-\frac12}3^{-\frac12}x_3^{\frac12};$$
$$I_3^{(n)}:\ (\nu t)^{\frac12}\kappa_1=2^{-\frac12}x_1^{\frac12}-2^{-1}x_3^{\frac12},$$
$$(\nu t)^{\frac12}\kappa_2=2^{-\frac12}x_2^{\frac12}-2^{-1}x_3^{\frac12}, \tag{A19}$$
$$(\nu t)^{\frac12}\kappa_3=2^{-1}x_3^{\frac12};$$
$$I_6^{(n)}:\ (\nu t)^{\frac12}\kappa_1=2^{-\frac12}x_1^{\frac12}-2^{-1}x_3^{\frac12},$$
$$(\nu t)^{\frac12}\kappa_2=2^{-1}x_3^{\frac12}+2^{-\frac12}x_2^{\frac12}, \tag{A20}$$
$$(\nu t)^{\frac12}\kappa_3=2^{-1}x_3^{\frac12}-2^{-\frac12}x_2^{\frac12}.$$

In going over to positive ranges of integration, the effect is as follows: Let $f(x_1^{\frac12},x_2^{\frac12},x_3^{\frac12})$ be the integrand of $I_j^{(n)}$ in terms of the x variables (not including the Jacobian). Then

$$I_j^{(n)}=\int_0^\infty\int_0^\infty\int_0^\infty[f(x_1^{\frac12},x_2^{\frac12},x_3^{\frac12})$$
$$+f(x_1^{\frac12},x_2^{\frac12},-x_3^{\frac12})+f(x_1^{\frac12},-x_2^{\frac12},x_3^{\frac12})$$
$$+f(-x_1^{\frac12},x_2^{\frac12},x_3^{\frac12})+f(x_1^{\frac12},-x_2^{\frac12},-x_3^{\frac12})$$
$$+f(-x_1^{\frac12},x_2^{\frac12},-x_3^{\frac12})+f(-x_1^{\frac12},-x_2^{\frac12},x_3^{\frac12})$$
$$+f(-x_1^{\frac12},-x_2^{\frac12},-x_3^{\frac12})]J\,dx_1\,dx_2\,dx_3, \tag{A21}$$

where $J=\partial(\kappa_1,\kappa_2,\kappa_3)/\partial(x_1,x_2,x_3)$.

The values of the $I_j^{(n)}$ are awkward, lengthy products. We have evaluated them for n from zero to three. The results are given in Table III, in terms of the quantities $\pi^{-\frac12}\sum_j'^{(n)}$, where the $\sum_j'^{(n)}$ are defined as the summations only in the above equations for the $I_j^{(n)}$; for example,

$$\sum_1'^{(n)}=(2\nu t)^{2n+7/2}I_1^{(n)}.$$

In Table III, the values given in terms of products of powers of whole numbers were evaluated by hand. Those given in decimal form were evaluated, from the formulas given, on the Boston University Computing Center 1620 computer.

TABLE III. Values of $\pi^{-\frac12}\sum'_j{}^{(n)}$.

n \ j	1	2	3	4	5	6
0	2^{-2}	$2^{-2}\cdot3^{-\frac12}$	$3\cdot2^{-4}$	0	$-2^{-3}\cdot3^{-2}$	$-9\cdot2^{-9/2}$
1	$2^{-3}\cdot3^2$	$2^{-3}\cdot3^{-1}\cdot13$	$2^{-8}\cdot3^2\cdot17$	$2^{-4}\cdot3^2$	$2^{-1}\cdot3^{-2}\cdot13$	$3^4\cdot2^{-9}\cdot17$
2	$2^{-6}\cdot3^2\cdot41$	$2^{-5}\cdot3^{\frac12}\cdot17$	3.4871	$2^{-7}\cdot3^4$	$-2^{-2}\cdot17$	-33.29
3	$2^{-5}\cdot3^2\cdot5\cdot181$	42.031	44.575	$2^{-8}\cdot3^4\cdot5\cdot37$	74.722	903

Algebraic Tabulation of Clebsch–Gordan Coefficients of SU_3 for the Product $(\lambda, \mu) \otimes (1, 1)$ of Representations of SU_3*

J. G. KURIYAN,† D. LURIÉ,‡ AND A. J. MACFARLANE†
Department of Physics and Astronomy, University of Rochester, Rochester, New York
(Received 7 October 1964)

An algebraic tabulation is made of the Clebsch–Gordan (CG) coefficients of SU_3 which occur in the reduction into irreducible representations of the direct product $(\lambda, \mu) \otimes (1, 1)$ of irreducible representations of SU_3. Full explanation is made of the method of handling the complications associated with the possible double occurrence of the representation (λ, μ) itself in the direct product. The phase convention employed is an explicitly stated generalization of the well-known Condon and Shortley phase convention for SU_2. The relationship of the CG coefficients associated with the direct product $(1, 1) \otimes (\lambda, \mu)$ to those coefficients already mentioned is also exhibited.

1. INTRODUCTION

THE purpose of this paper is to offer a manageable and internally consistent algebraic tabulation of certain Clebsch–Gordan (CG) coefficients of SU_3 for use in connection with the octet version[1] of the unitary symmetry theory. The CG coefficients we are concerned with are those that occur in the reduction of the direct product representation[2]

$$(\lambda, \mu) \otimes (1, 1), \tag{1.1}$$

of SU_3. The product (1.1) contains the following irreducible representations[3] (IR) of SU_3:

(I) $(\lambda+2, \mu-1)$ once unless $\mu=0$;
(II) $(\lambda-1, \mu-1)$ once unless $\lambda=0$ or $\mu=0$;
(III) $(\lambda-2, \mu+1)$ once unless $\lambda=0$ or $\lambda=1$;
(IV) $(\lambda+1, \mu+1)$ once;
(V) $(\lambda-1, \mu+2)$ once unless $\lambda=0$;
(VI) $(\lambda+1, \mu-2)$ once unless $\mu=0$ or $\mu=1$;
(VII) (λ, μ) twice unless $\lambda=0$ or $\mu=0$; once if $\lambda=0$ and $\mu\neq0$ or if $\lambda\neq0$ and $\mu=0$; not at all if $\lambda=\mu=0$. (1.2)

Our notation for the corresponding CG coefficients of SU_3 is a generalization

$$C(\lambda\mu\ 11\ \lambda'\mu'\gamma;\ IMY\ JNZ\ I'M'Y'), \tag{1.3}$$

for SU_3 of the notation $C(j_1j_2j;\ m_1m_2m)$ employed by Rose[4] for CG coefficients of SU_2. In (1.3), the labels λ' and μ' refer to the seven cases listed under (1.2). The labels I, M, and Y are respectively the isospin, z-component of isospin, and hypercharge labels of states of (λ, μ); the labels JNZ and $I'M'Y'$ refer similarly to (1, 1) and (λ', μ'), respectively. The label γ is effective only in Case (VII), where it is associated with the double appearance of the IR (λ, μ) in the product (1.1).[5] Its precise significance is as follows. Whenever (λ, μ) occurs twice in the product, it follows that we can construct two independent sets of orthonormal basis states $|\lambda\mu IMY\rangle$ out of products of basis states of (λ, μ) and (1, 1). Clearly we can arrange by use of the Schmidt procedure that the two sets of states be mutually orthogonal. We employ the notation $|\lambda\mu IMY\gamma\rangle$ where γ is a label that takes on two values, $\gamma = 1$ and 2 say, for such mutually orthogonal sets of basis states, and the notation

$$C(\lambda\mu\ 11\ \lambda\mu\gamma;\ IMY\ JNZ\ I'M'Y') \tag{1.4}$$

for the corresponding mutual orthogonal sets of CG coefficients. However, there remains considerable arbitrariness in the definition of such coefficients. We can make an arbitrary orthogonal transformation with respect to γ on some given sets of states $|\lambda\mu\ IMY\ \gamma\rangle$ and thereby obtain new sets of states of the same type and hence new sets of coefficients of the type (1.4). We resolve this arbitrariness by a method suggested by Biedenharn.[6] Herein, one demands that one of the sets of CG coefficients (1.4)

* Research supported in part by the U. S. Atomic Energy Commission.
† Present address: Department of Physics, Syracuse University, Syracuse, New York.
‡ Present address: Institute for Advanced Studies, Dublin, Ireland.
[1] Y. Ne'eman, Nucl. Phys. **26**, 222 (1961); M. Gell-Mann, Phys. Rev. **125**, 1067 (1962); S. Okubo, Prog. Theoret. Phys. (Kyoto) **27**, 944 (1962).
[2] We use the highest weight notation for IRs of SU_3. See, for example, A. J. Macfarlane, E. C. G. Sudarshan, and C. Dullemond, Nuovo Cimento **30**, 845 (1963).
[3] D. Lurié and A. J. Macfarlane, J. Math. Phys. **5**, 565 (1964). We refer to this paper as LM in what follows.
[4] M. E. Rose, *Elementary Theory of Angular Momentum* (John Wiley & Sons, Inc., New York, 1957).
[5] For a general discussion of the definition of CG coefficients for groups for which the direct product of two IRs may contain certain IRs more than once in its reduction, see M. Hamermesh, *Group Theory* (Addison-Wesley Publishing Company, Reading, Massachusetts, 1962), Chap. 5. See also A. J. Macfarlane, N. Mukunda, and E. C. G. Sudarshan, J. Math. Phys. **5**, 575 (1964).
[6] L. C. Biedenharn, Phys. Letters **3**, 254 (1963).

yield the matrix elements of the generators of SU_3 in the IR (λ, μ) interpreted in the light of the Wigner–Eckart theorem for SU_3. We attach the label $\gamma = 1$ to the set of CG coefficients (1.4) constructed according to this criterion, and determine the $\gamma = 2$ set (to within a phase) by orthogonality. Since Case (VII) involves two sets, it follows from (1.2) that we have to obtain and display tables of eight sets of CG coefficients (1.3) of SU_3. It can be proved, however, that the CG coefficients (1.3) factorize[7] according to

$$C(\lambda\mu\ 11\ \lambda'\mu'\gamma;\ IMY\ JNZ\ I'M'Y') = C(IJI';\ MNM')\ U(\lambda\mu\ 11\ \lambda'\mu'\gamma;\ IY\ JZ\ I'Y'), \tag{1.5}$$

where $C(IJI';\ MNM')$ refers to the isospin SU_2-subgroup of SU_3, and the isoscalar factor $U(\cdots)$ is independent of M, N, and M'. Hence, CG coefficient of SU_2 being well-known, we need only tabulate isoscalar factors.

Of vital importance in connection with our tables is the provision of a clear and complete statement of the manner in which we have disposed of the arbitrariness of phase that exists in the definition of the CG coefficients (1.3). Since our phase convention is essentially a generalization of the Condon and Shortley phase convention for CG coefficients of SU_2, it is convenient to begin by giving Messiah's statement[8] thereof. This involves the two requirements:

(a) $\langle j\ m+1\ |J_+|\ j\ m\rangle$, real and nonnegative,

(b) $C(j_1j_2j;\ j_1j-j_1j)$, real and positive.

Requirement (a) fixes the relative phases of the set of all CG coefficients of SU_2 of the form $C(j_1j_2j;\ m_1m_2m)$ for any nontrivial ordered set of allowed values j_1, j_2, and j, and requirement (b) then fixes the absolute phase of one canonically selected member of each such set. Similarly a full statement of our phase convention for CG coefficients of SU_3 divides into two parts, the first relating to relative phases of the set of all CG coefficients with a given nontrivial ordered triple of IR labels and the second fixing the absolute phase of one suitably selected member of each such set. For the first part of our phase convention, we adhere to the phase convention of Biedenharn[9] regarding the matrix elements of the generators[10] of SU_3. This involves two requirements:

(A-1)

$\langle\lambda\mu\ IM+1Y\ |I_+|\ \lambda\mu\ IMY\rangle$,

real and nonnegative, or equivalently (1.6)

$\langle\lambda\mu\ IM+1Y\ |I_{+1}|\ \lambda\mu\ IMY\rangle = C(I1\ I;\ M1\ M+1)\ (fI)$,

with $f(I)$ real and nonnegative;

(A-2)

$\langle\lambda\mu\ I\pm\tfrac{1}{2}\ M+\tfrac{1}{2}\ Y\ |F_+|\ \lambda\mu\ IMY\rangle = C(I\tfrac{1}{2}\ I+\tfrac{1}{2};\ M\tfrac{1}{2}\ M+\tfrac{1}{2})\ f_\pm(\lambda\mu\ IY)$, (1.7)

with $f_\pm(\lambda\mu IY)$ real nonnegative functions of the indicated arguments. It is important to observe that (A-2) differs from the demand[11] that the matrix elements of F_+ all be real and nonnegative, since $C(I\tfrac{1}{2}\ I-\tfrac{1}{2};\ M\tfrac{1}{2}\ M+\tfrac{1}{2})$ is nonpositive. To give our generalization for SU_3 of statement (b) for SU_2, we must consider the CG coefficient

$$C(\lambda\mu\ 11\ \lambda'\mu'\gamma;\ \bar{I}\bar{I}\bar{Y}\ JNZ\ \bar{I}'\bar{I}'\bar{Y}'), \tag{1.8}$$

where $\bar{I}\bar{I}\bar{Y}$ refer to the highest weight states of (λ, μ), i.e., $\bar{I} = \frac{1}{2}(\lambda+\mu)$, $\bar{Y} = \frac{1}{3}(\lambda-\mu)$, and $\bar{I}'\bar{I}'\bar{Y}'$ refer to the highest weight state of (λ', μ'). In each of the cases (I) to (VI) above, there is only one value of J for which the CG coefficient (1.8) does not vanish identically.[12] Further, in each case, the CG coefficient (1.3) with this value of J does not vanish accidentally[12] for any (λ, μ). Accordingly, in Cases (I) to (VI), we may demand that the CG coefficient (1.8) be real and positive. In Case (VII), however, the values 1 and 0 of J can each yield nonvanishing values for the coefficients (1.8). It might seem natural to demand that the coefficients

$$C(\lambda\mu\ 11\ \lambda'\mu'\gamma;\ \bar{I}\bar{I}\bar{Y}\ J00\ \bar{I}\bar{I}\bar{Y})$$

with the higher value of J be real and positive. This prescription is indeed satisfactory for the $\gamma = 1$

[7] G. Racah, Phys. Rev. **76**, 1352 (1949); A. R. Edmonds, Proc. Phys. Soc. London **A268**, 567 (1962).

[8] A. Messiah, *Quantum Mechanics* (North-Holland Publishing Company, Amsterdam, 1963), p. 1055.

[9] L. C. Biedenharn, Phys. Letters **3**, 69 (1962).

[10] We use the notation of LM for the generators of SU_3.

[11] The derivations of the matrix elements of the generators of SU_3 by M. Harvey and J. P. Elliot, Proc. Roy. Soc. London **A272**, 557 (1963) and K. T. Hecht. SU_3-reduction coefficients, fractional parentage coefficients etc., University of Michigan preprint, 1963, contain this alternative phase convention to our A2. The derivations of D. L. Pursey, Proc. Roy. Soc. London **A275**, 284 (1963), and of N. Mukunda and L. K. Pandit, J. Math. Phys. **6**, 746 (1965), employ the phase convention of Biedenharn.

[12] We say a CG coefficient vanishes identically if it vanishes because of a selection rule either on its IR labels or on its internal quantum-number labels. A CG coefficient is said to vanish accidentally if the vanishing is not a result of a selection rule.

TABLE I. The isoscalar factors of the CG coefficients of SU_3 which occur in Eq. (1.9).

(λ', μ')	(J, Z)	$U(\lambda\mu\ 11\ \lambda'\mu'\gamma;\ \bar{I}\bar{Y}\ JZ\ \bar{I}'\bar{Y}')$
$(\lambda+2, \mu-1)$	$(\frac{1}{2}, 1)$	$\left[\frac{\mu}{(\mu+1)}\right]^{\frac{1}{2}}$
$(\lambda-1, \mu-1)$	$(1, 0)$	$\left[\frac{\lambda\mu(\lambda+\mu)(\lambda+\mu+1)}{(\lambda+1)(\mu+1)(\lambda+\mu+2)(\lambda+\mu-1)}\right]^{\frac{1}{2}}$
$(\lambda-2, \mu+1)$	$(\frac{1}{2}, -1)$	$\left[\frac{(\lambda-1)(\lambda+\mu+1)^2}{(\lambda+1)(\lambda+\mu)(\lambda+\mu+2)}\right]^{\frac{1}{2}}$
$(\lambda+1, \mu+1)$	$(1, 0)$	1
$(\lambda-1, \mu+2)$	$(\frac{1}{2}, -1)$	$\left[\frac{\lambda}{(\lambda+1)}\right]^{\frac{1}{2}}$
$(\lambda+1, \mu-2)$	$(\frac{1}{2}, 1)$	$\left[\frac{(\mu-1)(\lambda+\mu+1)^2}{(\mu+1)(\lambda+\mu)(\lambda+\mu+2)}\right]^{\frac{1}{2}}$
(λ, μ) $\gamma=1$	$(1, 0)$	$\left[\frac{3(\lambda+\mu)(\lambda+\mu+2)}{4(\lambda^2+\mu^2+\lambda\mu+3\lambda+3\mu)}\right]^{\frac{1}{2}}$
(λ, μ) $\gamma=2$	$(0, 0)$	$\left[\frac{3\mu\lambda(\lambda+\mu+1)(\lambda+\mu+4)^2}{4(\lambda+2)(\mu+2)(\lambda+\mu+3)(\lambda^2+\mu^2+\lambda\mu+3\lambda+3\mu)}\right]^{\frac{1}{2}}$

set of CG coefficients, but it breaks down for the $\gamma = 2$ set, since $C(\lambda\mu\ 11\ \lambda\mu 2;\ \bar{I}\bar{I}\bar{Y}\ 100\ \bar{I}\bar{I}\bar{Y})$ vanishes, accidentally not identically, whenever $\lambda = \mu$. Accordingly in the latter case, we demand that the CG coefficient with $J = 0$ be real and positive. Thus, to summarize the second part of our phase convention for the CG coefficients (1.3), we demand (B) that each of the CG coefficients

$$
\begin{aligned}
&C(\lambda\mu\ 11\ \lambda{+}2\mu{-}1;\ \bar{I}\bar{I}\bar{Y}\ \tfrac{1}{2}\ \tfrac{1}{2}\ 1\ \bar{I}{+}\tfrac{1}{2}\ \bar{I}{+}\tfrac{1}{2}\ \bar{Y}{+}1),\\
&C(\lambda\mu\ 11\ \lambda{-}1\mu{-}1;\ \bar{I}\bar{I}\bar{Y}\ 1\ {-}1\ 0\ \bar{I}{-}1\ \bar{I}{-}1\ \bar{Y}),\\
&C(\lambda\mu\ 11\ \lambda{-}2\mu{+}1;\ \bar{I}\bar{I}\bar{Y}\ \tfrac{1}{2}\ {-}\tfrac{1}{2}\ {-}1\ \bar{I}{-}\tfrac{1}{2}\ \bar{I}{-}\tfrac{1}{2}\ \bar{Y}{-}1),\\
&C(\lambda\mu\ 11\ \lambda{+}1\mu{+}1;\ \bar{I}\bar{I}\bar{Y}\ 1\ 1\ 0\ \bar{I}{+}1\ \bar{I}{+}1\ \bar{Y}),\\
&C(\lambda\mu\ 11\ \lambda{-}1\mu{+}2;\ \bar{I}\bar{I}\bar{Y}\ \tfrac{1}{2}\ \tfrac{1}{2}\ {-}1\ \bar{I}{+}\tfrac{1}{2}\ \bar{I}{+}\tfrac{1}{2}\ \bar{Y}{-}1),\\
&C(\lambda\mu\ 11\ \lambda{+}1\mu{-}2;\ \bar{I}\bar{I}\bar{Y}\ \tfrac{1}{2}\ {-}\tfrac{1}{2}\ 1\ \bar{I}{-}\tfrac{1}{2}\ \bar{I}{-}\tfrac{1}{2}\ \bar{Y}{+}1),\\
&C(\lambda\mu\ 11\ \lambda\mu 1;\ \bar{I}\bar{I}\bar{Y}\ 100\ \bar{I}\bar{I}\bar{Y}),\\
&C(\lambda\mu\ 11\ \lambda\mu 2;\ \bar{I}\bar{I}\bar{Y}\ 000\ \bar{I}\bar{I}\bar{Y}),
\end{aligned}
\qquad (1.9)
$$

be real and positive. The corresponding isoscalar factors are displayed as functions of λ and μ in Table I.

We wish to exhibit that the structure of part (B) of our phase convention for the CG coefficients (1.3) appears quite natural when viewed in the light of known results concerning the Clebsch–Gordan series of SU_3. First it is to be noted that each state of the IR (1, 1) of SU_3 is mentioned once and once only in the set (1.9) of CG coefficients. Further, states belonging to simple weights occur in association with those IR (λ', μ') which occur only once in the product (1.1), and the two states belonging to the double weight $I_z = Y = 0$ occur in association with the two occurrences of (λ, μ) itself in the product. In other words, in statement (b) we have put the states of the octet in 1:1 correspondence with the irreducible constituents of the direct product (1.1). Thus we have, in an explicit and slightly sharpened form, a special case of a lemma due to Biedenharn.[6] Alternatively we may say that statement (B) illustrates the fact that addition of the highest weight of (λ, μ) to each weight of (1, 1) exactly yields the highest weights of each irreducible constituent of the product (1.1), at least [compare with (1.2)] for λ and μ large enough. This fact however is simply the principal on which Speiser's method[13] for the reduction of direct products of IRs of SU_3 depends. It appears that generalization, to the CG coefficients associated with an arbitrary direct product of IR of SU_3, of part (B) of our phase convention must rest heavily on Biedenharn's lemma. It is hoped that this generalization can be given in a future publication.

[13] D. R. Speiser, Proc. Istanbul Summer School, Istanbul, 1962, (to be published). See also Sec. 12 of J. J. deSwart, Rev. Mod. Phys. **35**, 916 (1963).

The discussion so far has been concerned with those points regarding the CG coefficients (1.3) which are of basic theoretical importance. Those technical details of the actual evaluation of the coefficients, which may be bypassed by the reader who wishes only to understand the nature of and make practical use of our tables, are given in Sec. 2. In order to increase the scope and practical utility of our results, it seems desirable that we show how to obtain from the CG coefficients (1.3) explicit formulas for the CG coefficients

$$C(11\ \lambda\mu\ \lambda'\mu'\gamma;\ JNZ\ IMY\ I'M'Y'). \qquad (1.10)$$

Discussion relating to this question is given in Sec. 3. We present our formulas for the isoscalar factors

$$U(\lambda\mu\ 11\ \lambda'\mu'\gamma;\ IY\ JZ\ I'Y') \qquad (1.11)$$

in Table II, which has been divided up into eight parts because of the complexity of various of the entries. Parts 1 to 8 exhibit the isoscalar factors (1.11) with $(JZ, I'Y')$, respectively, given by

$(\frac{1}{2}\ 1, I+\frac{1}{2}\ Y+1)$, $(1\ 0, I-1\ Y)$, $(\frac{1}{2}\ -1, I-\frac{1}{2}\ Y-1)$,

$(1\ 0, I+1\ Y)$, $(\frac{1}{2}\ -1, I+\frac{1}{2}\ Y-1)$, $(\frac{1}{2}\ 1, I-\frac{1}{2}\ Y+1)$,

$(10, IY)$, and $(00, IY)$. The order of the rows of each of Parts 1 to 8 of Table II corresponds exactly to the order of the lines of (1.2) with $\gamma = 1$ before $\gamma = 2$ under case (VII), and the order of the eight parts of Table II can be regarded as being exactly the same in the light of the correspondence referred to below (1.9). Associated with each isoscalar factor $U(\lambda\mu\ 11\ \lambda'\mu'\gamma;\ IY\ JZ\ I'Y')$ is the real phase χ by which it must be multiplied in order to yield the isoscalar factor (cf. Sec. 3) $U(11\,\lambda\mu\lambda'\mu'\gamma; JZ\,IY\,I'Y')$. We have followed this procedure because a general expression for χ in terms of λ, μ, I, Y, etc., has not been obtained.

We conclude this section with a brief survey of previous work on CG coefficients of SU_3. Various authors have made numerical tabulations, namely, Sawada and Yonezawa,[14] Edmonds,[7] Rashid,[15] Dothan and Harari,[16] Tarjanne,[17] deSwart,[13] and McNamee and Chilton.[18] The first-named source provides CG coefficients such as are used in connection with the Sakata version[19] of the unitary symmetry theory. Of the other tabulations, all of which relate to the octet version[1] of unitary symmetry, that of deSwart is the most extensive. Unlike some of the earlier tabulations it is given in association with a clear (but not quite complete) statement of its phase convention, which however differs from the present phase convention in that it uses, instead of (A-2), the same alternative to (A-2) as was used by Harvey and Elliott,[11] and by Hecht.[11] In addition to the numerical compilations just noted, there is the algebraic compilation of the CG coefficients of SU_3 for the products of (λ, μ) with (2, 0), (0, 2), (4, 0), and (2, 1) made by Hecht[11] for use in connection with the work of Elliott[20] on the nuclear shell model. Further, Moshinsky[21] has obtained an algebraic expression in the form of a finite series for the CG coefficients of SU_3 for the product $(\lambda, \mu) \otimes (\kappa, 0)$.

Note added in proof: Pandit and Mukunda, "Algebraic Tabulation of Clebsch–Gordan Coefficients of SU_3 for the Product $(\lambda, \mu) \otimes (3, 0)$ of Representations of SU_3," University of Rochester Report UR-875-73 (to be published), have recently obtained results closely related to those of the present work. They have used the same phase conventions as are employed here.

2. METHOD OF EVALUATION

In LM we introduced a set of operators $\mathcal{I}_Z$, $\mathcal{I}_\pm$, $\mathcal{Y}$, $\mathcal{F}_\pm$, and $\mathcal{G}_\pm$. These operators form a (nonstandard) set of components of an octet operator of SU_3, since they transform under SU_3 exactly as do the generators I_Z, $I_\pm$, Y, $F_\pm$, and $G_\pm$ of SU_3, i.e., according to the octet IR (1, 1). In LM, an evaluation by a purely algebraic method, of all the matrix elements of these operators was performed.[22] The results viewed in the light of the Wigner–Eckart theorem of SU_3 are the basis of our determination of the CG coefficients (1.3) of SU_3.

The first step of the analysis is to replace the set of components of an octet operator of SU_3 by

[14] S. Sawada and M. Yonezawa, Progr. Theoret. Phys. (Kyoto) **23**, 662 (1960).

[15] M. A. Rashid, Nuovo Cimento **26**, 118 (1962).

[16] Y. Dothan and H. Harari, Israel Atomic Energy Commission, Report IA–777, Soreq, Israel, 1963, unpublished.

[17] P. Tarjanne, Carnegie Institute of Technology Report NYO 9290A, Pittsburgh, 1963 (unpublished). See also P. Tarjanne, Ann. Acad. Sci. Fennicae, A VI Physica, 105 (1962).

[18] P. McNamee and F. Chilton, Stanford University Preprint, ITP–126 (1964).

[19] S. Sakata, Progr. Theoret. Phys. (Kyoto) **16**, 686 (1956); Y. Yamaguchi, Progr. Theoret. Phys. (Kyoto) Suppl. **11**, 1, 37 (1960); M. Ikeda, S. Ogawa, and Y. Ohnuki, Progr. Theoret. Phys. (Kyoto) **22**, 715 (1960); **23**, 1073 (1960).

[20] J. P. Elliott, Proc. Roy. Soc. London **A245**, 128, 562 (1958).

[21] M. Moshinsky, Rev. Mod. Phys. **34**, 813 (1962). See also M. Moshinsky, J. Math. Phys. **4**, 1138 (1963) and "Group Theory and the Many Body Problems" to be published in *Physics of Many Particle Systems*, edited by E. Meeron (Gordon and Breach, New York, 1964.

[22] Identical results have been obtained by N. Mukunda and L. K. Pandit by the tensorial method described in their recent paper.[11]

TABLE II. The isoscalar factors $U(\lambda\mu\ 11\ \lambda'\mu'\gamma;\ IY\ JZ\ I'Y')$. We use the following abbreviations:

$$x = I + \tfrac{1}{2}Y,\quad y = I - \tfrac{1}{2}Y,\quad p = (\lambda - \mu)/3,\quad q = (\lambda + 2\mu)/3,\quad r = (2\lambda + \mu)/3,$$

$$I_2 = (\lambda^2 + \mu^2 + \lambda\mu + 3\lambda + 3\mu)/9,\quad I_3 = (\lambda - \mu)(\lambda + 2\mu + 3)(2\lambda + \mu + 3)/162$$

The functions I_2 and I_3 of λ and μ are the eigenvalues in the IR (λ, μ) of SU_3 of the quadratic and the cubic invariant Casimir operators of SU_3.

Part 1

(λ', μ')		$U(\lambda\mu\ 11\ \lambda'\mu'\gamma;\ IY\ \frac{1}{2}\ 1\ I + \frac{1}{2}\ Y + 1)$	χ
$(\lambda + 2, \mu - 1)$	$(+)$	$\left[\frac{(y-p)(r+y+2)(p+x+1)(p+x+2)(q+x+2)}{(x+y+2)(\lambda+1)(\lambda+2)(\mu+1)(\lambda+\mu+2)}\right]^{\frac{1}{2}}$	$(-)$
$(\lambda - 1, \mu - 1)$	$(+)$	$\left[\frac{(q-y)(r-x)(r+y+1)(p+x+1)(r-x-1)}{(x+y+2)(\lambda+1)(\mu+1)(\lambda+\mu+1)(\lambda+\mu+2)}\right]^{\frac{1}{2}}$	$(+)$
$(\lambda - 2, \mu + 1)$	$(-)$	$\left[\frac{(r-x)(y+1-p)(y+1+r)(q+x+2)(r-x-1)}{(x+y+2)\lambda(\lambda+1)(\mu+1)(\lambda+\mu+2)}\right]^{\frac{1}{2}}$	$(+)$
$(\lambda + 1, \mu + 1)$	$(+)$	$\left[\frac{(r+y+2)(q-y+1)(p+x+1)(q+x+2)(q+x+3)}{(x+y+2)(\lambda+1)(\mu+1)(\lambda+\mu+2)(\lambda+\mu+3)}\right]^{\frac{1}{2}}$	$(+)$
$(\lambda - 1, \mu + 2)$	$(-)$	$\left[\frac{(r-x)(q+x+2)(q+x+3)(y+1-p)(q+1-y)}{(x+y+2)(\lambda+1)(\mu+1)(\mu+2)(\lambda+\mu+2)}\right]^{\frac{1}{2}}$	$(-)$
$(\lambda + 1, \mu - 2)$	$(+)$	$\left[\frac{(r-x)(y-p)(q-y)(p+x+1)(p+x+2)}{\mu(\mu+1)(\lambda+1)(\lambda+\mu+2)(x+y+2)}\right]^{\frac{1}{2}}$	$(+)$
(λ, μ) $\gamma = 1$	$(+)$	$\left[\frac{3(p+x+1)(r-x)(q+x+2)}{2(\lambda^2+\mu^2+\mu\lambda+3\lambda+3\mu)(x+y+2)}\right]^{\frac{1}{2}}$	$(-)$
(λ, μ) $\gamma = 2$	$(-)$	$\left[\frac{18I_2(p+x+1)(q+x+2)(r-x)}{(x+y+2)\lambda\mu(\lambda+2)(\mu+2)(\lambda+\mu+1)(\lambda+\mu+3)}\right]^{\frac{1}{2}}\left[y+\frac{1}{2}-\frac{3I_3}{I_2}\right]$	$(+)$

Part 2

(λ', μ')		$U(\lambda\mu\ 11\ \lambda'\mu'\gamma;\ IY\ 10\ I - IY)$	χ
$(\lambda + 2, \mu - 1)$	$(-)$	$\left[\frac{(y-p)(y-1-p)(q-y+1)(q+x+1)(r-x+1)(r-x+2)}{(x+y)(x+y-1)(\lambda+1)(\lambda+2)(\mu+1)(\lambda+\mu+2)}\right]^{\frac{1}{2}}$	$(-)$
$(\lambda - 1, \mu - 1)$	$(+)$	$\left[\frac{(y-p)(r+y)(p+x)(r+y+1)(q+x+1)(q+x)}{(x+y)(x+y-1)(\lambda+1)(\mu+1)(\lambda+\mu+1)(\lambda+\mu+2)}\right]^{\frac{1}{2}}$	$(+)$
$(\lambda - 2, \mu + 1)$	$(+)$	$\left[\frac{(p+x)(r+y)(p+x-1)(q+x+1)(q-y+1)(r+y+1)}{(x+y)(x+y-1)\lambda(\lambda+1)(\mu+1)(\lambda+\mu+2)}\right]^{\frac{1}{2}}$	$(+)$
$(\lambda + 1, \mu + 1)$	$(-)$	$\left[\frac{(p+x)(y-p)(r-x+1)(r-x+2)(q-y+2)(r-y+1)}{(x+y)(x+y-1)(\lambda+1)(\mu+1)(\lambda+\mu+2)(\lambda+\mu+3)}\right]^{\frac{1}{2}}$	$(+)$
$(\lambda - 1, \mu + 2)$	$(-)$	$\left[\frac{(p+x)(p+x-1)(r-x+1)(r+y+1)(q-y+2)(q-y+1)}{(x+y)(x+y-1)(\lambda+1)(\mu+1)(\mu+2)(\lambda+\mu+2)}\right]^{\frac{1}{2}}$	$(-)$
$(\lambda + 1, \mu - 2)$	$(-)$	$\left[\frac{(y-p)(q+x)(y-p-1)(r+y+1)(r-x+1)(q+x+1)}{(x+y)(x+y-1)\mu(\mu+1)(\lambda+1)(\lambda+\mu+2)}\right]^{\frac{1}{2}}$	$(+)$
(λ, μ) $\gamma = 1$		0	
(λ, μ) $\gamma = 2$	$(-)$	$\left[\frac{18I_2(p+x)(y-p)(q+x+1)(r-x+1)(q-y+1)(r+y+1)}{(x+y)(x+y-1)\lambda\mu(\lambda+2)(\mu+2)(\lambda+\mu+1)(\lambda+\mu+3)}\right]^{\frac{1}{2}}$	$(+)$

TABLE II. (Continued)

Part 3

(λ', μ')		$U(\lambda\mu\ 11\ \lambda'\mu'\gamma;\ I\ Y\ \frac{1}{2}\ -1\ I-\frac{1}{2}\ Y-1)$	χ
$(\lambda+2, \mu-1)$	$(-)$	$\left[\frac{(y-p)(r+y+2)(q+x+1)(r-x+1)(r-x+2)}{(x+y)(\lambda+1)(\lambda+2)(\mu+1)(\lambda+\mu+2)}\right]^{\frac{1}{2}}$	$(+)$
$(\lambda-1, \mu-1)$	$(-)$	$\left[\frac{(q-y)(p+x)(q+x)(r+y+1)(q+x+1)}{(x+y)(\lambda+1)(\mu+1)(\lambda+\mu+1)(\lambda+\mu+2)}\right]^{\frac{1}{2}}$	$(-)$
$(\lambda-2, \mu+1)$	$(+)$	$\left[\frac{(p+x)(y+1-p)(r+y+1)(p+x-1)(q+x+1)}{(x+y)\lambda(\lambda+1)(\mu+1)(\lambda+\mu+2)}\right]^{\frac{1}{2}}$	$(-)$
$(\lambda+1, \mu+1)$	$(-)$	$\left[\frac{(p+x)(r+y+2)(q-y+1)(r-x+2)(r-x+1)}{(x+y)(\lambda+1)(\mu+1)(\lambda+\mu+2)(\lambda+\mu+3)}\right]^{\frac{1}{2}}$	$(-)$
$(\lambda-1, \mu+2)$	$(-)$	$\left[\frac{(p+x)(p+x-1)(r-x+1)(y+1-p)(q-y+1)}{(x+y)(\lambda+1)(\mu+1)(\mu+2)(\lambda+\mu+2)}\right]^{\frac{1}{2}}$	$(+)$
$(\lambda+1, \mu-2)$	$(+)$	$\left[\frac{(q+x)(y-p)(q-y)(r-x+1)(q+x+1)}{(x+y)\mu(\mu+1)(\lambda+1)(\lambda+\mu+2)}\right]^{\frac{1}{2}}$	$(-)$
(λ, μ) $\gamma=1$	$(+)$	$\left[\frac{3(p+x)(q+x+1)(r-x+1)}{2(x+y)(\lambda^2+\mu^2+\mu\lambda+3\mu+3\lambda)}\right]^{\frac{1}{2}}$	$(+)$
(λ, μ) $\gamma=2$	$(-)$	$\left[\frac{18I_2(p+x)(q+x+1)(r-x+1)}{(x+y)\lambda\mu(\lambda+2)(\mu+2)(\lambda+\mu+1)(\lambda+\mu+3)}\right]^{\frac{1}{2}}\left[y+\frac{1}{2}-\frac{3I_3}{I_2}\right]$	$(-)$

Part 4

(λ', μ')		$U(\lambda\mu\ 11\ \lambda'\mu'\gamma;\ I\ Y\ 10\ I+I\ Y)$	χ
$(\lambda+2, \mu-1)$	$(-)$	$\left[\frac{(q-y)(p+x+1)(p+x+2)(q+x+2)(r+y+2)(r+y+3)}{(x+y+2)(x+y+3)(\lambda+1)(\lambda+2)(\mu+1)(\lambda+\mu+2)}\right]^{\frac{1}{2}}$	$(-)$
$(\lambda-1, \mu-1)$	$(-)$	$\left[\frac{(r-x)(q-y)(p+x+1)(r-x-1)(y+1-p)(r-y-1)}{(x+y+2)(x+y+3)(\lambda+1)(\mu+1)(\lambda+\mu+1)(\lambda+\mu+2)}\right]^{\frac{1}{2}}$	$(+)$
$(\lambda-2, \mu+1)$	$(+)$	$\left[\frac{(q-y)(r-x)(y+1-p)(y+2-p)(q+x+2)(r-x-1)}{(x+y+2)(x+y+3)\lambda(\lambda+1)(\mu+1)(\lambda+\mu+2)}\right]^{\frac{1}{2}}$	$(+)$
$(\lambda+1, \mu+1)$	$(+)$	$\left[\frac{(y+1-p)(r+y+2)(r+y+3)(p+x+1)(q+x+2)(q+x+3)}{(x+y+2)(x+y+3)(\lambda+1)(\mu+1)(\lambda+\mu+2)(\lambda+\mu+3)}\right]^{\frac{1}{2}}$	$(+)$
$(\lambda-1, \mu+2)$	$(-)$	$\left[\frac{(r-x)(y+1-p)(y+2-p)(r+y+2)(q+x+2)(q+x+3)}{(x+y+2)(x+y+3)(\lambda+1)(\mu+1)(\mu+2)(\lambda+\mu+2)}\right]^{\frac{1}{2}}$	$(-)$
$(\lambda+1, \mu-2)$	$(-)$	$\left[\frac{(r-x)(q-y)(p+x+1)(p+x+2)(r+y+2)(q-y-1)}{(x+y+2)(x+y+3)\mu(\mu+1)(\lambda+1)(\lambda+\mu+2)}\right]^{\frac{1}{2}}$	$(+)$
(λ, μ) $\gamma=1$		0	
(λ, μ) $\gamma=2$	$(+)$	$\left[\frac{18I_2(r-x)(q-y)(p+x+1)(q+x+2)(y-p+1)(r+y+2)}{(x+y+2)(x+y+3)\lambda\mu(\lambda+2)(\mu+2)(\lambda+\mu+1)(\lambda+\mu+3)}\right]^{\frac{1}{2}}$	$(+)$

TABLE II. (Continued)

Part 5

(λ', μ')		$U(\lambda\mu\ 11\ \lambda'\mu'\gamma;\ I\ Y\ \frac{1}{2}-1\ I+\frac{1}{2}\ Y-1)$	χ
$(\lambda+2, \mu-1)$	$(-)$	$\left[\frac{(q-y)(r+y+2)(r+y+3)(p+x+1)(r-x+1)}{(x+y+2)(\lambda+1)(\lambda+2)(\mu+1)(\lambda+\mu+2)}\right]^{\frac{1}{2}}$	$(-)$
$(\lambda-1, \mu-1)$	$(+)$	$\left[\frac{(r-x)(q-y)(q+x+1)(y+1-p)(r-y-1)}{(x+y+2)(\lambda+1)(\mu+1)(\lambda+\mu+1)(\lambda+\mu+2)}\right]^{\frac{1}{2}}$	$(+)$
$(\lambda-2, \mu+1)$	$(-)$	$\left[\frac{(q-y)(p+x)(r-x)(y+1-p)(y+2-p)}{(x+y+2)\lambda(\lambda+1)(\mu+1)(\lambda+\mu+2)}\right]^{\frac{1}{2}}$	$(+)$
$(\lambda+1, \mu+1)$	$(+)$	$\left[\frac{(r-x+1)(q+x+2)(y+1-p)(r+y+3)(r+y+2)}{(x+y+2)(\lambda+1)(\mu+1)(\lambda+\mu+2)(\lambda+\mu+3)}\right]^{\frac{1}{2}}$	$(+)$
$(\lambda-1, \mu+2)$	$(+)$	$\left[\frac{(p+x)(q+x+2)(y+1-p)(y+2-p)(r+y+2)}{(x+y+2)(\lambda+1)(\mu+1)(\mu+2)(\lambda+\mu+2)}\right]^{\frac{1}{2}}$	$(-)$
$(\lambda+1, \mu-2)$	$(+)$	$\left[\frac{(q-y)(p+x+1)(q+x+1)(r+y+2)(q-y-1)}{(x+y+2)\mu(\mu+1)(\lambda+1)(\lambda+\mu+2)}\right]^{\frac{1}{2}}$	$(+)$
(λ, μ) $\gamma=1$	$(-)$	$\left[\frac{3(y+1-p)(q-y)(r+y+2)}{2(x+y+2)(\lambda^2+\mu^2+\mu\lambda+3\mu+3\lambda)}\right]^{\frac{1}{2}}$	$(-)$
(λ, μ) $\gamma=2$	$(-)$	$\left[\frac{18I_2(q-y)(y+1-p)(r+y+2)}{(x+y+2)\lambda\mu(\lambda+2)(\mu+2)(\lambda+\mu+1)(\lambda+\mu+3)}\right]^{\frac{1}{2}}\left[x+\frac{1}{2}+\frac{3I_3}{I_2}\right]$	$(+)$

Part 6

(λ', μ')		$U(\lambda\mu\ 11\ \lambda'\mu'\gamma;\ I\ Y\ \frac{1}{2}\ 1\ I-\frac{1}{2}\ Y+1)$	χ
$(\lambda+2, \mu-1)$	$(+)$	$\left[\frac{(y-p)(y-1-p)(q-y+1)(p+x+1)(r-x+1)}{(x+y)(\lambda+1)(\lambda+2)(\mu+1)(\lambda+\mu+2)}\right]^{\frac{1}{2}}$	$(+)$
$(\lambda-1, \mu-1)$	$(+)$	$\left[\frac{(r-x)(y-p)(r+y)(q+x+1)(r+y+1)}{(x+y)(\lambda+1)(\mu+1)(\lambda+\mu+1)(\lambda+\mu+2)}\right]^{\frac{1}{2}}$	$(-)$
$(\lambda-2, \mu+1)$	$(+)$	$\left[\frac{(r+y)(p+x)(r-x)(q-y+1)(r+y+1)}{(x+y)\lambda(\lambda+1)(\mu+1)(\lambda+\mu+2)}\right]^{\frac{1}{2}}$	$(-)$
$(\lambda+1, \mu+1)$	$(+)$	$\left[\frac{(y-p)(r-x+1)(q+x+2)(q-y+2)(r-y+1)}{(x+y)(\lambda+1)(\mu+1)(\lambda+\mu+2)(\lambda+\mu+3)}\right]^{\frac{1}{2}}$	$(-)$
$(\lambda-1, \mu+2)$	$(+)$	$\left[\frac{(p+x)(q+x+2)(r+y+1)(q-y+1)(q-y+2)}{(x+y)(\lambda+1)(\mu+1)(\mu+2)(\lambda+\mu+2)}\right]^{\frac{1}{2}}$	$(+)$
$(\lambda+1, \mu-2)$	$(+)$	$\left[\frac{(y-p)(p+x+1)(q+x+1)(y-1-p)(r+y+1)}{(x+y)\mu(\mu+1)(\lambda+1)(\lambda+\mu+2)}\right]^{\frac{1}{2}}$	$(-)$
(λ, μ) $\gamma=1$	$(+)$	$\left[\frac{3(y-p)(q-y+1)(r+y+1)}{2(x+y)(\lambda^2+\mu^2+\mu\lambda+3\mu+3\lambda)}\right]^{\frac{1}{2}}$	$(+)$
(λ, μ) $\gamma=2$	$(+)$	$\left[\frac{18I_2(y-p)(q-y+1)(r+y+1)}{(x+y)\lambda\mu(\lambda+2)(\mu+2)(\lambda+\mu+1)(\lambda+\mu+3)}\right]^{\frac{1}{2}}\left[x+\frac{1}{2}+\frac{3I_3}{I_2}\right]$	$(-)$

TABLE II. (Continued)

Part 7

(λ', μ')		$U(\lambda\mu\ 11\ \lambda'\mu'\gamma;\ I\,Y\ 10\ I\,Y)$	χ
$(\lambda+2, \mu-1)$	$(-)$	$\left[\frac{2(y-p)(r+y+2)(p+x+1)(r-x+1)}{(x+y)(x+y+2)(\lambda+1)(\mu+1)(\lambda+2)(\lambda+\mu+2)}\right]^{\frac{1}{2}}[q+1+\frac{1}{2}(x-y)]$	$(+)$
$(\lambda-1, \mu-1)$	$(-)$	$\left[\frac{2(q-y)(r-x)(r+y+1)(q+x+1)}{(x+y)(x+y+2)(\lambda+1)(\mu+1)(\lambda+\mu+1)(\lambda+\mu+2)}\right]^{\frac{1}{2}}[\frac{1}{2}(y-x)-p]$	$(-)$
$(\lambda-2, \mu+1)$	$(-)$	$\left[\frac{2(p+x)(r-x)(y+1-p)(r+y+1)}{(x+y)(x+y+2)\lambda(\lambda+1)(\mu+1)(\lambda+\mu+2)}\right]^{\frac{1}{2}}[q+1+\frac{1}{2}(x-y)]$	$(-)$
$(\lambda+1, \mu+1)$	$(+)$	$\left[\frac{2(r+y+2)(q-y+1)(r-x+1)(q+x+2)}{(x+y)(x+y+2)(\lambda+1)(\mu+1)(\lambda+\mu+2)(\lambda+\mu+3)}\right]^{\frac{1}{2}}[\frac{1}{2}(y-x)-p]$	$(-)$
$(\lambda-1, \mu+2)$	$(+)$	$\left[\frac{2(p+x)(q+x+2)(y+1-p)(q+1-y)}{(x+y)(x+y+2)(\lambda+1)(\mu+1)(\mu+2)(\lambda+\mu+2)}\right]^{\frac{1}{2}}[r+1-\frac{1}{2}(x-y)]$	$(+)$
$(\lambda+1, \mu-2)$	$(-)$	$\left[\frac{2(y-p)(q-y)(p+x+1)(q+x+1)}{(x+y)(x+y+2)\mu(\mu+1)(\lambda+1)(\lambda+\mu+2)}\right]^{\frac{1}{2}}[r+1-\frac{1}{2}(x-y)]$	$(-)$
(λ, μ) $\gamma=1$	$(+)$	$\left[\frac{3(x+y)(x+y+2)}{4(\lambda^2+\mu^2+\lambda\mu+3\lambda+3\mu)}\right]^{\frac{1}{2}}$	$(+)$
(λ, μ) $\gamma=2$	$(+)$	$\left[\frac{9I_2(x+y)(x+y+2)}{\lambda\mu(\lambda+2)(\mu+2)(\lambda+\mu+1)(\lambda+\mu+3)}\right]^{\frac{1}{2}}$ $\times\left[\frac{1}{(x+y+1)}\left\{\frac{(y-p)(q-y+1)(r+y+1)}{(x+y)}\right.\right.$ $\left.\left.-\frac{(p+x+1)(q+x+2)(r-x)}{(x+y+2)}\right\}-(x-y)+\frac{3I_3}{I_2}-\frac{1}{2}\right]$	$(-)$

Part 8

(λ', μ')		$U(\lambda\mu\ 11\ \lambda'\mu'\gamma;\ I\,Y\ 00\ I\,Y)$	χ
$(\lambda+2, \mu-1)$	$(+)$	$\left[\frac{3(y-p)(r+y+2)(p+x+1)(r-x+1)}{2(\lambda+1)(\lambda+2)(\mu+1)(\lambda+\mu+2)}\right]^{\frac{1}{2}}$	$(-)$
$(\lambda-1, \mu-1)$	$(-)$	$\left[\frac{3(r+y+1)(q-y)(r-x)(q+x+1)}{2(\lambda+1)(\mu+1)(\lambda+\mu+1)(\lambda+\mu+2)}\right]^{\frac{1}{2}}$	$(+)$
$(\lambda-2, \mu+1)$	$(+)$	$\left[\frac{3(y+1-p)(r+y+1)(p+x)(r-x)}{2(\lambda+1)\lambda(\mu+1)(\lambda+\mu+2)}\right]^{\frac{1}{2}}$	$(+)$
$(\lambda+1, \mu+1)$	$(+)$	$\left[\frac{3(r+y+2)(q-y+1)(r-x+1)(q+x+2)}{2(\lambda+1)(\mu+1)(\lambda+\mu+2)(\lambda+\mu+3)}\right]^{\frac{1}{2}}$	$(+)$
$(\lambda-1, \mu+2)$	$(+)$	$\left[\frac{3(p+x)(q+x+2)(y+1-p)(q-y+1)}{2(\lambda+1)(\mu+1)(\mu+2)(\lambda+\mu+2)}\right]^{\frac{1}{2}}$	$(-)$
$(\lambda+1, \mu-2)$	$(-)$	$\left[\frac{3(p+x+1)(q+x+1)(y-p)(q-y)}{2\mu(\mu+1)(\lambda+1)(\lambda+\mu+2)}\right]^{\frac{1}{2}}$	$(+)$
(λ, μ) $\gamma=1$	$(-)$	$\left[\frac{9(x-y)^2}{4(\lambda^2+\mu^2+\lambda\mu+3\mu+3\lambda)}\right]^{\frac{1}{2}}$	$(-)$
(λ, μ) $\gamma=2$	$(+)$	$\frac{[27I_2]^{\frac{1}{2}}\left[(y-x)\left(x+\frac{1}{2}+\frac{3I_3}{I_2}\right)+x(x+1)-I_2\right]}{[\lambda\mu(\lambda+2)(\mu+2)(\lambda+\mu+1)(\lambda+\mu+3)]^{\frac{1}{2}}}$	$(+)$

a certain set of components, $\mathfrak{M}^{(1,1)}_{IMY}$ or simply $\mathfrak{M}_{IMY}$, which we refer to as a standard set. They are standard in the sense that their matrix elements—temporarily excluding from attention those of case (VII)—can be expressed in the form

$$\langle \lambda'\mu'\ I'M'Y'\ |\mathfrak{M}_{JNZ}|\ \lambda\mu\ IMY\rangle = C(\lambda\mu\ 11\ \lambda'\mu';\ IMY\ JNZ\ I'M'Y')\langle \lambda'\mu'\ ||\mathfrak{M}||\ \lambda\mu\rangle, \tag{2.1}$$

where the reduced matrix element is independent of IMY, JNZ, and $I'M'Y'$ and the CG coefficient satisfies the phase convention laid down as statements (A-1), (A-2), and (B) in the introduction. The results of LM cannot be immediately written in the form (2.1) for two reasons:

(i) the correct *relative* normalization of the operators $\mathcal{I}_Z$, $\mathcal{I}_\pm$, $\mathcal{Y}$, $\mathcal{F}_\pm$ and $\mathcal{G}_\pm$ is as yet undertermined,[23]

(ii) even after appropriate relative normalization factors have been associated with them, the functions $\mathcal{I}$, $\mathcal{Y}$, $\mathcal{F}$, and $\mathcal{G}$ tabulated in LM differ from the corresponding isoscalar factor by a *common* normalization factor which in each of the cases I to VI is a function only of λ and μ.

We consider point (i) first. From inspection of the selection rules on IMY obeyed by the operators $\mathcal{I}_Z$, $\mathcal{I}_\pm$, $\mathcal{Y}$, $\mathcal{F}_\pm$, and $\mathcal{G}_\pm$, it is clear that we must make associations of the form

$$\begin{aligned}
\mathfrak{M}_{IM0} &= a\mathcal{I}_M, \qquad M = 1, 0, -1,\\
&\qquad\qquad \mathcal{I}_{\pm1} = \mp\mathcal{I}_\pm/\sqrt{2},\ \mathcal{I}_0 = \mathcal{I}_Z,\\
\mathfrak{M}_{000} &= b\mathcal{Y},\\
\mathfrak{M}_{\frac{1}{2}\pm\frac{1}{2}1} &= c\mathcal{F}_\pm,\\
\mathfrak{M}_{\frac{1}{2}\pm\frac{1}{2}-1} &= d\mathcal{G}_\pm.
\end{aligned} \tag{2.2}$$

We can determine the ratios a, b, c, and d by examination of an appropriate special case of (2.1). We choose the case of $(\lambda, \mu) = (1, 1)$ and $(\lambda', \mu') = (3, 0)$, and use Eq. (2.2) and the entries of Table I of LM to evaluate explicitly the left side of (2.1). We can write the results in the form

$$\langle 30\ I'M'Y'\ |\mathfrak{M}_{JNZ}|\ 11\ IMY\rangle = f(IMY, JNZ, I'M'Y')\ \langle 30\ ||(1,1)||\ 11\rangle, \tag{2.3}$$

where the functions f are known functions of the indicated arguments, each one containing one of the numbers a, b, c, and d as a factor, and the notation of LM for the reduced matrix element has been used since this need not be the same as $\langle 30\ ||\mathfrak{M}||\ 11\rangle$. We must now compute the CG coefficients

$$C(11\ 11\ 30;\ IMY\ JNZ\ I'M'Y')$$

from first principles,[24] being careful to impose the phase convention specified in the introduction. Accordingly from (2.1), we have a result of the form

$$f(IMY,\ JNZ,\ I'M'Y') = x\ C(11\ 11\ 30;\ IMY\ JNZ\ I'M'Y'). \tag{2.4}$$

Herein x is a number independent of IMY, JNZ and $I'M'Y'$. It is, in fact, the ratio of the reduced matrix elements $\langle 30\ ||(1,1)||\ 11\rangle$ and $\langle 30\ ||\mathfrak{M}||\ 11\rangle$, and its presence reflects the fact [cf. (ii) above] that the functions f of (2.3) are not normalized. From (2.4) we get

$$a/2 = -b/\sqrt{3} = c/\sqrt{2} = d/\sqrt{2} = x. \tag{2.5}$$

In the special case of the generators of SU_3, the choice $x = 2\sqrt{3}$, allows the identifications

$$\begin{aligned}
M_{110} &= I_{+1}/\sqrt{3} = -E_1,\\
M_{100} &= I_z/\sqrt{3} = H_1,\\
M_{1-10} &= I_{-1}/\sqrt{3} = E_{-1},\\
M_{000} &= -Y/2 = -H_2,\\
M_{\frac{1}{2}\frac{1}{2}1} &= F_+/\sqrt{6} = E_2,\\
M_{\frac{1}{2}-\frac{1}{2}1} &= F_-/\sqrt{6} = E_3,\\
M_{\frac{1}{2}\frac{1}{2}-1} &= G_+/\sqrt{6} = -E_{-3},\\
M_{\frac{1}{2}-\frac{1}{2}-1} &= G_-/\sqrt{6} = E_{-2},
\end{aligned} \tag{2.6}$$

to be made, and we adopt it in general.[25] We refer to the components $\mathfrak{M}_{IMY}$ of an octet operator, related to the original components by the "script analog" of Eq. (2.6), as a standard set of components.[28] The results of LM under Cases (I) to (VI) can now be written in the form

[23] The analogous situation for SU_2 involves the replacement of the components $\mathcal{A}_\pm$ $\mathcal{A}_z$ of a vector operator by the spherical components $\mathcal{A}_m$ ($m = 1$, 0, and -1).

[24] See Sec. VI D of R. E. Behrends, J. Dreitlein, C. Fronsdal, and B. W. Lee, Rev. Mod. Phys. **34**, 1 (1962), for a description of the method of calculation. We always use the formulas of Biedenharn,[9] or the transcription thereof given in Sec. 2 of LM, for the matrix elements of the generators of SU_3, and hence obtain results differing from those obtained by these authors and by some signs.

[25] We have introduced in the third column the Weyl–Racah[26] generators. From the expression for the Casimir operator I_2 of SU_3 given by Baird and Biedenharn[27] in terms of these generators and Eq. (2.6), we see that we have $I_2 = \Sigma_{IMY} M_{IMY} M_{IMY}$.†

[26] H. Weyl, lecture notes, Institute for Advanced Study, Princeton, 1935 (unpublished); G. Racah, lecture notes, Institute for Advanced Study, Princeton, 1951 (unpublished).

[27] G. E. Baird and L. C. Biedenharn, J. Math. Phys. **4**, 1449 (1963).

[28] Essentially the same definition of a standard set of components of an octet operator has been given by N. Mukunda and L. K. Pandit.[11]

$$\langle \lambda'\mu' \; I'M'Y' \; |\mathfrak{M}_{JNZ}| \; \lambda\mu \; IMY\rangle$$
$$= C(IJI'; MNM')\mathfrak{M}(\lambda\mu \, 11 \, \lambda'\mu'; IY \, JZ \, I'Y')$$
$$\langle \lambda'\mu' \; ||(1,1)|| \; \lambda\mu\rangle, \quad (2.7)$$

where the functions f are known explicitly in virtue of the Tables I to VI of LM and the relation of $\mathfrak{M}_{JNZ}$ to the original set of generators. The result (2.7) differs from (2.1) only because the λ, μ, λ' and μ' dependent normalization factor necessary to convert the function $\mathfrak{M}$ of (2.7) into an isoscalar factor U has not been determined, or rather is absorbed in the reduced matrix element. To make (2.7) and (2.1) identical, we set

$$\mathfrak{M}(\cdots) = f_{\lambda'\mu'}(\lambda, \mu) \; U(\cdots),$$
$$\langle \lambda'\mu' \; ||\mathfrak{M}|| \; \lambda\mu\rangle = f_{\lambda'\mu'}(\lambda, \mu)\langle \lambda'\mu' \; ||(1,1)|| \; \lambda\mu\rangle, \quad (2.8)$$

and determine $f_{\lambda'\mu'}(\lambda, \mu)$ using the orthonormality of the isoscalar factors according to

$$[f_{\lambda'\mu'}(\lambda, \mu)]^2$$
$$= \sum [\mathfrak{M}(\lambda\mu \; 11 \; \lambda'\mu'; IY \; JZ \; I'Y')]^2, \quad (2.9)$$

the sum being over IY and JZ at fixed $I'Y'$. By judicious choice of $I'Y'$, e.g.,[29]

$$I' = 0, \qquad Y' = 2(\lambda' - \mu')/3, \quad (2.10)$$

we easily evaluate the functions $f_{\lambda'\mu'}(\lambda, \mu)$ up to a sign. At this point, the relative signs of all CG coefficients within any of cases I to VI are determined by virtue of our use of the Biedenharn phase convention (A-1) and (A-2) in LM, and in the derivation of Eq. (2.3). It remains only to impose (B) to determine the absolute phase of one member of each set, and the entries in the first six rows of Parts 1 to 8 of Table II are then accounted for.

We now turn to case (VII), which was explicitly excluded from the previous remarks of this section, in order to attend to those complications that arise as a result of the double appearance of (λ, μ) itself in the product (1.1). We employ the standard set $\mathfrak{M}_{IMY}$ of components of an octet operator and consider their matrix elements under case (VII), written in the form

$$\langle \lambda\mu \; I'M'Y' \; |\mathfrak{M}_{JNZ}| \; \lambda\mu \; IMY\rangle$$
$$= \sum\nolimits_\gamma C(\lambda\mu \; 11 \; \lambda\mu\gamma; IMY \; JNZ \; I'M'Y')$$
$$\langle \lambda\mu \; ||\mathfrak{M}|| \; \lambda\mu\rangle_\gamma. \quad (2.11)$$

The sum is over the two values 1 and 2 of the index γ which distinguishes an orthogonal pair of sets of CG coefficients and an independent pair of reduced matrix elements. Arbitrariness in the definition of the orthogonal pair of sets of CG coefficients is to be resolved by the demand that we can give the matrix elements of the generators (2.6) of SU_3 in the form

$$\langle \lambda\mu \; I'M'Y' \; |M_{JNZ}| \; \lambda\mu \; IMY\rangle$$
$$= C(\lambda\mu \; 11 \; \lambda\mu 1; IMY \; JNZ \; I'M'Y')$$
$$\langle \lambda\mu \; ||M|| \; \lambda\mu\rangle_1. \quad (2.12)$$

This expresses as a formula the corresponding statement, made in the introduction.

By the method of LM, or by use of formula (A.8) of a paper by Okubo[30] or by the method of Mukunda and Pandit,[11] we find explicit results of the form

$$\langle \lambda\mu \; I'M'Y' \; |\mathfrak{M}_{JNZ}| \; \lambda\mu \; IMY\rangle$$
$$= \sum_{\alpha=1}^{2} \mathfrak{M}(\lambda\mu, \; IMY \; JNZ \; I'M'Y', \alpha)$$
$$\times \langle \lambda\mu \; ||(1,1)|| \; \lambda\mu\rangle_\alpha. \quad (2.13)$$

It so happens that the results actually obtained are such that the $\alpha = 1$ functions already yield the matrix elements of the generators apart from an over-all factor. We can immediately find the $\gamma = 1$ coefficients by finding the normalization factor $f_1(\lambda, \mu)$ as in cases (I) to (VI) and demanding that $C(\lambda\mu \; 11 \; \lambda\mu 1; \bar{I}\bar{I}\bar{Y} \; 100 \; \bar{I}\bar{I}\bar{Y})$ be positive.[31] This accounts for the seventh row of parts 1 to 8 of Table II.

It is convenient for the remaining portion of our discussion to introduce the notation

$$v = (IMY), \qquad k = (JNZ), \qquad v' = (I'M'Y'),$$
$$k_\alpha = \langle \lambda\mu \; ||(1,1)|| \; \lambda\mu\rangle_\alpha, \qquad \alpha = 1, 2.$$

Now Eq. (2.13) can be written as

$$\langle \lambda\mu \; v' \; |\mathfrak{M}_k| \; \lambda\mu \; v\rangle = \sum_{k\alpha} \mathfrak{M}(\lambda\mu, vk\alpha', \alpha)k_\alpha \quad (2.14)$$

and

$$f_1^2 = \sum_{kv} [\mathfrak{M}(\lambda\mu, vkv', 1)]^2. \quad (2.15)$$

If we next compute

$$g = \sum_{kv} \mathfrak{M}(\lambda\mu, vkv', 1)\mathfrak{M}(\lambda\mu, vkv', 2), \quad (2.16)$$

[29] That any IR (λ', μ') of SU_3 contains a state $(I, Y) = (0, 2(\lambda'\mu')/3)$ follows from the general formulas[3] for the IY content of an IR of SU_3. Derivation of the forumlas in question can be found in C. R. Hagen and A. J. Macfarlane, J. Math. Phys. 5, 1335 (1964).

[30] S. Okubo, Progr. Theoret. Phys. (Kyoto) 27, 949 (1963).

[31] Comparing now with Eq. (2.12), we find $\langle \lambda\mu \; ||M \; ||\lambda\mu\rangle_1 = (\lambda^2 + \mu^2 + \mu\lambda + 3\mu + 3\lambda)/9$, which is the eigenvalue in (λ, μ) of the Casimir operator I_2 of SU_3[25,27]. Also the CRs of the Lie algebra of SU_3 can now be given in the concise form

$$[M_{I_1M_1Y_1}, \; M_{I_2M_2Y_2}] = \Sigma_I \langle 11IMY| \; M_{I_1M_1Y_1} \; |11I_2M_2Y_2\rangle M_{IMY}.$$

we find g to be a nonvanishing function of λ and μ but not v'. Hence normalization of the set $\mathfrak{M}(\cdots 2)$, i.e., dividing by f_2

$$f_2^2 = \sum_{kv} [\mathfrak{M}(\lambda\mu, vkv', 2)]^2, \tag{2.17}$$

does not yield a set of SU_3 CG coefficients orthogonal to the set just constructed from the functions $\mathfrak{M}(\cdots 1)$. However, it is admissible to rewrite Eq. (2.13) in the form

$$\langle \lambda\mu v' \, | \mathfrak{M}_k | \, \lambda\mu v \rangle = \sum_{\gamma=1}^{2} \mathfrak{M}'(\lambda\mu, vkv', \gamma) k'_\gamma, \tag{2.18}$$

where

$$\begin{aligned} \mathfrak{M}'(\cdots 1) &= \mathfrak{M}(\cdots 1), \\ k'_1 &= k_1 + xk_2, \\ \mathfrak{M}'(\cdots 2) &= \mathfrak{M}(\cdots 2) - x\mathfrak{M}(\cdots 1), \\ k'_2 &= k_2, \end{aligned} \tag{2.19}$$

and x is determined by the demand that we have

$$\sum_{vk} \mathfrak{M}'(\lambda\mu, vkv', 1)\mathfrak{M}'(\lambda\mu, vkv', 2) = 0. \tag{2.20}$$

From (2.19) and (2.20), we get

$$x = g/(f_1^2). \tag{2.21}$$

We now get the required $\gamma = 2$ set of CG coefficients by normalization of the functions $\mathfrak{M}'(\cdots 2)$ given by (2.19) and (2.21). These can be written, in terms of the original functions $\mathfrak{M}(\cdots \alpha)$ in the form

$$\begin{aligned} &C(\lambda\mu\ 11\ \lambda\mu 2; vkv') \\ &\quad = [f_1^2\mathfrak{M}(\lambda\mu, vkv', 2) - g\mathfrak{M}(\lambda\mu, vkv', 1)] \\ &\qquad \times [f_1(f_1^2 f_2^2 - g^2)^{\frac{1}{2}}]^{-1}. \end{aligned} \tag{2.22}$$

This is sufficient to account for the entries in the eighth row of Parts 1 to 8 of Table II, except for an over-all sign, which is determined by the requirement [cf. (1.9)] that

$$C(\lambda\mu\ 11\ \lambda\mu 2; \bar{I}\bar{I}\bar{Y}\ 000\ \bar{I}\bar{I}\bar{Y})$$

be positive.

3. THE CG COEFFICIENTS FOR THE PRODUCT (1,1) ⊗ (λ,μ)

We now show how to obtain the CG coefficients of SU_3 which arise in the reduction of the direct product

$$(1, 1) \otimes (\lambda, \mu), \tag{3.1}$$

of IRs of SU_3, from those for the direct product (1.1), whose evaluation has been described above. Clearly, the representations (1.1) and (3.1) have the same IR content so we may here refer to cases (I) to (VII) in the sense of (1.2). It is convenient to confine attention, at first, to cases (I) to (VI), which involve those IRs (λ', μ') which do not occur more than once in (3.1), and to employ the notation

$$\begin{aligned} &v = (IMY), \quad k = (JNZ), \quad v' = (I'M'Y'), \\ &\bar{k} = (110), \quad \bar{v}' = (\bar{I}'\bar{I}'\bar{Y}'). \end{aligned}$$

We consider the highest weight states

$$\begin{aligned} &|(\lambda, \mu)(1, 1); (\lambda', \mu')\bar{v}'\rangle \\ &\quad = \sum_{vk} C(\lambda\mu\ 11\ \lambda'\mu'; vk\bar{v}')\ |(\lambda, \mu)v\rangle\ |(1, 1)k\rangle, \end{aligned} \tag{3.2}$$

$$\begin{aligned} &|(1, 1)(\lambda, \mu); (\lambda'\mu')\bar{v}'\rangle \\ &\quad = \sum_{vk} C(11\ \lambda\mu\ \lambda'\mu'; k\bar{v}')\ |(1, 1)k\rangle\ |(\lambda, \mu)v\rangle. \end{aligned} \tag{3.3}$$

Obviously the states (3.2), (3.3) differ only by a real phase factor

$$\begin{aligned} &|(1, 1)(\lambda, \mu); (\lambda', \mu')\bar{v}'\rangle \\ &\qquad = \epsilon_{\lambda'\mu'}\ |(1, 1)(\lambda, \mu); (\lambda', \mu')\bar{v}'\rangle. \end{aligned} \tag{3.4}$$

Further, by applying lowering operators suitably to (3.4), we can, for any other states $|\cdots (\lambda', \mu')v'\rangle$, prove the result

$$\begin{aligned} &|(1, 1)(\lambda, \mu); (\lambda', \mu')v'\rangle \\ &\qquad = \epsilon_{\lambda'\mu'}\ |(1, 1)(\lambda, \mu); (\lambda', \mu')v'\rangle, \end{aligned} \tag{3.5}$$

involving the same real phase factor $\epsilon_{\lambda'\mu'}$ for all v', and hence deduce

$$\begin{aligned} &C(11\ \lambda\mu\ \lambda'\mu', kvv') \\ &\qquad = \epsilon_{\lambda'\mu'}\ C(\lambda\mu\ 11\ \lambda'\mu'; vkv'). \end{aligned} \tag{3.6}$$

The CG coefficients on the right side of (3.6) are thus determined from previous results and Eq. (3.6), to within an over-all real phase. For all of the cases at present in question there is only allowed set of values v for which $C(11\ \lambda\mu\ \lambda'\mu'; \bar{k}v\bar{v}')$ is not identically zero. Since the corresponding CG coefficients never vanish accidentally for any (λ', μ'), we may, in direct generalization of part b of the Condon and Shortley phase convention for SU_2 and in agreement with our discussion of Sec. 1, demand that

$$C(11\ \lambda\mu\ \lambda'\mu'; \bar{k}v\bar{v}')$$

be real and positive. Since $C(\lambda\mu\ 11\ \lambda'\mu'; v\bar{k}\bar{v}')$ is known, this requirement and Eq. (3.6) allows $\epsilon_{\lambda'\mu'}$ to be fixed. For purpose of tabulation of results, it proved convenient to write Eq. (3.6) in the form

$$U(\lambda\mu\ 11\ \lambda'\mu';\ IY\ JZ\ I'Y') = \chi U(11\ \lambda\mu\ \lambda'\mu';\ JZ\ IY\ I'Y'), \quad (3.7)$$

where[32]

$$\chi = (-)^{I+J-I'}\epsilon_{\lambda'\mu'}, \quad (3.8)$$

and to enter beside each $U(\lambda\mu\ 11\ \lambda'\mu';\ IY\ JZ\ I'Y')$ the value of χ by which it must be multiplied to yield the corresponding $U(11\ \lambda\mu\ \lambda'\mu';\ JZ\ IY\ I'Y')$. Next we turn to Case (VII) and the difficulties associated with the double appearance of (λ, μ) in the reductions of (1.1) and (3.1). We again consider the highest weight states

$$|(\lambda, \mu)(1, 1);\ (\lambda, \mu)\bar{v},\ \gamma\rangle = \sum_{vk} C(\lambda\mu\ 11\ \lambda\mu\gamma;\ vk\bar{v})\ |(\lambda, \mu)v\rangle\ |(1, 1)k\rangle, \quad (3.9)$$

where the significance of γ is as explained in Sec. 1, and any orthogonal pair of highest weight states

$$|(1, 1)(\lambda, \mu);\ (\lambda, \mu)\bar{v},\ \delta\rangle = \sum_{vk} C\ 11\ \lambda\mu\ \lambda\mu\delta;\ kv\bar{v})\ |(1, 1)k\rangle\ |(\lambda, \mu)v\rangle, \quad (3.10)$$

distinguished by a label δ with values 1 and 2, in the indicated space of product states. By an orthogonal transformation of the states (3.10) with respect to δ, we can reach states

$$|(1, 1)(\lambda, \mu);\ (\lambda, \mu)\bar{v},\ \gamma\rangle$$

such that

$$|(1, 1)(\lambda, \mu);\ (\lambda, \mu)\bar{v},\ \gamma\rangle = \epsilon_{\lambda\mu\gamma}\ |(\lambda, \mu)(1, 1);\ (\lambda, \mu)\bar{v},\ \gamma\rangle, \quad (3.11)$$

where $\epsilon_{\lambda\mu\gamma}$, $\gamma = 1, 2$, are real phase factors which are determined appropriately below. Since all states of the type $|(1, 1)(\lambda, \mu);\ (\lambda, \mu)v,\ \gamma\rangle$ are obtained by applying lowering operators suitably to those of highest weight, Eq. (3.11) allows the definition to within an over-all real phase factor, for each value of γ, of all such states and of the corresponding CG coefficients, for which we have the result

$$C(11\ \lambda\mu\ \lambda\mu\gamma;\ kvv') = \epsilon_{\lambda\mu\gamma}C(\lambda\mu\ 11\ \lambda\mu\gamma;\ vkv'), \quad (3.12)$$

with $\epsilon_{\lambda\mu\gamma}$ as before, independent of v, k, and v'. Essentially what we have done is to use Biedenharn's canonical resolution of the multiplicity problem for the product (3.1) to define a cononical resolution of the multiplicity problem for the product (3.1). We can fix the $\epsilon_{\lambda\mu\gamma}$ by consideration of the CG coefficients

$$C(11\ \lambda\mu\ \lambda\mu\gamma;\ \bar{k}v\bar{v}), \quad (3.13)$$

which vanish identically unless $v = (\bar{I}\bar{I}-1\bar{Y})$ or $v = (\bar{I}-1\bar{I}-1\bar{Y})$. We demand that

$$C(11\ \lambda\mu\ \lambda\mu 1;\ 110\ \bar{I}\bar{I}-1\bar{Y}\ \bar{I}\bar{I}\bar{Y}), \quad (3.14)$$

$$C(11\ \lambda\mu\ \lambda\mu 2;\ 110\ \bar{I}-1\bar{I}-1\bar{Y}\ \bar{I}\bar{I}\bar{Y}), \quad (3.15)$$

be real and positive, since, for $v = (\bar{I}\bar{I}-1\bar{Y})$, accidental vanishing of the $\gamma = 2$ coefficient (3.13) can occur. We can then determine $\epsilon_{\lambda\mu\gamma}$ using (3.12) and the known forms of $C(\lambda\mu\ 11\ \lambda\mu\gamma;\ v\bar{k}\bar{v})$. As above, we present results by tabulating along with

$$U(\lambda\mu\ 11\ \lambda\mu\gamma;\ IY\ JZ\ I'Y')$$

the factor $\chi = (-)^{I+J-I'}\epsilon_{\lambda\mu\gamma}$ by which it must be multiplied to yield

$$U(11\ \lambda\mu\ \lambda\mu\gamma;\ JZ\ IY\ I'Y').$$

[32] Equation (1.5) and

$$C(IJI', MNM') = (-)^{I+J-I'}C(JII', NMM')$$

have been used.

ACKNOWLEDGMENTS

It is a pleasure to thank Dr. N. Mukunda for many enlightening discussions. We also thank Dr. G. Costa for having drawn our attention to certain errors that were present in an earlier version of or tables.

On "Diagonal" Coherent-State Representations for Quantum-Mechanical Density Matrices*

JOHN R. KLAUDER AND JAMES MCKENNA
Bell Telephone Laboratories, Incorporated, Murray Hill, New Jersey
AND
DOUGLAS G. CURRIE
Department of Physics and Astronomy, University of Maryland, College Park, Maryland
(Received 25 September 1964)

It is proved that every density matrix is the limit, in the sense of weak operator convergence, of a sequence of operators each of which may be represented as an integral over projection operators onto coherent states (in the sense of Glauber) with a square-integrable weight function. This result is a special case of one that holds for all operators with trace and for overcomplete families of states other than just the coherent states. We prove our more general result, at no cost of complexity, within the more general framework of continuous-representation theory. The significance of our results for representing traces of operators is indicated.

I. INTRODUCTION

In this note we should like to establish a theorem pertaining to the representation of quantum-mechanical density operators as integrals over projection operators $|\alpha\rangle\langle\alpha|$ onto coherent state vectors $|\alpha\rangle$. Such "diagonal" representations have been introduced in the context of optical coherence by Glauber[1] and by Sudarshan,[2] and the latter author has also emphasized the exceptional generality of such representations.

In the notation of Dirac, let $|n\rangle$, $n = 0, 1, 2, \cdots$, denote the energy eigenstates for a harmonic oscillator of unit angular frequency. In terms of these, we define the coherent state

$$|\alpha\rangle \equiv e^{-\frac{1}{2}|\alpha|^2} \sum_{n=0}^{\infty} \frac{\alpha^n}{(n!)^{\frac{1}{2}}} |n\rangle, \tag{1}$$

and its adjoint

$$\langle\alpha| \equiv e^{-\frac{1}{2}|\alpha|^2} \sum_{n=0}^{\infty} \frac{\alpha^{*n}}{(n!)^{\frac{1}{2}}} \langle n|,$$

for all complex α. These states have unit norm, are not all independent, but are complete in the sense that the unit operator may be expressed in the form

$$1 = \pi^{-1} \int |\alpha\rangle\langle\alpha| \, d^2\alpha, \tag{2}$$

where for $\alpha = \alpha_r + i\alpha_i$, we set $d^2\alpha = d\alpha_r d\alpha_i$, the integral extending over all values of α_r and α_i. The form (2) for the decomposition of unity in terms of the states of Eq. (1) appears to have been first stated and used by one of the present authors in applications other than optical coherence.[3] By a double application of Eq. (2), any bounded operator $\mathcal{B}$ admits the decomposition

$$\mathcal{B} = \pi^{-2} \iint |\alpha\rangle\langle\alpha| \, \mathcal{B} \, |\beta\rangle\langle\beta| \, d^2\alpha \, d^2\beta. \tag{3}$$

However, every function $b(\alpha) \equiv b(\alpha_r, \alpha_i)$ defines an operator of the "diagonal" form

$$\pi^{-1} \int b(\alpha) \, |\alpha\rangle\langle\alpha| \, d^2\alpha \equiv \mathcal{B} \tag{4}$$

whenever $b(\alpha)$ is bounded (and measurable) or if it satisfies

$$\int |b(\alpha)|^2 \, d^2\alpha < \infty.$$

Such operators include certain "diagonal" quantum-mechanical density operators

$$\rho = \pi^{-1} \int P(\alpha) \, |\alpha\rangle\langle\alpha| \, d^2\alpha, \tag{5}$$

which have enjoyed considerable application in optical coherence studies.[1,2,4,5] Furthermore, the diagonal representation of (5) entails computational simplifications and imparts a c-number classical appearance to traces involving normal-ordered operators.[2] These properties give rise to the question of how broad a representation is afforded by (5). While not fully answering this question, we have

* An account of this paper was presented at the Conference on Quantum Electrodynamics of High Intensity Photon Beams, held at Durham, North Carolina, August 1964.

[1] R. J. Glauber, Phys. Rev. Letters **10**, 84 (1963); Phys. Rev. **131**, 2766 (1963).

[2] E. C. G. Sudarshan, Phys. Rev. Letters **10**, 277 (1963); Proceedings of the Symposium on Optical Masers (Polytechnic Press, Brooklyn, New York and J. Wiley & Sons, Inc., New York, 1963), p. 45.

[3] J. R. Klauder, Ann. Phys. (N. Y.) **11**, 123 (1960), Eqs. (4) and (5).

[4] L. Mandel, Phys. Letters **7**, 117 (1963); Phys. Rev. **134**, A10 (1964).

[5] C. L. Mehta and E. Wolf, Phys. Rev. **134**, A1149 (1964).

obtained the result embodied in the following theorem.

Theorem 1. Every density operator ρ is the weak limit of a sequence of bounded operators each of which admits a diagonal representation in the sense of (4). In particular, this means that for every density operator ρ, there exists a sequence of square-integrable functions $b_n(\alpha)$ such that

$$\langle 1|\ \rho\ |2\rangle = \lim_{n\to\infty} \int b_n(\alpha)\ \langle 1\ |\ \alpha\rangle\ \langle\alpha\ |\ 2\rangle\ d^2\alpha,$$

for all pairs of states $|1\rangle$ and $|2\rangle$.

When $b(\alpha)$ is a tempered distribution, then it can be found by taking the diagonal coherent matrix elements of (4) and applying the convolution theorem to the left-hand side. For some ρ, however, $b(\alpha)$ is not a tempered distribution (as shown in Appendix B); the crux of the theorem is that even these cases can be handled through a sequence of square-integrable $b_n(\alpha)$. A suitable set of functions $b_n(\alpha)$ is explicitly given in Sec. II in the course of our proof. We caution the reader that in general it is *not* legitimate to take the limit under the integral sign in the statement of Theorem 1.

We should like to emphasize that both the "diagonal" representation—in the sense of Theorem 1—and a generally distinct off-diagonal representation as in (3) exist simultaneously for a density matrix. This situation is made possible by the *overcompleteness* of the coherent states, and is roughly analogous to the various ways of representing a given electromagnetic field by changing the gauge freedom in the four-vector potential.

Instead of directly proving Theorem 1, we should like to prove a more general theorem whose truth implies Theorem 1 and which is the theorem we actually prove.

Formulation in Terms of Phase-Space Continuous Representations

Theorem 1 is actually a special case of another theorem which holds for any operator in the trace class and for representations more general than those afforded by coherent states. An operator $\mathfrak{G}$ is said to be in the trace class if, given any complete, orthonormal sequence $\{\Psi_i\}$, the series $\sum_{i=1}^{\infty}\ (\Psi_i,\ \mathfrak{G}\Psi_i)$ converges absolutely. When absolutely convergent, the value of the series is independent of the orthonormal sequence $\{\Psi_i\}$ and is defined to be the trace of $\mathfrak{G}$. To define the more general representations we refer to, we replace the set of coherent states by the notion of an overcomplete family of states[6] for a single quantum-mechanical degree of freedom,[7] the properties of which have been thoroughly studied elsewhere.[8] In standard Hilbert-space notation as used in quantum theory, the members of a phase-space overcomplete family of states are unit vectors defined by

$$\Phi[p,\ q] = e^{-iqP}e^{ipQ}\Phi_0 \tag{6}$$

for all real p and q. In (6), P and Q denote irreducible, self-adjoint momentum and position operators (we set $\hbar = 1$), and Φ_0 is a unit vector called the fiducial vector. For any fiducial vector, it has been rigorously established in CRT IV that the unit operator may be expressed as

$$1 = \iint \Phi[p,\ q]\Phi[p,\ q]^{\dagger}(dp\ dq/2\pi), \tag{7}$$

the integration extending over the entire phase space. Clearly, Eq. (7) is an analogue of Eq. (2). Indeed, if Φ_0 fulfills $(Q + iP)\Phi_0 = 0$, and thus corresponds to the ground state of a harmonic oscillator, then $\Phi[p,\ q]$ differs in an inessential way from a coherent state; namely, in this case

$$e^{ipq/2}\Phi[p,\ q] = |2^{-\frac{1}{2}}(q + ip)\rangle. \tag{8}$$

However, the oscillator ground state is but one of all possible unit vectors for which (7) holds true.

Just as in Eq. (4), we may define a bounded operator $\mathfrak{G}$ in the "diagonal" form

$$\mathfrak{G} = \iint b(p,\ q)\Phi[p,\ q]\Phi[p,\ q]^{\dagger}(dp\ dq/2\pi) \tag{9}$$

for every function $b(p,\ q) \in L^2(R \times R)$ or for every bounded and measurable function $b(p,\ q)$. This construction holds for an arbitrary fiducial vector Φ_0, and has been rigorously defined and analyzed in CRT IV, Sec. 3. Hence, we can generalize the earlier question concerning "diagonal" representations involving coherent states and investigate operators exhibiting a "diagonal" representation for a Φ_0 other than a oscillator ground state. In our generalization, we are led in this paper to restrict attention to those Φ_0 for which the reproducing kernel

$$\mathcal{K}(p,\ q;\ p',\ q') \equiv (\Phi[p,\ q],\ \Phi[p',\ q']) \tag{10}$$

never vanishes. This condition is fulfilled for an harmonic oscillator fiducial vector $\{|\langle\alpha\ |\ \beta\rangle| = \exp\ (-\frac{1}{2}\ |\alpha - \beta|^2)\}$, but there exist many other choices of Φ_0 for which the kernel $\mathcal{K}$ never vanishes. Numerous examples have been found by one of us

[6] J. R. Klauder, J. Math. Phys. 4, 1055 (1963).

[7] J. R. Klauder, J. Math. Phys. 4, 1058 (1963); 5, 177 (1964), referred to as CRT III.

[8] J. McKenna and J. R. Klauder, J. Math. Phys. 5, 878 (1964) referred to as CRT IV.

(JM) and were cited in CRT III, Sec. 2. We can now state our more general theorem precisely, for which Theorem 1 is a special case:

Theorem 2. Consider any fiducial vector Φ_0 whose associated reproducing kernel, Eq. (10), never vanishes. Then every operator in the trace class (hence every density matrix ρ) is the weak limit of a sequence of bounded operators each of which admits a diagonal representation in the sense of (9). In particular, this means that, if $\mathcal{K}$ never vanishes, for every operator $\mathcal{B}$ in the trace class there exists a sequence of square-integrable functions $b_n(p, q)$ such that

$$(\Omega, \mathcal{B}\Lambda) = \lim_{n\to\infty} \iint b_n(p, q)(\Omega, \Phi[p, q]) \times (\Phi[p, q], \Lambda)(dp\, dq/2\pi) \quad (11)$$

for all pairs of vectors Ω and Λ in Hilbert space.

It is perhaps worth remarking that the complete set of fiducial vectors whose reproducing kernels never vanish is not yet known; at present the only examples known are those cited in CRT III. Moreover, any special physical significance such fiducial vectors may have remains obscure at present.

While both Theorems 1 and 2 have been stated for a single degree of freedom it is clear that analogous theorems hold for any finite number of degrees of freedom; for simplicity we confine our discussion to a single degree of freedom. It is to the proof of Theorem 2 that the next section and Appendix A are devoted.

The significance of these results for representing quantum-mechanical traces is studied in Sec. III. The reader more interested in this aspect may proceed directly to Sec. III.

II. PROOF OF GENERAL THEOREM

A bounded operator $\mathcal{B}$ is in the trace class if and only if it has the polar decomposition

$$\mathcal{B} = \sum_{j=1}^{\infty} \beta_j \Psi_j \mathrm{X}_j^{\dagger} \quad (12)$$

with

$$\beta_j \geq 0, \quad (13a)$$

$$\sum_{j=1}^{\infty} \beta_j < \infty, \quad (13b)$$

where $\{\Psi_j\}$ and $\{\mathrm{X}_j\}$ are two orthonormal sequences.[9]

[9] R. Schatten, *Norm Ideals of Completely Continuous Operators* (Springer-Verlag, Berlin, 1960). It should be noted that $\mathrm{Tr}\, \mathcal{B} = \sum_{j=1}^{\infty} \beta_j (\mathrm{X}_j, \Psi_j)$. $\mathcal{B}$ is a density operator if and only if $\mathrm{X}_j = \Psi_j$ and $\sum_{j=1}^{\infty} \beta_j = 1$.

Let us assume first that the operator $\mathcal{B}$ of (12) can be written in the form (9) and evaluate the "diagonal matrix elements" $B(r, s) \equiv (\Phi[r, s], \mathcal{B}\Phi[r, s])$. It follows that

$$B(r, s) \equiv \sum_{j=1}^{\infty} \beta_j \psi_j(r, s)\chi_j^*(r, s)$$

$$= \int b(p, q)\, |\mathcal{K}(r - p, s - q; 0, 0)|^2\, (dp\, dq/2\pi) \quad (14)$$

where we have introduced the definitions

$$\psi_j(r, s) = (\Phi[r, s], \Psi_j), \quad (15a)$$

$$\chi_j(r, s) = (\Phi[r, s], \mathrm{X}_j), \quad (15b)$$

and have exploited the property

$$\mathcal{K}(p, q; r, s) = e^{ip(q-s)}\mathcal{K}^*(r - p, s - q; 0, 0)$$

of the reproducing kernel [CRT IV, Eqs. (5.4b), (5.6)].

Motivated by the difference form of the kernel in (14), we examine the existence and value of the Fourier transform of Eq. (14). The left-hand side of (14) is integrable and may be integrated term-by-term[10] since

$$\sum_{j=1}^{\infty} \beta_j \int |\psi_j(r, s)\chi_j^*(r, s)|\, (dr\, ds/2\pi) \leq \sum_{j=1}^{\infty} \beta_j < \infty$$

holds in virtue of Schwartz's inequality applied to the basic formula [CRT IV, Eq. (3.16)]

$$\int \chi_j^*(r, s)\psi_j(r, s)(dr\, ds/2\pi) = (\mathrm{X}_j, \Psi_j).$$

In consequence, we may compute the Fourier transform

$$\tilde{B}(x, k) = \iint e^{-irx+isk} B(r, s)(dr\, ds/2\pi)$$

of the left side of (14) term by term. According to the calculation carried out in Appendix A, this result is

$$\tilde{B}(x, k) = \mathcal{K}(k, x; 0, 0) \sum_{j=1}^{\infty} \beta_j J_{\psi_j \chi_j}(x, k), \quad (16)$$

where

$$J_{\psi_j \chi_j}(x, k) \equiv \int \psi_j(y) e^{iyk} \chi_j^*(y + x)\, dy = (\mathrm{X}_j, U[k, x]\Psi_j).$$

Here $\psi_j(y)$ and $\chi_j(y)$ denote the image in $L^2(R)$ of Ψ_j and X_j, respectively, in the Schrödinger representation.

[10] P. R. Halmos, *Measure Theory* (D. Van Nostrand Company, New York, 1950), p. 114.

We now assume that the weight function $b(p, q)$ in (14) is square integrable, which in consequence ensures that $\mathfrak{B}$ has the form (9) and supersedes our explicit assumption to that effect. With $b(p, q) \in L^2(R \times R)$, we can Fourier transform the right side of (14) using the convolution theorem[11] and equate the result to $\tilde{B}(x, k)$. Using Eq. (A4) in the special case (A6d), we find that

$$\tilde{B}(x, k) = \tilde{b}(x, k)\, |\mathcal{K}(k, x; 0, 0)|^2. \quad (17)$$

If we now restrict ourselves to fiducial vectors such that $\mathcal{K}$ never vanishes—a condition we henceforth assume fulfilled—then (16) leads to

$$\tilde{b}(x, k) = \sum_{i=1}^{\infty} \frac{\beta_i J_{\psi_i, \chi_i}(x, k)}{\mathcal{K}^*(k, x; 0, 0)}. \quad (18)$$

The numerator of (18),

$$N(x, k) = \sum_{i=1}^{\infty} \beta_i J_{\psi_i \chi_i}(x, k), \quad (19)$$

is bounded [from Eq. (13) and (A6a)] and square integrable because the $J_{\psi_i \chi_i}(x, k)$ form an orthonormal sequence in $L^2(R \times R)$ and $\sum_{i=1}^{\infty} \beta_i^2 < \infty$. Since $\mathcal{K}(k, x; 0, 0)$ never vanishes and is everywhere continuous (cf. CRT IV, Secs. 3,5), the denominator in (18) is bounded away from zero on every compact set. However, these conditions on N and $\mathcal{K}$ are insufficient to ensure that

$$\tilde{b}(x, k) \in L^2(R \times R),$$

a condition we have assumed in Fourier-transforming the right side of (14) with the aid of the convolution theorem. An example is given in Apprendix B of a $\tilde{b}(x, k) \notin L^2(R \times R)$.

To surmount this difficulty we appeal to the notion of weak operator convergence,[12] and try to construct our trace-class operator $\mathfrak{B}$ in that fashion. The approximating sequence is of course not unique, but our example serves to prove our theorem.

Let

$$\chi_n(x, k) = 1; \qquad |x| \leq n, \qquad |k| \leq n,$$
$$= 0; \qquad \text{otherwise},$$

and set

$$\tilde{b}_n(x, k) = \chi_n(x, k)\tilde{b}(x, k). \quad (20)$$

For each n, the function $\tilde{b}_n(x, k)$ is bounded and has compact support, hence it is both integrable and square integrable. Its Fourier transform

$$b_n(p, q) = \iint \tilde{b}_n(x, k) e^{ipx - iqk} \left(\frac{dx\, dk}{2\pi}\right)$$

is likewise square integrable, which ensures that

$$\mathfrak{B}_n \equiv \iint b_n(p, q)\Phi[p, q]\Phi[p, q]^\dagger \left(\frac{dp\, dq}{2\pi}\right) \quad (21)$$

is well defined. We now show that

$$\lim_{n\to\infty} (\Omega, \mathfrak{B}_n \Lambda) = (\Omega, \mathfrak{B}\Lambda)$$

for all Ω and Λ in the Hilbert space.

From Parseval's theorem and Eq. (A4) we find

$$(\Omega, \mathfrak{B}_n\Lambda) = \iint b_n(p, q)[\omega(p, q)\lambda^*(p, q)]^* \left(\frac{dp\, dq}{2\pi}\right)$$
$$= \iint \tilde{b}_n(x, k)\mathcal{K}^*(k, x; 0, 0) J^*_{\omega,\lambda}(x, k) \left(\frac{dx\, dk}{2\pi}\right).$$

From Eqs. (18), (19), and (20) for $\tilde{b}_n$ we find

$$(\Omega, \mathfrak{B}_n\Lambda) = \int_{-n}^{n}\int_{-n}^{n} N(x, k) J^*_{\omega,\lambda}(x, k) \left(\frac{dx\, dk}{2\pi}\right). \quad (22)$$

Since each factor in this integrand is in $L^2(R \times R)$ the limit of the integral as $n \to \infty$ exists and is

$$\lim_{n\to\infty} (\Omega, \mathfrak{B}_n\Lambda) = \iint N(x, k) J^*_{\omega,\lambda}(x, k) \left(\frac{dx\, dk}{2\pi}\right).$$

To evaluate the right-hand side we substitute in Eq. (19). Summation and integration may be interchanged since

$$\sum_{i=1}^{\infty} \beta_i \iint |J_{\psi_i \chi_i}(x, k) J^*_{\omega,\lambda}(x, k)| \left(\frac{dx\, dk}{2\pi}\right)$$
$$\leq \sum_{i=1}^{\infty} \beta_i\, ||\Omega||\, ||\Lambda|| < \infty,$$

in virtue of (A6b). Thus, we have from (A6b) and (12) that

$$\lim_{n\to\infty} (\Omega, \mathfrak{B}_n\Lambda) = \sum_{i=1}^{\infty} \beta_i \int J_{\psi_i \chi_i}(x, k) J^*_{\omega,\lambda}(x, k) \left(\frac{dx\, dk}{2\pi}\right)$$
$$= \sum_{i=1}^{\infty} \beta_i (\Omega, \Psi_i)(\mathrm{X}_i, \Lambda),$$
$$= \left(\Omega, \left[\sum_{i=1}^{\infty} \beta_i \Psi_i \mathrm{X}_i^\dagger\right]\Lambda\right),$$
$$= (\Omega, \mathfrak{B}\Lambda),$$

as was to be shown. This concludes the proof of Theorem 2 that $\mathfrak{B}$ is the weak limit of the operator sequence $\mathfrak{B}_n$ each member of which had a diagonal representation in the manner of (9), under the assumption that the reproducing kernel never vanishes.

[11] E. C. Titchmarsh, *Theory of Fourier Integrals* (Oxford University Press, London, 1948), p. 90.

[12] F. Riesz and B. Sz.-Nagy, *Functional Analysis* (Frederick Ungar Publishing Company, New York, 1955), p. 150.

Note added in manuscript: Although we have confined our discussion to weak operator convergence, the sequence $\{\mathcal{B}_n\}$ actually converges in norm.[12] From (22), it follows that

$$(\Omega, (\mathcal{B}_n - \mathcal{B}_m)\Lambda) = \iint [\chi_n(x, k) - \chi_m(x, k)]N(x, k) \times (\Lambda, U[k, x]\Omega)^*(dk\, dx/2\pi).$$

If we set $\Omega = (\mathcal{B}_n - \mathcal{B}_m)\Lambda$, then

$$||(\mathcal{B}_n - \mathcal{B}_m)\Lambda||^2 = \iint [\chi_n(x, k) - \chi_m(x, k)]N(x, k) \times (\Lambda, U[k, x](\mathcal{B}_n - \mathcal{B}_m)\Lambda)^* (dk\, dx/2\pi),$$

which by Schwartz's inequality and an application of CRT IV (3.16) leads to

$$||(\mathcal{B}_n - \mathcal{B}_m)\Lambda||^2 \leq N_{nm}\, ||(\mathcal{B}_n - \mathcal{B}_m)\Lambda|| \cdot ||\Lambda||,$$

where

$$N_{nm} \equiv \left\{\iint [\chi_n(x, k) - \chi_m(x, k)]^2\, |N(x, k)|^2 \left(\frac{dk\, dx}{2\pi}\right)\right\}^{\frac{1}{2}}.$$

Therefore, we find that

$$||(\mathcal{B}_n - \mathcal{B}_m)\Lambda||/||\Lambda|| \leq N_{nm}.$$

Since the N_{nm} are independent of Λ, this inequality implies that

$$\sup_{\Lambda} ||(\mathcal{B}_n - \mathcal{B}_m)\Lambda||/||\Lambda|| \equiv ||\mathcal{B}_n - \mathcal{B}_m|| \leq N_{nm}.$$

Finally, in view of the square integrability of $N(x, k)$, the sequence $N_{nm} \to 0$ as $m, n \to \infty$. Therefore, the sequence $\mathcal{B}_n$ converges in norm, as was to be shown. Moreover, it can be shown that the operators $\mathcal{B}_n$ are in the Hilbert–Schmidt class,[9] and converge in the Hilbert–Schmidt norm. In fact,

$$\mathrm{Tr}\, [(\mathcal{B}_n^\dagger - \mathcal{B}_m^\dagger)(\mathcal{B}_n - \mathcal{B}_m)] = N_{nm}^2,$$

which will be proved in a subsequent publication.

We thank Professor Sudarshan for a private communication in which the question of stronger forms of convergence was raised.

III. DISCUSSION

Lest there be some uncertainty as to the meaning of our general theorem, and its application to evaluating traces of quantum-mechanical operators, let us make some additional remarks. If ρ is in the trace class and $\mathcal{A}$ is an arbitrary bounded operator then $\mathrm{Tr}\, (\rho\mathcal{A})$ exists[9] and is always given by

$$\mathrm{Tr}\, (\rho\mathcal{A}) = \sum_{j=1}^{\infty} (\Psi_j, \rho\mathcal{A}\Psi_j) = \sum_{j=1}^{\infty} \lim_{n\to\infty} \iint b_n(p, q)A_j(p, q)\left(\frac{dp\, dq}{2\pi}\right), \quad (23)$$

where

$$A_j(p, q) \equiv (\Psi_j, \Phi[p, q])(\Phi[p, q], \mathcal{A}\Psi_j), \quad (24)$$

and $\{\Psi_j\}$ is a complete orthornormal sequence. This result is a direct consequence of Theorem 2 applied to each term in the sum defining the trace. It would be nice to be able to interchange the summation and limiting operations in (23), but in general, this is not possible. However, if in addition to being bounded, the operator $\mathcal{A}$ is itself in the trace class, we may appeal to the polar decomposition form for $\mathcal{A}$ to examine (23) more closely. A calculation patterned after that given in Sec. II then justifies the interchange of the summation with the remaining operations. Thus, in the case that ρ and $\mathcal{A}$ both are in the trace class we find that

$$\mathrm{Tr}\, (\rho\mathcal{A}) = \lim_{n\to\infty} \iint b_n(p, q)A(p, q)\left(\frac{dp\, dq}{2\pi}\right), \quad (25)$$

$$A(p, q) = \sum_{j=1}^{\infty} A_j(p, q) = (\Phi[p, q], \mathcal{A}\Phi[p, q]) \quad (26)$$

always holds.

It is natural to ask when the limiting operations on n are unnecessary in all our formulas. We give a sufficient condition in the important case of coherent-state representations, where Φ_0 is an oscillator ground state. Since $N(x, k)$, as defined in Eq. (19), is the uniformly convergent sum of continuous functions, it likewise is continuous. If in addition the continuous function $b(x, k)$, defined in Eq. (18), is bounded by a polynomial in x and k, then its Fourier transform $b(p, q)$ is a tempered distribution, and its operation on all functions

$$\omega^*(p, q)\lambda(p, q) \equiv (\Omega, \Phi[p, q])(\Phi[p, q], \Lambda)$$

is well defined. This implies that Eq. (9) defines a bona fide operator.

ACKNOWLEDGMENTS

Two of the authors (JRK and JM) acknowledge discussions with Professor Glauber and Professor Sudarshan that have largely prompted the present investigation. One of the authors (DGC) is grateful for the support of ARPA and Bell Telephone Laboratories during this work.

APPENDIX A

We wish to evaluate the expression

$$F(x, k) = \iint e^{-ipx+iqk}\psi(p, q)\chi^*(p, q)\left(\frac{dp\, dq}{2\pi}\right), \quad (A1)$$

for

$$\psi(p, q) \equiv (\Phi[p, q], \Psi) = \int \varphi_0^*(y)e^{-ipy}\psi(y + q)\, dy, \quad \text{(A2a)}$$

$$\chi(p, q) \equiv (\Phi[p, q], \mathrm{X}) = \int \varphi_0^*(z)e^{-ipz}\chi(z + q)\, dz, \quad \text{(A2b)}$$

where $\varphi_0(y)$, $\psi(y)$ and $\chi(y)$ denote the image in $L^2(R)$ in the Schrödinger representation of Φ_0, Ψ and X, respectively. The integral in (A1) is absolutely convergent so we may integrate over p first[13]:

$$F(x, k) = (2\pi)^{-1} \int e^{iqk}\, dq \int e^{-ipx}\psi(p, q)\chi^*(p, q)\, dp.$$

Since $\psi(p, q)$ and $e^{-ipx}\chi^*(p, q)$ are both square integrable in p and q, we can, for almost all q, evaluate the inner integral over p by Parseval's theorem and find

$$F(x, k) = \int dq\, e^{iqk} \times \int \varphi_0^*(y)\psi(y + q)\varphi_0(y + x)\chi^*(y + q + x)\, dy, \quad \text{(A3)}$$

since, e.g., $\psi(p, q)$ is already the Fourier transform over p of $\varphi_0^*(y)\psi(y + q) \in L^2(R)$ for almost all q, as was shown in CRT IV, Sec. 3. An interchange of the y and q integrations, as permitted by the absolute integrability of (A3), followed by a simple change of variables yields

$$F(x, k) = \int dy\, e^{-iyk}\varphi_0^*(y)\varphi_0(y + x) \times \int dq\, e^{iqk}\psi(q)\chi^*(q + x) = \mathcal{K}(k, x; 0, 0) \int dq\, e^{iqk}\psi(q)\chi^*(q + x), \quad \text{(A4)}$$

which is the desired expression for each term in (17).

If we let

$$J_{\psi\chi}(x, k) = \int \psi(y)e^{iyk}\chi^*(y + x)\, dy = (\mathrm{X}, U[k, x]\Psi), \quad \text{(A5)}$$

then the results of CRT IV readily show that

(a) $$|J_{\psi\chi}(x, k)| \leq ||\Psi||\, ||\mathrm{X}||, \quad \text{(A6a)}$$

(b) $$\iint J_{\psi\chi}^*(x, k)J_{\psi'\chi'}(x, k)(dx\, dk/2\pi) = (\Psi, \Psi')(\mathrm{X}', \mathrm{X}), \quad \text{(A6b)}$$

while the following arise by definition,

(c) $$J_{\varphi_0\psi}(x, k) = \psi^*(k, x), \quad \text{(A6c)}$$

(d) $$J_{\varphi_0\varphi_0}(x, k) = \mathcal{K}^*(k, x; 0, 0), \quad \text{(A6d)}$$

where $\varphi_0(y)$ denotes the image of the fiducial vector.

APPENDIX B

An example for which $\tilde{b}(x, k)$ is not square integrable is readily found. Choose the "coherent-state representation" afforded by requiring $(Q+iP)\Phi_0=0$. For the operator $\mathfrak{B}$, choose $\mathfrak{B} = \Psi\Psi^\dagger$, where Ψ is the normalized solution of $(\omega Q + iP)\Psi = 0$, i.e., an oscillator ground state with frequency $\omega \neq 1$. In this case, it readily follows that the "diagonal" elements

$$B(r, s) = |\psi(r, s)|^2 = 2\omega^{\frac{1}{2}}(1 + \omega)^{-1} \exp[-(1 + \omega)^{-1}(\omega s^2 + r^2)], \quad \text{(B1)}$$

and thus that

$$\tilde{B}(x, k) = \exp[-\tfrac{1}{4}(1 + \omega)(x^2 + \omega^{-1}k^2)]. \quad \text{(B2)}$$

For this case, Eq. (17) yields

$$\tilde{b}(x, k) = \exp\{-\tfrac{1}{4}[x^2(\omega - 1) + k^2(\omega^{-1} - 1)]\}, \quad \text{(B3)}$$

and expression which *grows exponentially* in either x^2 or k^2, depending on the size of ω. Hence, $\tilde{b}(x, k)$ for this density matrix is not square integrable.

Moreover, since $\tilde{b}(x, k)$ grows faster than a polynomial in x or k, the possibility that $b(p, q)$ can be interpreted as a tempered distribution in the sense of Schwartz[14] is ruled out in this case. For if $b(p, q)$ were a tempered distribution, the convolution theorem would still be valid, and $\tilde{b}(x, k)$ would also be a tempered distribution.[14] However, in the example just given, $\tilde{b}(x, k)$ cannot be a tempered distribution since it grows faster than a polynomial at infinity.[14] The example treated in this Appendix may shed some light on the nature of the expression for $b(\alpha)$ given by Sudarshan.[2]

[13] E. J. McShane, *Integration* (Princeton University Press, Princeton, New Jersey 1944), p. 145.

[14] L. Schwartz, *Theorie des distributions*, (Hermann & Cie., Paris, 1957), Vol. II.

S-Operator Theory. I. Formulation*

ROBERT E. PUGH
University of Toronto, Toronto, Canada
(Received 26 August 1964)

It is postulated that the S operator is dynamically independent of the wavefunctions of the asymptotic states. From this postulate a functional differential equation for the S operator is developed. Its solutions include both the renormalized and unrenormalized Feynman–Dyson S operators; appropriate boundary conditions distinguish the two solutions. Interpolating quantum fields are defined in terms of the S operator and are only of secondary importance in this theory. Calculations in this theory are not appreciably more difficult than the corresponding calculations with the unrenormalized S operator.

A. INTRODUCTION

In the past ten years several formulations of quantum field theory that avoid ultraviolet divergences have been developed. The basic formulation is due to Lehmann, Symanzik, and Zimmermann[1]; others[2,3] are extensions of the ideas introduced in the basic formulation. In each of these formulations one obtains an infinite set of integrodifferential equations for the S-matrix elements off the mass shell.

In this paper the infinite set of equations for S-matrix elements is replaced by one equation for the S operator. This S-operator equation has the form of a second-order differential equation wherein the derivatives are functional derivatives with respect to free field operators.

For simplicity we deal with Hermitian scalar fields throughout this paper. The extension to other types of fields is straightforward; the special problems associated with quantum electrodynamics form the substance of a subsequent paper.

The functional differential calculus of free fields is discussed in Sec. B. In C, the basic formulation of S-operator theory is developed, and in D this formulation is shown to be equivalent to a previous formulation.[3] In E it is shown that the unrenormalized Feynman–Dyson S operator is a solution. A sample calculation is shown in Sec. F.

B. MATHEMATICAL INTRODUCTION

Given an Hermitian scalar field, $\varphi_{\text{in}}(x)$, satisfying

$$K_x\varphi_{\text{in}}(x) = (\Box_x - m^2)\varphi_{\text{in}}(x) = 0, \tag{1}$$

$$[\varphi_{\text{in}}(x), \varphi_{\text{in}}(y)] = -i\Delta(x - y), \tag{2}$$

$$\varphi_{\text{in}}^{(+)}(x)\,|0\rangle = 0, \tag{3}$$

we wish to define a functional derivative having the properties[4]

$$\delta\varphi_{\text{in}}(x)/\delta\varphi_{\text{in}}(y) = \delta(x - y), \tag{4}$$

$$\frac{\delta}{\delta\varphi_{\text{in}}(y)} :\varphi_{\text{in}}(x_1) \cdots \varphi_{\text{in}}(x_n):$$

$$= \sum_i \delta(x_i - y)\, :\varphi_{\text{in}}(x_1) \cdots \Lambda_i \cdots \varphi_{\text{in}}(x_n):. \tag{5}$$

Here Λ_i indicates that $\varphi_{\text{in}}(x_i)$ is omitted from the product. (In general the symbol $\Lambda_{i,j,\ldots,k}$ will indicate that the ith, jth, $\cdots$, kth terms of a product or sequence are all absent.) To streamline the notation we shall write $\delta/\delta x_i$ for $\delta/\delta\varphi_{\text{in}}(x_i)$.

In general a functional of the free field $\varphi_{\text{in}}(x)$ has the form

$$F[f] = \sum_{n=0}^{\infty} F_n[f_n] \tag{6}$$

with

$$F_n[f_n] = \frac{1}{n!}\int d^4\xi_1 \cdots d^4\xi_n\, f_n(\xi_1 \cdots \xi_n) \times :\varphi_{\text{in}}(\xi_1) \cdots \varphi_{\text{in}}(\xi_n):. \tag{7}$$

Without loss of generality we assume that f_n is a symmetric function of its variables. The restriction, $K\varphi_{\text{in}} = 0$, is equivalent to the invariance of F under the transformations

$$f_n \to f_n + df_n, \tag{8}$$

where $df_n(\xi_1 \cdots \xi_n)$ is any generalized function that vanishes on the mass shells of the variables, $\xi_1, \cdots, \xi_n$. An important class of functions, Df_n, is of the form

$$D_{ij}f_n(\xi_1 \cdots \xi_n) = g(\xi_1 \cdots \Lambda_{ij} \cdots \xi_n)K_i\delta(\xi_i - \xi_j). \tag{9}$$

Other members of this class may have more δ functions and more K's. The importance of this class of functions is that given any function f one can always separate it into parts of the form (9) and parts that are not of the form (9). This separation

* This work was supported by the National Research Council of Canada.

[1] H. Lehmann, R. Symanzik, and W. Zimmermann, Nuovo Cimento **1**, 205 (1955).

[2] K. Nishijima, Phys. Rev. **119**, 485 (1960).

[3] R. Pugh, Ann. Phys. (N. Y.) **23**, 335 (1963).

[4] F. Rohrlich, J. Math. Phys. **5**, 324 (1964).

is unique:

$$f = \tilde{f} + Df. \tag{10}$$

In a similar fashion a functional F_n can be split into two parts,

$$F_n = \tilde{F}_n + DF_n, \tag{11}$$

with $\tilde{F}_n = F_n[\tilde{f}_n]$ and $DF_n = F_n[Df_n]$. The functional derivatives of $\tilde{F}_n$ and DF_n are defined differently: We define

$$\frac{\delta\tilde{F}_n}{\delta x} = [(n-1)!]^{-1} \int d^4\xi_1 \cdots d^4\xi_{n-1} \tilde{f}_n(\xi_1 \cdots \xi_{n-1}x) \times :\varphi_{\text{in}}(\xi_1) \cdots \varphi_{\text{in}}(\xi_{n-1}):, \tag{12}$$

but DF_n is defined to be a null functional and all its functional derivatives are defined to be zero. It follows then that functional differentiation does not commute with space–time integration except when the functional is of the $\tilde{F}$ form.

As a trivial, but important, example consider the functional

$$F = K_x\varphi_{\text{in}}(x) = \int d^4y[K_y\delta(x-y)]\varphi_{\text{in}}(y).$$

By the above definition this is a null functional so its functional derivatives are zero. If, on the other hand, one were to apply (12) to calculate the functional derivative of $K_x\varphi_{\text{in}}(x)$ one would get

$$\delta F/\delta z = K_z\delta(x-z)$$

rather than zero.

As far as possible we shall deal only with functions f for which $Df = 0$, and so we shall not need to make explicit separations like (10) in the following.

It should be pointed out that there are many other functions df_n that vanish on the mass shell besides those of the form (9). In general these other functions cannot be separated uniquely from a given function f. Hence two functionals $F_n[f_n]$ and $F_n[g_n]$ may be equal even though $f_n \neq g_n$. Since

$$f_n(x_1 \cdots x_n) = \langle 0|\ \delta^n F[f]/\delta x_1 \cdots \delta x_n\ |0\rangle, \tag{13}$$

it follows that $F[f] = F[g]$ does *not* imply that the functional derivatives of $F[f]$ are equal to the functional derivatives of $F[g]$. In the case that all the functional derivatives of $F[f]$ are equal to the respective functional derivatives of $F[g]$ we say that $F[f]$ and $F[g]$ are *identical* and write $F[f] \equiv F[g]$. Hence we have

$$F[f] \equiv F[g] \Leftrightarrow \frac{\delta^n F[f]}{\delta x_1 \cdots \delta x_n} \equiv \frac{\delta^n F[g]}{\delta x_1 \cdots \delta x_n} \Leftrightarrow f_n = g_n \quad \text{for all} \quad n. \tag{14}$$

One must be careful to distinguish equality from identity in dealing with functionals of free fields.

To conclude this section we investigate the functional differentiation of time-ordered products that arise in quantum field theory. The functional differentiation of an ordinary time-ordered product is straightforward:

$$(\delta/\delta x)(\varphi(x_1) \cdots \varphi(x_n))_+ \equiv \sum_{i=1}^{n} \left(\varphi(x_1) \cdots \frac{\delta\varphi(x_i)}{\delta x} \cdots \varphi(x_n)\right)_+, \tag{15}$$

but if Klein–Gordon operators are present, some care must be taken since functional differentiation and space–time differentiation do not, in general, commute. Consider the following:

$$N = (\delta/\delta x)K_1 \cdots K_n(\varphi(x_1) \cdots \varphi(x_n))_+.$$

Let us assume that $\varphi(y)$ is an interpolating quantum field with the properties

$$\varphi(y) \equiv \sum_{k=1}^{\infty} \varphi_k(y),$$

$$\varphi_k(y) \equiv \int d^4\eta_1 \cdots d^4\eta_k\ \Phi(y;\eta_1 \cdots \eta_k) \times :\varphi_{\text{in}}(\eta_1) \cdots \varphi_{\text{in}}(\eta_k):,$$

$$K_y\varphi_1(y) \equiv 0,$$

and let

$$\chi(y) \equiv \varphi(y) - \varphi_1(y). \tag{16}$$

$\varphi_1(y)$ is a free field. Now examine the contribution to N of $\delta\varphi_1(x_i)/\delta x$. Imagine the time-ordered product in N written out as a sum of integrals of normal products of free fields. In general, any free field appearing in $\varphi(x_i)$ will either appear in one of the normal products or it will be contracted into some other free field; in particular, this is true for $\varphi_1(x_i)$. Only if $\varphi_1(x_i)$ is contracted does the time ordering introduce a θ function depending on x_i^0: those terms in which $\varphi_1(x_i)$ appears in a normal product, depend on x_i^0 only through $\varphi_1(x_i)$. When K_i is applied to those terms we get zero. Thus we conclude that $\varphi_1(x_i)$ contributes to $K_1 \cdots K_n(\varphi(x_1) \cdots \varphi(x_n))_+$ only through its contractions with other free field operators. Hence $\delta\varphi_1(x_i)/\delta x$ does not appear in N. Therefore we have the result

$$(\delta/\delta x)K_1 \cdots K_n(\varphi(x_1) \cdots \varphi(x_n))_+ = K_1 \cdots K_n \sum_{i=1}^{n} \left(\varphi(x_1) \cdots \frac{\delta\chi(x_i)}{\delta x} \cdots \varphi(x_n)\right)_+. \tag{7}$$

This result tells us how to commute functional derivatives and Klein–Gordon operators in this common field-theoretic expression.

C. FORMULATION

In this section we derive the equation for the S operator. The fundamental dynamical axiom [Eq. (31) or (32) below] is stated in mathematical language and no attempt is made to interpret this axiom physically.

The free particle state vectors are obtained by repeatedly applying the creation operators

$$\varphi^{\alpha}_{\text{in}} \equiv -i \int d^3x f_\alpha(x) \overleftrightarrow{\partial}_0 \varphi_{\text{in}}(x), \quad \alpha = 1, 2, \cdots, \tag{18}$$

to the vacuum state and normalizing the result. Let $|\alpha)$ be the n-particle state $|\alpha_1 \cdots \alpha_n)$ and let $|\alpha/\alpha_i)$ denote the $(n-1)$-particle state $|\alpha_1 \cdots \Lambda_i \cdots \alpha_n)$. With all the α's different we have

$$|\alpha) = \varphi^{\alpha_i}_{\text{in}} |\alpha/\alpha_i), \tag{19}$$

while if some of the α's are equal, there will be a normalization factor in (19). This normalization factor will be ignored since it does not affect our final result: we therefore assume that all α's are different. Under this assumption one has

$$\varphi^{\gamma\dagger}_{\text{in}} |\alpha) = |\alpha/\gamma). \tag{20}$$

We define $|\alpha/\gamma) = 0$ in the case that no α_i is equal to γ.

Next observe that if γ belongs to the set β,

$$\begin{aligned}(\alpha| S |\beta) &= (\alpha| S\varphi^{\gamma}_{\text{in}} |\beta/\gamma) \\ &= (\alpha| [S, \varphi^{\gamma}_{\text{in}}] |\beta/\gamma) + (\alpha| \varphi^{\gamma}_{\text{in}} S |\beta/\gamma) \\ &= \int d^4x f_\gamma(x)(\alpha| \frac{\delta S}{\delta x} |\beta/\gamma) + (\alpha/\gamma| S |\beta/\gamma). \end{aligned} \tag{21}$$

In the last step we used

$$[\varphi_{\text{in}}(x), F] = -i \int d^4\xi \, \Delta(x-\xi) \frac{\delta F}{\delta \xi}. \tag{22}$$

On the other hand, if there exists an operator $\varphi(x)$ that satisfies the LSZ asymptotic conditions, one gets the well-known result[1]

$$(\alpha| S |\beta) = -i \int d^4x f_\gamma(x)(\alpha| SK_x\varphi(x) |\beta/\gamma) + (\alpha/\gamma| S |\beta/\gamma). \tag{23}$$

[$\varphi(x)$ will be defined below, so we are not making its existence or its asymptotic behavior axioms of the formulation.] If we compare (21) and (23) and if we make use of the fact that the states $(\alpha|$ and $|\beta/\gamma)$ are competely arbitrary, we get

$$\int d^4x f_\gamma(x) \left\{ \frac{\delta S}{\delta x} + iSK_x\varphi(x) \right\} = 0. \tag{24}$$

One may continue the above reduction as was done in Ref. 1 to get

$$\int d^4x_1 \cdots d^4x_n f_{\gamma_1}(x_1) \cdots f_{\gamma_n}(x_n) \left\{ \frac{\delta^n S}{\delta x_1 \cdots \delta x_n} - (-i)^n K_1 \cdots K_n S\varphi(\varphi(x_1) \cdots \varphi(x_n)) \right\} = 0. \tag{25}$$

The φ product is defined in Eq. (38) below. Here the f's can be either positive or negative energy wavefunctions: positive if $|\beta)$ was reduced and negative if $(\alpha|$ was reduced.

Equation (25) is trivial for $n = 0$; for $n = 1$ it can be satisfied without making use of any of the properties of the wavefunctions by *defining* $\varphi(x)$ so that

$$\delta S/\delta x \equiv -iSK_x\varphi(x). \tag{26}$$

(Since this is an equation of definition, we are at liberty to say that the two sides are identical in the sense of Sec. B.) We may integrate (26) in such a way that the asymptotic conditions are satisfied.[5] The result is

$$\varphi(x) \equiv \varphi_{\text{in}}(x) - iS^\dagger \int d^4\xi \, \Delta_R(x-\xi) \frac{\delta S}{\delta \xi}. \tag{27}$$

If we demand that $\varphi(x)$ be Hermitian, then S is restricted by

$$S^\dagger \, \delta S/\delta\xi \equiv -(\delta S^\dagger/\delta\xi)S. \tag{28}$$

Taken together with ordinary unitarity

$$S^\dagger S = 1, \tag{29}$$

this implies

$$S^\dagger S \equiv 1, \tag{30}$$

which is *generalized* unitarity.[2]

Just as we defined $\varphi(x)$ so that the $n = 1$ part of (25) was satisfied identically, we may postulate that

$$\delta^2 S/\delta x_1 \, \delta x_2 \equiv -SK_1K_2\varphi(\varphi(x_1)\varphi(x_2)), \tag{31}$$

so that the $n = 2$ part of (25) is satisfied identically. Both (26) and (31) could be derived from the postulate that (25) is true with arbitrary functions $f_{\gamma_i}(x_i)$ (functions that are not necessarily solutions of the Klein–Gordon equation). It will be shown below that (26) and (31) imply that (25) is satisfied automatically for all n. First we note that

$$\begin{aligned}\delta\varphi(x)/\delta y &\equiv -iK_y[\varphi(x), \varphi(y)]_R \\ &\equiv -iK_y\{\theta(x-y)[\varphi(x), \varphi(y)]\} \end{aligned} \tag{32}$$

[5] G. Källén, in *Handbuch der Physik*, edited by S. Flügge (Springer-Verlag, Berlin, 1958), Vol. V/1, p. 232; W. Zimmermann, Nuovo Cimento **10**, 597 (1958); H. Rollnik, B. Stech, and E. Nunnemann, Z. Physik **159**, 482 (1960).

is an immediate consequence of (27) and (31). Either (31) or (32) may be looked upon as the fundamental *dynamical axiom* of this formulation of quantum field theory.

It will now be shown that (25) is satisfied for all n by Eqs. (27), (30), and (31). We prove this by induction. Assume

$$\delta^n S/\delta x_1 \cdots \delta x_n \equiv (-i)^n S K_1 \cdots K_n \varphi(\varphi_1 \cdots \varphi_n) \quad (33)$$

for some n. Here we have written φ_i for $\varphi(x_i)$. We differentiate (33) to get a similar equation for the $(n+1)$th derivative of S. First examine the functional derivative of $K_1 \cdots K_n (\varphi_1 \cdots \varphi_n)_+$. We use Eq. (17) and

$$\begin{aligned}\delta\chi_i/\delta x_{n+1} &\equiv -iK_{n+1}[\varphi_i, \varphi_{n+1}]_R - \delta(x_i - x_{n+1}) \\ &\equiv -iK_{n+1}\{[\varphi_i, \varphi_{n+1}]_R + i\Delta_c(x_i - x_{n+1})\} \end{aligned} \quad (34)$$

to get

$$\begin{aligned}&\delta/\delta x_{n+1} K_1 \cdots K_n(\varphi_1 \cdots \varphi_n)_+ \\ &\equiv -iK_1 \cdots K_{n+1} \sum_{i=1}^{n} \{ \sum_{(1\cdots n)} \theta_{12} \cdots \theta_{n-1,n} \\ &\times \theta_{i,n+1}\varphi_1 \cdots [\varphi_i, \varphi_{n+1}] \cdots \varphi_n \\ &+ i\Delta_c(x_i - x_{n+1})(\varphi_1 \cdots \Lambda_i \cdots \varphi_n)_+\}, \end{aligned} \quad (35)$$

with $\theta_{ij} \equiv \theta(x_i - x_j)$, and $\sum_{(1\cdots n)}$ indicating a sum over all permutations of the variables $(x_1 \cdots x_n)$. Collecting coefficients in (35) one gets

$$\begin{aligned}&\frac{\delta}{\delta x_{n+1}} K_1 \cdots K_n(\varphi_1 \cdots \varphi_n)_+ \\ &\equiv -iK_1 \cdots K_{n+1}\Big\{ \sum_{(1\cdots n)} \sum_{i=1}^{n-1} \theta_{12} \cdots \theta_{n-1,n} \\ &\times (\theta_{i,n+1} - \theta_{i+1,n+1})\varphi_1 \cdots \varphi_i\varphi_{n+1}\varphi_{i+1} \cdots \varphi_n \\ &+ \sum_{(1\cdots n)} [\theta_{12} \cdots \theta_{n,n+1}\varphi_1 \cdots \varphi_{n+1} \\ &- \theta_{1,n+1}\theta_{12} \cdots \theta_{n-1,n}\varphi_{n+1}\varphi_1 \cdots \varphi_n] \\ &+ \sum_{i=1}^{n} i\Delta_c(x_i - x_{n+1})(\varphi_1 \cdots \Lambda_i \cdots \varphi_n)_+\Big\} \\ &\equiv -iK_1 \cdots K_{n+1}\{(\varphi_1 \cdots \varphi_{n+1})_+ - \varphi_{n+1}(\varphi_1 \cdots \varphi_n)_+ \\ &+ \sum_i i\Delta_c(x_i - x_{n+1})(\varphi_1 \cdots \Lambda_i \cdots \varphi_n)_+\}. \end{aligned} \quad (36)$$

When $SK_1 \cdots K_n(\varphi_1 \cdots \varphi_n)_+$ is differentiated, the contribution from $\delta S/\delta x_{n+1}$ just cancels the second term in (36):

$$\begin{aligned}&(\delta/\delta x_{n+1})[SK_1 \cdots K_n(\varphi_1 \cdots \varphi_n)_+] \\ &\equiv -iSK_1 \cdots K_{n+1}\{(\varphi_1 \cdots \varphi_{n+1})_+ \\ &+ \sum_i i\Delta_c(x_i - x_{n+1})(\varphi_1 \cdots \Lambda_i \cdots \varphi_n)_+\}. \end{aligned} \quad (37)$$

Finally when we compare the φ product with the time-ordered product,

$$\begin{aligned}\varphi(\varphi_1 \cdots \varphi_n) &\equiv (\varphi_1 \cdots \varphi_n)_+ \\ &+ \sum_{\text{pairs}} [i\Delta_c(x_i - x_j)](\varphi_1 \cdots \Lambda_{ij} \cdots \varphi_n)_+ \\ &+ \sum_{\text{two pairs}} [i\Delta_c(x_i - x_j)][i\Delta_c(x_k - x_l)] \\ &\times (\varphi_1 \cdots \Lambda_{ijkl} \cdots \varphi_n)_+ + \text{etc.}, \end{aligned} \quad (38)$$

we see that the last term in (37) just supplies the extra contractions that appear in $\varphi(\varphi_1 \cdots \varphi_{n+1})$ but do not appear in $\varphi(\varphi_1 \cdots \varphi_n)$. Hence we have the result

$$\begin{aligned}&(\delta/\delta x_{n+1})[SK_1 \cdots K_n\varphi(\varphi_1 \cdots \varphi_n)] \\ &\equiv -iSK_1 \cdots K_{n+1}\varphi(\varphi_1 \cdots \varphi_{n+1}) \end{aligned} \quad (39)$$

so that

$$\begin{aligned}&\delta^{n+1}S/\delta x_1 \cdots \delta x_{n+1} \\ &\equiv (-i)^{n+1} SK_1 \cdots K_{n+1}\varphi(\varphi_1 \cdots \varphi_{n+1}). \end{aligned} \quad (40)$$

This completes the proof that (25) is satisfied for all values of n by Eqs. (27), (30) and (31).

We conclude this section by eliminating the interpolating field from Eq. (31). Writing $\varphi \equiv \varphi_{\text{in}} + \chi$, we have

$$\varphi(\varphi_1\varphi_2) \equiv :\varphi_1^{\text{in}}\varphi_2^{\text{in}}: + (\varphi_1^{\text{in}}\chi_2)_+ + (\chi_1\varphi_2^{\text{in}})_+ + (\chi_1\chi_2)_+.$$

When K_1K_2 is applied to this φ product, the first term gives zero, and the second and third terms contribute only when the φ_{in}'s contract on the χ's:

$$\begin{aligned}&K_1K_2(\varphi_1^{\text{in}}\chi_2)_+ \equiv -K_1K_2\{\theta_{21}[\varphi_1^{\text{in}}, \chi_2]\} \\ &\equiv iK_1K_2\Big\{\theta_{21}\int d^4\xi_1\, \Delta(x_1 - \xi_1)\frac{\delta\chi_2}{\delta\xi_1}\Big\}. \\ &\equiv K_1K_2\Big\{\theta_{21}\int d^4\xi_1\, d^4\xi_2\, \Delta(x_1 - \xi_1)\, \Delta_R(x_2 - \xi_2) \\ &\times \Big[S^\dagger \frac{\delta^2 S}{\delta\xi_1\, \delta\xi_2} - S^\dagger \frac{\delta S}{\delta\xi_1} S^\dagger \frac{\delta S}{\delta\xi_2}\Big]\Big\}. \end{aligned} \quad (41)$$

Using $\Delta_A(x) = \Delta_R(x) - \Delta(x)$, one finds

$$\begin{aligned}&K_1K_2\varphi(\varphi_1\varphi_2) \equiv K_1K_2\int d^4\xi_1\, d^4\xi_2\{\theta_{12}\Delta_R(x_1 - \xi_1)\Delta(x_2 - \xi_2) + \theta_{21}\Delta(x_1 - \xi_1)\Delta_R(x_2 - \xi_2)\}S^\dagger\, \delta^2 S/\delta\xi_1\, \delta\xi_2 \\ &- K_1K_2\int d^4\xi_1\, d^4\xi_2\Big\{\theta_{12}\Delta_A(x_1 - \xi_1)\Delta_R(x_2 - \xi_2)S^\dagger\frac{\delta S}{\delta\xi_1}S^\dagger\frac{\delta S}{\delta\xi_2} + \theta_{21}\Delta_R(x_1 - \xi_1)\Delta_A(x_2 - \xi_2)S^\dagger\frac{\delta S}{\delta\xi_2}S^\dagger\frac{\delta S}{\delta\xi_1}\Big\}, \end{aligned} \quad (42)$$

so that

$$(1 - B)S^\dagger \delta^2 S/\delta x_1\, \delta x_2 \equiv K_1K_2 \int d^4\xi_1\, d^4\xi_2 \times \{\theta_{12}\Delta_A(x_1 - \xi_1)\Delta_R(x_2 - \xi_2)S^\dagger(\delta S/\delta\xi_1)S^\dagger \delta S/\delta\xi_2 + \theta_{21}\Delta_R(x_1 - \xi_1)\Delta_A(x_2 - \xi_2)S^\dagger(\delta S/\delta\xi_2)S^\dagger \delta S/\delta\xi_1\}. \quad (43)$$

Here we have put

$$B\, \delta^2 S/\delta x_1\, \delta x_2 \equiv -K_1K_2 \int d^4\xi_1\, d^4\xi_2 \times [\theta(x_1 - \xi_1)\theta(x_1 - x_2) + \theta(x_2 - \xi_2)\theta(x_2 - x_1)] \Delta(x_1 - \xi_1)\Delta(x_2 - \xi_2)\, \delta^2 S/\delta\xi_1\, \delta\xi_2. \quad (44)$$

The properties of the integral operator, B, are well known.[3] It is a projection operator

$$B^2 = B, \quad (45)$$

and its eigenfunctions of eigenvalue unity

$$Bb = b \quad (46)$$

are of the form

$$b(x_1, x_2) = (\partial^n/\partial x_1^{0n})(\partial^m/\partial x_2^{0m})[\delta(x_1^0 - x_2^0)g(x_1, x_2)], \quad (47)$$

with n and m restricted by

$$n + m < 4. \quad (48)$$

It can easily be shown[3] that the right side of (43) satisfies $B\lambda = 0$; therefore the integral equation, (43), can be solved. Its general solution is

$$S^\dagger\, \delta^2 S/\delta x_1\, \delta x_2 \equiv b'(x_1, x_2) + K_1K_2 \int d^4\xi_1\, d^4\xi_2 \times \{\theta_{12}\Delta_A^{(1)}\Delta_R^{(2)}S^\dagger(\delta S/\delta\xi_1)S^\dagger \delta S/\delta\xi_2 + \theta_{21}\Delta_R^{(1)}\Delta_A^{(2)}S^\dagger(\delta S/\delta\xi_2)S^\dagger \delta S/\delta\xi_1\} \quad (49)$$

or

$$\delta^2 S/\delta x_1\, \delta x_2 \equiv b(x_1, x_2) + K_1K_2 \int d^4\xi_1\, d^4\xi_2 \times \{\theta_{12}\Delta_A^{(1)}\Delta_R^{(2)}(\delta S/\delta\xi_1)S^\dagger \delta S/\delta\xi_2 + \theta_{21}\Delta_R^{(1)}\Delta_A^{(2)}(\delta S/\delta\xi_2)S^\dagger \delta S/\delta\xi_1\}. \quad (50)$$

Here we have put $\Delta_j^{(i)} = \Delta_j(x_i - \xi_i)$, for $j = R, A$ and $i = 1, 2$ and $b = Sb'$.

In Eq. (50) the left side is linear in S while the right side is nonlinear. It is therefore a convenient equation for a perturbation treatment of quantum field theory. In Sec. F, we shall solve (50) for S in powers of a coupling constant.

D. EQUIVALENCE WITH S-MATRIX THEORY

A previous formulation of quantum field theory[3] was based on (27), (30) and a set of equations for S-matrix elements which can now be written in the form

$$\langle 0|\, \delta^n S/\delta x_1 \cdots \delta x_n\, |0\rangle = (-i)^n K_1 \cdots K_n \langle 0|\, \varphi(\varphi_1 \cdots \varphi_n)\, |0\rangle, \quad n = 2, 3, \cdots. \quad (51)$$

Using the results of Secs. B and C, one can easily see that (51) is equivalent to the present formulation. Since $S\,|0) = |0)$, Eq. (51) may be written as

$$\langle 0|\, \frac{\delta^n}{\delta y_1 \cdots \delta y_n} \frac{\delta^2 S}{\delta x_1\, \delta x_2}\, |0\rangle = \langle 0|\, \frac{\delta^n}{\delta y_1 \cdots \delta y_n} (-i)^2 K_1K_2 S\varphi(\varphi_1\varphi_2)\, |0\rangle, \quad \text{for all } n. \quad (52)$$

According to (14), this implies that

$$\delta^2 S/\delta x_1\, \delta x_2 \equiv (-i)^2 K_1K_2 S\varphi(\varphi_1\varphi_2), \quad (53)$$

which is the basic equation of this present formulation.

E. THE FEYNMAN-DYSON S OPERATOR

Equation (50) may be written in a still simpler form if one is willing to risk introducing possibly divergent integrals into the theory. First notice that if the K_1K_2 differentiation is carried out in (50), any term in which θ_{12} or θ_{21} is differentiated satisfies Eq. (46). Hence all these terms may be separated off and absorbed into the function $b(x_1, x_2)$. (It is precisely this separation that introduces divergences into the theory since these terms are, in general, divergent.) In effect, we can therefore commute the K's and the θ's:

$$S^\dagger \frac{\delta^2 S}{\delta x_1\, \delta x_2} \equiv b(x_1, x_2) + \left\{\left(S^\dagger \frac{\delta S}{\delta x_1}\right)\left(S^\dagger \frac{\delta S}{\delta x_2}\right)\right\}_+. \quad (54)$$

It will now be shown that the Feynman–Dyson S operator satisfies (54) and, therefore, (50) also. The Feynman–Dyson S operator may be written

$$S_0 \equiv (e^{-iH})_+, \quad (55)$$

with

$$H = g \int d^4x\, :\varphi_{\text{in}}(x)\varphi_{\text{in}}(x) \cdots \varphi_{\text{in}}(x):. \quad (56)$$

Since $\delta H/\delta x_1$ is a function only of $\varphi_{\text{in}}(x_1)$, we have

$$\frac{\delta S_0}{\delta x_1} \equiv -i\left(\frac{\delta H}{\delta x_1} e^{-iH}\right)_+ \quad (57)$$

and

$$\frac{\delta^2 S}{\delta x_1\,\delta x_2} \equiv -i\left(\frac{\delta^2 H}{\delta x_1\,\delta x_2}\,e^{-iH}\right)_+ + (-i)^2\left(\frac{\delta H}{\delta x_1}\frac{\delta H}{\delta x_2}\,e^{-iH}\right)_+. \tag{58}$$

Noting the identity

$$\left(\frac{\delta H}{\delta x_1}\frac{\delta H}{\delta x_2}\,e^{-iH}\right)_+ \equiv \theta_{12}\left(\frac{\delta H}{\delta x_1}\,e^{-iH}\right)_+(e^{iH})_-\left(\frac{\delta H}{\delta x_2}\,e^{-iH}\right)_+ + \theta_{21}\left(\frac{\delta H}{\delta x_2}\,e^{-iH}\right)_+(e^{iH})_-\left(\frac{\delta H}{\delta x_1}\,e^{-iH}\right)_+ \tag{59}$$

we have

$$S_0^\dagger\frac{\delta^2 S_0}{\delta x_1\,\delta x_2} \equiv -iS_0^\dagger\left(\frac{\delta^2 H}{\delta x_1\,\delta x_2}\,e^{-iH}\right)_+ + \left\{\left(S_0^\dagger\frac{\delta S_0}{\delta x_1}\right)\left(S_0^\dagger\frac{\delta S_0}{\delta x_2}\right)\right\}_+. \tag{60}$$

Finally, since

$$(1 - B)(\delta^2 H/\delta x_1\,\delta x_2\,e^{-iH})_+ = 0, \tag{61}$$

Eq. (60) shows that S_0 satisfies (54).

F. A SAMPLE CALCULATION

In this section we calculate the S operator to second order with a φ^3 interaction. Of course the results are well known; the point to be demonstrated is that the calculation is not appreciably more difficult than the corresponding calculation in the Lagrangian formulation of quantum field theory.

Imagine $b(x_1, x_2)$ in (50) expanded in a power series in a coupling constant, g:

$$b(x_1, x_2) = \sum_{n=1}^{\infty} g^n b^{(n)}(x_1, x_2). \tag{62}$$

For a φ^3 interaction we take

$$b^{(1)}(x_1, x_2) = i\,\delta(x_1, x_2)\varphi_{\text{in}}(x_2). \tag{63}$$

Since derivatives of S are of at least first order in g, $b^{(1)}$ is the only first-order term on the right of (50). Hence with

$$S = \sum_{n=0}^{\infty} g^n S^{(n)}, \qquad S^{(0)} = 1, \tag{64}$$

we have

$$\delta^2 S^{(1)}/\delta x_1\,\delta x_2 \equiv i\,\delta(x_1 - x_2)\varphi_{\text{in}}(x_2). \tag{65}$$

Integrating formally we have

$$S^{(1)} \equiv \frac{i}{6}\int d^4x\,:\varphi_{\text{in}}(x)\varphi_{\text{in}}(x)\varphi_{\text{in}}(x):\tag{66}$$

and

$$\delta S^{(1)}/\delta x \equiv \tfrac{1}{2}i\,:\varphi_{\text{in}}(x)\varphi_{\text{in}}(x):. \tag{67}$$

Hence the equation for $S^{(2)}$ is

$$\frac{\delta^2 S^{(2)}}{\delta x_1\,\delta x_2} \equiv b^{(2)}(x_1, x_2) - \tfrac{1}{4}K_1K_2\int d^4\xi_1\,d^4\xi_2 \times\{\theta_{12}\Delta_A^{(1)}\Delta_R^{(2)}\,:\varphi_{\text{in}}(\xi_1)\varphi_{\text{in}}(\xi_1):\,:\varphi_{\text{in}}(\xi_2)\varphi_{\text{in}}(\xi_2): + \theta_{21}\Delta_R^{(1)}\Delta_A^{(2)}\,:\varphi_{\text{in}}(\xi_2)\varphi_{\text{in}}(\xi_2)\,::\varphi_{\text{in}}(\xi_1)\varphi_{\text{in}}(\xi_1):\} \tag{68}$$

$$\equiv b^{(2)}(x_1, x_2) - \tfrac{1}{4}\,:\varphi_{\text{in}}(x_1)\varphi_{\text{in}}(x_1)\varphi_{\text{in}}(x_2)\varphi_{\text{in}}(x_2): + iK_1K_2\int d^4\xi_1\,d^4\xi_2\{\theta_{12}\Delta_A^{(1)}\Delta_R^{(2)}\Delta_+(\xi_1 - \xi_2) + \theta_{21}\Delta_R^{(1)}\Delta_A^{(2)}\Delta_+(\xi_2 - \xi_1)\}\,:\varphi_{\text{in}}(\xi_1)\varphi_{\text{in}}(\xi_2): + \tfrac{1}{2}K_1K_2\int d^4\xi_1\,d^4\xi_2\{\theta_{12}\Delta_A^{(1)}\Delta_R^{(2)}\Delta_+^2(\xi_1 - \xi_2) + \theta_{21}\Delta_R^{(1)}\Delta_A^{(2)}\Delta_+^2(\xi_2 - \xi_1)\}. \tag{69}$$

Integrating formally we get

$$S^{(2)} \equiv A - \frac{1}{72}\int d^4\xi_1\,d^4\xi_2 \times\,:\varphi_{\text{in}}(\xi_1)\varphi_{\text{in}}(\xi_1)\varphi_{\text{in}}(\xi_1)\varphi_{\text{in}}(\xi_2)\varphi_{\text{in}}(\xi_2)\varphi_{\text{in}}(\xi_2): + \frac{i}{8}\int d^4\xi_1\,d^4\xi_2\,:\varphi_{\text{in}}(\xi_1)\varphi_{\text{in}}(\xi_1)\Delta_c(\xi_1 - \xi_2)\varphi_{\text{in}}(\xi_2)\varphi_{\text{in}}(\xi_2): - \frac{1}{2}\int d^4\xi_1\,d^4\xi_2\,\omega^{(2)}(\xi_1, \xi_2)\,:\varphi_{\text{in}}(\xi_1)\varphi_{\text{in}}(\xi_2):. \tag{70}$$

In the second integral we commuted the K's with the θ's since the resulting expression is finite. In the last term we have put

$$\omega^{(2)}(\xi_1, \xi_2) = -\frac{1}{2}K_1K_2\int d^4\eta_1\,d^4\eta_2 \times\{\theta_{12}\Delta_A^{(1)}\Delta_R^{(2)}\Delta_+^2(\eta_1 - \eta_2) + \theta_{21}\Delta_R^{(1)}\Delta_A^{(2)}\Delta_+^2(\eta_2 - \eta_1)\}. \tag{71}$$

This is the "renormalized" propagator. A is any operator that satisfies

$$(1 - B)\delta^2 A/\delta x_1\,\delta x_2 = 0. \tag{72}$$

The operator A is determined by imposing the following boundary condition on S: the S operator can be written as a sum of terms $S = \sum_i X_i$ such that each X_i satisfies *one* of the following conditions: (a) $X_i = \delta X_i/\delta y = 0$. This ensures that such terms do not contribute to higher-order parts of the S operator. (b) X_i is the vertex part of the S operator. (c) The Fourier transform of $\delta^n X_i/\delta x_1 \cdots \delta x_n$ ($n = 1, 2, \cdots$) vanishes for large energies.

The first and third terms in (70) satisfy condition (a); the middle term satisfies condition (c). Hence $A = 0$ and we have

$$S^{(2)} = +\frac{i}{8}\int d^4\xi_1\, d^4\xi_2 \times :\varphi_{in}(\xi_1)\varphi_{in}(\xi_1)\Delta_c(\xi_1 - \xi_2)\varphi_{in}(\xi_2)\varphi_{in}(\xi_2): . \quad (73)$$

G. DISCUSSION

Let us review the main points of the above formalism: Assuming that the properties of the S operator and its functional derivatives are independent of the kinematics of the free-particle wavefunctions we arrived at the relationships

$$S^\dagger\, \delta^n S/\delta x_1 \cdots \delta x_n \equiv (-i)^n K_1 \cdots K_n \varphi(\varphi_1 \cdots \varphi_n), \qquad n = 0, 1, 2, \cdots .$$

For $n = 0$, this gives generalized unitarity. For $n = 1$ it defines the interpolating field and for $n = 2$, it specifies the dynamics of quantum field theory. The relationships with $n > 2$ are derivable from the $n = 0$, 1, and 2, cases and hence they say nothing new. We found that the solution of the equation in perturbation theory was not any more difficult than a similar solution using Feynman techniques; in fact the solution in the S-operator formalism is, perhaps, more palatable since all integrals are convergent.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 6, NUMBER 5 MAY 1965

Tensor Methods and a Unified Representation Theory of SU_3*

N. MUKUNDA† AND L. K. PANDIT‡
Department of Physics and Astronomy, University of Rochester, Rochester, New York
(Received 19 June 1964; final manuscript received 4 September 1964)

Starting with irreducible tensors, we develop an explicit construction of orthonormal basic states for an arbitrary unitary irreducible representation (λ, μ) of the group SU_3. A knowledge of the simple properties of the irreducible tensors can then be exploited to obtain a variety of results, which ordinarily require more abstract algebraic methods for their derivation. As illustrative applications, we (i) derive Biedenharn's expressions for the matrix elements of the generators of SU_3, (ii) compute the matrix elements of octet-type operators for the case $(\lambda, \mu) \to (\lambda, \mu)$, and (iii) develop an explicit unitary transformation connecting the isospin and the U-spin states in any arbitrary irreducible representation.

1. INTRODUCTION

THE group SU_3 has, after a rather long search, emerged as the group providing us with a powerful tool for bringing order into the ever increasing complexity of high-energy physics.[1] As a result, interest in the study of the mathematical properties of SU_3 and its unitary irreducible representations (UIR's) has grown immensely. Two fairly distinct methods exist in dealing with this group; namely, the tensor method, and the canonical method of Weyl, Cartan, and Racah.[2] The former has been exploited chiefly by Okubo[3] in many elegant applications demonstrating the power of this approach. On the other hand, for deriving many detailed properties of SU_3 and its representations, it appears that the second method must be used. As examples, we may mention the derivation by Biedenharn[4] of the matrix elements of the generators of SU_3 in an arbitrary UIR, and the evaluation of certain Wigner coefficients of the group by Moshinsky.[5] The second

* Research supported in part by the U. S. Atomic Energy Commission.

† On leave of absence from Atomic Energy Establishment, Bombay, India.

‡ On leave of absence from Tata Institute of Fundamental Research, Bombay, India.

[1] M. Gell-Mann, Phys. Rev. **125**, 1067 (1962); Y. Ne'eman, Nucl. Phys. **26**, 222 (1961).

[2] See G. Racah, *Group Theory and Spectroscopy*, Institute for Advanced Study Lecture Notes, Princeton, New Jersey (1951). See also: L. C. Biedenharn, *Group Theoretical Approaches to Nuclear Spectroscopy, Lectures in Theoretical Physics*, (University of Colorado, Boulder, Colorado 1962), Vol. 5.

[3] S. Okubo, Progr. Theoret. Phys. (Kyoto) **27**, 949 (1962) and later papers. See also M. Ikeda, S. Ogawa, and Y. Ohnuki, Progr. Theoret. Phys. (Kyoto) **22**, 715 (1959); A. R. Edmonds, Proc. Roy. Soc. **A268**, 567 (1962).

[4] L. C. Biedenharn, Phys. Letters **3**, 69 (1962); **3**, 254 (1963). See also M. Harvey and J. P. Elliott, Proc. Roy. Soc. London **A272**, 557 (1963); K. T. Hecht, "SU_3 Reduction Coefficients, Fractional Parentage Coefficients, etc.," University of Michigan preprint (1963); D. L. Pursey, Proc. Roy. Soc. London **A275**, 284 (1963).

[5] M. Moshinsky, Rev. Mod. Phys. **34**, 813 (1962). See also M. A. Rashid, Nuovo Cimento **26**, 118 (1962); J. J. de Swart, Rev. Mod. Phys. **35**, 916 (1963); K. T. Hecht, Ref. 4.

The first and third terms in (70) satisfy condition (a); the middle term satisfies condition (c). Hence $A = 0$ and we have

$$S^{(2)} = +\frac{i}{8}\int d^4\xi_1\, d^4\xi_2 \times :\varphi_{in}(\xi_1)\varphi_{in}(\xi_1)\Delta_c(\xi_1 - \xi_2)\varphi_{in}(\xi_2)\varphi_{in}(\xi_2): . \quad (73)$$

G. DISCUSSION

Let us review the main points of the above formalism: Assuming that the properties of the S operator and its functional derivatives are independent of the kinematics of the free-particle wavefunctions we arrived at the relationships

$$S^\dagger\, \delta^n S/\delta x_1 \cdots \delta x_n \equiv (-i)^n K_1 \cdots K_n \varphi(\varphi_1 \cdots \varphi_n), \qquad n = 0, 1, 2, \cdots .$$

For $n = 0$, this gives generalized unitarity. For $n = 1$ it defines the interpolating field and for $n = 2$, it specifies the dynamics of quantum field theory. The relationships with $n > 2$ are derivable from the $n = 0$, 1, and 2, cases and hence they say nothing new. We found that the solution of the equation in perturbation theory was not any more difficult than a similar solution using Feynman techniques; in fact the solution in the S-operator formalism is, perhaps, more palatable since all integrals are convergent.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 6, NUMBER 5 MAY 1965

Tensor Methods and a Unified Representation Theory of SU_3*

N. MUKUNDA† AND L. K. PANDIT‡
Department of Physics and Astronomy, University of Rochester, Rochester, New York
(Received 19 June 1964; final manuscript received 4 September 1964)

Starting with irreducible tensors, we develop an explicit construction of orthonormal basic states for an arbitrary unitary irreducible representation (λ, μ) of the group SU_3. A knowledge of the simple properties of the irreducible tensors can then be exploited to obtain a variety of results, which ordinarily require more abstract algebraic methods for their derivation. As illustrative applications, we (i) derive Biedenharn's expressions for the matrix elements of the generators of SU_3, (ii) compute the matrix elements of octet-type operators for the case $(\lambda, \mu) \to (\lambda, \mu)$, and (iii) develop an explicit unitary transformation connecting the isospin and the U-spin states in any arbitrary irreducible representation.

1. INTRODUCTION

THE group SU_3 has, after a rather long search, emerged as the group providing us with a powerful tool for bringing order into the ever increasing complexity of high-energy physics.[1] As a result, interest in the study of the mathematical properties of SU_3 and its unitary irreducible representations (UIR's) has grown immensely. Two fairly distinct methods exist in dealing with this group; namely, the tensor method, and the canonical method of Weyl, Cartan, and Racah.[2] The former has been exploited chiefly by Okubo[3] in many elegant applications demonstrating the power of this approach. On the other hand, for deriving many detailed properties of SU_3 and its representations, it appears that the second method must be used. As examples, we may mention the derivation by Biedenharn[4] of the matrix elements of the generators of SU_3 in an arbitrary UIR, and the evaluation of certain Wigner coefficients of the group by Moshinsky.[5] The second

* Research supported in part by the U. S. Atomic Energy Commission.

† On leave of absence from Atomic Energy Establishment, Bombay, India.

‡ On leave of absence from Tata Institute of Fundamental Research, Bombay, India.

[1] M. Gell-Mann, Phys. Rev. **125**, 1067 (1962); Y. Ne'eman, Nucl. Phys. **26**, 222 (1961).

[2] See G. Racah, *Group Theory and Spectroscopy*, Institute for Advanced Study Lecture Notes, Princeton, New Jersey (1951). See also: L. C. Biedenharn, *Group Theoretical Approaches to Nuclear Spectroscopy, Lectures in Theoretical Physics*, (University of Colorado, Boulder, Colorado 1962), Vol. 5.

[3] S. Okubo, Progr. Theoret. Phys. (Kyoto) **27**, 949 (1962) and later papers. See also M. Ikeda, S. Ogawa, and Y. Ohnuki, Progr. Theoret. Phys. (Kyoto) **22**, 715 (1959); A. R. Edmonds, Proc. Roy. Soc. **A268**, 567 (1962).

[4] L. C. Biedenharn, Phys. Letters **3**, 69 (1962); **3**, 254 (1963). See also M. Harvey and J. P. Elliott, Proc. Roy. Soc. London **A272**, 557 (1963); K. T. Hecht, "SU_3 Reduction Coefficients, Fractional Parentage Coefficients, etc.," University of Michigan preprint (1963); D. L. Pursey, Proc. Roy. Soc. London **A275**, 284 (1963).

[5] M. Moshinsky, Rev. Mod. Phys. **34**, 813 (1962). See also M. A. Rashid, Nuovo Cimento **26**, 118 (1962); J. J. de Swart, Rev. Mod. Phys. **35**, 916 (1963); K. T. Hecht, Ref. 4.

method is, however, often felt to be more demanding in requiring a knowledge of advanced group theory.

In the course of a study of some mathematical properties of the group G_2, one of us (NM) has developed methods,[6] which turn out to be of even greater use in dealing with SU_3 on account of the greater inherent simplicity of this group. The technique of this approach establishes a direct link between the two methods referred to above, and thereby profits from the advantages of both. It is our purpose in the present paper to give an exposition of this method along with some applications. Our work is, in a sense, a "spinor calculus" for SU_3, in analogy to that developed by Van der Waerden[7] for SU_2.

In the following we assume that the reader is familiar with the parts of the excellent article by Behrends *et al.*,[8] dealing with SU_3. For the rest we use the well-known ideas drawn from the theory of the group SU_2, or the algebra of angular momentum.[9,10]

The material of this paper is arranged as follows. In Sec. I, we review very briefly the definition of SU_3 and its two fundamental UIR's, and also identify its important subgroups. Section II contains a definition of a general UIR, and a statement of the solution of the state-labeling problem. Section III is the heart of the paper. Here the explicit connection between irreducible tensors and orthonormal basic functions is established. In Sec. IV we give the first simple application of the results of Sec. III. The results of Biedenharn[11] for the matrix elements of the generators of SU_3 in an arbitrary UIR are derived here. The next application is given in Sec. V devoted to the evaluation of a certain class of matrix elements of the "octet" ("regular") operators of the group. Lastly, in Sec. VI, we are concerned with the U-spin formalism of Levinson *et al.*[12] We develop here the explicit unitary transformation connecting the isospin (I-spin) states to the U-spin states. Appendix A contains an important algebraic computation required in Sec. III. In Appendix B we give a compilation of the most useful results of this work.

We give all our final results in such a way as to be consistent with the phase conventions set up by Biedenharn.[11]

I. COMMUTATION RULES AND FUNDAMENTAL REPRESENTATIONS

The group SU_3 is the group of all unitary unimodular matrices in three dimensions. It possesses eight generators usually denoted as below:

$$H_1, H_2; E_{\pm\alpha} \quad (\alpha = 1, 2, 3); \qquad H_i^\dagger = H_i, \quad E_\alpha^\dagger = E_{-\alpha}. \tag{I.1}$$

The commutation rules (CR's) of these generators with each other can be found in the paper by Behrends *et al.*,[13] and are not repeated here. The generators I_i of isospin rotations and of hypercharge gauge transformations are identified as follows:

$$I_z = 3^{\frac{1}{2}}H_1, \qquad Y = 2H_2, \qquad I_\pm = 6^{\frac{1}{2}}E_{\pm 1}. \tag{I.2}$$

Following Lurié and Macfarlane,[14] we can arrange the rest of the generators into tensor operators with respect to the isospin subgroup of SU_3. Thus we have that $F_+ = 6^{\frac{1}{2}}E_2$ and $F_- = 6^{\frac{1}{2}}E_3$ form the $I_z = +\frac{1}{2}$ and $I_z = -\frac{1}{2}$ components, respectively, of a spherical tensor operator of rank $\frac{1}{2}$; similarly for $G_+ = -6^{\frac{1}{2}}E_{-3}$ and $G_- = 6^{\frac{1}{2}}E_{-2}$.

The two *basic* or *fundamental representations* of SU_3 are both three dimensional and are denoted by (1, 0) and (0, 1). The former is the defining representation and the latter its complex conjugate. We give, in Fig. 1, the states that appear in these unitary irreducible representations (UIR) in terms of the eigenvalues of the simultaneously diagonal operators I_z and Y. The explicit forms of the generators E_α in these UIR's are

for (1, 0):

$$E_1 = 6^{-\frac{1}{2}}\,|1\rangle\langle 2|, \quad E_2 = 6^{-\frac{1}{2}}\,|1\rangle\langle 3|, \quad E_3 = 6^{-\frac{1}{2}}\,|2\rangle\langle 3|; \tag{I.3}$$

for (0, 1):

$$E_1 = 6^{-\frac{1}{2}}\,|2\rangle\langle 1|, \; E_2 = 6^{-\frac{1}{2}}\,|3\rangle\langle 1|, \; E_3 = -6^{-\frac{1}{2}}\,|3\rangle\langle 2|. \tag{I.4}$$

Each of the UIR's (1, 0) and (0, 1) contains one isodoublet and one isosinglet. There is some freedom

[6] N. Mukunda, Ph.D. thesis, University of Rochester, Rochester, New York, (1964).
[7] B. L. van der Waerden, *Die Gruppentheoretische Methode in der Quantenmechanik* (Springer-Verlag, Berlin, 1932).
[8] R. E. Behrends, J. Dreitlein, C. Fronsdal, and B. W. Lee, Rev. Mod. Phys. **34**, 1 (1962). See also D. R. Speiser and J. Tarski, J. Math. Phys. **4**, 588 (1963).
[9] M. Hamermesh, *Group Theory* (Addison-Wesley Publishing Company, Inc., Reading, Massachusetts, 1961).
[10] See, for example, M. E. Rose, *Elementary Theory of Angular Momentum* (John Wiley & Sons, Inc., New York, 1957).
[11] L. C. Biedenharn, Ref. 4.
[12] C. A. Levinson, H. J. Lipkin, and S. Meshkov, Phys. Letters **1**, 44, 125, 307 (1962); Nuovo Cimento **23**, 236 (1962); Phys. Rev. Letters **10**, 361 (1962). See also A. J. Macfarlane, E. C. G. Sudarshan, and C. Dullemond, Nuovo Cimento **30**, 845 (1963).
[13] R. E. Behrends *et al.*, Ref. 8.
[14] D. Lurié and A. J. Macfarlane, J. Math. Phys. **5**, 565 (1964).

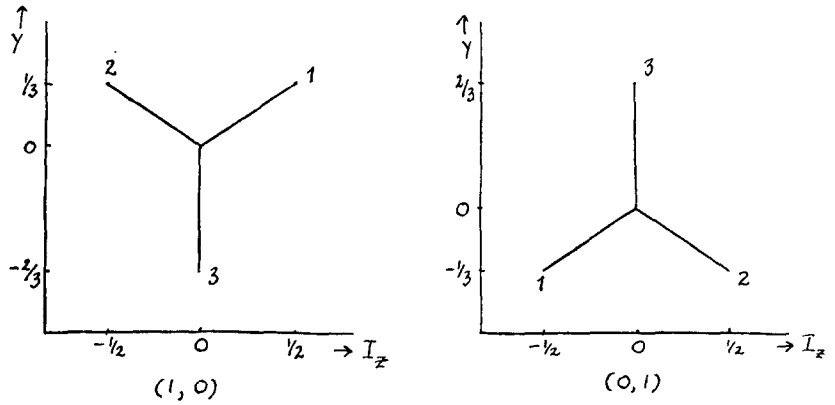


FIG. 1. States of the fundamental representations.

in the choice of signs in (I.3) and (I.4); we have taken care only to obey the CR's and to obey the Condon–Shortley phase convention as far as isospin is concerned.

We look briefly at the actual matrices that occur in the UIR's (1, 0) and (0, 1), for finite group elements of SU_3. If **a** is some element of SU_3, it induces the following transformation on a set of three quantities x^m (complex numbers):

$$\mathbf{a}: \quad x^m \to x'^m = \sum_{m'=1}^{3} A^m_{m'} x^{m'}; \quad m = 1, 2, 3. \tag{I.5}$$

Here $||A^m_{m'}||$ is a unitary unimodular matrix; and as **a** ranges over the whole of SU_3, we get *all* unitary unimodular matrices $||A||$. The correspondence $\mathbf{a} \to ||A||$ constitutes the UIR (1, 0). The UIR (0, 1) is obtained essentially (that is, apart from trivial changes due to phase conventions) by complex conjugation from (I.5):

$$\mathbf{a}: \quad y_n \to y'_n = \sum_{n'} A^{n*}_{n'} y_{n'}. \tag{I.6}$$

We call x^m a quantity of type (1, 0), and y_n of type (0, 1). We also use upper indices m and lower indices n as in (I.5) and (I.6). From x^m and y_n we can form one invariant bilinear quantity:

$$\sum_{m,n} g^n_m x^m y_n \equiv x^1 y_1 - x^2 y_2 - x^3 y_3. \tag{I.7}$$

The form (I.7) is fixed by (I.3) and (I.4); and from the structure of E_1 in (I.3) and (I.4), we can recognize the first two terms in the form (I.7) as consisting of two spin-$\frac{1}{2}$ quantities coupled to spin zero.

It is useful to record here the relationship between the generators $I_\pm$, I_z, Y, $F_\pm$, $G_\pm$ obeying the CR's given by Behrends *et al.*,[13] and the same generators looked upon as forming a tensor with respect to SU_3. The CR's given by Okubo[15] are in the latter form. Since the manner in which we define our invariant is given by (I.7) in terms of a metric tensor g^n_m which is *not the same* as a Kronecker symbol δ^n_m we must use the former, and not the latter, as an invariant tensor. Following Okubo,[15] the generators of SU_3 can be written as B^i_j with

$$\sum_{i,j} g^i_j B^i_j \equiv B^1_1 - B^2_2 - B^3_3 = 0; \qquad (B^i_j)^\dagger = \sum_{i',j'} g^j_{j'} g^{i'}_i B^{j'}_{i'}; \tag{I.8}$$

and

$$[B^i_j, B^k_l] = g^k_j B^i_l - g^i_l B^k_j. \tag{I.9}$$

These rules are then consistent with the CR's of Behrends *et al.* if we make the identifications

$$I_+ = -B^1_2, \quad I_- = B^2_1, \quad I_z = \tfrac{1}{2}(B^1_1 + B^2_2),$$
$$Y = B^3_3, \quad F_+ = -B^1_3, \quad F_- = -B^2_3, \tag{I.10}$$
$$G_+ = B^3_2, \quad G_- = B^3_1.$$

It must always be borne in mind that the indices i, j in B^i_j belong to the UIR's (1, 0), (0, 1) respectively, transforming accordingly, and carrying the isospin and hypercharge values given by Fig. 1.

II. GENERAL UNITARY REPRESENTATION AND STATE LABELING

An arbitrary UIR of SU_3 is denoted by two integers, each zero or positive: (λ, μ). The UIR (λ, μ) is obtained by taking the following direct product of the fundamental representations:

$$\underbrace{(1,0) \otimes (1,0) \otimes \cdots \otimes (1,0)}_{\lambda} \otimes \underbrace{(0,1) \otimes (0,1) \otimes \cdots \otimes (0,1)}_{\mu}, \tag{II.1}$$

[15] S. Okubo, Ref. 3. It should be noted that our upper index corresponds to Okubo's lower index.

reducing this product into irreducible constituents, and taking the "largest piece" [equivalently, the irreducible part that contains the highest I_z that can occur in the product (II.1)].

We define (λ, μ) in terms of tensors. A tensor T of type (λ, μ) is a set of quantities $T^{m_1 m_2 \cdots m_\lambda}_{n_1 n_2 \cdots n_\mu}$ that transform as follows, when x^m and y_n of the last section undergo the transformations (I.5) and (I.6):

$$\text{a:}\quad T^{m_1 \cdots m_\lambda}_{n_1 \cdots n_\mu} \to T'^{m_1 \cdots m_\lambda}_{n_1 \cdots n_\mu}$$
$$= \sum_{\{m'\}} \sum_{\{n'\}} A^{m_1}_{m_1'} \cdots A^{m_\lambda}_{m_\lambda'} A^{n_1}_{n_1'}{}^{*} \cdots A^{n_\mu}_{n_\mu'}{}^{*} T^{m_1' \cdots m_\lambda'}_{n_1' \cdots n_\mu'}. \quad \text{(II.2)}$$

[Again, (II.2) is true apart from trivial changes due to particular choice of phases in (I.3) and (I.4.] If a tensor T of type (λ, μ) has the following properties:

(i) symmetry among the upper, or (1, 0)-type, indices,

(ii) symmetry among the lower, or (0, 1)-type, indices,

(iii) "tracelessness," i.e.,

$$\sum_{m_1, n_1} g_{m_1 n_1} T^{m_1 m_2 \cdots}_{n_1 n_2 \cdots}$$
$$\equiv T^{1 m_2 \cdots}_{1 n_2 \cdots} - T^{2 m_2 \cdots}_{2 n_2 \cdots} - T^{3 m_2 \cdots}_{3 n_2 \cdots} = 0, \quad \text{(II.3)}$$

then T is called an *irreducible tensor*; and the linearly independent components of T supply the UIR (λ, μ) of SU_3.

The orthonormal basic states occurring in the UIR (λ, μ) can be completely labeled by three quantities: I, I_z, Y. These form a complete commuting set of operators, within a UIR. The values of I and Y that appear in (λ, μ) were essentially determined long ago by Weyl[16] in his work on the unitary groups SU_n, and can be stated as follows. For each pair of integers f, g obeying

$$\lambda + \mu \geq f \geq \mu \geq g \geq 0, \quad \text{(II.4)}$$

we find, once, the (I, Y) multiplet

$$I = \tfrac{1}{2}(f - g), \quad Y = f + g - \tfrac{2}{3}(\lambda + 2\mu). \quad \text{(II.5)}$$

We note that the results embodied in (II.4) and (II.5) can be derived in a very simple way, from our approach [see the parenthetic remark preceding (III.10)].

III. CONSTRUCTION OF ORTHONORMAL BASIC STATES

In this section we carry out our program of explicitly constructing an orthonormal basis for the UIR (λ, μ), starting from an irreducible tensor T of type (λ, μ).*

Let the range of values of I, I_z, Y, as given by (II.4) and (II.5) be denoted by R. (Also, we shall replace I_z by the symbol M.) Let T be an irreducible tensor of type (λ, μ). Then since every index of T, upper or lower, transforms in a unitary manner, we can set up a quantity that is obviously invariant under SU_3:

$$\mathfrak{M} = \sum_{\{m\}} \sum_{\{n\}} |T^{m_1 \cdots m_\lambda}_{n_1 \cdots n_\mu}|^2. \quad \text{(III.1)}$$

Since T is irreducible and supplies us with the UIR (λ, μ) we expect to be able to define a set of quantities $\Psi^{I,Y}_M$ linearly in terms of T, such that

$$\mathfrak{M} = \sum_R |\Psi^{I,Y}_M|^2. \quad \text{(III.2)}$$

Our purpose is to compare (III.1) and (III.2) and deduce the connection between $\Psi^{I,Y}_M$ and T; from the form of $\mathfrak{M}$ in (III.2), it is legitimate to speak of the $\Psi^{I,Y}_M$ as an "orthonormal basis."

We shall rewrite (III.1) in a series of steps. Each index m_i takes on three values corresponding to the three states that appear in the UIR (1, 0) (see Fig. 1); similarly for the n_i we use the states of (0, 1). By the symmetry properties of T, any component of T is completely specified by stating how many 1's, 2's, and 3's there are among the upper indices, and similarly for the lower indices. The position where a particular index occurs is of no consequence. We can then always write components of T in the form

$$T^{111\cdots 222\cdots 33\cdots}_{111\cdots\ 22\cdots 33\cdots}.$$

Let us pick a component of T with the first $(j_1 + m_1)$ upper indices and $(j_2 - m_2)$ lower indices equal to 1, the next $(j_1 - m_1)$ upper and $(j_2 + m_2)$ lower indices equal to 2, and the remaining $(\lambda - 2j_1)$ upper and $(\mu - 2j_2)$ lower indices equal to 3. We can unambiguously *denote* this component of T by $T^{j_1 m_1}_{j_2 m_2}$. The ranges of j_1, j_2, m_1, m_2 are:

$$|m_1| \leq j_1; \quad |m_2| \leq j_2; \quad j_1 = 0, \tfrac{1}{2}, 1, \cdots \tfrac{1}{2}\lambda; \quad j_2 = 0, \tfrac{1}{2}, 1, \cdots \tfrac{1}{2}\mu. \quad \text{(III.3)}$$

Then, again because of the symmetry properties of T we can write:

$$\mathfrak{M} = \sum_{j_1 m_1} \sum_{j_2 m_2} |\psi^{j_1 m_1}_{j_2 m_2}|^2, \quad \text{(III.4)}$$

where we have set

$$\psi^{j_1 m_1}_{j_2 m_2} = N_1(j_1, m_1, j_2, m_2) T^{j_1 m_1}_{j_2 m_2},$$

[16] H. Weyl, *Group Theory and Quantum Mechanics* (Dover Publications, Inc., New York, 1931). See also J. E. Wess, Nuovo Cimento **15**, 52 (1960); M. Ikeda *et al.*, Ref. 3; S. Okubo, Ref. 3; A. J. Macfarlane *et al.*, Ref. 12; C. R. Hagen and A. J. Macfarlane, J. Math. Phys. **5**, 1335 (1964).

$$N_1(j_1, m_1, j_2, m_2) = [\lambda!\mu!]^{\frac{1}{2}}[(j_1 + m_1)!(j_1 - m_1)!(\lambda - 2j_1)! \times (j_2 + m_2)!(j_2 - m_2)!(\mu - 2j_2)!]^{-\frac{1}{2}}. \quad \text{(III.5)}$$

The significance of the variables j_1, m_1, j_2, m_2 is the following. From the definition of $T^{j_1 m_1}_{j_2 m_2}$, we know the $2j_1$ of the upper indices are of the spin-$\frac{1}{2}$ type, and $2j_2$ of the lower indices are of the spin-$\frac{1}{2}$ type (note, that on account of their formal equivalence, we use the terms spin and isospin interchangeably). Further we know that T is separately symmetric among the upper and the lower indices. Now it is a well known theorem in ordinary angular momentum theory, that if any number, $2j$, of spin-$\frac{1}{2}$ states are coupled in a *completely symmetric* manner, we automatically obtain a definite total spin, which is, in fact, the maximum possible value, j. Thus j_1 is the value of the isospin contributed by the upper indices by themselves, and j_2 that by the lower. Also, the variables m_1, m_2 were set up so as to correspond to the projection quantum numbers associated with j_1 and j_2 respectively. Finally, since we have taken care to define the operators $I_\pm$ according to the standard Condon–Shortley phase convention[17] in the UIR's (1, 0) and (0, 1) the quantities $\psi^{j_1 m_1}_{j_2 m_2}$ defined in (III.5) also are consistent with the standard phase convention.

From the form of the UIR's (1, 0) and (0, 1) as given by Fig. 1, we see that $\psi^{j_1 m_1}_{j_2 m_2}$, or $T^{j_1 m_1}_{j_2 m_2}$, possesses a definite value of hypercharge Y:

$$Y = 2(j_1 - j_2) - \tfrac{2}{3}(\lambda - \mu). \quad \text{(III.6)}$$

Also, $\psi^{j_1 m_1}_{j_2 m_2}$ possesses a definite value of total I_z:

$$I_z \equiv M = m_1 + m_2. \quad \text{(III.7)}$$

Comparing (III.6) and (III.7) with (III.5), we see that we have already succeeded in introducing two of the quantum numbers needed to specify the states of (λ, μ), and are getting close to the form (III.2) for $\mathfrak{M}$. The quantum number still to be introduced is the total isospin, I. It is clear that I must arise from a suitable coupling of j_1 and j_2. So let us define a new set of quantities, that takes us from the (m_1, m_2) representation to an (I, M) representation, viz.,

$$\psi^{I(j_1 j_2)}_M = \sum_{m_1 m_2} C(j_1, j_2, I; m_1, m_2, M)\psi^{j_1 m_1}_{j_2 m_2}. \quad \text{(III.8)}$$

Then we can write (III.4) as

$$\mathfrak{M} = \sum_{j_1, j_2} \sum_I \sum_M |\psi^{I(j_1 j_2)}_M|^2. \quad \text{(III.9)}$$

[17] E. U. Condon and G. H. Shortley, *The Theory of Atomic Spectra* (Cambridge University Press, Cambridge, England 1935).

We have yet to take into account the *tracelessness condition* which T has to satisfy. We set out to do this next.

Consider the quantities $\psi^{I(j_1 j_2)}_M$, and their construction from $T^{j_1 m_1}_{j_2 m_2}$. For the case $I = j_1 + j_2$, it is clear that, not only are any two upper spin-$\frac{1}{2}$ indices necessarily coupled to spin 1 (by symmetry), and similarly for any two lower spin-$\frac{1}{2}$ indices, but also any one upper and one lower spin-$\frac{1}{2}$ index must be coupled only to spin 1. Otherwise, we would not be able to get the maximum value $j_1 + j_2$ for I. Figuratively speaking, this is the "stretched-out configuration." This immediately suggests, that if we wish to obtain a value of I less than $j_1 + j_2$, then we essentially have to couple a sufficient number of upper and lower spin-$\frac{1}{2}$ indices to spin zero, pairwise, and then couple *all* the remaining spin-$\frac{1}{2}$ indices to *their maximum*. For, with given j_1, j_2, each I value occurs just once; and we have exhibited a mechanism to obtain such an I value.

Now from (I.7), and the remarks following it, we see that the *tracelessness* condition, (II.3iii) implies that coupling an upper spin-$\frac{1}{2}$ index and a lower spin-$\frac{1}{2}$ index to total spin zero, is numerically equivalent to replacing each of these indices by the index 3, which carries no isospin. From this and the observations made in the foregoing paragraph we realize, that in looking for *independent* linear combinations of $T^{m_1 \cdots m_\lambda}_{n_1 \cdots n_\mu}$ we need only couple j_1, j_2 in $\psi^{j_1 m_1}_{j_2 m_2}$ or $T^{j_1 m_1}_{j_2 m_2}$ to a *maximum*, $I = j_1 + j_2$; taking any lower value of I is equivalent, apart from a numerical factor, to suitably reducing j_1 and j_2 by equal amounts to $j_1^{\min}$, $j_2^{\min}$, say, and then coupling $j_1^{\min}$ and $j_2^{\min}$ to their maximum to give the desired I value: We may *remark* parenthetically, that these observations immediately lead to the rules (II.4) and (II.5). In this process of reducing j_1, j_2 to $j_1^{\min}$, $j_2^{\min}$, the hypercharge value does not change; and, for given I, Y, we easily obtain

$$\begin{aligned} Y &= 2(j_1 - j_2) - \tfrac{2}{3}(\lambda - \mu) \\ &= 2(j_1^{\min} - j_2^{\min}) - \tfrac{2}{3}(\lambda - \mu); \\ j_1^{\min} &= \tfrac{1}{2}(j_1 - j_2 + I) \\ &= \tfrac{1}{2}I + \tfrac{1}{4}Y + \tfrac{1}{6}(\lambda - \mu); \\ j_2^{\min} &= \tfrac{1}{2}(j_2 - j_1 + I) \\ &= \tfrac{1}{2}I - \tfrac{1}{4}Y - \tfrac{1}{6}(\lambda - \mu). \end{aligned} \quad \text{(III.10)}$$

In our scheme, then, tracelessness can be expressed in the following manner:

$$\psi^{I(j_1 j_2)}_M = N_2(j_1, j_2, I)\psi^{I(\frac{1}{2}[j_1 - j_2 + I], \frac{1}{2}[j_2 - j_1 + I])}_M. \quad \text{(III.11)}$$

Clearly, the factor N_2 is independent of M. We now proceed to evaluate this factor.

As a first step we have to solve a problem connected with the SU_2 group, and coupling of angular momenta. For this purpose it proves helpful to use the method of boson operators introduced by Schwinger and widely used by Bargmann and Moshinsky.[18] Consider two boson oscillators with creation and destruction operators $a_\pm^*$, $a_\pm$, respectively. They obey the standard CR's.

$$[a_+, a_\mp^*] = [a_-, a_\pm^*] = 1, \qquad [a_+, a_-] = [a_+, a_\pm^*] = \cdots = 0 \tag{III.12}$$

We define the vacuum state $|0\rangle$ by

$$a_+ |0\rangle = a_- |0\rangle = 0; \qquad \langle 0|0\rangle = 1. \tag{III.13}$$

Now define three operators $J_1^{(a)}$, $J_2^{(a)}$, $J_3^{(a)}$ by

$$J_1^{(a)} \pm iJ_2^{(a)} = J_\pm^{(a)} = a_\pm^* a_\mp; \; J_3^{(a)} = \tfrac{1}{2}(a_+^* a_+ - a_-^* a_-). \tag{III.14}$$

Then, using (III.12), one can easily show that the operators $J_i^{(a)}$ obey the CR's of the SU_2 algebra, that is the algebra of angular momentum. Further, with respect to $J_i^{(a)}$, the pair $a_\pm^*$ and $a_\pm^*$ form the $+\frac{1}{2}$ and $-\frac{1}{2}$ components of a spherical tensor of rank $\frac{1}{2}$.

We may now proceed to arrange the states of the two oscillators so that they provide IR's of the operators $J_i^{(a)}$. It is easily shown that the $(2j+1)$ normalized states

$$|j, m\rangle = \frac{(a_+^*)^{j+m}(a_-^*)^{j-m}}{[(j+m)!(j-m)!]^{\frac{1}{2}}} |0\rangle; \qquad m = -j, \cdots j-1, j; \tag{III.15}$$

provide us with the representation of $J_i^{(a)}$ corresponding to spin j. In this way, each UIR of SU_2 appears once among the states of the two oscillators.

Let us now add two more oscillators with operators $b_\pm^*$, $b_\pm$ obeying CR's exactly like (III.12). Further, let all the a-type operators commute with the b-type operators, and let them possess a common vacuum. Analogous to (III.14), define operators $J_i^{(b)}$ and set

$$J_i = J_i^{(a)} + J_i^{(b)}. \tag{III.14'}$$

Then the J_i also obey the CR's of angular momentum. With respect to J_i, both $(a_\pm^*, a_\pm^*)$ and $(b_\pm^*, b_\pm^*)$ form spin-$\frac{1}{2}$ systems. We can check that the operators J_i commute with the combination

$$(a_\pm^* b_\mp^* - a_\mp^* b_\pm^*). \tag{III.16}$$

The operator expression (III.16) is a scalar operator.

In an obvious way we define the following normalized states:

$$|j_1, m_1, j_2, m_2\rangle = \frac{(a_+^*)^{j_1+m_1}(a_-^*)^{j_1-m_1}(b_+^*)^{j_2+m_2}(b_-^*)^{j_2-m_2}|0\rangle}{[(j_1+m_1)!(j_1-m_1)!(j_2+m_2)!(j_2-m_2)!]^{\frac{1}{2}}}, \tag{III.17}$$

$$|IM(j_1 j_2)\rangle = \sum_{m_1 m_2} C(j_1, j_2, I; m_1, m_2, M)\, |j_1, m_1, j_2, m_2\rangle. \tag{III.18}$$

Then the foregoing statements concerning coupling of two angular momenta j_1 and j_2 to a total $I < j_1 + j_2$ are expressed by the equation

$$|IM(j_1 j_2)\rangle = \Lambda(j_1 j_2 I)(a_\pm^* b_\mp^* - a_\mp^* b_\pm^*)^{j_1+j_2-I} \left| IM\left(\frac{j_1 - j_2 + I}{2}, \frac{j_2 - j_1 + I}{2}\right)\right\rangle, \tag{III.19}$$

We shall first evaluate the quantity Λ, and then derive N_2 of (III.11) using Λ. To find Λ, take the case $M = I$ in (III.19), and substitute from (III.17) and (III.18) on both sides of (III.19). Then on the rhs, we have a very simple CG coefficient whose value is unity, while on the left, we have a sum over m_1:

$$\sum_{m_1 m_2} C(j_1 j_2 I; m_1 m_2 I) \frac{(a_+^*)^{j_1+m_1}(a_-^*)^{j_1-m_1}(b_+^*)^{j_2+m_2}(b_-^*)^{j_2-m_2}|0\rangle}{[(j_1+m_1)!(j_1-m_1)!(j_2+m_2)!(j_2-m_2)!]^{\frac{1}{2}}}$$
$$= \Lambda(j_1 j_2 I)(a_\pm^* b_\mp^* - a_\mp^* b_\pm^*)^{j_1+j_2-I} \frac{(a_+^*)^{j_1-j_2+I}(b_+^*)^{j_2-j_1+I}|0\rangle}{[(j_1-j_2+I)!(j_2-j_1+I)!]^{\frac{1}{2}}}. \tag{III.20}$$

The special CG coefficient in (III.20) can be obtained from the general expression for $C(j_1 j_2 I; m_1 m_2 M)$ and consists of a single term. Comparing both sides of (III.20), one finds

$$\Lambda(j_1 j_2 I) = [(2I+1)!]^{\frac{1}{2}} \times [(j_1 + j_2 - I)!(j_1 + j_2 + I + 1)!]^{-\frac{1}{2}}. \tag{III.21}$$

[18] J. Schwinger, *On Angular Momentum*, USAEC, NYO 3071, 1952 (unpublished). V. Bargmann, Rev. Mod. Phys. **34**, 829 (1962); V. Bargmann and M. Moshinsky, Nucl. Phys. **18**, 697 (1960); **23**, 177 (1961).

We find N_2 from Λ as follows. (III.19) is a property of the SU_2 group. In (III.19) we can use components of our irreducible tensor $T^{j_1 m_1}_{j_2 m_2}$ in place of the states and operators appearing there. The $a^*_\pm$ operators play then the role of upper spin-$\frac{1}{2}$ indices, the $b^*_\pm$ of lower spin-$\frac{1}{2}$ indices. The operator expression on the rhs of (III.19) expresses explicitly the "pairwise zero coupling" of upper and lower spin-$\frac{1}{2}$ indices in $T^{j_1 m_1}_{j_2 m_2}$. [(III.17, 18) are analogs of (III.5) and (III.8) respectively.] From (II.3iii) we know that the "zero coupling" appearing on the rhs of (III.19) allows us to just increase the number of indices in $T^{j_1 m_1}_{j_2 m_2}$ that have the value 3. Thus beginning on the lhs of (III.19) with $(\lambda - 2j_1)$ upper indices equal to 3, and $(\mu - 2j_2)$ lower indices equal to 3, we use the tracelessness of $T^{m_1 \cdots m_\lambda}_{n_1 \cdots n_\mu}$ to obtain on the rhs $(\lambda - 2j_1 + j_1 + j_2 - I)$ upper and $(\mu - 2j_2 + j_1 + j_2 - I)$ lower indices equal to 3. At this point, we must remember that N_1 in (III.5) has combinatorial factors

$$[(\lambda - 2j_1)!(\mu - 2j_2)!]^{-\frac{1}{2}}$$

that do not appear in (III.17, III.18). This is because the indices in $T^{m_1 \cdots m_\lambda}_{n_1 \cdots n_\mu}$ have also the value 3 in their range. Naturally such factors do not appear in (III.17)–(III.21) which deal with SU_2. Remembering that the number of 3's changes when we use tracelessness, we combine (III.5), (III.8), (III.17)–(III.19), and (III.24) to get

$$\begin{aligned}\psi_M^{I(j_1 j_2)} &= [(\lambda - 2j_1)!(\mu - 2j_2)!]^{-\frac{1}{2}} \Lambda(j_1 j_2 I) \\ &\quad \times [(\lambda - j_1 + j_2 - I)!(\mu - j_2 + j_1 - I)!]^{\frac{1}{2}} \psi_M^{I(\frac{1}{2}[j_1 - j_2 + I], \frac{1}{2}[j_2 - j_1 + I])} \\ &= N_2(j_1, j_2, I) \psi_M^{I(\frac{1}{2}[j_1 - j_2 + I], \frac{1}{2}[j_2 - j_1 + I])}, \end{aligned} \tag{III.22}$$

with

$$N_2(j_1, j_2, I) = \left[\frac{(2I + 1)!(\lambda - j_1 + j_2 - I)!(\mu - j_2 + j_1 - I)!}{(j_1 + j_2 - I)!(j_1 + j_2 + I + 1)!(\lambda - 2j_1)!(\mu - 2j_2)!} \right]^{\frac{1}{2}}. \tag{III.23}$$

This completes the evaluation of N_2.

Let us now return to our basic equation for the invariant $\mathfrak{M}$:

$$\mathfrak{M} = \sum_{j_1 j_2} \sum_I \sum_M |\psi_M^{I(j_1 j_2)}|^2. \tag{III.9}$$

Let us choose some value for Y; this fixes the difference $(j_1 - j_2)$, by (III.10). Let us next choose some value for I. Then all the terms in (III.9) for this pair of values of I and Y are multiples, by (III.22), of the "basic" term.

$$\psi_M^{I(j_1^{\min}, j_2^{\min})}, \tag{III.24}$$

where $j_1^{\min}$ and $j_2^{\min}$ are given in (III.10). We can therefore write

$$\mathfrak{M} = \sum_R [N_3(I, Y)]^2 \, |\psi_M^{I(j_1^{\min}, j_2^{\min})}|^2. \tag{III.25}$$

The quantity $[N_3]^2$ is clearly a sum of factors $[N_2]^2$ over (j_1, j_2) for fixed I, Y, i.e., fixed $(j_1 - j_2)$. We evaluate N_3 in the Appendix A; the result is

$$N_3(I, Y) = [(2I + 1)!(\lambda + \mu + 1)!]^{\frac{1}{2}} \left[\left(\frac{\lambda + 2\mu}{3} + I + \frac{Y}{2} + 1 \right)! \left(\frac{2\lambda + \mu}{3} + I - \frac{Y}{2} + 1 \right)! \right]^{-\frac{1}{2}}. \tag{III.26}$$

Having obtained the form (III.25) our task is done. Apart from a phase, which we shall choose suitably in the next section, we set

$$\Psi_M^{I,Y} = N_3(I, Y) \psi_M^{I(j_1^{\min} j_2^{\min})} = N_3(I, Y) \sum_{m_1 m_2} C(j_1^{\min}, j_2^{\min}, I; m_1 m_2 M) N_1(j_1^{\min} m_1, j_2^{\min} m_2) T^{j_1^{\min} m_1}_{j_2^{\min} m_2}, \tag{III.27}$$

where $j_1^{\min}$, $j_2^{\min}$, are defined in Eq. (III.10). (III.27) is easily inverted to express T in terms of Ψ:

$$\begin{aligned} T^{j_1 m_1}_{j_2 m_2} &= [N_1(j_1 m_1 j_2 m_2)]^{-1} \psi^{j_1 m_1}_{j_2 m_2} = [N_1(j_1 m_1 j_2 m_2)]^{-1} \sum_I C(j_1 j_2 I; m_1 m_2 M) \psi_M^{I(j_1 j_2)} \\ &= [N_1(j_1 m_1 j_2 m_2)]^{-1} \sum_I C(j_1 j_2 I; m_1 m_2 M) N_2(j_1 j_2 I) \psi_M^{I(\frac{1}{2}[j_1 - j_2 + I], \frac{1}{2}[j_2 - j_1 + I])} \\ &= [N_1(j_1 m_1 j_2 m_2)]^{-1} \sum_I C(j_1 j_2 I; m_1 m_2 M) N_2(j_1 j_2 I) [N_3(I, Y)]^{-1} \Psi_M^{I,Y}; \\ Y &= 2(j_1 - j_2) - \tfrac{2}{3}(\lambda - \mu). \end{aligned} \tag{III.28}$$

(III.27)–(III.28) are the basic equations of our method. (III.27) allows us to express orthonormal basic states of a UIR (λ, μ) in terms of irreducible tensors of type (λ, μ). It is useful in defining physical particles, in terms of appropriate tensors. As a simple application, we consider the octet representation of SU_3, $\lambda = \mu = 1$, and take $I = M = Y = 0$ on the lhs of (III.27). Then $j_1^{\min} = j_2^{\min} = 0$, and we get

$$\Psi_0^{0,0} = N_3(0, 0)N_1(0, 0; 0, 0)T_3^3 = (3!/2!\,2!)^{\frac{1}{2}}T_3^3 = 3T_3^3/6^{\frac{1}{2}}. \quad \text{(III.29)}$$

For $I = M = 1$, $Y = 0$, for instance, $j_1^{\min} = j_2^{\min} = \frac{1}{2}$ and

$$\Psi_1^{1,0} = N_3(1, 0)N_1(\tfrac{1}{2}, \tfrac{1}{2}; \tfrac{1}{2}, \tfrac{1}{2})T_2^1 = T_2^1. \quad \text{(III.30)}$$

(III.29–III.30) are quite familiar (apart from phases) in connection with tensor representations of the Λ^0 and Σ^+ states of the octet model based on SU_3.[19]

As we have already stated, there is a great deal of freedom in the choice of phase in (III.24). We make a definite choice in the following section when we adopt Biedenharn's phase convention.

Note added in manuscript: It has been pointed out to us by the referee that the work of M. Moshinsky in constructing polynomial bases for irreducible representations of SU_3 [J. Math. Phys. **4**, 1128(1963); see also M. Moshinsky, Ref. (5)], can be related to the present work as follows. Let us momentarily use a metric in which the invariant of (I.7) is constructed with a Kronecker delta δ_j^i. One starts with two "vectors" x^m, y_n of types (1, 0), (0, 1) respectively, obeying the restriction

$$\sum_m x^m y_m = 0. \quad \text{(i)}$$

With the help of these two vectors, one can construct a *particular* irreducible tensor T of type (λ, μ) by defining

$$T_{n_1 \cdots n_\mu}^{m_1 \cdots m_\lambda} = x^{m_1} \cdots x^{m_\lambda} y_{n_1} \cdots y_{n_\mu}. \quad \text{(ii)}$$

[Note however that every irreducible tensor T of type (λ, μ) is *not* of the form (ii).] If one wants to build T from vectors of type (1, 0) alone, define

$$y_n = \sum_{m'm''} \epsilon_{nm'm''} x^{m'} z^{m''}, \quad \text{(iii)}$$

where $\epsilon_{nm'm''}$ is the completely antisymmetric tensor, and $z^{m''}$ is an arbitrary vector of type (1, 0). Then (i) is automatically obeyed. Now denote the vectors $x^{m'}$, $z^{m''}$ by

$$x^{m'} \equiv x_{(1)}^{m'}, \qquad z^{m''} = x_{(2)}^{m''} \quad \text{(iv)}$$

and introduce the determinantal notation:

$$\Delta_S^m \equiv x_{(S)}^m, \quad \Delta_{S'S''}^{m'm''} \equiv x_{(S')}^{m'} x_{(S'')}^{m''} - x_{(S')}^{m''} x_{(S'')}^{m'}. \quad \text{(v)}$$

From (iii) one has

$$y_1 = \Delta_{12}^{23}, \quad y_2 = \Delta_{12}^{31}, \quad y_3 = \Delta_{12}^{12}, \quad x^m = \Delta_1^m. \quad \text{(vi)}$$

With the above notations, and definition (ii) for T, the component of T denoted by us as [see above (III.3)]

$$T_{j_2 m_2}^{j_1 m_1}$$

becomes

$$(\Delta_1^1)^{j_1+m_1}(\Delta_1^2)^{j_1-m_1}(\Delta_1^3)^{\lambda-2j_1}(\Delta_{12}^{23})^{j_2-m_2} (\Delta_{12}^{31})^{j_2+m_2}(\Delta_{12}^{12})^{\mu-2j_2}. \quad \text{(vii)}$$

Expression (vii) is the one given by Moshinsky (except for interchange of upper and lower indices) for the basis of irreducible representations of SU_3. We note that Moshinsky uses the formalism of boson creation and destruction operators, and this technique has been used by him to construct bases for irredicible representations of the unitary groups SU_n, for all n.

IV. MATRIX ELEMENTS OF GENERATORS

In this section we show how to derive the matrix elements of the generators of SU_3, in the UIR (λ, μ), in an orthonormal basis $\Psi_M^{I,Y}$ labeled by I, M, Y.

The forms of the operators I_z, $I_\pm$ are well known from the theory of SU_2 and read:

$$\begin{aligned} I_z\Psi_M^{I,Y} &= M\Psi_M^{I,Y}, \\ I_\pm\Psi_M^{I,Y} &= [(I \mp M)(I \pm M + 1)]^{\frac{1}{2}}\Psi_{M\pm1}^{I,Y}. \end{aligned} \quad \text{(IV.1)}$$

For H_2, or Y, we have

$$Y\Psi_M^{I,Y} = Y\Psi_M^{I,Y}. \quad \text{(IV.2)}$$

Using the Wigner–Eckart theorem for SU_2, and the fact that $F_\pm$ form the $\pm\frac{1}{2}$ components of a spherical tensor of rank $\frac{1}{2}$, all matrix elements of $F_\pm$ can be expressed in terms of two reduced matrix elements, which are functions of I and Y. We write this as follows, denoting basic states of the UIR (λ, μ) for the moment, as $|(\lambda, \mu); IMY\rangle$,

[19] S. Okubo, Ref. 3.

$$\langle(\lambda,\mu);\, I+\tfrac{1}{2},\, M\pm\tfrac{1}{2},\, Y+1\,|F_\pm|\,(\lambda,\mu);\, IMY\rangle = C(I,\tfrac{1}{2},I+\tfrac{1}{2};\, M,\pm\tfrac{1}{2},M\pm\tfrac{1}{2})f_1(\lambda\mu;\, IY),$$
$$\langle(\lambda,\mu);\, I-\tfrac{1}{2},\, M\pm\tfrac{1}{2},\, Y+1\,|F_\pm|\,(\lambda,\mu);\, IMY\rangle = C(I,\tfrac{1}{2},I-\tfrac{1}{2};\, M,\pm\tfrac{1}{2},M\pm\tfrac{1}{2})f_2(\lambda\mu;\, IY). \tag{IV.3}$$

The analogous relations for $G_\pm$ can be derived by complex conjugation from (IV.3) and relabeling of variables I, Y. They read:

$$\begin{aligned}\langle(\lambda,\mu);\, I+\tfrac{1}{2},\, M\pm\tfrac{1}{2},\, Y-1|\,G_\pm\,|(\lambda,\mu);\, IMY\rangle \\ = -[(2I+1)/(2I+2)]^{\frac{1}{2}}C(I,\tfrac{1}{2},I+\tfrac{1}{2};\, M,\pm\tfrac{1}{2},M\pm\tfrac{1}{2})f_2^*(\lambda\mu;\, I+\tfrac{1}{2},\, Y-1),\\ \langle(\lambda,\mu);\, I-\tfrac{1}{2},\, M\pm\tfrac{1}{2},\, Y-1|\,G_\pm\,|(\lambda,\mu),\, IMY\rangle \\ = [(2I+1)/2I]^{\frac{1}{2}}C(I,\tfrac{1}{2},I-\tfrac{1}{2};\, M,\pm\tfrac{1}{2},M\pm\tfrac{1}{2})f_1^*(\lambda\mu;\, I-\tfrac{1}{2},\, Y-1).\end{aligned} \tag{IV.4}$$

The object is to find the functions f_1 and f_2. We shall in the sequel omit the variables $(\lambda\mu)$ in states and in f_1 and f_2, as they do not change in the course of our discussion.

The simplest way to find f_1 and f_2 is to take a special case of (IV.3), say, and rewrite it as an equation for states Ψ related to an irreducible tensor T by (III.27) and (III.28). Choose $M = I$ and take the operator F_- in (IV.3). We can then write

$$\begin{aligned}F_-\Psi_I^{I,Y} &= C(I,\tfrac{1}{2},I+\tfrac{1}{2};\, I,-\tfrac{1}{2},I-\tfrac{1}{2})\\ &\quad\times f_1(I,\, Y)\Psi_{I-\frac{1}{2}}^{I+\frac{1}{2},\,Y+1}\\ &\quad+ C(I,\tfrac{1}{2},I-\tfrac{1}{2};\, I,-\tfrac{1}{2},I-\tfrac{1}{2})\\ &\quad\times f_2(I,\, Y)\Psi_{I-\frac{1}{2}}^{I-\frac{1}{2},\,Y+1}.\end{aligned} \tag{IV.5}$$

We will now explicitly evaluate the lhs of (IV.5) and identify f_1 and f_2; starting with a given I and Y, we introduce the quantities $j_1^{\min}$, $j_2^{\min}$ and refer to them from now on as j_1, j_2:

$$\begin{aligned}j_1 &= \tfrac{1}{4}Y + \tfrac{1}{6}(\lambda-\mu) + \tfrac{1}{2}I,\\ j_2 &= -\tfrac{1}{4}Y - \tfrac{1}{6}(\lambda-\mu) + \tfrac{1}{2}I.\end{aligned} \tag{III.10}$$

We have, using (III.27):

$$\begin{aligned}F_-\Psi_I^{I,Y} &= F_-\{N_3(I,\, Y)N_1(j_1j_1j_2j_2)T_{j_2j_2}^{j_1j_1}\}\\ &= N_3(I,\, Y)N_1(j_1j_1j_2j_2)\{F_-T_{j_2j_2}^{j_1j_1}\}\\ &= N_3(I,\, Y)N_1(j_1j_1j_2j_2)\{6^{\frac{1}{2}}E_{+3}T_{j_2j_2}^{j_1j_1}\}.\end{aligned} \tag{IV.6}$$

To evaluate the curly bracket term in (IV.6), we remember from (I.3) and (I.4) that the operator $6^{\frac{1}{2}}E_{+3}$ has the following effect: acting on an upper index, it changes a 3 into a 2; acting on a lower index it changes a 2 into a 3, and gives an additional minus sign. Since we also know that the generators for a direct product of many UIR's are the sums of the generators of the individual UIR's, we see that $6^{\frac{1}{2}}E_{+3}$ acting on $T_{j_2j_2}^{j_1j_1}$ changes every upper index 3 into a 2, one at a time, and then each lower index 2 into a 3, again one at a time; the latter terms have an extra minus sign. Using again the fact that T is symmetric, we have

$$\begin{aligned}&6^{\frac{1}{2}}E_{+3}T_{j_2j_2}^{j_1j_1}\\ &\quad= (\lambda-2j_1)T_{j_2j_2}^{j_1+\frac{1}{2},\,j_1-\frac{1}{2}} - 2j_2T_{j_2-\frac{1}{2},\,j_2-\frac{1}{2}}^{j_1j_1},\end{aligned} \tag{IV.7}$$

so that

$$\begin{aligned}F_-\Psi_I^{I,Y} &= N_3(I,\, Y)N_1(j_1j_1j_2j_2)\\ &\quad\times\{(\lambda-2j_1)T_{j_2j_2}^{j_1+\frac{1}{2},\,j_1-\frac{1}{2}} - 2j_2T_{j_2-\frac{1}{2},\,j_2-\frac{1}{2}}^{j_1j_1}\}.\end{aligned} \tag{IV.8}$$

We now use (III.28) to write the T in (IV.8) in terms of Ψ. We find

$$\begin{aligned}T_{j_2j_2}^{j_1+\frac{1}{2},\,j_1-\frac{1}{2}} &= \frac{N_2(j_1+\frac{1}{2},\, j_2,\, I+\frac{1}{2})}{N_1(j_1+\frac{1}{2},\, j_1-\frac{1}{2},\, j_2,\, j_2)}\\ &\times\frac{C(j_1+\frac{1}{2},\, j_2,\, I+\frac{1}{2};\, j_1-\frac{1}{2},\, j_2,\, I-\frac{1}{2})}{N_3(I+\frac{1}{2},\, Y+1)}\,\Psi_{I-\frac{1}{2}}^{I+\frac{1}{2},\,Y+1}\\ &+\frac{N_2(j_1+\frac{1}{2},\, j_2,\, I-\frac{1}{2})}{N_1(j_1+\frac{1}{2},\, j_1-\frac{1}{2},\, j_2,\, j_2)}\\ &\times\frac{C(j_1+\frac{1}{2},\, j_2,\, I-\frac{1}{2};\, j_1-\frac{1}{2},\, j_2,\, I-\frac{1}{2})}{N_3(I-\frac{1}{2},\, Y+1)}\,\Psi_{I-\frac{1}{2}}^{I-\frac{1}{2},\,Y+1}.\end{aligned} \tag{IV.9}$$

$$\begin{aligned}T_{j_2-\frac{1}{2},\,j_2-\frac{1}{2}}^{j_1j_1} &= \frac{N_2(j_1,\, j_2-\frac{1}{2},\, I-\frac{1}{2})}{N_1(j_1,\, j_1,\, j_2-\frac{1}{2},\, j_2-\frac{1}{2})}\\ &\times\frac{1}{N_3(I-\frac{1}{2},\, Y+1)}\cdot\Psi_{I-\frac{1}{2}}^{I-\frac{1}{2},\,Y+1}.\end{aligned} \tag{IV.10}$$

Using (IV.9)–(IV.10) in (IV.8), and comparing the terms $\Psi_{I-\frac{1}{2}}^{I+\frac{1}{2},\,Y+1}$ on both sides of (IV.5) we get

$$\begin{aligned}&C(I,\tfrac{1}{2},I+\tfrac{1}{2};\, I,-\tfrac{1}{2},I-\tfrac{1}{2})f_1(I,\, Y)\\ &=\frac{N_3(I,\, Y)}{N_3(I+\frac{1}{2},\, Y+1)}\cdot\frac{N_1(j_1j_1j_2j_2)}{N_1(j_1+\frac{1}{2},\, j_1-\frac{1}{2},\, j_2,\, j_2)}\\ &\times(\lambda-2j_1)N_2(j_1+\tfrac{1}{2},\, j_2,\, I+\tfrac{1}{2})\\ &\times C(j_1+\tfrac{1}{2},\, j_2,\, I+\tfrac{1}{2};\, j_1-\tfrac{1}{2},\, j_2,\, I-\tfrac{1}{2}).\end{aligned} \tag{IV.11}$$

The ratios of N factors appearing in (IV.11) are easily evaluated; the factor N_2 turns out to be unity;

the special CG coefficients present in (IV.11) also have simple forms. With the values of j_1, j_2, given above, one then gets:

$$f_1(I, Y) = \frac{1}{(2I+2)^{\frac{1}{2}}}\left\{\left(\frac{\lambda+2\mu}{3}+I+\frac{Y}{2}+2\right)\right.$$
$$\left.\times\left(\frac{\lambda-\mu}{3}+I+\frac{Y}{2}+1\right)\left(\frac{2\lambda+\mu}{3}-I-\frac{Y}{2}\right)\right\}^{\frac{1}{2}}. \quad \text{(IV.12)}$$

In an analogous manner one gets

$$f_2(I, Y) = -\frac{1}{(2I)^{\frac{1}{2}}}\left\{\left(\frac{2\lambda+\mu}{3}+I-\frac{Y}{2}+1\right)\right.$$
$$\left.\times\left(\frac{\mu-\lambda}{3}+I-\frac{Y}{2}\right)\left(\frac{\lambda+2\mu}{3}-I+\frac{Y}{2}+1\right)\right\}^{\frac{1}{2}}. \quad \text{(IV.13)}$$

Equations (IV.1)–(IV.4) and (IV.12)–(IV.13) constitute the solution to the problem of evaluating the matrix elements of the generators of SU_3 in a arbitrary UIR (λ, μ). They differ from the results given by Biedenharn,[20] only in that Biedenharn's function f_2 has an opposite sign to ours. Since we wish to be in accord with Biedenharn's phase conventions, at this point we make a definite choice of phase in Eq. (III.27). As was mentioned there, we can always introduce a phase dependent on λ, μ, I, Y in the connection between $\Psi_M^{I,Y}$ and $T^{j_1 m_1}_{j_2 m_2}$; for such a phase would never appear in the invariant $\mathfrak{M}$. In terms of the $\Psi_M^{I,Y}$ of (III.27), we define new basic functions $\Phi_M^{I,Y}$ as follows:

$$\Phi_M^{I,Y} = (-1)^{I-\frac{1}{2}Y-\frac{1}{2}(\lambda+2\mu)}\Psi_M^{I,Y}. \quad \text{(IV.14)}$$

This phase is chosen so as to be *real*, and to be unity when I is maximum. If we now define functions $f_1(I, Y)$ and $f_2(I, Y)$ as reduced matrix elements of $F_\pm$ taken between states $\Phi_M^{I,Y}$ then f_1 is the same as before, (IV.12), while f_2 changes sign, as compared with (IV.13). In comformity with Biedenharn, we have then

$$\langle(\lambda\mu); I+\tfrac{1}{2}, M\pm\tfrac{1}{2}, Y+1|\, F_\pm\, |(\lambda\mu); IMY\rangle = C(I, \tfrac{1}{2}, I+\tfrac{1}{2}; M, \pm\tfrac{1}{2}, M\pm\tfrac{1}{2})(2I+2)^{-\frac{1}{2}}$$
$$\times\left\{\left(\frac{\lambda+2\mu}{3}+I+\frac{Y}{2}+2\right)\left(\frac{\lambda-\mu}{3}+I+\frac{Y}{2}+1\right)\left(\frac{2\lambda+\mu}{3}-I-\frac{Y}{2}\right)\right\}^{\frac{1}{2}},$$
$$\langle(\lambda\mu); I-\tfrac{1}{2}, M\pm\tfrac{1}{2}, Y+1|\, F_\pm\, |(\lambda\mu); IMY\rangle = C(I, \tfrac{1}{2}, I-\tfrac{1}{2}; M, \pm\tfrac{1}{2}, M\pm\tfrac{1}{2})(2I)^{-\frac{1}{2}}$$
$$\times\left\{\left(\frac{\lambda+2\mu}{3}-I+\frac{Y}{2}+1\right)\left(\frac{\mu-\lambda}{3}+I-\frac{Y}{2}\right)\left(\frac{2\lambda+\mu}{3}+I-\frac{Y}{2}+1\right)\right\}^{\frac{1}{2}},$$
$$|(\lambda, \mu); IMY\rangle \equiv \Phi_M^{I,Y}. \quad \text{(IV.15)}$$

It is worth stressing that our final choice of $\Phi_M^{I,Y}$ defined in terms of T must be taken in conjunction with the definitions (I.3–I.4) of the fundamental UIR's (1, 0), (0, 1).

V. MATRIX ELEMENTS OF OCTET OPERATORS

In a recent paper, Lurié and Macfarlane[21] have calculated the matrix elements of the *regular* (or *octet*) operators of SU_3. They have employed algebraic methods making a very judicious use of the commutation relations of the octet operators with the generators of the group, and also the results of Biedenharn for the matrix elements of the generators. In the present section, we propose to use the method of Sec. III for obtaining the matrix elements of the octet operators as an illustration of the simplicity of calculation made possible by this method. Knowledge of the matrix elements of the generators is not required for this calculation. We shall concern ourselves with the one special case not treated by Lurié and Macfarlane in their paper; so that the results of this section will be a natural supplement to their results. We rely heavily for notations and general observations on their paper so as to avoid repetition.

Lurié and Macfarlane arrange the octet operators in the set

$$(\mathcal{I}_z, \mathcal{Y}, \mathcal{I}_\pm, \mathcal{F}_\pm, \mathcal{G}_\pm), \quad \text{(V.1)}$$

corresponding to the set of generators [cf. (I.10)]

$$(I_z, Y, I_\pm, F_\pm, G_\pm), \quad \text{(V.2)}$$

where the former transform under SU_3 exactly like the latter. For the purpose of our method, which makes use of tensors, we look upon the octet operators as components of a traceless tensor S^p_q, and make the following identifications:

$$\mathcal{I}_z = \tfrac{1}{2}(S^1_1 + S^2_2), \quad \mathcal{Y} = S^3_3, \quad \mathcal{I}_+ = -S^1_2,$$
$$\mathcal{I}_- = S^2_1, \quad \mathcal{F}_+ = -S^1_3, \quad \mathcal{F}_- = -S^2_3, \quad \text{(V.3)}$$
$$\mathcal{G}_+ = S^3_2, \quad \mathcal{G}_- = S^3_1.$$

[20] L. C. Biedenharn, Ref. 4. See also G. E. Baird and L. C. Biedenharn, J. Math. Phys. 4, 1449 (1963).
[21] D. Lurié and A. J. Macfarlane, Ref. 14.

The operator S^p_q, which transforms under SU_3 according to the UIR (1, 1), will connect a UIR (λ, μ) to UIR (λ', μ') occurring in the reduction of the Kronecker product

$$(1, 1) \otimes (\lambda, \mu). \tag{V.4}$$

Lurié and Macfarlane note that this product contains

(1) $(\lambda+2, \mu-1)$ once, unless $\mu=0$,

(2) $(\lambda-1, \mu-1)$ once, unless $\mu=0$ or $\lambda=0$,

(3) $(\lambda-2, \mu+1)$ once, unless $\lambda=0$ or 1,

(4) $(\lambda+1, \mu+1)$ once,

(5) $(\lambda-1, \mu+2)$ once, unless $\lambda=0$,

(6) $(\lambda+1, \mu-2)$ once, unless $\mu=0$ or 1,

(7) (λ, μ) twice, if $\lambda, \mu \neq 0$;
once, if $\lambda=0, \mu\neq0$, or $\mu=0, \lambda\neq0$;
not at all if $\lambda=\mu=0$. (V.5)

They have treated the first six of the above cases by their method. In the seventh case the UIR (λ, μ) occurs twice, in general. It is this case we treat here.

There is a rather unique way of distinguishing the two occurrences of (λ, μ) in $(1, 1) \otimes (\lambda, \mu)$ in the tensor method which we employ. This may be considered as an added advantage of the present approach. What we have to do is to construct irreducible tensors of type (λ, μ) from the traceless tensors S^p_q and the irreducible tensor $T^{m_1m_2\cdots m_\lambda}_{n_1n_2\cdots n_\mu}$ which, being irreducible, is symmetric in all of its λ upper indices, symmetric in all of its μ lower indices, and is traceless. From these two tensors we can construct two independent tensors with λ upper and μ lower indices in the following manner.

$$\begin{aligned} &\text{(I)} && S^p_q T^{m_1m_2\cdots m_\lambda}_{\mathbf{n_1n_2}\cdots n_\mu}, \\ &\text{(II)} && S^p_q T^{\mathbf{m_1}m_2\cdots m_\lambda}_{n_1\mathbf{n_2}\cdots n_\mu}, \end{aligned} \tag{V.6}$$

where we *contract* the two boldface indices. These two different ways of contracting indices lead us to the two independent (λ, μ) occurrences mentioned above, after we have carried out the symmetrization of the upper, and of the lower indices and made the resulting tensors traceless. Once this has been done the results of Sec. III can be used in almost a mechanical way to obtain the desired results for the matrix elements of the octet operators for the two cases under consideration. One can choose suitable linear combinations of the two sets of results if desired. It may be mentioned that the above construction does not, in general, lead automatically to two mutually orthogonal representation spaces.

In the following we are not concerned with any normalization factors, and are satisfied with giving the (I, M, Y) dependence of the matrix elements.

Before proceeding further with our calculations, we refer to the following observations made by Lurié and Macfarlane. The operators $\mathcal{Y}$, $\mathcal{I}_z$, $\mathcal{I}_\pm$ conserve Y; $\mathcal{F}_\pm$ raise Y to $Y+1$; $\mathcal{G}_\pm$ lowers Y to $Y-1$. Further, the set

$$(-\mathcal{I}_+/\sqrt{2}, \mathcal{I}_z, \mathcal{I}_-/\sqrt{2}) \tag{V.7}$$

can be looked upon as the $M = 1, 0, -1$ components of a spherical vector with respect to $R_3(I)$. $\mathcal{F}_\pm$ behave as the $M = +\frac{1}{2}, -\frac{1}{2}$ components of a spherical tensor of rank $\frac{1}{2}$ with respect to $R_3(I)$; so also the $\mathcal{G}_\pm$. These observations immediately lead to a major simplification, in that we can use the Wigner–Eckart theorem for $R_3(I)$ and write, with Lurié and Macfarlane,

$$\langle \lambda'\mu'; I'MY\ |\mathcal{I}_z|\ \lambda\mu; IMY\rangle = C(I, 1, I'; M, 0, M)\mathcal{I}(\lambda\mu IY; \lambda'\mu' I'Y)\langle\lambda'\mu'||\ (1, 1)\ ||\lambda\mu\rangle, \tag{V.8i}$$

$$\langle \lambda'\mu'; I', M \pm 1, Y|\ \mathcal{I}_\pm\ |\lambda\mu; IMY\rangle = \mp\sqrt{2}C(I, 1, I'; M, \pm1, M \pm 1)\mathcal{I}(\lambda\mu IY; \lambda'\mu' I'Y)\langle\lambda'\mu'||\ (1, 1)\ ||\lambda\mu\rangle, \tag{V.8ii}$$

$$\begin{aligned}\langle \lambda'\mu'; I', M \pm \tfrac{1}{2}, Y+1|\ \mathcal{F}_\pm\ |\lambda\mu; IMY\rangle \\ = C(I, \tfrac{1}{2}, I'; M, \pm\tfrac{1}{2}, M \pm \tfrac{1}{2})\mathcal{F}(\lambda\mu IY; \lambda'\mu' I', Y+1)\langle\lambda'\mu'||\ (1, 1)\ ||\lambda\mu\rangle,\end{aligned} \tag{V.8iii}$$

$$\begin{aligned}\langle \lambda'\mu'; I', M \pm \tfrac{1}{2}, Y-1|\ \mathcal{G}_\pm\ |\lambda\mu; IMY\rangle \\ = C(I, \tfrac{1}{2}, I'; M, \pm\tfrac{1}{2}, M \pm \tfrac{1}{2})\mathcal{G}(\lambda\mu IY; \lambda'\mu' I', Y-1)\langle\lambda'\mu'||\ (1, 1)\ ||\lambda\mu\rangle,\end{aligned} \tag{V.8iv}$$

$$\langle \lambda'\mu'; IMY|\ \mathcal{Y}\ |\lambda\mu; IMY\rangle = \mathcal{Y}(\lambda\mu IY; \lambda'\mu' IY)\langle\lambda'\mu'||\ (1, 1)\ ||\lambda\mu\rangle. \tag{V.8v}$$

In the above relations, the M dependence is carried by the CG coefficients. Our problem then is to obtain the (I, Y) dependence, that is, we have to determine the functions $\mathcal{I}$, $\mathcal{F}$, $\mathcal{G}$, and $\mathcal{Y}$; the *reduced* matrix element $\langle\lambda'\mu'||\ (1, 1)\ ||\lambda\mu\rangle$ is thereby defined by these relations.

We, thus, really have to calculate only the (I, Y) dependence of the matrix elements of the suitably chosen operators S^1_2, S^1_3, S^3_2, S^1_1, S^2_2 and S^3_3. Actually, we do not have to deal with tensor *operators*, but can equally well work with a set of *states* having the same transformation laws under SU_3 as the

operators. The (I, Y)-dependent factors that appear on the right-hand side of (V.8) are just the ones that we should use to couple states $|\lambda\mu; IMY\rangle$, and states $|11; IMY\rangle$, transforming as S^p_q, to obtain product states of type $|\lambda'\mu'; I'M'Y'\rangle$. Formally speaking, the only difference between dealing with matrix elements of tensor operators, and coupling of states is, that in the former case we always deal with both the initial and final states in the *same space*, while in the latter we have two distinct spaces to start with and the final states are defined in the *product space*. Beyond this, the former, quantities involve "reduced matrix elements" which have no analog in the latter. The important point for us is that both involve the same Wigner coupling coefficients. We use this fact in our calculations below.

We divide the computation into two cases, corresponding to the two occurrences of (λ, μ) in $(1, 1) \otimes (\lambda, \mu)$ which we denote by $(\lambda, \mu)_1$ and $(\lambda, \mu)_2$, respectively.

Case I. $(\lambda, \mu) \rightarrow (\lambda, \mu)_1$

This corresponds to (I) of (V.6). Let us define the tensor with λ upper and μ lower indices of this case by

$$R^{m_1 \cdots m_\lambda}_{q, n_2 \cdots n_\mu} = g^\alpha_\beta S^\beta_q T^{m_1 \cdots m_\lambda}_{\alpha n_2 \cdots n_\mu}, \tag{V.9}$$

where R is already symmetric in all of its upper indices, and all of its lower indices appearing after the comma. Also note that the only nonvanishing traces of R correspond to contractions of the lower index written before the comma and an upper index.

Let us rename q as n_1 and define the tensor obtained by symmetrization of R with respect to all lower indices, namely,

$$\bar{R}^{m_1 \cdots m_\lambda}_{n_1 \cdots n_\mu} = \sum_{i=1}^{\mu} R^{m_1 \cdots m_\lambda}_{n_i, n_1 \cdots n_{i-1} n_{i+1} \cdots n_\mu}. \tag{V.10}$$

The nonzero traces of $\bar{R}$ are of the form

$$g^\alpha_\beta \bar{R}^{\beta m_1 \cdots m_{j-1} m_{j+1} \cdots m_\lambda}_{\alpha n_1 \cdots n_{i-1} n_{i+1} \cdots n_\mu} = g^\alpha_\beta R^{\beta m_1 \cdots m_{j-1} m_{j+1} \cdots m_\lambda}_{\alpha, n_1 \cdots n_{i-1} n_{i+1} \cdots n_\mu)}$$
$$(i = 1, \cdots, \mu; j = 1, \cdots, \lambda). \tag{V.11}$$

This observation immediately leads us to the construction of the following tensor:

$$P^{m_1 \cdots m_\lambda}_{n_1 \cdots n_\mu} = \bar{R}^{m_1 \cdots m_\lambda}_{n_1 \cdots n_\mu} - \frac{1}{\lambda + \mu + 1} \times \sum_{i=1}^{\mu} \sum_{j=1}^{\lambda} g^{m_j}_{n_i} g^\alpha_\beta \bar{R}^{\beta m_1 \cdots m_{j-1} m_{j+1} \cdots m_\lambda}_{\alpha n_1 \cdots n_{i-1} n_{i+1} \cdots n_\mu}, \tag{V.12}$$

which is separately symmetric in its upper and lower indices and is also *traceless*. It thus provides us with a basis of our UIR $(\lambda, \mu)_1$. In terms of tensors S and T, we have

$$P^{m_1 \cdots m_\lambda}_{n_1 \cdots n_\mu} = \sum_{k=1}^{\mu} g^\beta_\alpha S^\alpha_{n_k} T^{m_1 \cdots m_\lambda}_{\beta n_1 \cdots n_{k-1} n_{k+1} \cdots n_\mu} - \frac{1}{\lambda + \mu + 1} \times \sum_{i=1}^{\mu} \sum_{j=1}^{\lambda} g^{m_j}_{n_i} g^\alpha_\gamma g^\beta_\delta S^\gamma_\beta T^{\delta m_1 \cdots m_{j-1} m_{j+1} \cdots m_\lambda}_{\alpha n_1 \cdots n_{i-1} n_{i+1} \cdots n_\mu}. \tag{V.13}$$

Case II. $(\lambda, \mu) \rightarrow (\lambda, \mu)_2$

For this case, we construct similarly the basic tensors of the UIR $(\lambda, \mu)_2$, corresponding to (II) of (V.6), namely

$$Q^{m_1 \cdots m_\lambda}_{n_1 \cdots n_\mu} = \sum_{i=1}^{\lambda} g^\alpha_\beta S^{m_i}_\alpha T^{\beta m_1 \cdots m_{i-1} m_{i+1} \cdots m_\lambda}_{n_1 \cdots n_\mu} - \frac{1}{\lambda + \mu + 1} \times \sum_{i=1}^{\lambda} \sum_{j=1}^{\mu} g^{m_i}_{n_j} g^\alpha_\gamma g^\beta_\delta S^\gamma_\beta T^{\delta m_1 \cdots m_{i-1} m_{i+1} \cdots m_\lambda}_{\alpha n_1 \cdots n_{j-1} n_{j+1} \cdots n_\mu}. \tag{V.14}$$

To illustrate the procedure, we shall give details of the calculations of the matrix elements of the operator S^1_3 for case I. Since all calculations follow the same simple pattern, we shall only quote the results of the remaining cases.

According to (III.27), the state with definite I, M, Y in the UIR $(\lambda, \mu)_1$ is given by

$$\Psi^{I,Y}_M = N_3(I, Y) \sum_{m_1, m_2} C(j_1, j_2, I; m_1, m_2, M) \times N_1(j_1 m_1 j_2 m_2) P^{(j_1+m_1, j_1-m_1, \lambda-2j_1)}_{(j_2-m_2, j_2+m_2, \mu-2j_2)}, \tag{V.15}$$

where

$$j_1 = \frac{I}{2} + \frac{Y}{4} + \frac{\lambda - \mu}{6}, \qquad j_2 = \frac{I}{2} - \frac{Y}{4} - \frac{\lambda - \mu}{6}, \qquad I = j_1 + j_2, \tag{V.16}$$

and where we have introduced the convenient *notation*

$$P^{(j_1+m_1, j_1-m_1, \lambda-2j_1)}_{(j_2-m_2, j_2+m_2, \mu-2j_2)} \tag{V.17}$$

for a component of the irreducible tensor P having 1 as the first $j_1 + m_1$ upper and $j_2 - m_2$ lower indices, 2 as the next $j_1 - m_1$ upper and $j_2 + m_2$ lower indices, and 3 as the remaining $\lambda - 2j_1$ upper and $\mu - 2j_2$ lower indices.

Since we are looking only for the (I, Y) dependence of the matrix elements, we may introduce a great simplification by choosing M equal to the maximum value I; then

$$\Psi^{I,Y}_I = N_3(I, Y) N_1(j_1, j_1, j_2, j_2) P^{(2j_1, 0, \lambda-2j_1)}_{(0, 2j_2, \mu-2j_2)}. \tag{V.18}$$

We now have only to look for the term containing S_3^1 in the right-hand side of the above equation, using the explicit expression for the component of tensor P occurring there in terms of the tensors T and S given by (V.13). This term is easily seen to be

$$N_3(I, Y)N_1(j_1j_1j_2j_2)\frac{(\mu - 2j_2)(\mu + 1 + 2j_1)}{\lambda + \mu + 1} \times S_3^1 T^{(2j_1,0,\lambda-2j_1)}_{(1,2j_2,\mu-2j_2-1)}. \quad \text{(V.19)}$$

Again using (III.28), we have in the space of the tensors T, in terms of states $\chi_M^{I,Y}$ of this space,

$$T^{(2j_1,0,\lambda-2j_1)}_{(1,2j_2,\mu-2j_2-1)} = \frac{1}{N_1(j_1, j_1, j_2 + \frac{1}{2}, j_2 - \frac{1}{2})} \times \sum_{I'} C(j_1, j_2 + \tfrac{1}{2}, I'; j_1, j_2 - \tfrac{1}{2}, I - \tfrac{1}{2}) \times \frac{N_2(j_1, j_2 + \frac{1}{2}, I')}{N_3(I', Y')} \chi^{I',Y'}_{I-\frac{1}{2}}. \quad \text{(V.20)}$$

From (V.16), and the CG coefficient appearing in the last equation, we see that

$$Y' = Y - 1, \qquad I' = I + \tfrac{1}{2}, I - \tfrac{1}{2}. \quad \text{(V.21)}$$

Thus the S_3^1 term of $\Psi_I^{I,Y}$ is, finally, given by

$$S_3^1 N_3(I, Y)\frac{N_1(j_1j_1j_2j_2)}{N_1(j_1, j_1, j_2 + \frac{1}{2}, j_2 - \frac{1}{2})}\cdot\frac{(\mu - 2j_2)(\mu + 1 + 2j_1)}{\lambda + \mu + 1} \times \Big\{C(j_1, j_2 + \tfrac{1}{2}, I + \tfrac{1}{2}; j_1, j_2 - \tfrac{1}{2}, I - \tfrac{1}{2})\cdot\frac{N_2(j_1, j_2 + \frac{1}{2}, I + \frac{1}{2})}{N_3(I + \frac{1}{2}, Y - 1)}\cdot\chi^{I+\frac{1}{2},Y-1}_{I-\frac{1}{2}} + C(j_1, j_2 + \tfrac{1}{2}, I - \tfrac{1}{2}; j_1, j_2 - \tfrac{1}{2}, I - \tfrac{1}{2})\frac{N_2(j_1, j_2 + \frac{1}{2}, I - \frac{1}{2})}{N_3(I - \frac{1}{2}, Y - 1)}\chi^{I-\frac{1}{2},Y-1}_{I-\frac{1}{2}}\Big\}. \quad \text{(V.22)}$$

Evaluating the various N factors (cf. Appendix B), using the known values of the CG coefficients, and using (V.16) to express the results in terms of I and Y, we obtain the result

$$\langle(\lambda\mu)_1; IIY|\, S_3^1\, |\lambda\mu; I + \tfrac{1}{2}, I - \tfrac{1}{2}, Y - 1\rangle = \frac{1}{\lambda + \mu + 1}\left[\frac{\lambda + 2\mu}{3} + I + \frac{Y}{2} + 1\right]\left\{\left[\frac{2\lambda + \mu}{3} + I - \frac{Y}{2} + 2\right]\left[\frac{\lambda + 2\mu}{3} - \left(I - \frac{Y}{2}\right)\right] \times \left[\frac{\mu - \lambda}{3} + I - \frac{Y}{2} + 1\right]\right\}^{\frac{1}{2}}[(2I + 1)(2I + 2)]^{-\frac{1}{2}}\langle\lambda\mu||(1, 1)||\lambda\mu\rangle_1, \quad \text{(V.23)}$$

$$\langle(\lambda\mu)_1; IIY|\, S_3^1\, |\lambda\mu; I - \tfrac{1}{2}, I - \tfrac{1}{2}, Y - 1\rangle = \frac{1}{\lambda + \mu + 1}\left[\frac{\lambda + 2\mu}{3} - \left(I - \frac{Y}{2}\right)\right]\left\{\left[\frac{\lambda + 2\mu}{3} + I + \frac{Y}{2} + 1\right]\left[\frac{2\lambda + \mu}{3} - \left(I + \frac{Y}{2}\right) + 1\right] \times \left[\frac{\lambda - \mu}{3} + I + \frac{Y}{2}\right]\right\}^{\frac{1}{2}}(2I + 1)^{-\frac{1}{2}}\langle\lambda\mu||(1, 1)||\lambda\mu\rangle_1. \quad \text{(V.24)}$$

We now incorporate the choice of phase made in (IV.14) and also write the result in the following standard form:

$$\langle(\lambda\mu)_1; I - \tfrac{1}{2}, I - \tfrac{1}{2}, Y + 1|\, S_3^1\, |\lambda\mu; I, I - 1, Y\rangle = -\frac{[\frac{1}{3}(\lambda + 2\mu) + I + \frac{1}{2}Y + 1]}{[2I(2I + 1)]^{\frac{1}{2}}(\lambda + \mu + 1)} \times \left\{\left[\frac{2\lambda + \mu}{3} + I - \frac{Y}{2} + 1\right]\left[\frac{\lambda + 2\mu}{3} - \left(I - \frac{Y}{2} - 1\right)\right]\left[\frac{\mu - \lambda}{3} + I - \frac{Y}{2}\right]\right\}^{\frac{1}{2}}\cdot\langle\lambda\mu||(1, 1)||\lambda\mu\rangle_1, \quad \text{(V.25)}$$

$$\langle(\lambda\mu)_1; I + \tfrac{1}{2}, I + \tfrac{1}{2}, Y + 1|\, S_3^1\, |\lambda\mu; I, I, Y\rangle = \frac{[\frac{1}{3}(\lambda + 2\mu) - (I - \frac{1}{2}Y)]}{(2I + 2)^{\frac{1}{2}}(\lambda + \mu + 1)} \times \left\{\left[\frac{\lambda + 2\mu}{3} + I + \frac{Y}{2} + 2\right]\left[\frac{2\lambda + \mu}{3} - \left(I + \frac{Y}{2}\right)\right]\left[\frac{\lambda - \mu}{3} + I + \frac{Y}{2} + 1\right]\right\}^{\frac{1}{2}}\langle\lambda\mu||(1, 1)||\lambda\mu\rangle_1. \quad \text{(V.26)}$$

Remembering that $-S_3^1$ is $\mathfrak{T}_+$ [see (V.3)], and comparing with (V.8iii), we finally have

$$\mathfrak{F}(\lambda\mu IY;(\lambda\mu)_1, I-\tfrac{1}{2}, Y+1) = \frac{-1}{[2I]^{\frac{1}{2}}}\cdot\frac{1}{\lambda+\mu+1}\left[\frac{\lambda+2\mu}{3}+I+\frac{Y}{2}+1\right]$$
$$\times\left\{\left[\frac{2\lambda+\mu}{3}+I-\frac{Y}{2}+1\right]\left[\frac{\lambda+2\mu}{3}-\left(I-\frac{Y}{2}\right)+1\right]\left[\frac{\mu-\lambda}{3}+I-\frac{Y}{2}\right]\right\}^{\frac{1}{2}},$$

$$\mathfrak{F}(\lambda\mu IY;(\lambda\mu)_1, I+\tfrac{1}{2}, Y+1) = \frac{-1}{[2I+2]^{\frac{1}{2}}}\cdot\frac{1}{\lambda+\mu+1}\cdot\left[\frac{\lambda+2\mu}{3}-\left(I-\frac{Y}{2}\right)\right]$$
$$\times\left\{\left[\frac{\lambda+2\mu}{3}+I+\frac{Y}{2}+2\right]\left[\frac{2\lambda+\mu}{3}-\left(I+\frac{Y}{2}\right)\right]\left[\frac{\lambda-\mu}{3}+I+\frac{Y}{2}+1\right]\right\}^{\frac{1}{2}}.$$

In this manner we obtain the results given in the following list:

$$\mathfrak{g}(\lambda\mu IY;(\lambda\mu)_1, I+1, Y) = \frac{-1}{2}\left[\frac{2I+3}{I+1}\right]^{\frac{1}{2}}\left\{1+\frac{\left[\frac{2\lambda+\mu}{3}-\left(I+\frac{Y}{2}\right)-1\right]\left[\frac{\lambda+2\mu}{3}-\left(I-\frac{Y}{2}\right)-1\right]}{(2I+3)(\lambda+\mu+1)}\right\}$$
$$\times\left\{\frac{\left[\frac{2\lambda+\mu}{3}-\left(I+\frac{Y}{2}\right)\right]\left[\frac{\lambda+2\mu}{3}-\left(I-\frac{Y}{2}\right)\right]\left[\frac{\lambda-\mu}{3}+I+\frac{Y}{2}+1\right]\left[\frac{\mu-\lambda}{3}+I-\frac{Y}{2}+1\right]}{[\frac{1}{3}(\lambda+2\mu)+I+\frac{1}{2}Y+2][\frac{1}{3}(2\lambda+\mu)+I-\frac{1}{2}Y+2]}\right\}^{\frac{1}{2}},$$

$$\mathfrak{g}(\lambda\mu IY;(\lambda\mu)_1 IY) = -\frac{1}{2}\left[\frac{I+1}{I}\right]^{\frac{1}{2}}\left[\frac{\mu-\lambda}{3}+I-\frac{Y}{2}\right]$$
$$+\frac{1}{2}\frac{1}{[I(I+1)]^{\frac{1}{2}}}\cdot\frac{1}{\lambda+\mu+1}\left[\frac{\lambda-\mu}{3}+\frac{Y}{2}\right]\left[\frac{2\lambda+\mu}{3}-\left(I+\frac{Y}{2}\right)\right]\left[\frac{\lambda+2\mu}{3}-\left(I-\frac{Y}{2}\right)\right],$$

$$\mathfrak{g}(\lambda\mu IY;(\lambda\mu)_1, I-1, Y) = \frac{1}{2[I(2I-1)]^{\frac{1}{2}}}\cdot\frac{1}{\lambda+\mu+1}\left\{\left[\frac{\lambda+2\mu}{3}+I+\frac{Y}{2}+1\right]\right.$$
$$\times\left[\frac{2\lambda+\mu}{3}+I-\frac{Y}{2}+1\right]\left[\frac{2\lambda+\mu}{3}-\left(I+\frac{Y}{2}\right)+1\right]\left[\frac{\lambda+2\mu}{3}-\left(I-\frac{Y}{2}\right)+1\right]$$
$$\left.\times\left[\frac{\lambda-\mu}{3}+I+\frac{Y}{2}\right]\left[\frac{\mu-\lambda}{3}+I-\frac{Y}{2}\right]\right\}^{\frac{1}{2}},$$

$$\mathfrak{F}(\lambda\mu IY;(\lambda\mu)_1, I+\tfrac{1}{2}, Y+1) = \frac{-1}{(2I+2)^{\frac{1}{2}}}\cdot\frac{1}{\lambda+\mu+1}\cdot\left[\frac{\lambda+2\mu}{3}-\left(I-\frac{Y}{2}\right)\right]$$
$$\times\left\{\left[\frac{\lambda+2\mu}{3}+I+\frac{Y}{2}+2\right]\left[\frac{2\lambda+\mu}{3}-\left(I+\frac{Y}{2}\right)\right]\left[\frac{\lambda-\mu}{3}+I+\frac{Y}{2}+1\right]\right\}^{\frac{1}{2}},$$

$$\mathfrak{F}(\lambda\mu IY;(\lambda\mu)_1, I-\tfrac{1}{2}, Y+1) = \frac{-1}{(2I)^{\frac{1}{2}}}\cdot\frac{1}{\lambda+\mu+1}\left[\frac{\lambda+2\mu}{3}+I+\frac{Y}{2}+1\right]$$
$$\times\left\{\left[\frac{2\lambda+\mu}{3}+I-\frac{Y}{2}+1\right]\left[\frac{\lambda+2\mu}{3}-\left(I-\frac{Y}{2}\right)+1\right]\left[\frac{\mu-\lambda}{3}+I-\frac{Y}{2}\right]\right\}^{\frac{1}{2}},$$

$$\mathfrak{G}(\lambda\mu IY;(\lambda\mu)_1, I+\tfrac{1}{2}, Y-1) = +(2I+2)^{\frac{1}{2}}\left\{1+\frac{\left[\frac{2\lambda+\mu}{3}-\left(I+\frac{Y}{2}\right)\right]\left[\frac{\lambda+2\mu}{3}-\left(I-\frac{Y}{2}\right)-1\right]}{(2I+2)(\lambda+\mu+1)}\right\}$$
$$\times\left\{\frac{\left[\frac{\mu-\lambda}{3}+I-\frac{Y}{2}+1\right]\left[\frac{\lambda+2\mu}{3}-\left(I-\frac{Y}{2}\right)\right]}{[\frac{1}{3}(2\lambda+\mu)+I-\frac{1}{2}Y+2]}\right\}^{\frac{1}{2}},$$

$$\mathfrak{G}(\lambda\mu IY;(\lambda\mu)_1, I-\tfrac{1}{2}, Y-1) = \frac{-1}{(2I)^{\frac{1}{2}}}\cdot\frac{1}{(\lambda+\mu+1)}\left[\frac{\lambda+2\mu}{3}-\left(I-\frac{Y}{2}\right)\right]$$
$$\times\left\{\left[\frac{\lambda+2\mu}{3}+I+\frac{Y}{2}+1\right]\left[\frac{2\lambda+\mu}{3}-\left(I+\frac{Y}{2}\right)+1\right]\left[\frac{\lambda-\mu}{3}+I+\frac{Y}{2}\right]\right\}^{\frac{1}{2}},$$

$$\mathcal{Y}(\lambda\mu I Y;(\lambda\mu)_1 I Y)$$
$$= -\frac{1}{2(\lambda+\mu+1)}\left\{\frac{1}{3}(3\lambda+3\mu+\lambda^2+\mu^2+\lambda\mu)+(3+4\mu+2\lambda)\frac{Y}{2}-3\left[I(I+1)-\frac{Y^2}{4}\right]\right\}. \quad \text{(V.27)}$$

For case II we find

$$\mathcal{g}(\lambda\mu I Y;(\lambda\mu)_2, I+1, Y) = -\frac{1}{2}\left[\frac{2I+3}{I+1}\right]^{\frac{1}{2}}\left\{1+\frac{\left[\frac{2\lambda+\mu}{3}-\left(I+\frac{Y}{2}\right)-1\right]\left[\frac{\lambda+2\mu}{3}-\left(I-\frac{Y}{2}\right)-1\right]}{(2I+3)(\lambda+\mu+1)}\right\}$$
$$\times\left\{\frac{\left[\frac{2\lambda+\mu}{3}-\left(I+\frac{Y}{2}\right)\right]\left[\frac{\lambda+2\mu}{3}-\left(I-\frac{Y}{2}\right)\right]\left[\frac{\lambda-\mu}{3}+I+\frac{Y}{2}+1\right]\left[\frac{\mu-\lambda}{3}+I-\frac{Y}{2}+1\right]}{[\frac{1}{3}(\lambda+2\mu)+I+\frac{1}{2}Y+2][\frac{1}{3}(2\lambda+\mu)+I-\frac{1}{2}Y+2]}\right\}^{\frac{1}{2}},$$

$$\mathcal{g}(\lambda\mu I Y;(\lambda\mu)_2 I Y) = \frac{-1}{2}\left[\frac{I+1}{I}\right]^{\frac{1}{2}}\left[\frac{\mu-\lambda}{3}-\left(I+\frac{Y}{2}\right)\right]$$
$$+\frac{1}{2}\frac{1}{[I(I+1)]^{\frac{1}{2}}}\cdot\frac{1}{\lambda+\mu+1}\left[\frac{\lambda-\mu}{3}+\frac{Y}{2}\right]\left[\frac{2\lambda+\mu}{3}-\left(I+\frac{Y}{2}\right)\right]\left[\frac{\lambda+2\mu}{3}-\left(I-\frac{Y}{2}\right)\right],$$

$$\mathcal{g}(\lambda\mu I Y;(\lambda\mu)_2, I-1, Y) = +\frac{1}{2}\cdot\frac{1}{\lambda+\mu+1}\cdot\frac{1}{[I(2I-1)]^{\frac{1}{2}}}$$
$$\times\left\{\left[\frac{\lambda+2\mu}{3}+I+\frac{Y}{2}+1\right]\left[\frac{2\lambda+\mu}{3}+I-\frac{Y}{2}+1\right]\left[\frac{2\lambda+\mu}{3}-\left(I+\frac{Y}{2}\right)+1\right]\right.$$
$$\left.\times\left[\frac{\lambda+2\mu}{3}-\left(I-\frac{Y}{2}\right)+1\right]\left[\frac{\lambda-\mu}{3}+I+\frac{Y}{2}\right]\left[\frac{\mu-\lambda}{3}+I-\frac{Y}{2}\right]\right\}^{\frac{1}{2}},$$

$$\mathcal{G}(\lambda\mu I Y;(\lambda\mu)_2, I+\tfrac{1}{2}, Y-1) = \frac{-1}{[2I+2]^{\frac{1}{2}}}\cdot\frac{1}{\lambda+\mu+1}\left[\frac{2\lambda+\mu}{3}-\left(I+\frac{Y}{2}\right)\right]$$
$$\times\left\{\left[\frac{2\lambda+\mu}{3}+I-\frac{Y}{2}+2\right]\left[\frac{\lambda+2\mu}{3}-\left(I-\frac{Y}{2}\right)\right]\left[\frac{\mu-\lambda}{3}+I-\frac{Y}{2}+1\right]\right\}^{\frac{1}{2}},$$

$$\mathcal{G}(\lambda\mu I Y;(\lambda\mu)_2, I-\tfrac{1}{2}, Y-1) = +\frac{1}{(2I)^{\frac{1}{2}}}\frac{1}{\lambda+\mu+1}\left[\frac{2\lambda+\mu}{3}+I-\frac{Y}{2}+1\right]$$
$$\times\left\{\left[\frac{2\lambda+\mu}{3}-\left(I+\frac{Y}{2}\right)+1\right]\left[\frac{\lambda+2\mu}{3}+I+\frac{Y}{2}+1\right]\left[\frac{\lambda-\mu}{3}+I+\frac{Y}{2}\right]\right\}^{\frac{1}{2}},$$

$$\mathcal{F}(\lambda\mu I Y;(\lambda\mu)_2, I+\tfrac{1}{2}, Y+1) = +(2I+2)^{\frac{1}{2}}\left\{1+\frac{\left[\frac{2\lambda+\mu}{3}-\left(I+\frac{Y}{2}\right)-1\right]\left[\frac{\lambda+2\mu}{3}-\left(I-\frac{Y}{2}\right)\right]}{(2I+2)(\lambda+\mu+1)}\right\}$$
$$\times\left\{\frac{\left[\frac{\lambda-\mu}{3}+I+\frac{Y}{2}+1\right]\left[\frac{2\lambda+\mu}{3}-\left(I+\frac{Y}{2}\right)\right]}{[\frac{1}{3}(\lambda+2\mu)+I+\frac{1}{2}Y+2]}\right\}^{\frac{1}{2}},$$

$$\mathcal{F}(\lambda\mu I Y;(\lambda\mu)_2, I-\tfrac{1}{2}, Y+1) = +\frac{1}{(2I)^{\frac{1}{2}}}\cdot\frac{1}{\lambda+\mu+1}\cdot\left[\frac{2\lambda+\mu}{3}-\left(I+\frac{Y}{2}\right)\right]$$
$$\times\left\{\left[\frac{2\lambda+\mu}{3}+\left(I-\frac{Y}{2}\right)+1\right]\left[\frac{\lambda+2\mu}{3}-\left(I-\frac{Y}{2}\right)+1\right]\left[\frac{\mu-\lambda}{3}+I-\frac{Y}{2}\right]\right\}^{\frac{1}{2}},$$

$$\mathcal{Y}(\lambda\mu I Y;(\lambda\mu)_2 I Y)$$
$$= -\frac{1}{2(\lambda+\mu+1)}\left\{\frac{1}{3}(3\lambda+3\mu+\lambda^2+\mu^2+\lambda\mu)-(3+4\lambda+2\mu)\frac{Y}{2}-3\left[I(I+1)-\frac{Y^2}{4}\right]\right\}. \quad \text{(V.28)}$$

It should be noted that a different reduced matrix element, $\langle\lambda\mu||\ (1,\ 1)\ ||\lambda\mu\rangle_2$, will appear in relations (V.8) for this case. Also, in evaluating the functions $\mathcal{Y}$, it must be remembered that we have contributions from all three terms of type S_1^1, S_2^2, S_3^3, when we expand $\Psi_I^{I,Y}$ according to (V.15).

We make the following observations concerning the above results.

(i) Excepting for over-all signs the ordered list of functions (V.27) of case I go over to the ordered list of functions (V.28) of case II, when we interchange λ and μ and replace Y by $-Y$.

(ii) If we take the differences of corresponding functions in cases I and II, we obtain the same functions that appear in the matrix elements of the generators, (IV.1), (IV.2), (IV.15). Thus.

$$\begin{aligned}
&\mathcal{F}(\lambda\mu IY;\ (\lambda\mu)_1,\ I \pm \tfrac{1}{2},\ Y+1)\\
&\quad - \mathcal{F}(\lambda\mu IY;\ (\lambda\mu)_2, I \pm \tfrac{1}{2},\ Y+1) = -f_{1,2}(\lambda\mu; IY);\\
&\mathcal{g}(\lambda\mu IY;\ (\lambda\mu)_1,\ I \pm 1,\ Y)\\
&\quad - \mathcal{g}(\lambda\mu IY;\ (\lambda\mu)_2,\ I \pm 1,\ Y) = 0;\\
&\mathcal{g}(\lambda\mu IY;\ (\lambda\mu)_1 IY)\\
&\quad - \mathcal{g}(\lambda\mu IY;\ (\lambda\mu)_2 IY) = -[I(I+1)]^{\frac{1}{2}};\\
&\mathcal{Y}(\lambda\mu IY;\ (\lambda\mu)_1 IY)\\
&\quad - \mathcal{Y}(\lambda\mu IY;\ (\lambda\mu)_2 IY) = -\tfrac{3}{2}Y. \qquad \text{(V.29)}
\end{aligned}$$

Thus this particular combination of our expressions corresponds to using the matrix elements of the generators as CG coefficients. This is the canonical way pointed out by Biedenharn.[4] Put somewhat differently, the coupling scheme described by

$$S_q^p T_{n_1}^{m_1}{}^{\cdots} - S_q^p T_{n_2}^{m_1}{}^{\cdots}$$

is the same as Biedenharn's uniquely defined coupling scheme using the matrix elements of the generators themselves. We note that the two independent schemes used here are in the tensor framework, "canonical."

At this point it is worth noting that in (V.29), Y appears multiplied by $(-\frac{3}{2})$ in the right-hand side of the last equation, whereas in the first equation $-f_{1,2}$ appear without any such factors. The reason for this is, that in defining an octet operator so that it transforms exactly like an octet of normalized states, we must take $-3 \times 6^{-\frac{1}{2}}\mathcal{Y}$ with $\mathcal{F}_\pm$, and not $\mathcal{Y}$ with $\mathcal{F}_\pm$ [cf. (B.11)]. This is in analogy with (V.7). Thus with f_1, f_2 we must take $-3 \times 6^{-\frac{1}{2}}Y$ in defining a Wigner coefficient up to an *over-all* normalization. Also in the expansion (V.15) of $\Psi_I^{I,Y}$, we must remember that $-3 \times 6^{-\frac{1}{2}}S_3^3$, and not S_3^3, is the normalized state.

(iii) The two expressions for $\mathcal{Y}$ in (V.27) and (V.28) have the forms

$$\begin{aligned}
&\mathcal{Y}_{1,2} = a[\tfrac{1}{4}Y^2 - I(I+1) + \mathcal{C}(\lambda,\mu)] + b_{1,2}Y,\\
&\mathcal{C} = \tfrac{1}{9}(\lambda^2 + \mu^2 + \lambda\mu + 3\lambda + 3\mu),
\end{aligned} \qquad \text{(V.30)}$$

where a, b_1, b_2 are functions of λ and μ. Since, according to the above considerations, two unknown parameters enter (as two independent reduced matrix elements) in the general matrix element

$$\langle\lambda\mu;\ IMY|\ S_3^3\ |\lambda\mu;\ IMY\rangle,$$

it is clear that the latter will have the form of the well known two-parameter *mass formula* of Okubo.[19]

VI. WEYL REFLECTIONS AND *U*-SPIN

In discussing electromagnetic processes in the context of SU_3 invariance for the strong interactions, it is useful to employ the language of U spin, introduced by Levinson, *et al.*[22] The electromagnetic interactions conserve U spin, and violate I-spin conservation. Macfarlane, Sudarshan, and Dullemond[23] have used the method of Weyl reflections, one of which takes I spin into U spin. We use the approach of the latter authors in deriving an explicit and general connection between I-spin and U-spin states.

It is convenient in this section to dinstinguish between operators and eigenvalues by using carets on the former. The U-spin subgroup of SU_3 is an SU_2 subgroup generated by $(6^{\frac{1}{2}}\hat{E}_{\pm 3},\ \frac{3}{4}\hat{Y} - \frac{1}{2}\hat{I}_z)$. All states belonging to a U-spin multiplet, have the same eigenvalue for the electric charge operator $\hat{Q}$ where, as usual,

$$\hat{Q} = \hat{I}_z + \tfrac{1}{2}\hat{Y}. \qquad \text{(VI.1)}$$

The operator $-\hat{Q}$ bears the same relationship to the U-spin subgroup of SU_3, as does hypercharge $\hat{Y}$ to the isospin subgroup.

The Weyl reflections form a finite subgroup of SU_3, and have the same structure as S_3, the permutation group on three objects. Following Okubo,[24] we concentrate on the operator W_{123} (W, for short) which cyclically permutes the states 1, 2, 3 in the UIR (1, 0). Using the variables x^m of Sec. I, which transform according to (1, 0), W is defined as follows.

$$W\colon x^1 \to x^2, \qquad x^2 \to x^3, \qquad x^3 \to x^1. \qquad \text{(VI.2)}$$

[22] C. A. Levinson *et al.*, Ref. 12.
[23] A. J. Macfarlane *et al.*, Ref. 12.
[24] S. Okubo, University of Rochester lectures (unpublished) (1964).

If we write W as a 3×3 matrix, it is unitary unimodular, so that $W \in SU_3$. Further, from (VI.2) above and (I.3) we have

$$W\hat{E}_{\pm 1}W^\dagger = \hat{E}_{\pm 3}, \quad W\hat{I}_z W^\dagger = \tfrac{3}{4}\hat{Y} - \tfrac{1}{2}\hat{I}_z \equiv \hat{U}_3,$$

$$W\hat{Y}W^\dagger = -\hat{I}_z - \frac{\hat{Y}}{2} = -\hat{Q}, \qquad \text{(VI.3)}$$

$$\therefore \quad W(\hat{I}^2, \hat{I}_z, \hat{Y})W^\dagger = (\hat{U}^2, \hat{U}_3, -\hat{Q}).$$

To derive the form of W in (0, 1) we note that (VI.3) must hold for the generators in any UIR. Since these are known for (0, 1) in (I.4), and since W must be unitary unimodular, we derive:

$$W: y_1 \to -y_2, \qquad y_2 \to y_3, \qquad y_3 \to -y_1. \qquad \text{(VI.4)}$$

Equations (VI.3)–(VI.4) tell us how W acts on upper and lower indices respectively in any tensor T. Our purpose is to derive the effect of W in any UIR (λ, μ).

The orthonormal basis $\Phi_M^{I,Y}$ introduced in the UIR (λ, μ) in Secs. III and IV is labeled by the quantum numbers (I, M, Y), which are the eigenvalues for the diagonal operators $(\hat{I}^2, \hat{I}_z, \hat{Y})$. An alternative basis in this same space is $\chi_{U_3}^{U,-Q}$ with the operators $(\hat{U}^2, \hat{U}_3, -\hat{Q})$ diagonal and having eigenvalues $(U, U_3, -Q)$. Since the two sets of operators are unitarily related by W, (VI.3), the ranges of eigenvalues of the two sets are exactly the same; and the states $\chi_{U_3}^{U,-Q}$, with respect to which $(\hat{U}^2, \hat{U}_3, -\hat{Q})$ are diagonal, are gotten by applying the unitary operator W to a state $\Phi_{U_3}^{U,-Q}$ with the *eigenvalues* $(U, U_3, -Q)$ for the *operators* $\hat{I}^2, \hat{I}_z, \hat{Y})$. Thus, using the notation (V.17) for components of an irreducible tensor T, we develop:

$$\chi_{U_3}^{U,-Q} = W\Phi_{U_3}^{U,-Q} = (-1)^{U+\frac{1}{2}Q-\frac{1}{3}(\lambda+2\mu)}N_3(U,-Q)\sum_{m_1 m_2} C(j_1 j_2 U; m_1 m_2 U_3)N_1(j_1 m_1; j_2 m_2)WT^{(j_1+m_1, j_1-m_1, \lambda-2j_1)}_{(j_2-m_2, j_2+m_2, \mu-2j_2)};$$

$$j_1 = \frac{U}{2} - \frac{Q}{4} + \frac{\lambda-\mu}{6}, \qquad j_2 = \frac{U}{2} + \frac{Q}{4} - \frac{\lambda-\mu}{6}. \qquad \text{(VI.5)}$$

By means of (VI.2) and (VI.4):

$$\chi_{U_3}^{U,-Q} = (-1)^{U+\frac{1}{2}Q-\frac{1}{3}(\lambda+2\mu)}N_3(U,-Q)\sum_{m_1 m_2} C(j_1 j_2 U; m_1 m_2 U_3)N_1(j_1 m_1; j_2 m_2)(-1)^{\mu-j_2-m_2}T^{(\lambda-2j_1, j_1+m_1, j_1-m_1)}_{(\mu-2j_2, j_2-m_2, j_2+m_2)}. \qquad \text{(VI.6)}$$

Re-express T in (VI.6) in terms of Φ using (III.28) and (IV.14):

$$\begin{aligned}\chi_{U_3}^{U,-Q} = {} & (-1)^{U+\frac{1}{2}Q-\frac{1}{3}(\lambda+2\mu)}N_3(U,-Q)\sum_{m_1 m_2}(-1)^{\mu-j_2-m_2}C(j_1 j_2 U; m_1 m_2 U_3) \\ & \times N_1(j_1 m_1, j_2 m_2)\left[N_1\left(\frac{\lambda-j_1+m_1}{2}, \frac{\lambda-3j_1-m_1}{2}; \frac{\mu-j_2-m_2}{2}, \frac{3j_2-\mu-m_2}{2}\right)\right]^{-1} \\ & \times \sum_I C\left(\frac{\lambda-j_1+m_1}{2}, \frac{\mu-j_2-m_2}{2}, I; \frac{\lambda-3j_1-m_1}{2}, \frac{3j_2-\mu-m_2}{2}, M\right) \\ & \times N_2\left(\frac{\lambda-j_1+m_1}{2}, \frac{\mu-j_2-m_2}{2}, I\right)[N_3(I,Y)]^{-1}(-1)^{-I+\frac{1}{2}Y+\frac{1}{3}(\lambda+2\mu)}\Phi_M^{I,Y};\end{aligned}$$

$$M = \tfrac{3}{4}Q - \tfrac{1}{2}U_3, \qquad Y = U_3 + \tfrac{1}{2}Q. \qquad \text{(VI.7)}$$

This rather forbidding expression can be somewhat simplified: the factors N_1 are identical and cancel out; while the factors N_2/N_3 can be rewritten as a CG coefficient. We then have:

$$\begin{aligned}\chi_{U_3}^{U,-Q} = {} & N_3(U,-Q)\sum_{m_1 m_2 I}(-1)^{\mu-j_2-m_2+U-I+\frac{1}{2}(Q+Y)}C(j_1 j_2 U; m_1 m_2 U_3) \\ & \times C\left(\frac{\lambda-j_1+m_1}{2}, \frac{\mu-j_2-m_2}{2}, I; \frac{\lambda-3j_1-m_1}{2}, \frac{3j_2-\mu-m_2}{2}, M\right) \\ & \times C\left(\frac{2\lambda+\mu}{6}+\frac{I+1}{2}-\frac{Y}{4}, \frac{\lambda+2\mu}{6}-\frac{I}{2}+\frac{Y}{4}, \frac{\lambda+\mu+1}{2}; \frac{\lambda+2\mu}{6}+\frac{I-U+1}{2}\right. \\ & \left. +\frac{Y}{4}+m_1-m_2, -\frac{2\lambda+\mu}{6}+\frac{I+U}{2}+\frac{Y}{4}+m_2-m_1, \frac{\mu-\lambda}{6}+I+\frac{Y}{2}+\frac{1}{2}\right)\Phi_M^{I,Y}. \qquad \text{(VI.8)}\end{aligned}$$

[It may be noted that the last CG coefficient in (VI.8) is numerically unaltered on replacing I by $-I-1$.]

Equation (VI.8) gives explicitly the unitary transformation that takes us from the (I, Y) basis to the (U, Q) basis. It may be noted that as W takes the I-spin SU_2 subgroup into the U-spin SU_2 subgroup with no phase changes of any kind, the U-spin wavefunctions χ of (VI.8) obey the standard Condon–Shortley convention as far as U spin is concerned.

The Weyl operator W used by us has the property

$$W^3 = 1. \tag{VI.9}$$

Macfarlane, Sudarshan, and Dullemond[23] use an operator W^1 with the property

$$(W^1)^2 = 1, \tag{VI.10}$$

to relate U spin to I spin. W^1 is defined in the UIR (1, 0) as interchanging states 1 and 3, leaving 2 alone. It turns our that the U-spin eigenstates obtained by these authors differ from ours, essentially, only in the sign of U_3 and in the interchange of U-spin raising and lowering operators.

Finally, we make a few comments concerning the *triangular representations*, i.e., UIR's of type $(\lambda, 0)$ or $(0, \mu)$. We shall look at $(\lambda, 0)$. The corresponding irreducible tensors

$$T^{m_1 m_2 \cdots m_\lambda}$$

have only upper indices in which they are symmetric; and there is no problem of tracelessness. Now in the UIR (1, 0), the states 1 and 2 form an isodoublet while 3 is an isosinglet. Similarly, 2 and 3 form an U-spin doublet, 1 is a U singlet. It is then obvious that in such a representation, any component $T^{i_1 m_2}$ of T simultaneously possesses a definite isospin and a definite U spin:

$$I = j_1, \qquad U = \tfrac{1}{2}(\lambda - j_1 - m_1). \tag{VI.11}$$

In other words, in such representations the operators $\hat{I}^2$, $\hat{I}_z$, $\hat{Y}$, $\hat{U}^2$, $\hat{U}_3$, $\hat{Q}$ are all simultaneously diagonal. We can also easily establish the equations:

$$\begin{aligned} Y &= \tfrac{1}{3}(2I) - \tfrac{2}{3}(\lambda - 2I) = 2I - \tfrac{2}{3}\lambda, \\ Q &= -\tfrac{1}{3}(2U) + \tfrac{2}{3}(\lambda - 2U) = -2U + \tfrac{2}{3}\lambda, \end{aligned} \tag{VI.12}$$

by essentially counting indices in T. The above remarks are meant to illustrate the fact that these simple properties of triangular representations[25] are trivially obvious when we think of these representations in terms of tensors.

[25] S. P. Rosen, J. Math. Phys. 5, 289 (1964).

ACKNOWLEDGMENTS

We would like to express our thanks to Professors D. Lurié, A. J. Macfarlane, S. Okubo, and E. C. G. Sudarshan for providing us with much stimulation and encouragement. One of us (LKP) wishes to thank the U. S. Educational Foundation in India for a Fulbright Travel Grant.

APPENDIX A

This appendix contains the calculation leading to the factor N_3 appearing in (III.26).

According to (III.9) the invariant $\mathfrak{M}$ has the form

$$\mathfrak{M} = \sum_{j_1 j_2} \sum_{IM} |\psi_M^{I(j_1 j_2)}|^2. \tag{A1}$$

Considering all terms in (A1) corresponding to fixed I, Y, and M, there is a pair of minimum values of j_1 and j_2 compatible with the given Y that can lead to the given I; from (III.10), namely,

$$\begin{aligned} j_1^{\min} &= \tfrac{1}{2}I + \tfrac{1}{4}Y + \tfrac{1}{6}(\lambda - \mu), \\ j_2^{\min} &= \tfrac{1}{2}I - \tfrac{1}{4}Y - \tfrac{1}{6}(\lambda - \mu). \end{aligned} \tag{A2}$$

If we consider the pairs of values

$$\begin{aligned} &(j_1, j_2) = (j_1^{\min}, j_2^{\min}), (j_1^{\min} + \tfrac{1}{2}, j_2^{\min} + \tfrac{1}{2}), \\ &(j_1^{\min} + 1, j_2^{\min} + 1), \cdots, (\tfrac{1}{2}\lambda, \tfrac{1}{2}\lambda + j_2^{\min} - j_1^{\min}), \\ &\qquad \text{or} \quad (\tfrac{1}{2}\mu + j_1^{\min} - j_2^{\min}, \tfrac{1}{2}\mu), \end{aligned} \tag{A3}$$

then each pair has the same Y, and these are all the pairs that will lead to the given I in (A1). However, the subsidiary condition (III.22) allows us to express all the terms $\psi_M^{I(j_1 j_2)}$ in (A1) with $(j_1 j_2)$ belonging to the set (A3) as multiples of the "lowest" or basic term $\psi_M^{I(j_1{}^{\min}, j_2{}^{\min})}$. In this way, the coefficient of $|\psi_M^{I(j_1{}^{\min} j_2{}^{\min})}|^2$ in (A1) is

$$[N_3(I, Y)]^2 = \sum_{j_1 j_2} |N_2(j_1 j_2 I)|^2, \quad j_1 j_2 \text{ as in (A3)}. \tag{A4}$$

The upper limit to the summation in (A4) is $j_1 = \frac{1}{2}\lambda$ or $j_2 = \frac{1}{2}\mu$, whichever occurs earlier in (A3). Thus we have two cases to consider:

Case I—upper limit $j_1 = \frac{1}{2}\lambda$. In this case we have

$$j_2^{\min} + \tfrac{1}{2}\lambda - j_1^{\min} \leq \tfrac{1}{2}\mu, \tag{A5}$$

i.e.,

$$Y \geq \tfrac{1}{3}(\lambda - \mu);$$

substituting for j_2 in terms of j_1 and Y,

$$j_2 = j_1 - \tfrac{1}{2}Y - \tfrac{1}{3}(\lambda - \mu), \tag{A6}$$

and taking N_2 from (III.23), we have

$$[N_3(IY)]^2$$

$$= \sum_{j_1=j_{1\min}}^{\lambda/2} \frac{(2I+1)![\frac{1}{3}(2\lambda+\mu)-I-\frac{1}{2}Y]![\frac{1}{3}(\lambda+2\mu)-I+\frac{1}{2}Y]!}{[2j_1-I-\frac{1}{2}Y-\frac{1}{3}(\lambda-\mu)]!(\lambda-2j_1)![2j_1+I+1-\frac{1}{2}Y-\frac{1}{3}(\lambda-\mu)]![\mu+Y+\frac{2}{3}(\lambda-\mu)-2j_1]!}. \tag{A7}$$

Writing $t = 2j_1$, we must compute

$$F = \sum_{t=t_0}^{\lambda} \left[\left(t-I-\frac{Y}{2}-\frac{\lambda-\mu}{3}\right)!(\lambda-t)!\left(t+I+1-\frac{Y}{2}-\frac{\lambda-\mu}{3}\right)!(\mu+Y+\tfrac{2}{3}(\lambda-\mu)-t)!\right]^{-1},$$

$$t_0 = I+\tfrac{1}{2}Y+\tfrac{1}{3}(\lambda-\mu). \tag{A8}$$

In (A8), t increases in steps of unity, and the range of t is all integers that makes the argument of each factorial nonnegative. If we call the factors in (A8) as $p!\,q!\,r!\,s!$ we find the following rules:

$$\begin{aligned} p+q &= \tfrac{1}{3}(2\lambda+\mu)-I-\tfrac{1}{2}Y, \qquad p+s = \tfrac{1}{3}(\lambda+2\mu)-I+\tfrac{1}{2}Y, \\ q+r &= \tfrac{1}{3}(2\lambda+\mu)+I+1-\tfrac{1}{2}Y, \qquad r+s = \tfrac{1}{3}(\lambda+2\mu)+I+1+\tfrac{1}{2}Y, \\ p+q+r+s &= \lambda+\mu+1. \end{aligned} \tag{A9}$$

The last sum rule in (A9) suggests that we look at the multinomial expansion

$$(a+b+c+d)^{\lambda+\mu+1} = \sum_{p+q+r+s=\lambda+\mu+1} \frac{(\lambda+\mu+1)!a^pb^qc^rd^s}{p!q!r!s!}. \tag{A10}$$

The remaining sum rules in (A9) suggest that we take

$$a = xy, \qquad b = xz, \qquad c = wz, \qquad d = wy$$

and evaluate F by looking for the coefficient of

$$\xi = x^{\frac{1}{3}(2\lambda+\mu)-I-\frac{1}{2}Y}y^{\frac{1}{3}(\lambda+2\mu)-I+\frac{1}{2}Y}z^{\frac{1}{3}(2\lambda+\mu)+I+1-\frac{1}{2}Y}w^{\frac{1}{3}(\lambda+2\mu)+I+1+\frac{1}{2}Y}$$

in (A10). However $a+b+c+d = (x+w)(y+z)$ and we get

$$F = \frac{1}{[(\lambda+\mu+1)!]} \text{ coefficient of } \xi \text{ in } [(x+w)(y+z)]^{\lambda+\mu+1}$$

$$= \frac{(\lambda+\mu+1)!}{\left(\frac{2\lambda+\mu}{3}-I-\frac{Y}{2}\right)!\left(\frac{\lambda+2\mu}{3}-I+\frac{Y}{2}\right)!\left(\frac{2\lambda+\mu}{3}+I+1-\frac{Y}{2}\right)!\left(\frac{\lambda+2\mu}{3}+I+1+\frac{Y}{2}\right)!}. \tag{A11}$$

Using (A11) for F in (A7), we obtain N_3 for case I. This agrees with (III.26). However, this expression for N_3 is invariant under the replacement $\lambda \leftrightarrow \mu$, $Y \rightarrow -Y$, which is just what distinguishes the cases I and II. Therefore our solution for N_3 holds for both the cases.

APPENDIX B

For ready reference, we list below the main results of Sec. III. The *normalized basic state* with proper choice of phase in the UIR (λ, μ) is

$$|\lambda\mu; IMY\rangle = \Phi_M^{I,Y} = (-1)^{I-\frac{1}{2}Y-\frac{1}{3}(\lambda+2\mu)}\Psi_M^{I,Y}, \tag{B1}$$

$$\Psi_M^{I,Y} = N_3(I,Y) \sum_{m_1,m_2} C(j_1, j_2, I; m_1, m_2, M) \times N_1(j_1m_1j_2m_2)T^{(j_1m_1)}_{(j_2m_2)}, \tag{B2}$$

where

$$T^{(j_1m_1)}_{(j_2m_2)} \equiv T^{(j_1+m_1,\, j_1-m_1,\, \lambda-2j_1)}_{(j_2-m_2,\, j_2+m_2,\, \mu-2j_2)}, \tag{B3}$$

and

$$\begin{aligned} j_1 &\equiv \tfrac{1}{2}I+\tfrac{1}{4}Y+\tfrac{1}{6}(\lambda-\mu), \\ j_2 &\equiv \tfrac{1}{2}I-\tfrac{1}{4}Y-\tfrac{1}{6}(\lambda-\mu). \end{aligned} \tag{B4}$$

The inverse to (B2) reads

$$T^{(j_1m_1)}_{(j_2m_2)} = [N_1(j_1m_1j_2m_2)]^{-1} \sum_I C(j_1j_2I; m_1m_2M) \times N_2(j_1j_2I)[N_3(I,Y)]^{-1}\Psi_M^{I,Y}, \tag{B5}$$

where

$$Y = 2(j_1-j_2)-\tfrac{2}{3}(\lambda-\mu). \tag{B6}$$

The N-factors:

$$N_1(j_1m_1j_2m_2) = [(\lambda!\mu!)]^{\frac{1}{2}}[(j_1 + m_1)!(j_1 - m_1)!(\lambda - 2j_1)!(j_2 + m_2)!(j_2 - m_2)!(\mu - 2j_2)!]^{-\frac{1}{2}}, \tag{B7}$$

$$N_2(j_1j_2I) = [(2I + 1)!(\lambda - j_1 + j_2 - I)!(\mu - j_2 + j_1 - I)!]^{\frac{1}{2}} \times [(j_1 + j_2 - I)!(j_1 + j_2 + I + 1)!(\lambda - 2j_1)!(\mu - 2j_2)!]^{-\frac{1}{2}}, \tag{B8}$$

$$N_3(I, Y) = [(2I + 1)!(\lambda + \mu + 1)!]^{\frac{1}{2}}\left[\left(\frac{\lambda + 2\mu}{3} + I + \frac{Y}{2} + 1\right)!\left(\frac{2\lambda + \mu}{3} + I - \frac{Y}{2} + 1\right)!\right]^{-\frac{1}{2}}. \tag{B9}$$

Standard octet operator: In analogy with the standard form of the components

$$(-g_+/\sqrt{2},\ g_z,\ g_-/\sqrt{2}) \tag{B10}$$

of the spherical vector operator, the components of the octet operator of SU_3 should be arranged as:

$$(\mathfrak{F}_+,\ \mathfrak{F}_-;\ -g_+,\ \sqrt{2}g_z,\ g_-;\ -3\mathfrak{Y}/6^{\frac{1}{2}};\ \mathfrak{G}_+,\ \mathfrak{G}_-) \tag{B11}$$

so as to transform exactly like a normalized octet of states obeying our final choice of phases.

Multipole Matrix Elements of the Translation Operator

M. DANOS AND L. C. MAXIMON
National Bureau of Standards, Washington, D. C.
(Received 16 June 1964)

Formulas are given for the expansion of multipole fields of arbitrary tensorial character into multipole fields about a shifted origin. The expansion coefficients are given as matrix elements of the translation operator. In analogy to the matrix elements of the rotation operator, we introduce for these matrix elements a standard form which represents a parallel displacement of the coordinate system along the z axis. Any arbitrary translation of the coordinate system then consists of a consecutive application of a rotation, a standard translation, and a rotation. Since the multipole fields form a complete set any arbitrary function can in principle be expressed in a shifted coordinate system by means of the given formulas. All mathematical derivations are given.

I. INTRODUCTION

An elaborate theory has evolved about the angular momentum conservation law.[1-11] The reason for this is the prevalence of central potentials in problems of interest in quantum mechanics. Thus, for example, all electrons move under the influence of a common central potential, while the mutual interaction between electrons, which is not spherically symmetric about the center of mass of the atom and leads to exchange of angular momentum between the individual electrons, can be treated as a perturbation.

The conservation of linear momentum, on the other hand, has not led to a comparable development of mathematical tools. The transformation to center-of-mass coordinates in general exhausts the translational symmetries of the problem: a spherically symmetric potential remains invariant under rotation of the coordinate system, while it changes form when the origin of the coordinate system is shifted away from, say, the location of the nucleus in an atom. Nevertheless, there exists a large family of problems for which the expansion of wavefunctions or wave fields around some point other than the fixed origin or center of mass, is needed. For example, in the quasi-deuteron effect, a proton and neutron colliding within a nucleus absorb a photon and are as a result emitted as fast particles. This process is an electric dipole transition in the center-of-mass system of the quasi-deuteron, the system consisting of the involved proton and neutron, but it is of arbitrary multipolarity in the center-of-mass system of the whole nucleus. Similarly, the photoproduction of π mesons on nucleons is mostly a magnetic dipole transition in the π meson–nucleon system and, again, it is of arbitrary multipolarity in the center-of-mass system of the whole nucleus. Another set of problems is exemplified by the α decay of heavy nuclei where one has to find the probability for an α particle to be formed at the surface of the nucleus. Here one has to expand the nuclear wavefunction, usually approximated by a harmonic oscillator wavefunction centered at the center of mass of the nucleus, into other wavefunctions describing the α particle, centered at a point at the nuclear surface. In a sense, solid-state physics can also be considered to belong to this family of problems. One could even call solid-state physics "The problem of the shifted origins" *par excellence:* a wave traveling through a crystal is scattered by an atom producing outgoing spherical multipole waves which in turn are scattered. When considering the rescattering of these waves by some other atom one can expand them into multipole waves about that other atom. The first two examples involve the expansion of a free field in a shifted coordinate system, while the last two involve more general functions. Naturally, the latter can be expanded in terms of a complete set of free-field solutions.

[1] E. P. Wigner, *Group Theory and Its Application to the Quantum Mechanics of Atomic Spectra* (Academic Press Inc., New York, 1959).
[2] H. Weyl, *The Theory of Groups and Quantum Mechanics* (Dover Publications, Inc., New York, 1931).
[3] C. Eckart, Rev. Mod. Phys. **2**, 305 (1930).
[4] H. B. G. Casimir, "Rotation of a Rigid Body in Quantum Mechanics," thesis, Leyden, 1931.
[5] B. L. van der Waerden, *Die Gruppentheoretische Methode in der Quantenmechanik* (Julius Springer-Verlag, Berlin, 1932).
[6] U. Fano and G. Racah, *Irreducible Tensorial Sets* (Academic Press Inc., New York, 1959).
[7] A. R. Edmonds, *Angular Momentum in Quantum Mechanics* (Princeton University Press, Princeton, New Jersey, 1957).
[8] M. E. Rose, *Elementary Theory of Angular Momentum* (John Wiley & Sons, Inc., New York, 1957).
[9] P. H. E. Meijer and E. Bauer, *Group Theory; The Application to Quantum Mechanics* (Interscience Publishers, Inc., New York, 1962).
[10] D. M. Brink and G. R. Satchler, *Angular Momentum* (Oxford University Press, Oxford, England, 1962).
[11] A. P. Yutsis, I. B. Levinson, and V. V. Vanagas, *Mathematical Apparatus of the Theory of Angular Momentum* (Translated from Russian) (Published for National Science Foundation, Washington, D. C., 1962).

These examples have the following feature in common: a function consisting of one or of few multipoles (angular momentum eigenstates) has to be expanded about a shifted origin. The "natural" mathematical technique for performing a displacement, the Fourier transform in Cartesian coordinates, is, however, inappropriate since it mixes all multipoles. A more direct method which preserves the multipole character seems called for.

In the present paper we derive formulas for the displacement of free multipole fields. The term "multipole fields" is used here and throughout this paper to denote solutions of the wave equation $(\nabla^2 + k^2)\,\psi = 0$ which are eigenstates of the total angular momentum operator. The fields may have scalar, spinor, vector, or any other tensorial character. Calling the "orbital" angular momentum L, the "tensorial" angular momentum S, and the total angular momentum $J = L + S$, a multipole field is thus specified by the quantum numbers J, L, S, J_z, and k. For a scalar field $S = 0$ and, following the general usage, we then write l, m, and k for the quantum numbers.

Many of our results are contained in the literature. They are, however, not easily available. As a matter of fact, we found the references only after having rederived most of them. The characteristic of being hidden was re-emphasized in that an additional genealogic tree of references was called to our attention[12] after having submitted the paper to the Editor. It is very likely that still further references exist in the literature.

In addition to reporting some results which we still have not found in the literature, we have two purposes in mind in presenting this paper: one concerns the point of view with which we propose to look at the problem; the other concerns the compactness of the form in which the results are presented and the simplification of the derivations. Beginning with the first, we would like to focus on the representation of the translation by an operator. We thus introduce a "standard translation matrix," $T(k\rho)$, Eq. (4) below, which we choose to describe a translation of the coordinate system along the z axis by a distance ρ. Any general shift of the coordinate system is then given by $D(R_2)T(k\rho)D(R_1)$ where D is the usual rotation operator: one first rotates the coordinate system so that the z axis points in the direction of the displacement (rotation R_1), one then shifts along the z axis, and lastly rotates into the final position (rotation R_2).

[12] Referee (anonymous), who indicated Refs. 17, 18, and 32.

The previous treatments have dealt instead with the parallel displacement which is a special case of the general displacement (parallel displacement plus a rotation) in that the second rotation, R_2, is the inverse of the first rotation, R_1. We would like to denote the parallel displacement operator by S, which is thus defined as $S = D(R_1^{-1})TD(R_1)$; see Eq. (19) below. This is evidently a more complicated quantity than T itself. However, we use the operator S where it is advantageous to do so. As a matter of fact, the derivations of the matrix elements of S are not more involved than those for the matrix elements of T. The reason for this is that S explicitly manifests its angular momentum properties [in the $Y_{\lambda\mu}(\eta, \psi)$ in (19′)], while they are somewhat obscured in T. In fact, we derive the matrix elements of T from those of S by the specialization to vanishing rotation angles, which then results in a translation along the z axis. The reasons for choosing the z axis in the definition of the "standard T" are almost self-evident; after all, it is the usual quantization axis for the spherical functions.

The T's form a representation of a Lie-group.[1,2] We are not concerned with this aspect here.

Concerning the second point, the achieved simplifications are demonstrated by the fact that all proofs are contained in this paper. As a matter of fact, most of the mathematical difficulties are associated with the case of incoming and outgoing spherical waves, where the radial part of the multipole fields is given by the spherical Hankel functions $h_l^{(\alpha)}(kr)$, $\alpha = 1, 2$. The case of standing waves, which contain the spherical Bessel functions $j_l(kr)$, is treated in a few lines [Eqs. (10)–(14)]. The results are also very simple and compact when written in terms of 3-j and 6-j coefficients. It is both gratifying and desirable to have the expressions in this form; the latter because of the many known characteristics of these coefficients which could come in handy in manipulations one may have to perform in applications.

Another large family of problems, which do not involve explicitly a translation operator, is very closely related with the subject of the present paper and the formulas given here can easily be adapted to be applicable to them also. The characteristic feature of these problems is the requirement to expand a function of a vector $\mathbf{r}$ in terms of the vectors $\boldsymbol{\rho}_1$ and $\boldsymbol{\rho}_2$ which fulfill the relation $\mathbf{r} = \boldsymbol{\rho}_2 - \boldsymbol{\rho}_1$. After the expansion of the function of $\mathbf{r}$ into multipoles one returns to the problem of this paper when one formally introduces $\boldsymbol{\rho}_1 = -\mathbf{R}$ and considers $\boldsymbol{\rho}_2$

to be the displacement vector [see Eq. (10) below]. The point of view is just somewhat different in that in the original problem ϱ_1 and ϱ_2 have the same character, in contrast to the concept of the translation operator where the "shift vector" and the "radius vector" have a different meaning. Although every function of two vectors can be expanded into the complete set of products of two multipole fields, this expansion will in general have very little in common with the translation operator. It is the "geometry" $\mathbf{r} = \varrho_2 - \varrho_1$ linking the three involved vectors which establishes this connection. The simplest example of a problem of this kind is the Coulomb interaction between two electrons in an atom which, being spherically symmetric, involves just the expansion of a scalar monopole field. In the nuclear shell model, the force between two nucleons can also be expanded in a similar series. However, because of the spin dependence of the nuclear forces the field has not only scalar but also tensor character, i.e., $S \neq 0$ in the notation of Eq. (35). It involves of necessity, therefore, components with $l \neq 0$ in order to be able to couple to $J = 0$ [see Eq. (35)] which is necessary since the potential energy must be a scalar. Also, here not only the limit $k = 0$ is needed, since the force does not have a $1/r$ form; rather a Fourier–Bessel transform over k has to be performed. Another very important case in which the expansion of higher multipole fields appears is in the Green's function of the wave equation, $G(\varrho_1, \varrho_2)$, in which the singular part again is only a function of $\mathbf{r}$.

Although the problems of this family do not conform to our point of view in that they do not contain the translation operator, we give the relevant expansion formulas for completeness.

In Sec. II we give a short summary of the literature on this subject as far as we are aware of it. In Sec. III we define the basic quantities, i.e., the relevant coordinate systems, the operators, and the matrix elements, and we derive the basic formulas for the translation of standing scalar wave fields of arbitrary multipolarity. In Sec. IV we discuss the extension to the case of incoming and outgoing waves. This part requires the heaviest mathematics. Some of the mathematical points concerning this section are treated in Appendices A and B. Finally, in Sec. V we derive the translation formulas for multipole fields of arbitrary tensorial character.

II. HISTORICAL SURVEY

In this section we review the historical development of the mathematical treatment of the translation transformation. We do not, however, concern ourselves at all with reviewing those many papers whose contribution lies in the application of the mathematical tools discussed in this paper but contain no new development of the mathematics.

The first translation transformation is due to Legendre[13] whose series $1/r = \sum (\rho^L / R^{L+1}) P_L(\cos\Theta)$ being the expansion of the monopole potential $1/r$ into multipoles around the point $z = -\rho$ in the old coordinate system, is nothing but the lowest term in the expansion in powers of k of the monopole field $h_0^{(\alpha)}(kr) Y_{00}(\theta, \varphi)$ which in our notation is

$$kh_0^{(\alpha)}(kr) Y_{00}(\theta, \varphi)|_{k=0} = \sum_L kh_L^{(\alpha)}(kR) Y_{L0}(\Theta, \varphi) T_{L0}^{(0)}(-k\rho)|_{k=0}. \quad \text{(I)}$$

The generalization of this relation to arbitrary multipolarity in the limit of $k = 0$, i.e., the expansion of the so-called solid harmonics, was performed by Carlson and Rushbrooke[14] and independently by Rose.[15] Recently, this subject has been treated again by Sack.[16]

There exist, however, earlier papers in which the translation of multipole fields was either treated directly or used in the solution of specific problems, sometimes in the form of a translation operation and sometimes as a product decomposition for the case $\mathbf{r} = \varrho_1 - \varrho_2$, discussed in the introduction.

The first of those of which we are aware is by Lord Rayleigh.[17] In that paper he treats the problem of what would now be called an "artificial dielectric," i.e., of an arrangement of spheres of one material in a rectangular lattice embedded in another material. One may consider this paper to constitute the beginning of modern solid-state physics. He treats the problem in the limit of long wavelengths by expanding about a given sphere the potential associated with the neighboring scattering spheres and summing explicitly the contributions from the nearest neighbors. This paper thus contains implicitly the expansion of the regular and the irregular solid harmonics about a shifted origin, and the result is obtained by direct computation. The generalization of the same problem to finite wavelength was

[13] A. M. Legendre, Mem. Math. Phys. Acad. Roy. Sci. Paris **10** (1785); for historical details see, e.g., E. W. Hobson, *The Theory of Spherical and Ellipsoidal Harmonics* (Cambridge University Press, London, 1931), p. 16.

[14] B. C. Carlson and G. S. Rushbrooke, Proc. Cambridge Phil. Soc. **46**, 626 (1950). Note also extensions of this work by R. J. Buehler and J. O. Hirschfelder, Phys. Rev. **83**, 628 (1951); **85**, 149 (1952), and by P. R. Fontana, J. Math. Phys. **2**, 825 (1961).

[15] M. E. Rose, J. Math. and Phys. **37**, 215 (1958).

[16] R. A. Sack, J. Math. Phys, **5**, 245, 252, 260 (1964).

[17] Lord Rayleigh, Phil. Mag. (5) **34**, 481 (1892).

performed by Kasterin[18] extending operator techniques introduced by Lord Rayleigh.[19] This paper thus contains, but only implicitly, the expansion of arbitrary multipole fields about a shifted origin.

Explicit formulas for the translation of multipole fields were developed in a series of papers beginning with a paper by Sato.[20] He derives the matrix elements of T for $h_l^{(2)}(kr)$ in terms of a power series in $k\rho$. He gives only recursion relations for the coefficients of the series. The derivation is based on the integral representation Eq. (29′) derived earlier by Syono[21] on the basis of a theorem by Weyl.[22] Sato rederives (29′) by Fourier transform techniques, which turns out here to be a straightforward but clumsy method. Friedman and Russek[23] generalize Sato's work to displacements into an arbitrary direction, i.e., they calculate the matrix elements of the parallel displacement operator S for both $j_l(kr)$ and $h_l^{(\alpha)}(kr)$ which they give in terms of $j_\lambda(k\rho)$; they give explicit expressions involving summations for the expansion coefficients. They refer for the integral representations (29′) and (30′) to Stratton,[24] who, unfortunately, sweeps under the rug the essential point of Weyl's theorem, namely, the problems associated with the continuation of the region of integration into the complex domain. The work was continued by Stein[25] who recognized that the expansion coefficients of Friedman and Russek[23] contain 3-j coefficients. He also generalizes the previous results to include vector spherical harmonics; he uses Stratton's definitions for the vector spherical harmonics. Stein's expansion coefficients still are sums over several 3-j coefficients. Cruzan[26] sums them up to one term for which he gives an explicit expression. We find that they can be written simply as a product of 3-j and 6-j coefficients.

The integral representation (30′) has also been derived by Erdelyi[27] using the differential operator technique going back to Lord Rayleigh[19] and Maxwell.[28] His paper is based on previous work by Whittaker[29] and Van der Pol.[30]

There exists also an independent explicit calculation for the shift of spherical vector functions with $J = 0$, 1, and 2 by Ament.[31]

The product decomposition for the regular multipole fields $j_l(kr)Y_{lm}(\theta, \varphi)$ as a function of ϱ_1 and ϱ_2 for $\mathbf{r} = \varrho_1 - \varrho_2$ was obtained by Kohn and Rostoker[32] by essentially the same method as used in this paper [see Eq. (10) and following]. They used it in expanding the free-wave Green's function in connection with the solution of the Schrödinger equation for crystals. Their work is a continuation of the method introduced by Lord Rayleigh.[17]

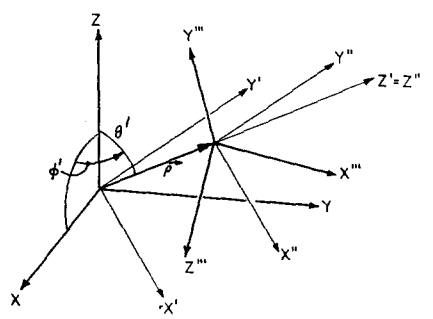


FIG. 1. Coordinates in a rotation–translation transformation.

III. THE BASIC FORMULAS

In a rotation–translation transformation (Fig. 1), one can always perform the operation in three consecutive steps: (1) Rotate the coordinate system (xyz) so that in the new orientation $(x'y'z')$ the z axis is parallel to the direction of the translation [Euler angles $(0, \theta', \varphi')$]; (2) translate the coordinate system to position $(x''y''z'')$ (shift vector ϱ); (3) rotate the coordinate system to bring it into the final position $(x'''y'''z''')$ [Euler angles $(\alpha\beta\gamma)$, not indicated in Fig. 1].

The steps (1) and (3) are performed by means of the well-known rotation operators D. We can thus concern ourselves exclusively with the translation operators. We would like to call these operators T, which can be considered to stand for "translation operator" or "Taylor's theorem."

In the shift along the z axis (Fig. 2) the angle φ does not change. The φ dependence thus remains unchanged and we omit reference to it for the time being. Our aim is to express the function $f(\theta, r)$ in terms of the new coordinates (Θ, R).

Since in the translation the values of the function

[18] N. Kasterin, Verslagen Kon. Akad. van Wetensch., Amsterdam, Wis-en Natuurk. Afd. **6**, 460 (1897/1898).
[19] Lord Rayleigh, *Theory of Sound* (Dover Publications, Inc., New York, 1945), Vol. 2, 2nd ed., Secs. 329, 330.
[20] Y. Sato, Bull. Earthquake Res. Inst. Tokyo Univ. **28**, 1, 175 (1950).
[21] S. Syono, Proc. Phys. Math. Soc. Japan, 3rd Ser., **20**, 100 (1938); Geophysical Magazine **12**, 67 (1938–1939).
[22] H. Weyl, Ann. Physik **60**, 481 (1919).
[23] B. Friedman and J. Russek, Quart. Appl. Math. **12**, 13 (1954).
[24] J. A. Stratton, *Electromagnetic Theory* (McGraw-Hill Book Company, Inc., New York, 1941), Sec. 9.29, pp. 577, 578.
[25] S. Stein, Quart. Appl. Math. **19**, 15 (1961).
[26] O. R. Cruzan, Quart. Appl. Math. **20**, 33 (1962).
[27] A. Erdelyi, Physica **4**, 107 (1937).
[28] J. C. Maxwell, *Electricity and Magnetism.* (Clarendon Press, Oxford, England, 1873), Part 1, Chap. 9.
[29] E. T. Whittaker, Math. Ann. **57**, 333 (1903).
[30] B. Van der Pol, Physica **3**, 385, 393 (1936).
[31] W. S. Ament, NRL Report 5307 (April 1959).
[32] W. Kohn and N. Rostoker, Phys. Rev. **94**, 1111 (1954).

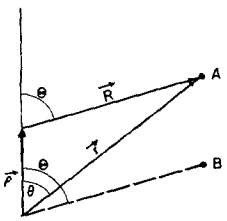


FIG. 2. Coordinates in a pure translation transformation.

at a particular point in space are to remain unchanged, we have

$$f(r, \theta) = \bar{f}(R, \Theta) \tag{1}$$

if R, Θ fulfill the "geometry" of Fig. 2. We would like to express the "new" function $\bar{f}$ by means of the translation operator in terms of the "old" function f. We are thus looking for an operator which would give

$$\bar{f}(R, \Theta) = Tf(R, \Theta). \tag{2}$$

Together with (1), this says that we are looking for an operator T which gives the value of the function f at the point A in terms of the values of the function f in the vicinity of point B. Clearly, this is fulfilled by

$$T(\rho) = e^{\rho \cdot \mathrm{grad}}, \tag{3}$$

i.e., by Taylor's theorem. We used above the expression "in the vicinity of" instead of "at" because of the appearance of derivatives of all orders in (3). Thus the values of the function f in an arbitrarily small but finite region around point B are needed. As is commonly done, we replace the operator grad by the Hermitian operator $\mathbf{p} = -i$ grad and obtain

$$T(\rho) = e^{i\rho \cdot \mathbf{p}}. \tag{3'}$$

Actually the form (3) or (3') is valid in general, for arbitrary directions of the displacement ρ. For T, in contrast to S, only the case in which $\rho_x = \rho_y = 0$ is needed.

We now apply the translation operator to a scalar multipole field. We thus want to evaluate:

$$\begin{aligned} j_l(kr)Y_{lm}(\theta, \varphi) &= Tj_l(kR)Y_{lm}(\Theta, \varphi) \\ &= \sum_L j_L(kR)Y_{Lm}(\Theta, \varphi)T^{(m)}_{Ll}(k\rho), \end{aligned} \tag{4}$$

where we have introduced the abbreviation

$$T^{(m)}_{Ll}(k\rho) = \langle j_L(kR)Y_{Lm}(\Theta, \varphi), e^{i\rho \cdot \mathbf{p}} j_l(kR)Y_{lm}(\Theta, \varphi)\rangle \tag{5}$$

for the matrix elements of the displacement operator. We have chosen this abbreviation in analogy to the notation of the rotation operators; in the matrix elements of T it is m that remains unchanged while in the matrix elements of D it is j that stays fixed.

We note some properties of T. From (3') it follows immediately that T is a unitary operator for real displacements. Thus the effect of two consecutive displacements x and y is simply

$$T^{(m)}_{Ll}(x + y) = \sum_{l'} T^{(m)}_{Ll'}(x)T^{(m)}_{l'l}(y). \tag{6}$$

There appear no normalization constants. The inverse shift $(-x)$ is represented by

$$T^{(m)}_{Ll}(-x) = T^{(m)}_{lL}(x)^* = T^{(m)}_{lL}(x), \tag{7}$$

the last equation being a consequence of the reality of the matrix elements $T^{(m)}_{Ll}$. The reality depends on the fact that the operator does not contain any φ dependence, and on the definition of the function $f_{lm} = Y_{lm}(\theta, \varphi)j_l(kr)$; additional phase factors in f_{lm} could lead to nonreal matrix elements.

Finally, the commutation relations of the operators T and D for infinitesimal translations and rotations are most easily obtained from the well-known commutation relations of the linear and angular momentum operators. We do not need them here.

We now proceed to the evaluation of the matrix elements of T. From (4) we find immediately

$$\begin{aligned} T^{(m)}_{Ll}(k\rho) = \int d\cos\Theta\, d\varphi\, R^2\, dR \\ \times j_L(kR)Y^*_{Lm}(\Theta, \varphi)j_l(kr)Y_{lm}(\theta, \varphi), \end{aligned} \tag{8}$$

where we have to express r and θ as functions of R and Θ. We have (see Fig. 2)

$$\begin{aligned} r &= (\rho^2 + R^2 + 2\rho R\cos\Theta)^{\frac{1}{2}}, \\ \cos\theta &= (\rho + R\cos\Theta)/(\rho^2 + R^2 + 2\rho R\cos\Theta)^{\frac{1}{2}}. \end{aligned} \tag{9}$$

The direct integration of (8) is not simple. It can, however, be circumvented by a method which one could call "Fourier transform in spherical coordinates." We start with the vector equation

$$\mathbf{r} = \boldsymbol{\rho} + \mathbf{R} \tag{10}$$

and consider the expansion of

$$e^{i\mathbf{k} \cdot \mathbf{r}} = e^{i\mathbf{k} \cdot \boldsymbol{\rho}} e^{i\mathbf{k} \cdot \mathbf{R}} \tag{11}$$

with arbitrary $\mathbf{k}$ into spherical waves with respect to an arbitrary coordinate system. Let the components of the vectors in this system be (r, θ, φ), (R, Θ, Φ), (ρ, η, ψ), and (k, θ', φ'). Then we have

$$\begin{aligned} \sum_{l'm'} 4\pi i^{l'} Y^*_{l'm'}(\theta', \varphi')Y_{l'm'}(\theta, \varphi)j_{l'}(kr) \\ = \sum_{LM} 4\pi i^L Y^*_{LM}(\theta', \varphi')Y_{LM}(\Theta, \Phi)j_L(kR) \\ \times \sum_{\lambda\mu} 4\pi i^\lambda Y^*_{\lambda\mu}(\theta', \varphi')Y_{\lambda\mu}(\eta, \psi)j_\lambda(k\rho). \end{aligned} \tag{12}$$

We now multiply both sides of (12) by $Y_{lm}(\theta', \varphi')$ and integrate over the direction of k, i.e., over φ', θ', which gives

$$
\begin{aligned}
&i^l Y_{lm}(\theta, \varphi) j_l(kr) \\
&= \sum_{\lambda\mu} \sum_{LM} 4\pi i^{\lambda+L} Y_{\lambda\mu}(\eta, \psi) j_\lambda(k\rho) Y_{LM}(\Theta, \Phi) j_L(kR) \\
&\times \int Y_{lm}(\theta', \varphi') Y^*_{LM}(\theta', \varphi') Y^*_{\lambda\mu}(\theta', \varphi') d\cos\theta' \, d\varphi' \\
&= \sum_{\lambda\mu} \sum_{LM} 4\pi i^{\lambda+L} Y_{LM}(\Theta, \Phi) j_L(kR) Y_{\lambda\mu}(\eta, \psi) j_\lambda(k\rho) \\
&\times (-1)^m [(2l+1)(2L+1)(2\lambda+1)/4\pi]^{\frac{1}{2}} \\
&\times \begin{pmatrix} l & L & \lambda \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l & L & \lambda \\ -m & M & \mu \end{pmatrix} \\
&\equiv i^l \sum_{LM} Y_{LM}(\Theta, \Phi) j_L(kR) S_{LM,lm}(k\rho, \eta, \psi). \qquad (13)
\end{aligned}
$$

We have here obtained the expression for the matrix element of the operator for the parallel displacement of the coordinate system along an arbitrary direction, which we call $S_{LM,lm}(k\rho, \eta, \psi)$. To obtain (4) we specialize to $\eta = \psi = 0$ which enforces $\Phi = \varphi$ and $\mu = 0$ and then also $M = m$. This yields:

$$
\begin{aligned}
T^{(m)}_{Ll}(k\rho) &= \sum_\lambda i^{\lambda+L-l}(-1)^m(2\lambda+1)[(2l+1)(2L+1)]^{\frac{1}{2}} \\
&\times \begin{pmatrix} l & L & \lambda \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l & L & \lambda \\ -m & m & 0 \end{pmatrix} j_\lambda(k\rho). \qquad (14)
\end{aligned}
$$

From (14) we can verify (7) explicitly. Owing to the properties of the 3-j coefficients, $L + \lambda + l$ must be an even number. Thus λ can change only in steps of 2, and $\lambda = (L + l) \bmod 2$. The change $\rho \to -\rho$ multiplies the spherical Bessel functions by $(-1)^\lambda$, and therefore we have

$$
T^{(m)}_{Ll}(-k\rho) = (-1)^{l-L} T^{(m)}_{Ll}(k\rho). \qquad (7')
$$

We note two more forms for the T's [see discussion following Eq. (26)]:

$$
\begin{aligned}
T^{(m)}_{Ll}(k\rho) &= (-1)^L [(2L+1)(L+m)!(L-m)! \\
&\times (2l+1)(l+m)!(l-m)!]^{\frac{1}{2}} \sum_p (-1)^p \frac{(2p)!}{2^p p!} \qquad (15) \\
&\times [(p+m)!(p-m)!(L-p)!(l-p)!]^{-1} \frac{j_{L+l-p}(k\rho)}{(k\rho)^p}
\end{aligned}
$$

and

$$
\begin{aligned}
T^{(m)}_{Ll}(k\rho) &= (-1)^{L-l+m}[(2l+1)(2L+1)]^{\frac{1}{2}} \\
&\times \left\{ \frac{(l+m)!(l-m)!}{(L+m)!(L-m)!} \right\}^{\frac{1}{2}} \sum_p (-1)^p \frac{(2p)!}{2^p p!} \\
&\times [(p+m)!(p-m)!]^{-1} \frac{(L+p)!}{(l-p)!} \frac{j_{L-l+p}(k\rho)}{(k\rho)^p}. \qquad (16)
\end{aligned}
$$

The general rotation–translation transformation in which the coordinate system is displaced from its original position to the final position (see Fig. 1) is then given in terms of the translation operator $T^{(m)}_{Ll}$ by

$$
\begin{aligned}
j_l(kr) Y_{lm}(\theta, \varphi) &= \sum_{ML m'} \mathfrak{D}^{(L)}_{Mm'}(\alpha, \beta, \gamma) T^{(m')}_{Ll}(k\rho) \\
&\times \mathfrak{D}^{(l)}_{m'm}(0, \theta, \varphi) j_L(kR) Y_{LM}(\Theta, \Phi) \sum_{ML m'}. \qquad (17)
\end{aligned}
$$

We emphasize again that we displace and rotate the coordinate system and not the "body."

We now show that $S_{LM,lm}$ indeed follows from (17). This is very simple since the second rotation, i.e., the rotation after the shift, is just the inverse of the first rotation. With

$$
\mathfrak{D}^{(L)}_{Mm}(\omega^{-1}) = \mathfrak{D}^{(L)}_{mM}(\omega)^* = (-1)^{m-M} \mathfrak{D}^{(L)}_{-m-M}(\omega), \qquad (18)
$$

we have

$$
\begin{aligned}
S_{LM,lm}(k\rho, \eta, \psi) &\equiv \sum_{m'} (-1)^{M-m'} \mathfrak{D}^{(L)}_{-m'-M}(0, \eta, \psi) \\
&\times \mathfrak{D}^{(l)}_{m'm}(0, \eta, \psi) T^{(m')}_{Ll}(k\rho) = \sum_{m'} (-1)^{M-m'} \\
&\times \sum_{\lambda\mu} (2\lambda+1) \begin{pmatrix} L & l & \lambda \\ -m' & m' & 0 \end{pmatrix} \\
&\times \begin{pmatrix} L & l & \lambda \\ -M & m & \mu \end{pmatrix} \mathfrak{D}^{(\lambda)}_{0\mu}(0, \eta, \psi)^* \\
&\times \sum_{\lambda'} i^{\lambda'+L-l}(-1)^{m'}(2\lambda'+1)[(2l+1)(2L+1)]^{\frac{1}{2}} \\
&\times \begin{pmatrix} L & l & \lambda' \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} L & l & \lambda' \\ -m' & m' & 0 \end{pmatrix} j_{\lambda'}(k\rho) \qquad (19)
\end{aligned}
$$

The summation over m' gives $(2\lambda + 1)^{-1}\delta_{\lambda\lambda'}$. Replacing the remaining $\mathfrak{D}^*$ by an appropriate Y one obtains back $S_{LM,lm}$ as defined by (13), as is necessary:

$$
\begin{aligned}
&S_{LM,lm}(k\rho, \eta, \psi) \\
&= \sum_{\lambda\mu} i^{L+\lambda-l}(-1)^m [4\pi(2L+1)(2l+1)(2\lambda+1)]^{\frac{1}{2}} \\
&\times \begin{pmatrix} L & l & \lambda \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} L & l & \lambda \\ M & -m & \mu \end{pmatrix} Y_{\lambda\mu}(\eta, \psi) j_\lambda(k\rho). \qquad (19')
\end{aligned}
$$

We would like to emphasize the simplicity of the dependence of $T^{(m)}_{Ll}$ and in particular of $S_{LM,lm}$ on $k\rho$, which is the consequence of the symmetry (11) between ρ and R. This symmetry is not at all evident from the considerations leading to (5). There one sees the symmetry between r and R; they are both associated with the "wavefunction" while ρ is associated with the "operator." A lack of symmetry, as exemplified in the forms (15) and (16), would thus offhand not be unexpected. One can obtain a

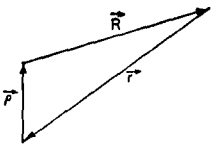


FIG. 3. Symmetrized translation transformation.

form exhibiting the complete symmetry between the three vectors if one considers the closed triangle of Fig. 3:

$$\mathbf{R} + \boldsymbol{\rho} + \mathbf{r} = 0$$

and expands

$$1 = e^{i\mathbf{k}\cdot\mathbf{R}} e^{i\mathbf{k}\cdot\boldsymbol{\rho}} e^{i\mathbf{k}\cdot\mathbf{r}}$$

$$= \sum_{lm} \sum_{\lambda\mu} \sum_{LM} i^{l+\lambda+L} [(4\pi)^3 (2l+1)(2\lambda+1)(2L+1)]^{\frac{1}{2}}$$

$$\times \begin{pmatrix} l & \lambda & L \\ m & \mu & M \end{pmatrix} \begin{pmatrix} l & \lambda & L \\ 0 & 0 & 0 \end{pmatrix}$$

$$\times j_l(kr) Y_{lm}(\theta,\varphi) j_L(kR) Y_{LM}(\Theta,\Phi) j_\lambda(k\rho) Y_{\lambda\mu}(\eta,\psi). \quad (20)$$

Here naturally all three vectors are equivalent and enter in a fully symmetric way.

In some applications one needs the expansion of the so-called "solid harmonics," $r^l Y_{lm}(\theta, \varphi)$, about a shifted origin. Traditionally this is done in analogy to the definition of the Legendre polynomials where the shift is "down" along the z axis. [See Introduction and Eq. (I).] We can obtain the appropriate formulas by multiplying both sides of Eq. (4) by k^{-l} and going to the limit $k = 0$; evidently l is the lowest power of k on both sides of (4). We now determine the lowest power of k for both factors of the right-hand side of (4) separately. It follows immediately from the 3-j coefficients of (14) that the sum over λ in (14) has only a finite number of terms and begins at $\lambda_{\min} = l - L$. The phase factor, from (14) and (7'), is then $i^{\lambda_{\min}+l-L} = +1$ if $l - L$ is an even number and -1 if $l - L$ is odd. Thus we have

$$T_{Ll}^{(m)}(-k\rho) \to (-1)^{l-L+m}(2l - 2L + 1)$$

$$\times [(2l+1)(2L+1)]^{\frac{1}{2}} \begin{pmatrix} l & L & l-L \\ 0 & 0 & 0 \end{pmatrix}$$

$$\times \begin{pmatrix} l & L & l-L \\ -m & m & 0 \end{pmatrix} \frac{(k\rho)^{l-L}}{(2l-2L+1)!!}$$

and, using the explicit expression for the first of the (3-j) symbols and replacing the other by a Wigner coefficient, we obtain the expression

$$r^l Y_{lm}(\theta,\varphi) = \sum_{LM} B_{LM,lm}$$

$$\times \rho^{l-L} Y_{l-L,m-M}(\eta,\psi) R^L Y_{LM}(\Theta,\Phi), \quad (21)$$

with

$$B_{LM,lm} = (-1)^{L-l} \left[\frac{4\pi(2l+1)!}{(2L+1)!(2l-2L+1)!} \right]^{\frac{1}{2}}$$

$$\times (L \quad M \quad l-L \quad m-M \mid lm) \quad (22)$$

IV. TRANSLATION OF INCOMING AND OUTGOING MULTIPOLE FIELDS

When one looks at Eq. (4) or at the analogous equation for parallel displacement involving $S_{LM,lm}(k\rho, \eta, \psi)$ it is tempting to split it up into incoming and outgoing waves:

$$h_l^{(1)}(kr) Y_{lm}(\theta,\varphi) + h_l^{(2)}(kr) Y_{lm}(\theta,\varphi)$$

$$= \sum_L h_L^{(1)}(kR) Y_{Lm}(\Theta,\Phi) T_{Ll}^{(m)}(k\rho)$$

$$+ \sum_L h_L^{(2)}(kR) Y_{Lm}(\Theta,\Phi) T_{Ll}^{(m)}(k\rho). \quad (23)$$

Since one should think that incoming and outgoing waves are quite distinct and should not mix it would seem that it must be possible to "split (23) down the middle," i.e., that there should hold separately, for $\alpha = 1$ and for $\alpha = 2$,

$$h_l^{(\alpha)}(kr) Y_{lm}(\theta,\varphi)$$

$$= \sum_L h_L^{(\alpha)}(kR) Y_{Lm}(\Theta,\Phi) T_{Ll}^{(m)}(k\rho), \quad (24)$$

$$\alpha = 1, 2.$$

This is indeed the case. In the most elementary way it can be seen by considering the formula ($Z_{l+\frac{1}{2}}$ is a general Bessel function)

$$Z_{l+\frac{1}{2}}(kr) Y_{ll}(\theta,\varphi)$$

$$= 2^{l+\frac{1}{2}} \Gamma(l+\tfrac{1}{2}) \sum_{n=0}^{\infty} (l+n+\tfrac{1}{2}) \frac{J_{l+n+\frac{1}{2}}(k\rho)}{(k\rho)^{l+\frac{1}{2}}}$$

$$\times Z_{l+n+\frac{1}{2}}(kR) Y_{ll}(\Theta',\varphi) C_n^{(l+\frac{1}{2})}(\cos\Theta'), \quad (25)$$

which has been obtained from the Bessel function addition theorem[33] by observing that

$$R/r = \sin\theta/\sin\Theta', \quad (26)$$

where $\Theta' = \pi - \Theta$. Equation (25) is nothing but the expansion (4) for $m = l$ "in raw form." It is valid for $Z_{l+\frac{1}{2}} = J_{l+\frac{1}{2}}$ without restrictions, for $Z_{l+\frac{1}{2}} = H_{l+\frac{1}{2}}^{(1,2)}$ only for $r, R > \rho$. Otherwise the expansion is identical for $J_{l+\frac{1}{2}}$ and for $H_{l+\frac{1}{2}}^{(1,2)}$. The expansion for $m < l$ can be obtained from (25) by means of the lowering operator L_-. As a matter of fact the forms (15) and (16) for T were obtained this way. This procedure is however rather clumsy. We now give a more transparent derivation.

[33] W. Magnus and F. Oberhettinger, *Formulas and Theorems for the Functions of Mathematical Physics* (Chelsea Publishing Company, New York, 1954), p. 21.

We start again with (11) but in contrast to (12) use the expansion of a plane wave into multipoles L, M only for $e^{i\mathbf{k}\cdot\varrho}$, multiply by $Y_{lm}(\theta', \varphi')$ and integrate over φ' and θ', the integration path S for θ' remaining unspecified for the time being:

$$\int_0^{2\pi} d\varphi' \int_S d\cos\theta' Y_{lm}(\theta', \varphi') e^{i\mathbf{k}\cdot\mathbf{r}}$$
$$= 4\pi \sum_{\lambda\mu} i^\lambda Y_{\lambda\mu}(\eta, \psi) j_\lambda(k\rho)$$
$$\times \int_0^{2\pi} d\varphi' \int_S d\cos\theta' Y^*_{\lambda\mu}(\theta', \varphi') Y_{lm}(\theta', \varphi') e^{i\mathbf{k}\cdot\mathbf{R}}$$
$$= 4\pi \sum_{\lambda\mu} \sum_{LM} i^\lambda (-1)^{M+\mu} Y_{\lambda\mu}(\eta, \psi) j_\lambda(k\rho)$$
$$\times \left[\frac{(2L+1)(2l+1)(2\lambda+1)}{4\pi}\right]^{\frac{1}{2}}$$
$$\times \begin{pmatrix} L & l & \lambda \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} L & l & \lambda \\ M & -m & \mu \end{pmatrix}$$
$$\times \int_0^{2\pi} d\varphi' \int_S d\cos\theta' Y_{LM}(\theta', \varphi') e^{i\mathbf{k}\cdot\mathbf{R}}. \tag{27}$$

We recognize the factor multiplying the integral on the right-hand side to be $i^{l-L} S_{LM,lm}(k\rho, \eta, \psi)$. We now have only to show that by choosing a suitable integration path S we can obtain the spherical Bessel functions, i.e.,

$$\int_0^{2\pi} d\varphi' \int_S d\cos\theta' Y_{lm}(\theta', \varphi') e^{i\mathbf{k}\cdot\mathbf{r}}$$
$$= C_S z_l(kr) Y_{lm}(\theta, \varphi). \tag{28}$$

If we choose the path S to go from $\cos\theta' = -1$ to $\cos\theta' = +1$ then we indeed obtain $C_S = 4\pi i^l$ and $z_l(kr) = j_l(kr)$. It is somewhat more intricate to show how to obtain the spherical Hankel functions. We begin by proving the special cases

$$\int_{-1}^{i\infty} d\cos\theta P_l(\cos\theta) e^{ikr\cos\theta} = i^l h_l^{(2)}(kr) \tag{29}$$

and

$$\int_{i\infty}^{+1} d\cos\theta P_l(\cos\theta) e^{ikr\cos\theta} = i^l h_l^{(1)}(kr). \tag{30}$$

The specialization here consists in taking r to lie along the z axis, in which case only $m = 0$ remains. We will now show that the left-hand sides obey the differential equation for spherical Bessel functions and then determine the kind and normalization of the spherical Bessel functions by finding the asymptotic behavior[34]:

[34] E. T. Whittaker, Proc. London Math. Soc. **35**, 198 (1903).

$$\left[\frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r} + k^2 - \frac{l(l+1)}{r^2}\right]$$
$$\times \int_a^b dx P_l(x) e^{ikrx} = \int_a^b \left[-k^2x^2 + \frac{2ikx}{r}\right.$$
$$\left. + k^2 - \frac{l(l+1)}{r^2}\right] dx P_l(x) e^{ikrx}. \tag{31}$$

Using the identity

$$e^{ikrx} = (1/ikr)\, de^{ikrx}/dx$$

and integrating by parts we obtain Eq. (31) equal to

$$\frac{k}{ir}(1-x^2)P_l(x)e^{ikrx}\Big|_a^b$$
$$+ \frac{1}{r^2}\frac{d}{dx}[(1-x^2)P_l(x)]e^{ikrx}\Big|_a^b + \frac{2x}{r}P_l(x)e^{ikrx}\Big|_a^b$$
$$- \int_a^b \frac{1}{r^2} e^{ikrx} dx \left\{\frac{d^2}{dx^2}[(1-x^2)P_l(x)]\right.$$
$$\left. + 2\frac{d}{dx}[xP_l(x)] + l(l+1)P_l(x)\right\}. \tag{32}$$

This expression vanishes at the boundaries $x = \pm 1$ and $x = i\infty$; the expression in the curly brackets in the integral vanishes identically because it is just the differential equation for $P_l(x)$. We now check the asymptotic behavior:

$$\int_a^b dx P_l(x) e^{ikrx} = \int_a^b dx P_l(x) \frac{d}{dx}\frac{e^{ikrx}}{ikr}$$
$$= P_l(x)\frac{e^{ikrx}}{ikr}\Big|_a^b - \int_a^b dx \frac{e^{ikr}}{ikr}\frac{dP_l}{dx}$$
$$= P_l(x)\frac{e^{ikrx}}{ikr}\Big|_a^b - \frac{dP_l(x)}{dx}\frac{e^{ikrx}}{(ikr)^2}\Big|_a^b + \cdots. \tag{33}$$

Inserting the appropriate integration boundaries in (33) and inserting the appropriate asymptotic expansion for $h_l^{(1,2)}$ in (29) and (30) we verify immediately the correctness of Eqs. (29) and (30).

We now return to (28). In order to connect this equation with (29) and (30) we rotate the coordinate system in (28) so that the new z axis coincides with the direction of r, and obtain,

$$\int d\varphi' d\cos\theta' Y_{lm}(\theta', \varphi') e^{i\mathbf{k}\cdot\mathbf{r}} = \sum_{m'} \mathfrak{D}^{(l)}_{m'm}(0, \theta, \varphi)$$
$$\times \int d\varphi\, d\cos\theta'' Y_{lm'}(\theta'', \varphi'') e^{ikr\cos\theta''}$$
$$= \sum_{m'} \mathfrak{D}^{(l)}_{m'm}(0, \theta, \varphi) 2\pi\delta_{m'0}$$
$$\times \int d\cos\theta'' Y_{l0}(\theta'', \varphi'') e^{ikr\cos\theta''}$$
$$= 2\pi Y_{lm}(\theta, \varphi) \int d\cos\theta'' P_l(\cos\theta'') e^{ikr\cos\theta''}. \tag{28'}$$

These transformations are perfectly all right if one integrates over the full surface of a unit sphere, which then leads directly to $j_l(kr)$. It has to be shown in detail that this procedure is still applicable with the region of integration indicated in (29) and (30). We devote Appendix A to the proof of (28′). We then obtain

$$\int_0^{2\pi} d\varphi' \int_{-1}^{i\infty} d\cos\theta' Y_{lm}(\theta', \varphi') e^{i\mathbf{k}\cdot\mathbf{r}} = 2\pi i^l Y_{lm}(\theta, \varphi) h_l^{(2)}(kr) \tag{29′}$$

and

$$\int_0^{2\pi} d\varphi' \int_{i\infty}^{1} d\cos\theta' Y_{lm}(\theta', \varphi') e^{i\mathbf{k}\cdot\mathbf{r}} = 2\pi i^l Y_{lm}(\theta, \varphi) h_l^{(1)}(kr). \tag{30′}$$

The expansion (24) is valid only for $\rho < R$. Since we want to be able to use the inverse transformation, we then want to impose also the equivalent condition $\rho < r$. The way in which this restriction appears is shown in Appendix B.

For completeness we would like to write down another form for the expansion of $h_l^{(\alpha)}(kr) Y_{lm}(\theta, \varphi)$, even though it does not conform to the point of view of this article in that the notion of the translation operator has to be given up. We introduce $r_<$ and $r_>$ to denote the smaller and larger of ρ and R, respectively, and we call the angles associated with $r_<$ and $r_>$ in an analogous way $\theta_<$, $\varphi_<$ and $\theta_>$, $\varphi_>$, respectively. Then we can write an equation which is formally valid without restrictions:

$$h_l^{(\alpha)}(kr) Y_{lm}(\theta, \varphi) = \sum_{LM\lambda\mu} i^{L+\lambda-l}(-1)^m [4\pi(2L+1)(2l+1)(2\lambda+1)]^{\frac{1}{2}} \times \begin{pmatrix} L & l & \lambda \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} L & l & \lambda \\ M & -m & \mu \end{pmatrix} \times Y_{LM}(\theta_>, \varphi_>) h_L^{(\alpha)}(kr_>) Y_{\lambda\mu}(\theta_<, \varphi_<) j_\lambda(kr_<). \tag{34}$$

The formula (I) of the Introduction corresponds to the limiting case $k \to 0$, analogous to Eq. (21) discussed in Sec. II, applied, however, to (24) instead of (4). Except for the limitations on the values of ρ and R, discussed above, Eq. (24) reduces, in the limit $k \to 0$, to Eqs. (21) and (22) provided that in (21) one replaces r^l by r^{-l-1} and R^L by R^{-L-1}. With these replacements Eq. (21) then describes the expansion of static multipole fields about a "down"-shifted origin.

V. TRANSLATION OF TENSOR MULTIPOLE FIELDS

The tensor wave fields are defined by

$$j_l(kr) Y_{lSM}^{[J]}(\theta, \varphi) = \sum_{m's'} (lm'Ss' \mid JM) Y_{lm'}(\theta, \varphi) \hat{e}_{s'}^{[S]} j_l(kr) \equiv [\mathbf{Y}^{[l]}(\hat{\mathbf{r}}) \times \hat{\mathbf{e}}^{[S]}]_M^{[J]} j_l(kr), \tag{35}$$

where we follow the notation of Fano and Racah,[6] and where the angles of the unit vector $\hat{\mathbf{r}}$ are θ, φ. The quantities $\hat{e}_s^{[S]}$ are unit tensors which, in the nomenclature of Ref. 6, transform under rotations in "contra-standard" fashion. For example, for $S=1$, the quantities $Y_{l1M}^{[J]}$ are the usual vector spherical harmonics, and the unit tensors are, in terms of the Cartesian unit vectors $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$, $\hat{\mathbf{z}}$,

$$\hat{e}_{+1}^{[1]} = (-\hat{\mathbf{x}} - i\hat{\mathbf{y}})/\sqrt{2}$$
$$\hat{e}_0^{[1]} = \hat{\mathbf{z}}$$
$$\hat{e}_{-1}^{[1]} = (\hat{\mathbf{x}} - i\hat{\mathbf{y}})/\sqrt{2}. \tag{36}$$

The phase convention of (36) is that of Condon and Shortley,[35] which is the one commonly used in the definition of the functions $Y_{lm}(\theta, \varphi)$.

The translation formulas are obtained immediately by observing that the unit tensors do not change under translation.

Since (35) involves coupling of angular momenta, the use of the displacement operator S is more convenient for the derivation than the use of T. We rewrite Eq. (13) as

$$j_l(kr) Y_{lm}(\hat{r}) = \sum_{l'\lambda} (4\pi)^{\frac{1}{2}} i^{-\lambda-l'-l} \begin{pmatrix} \lambda & l' & l \\ 0 & 0 & 0 \end{pmatrix} j_\lambda(k\rho) j_{l'}(kR) \times [(2l'+1)(2\lambda+1)]^{\frac{1}{2}} [\mathbf{Y}^{[\lambda]}(\hat{\rho}) \times \mathbf{Y}^{[l']}(\hat{R})]_m^{[l]} \equiv \sum_{\lambda l'} j_\lambda(k\rho) j_{l'}(kR) [\mathbf{Y}^{[\lambda]}(\hat{\rho}) \times \mathbf{Y}^{[l']}(\hat{R})]_m^{[l]} t_{\lambda l';l}. \tag{37}$$

We now couple $\hat{e}_s^{[S]}$ to both sides of (37) to give a total angular momentum J:

$$\begin{aligned} j_l(kr) [\mathbf{Y}^{[l]}(\hat{r}) \times \hat{\mathbf{e}}^{[S]}]_M^{[J]} &= \sum_{\lambda l'} j_\lambda(k\rho) j_{l'}(kR) \times [[\mathbf{Y}^{[\lambda]}(\hat{\rho}) \times \mathbf{Y}^{[l']}(\hat{R})]^{[l]} \times \hat{\mathbf{e}}^{[S]}]_M^{[J]} t_{\lambda l';l} \\ &= \sum_{\lambda l'} \sum_{J'} j_\lambda(k\rho) j_{l'}(kR) \times [\mathbf{Y}^{[\lambda]}(\hat{\rho}) \times [\mathbf{Y}^{[l']}(\hat{R}) \times \hat{\mathbf{e}}^{[S]}]^{[J']}]_M^{[J]} \times \langle \lambda, (l'S)J' \mid (\lambda l')l, S\rangle^{(J)} t_{\lambda l';l} \\ &\equiv \sum_{l'J'M'} j_{l'}(kR) Y_{l'SM'}^{[J']}(\Theta, \Phi) \times S_{l'J'M',lJM}^{(S)}(k\rho, \eta, \psi), \end{aligned} \tag{38}$$

[35] E. U. Condon and G. H. Shortley, *The Theory of Atomic Spectra* (Cambridge University Press, London, 1951).

where

$$S^{(S)}_{l'J'M',lJM}(k\rho,\eta,\psi)=\sum_{\lambda\mu}(4\pi)^{\frac{1}{2}}i^{l'-l-\lambda}(-1)^{S+J+J'-M}\times\begin{pmatrix}\lambda & l' & l\\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix}\lambda & J' & J\\ \mu & M' & -M\end{pmatrix}\begin{Bmatrix}\lambda & J' & J\\ S & l & l'\end{Bmatrix}\times j_\lambda(k\rho)Y_{\lambda\mu}(\eta,\psi). \quad (39)$$

Equations (38) and (39) are the generalizations of Eqs. (13) and (19′) to tensor multipole fields, and Eq. (39) reduces to Eq. (19′) for $S=0$, viz.,

$$S^{(0)}_{l'l'M',llM}(k\rho,\eta,\psi)=S_{l'M',lM}(k\rho,\eta,\psi).$$

Finally, we obtain the generalization of T to the case of tensor multipole fields by putting $\eta=\psi=0$ in (39), which yields

$$j_l(kr)Y^{[J]}_{lSM}(\theta,\varphi)=\sum_{l'J'}j_{l'}(kR)Y^{[J']}_{l'SM}(\Theta,\varphi)\mathcal{T}^{(S;M)}_{l'J',lJ}(k\rho) \quad (40)$$

with

$$\mathcal{T}^{(S;M)}_{l'J',lJ}(k\rho)=\sum_\lambda i^{l'-l-\lambda}(-1)^{S+J+J'-M}\times(2\lambda+1)[(2l+1)(2J+1)(2J'+1)(2l'+1)]^{\frac{1}{2}}\times\begin{pmatrix}\lambda & l' & l\\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix}\lambda & J' & J\\ 0 & M & -M\end{pmatrix}\begin{Bmatrix}\lambda & J' & J\\ S & l & l'\end{Bmatrix}j_\lambda(k\rho) \quad (41)$$

Again, $\mathcal{T}^{(0,M)}_{l'J',lJ}(k\rho)=T^{(M)}_{l'l}(k\rho)$.

With the limitations on the values of R and ρ discussed in Sec. IV, Eq. (41) is valid also for incoming and outgoing waves, i.e., for the case in which $j_l(kr)$ and $j_{l'}(kR)$ in Eq. (40) are replaced by $h^{(\alpha)}_l(kr)$ and $h^{(\alpha)}_{l'}(kR)$, respectively; $\alpha=1,2$.

For completeness, we give the expressions for the expansion of a tensorial multipole field in terms of two other tensorial multipole fields, i.e.,

$$j_l(kr)Y^{[J]}_{lSM}(\theta,\varphi)=\sum_{l_1J_1M_1l_2J_2M_2}j_{l_1}(kR_1)\times Y^{[J_1]}_{l_1S_1M_1}(\Theta_1,\Phi_1)j_{l_2}(kR_2)Y^{[J_2]}_{l_2S_2M_2}(\Theta_2,\Phi_2)\times A^{(S_1,S_2;S)}_{l_1J_1,l_2J_2;lJ}(J_1M_1J_2M_2\mid JM), \quad (42)$$

where the vectors fulfill the relation:

$$\mathbf{r}=\mathbf{R}_1-\mathbf{R}_2 \quad (43)$$

and

$$\mathbf{S}=\mathbf{S}_1+\mathbf{S}_2,\qquad \hat{e}^{[S]}_M=[\hat{\mathbf{e}}^{[S_1]}\times\hat{\mathbf{e}}^{[S_2]}]^{[S]}_M, \quad (44)$$

with given S, S_1, and S_2. To derive (42) we again begin with (37). In order to take into account the geometry (43) which differs from (10) by the direction of $\mathbf{R}_2$, we write:

$$j_l(kr)Y_{lM}(\hat{r})=\sum_{l_1l_2}j_{l_1}(kR_1)j_{l_2}(kR_2)\times[\mathbf{Y}^{[l_1]}(\hat{R}_1)\times\mathbf{Y}^{[l_2]}(\hat{R}_2)]^{[l]}\mathcal{l}_{l_1l_2;l}, \quad (45)$$

where

$$\mathcal{l}_{l_1l_2;L}=(-)^{l_2}t_{l_1l_2;L}. \quad (46)$$

We now couple the tensorial part, (44), to both sides of (45) to obtain:

$$\begin{aligned}j_l(kr)[\mathbf{Y}^{[l]}(\hat{r})\times\hat{\mathbf{e}}^{[S]}]^{[J]}_M &= \sum j_{l_1}(kR_1)j_{l_2}(kR_2)\\ &\times\big[[\mathbf{Y}^{[l_1]}(\hat{R}_1)\times\mathbf{Y}^{[l_2]}(\hat{R}_2)]^{[l]}\\ &\times[\hat{\mathbf{e}}^{[S_1]}\times\hat{\mathbf{e}}^{[S_2]}]^{[S]}\big]^{[J]}_M\mathcal{l}_{l_1l_2;l}\\ &=\sum j_{l_1}(kR_1)j_{l_2}(kR_2)\big[[\mathbf{Y}^{[l_1]}(\hat{R}_1)\times\hat{\mathbf{e}}^{[S_1]}]^{[J_1]}\\ &\times[\mathbf{Y}^{[l_2]}(\hat{R}_2)\times\hat{\mathbf{e}}^{[S_2]}]^{[J_2]}\big]^{[J]}_M\\ &\times\langle(l_1S_1)J_1,(l_2S_2)J_2\mid(l_1l_2)l,(S_1S_2)S\rangle^{(J)}\mathcal{l}_{l_1l_2;l}. \end{aligned}\quad (47)$$

Thus, the expansion coefficients of (42) are:

$$A^{(S_1,S_2;S)}_{l_1J_1,l_2J_2;lJ}=i^{l_1-l_2+l}\times[4\pi(2l_1+1)(2l_2+1)(2l+1)(2J_1+1)(2J_2+1)]^{\frac{1}{2}}\times\begin{pmatrix}l_1 & l_2 & l\\ 0 & 0 & 0\end{pmatrix}\begin{Bmatrix}l_1 & l_2 & l\\ S_1 & S_2 & S\\ J_1 & J_2 & J\end{Bmatrix}. \quad (48)$$

APPENDIX A

The theorem by Weyl[22] which contains the proof of (32) is more general than needed here. We are going to prove a simplified version of it which is sufficient for our purposes.

The transformation (29′) consists of two distinct steps. The first step consists in changing the coordinate system (θ,φ) into a rotated coordinate system (Euler angles $\alpha,\beta,0$) in which the polar axis points into the direction $\mathbf{r}$, i.e., (α,β) are the polar coordinates of $\mathbf{r}$ in the (θ,φ) system (we have changed the notation here for convenience). We thus write

$$\int_0^{2\pi}d\varphi\int_S d\cos\theta\,Y_{lm}(\theta,\varphi)e^{i\mathbf{k}\cdot\mathbf{r}}=\sum_{m'}\mathfrak{D}^{(l)}_{m'm}(0,\beta,\alpha)\times\int_{L'}d\psi\int_{S'}d\cos\eta\,Y_{lm'}(\eta,\psi)e^{ikr\cos\eta}. \quad (A1)$$

This step is valid generally since it merely expresses a transformation into a rotated system of coordinates. As long as α and β are real the functions $\mathfrak{D}$ are regular and one may safely embark on an analytical continuation of the angles φ,θ,ψ,η into the complex domain. The new regions of integration L' and S' are given by expressing the old region of integration in terms of the new variables.

The essential second step consists in changing the

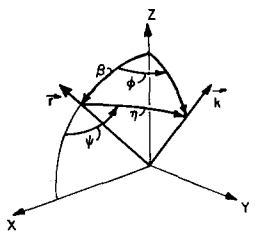


FIG. 4. Definition of the relevant angles.

region of integration back to the original region, i.e., in replacing the paths L' and S' by L (in which the variable goes from 0 to 2π) and S.

For simplicity (evidently without loss of generality) we put $\alpha = 0$. There then hold (see Fig. 4)

$$\cos\theta = \cos\beta\cos\eta - \sin\beta\sin\eta\cos\psi \qquad \text{(A2)}$$

$$\sin\theta\cos\varphi = \sin\beta\cos\eta + \cos\beta\sin\eta\cos\psi \qquad \text{(A3)}$$

$$\sin\theta\sin\varphi = \sin\eta\sin\psi \qquad \text{(A4)}$$

$$\cos\eta = \cos\beta\cos\theta + \sin\beta\sin\theta\cos\varphi \qquad \text{(A5)}$$

$$\sin\eta\cos\psi = -\sin\beta\cos\theta + \cos\beta\sin\theta\cos\varphi. \qquad \text{(A6)}$$

We take S to be the path shown in Fig. 5. Then the total region of integration in θ and φ can be depicted as in Fig. 6, where we have supplemented the cos θ plane by a third axis which is used to show the variable φ. The real region $0 \leq \varphi \leq 2\pi$, $0 \leq \cos\theta \leq 1$, will map into a real region in η, ψ, which is shown in Fig. 7 where the grid θ = const and φ = const is also indicated. The singular curve $\theta = \beta$ is indicated by a heavy line. The rest of the region of integration, i.e., where cos θ varies between 0 and $i\infty$, can not be shown as simply since there ψ is complex. We thus have to use two complex planes there. In Fig. 8 and Fig. 9 we have shown the variations of ψ and cos η if we keep θ fixed at some value $\theta = \frac{1}{2}\pi - i\nu$, so that $\cos\theta = i\mu$, and vary φ between 0 and 2π, as obtained in a straightforward manner from (A2)–(A6). {The characteristics of the curve of Fig. 8 can be most easily obtained from the parametric representation

$$\tan(2\,\mathrm{Re}\,\psi) = 2\sin\varphi\cos\varphi\cos\beta \times [\cos^2\beta + \sin^2\beta\tanh^2\nu + \sin^2\varphi(1+\cos^2\beta)]^{-1}$$

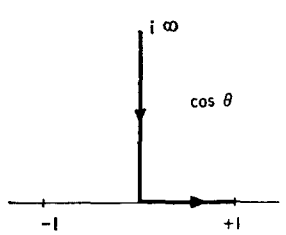


FIG. 5. Path of integration in the cos θ plane.

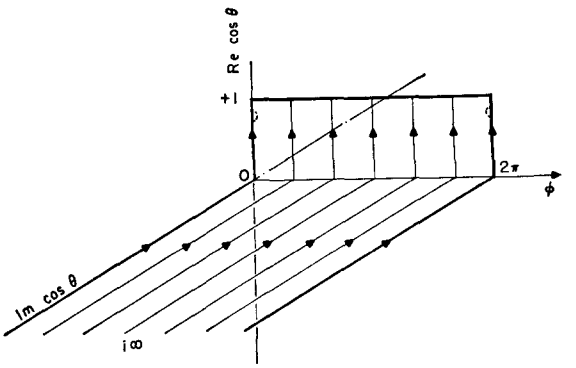


FIG. 6. φ–cos θ integration region.

$$\tanh(2\,\mathrm{Im}\,\psi) = 2\sin\varphi\tanh\nu\sin\beta \times [\cos^2\beta + \sin^2\beta\tanh^2\nu + \sin^2\varphi(1-\cos^2\beta)]^{-1};$$

artanh ≡ inverse hyperbolic tangent.}

We recall first that asymptotically η has to keep out of the crosshatched regions in Fig. 10 in order that (A1) converge at the limit $i\infty$. One sees from (A5) that cos η changes linearly between cos η = cos $(\theta + \beta)$ and cos η = cos $(\theta - \beta)$ if cos φ changes between $+1$ and -1. Choosing the branch of the arccos function indicated in Fig. 10 so that $\eta \to \theta$ and $\psi \to \varphi$ for $\beta \to 0$, we see that this implies that $\beta < \pi/2$. We require this to be true. Then in the mapping of Fig. 6 the singular curve $\theta = \beta$ will have the qualitative features as shown and the mapping covers in a one-to-one fashion a continuous region in the real ψ–η plane. We now want to change the paths of integration in the variables ψ and η. First we want to let ψ become real, and then we want to change, for any fixed ψ, the path in the cos η

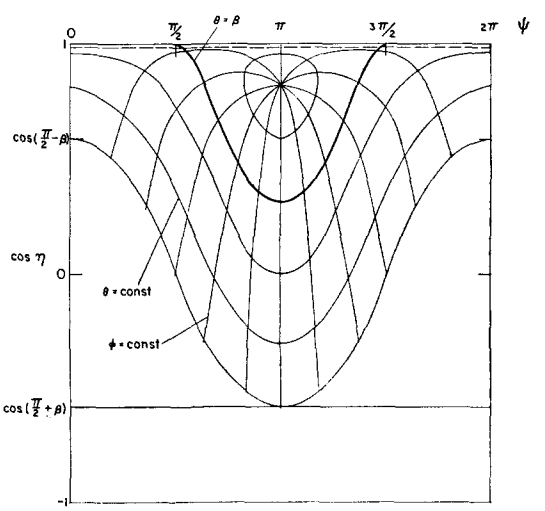


FIG. 7. Projection of the real integration region of Fig. 6 onto the ψ–cos η plane.

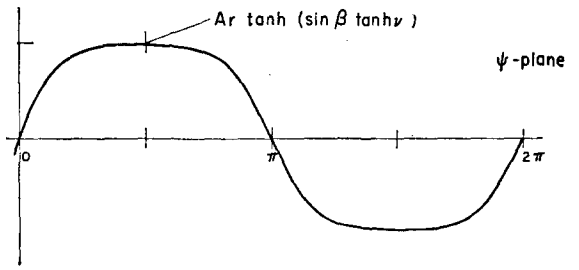


FIG. 8. Variation of ψ as function of φ for a fixed $\cos\theta = i\mu$.

integration to the path of Fig. 5. In order to be able to do this two conditions must be satisfied: (a) the integration endpoints must remain fixed, and (b) one must not sweep over any singularities while changing the path of integration. We first consider the requirement (a).

In the ψ integration the endpoints do not pose any problem: we have $\psi = \varphi$ for $\varphi = n\pi$, n being integer. In the $\cos\eta$ integration one can perform the usual rearrangement of the integrations in the real region of integration, for $0 \leq \varphi \leq 2\pi$, $0 \leq \theta \leq \pi/2$, i.e., we can integrate so that the endpoint lies at $\cos\eta = 1$. The other endpoint of the $\cos\eta$ integration can be fixed at $\cos\eta = i\infty$ since it lies within the convergent uncrosshatched region of Fig. 10. We have thus succeeded in rewriting the integrations so that the endpoints in the variables η and ψ are now the same as those in the variables θ and φ, i.e., requirement (a) is fulfilled.

In order to determine whether we can shift the paths of integration we have to find the location of the singularities of the integrand of the right-hand side of (A1). The singularities are: branchpoints at $\cos\eta = \pm 1$, and an essential singularity at $|\cos\eta| = \infty$. The last singularity has been taken care of by requiring that the integration path lie within the noncrosshatched region of Fig. 10. Further, we see (Fig. 9) that the condition $\beta < \pi/2$ also insures that the integration paths always stay away from the branchpoints $\cos\eta = \pm 1$, except at the endpoint of the integration which lies in the real integration region and where we have already specified on which branch

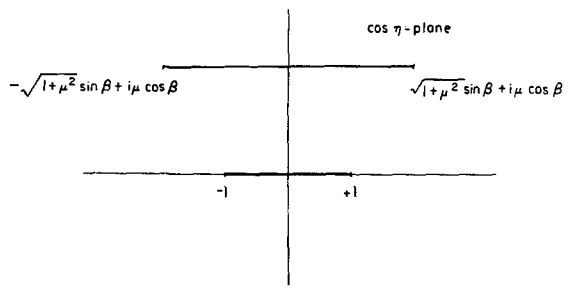


FIG. 9. Variation of $\cos\eta$ as function of φ for a fixed $\cos\theta = i\mu$.

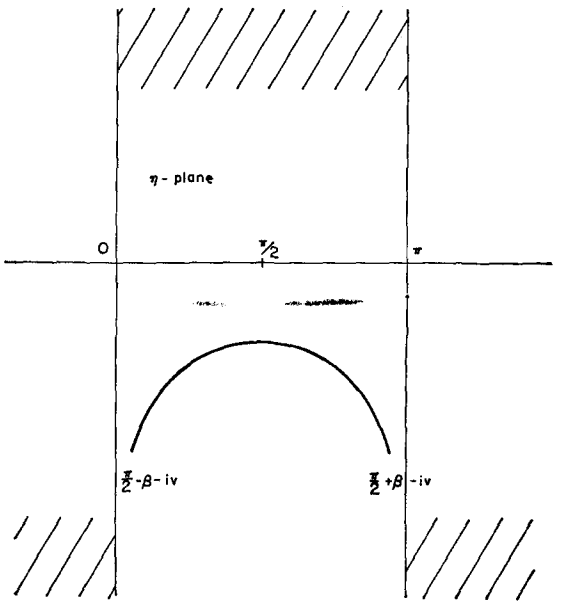


FIG. 10. Variation of η as function of φ for a fixed $\cos\theta = i\mu$, $\theta = \pi/2 - i\nu$.

of the function we should stay. To be totally precise, one has to consider the limit $\rho \to 0$ of an integration with the endpoint $\cos\eta = 1$ replaced by $\cos\eta = 1 - \rho$ (dotted line in Fig. 7). This would mean that in the region of integration in Fig. 6 the two little halfcircles with radius $r = \rho\cos\beta$ at $\varphi = 0$ and 2π and at $\cos\theta = \cos\beta$ (dotted halfcircles in Fig. 6) would have to be left out before going to the limit $\rho = 0$. One sees that this does not lead to any complications and coincides with the usual procedure one has to follow when one integrates over the whole surface of the real unit sphere in terms of a rotated system of coordinates. We thus see that condition (b) also is fulfilled in our case and that the essential second step is fully justified. We can therefore write

$$\int_{L'} d\psi \int_{S'} d\cos\eta\, Y_{lm'}(\eta, \psi) e^{ikr\cos\eta}$$
$$= \int_0^{2\pi} d\psi \int_{i\infty}^{1} d\cos\eta\, Y_{lm'}(\eta, \psi) e^{ikr\cos\eta}$$
$$= 2\pi\delta_{m'0} \int_{i\infty}^{1} d\cos\eta\, Y_{l0}(\eta, \psi) e^{ikr\cos\eta}, \qquad \text{(A7)}$$

which we can then insert in (A1) to obtain finally

$$\int_0^{2\pi} d\varphi \int_{i\infty}^{1} d\cos\theta\, Y_{lm}(\theta, \varphi) e^{i\mathbf{k}\cdot\mathbf{r}}$$
$$= 2\pi Y_{lm}(\beta, \alpha) \int_{i\infty}^{1} d\cos\eta P_l(\cos\eta) e^{ikr\cos\eta}$$
$$= 2\pi i^l Y_{lm}(\beta, \alpha) h_l^{(1)}(kr). \qquad \text{(A8)}$$

A completely analogous proof holds for the complementary path leading to $h_l^{(2)}$. In (A8) we have

dropped the assumption $\alpha = 0$, which obviously does not influence the proof.

APPENDIX B

In the expansion

$$h_l^{(\alpha)}(kr)\,Y_{lm}(\theta,\varphi) = \sum_L h_L^{(\alpha)}(kR)\,Y_{Lm}(\Theta,\varphi)T_{Ll}^{(m)}(k\rho) \qquad \alpha = 1, 2 \tag{B1}$$

one has to impose certain restrictions on r, R, and ρ. These restrictions come about in the following manner.

In Appendix A it was shown that the integrals (29′) and (30′) are defined as long as the angle between the polar axis and the direction of $\mathbf{r}$ is less than $\pi/2$. We now consider (27), i.e.,

$$\int d\varphi'\, d\cos\theta' Y_{lm}(\theta',\varphi')e^{i\mathbf{k}\cdot\mathbf{r}} = \int d\varphi'\, d\cos\theta' \times \sum_{LM} i^{l-L}Y_{LM}(\theta',\varphi')S_{LM,lm}(k\rho,\eta,\psi)e^{i\mathbf{k}\cdot\mathbf{R}}. \tag{B2}$$

One will be most economical in rotations if one rotates the coordinate system so that the new polar axis coincides with the direction of either $\mathbf{r}$ or $\mathbf{R}$ *before* deciding on the integration region. Then one will have to perform the transformation discussed in Appendix A only for an angle β equal to the angle between the directions of $\mathbf{R}$ and $\mathbf{r}$. The expansion thus obtained will be valid as long as one can interchange the order of summation and integration in (B2). This is allowed if the resulting series converges uniformly.

The first restriction is thus the requirement that the angle between the directions of R and $\mathbf{r}$ be less than $\pi/2$. However, the requirement of the convergence of (B1) yields a more stringent restriction. We thus investigate the magnitude of the terms for large L, i.e., $L \to \infty$.

A term of the series (B1) written out is:

$$A_L = h_L^{(\alpha)}(kR)\,Y_{Lm}(\Theta,\varphi)T_{Ll}^{(m)}(k\rho) = \sum_\lambda h_L^{(\alpha)}(kR)\,Y_{Lm}(\Theta,\varphi)i^{\lambda+L-l}(-1)^m \times (2\lambda+1)[(2L+1)(2l+1)]^{\frac{1}{2}} \times \begin{pmatrix} l & L & \lambda \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} l & L & \lambda \\ -m & m & 0 \end{pmatrix} j_\lambda(k\rho) \equiv \sum_\lambda A_L^{(\lambda)}. \tag{B3}$$

Because of the triangular conditions, for a given l the sum over λ in (B3) will, for arbitrarily large L, have only l terms. We thus have

$$|A_L| \le l\,|(A_L^{(\lambda)})_{\max}|. \tag{B4}$$

We now turn to the individual factors in $A_L^{(\lambda)}$. The behavior of the 3-j symbols follows from their explicit form:

$$\begin{pmatrix} l & L & \lambda \\ m & -m & 0 \end{pmatrix} = [(l+L-\lambda)!(L+\lambda-l)!(\lambda+l-L)!]^{\frac{1}{2}}[(l+L+\lambda+1)!]^{-\frac{1}{2}} \times \sum_p \frac{(-1)^{2L+\lambda+p}[(l+m)!(l-m)!(L+m)!(L-m)!\lambda!\lambda!]^{\frac{1}{2}}}{p!(l+L-\lambda-p)!(l-m-p)!(\lambda-L+m+p)!(L-m-p)!(\lambda-l+m+p)!}. \tag{B5}$$

This sum is again limited to a small number of terms, namely, at most to l terms. Equation (B5) can be estimated by approximating the factorials which have an argument of order L by Sirling's formula. In this way one obtains in a straightforward manner

$$\left|\begin{pmatrix} l & L & \lambda \\ m & -m & 0 \end{pmatrix}\right| = \frac{c(l,m)}{(L)^{\frac{1}{2}}}\left[1 + O\left(\frac{1}{L}\right)\right] \qquad \text{for} \quad L \gg l,\ l = O(1), \tag{B6}$$

where $c(l, m) = O(1)$, which is to be expected from the relation of the 3-j symbols to the Wigner coefficients, the latter being orthonormal.

The spherical harmonics remain bounded:

$$|Y_{Lm}(\theta,\varphi)| = O(1)[1 + O(1/L)] \qquad \text{for} \quad L \gg 1,\ m = O(1), \tag{B7}$$

a result also to be expected since the Y's are normalized. Finally, we have

$$j_\lambda(z) = \frac{1}{2}\left(\frac{z}{2}\right)^\lambda \frac{\pi^{\frac{1}{2}}}{\Gamma(\lambda+1+\frac{1}{2})}\left[1 + O\left(\frac{1}{\lambda}\right)\right] \quad \text{for } \lambda \gg z, \tag{B8}$$

$$|h_L(z)| = \frac{1}{2}\left(\frac{z}{2}\right)^{-L-1} \frac{\pi^{\frac{1}{2}}}{\Gamma(\frac{1}{2}-L)}\left[1 + O\left(\frac{1}{L}\right)\right] \qquad \text{for} \quad L \gg z. \tag{B9}$$

According to (B4) we have to take the largest possible term. This leads to $\lambda = L - l$. Inserting (B6), (B7), and (B8) in (B3) we obtain

$$A_L = O[(\rho/R)^L L^{l-\frac{1}{2}}]$$

which for $\rho/R < 1$ decreases faster than any power of L, for $\rho/R > 1$ goes to infinity. Therefore the series (B1) converges for $\rho/R < 1$.

The equivalent series for j_l instead of h_l converges for all values of ρ/R.

Variational Principle for Saturated Magnetoelastic Insulators

H. F. TIERSTEN
Bell Telephone Laboratories, Incorporated, Murray Hill, New Jersey
(Received 25 September 1964)

In a previous paper, a system of nonlinear differential equations and boundary conditions governing the macroscopic behavior of arbitrarily anisotropic nonconducting magnetically saturated media undergoing large deformations, was derived. The derivation utilized the classical procedure of defining field vectors and determining the equations relating them by applying energy and momentum conservation theorems. The macroscopic effect of the quantum mechanical exchange interaction was included as was dissipation and the associated thermodynamics. The magnetic field was assumed to be quasistationary. In this paper a variational principle is presented, which is shown to yield the aforementioned system of equations and boundary conditions in the absence of dissipation and heat flow.

I. INTRODUCTION

IN a previous paper[1] (hereafter referred to as I), the differential equations and boundary conditions governing the macroscopic behavior of arbitrarily anisotropic nonconducting magnetically saturated media undergoing large deformations were derived by means of a systematic and consistent application of the laws of continuum physics to a model consisting of an electronic spin continuum coupled to a lattice continuum. The macroscopic effect of the quantum-mechanical exchange interaction was included as was dissipation and the associated thermodynamics. As usual in such derivations, certain basic interacting field vectors were defined and the equations relating them were determined by means of an application of the well-known momentum and energy conservation theorems and irreversible thermodynamic principles. The precise definitions of all the variables with detailed discussion pertaining to them and the equations relating them are, of course, contained in I; and consequently are not presented here. Nevertheless, this paper is reasonably self-contained.

Now, it is commonly agreed that it is both interesting and useful to have a variational principle which reproduces an entire system of equations which were obtained previously in the alternative manner of defining field variables and applying conservation theorems. It is interesting from the standpoint of having a single statement embodying the entire theory in a manner which exhibits energies of interaction; and useful for the purpose of obtaining approximate solutions of the equations or generalizing the theory.

In this paper a variational principal is presented which yields the system of equations—minus dissipation and heat flow—previously derived in I in the aforementioned straightforward manner. Both the differential equations and the boundary conditions at material surfaces of discontinuity are obtained from this variational formulation.

[1] H. F. Tiersten, J. Math. Phys. 5, 1298 (1964).

2. PRELIMINARY CONSIDERATIONS

In order that this paper be somewhat self-contained we must—before proceeding with the variational principle—briefly repeat certain definitions and conditions employed in I. The symbolism we use is identical with that of I. Let x_i denote the Cartesian components of a material particle at some reference time t_0, and y_i the components of the same particle at some arbitrary time t. As in I, the x_i are referred to as material coordinates, and the y_i as spatial coordinates. The deformation (or motion) of the body is described by the mapping[2]

$$y_i = y_i(x_k, t), \tag{2.1}$$

which is one-to-one and differentiable as often as required except possibly at some singular points, curves, and surfaces. Let ρ_0 be the mass density in the reference configuration, dV an element of deformed volume, and dV_0 the corresponding element of undeformed volume. Then the conservation of mass takes the form

$$\rho dV = \rho_0 dV_0, \tag{2.2}$$

for any differential element of matter. This will be an important constraint in the variation. In view of the well-known geometric relation

$$J dV_0 = dV, \tag{2.3}$$

[2] For a clear discussion of deformation theory see Ref. 3, Secs. 13, 15, and 16.

[3] C. Truesdell and R. A. Toupin, "The Classical Field Theories" in *Encyclopedia of Physics*, edited by S. Flügge (Springer-Verlag, Berlin, 1960), Vol. III/I.

where J is the Jacobian of the transformation (2.1), and

$$J = \det ||\partial y_i/\partial x_i|| > 0, \tag{2.4}$$

the conservation of mass may be written in the form

$$\rho J = \rho_0. \tag{2.5}$$

At each material point y_i the continuum possesses a magnetic moment per unit mass μ_i, which is related to the ordinary magnetization vector M_i by the relation

$$M_i = \rho\mu_i. \tag{2.6}$$

Since the material is magnetically saturated and the mass is conserved, the magnitude of the magnetic moment per unit mass is conserved, which gives us the important relation

$$\mu_k\mu_k = \mu_S^2, \tag{2.7}$$

where μ_S is constant in a homogeneous material. Taking the material gradient of (2.7), we obtain

$$\mu_k\partial\mu_k/\partial x_m = 0. \tag{2.8}$$

Equations (2.7) and (2.8) will be important constraints in the variation. As a result of the existence of electron spin, we have associated with the magnetization vector $\mathbf{M}$ at each material point a volume density of angular momentum $\mathbf{J}$, given by

$$\mathbf{J} = \gamma^{-1}\rho\boldsymbol{\mu}, \tag{2.9}$$

where γ is the gyromagnetic ratio and is constant for a particular material. Also, at each point of space y_i, we have the magnetic scalar potential φ, from which the Maxwell magnetic field vector $\mathbf{H}^M$ may be determined from the relation

$$H_i^M = -\partial\varphi/\partial y_i. \tag{2.10}$$

Thus, it is clear that electromagnetic propagation is excluded and here, as in I, the magnetic field is quasistationary. Naturally, $\boldsymbol{\mu}$ and $\mathbf{H}^M$ are axial vectors and φ is an axial scalar. It should be noted that the magnetic scalar potential φ exists at all points of space, even those not occupied by matter; whereas $\boldsymbol{\mu}$ exists only at those points of space occupied by matter. The mapping (2.1), however, is defined only for those points occupied by matter since the material particle is preserved in the mapping. In the points of space not occupied by matter there are no material particles, and for those points the mapping becomes meaningless, and with the exception of boundary points, the y_i take the place of the x_i in matter and remain fixed in an *actual* motion but they may be subject to a variation in a *virtual* motion.

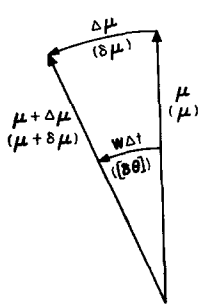


FIG. 1. Incremental motion of magnetization vector.

The velocity v_i of a material particle is given by

$$v_i = dy_i/dt = \partial y_i(x_k, t)/\partial t, \tag{2.11}$$

where d/dt denotes the material time derivative. Similarly, the time rate of change of magnetic moment per unit mass is given by

$$d\mu_i/dt = \partial\mu_i(x_k, t)/\partial t. \tag{2.12}$$

Now it should be observed that whereas any value of $\mathbf{v}$ is kinematically possible, only those values of $d\boldsymbol{\mu}/dt$ are possible which are consistent with (2.7), i.e., which satisfy the relation

$$\mu_i d\mu_i/dt = 0. \tag{2.13}$$

Consideration of this fact reveals that $\boldsymbol{\mu}$ may be instantaneously translating with velocity $\mathbf{v}$ and rotating with angular velocity $\mathbf{w}$. The angle $\mathbf{w}\Delta t$, through which $\boldsymbol{\mu}$ turns in time Δt is shown in Fig. 1. From Fig. 1 it is clear that

$$\mathbf{w}\Delta t = \mu_S^{-2}\boldsymbol{\mu}\times\Delta\boldsymbol{\mu}. \tag{2.14}$$

Now, the angular velocity $\mathbf{w}$ is a so-called "kinematical (or nonholonomic) vector," i.e., a vector which is not the time derivative of an actual vector function. Under these circumstances, $[\boldsymbol{\Delta\theta}] = \mathbf{w}\Delta t$ exists, but is merely an infinitesimal vectorial change in angle and not the differential of a vector function (see Ref. 4). It is well known that the negative of the time rate of change of the spin angular momentum $\gamma^{-1}\rho\boldsymbol{\mu}dV$ may be regarded as a d'Alembertian inertial couple. The rate at which work is done by this couple in an *actual* motion vanishes by virtue of the relation

$$\frac{d}{dt}\left(\frac{1}{\gamma}\rho\boldsymbol{\mu}\,dV\right)\cdot\mathbf{w} = \frac{1}{\gamma\mu_S^2}\rho\,dV\,\frac{d\boldsymbol{\mu}}{dt}\cdot\boldsymbol{\mu}\times\frac{d\boldsymbol{\mu}}{dt} = 0. \tag{2.15}$$

In other words, as the spin angular momentum changes in time, the energy associated with it does not. As a consequence of the previous considerations,

[4] C. Lanczos, *The Variational Principles of Mechanics*, (University of Toronto Press, Toronto, 1949), Chap. IV, Sec. 2.

the *virtual* work done by the inertial couple of the time rate of change of spin angular momentum will be computed directly in the variational principle and not determined as a variation of an energy function.

In the variational principle it will be important to remember that the variables x_i and t are held fixed in the variations considered for points in the material body, and the actual point and t are held fixed for points in free space where x_i has no meaning. However, it must be remembered that those points of free space which abut the boundary of a material region, are subject to the same variation of position as the material points of the boundary, i.e., the δy_i. That is why the positions of the points of free space must be subject to a variation. As usual, the variations are virtual (as opposed to actual) changes in the variables which are consistent with the constraints. In the present case there are three constraint conditions for points in the material body, which are given by (2.2), (2.7), and (2.8), respectively. To assure that these constraints are not violated, the variations must be such that

$$\delta(\rho dV) = 0, \tag{2.16a}$$

$$\mu_k \delta\mu_k = 0, \tag{2.16b}$$

$$\mu_k \delta(\partial\mu_k/\partial x_m) + (\partial\mu_k/\partial x_m)\delta\mu_k = 0. \tag{2.16c}$$

In a variation $\delta\mathfrak{y}$, $\mathfrak{y}$ turns through an angle $[\delta\theta]$, which is given by

$$[\delta\theta] = \mu_S^{-2}\mathfrak{y} \times \delta\mathfrak{y}, \tag{2.17}$$

as shown in Fig. 1. Naturally, the vector quantity $[\delta\theta]$ is an infinitesimal nonholonomic vector, and not the variation of an actual vector function, just as is $\mathbf{w}\Delta t$. Although the work done by the d'Alembertian inertial couple $-d(\gamma^{-1}\rho\mathfrak{y}dV)/dt$ in an actual motion vanishes, in a virtual motion it need not. The virtual work δW done by the inertial couple in an arbitrary variation $[\delta\theta]$ is given by

$$\begin{aligned}\delta W &= -\gamma^{-1}[d(\rho\mathfrak{y}\, dV)/dt]\cdot[\delta\theta] \\ &= -\gamma^{-1}\mu_S^{-2}\rho\, dV\,(d\mathfrak{y}/dt)\cdot\mathfrak{y}\times\delta\mathfrak{y}.\end{aligned} \tag{2.18}$$

From (2.17), with the help of (2.16b), we obtain

$$\delta\mathfrak{y} = -\mathfrak{y}\times[\delta\theta], \tag{2.19a}$$

or, in Cartesian tensor notation

$$\delta\mu_l = -e_{lmi}\mu_m[\delta\theta]_i. \tag{2.19b}$$

In Secs. 5 and 6 of I, it is shown that there exists at each material point a free energy density F which, in this nondissipative case, corresponds to an internal energy per unit mass U, of the form

$$U = U(\partial y_i/\partial x_m,\ \mu_i,\ \partial\mu_i/\partial x_m). \tag{2.20}$$

In addition, at each material point there exists a kinetic energy per unit mass T given by

$$T = \tfrac{1}{2}v_k v_k. \tag{2.21}$$

3. THE VARIATIONAL PRINCIPLE

Before formulating the variational principle we define the Lagrangian function

$$\begin{aligned}\mathfrak{L} = \int_{\mathfrak{V}} \left(\rho(T-U) - \frac{\partial\varphi}{\partial y_k}M_k + \frac{1}{8\pi}\frac{\partial\varphi}{\partial y_k}\frac{\partial\varphi}{\partial y_k}\right) dV \\ + \frac{1}{8\pi}\int_{\mathfrak{R}} \frac{\partial\varphi}{\partial y_k}\frac{\partial\varphi}{\partial y_k}\, dV,\end{aligned} \tag{3.1}$$

where $\mathfrak{V} = \sum_n V_n$ stands for the sum of all material regions and $\mathfrak{R}$ for free space. All variables appearing in (3.1) have already been defined in Sec. 2. The first term on the rhs of (3.1) represents the mechanical kinetic energy of the material body and the second term the internal (stored) energy of deformation, magnetization, and exchange. This latter term is quite analogous to the energy stored in a purely elastic solid. The spin kinetic energy is not contained in (3.1), since the variation of it vanishes by virtue of (2.7). The third term represents the interaction energy of the dipole moment with the magnetic field, the fourth term the magnetic field energy in the material regions, and the last term the magnetic field energy in free space. The variational principle may be written in the form

$$\begin{aligned}\delta\int_{t_0}^{t}\mathfrak{L}\, dt + \int_{t_0}^{t}\int_{\mathfrak{V}}\delta W\, dt + \int_{t_0}^{t}\int_{\mathfrak{V}}\Big[\lambda_k\delta\mu_k \\ + L_m\Big(\mu_k\delta\Big(\frac{\partial\mu_k}{\partial x_m}\Big) + \frac{\partial\mu_k}{\partial x_m}\delta\mu_k\Big)\Big] dV\, dt = 0,\end{aligned} \tag{3.2}$$

where λ and L_m are four Lagrangian undetermined multipliers which have been introduced to assure that the variations are consistent with the constraint conditions (2.16b, c) and δW is the virtual work done by the inertial couple of the time rate of change of the spin angular momentum, and is given by (2.18). {The third integral in (3.2) could equally readily have been combined with the first by defining an

$$\mathfrak{L}' = \mathfrak{L} + \int_{\mathfrak{V}} [\lambda(\mu_k\mu_k - \mu_S^2) + L_m\mu_k\,\partial\mu_k/\partial x_m]\, dV.$$

However, the procedure adopted seems more direct and preferable.} In (3.2) the variables subject to independent variations are y_i, μ_i, and φ for material points and y_i and φ for points in free space. The variations $\delta\mu_i$ and $\delta(\partial\mu_i/\partial x_m)$ in (3.2) may be treated

as if they are independent even though they are constrained by (2.16b, c) by virtue of the fact that the four conditions (2.16b, c) along with the proper number of undetermined multipliers have been introduced into (3.2) in the manner shown. As usual, the variations vanish at the endpoints of the interval, i.e., at t_0 and t.

In order to demonstrate that (3.2) does indeed yield the system of equations derived in *I*—minus dissipation and heat flow and containing *material* surfaces of discontinuity only—we shall need certain very useful relations which we now briefly introduce. First from (2.1) and the fact that it is one-to-one and continuously differentiable, we have

$$\frac{\partial}{\partial y_i} = \frac{\partial x_m}{\partial y_i}\frac{\partial}{\partial x_m}, \qquad \frac{\partial}{\partial x_i} = \frac{\partial y_m}{\partial x_i}\frac{\partial}{\partial y_m}. \tag{3.3a, b}$$

Since $\partial y_i/\partial y_j = \delta_{ij}$ and $\partial x_i/\partial x_j = \delta_{ij}$, where δ_{ij} is the Kronecker delta, we have from (3.3)

$$\frac{\partial y_i}{\partial x_m}\frac{\partial x_m}{\partial y_j} = \delta_{ij}, \qquad \frac{\partial x_i}{\partial y_m}\frac{\partial y_m}{\partial x_i} = \delta_{ij}. \tag{3.4a, b}$$

From the theory of 3×3 determinants (see Ref. 5)

$$e_{ijk}\frac{\partial y_i}{\partial x_l}\frac{\partial y_j}{\partial x_m}\frac{\partial y_k}{\partial x_n} = J e_{lmn}, \tag{3.5}$$

where e_{ijk} is the alternating tensor. From (3.5), with the aid of the well-known identity $e_{lmn}e_{lmn} = 6$, we find

$$J = \frac{1}{6}e_{ijk}e_{lmn}\frac{\partial y_i}{\partial x_l}\frac{\partial y_j}{\partial x_m}\frac{\partial y_k}{\partial x_n}. \tag{3.6}$$

Operating on (3.5) with e_{rmn}, using the well-known tensor identity $e_{lmn}e_{rmn} = 2\delta_{lr}$ and (3.4b), we obtain

$$\frac{1}{2}e_{ijk}e_{rmn}\frac{\partial y_j}{\partial x_m}\frac{\partial y_k}{\partial x_n} = J\frac{\partial x_r}{\partial y_i}. \tag{3.7}$$

We will also need the relations between the basic variations δy_i, $\delta\mu_i$, and $\delta\varphi$ and the variations of the other variables appearing in (3.2). Some immediate relations of this nature, arising from the fact that the δ-process commutes with partial differentiation with respect to variables held fixed in the variation, are

$$\delta\left(\frac{dy_i}{dt}\right) = \delta\left(\frac{\partial y_i(x_k, t)}{\partial t}\right) = \frac{\partial(\delta y_i)}{\partial t} = \frac{d(\delta y_i)}{dt}, \tag{3.8a}$$

$$\delta(\partial y_i/\partial x_m) = \partial(\delta y_i)/\partial x_m, \tag{3.8b}$$

$$\delta(\partial\mu_i/\partial x_m) = \partial(\delta\mu_i)/\partial x_m, \tag{3.8c}$$

$$\delta(\partial\varphi/\partial x_m) = \partial(\delta\varphi)/\partial x_m. \tag{3.8d}$$

[5] A. J. McConnell, *Applications of Tensor Analysis* (Dover Publications, Inc., New York, 1957), Sec. 6.

From (3.4b), with the aid of (3.4a) and (3.8b), we obtain

$$\delta\left(\frac{\partial x_m}{\partial y_k}\right) = -\frac{\partial x_m}{\partial y_l}\frac{\partial x_n}{\partial y_k}\frac{\partial}{\partial x_n}(\delta y_l). \tag{3.9}$$

Application of the δ process to (3.6) yields

$$\delta J = \frac{1}{2}e_{ijk}e_{lmn}\frac{\partial y_i}{\partial x_l}\frac{\partial y_j}{\partial x_m}\delta\left(\frac{\partial y_k}{\partial x_n}\right), \tag{3.10}$$

from which, with (3.7) and (3.8b), we obtain

$$\delta J = J(\partial x_n/\partial y_k)\,\partial(\delta y_k)/\partial x_n. \tag{3.11}$$

Since

$$\frac{\partial\varphi}{\partial y_k} = \frac{\partial x_m}{\partial y_k}\frac{\partial\varphi}{\partial x_m},$$

application of the δ process to $\partial\varphi/\partial y_k$ yields

$$\delta\left(\frac{\partial\varphi}{\partial y_k}\right) = \frac{\partial x_m}{\partial y_k}\delta\left(\frac{\partial\varphi}{\partial x_m}\right) + \frac{\partial\varphi}{\partial x_m}\delta\left(\frac{\partial x_m}{\partial y_k}\right), \tag{3.12}$$

from which, with (3.3a), (3.8b), (3.8d), and (3.9), we obtain

$$\delta\left(\frac{\partial\varphi}{\partial y_k}\right) = \frac{\partial(\delta\varphi)}{\partial y_k} - \frac{\partial\varphi}{\partial y_l}\frac{\partial(\delta y_l)}{\partial y_k}. \tag{3.13}$$

Although (3.13) has been derived for points in material regions, it is valid for points in free space also, since a virtual displacement of the position of these points is possible—even though an actual displacement is meaningless—by means of defining a virtual mapping analogous to (2.1), but not associated with a material particle.

We are now in a position to take the variation called for in the first term of (3.2). First, from (2.21) with (2.11), (2.16a), (3.8a) and an integration with respect to time, we obtain

$$\delta\int_{t_0}^{t} dt\int_{\mathcal{V}}\frac{1}{2}\rho\frac{dy_i}{dt}\frac{dy_i}{dt}\,dV = \int_{\mathcal{V}}\rho\,dV\left[\frac{dy_i}{dt}\,\delta y_i\right]_{t_0}^{t} - \int_{t_0}^{t} dt\int_{\mathcal{V}}\rho\frac{d^2y_i}{dt^2}\,\delta y_i\,dV,$$

from which, since δy_i vanishes at t_0 and t, we find

$$\delta\int_{t_0}^{t} dt\int_{\mathcal{V}}\frac{1}{2}\rho\frac{dy_i}{dt}\frac{dy_i}{dt}\,dV = -\int_{t_0}^{t} dt\int_{\mathcal{V}}\rho\frac{d^2y_i}{dt^2}\,\delta y_i\,dV. \tag{3.14}$$

Now, from (2.20), with (2.16a), and (3.8b, c), we note that

$$\delta\int_{t_0}^{t} dt\int_{\mathcal{V}}\rho U\,dV = \int_{t_0}^{t} dt\int_{\mathcal{V}}\rho\left(\frac{\partial U}{\partial(\partial y_i/\partial x_m)}\frac{\partial(\delta y_i)}{\partial x_m} + \frac{\partial U}{\partial(\partial\mu_i/\partial x_m)}\frac{\partial(\delta\mu_i)}{\partial x_m} + \frac{\partial U}{\partial\mu_i}\delta\mu_i\right)dV. \tag{3.15}$$

The first volume integral on the rhs of (3.15) yields

$$\int_v \rho \frac{\partial U}{\partial(\partial y_i/\partial x_m)} \frac{\partial(\delta y_i)}{\partial x_m} dV$$
$$= \int_v \rho \frac{\partial U}{\partial(\partial y_i/\partial x_m)} \frac{\partial y_i}{\partial x_m} \frac{\partial(\delta y_i)}{\partial y_i} dV$$
$$= \int_v \frac{\partial}{\partial y_i}\left(\rho \frac{\partial y_i}{\partial x_m} \frac{\partial U}{\partial(\partial y_i/\partial x_m)} \delta y_i\right) dV$$
$$- \int_v \frac{\partial}{\partial y_i}\left(\rho \frac{\partial y_i}{\partial x_m} \frac{\partial U}{\partial(\partial y_i/\partial x_m)}\right)\delta y_i \, dV.$$

Employing the divergence theorem, we obtain

$$\int_v \rho \frac{\partial U}{\partial(\partial y_i/\partial x_m)} \frac{\partial(\delta y_i)}{\partial x_m} dV$$
$$= \int_s n_i \rho \frac{\partial y_i}{\partial x_m} \frac{\partial U}{\partial(\partial y_i/\partial x_m)} \delta y_i \, dV$$
$$+ \int_{\mathfrak{G}} n_i \rho \frac{\partial y_i}{\partial x_m} \frac{\partial U}{\partial(\partial y_i/\partial x_m)} \delta y_i \, dS$$
$$- \int_v \frac{\partial}{\partial y_i}\left(\rho \frac{\partial y_i}{\partial x_m} \frac{\partial U}{\partial(\partial y_i/\partial x_m)}\right)\delta y_i \, dV,$$

where n_i denotes the Cartesian components of the outwardly directed unit normal across *a* surface of discontinuity, $\mathfrak{s}$ the sum of those surfaces of discontinuity which separate two distinct material regions, and $\mathfrak{G}$ the sum of those surfaces of discontinuity which separate material regions from free space. The sum $\mathfrak{s}$, of course, counts all such surfaces twice, once from each side. If we agree to denote one side of a material surface of discontinuity by a plus and the other by a minus and count the surface only once, we may write

$$\int_v \rho \frac{\partial U}{\partial(\partial y_i/\partial x_m)} \frac{\partial}{\partial x_m}(\delta y_i) \, dV$$
$$= \int_{s+} n_i^+\left[\rho \frac{\partial y_i}{\partial x_m} \frac{\partial U}{\partial(\partial y_i/\partial x_m)} \delta y_i\right] dS$$
$$+ \int_{\mathfrak{G}} n_i \rho \frac{\partial y_i}{\partial x_m} \frac{\partial U}{\partial(\partial y_i/\partial x_m)} \delta y_i \, dS$$
$$- \int_v \frac{\partial}{\partial y_i}\left(\rho \frac{\partial y_i}{\partial x_m} \frac{\partial U}{\partial(\partial y_i/\partial x_m)}\right)\delta y_i \, dV, \quad (3.16)$$

where n_i^+ denotes the Cartesian components of the unit normal to the surface of discontinuity directed from the + region to the − region, and we have introduced the conventional notation $[C_i]$ for $C_i^+ - C_i^-$. By means of the same procedure, the second volume integral on the rhs of (3.15) gives us

$$\int_v \rho \frac{\partial U}{\partial(\partial \mu_i/\partial x_m)} \frac{\partial}{\partial x_m}(\delta \mu_i) \, dV$$
$$= \int_{s+} n_i^+\left[\rho \frac{\partial y_i}{\partial x_m} \frac{\partial U}{\partial(\partial \mu_i/\partial x_m)} \delta \mu_i\right] dS$$
$$+ \int_{\mathfrak{G}} n_i \rho \frac{\partial y_i}{\partial x_m} \frac{\partial U}{\partial(\partial \mu_i/\partial x_m)} \delta \mu_i \, dS$$
$$- \int_v \frac{\partial}{\partial y_i}\left(\rho \frac{\partial y_i}{\partial x_m} \frac{\partial U}{\partial(\partial \mu_i/\partial x_m)}\right)\delta \mu_i \, dV. \quad (3.17)$$

The variation of the third term in (3.1), with the aid of (2.6), (2.16a), and (3.13), yields

$$\delta \int_v \frac{\partial \varphi}{\partial y_k} M_k \, dV$$
$$= \int_v \left(\mu_k \frac{\partial(\delta\varphi)}{\partial y_k} - \mu_k \frac{\partial \varphi}{\partial y_l} \frac{\partial(\delta y_l)}{\partial y_k} + \frac{\partial \varphi}{\partial y_k} \delta \mu_k\right)\rho \, dV$$
$$= \int_v \frac{\partial}{\partial y_k}\left(M_k \delta\varphi - M_k \frac{\partial \varphi}{\partial y_l} \delta y_l\right) dV$$
$$+ \int_v \left(-\frac{\partial M_k}{\partial y_k} \delta\varphi + \frac{\partial M_k}{\partial y_k} \frac{\partial \varphi}{\partial y_l} \delta y_l\right.$$
$$\left. + M_k \frac{\partial^2 \varphi}{\partial y_k \, \partial y_l} \delta y_l + \frac{\partial \varphi}{\partial y_k} \rho \delta \mu_k\right) dV,$$

from which, with the aid of the divergence theorem and the introduction of the notation agreements surrounding Eq. (3.16), we obtain

$$\delta \int_v \frac{\partial \varphi}{\partial y_k} M_k \, dV = \int_{s+} n_k^+\left[M_k \delta\varphi - M_k \frac{\partial \varphi}{\partial y_l} \delta y_l\right] dS$$
$$+ \int_{\mathfrak{G}} n_k\left(M_k \delta\varphi - M_k \frac{\partial \varphi}{\partial y_l} \delta y_l\right) dS$$
$$+ \int_v \left(-\frac{\partial M_k}{\partial y_k} \delta\varphi + \frac{\partial M_k}{\partial y_k} \frac{\partial \varphi}{\partial y_l} \delta y_l\right.$$
$$\left. + M_k \frac{\partial^2 \varphi}{\partial y_k \, \partial y_l} \delta y_l + \frac{\partial \varphi}{\partial y_k} \rho \delta \mu_k\right) dV. \quad (3.18)$$

For the variation of the fourth term in (3.1) with the aid of (2.3), (3.3a), (3.11), (3.13), the divergence theorem and the aforementioned notation conventions, we obtain

$$\delta \int_v \frac{1}{8\pi} \frac{\partial \varphi}{\partial y_k} \frac{\partial \varphi}{\partial y_k} dV$$
$$= \frac{1}{8\pi} \int_{s+} n_k^+\left[2 \frac{\partial \varphi}{\partial y_k} \delta\varphi - 2 \frac{\partial \varphi}{\partial y_k} \frac{\partial \varphi}{\partial y_l} \delta y_l + \frac{\partial \varphi}{\partial y_l} \frac{\partial \varphi}{\partial y_l} \delta y_k\right] dS$$
$$+ \frac{1}{8\pi} \int_{\mathfrak{G}} n_k\left(2 \frac{\partial \varphi}{\partial y_k} \delta\varphi - 2 \frac{\partial \varphi}{\partial y_k} \frac{\partial \varphi}{\partial y_l} \delta y_l + \frac{\partial \varphi}{\partial y_l} \frac{\partial \varphi}{\partial y_l} \delta y_k\right) dS$$
$$+ \frac{1}{8\pi} \int_v \left(-2 \frac{\partial^2 \varphi}{\partial y_k \, \partial y_k} \delta\varphi + 2 \frac{\partial^2 \varphi}{\partial y_k \, \partial y_k} \frac{\partial \varphi}{\partial y_l} \delta y_l\right) dV. \quad (3.19)$$

Similarly, the variation of the free space magnetic field energy yields

$$\delta \int_{\mathcal{R}} \frac{1}{8\pi} \frac{\partial \varphi}{\partial y_k} \frac{\partial \varphi}{\partial y_k} dV = \frac{1}{8\pi} \int_{\mathcal{B}} n_k \left(2 \frac{\partial \varphi}{\partial y_k} \delta\varphi - 2 \frac{\partial \varphi}{\partial y_k} \frac{\partial \varphi}{\partial y_l} \delta y_l + \frac{\partial \varphi}{\partial y_l} \frac{\partial \varphi}{\partial y_l} \delta y_k \right) dS + \frac{1}{8\pi} \int_{\mathcal{R}} \left(-2 \frac{\partial^2 \varphi}{\partial y_k \partial y_k} \delta\varphi + 2 \frac{\partial^2 \varphi}{\partial y_k \partial y_k} \frac{\partial \varphi}{\partial y_l} \delta y_l \right) dV, \tag{3.20}$$

in which δy_l in the surface integral is the δy_l of the boundary of the adjacent material region, δy_l in the volume integral is arbitrary, and the variations δy_l and $\delta\varphi$ in the surface integral over the boundary at infinity have already been assumed to vanish.

By means of a procedure similar to the one employed in going from the first term on the rhs of (3.15) to (3.16), we obtain

$$\int_{\mathcal{V}} L_m \mu_k \delta\left(\frac{\partial \mu_k}{\partial x_m}\right) \rho \, dV = \int_{\mathcal{S}^+} n_i^+ \left[\frac{\partial y_i}{\partial x_m} L_m \rho \mu_k \delta\mu_k \right] dS + \int_{\mathcal{B}} n_i \frac{\partial y_i}{\partial x_m} L_m \rho \mu_k \delta\mu_k \, dS - \int_{\mathcal{V}} \left[\frac{\partial}{\partial y_i} \left(\frac{\partial y_i}{\partial x_m} L_m \rho \right) \mu_k + L_m \rho \frac{\partial \mu_k}{\partial x_m} \right] \delta\mu_k \, dV. \tag{3.21}$$

Substituting first from (3.16) and (3.17) into (3.15), then from (2.18), (3.14), (3.15), and (3.18)–(3.21) into (3.2) while introducing the rhs of (2.19) for $\delta\mathfrak{y}$ wherever it occurs and recombining terms, we obtain

$$\begin{aligned}
&\int_{t_0}^{t} dt \int_{\mathcal{V}} \left[-\rho \frac{d^2 y_i}{dt^2} + \frac{\partial}{\partial y_i} \left(\rho \frac{\partial y_i}{\partial x_m} \frac{\partial U}{\partial (\partial y_i / \partial x_m)} \right) - M_k \frac{\partial^2 \varphi}{\partial y_k \partial y_i} + \frac{\partial \varphi}{\partial y_j} \left(\frac{1}{4\pi} \frac{\partial^2 \varphi}{\partial y_k \partial y_k} - \frac{\partial M_k}{\partial y_k} \right) \right] \delta y_i \, dV \\
&- \int_{t_0}^{t} dt \int_{\mathcal{V}} \left[e_{ikl} \mu_k \frac{\partial}{\partial y_i} \left(\rho \frac{\partial y_i}{\partial x_m} \frac{\partial U}{\partial (\partial \mu_i / \partial x_m)} \right) - e_{ikl} \mu_k \left(\rho \frac{\partial U}{\partial \mu_j} - \lambda' \rho \mu_i + \rho \frac{\partial \varphi}{\partial y_j} \right) + \frac{1}{\gamma} \rho \frac{d\mu_l}{dt} \right] [\delta\theta]_l \, dV \\
&+ \int_{t_0}^{t} dt \int_{\mathcal{V}} \left(\frac{\partial M_k}{\partial y_k} - \frac{1}{4\pi} \frac{\partial^2 \varphi}{\partial y_k \partial y_k} \right) \delta\varphi \, dV - \frac{1}{4\pi} \int_{t_0}^{t} dt \int_{\mathcal{R}} \frac{\partial^2 \varphi}{\partial y_k \partial y_k} \left(\delta\varphi - \frac{\partial \varphi}{\partial y_l} \delta y_l \right) dV \\
&- \int_{t_0}^{t} dt \int_{\mathcal{S}^+} n_i^+ \left[\left(\rho \frac{\partial y_i}{\partial x_m} \frac{\partial U}{\partial (\partial y_i / \partial x_m)} - M_i \frac{\partial \varphi}{\partial y_i} + \frac{1}{4\pi} \frac{\partial \varphi}{\partial y_i} \frac{\partial \varphi}{\partial y_i} - \frac{1}{8\pi} \frac{\partial \varphi}{\partial y_k} \frac{\partial \varphi}{\partial y_k} \delta_{ii} \right) \delta y_i \right] dS \\
&- \int_{t_0}^{t} dt \int_{\mathcal{B}} n_i^+ \left[\rho \frac{\partial y_i}{\partial x_m} \frac{\partial U}{\partial (\partial y_i / \partial x_m)} - M_i \frac{\partial \varphi^+}{\partial y_i} + \frac{1}{4\pi} \left[\frac{\partial \varphi}{\partial y_i} \frac{\partial \varphi}{\partial y_i} - \frac{1}{2} \frac{\partial \varphi}{\partial y_k} \frac{\partial \varphi}{\partial y_k} \delta_{ii} \right] \right] \delta y_i \, dS \\
&+ \int_{t_0}^{t} dt \int_{\mathcal{S}^+} n_i^+ \left[e_{ikl} \mu_k \left(\rho \frac{\partial y_i}{\partial x_m} \frac{\partial U}{\partial (\partial \mu_j / \partial x_m)} - L_m \rho \frac{\partial y_i}{\partial x_m} \mu_j \right) [\delta\theta]_l \right] dS \\
&+ \int_{t_0}^{t} dt \int_{\mathcal{B}} n_i^+ e_{ikl} \mu_k \left(\rho \frac{\partial y_i}{\partial x_m} \frac{\partial U}{\partial (\partial \mu_j / \partial x_m)} - L_m \rho \frac{\partial y_i}{\partial x_m} \mu_j \right) [\delta\theta]_l \, dS \\
&- \int_{t_0}^{t} dt \int_{\mathcal{S}^+} n_i^+ \left[\left(M_i - \frac{1}{4\pi} \frac{\partial \varphi}{\partial y_i} \right) \delta\varphi \right] - \int_{t_0}^{t} dt \int_{\mathcal{B}} n_i^+ \left(M_i \delta\varphi - \frac{1}{4\pi} \left[\frac{\partial \varphi}{\partial y_i} \delta\varphi \right] \right) dS = 0,
\end{aligned}$$

where

$$\lambda' = \lambda - \frac{1}{\rho} \frac{\partial}{\partial y_i} \left(\frac{\partial y_i}{\partial x_m} L_m \rho \right), \tag{3.22}$$

and we have taken the unit normal $\mathbf{n}^+$ on the boundary $\mathcal{B}$ between the material regions and free space to be pointing out of the material region. Then, for consistency with the system of notation previously introduced for the surfaces $\mathcal{S}^+$, the free space side of $\mathcal{B}$ is denoted by $-$ and the material side by $+$. Since we have introduced the proper number of undetermined multipliers in the usual Lagrangian manner, we may treat all variations appearing in (3.22) as if they are independent, even though the three components $[\delta\theta]_i$ are not. Consequently, if we make the identifications

$$\tau_{ij} = \rho \frac{\partial y_i}{\partial x_m} \frac{\partial U}{\partial (\partial y_i / \partial x_m)}, \tag{3.23a}$$

$$A_{ij} = -\frac{\partial y_i}{\partial x_m} \frac{\partial U}{\partial (\partial \mu_j / \partial x_m)}, \tag{3.23b}$$

$$H_i^{\mathrm{L}} = -\frac{\partial U}{\partial \mu_i} + \lambda' \mu_i, \tag{3.23c}$$

substitute from (2.10) and introduce the relation $B_k = H_k + 4\pi M_k$, we obtain the system of equations

$$\frac{\partial \tau_{ij}}{\partial y_i} + M_k \frac{\partial H_j^{\mathrm{M}}}{\partial y_k} = \rho \frac{d^2 y_j}{dt^2}, \tag{3.24a}$$

$$e_{kjl}\mu_k(-\partial(\rho A_{ij})/\partial y_i + \rho H_j^{\mathrm{L}} + \rho H_j^{\mathrm{M}}) = \gamma^{-1}\rho\, d\mu_l/dt, \tag{3.24b}$$

$$\partial B_k/\partial y_k = 0, \tag{3.24c}$$

at each material point, and

$$\partial H_k^{\mathrm{M}}/\partial y_k = 0, \tag{3.25}$$

at each point of free space; and the discontinuity conditions[6]

$$n_i^+\left[\left(\tau_{ij} + \frac{1}{4\pi}B_iH_j^{\mathrm{M}} - \frac{1}{8\pi}H_k^{\mathrm{M}}H_k^{\mathrm{M}}\delta_{ij}\right)\delta y_j\right] = 0,$$
$$n_i^+[e_{kjl}\mu_k\rho A_{ij}[\delta\theta]_l] = 0, \tag{3.26}$$
$$n_i^+[B_i\delta\varphi] = 0,$$

across discontinuity surfaces separating material regions, and

$$n_i^+(\tau_{ij} + M_iH_j^{\mathrm{M}^+}) + \tfrac{1}{4}\pi^{-1}n_i^+[H_i^{\mathrm{M}}H_j^{\mathrm{M}} - \tfrac{1}{2}H_k^{\mathrm{M}}H_k^{\mathrm{M}}\delta_{ij}] = 0, \tag{3.27a}$$

$$n_i^+e_{kjl}\mu_k\rho A_{ij} = 0, \tag{3.27b}$$

$$n_i^+M_i + \tfrac{1}{4}\pi^{-1}n_i^+[H_i^{\mathrm{M}}] = 0, \tag{3.27c}$$

across the discontinuity surface $\mathcal{S}$ separating material regions from free space. In obtaining (3.27c) from (3.22) we have assumed that the magnetic scalar potential φ is continuous across $\mathcal{S}$. If we further assume that across discontinuity surfaces separating material regions, adjacent regions are attached, the magnetic scalar potential is continuous and the variation in magnetization direction, given by the nonholonomic vector $[\delta\theta]$ is continuous, the variations in (3.26) can be factored out of the discontinuity brackets, and we obtain the boundary conditions

$$n_i^+[\tau_{ij} + \tfrac{1}{4}\pi^{-1}(B_iH_j^{\mathrm{M}} - \tfrac{1}{2}H_k^{\mathrm{M}}H_k^{\mathrm{M}}\delta_{ij})] = 0,$$
$$n_i^+[e_{kjl}M_kA_{ij}] = 0, \quad n_i^+[B_i] = 0, \tag{3.28}$$
$$[y_i] = 0, \quad [\mu_{\mathrm{S}}^{-2}e_{ijk}\mu_j\, d\mu_k/dt] = 0, \quad [\varphi] = 0,$$

across material surfaces of discontinuity. Equations (3.24a–c), (3.23a–c), and (3.28a–c) correspond, respectively, with Eqs. (8.1a), (8.1d), (8.1f), (6.14a), (6.14c), (6.14b), and (8.6a, b), (8.6d) of I. However, certain additional considerations are necessary to demonstrate that the variational principle presented in this paper reproduces the complete system of equations derived in I—minus dissipation and heat conduction. These considerations are discussed in the next section.

[6] At this stage it should be noted that λ' could have been eliminated just as readily as L_m, since it too yields a zero term wherever it appears. The reason for retaining λ' and not L_m will be made clear in the next section.

4. INVARIANCE REQUIREMENTS

As noted in Sec. 6 of I, the internal energy function U cannot be an arbitrary function of the variables listed on the rhs of (2.20) because it must be invariant under all rigid motions of the deformed and magnetized material. As shown in Sec. 6 of I, this requirement of invariance under rigid motions enables U to be reduced from the form shown in (2.20) to the form

$$U = U(C_{rs}, N_r, D_{rs}), \tag{4.1}$$

where

$$C_{rs} = (\partial y_k/\partial x_r)\partial y_k/\partial x_s,$$
$$N_r = (\partial y_k/\partial x_r)\mu_k, \tag{4.2}$$
$$D_{rs} = (\partial\mu_k/\partial x_r)\partial y_k/\partial x_s.$$

In Sec. 4 of I, a condition—Eq. (4.15)—was introduced so that the system of equations there derived was consistent with the saturation condition. It was further shown, in Sec. 6 of I, that this condition served to make the exchange energy invariant in a rigid rotation of the spin continuum. (The terms spin continuum and lattice continuum are discussed thoroughly in Sec. 2 of I.) It is well known, of course, that the exchange energy is so invariant in the quantum mechanical description.[7,8] It should be observed that in the macroscopic description employed here and in I, the exchange energy refers to all terms in U containing $\partial\mu_i/\partial x_j$. The above mentioned condition is not obtained as a direct consequence of the variational principle presented in this paper. However, we can show that Eq. (4.15) of I is satisfied identically if we require (as a physical principle now) the exchange energy to be invariant in a rigid rotation of the spin system with respect to the lattice configuration. In such a rotation the vectorial components μ_i and $\partial\mu_i/\partial x_m$ are transformed to μ_j', $\partial\mu_j'/\partial x_m$, respectively, while $\partial y_i/\partial x_m$ are held fixed, where

$$\mu_j' = R_{ji}\mu_i$$
$$\partial\mu_j'/\partial x_m = R_{ji}\,\partial\mu_i/\partial x_m, \tag{4.3}$$

and R_{ij} represents an orthogonal transformation. Since only the exchange terms—and not the magnetic terms—in the energy function must be in-

[7] J. H. Van Vleck, Rev. Mod. Phys. **17**, 27 (1945).
[8] C. Herring and C. Kittel, Phys. Rev. **81**, 869 (1951).

variant under the aforementioned rigid rotation, we have

$$U\left(\frac{\partial y_i}{\partial x_m}, \mu_i, \frac{\partial \mu_i}{\partial x_m}\right) \neq U'\left(\frac{\partial y_i}{\partial x_m}, \mu_i', \frac{\partial \mu_i'}{\partial x_m}\right), \tag{4.4a}$$

$$U\left(\frac{\partial y_i}{\partial x_m}, \mu_i, \frac{\partial \mu_i}{\partial x_m}\right) = U'\left(\frac{\partial y_i}{\partial x_m}, \mu_i, \frac{\partial \mu_i'}{\partial x_m}\right). \tag{4.4b}$$

Inasmuch as invariance under infinitesimal rotations is sufficient to insure invariance under finite rotations (this is discussed in Ref. 9; also see Ref. 10) it will suffice for us to consider an infinitesimal rigid rotation which is given by

$$R_{ij} = \delta_{ij} + \eta_{ij}, \tag{4.5}$$

where η_{ij} is infinitesimal and antisymmetric, and where (4.4b) is satisfied. Thus, we may write

$$U' - U = \Delta U$$

$$= \frac{\partial U'}{\partial(\partial \mu_i'/\partial x_m)} \frac{\partial(\partial \mu_i'/\partial x_m)}{\partial R_{kl}}\Bigg]_{R_{kl}=\delta_{kl}} (R_{kl} - \delta_{kl}) = 0,$$

from which, with (4.3) and (4.5), we obtain

$$\Delta U = \frac{\partial U}{\partial(\partial \mu_i/\partial x_m)} \frac{\partial \mu_l}{\partial x_m} \eta_{il} = 0. \tag{4.6}$$

Since the η_{il} constitute an *arbitrary* antisymmetric tensor, the coefficients of the η_{il} in (4.6) must be symmetric, and we obtain the condition

$$\frac{\partial U}{\partial(\partial \mu_i/\partial x_m)} \frac{\partial \mu_l}{\partial x_m} = \frac{\partial U}{\partial(\partial \mu_l/\partial x_m)} \frac{\partial \mu_i}{\partial x_m}. \tag{4.7}$$

Any arbitrary function U of **C**, **N**, and **D** will not necessarily satisfy (4.7). To find the additional restrictions on U which are engendered by (4.7), we first note that

$$\frac{\partial U}{\partial(\partial \mu_i/\partial x_m)} = \frac{\partial U}{\partial D_{ms}} \frac{\partial y_i}{\partial x_s}, \tag{4.8}$$

and that

$$\partial \mu_k/\partial x_m = D_{mn}(\partial x_n/\partial y_k), \tag{4.9}$$

and then substitute from (4.8) and (4.9) into (4.7), while employing (3.4b) and introducing the symbol $\mathbf{C}^{-1}$ for the reciprocal of **C**, with the result

$$(\partial U/\partial D_{mq}) D_{mn} C_{nr}^{-1} = (\partial U/\partial D_{mr}) D_{mn} C_{nq}^{-1}. \tag{4.10}$$

These comprise a system of three independent differential equations in the 18 variables **C**, **N**, and **D**, which must be satisfied by U. Consequently, U must reduce to an arbitrary function of any $18 - 3 = 15$ functionally independent solutions of (4.10), which must be composed of **C**, **N**, and **D**. It is obvious that **C** and **N** constitute nine such solutions. Six additional solutions are given by $D_{ia} C_{ab}^{-1} D_{jb} \equiv \Gamma_{ij}$, as may be verified by first noting that for any function f of $\mathbf{\Gamma}$

$$\frac{\partial f}{\partial D_{mq}} = \frac{\partial f}{\partial \Gamma_{mj}} C_{qb}^{-1} D_{jb} + \frac{\partial f}{\partial \Gamma_{jm}} D_{jb} C_{bq}^{-1}, \tag{4.11}$$

and then substituting from (4.11) into (4.10) to obtain

$$\frac{\partial f}{\partial \Gamma_{mj}} C_{qb}^{-1} D_{jb} D_{mn} C_{nr}^{-1} + \frac{\partial f}{\partial \Gamma_{jm}} D_{jb} C_{bq}^{-1} D_{mn} C_{nr}^{-1}$$

$$- \frac{\partial f}{\partial \Gamma_{mj}} C_{rb}^{-1} D_{jb} D_{mn} C_{nq}^{-1} - \frac{\partial f}{\partial \Gamma_{jm}} D_{jb} C_{br}^{-1} D_{mn} C_{nq}^{-1} = 0,$$

which is satisfied identically. From the definition of $\mathbf{\Gamma}$ and (6.20b) of I, it is clear that $\mathbf{\Gamma} \equiv \mathbf{G}$ of I; and from (6.16c) of I we see that

$$G_{ij} = (\partial \mu_k/\partial x_i)\partial \mu_k/\partial x_j. \tag{4.12}$$

Thus we have shown that U may be reduced to the form

$$U = U(\mathbf{C}, \mathbf{N}, \mathbf{G}). \tag{4.13}$$

in place of the form shown in (4.1).

In order to show that (4.7) is equivalent to (4.15) of I, we first solve (3.23b) for $\partial U/\partial(\partial \mu_i/\partial x_m)$, with the result

$$\frac{\partial U}{\partial(\partial \mu_i/\partial x_m)} = -\frac{\partial x_m}{\partial y_i} A_{ij}, \tag{4.14}$$

and then substitute from (4.14) into (4.7) while employing (3.3a), to obtain the conditions

$$A_{ij}(\partial \mu_l/\partial y_j) = A_{jl}(\partial \mu_i/\partial y_j), \tag{4.15}$$

thereby showing that Eq. (4.15) of I is satisfied identically if the exchange energy is invariant to rigid rotations of the spin continuum with respect to the lattice continuum.

Noting that

$$\frac{\partial U}{\partial(\partial \mu_i/\partial x_m)} = 2 \frac{\partial U}{\partial G_{rm}} \frac{\partial \mu_j}{\partial x_r}, \tag{4.16}$$

and substituting from (4.16) into (3.23b), we find

$$A_{ij} = -2 \frac{\partial y_i}{\partial x_m} \frac{\partial U}{\partial G_{mr}} \frac{\partial \mu_j}{\partial x_r}, \tag{4.17}$$

which is identical with (6.35c) of I. From (4.17), (2.8), and (2.6), we obtain the important condition

[9] H. Weyl, *The Classical Groups, Their Invariants and Representations* (Princeton University Press, Princeton, New Jersey, 1946), Chap. 2, Sec. 13.

[10] L. F. Eisenhart, *Continuous Groups of Transformations* (Princeton University Press, Princeton, New Jersey, 1933), (reissued by Dover Publications, Inc., New York, 1961), Sec. 17.

$$A_{i,i}M_i = 0, \tag{4.18}$$

which is identical with Eq. (4.9) of I, and reduces from 9 to 6 the number of possible components of **A**. From (4.15) we may conclude that

$$e_{kjl}A_{i,j}(\partial\mu_k/\partial y_i) = 0. \tag{4.19}$$

Multiplying (4.19) by ρ, subtracting the resulting equation from (3.24b) and rearranging terms, we find

$$-e_{kjl}\,\partial(\rho A_{i,j}\mu_k)/\partial y_i + e_{kjl}\rho\mu_k(H_j^{\mathrm{L}} + H_j^{\mathrm{M}}) = \gamma^{-1}\rho\, d\mu_l/dt. \tag{4.20}$$

Multiplying (4.20) by dV, integrating over an *arbitrary* material region V, bounded by a surface S, and employing the divergence theorem, (2.2) and (2.6) we obtain the integral form

$$-\int_S n_i e_{kjl}A_{i,j}M_k\, dS + \int_V e_{kjl}M_k(H_j^{\mathrm{L}} + H_j^{\mathrm{M}})\, dV = \frac{d}{dt}\int_V \frac{1}{\gamma} M_l\, dV, \tag{4.21a}$$

which in invariant vector (dyadic) notation, takes the form

$$\int_S \mathbf{n}\cdot\mathbf{A}\times\mathbf{M}\, dS + \int_V \mathbf{M}\times(\mathbf{H}^{\mathrm{L}} + \mathbf{H}^{\mathrm{M}})\, dV = \frac{d}{dt}\int \frac{1}{\gamma}\mathbf{M}\, dV, \tag{4.21b}$$

which is identical with Eq. (4.2) of I, and shows that the angular momentum, of an arbitrary portion of the spin continuum, is conserved.

Since $\mathbf{H}^{\mathrm{L}}$ occurs in the equations always in the form $\mathbf{M}\times\mathbf{H}^{\mathrm{L}}$, only that portion of $\mathbf{H}^{\mathrm{L}}$ which is perpendicular to $\mathbf{M}$ is effective and meaningful, and we may take

$$\mathbf{H}^{\mathrm{L}}\cdot\mathbf{M} = 0, \tag{4.22}$$

without loss in generality. [The multiplier λ' was retained so that (4.22) would be satisfied identically for arbitrary U. Such a retention was not necessary with the L_m.] Noting that

$$\frac{\partial U}{\partial\mu_i} = \frac{\partial U}{\partial N_l}\frac{\partial y_i}{\partial x_l}, \tag{4.23}$$

and substituting from (4.23) into (3.23c), and then form (3.23c) into (4.22), we obtain

$$\lambda' = \frac{1}{\mu_{\mathrm{S}}^2}\frac{\partial U}{\partial N_l}\frac{\partial y_k}{\partial x_l}\mu_k. \tag{4.24}$$

It should be noted that the equation of the conservation of angular momentum of the lattice continuum—Eq. (3.10) of I—is not obtained as a direct consequence of the variational principle. However, we shall show that Eq. (3.10) of I is satisfied identically by virtue of the fact that the internal energy has been made invariant under rigid rotations of the deformed and magnetized material. (A similar observation was made and demonstration given by Toupin, Ref. 11, pp. 884–885.) Since, from (4.1) and (4.2) ,we have

$$\frac{\partial U}{\partial(\partial y_i/\partial x_m)} = 2\frac{\partial U}{\partial C_{rm}}\frac{\partial y_j}{\partial x_r} + \frac{\partial U}{\partial N_m}\mu_i, \qquad \frac{\partial U}{\partial\mu_i} = \frac{\partial U}{\partial N_l}\frac{\partial y_i}{\partial x_l}, \tag{4.25}$$

in which we have introduced the convention

$$\partial U/\partial C_{rs} = \partial U/\partial C_{sr}.$$

The substitution of (4.25) into (3.23a) and (3.23c), respectively yields with the aid of (4.24)

$$\tau_{ij} = 2\rho\frac{\partial y_i}{\partial x_m}\frac{\partial U}{\partial C_{mr}}\frac{\partial y_j}{\partial x_r} + \rho\frac{\partial y_i}{\partial x_m}\frac{\partial U}{\partial N_m}\mu_j, \tag{4.26a}$$

$$H_i^{\mathrm{L}} = -\frac{\partial U}{\partial N_l}\frac{\partial y_i}{\partial x_l} + \frac{1}{\mu_{\mathrm{S}}^2}\frac{\partial U}{\partial N_l}\frac{\partial y_k}{\partial x_l}\mu_k\mu_i. \tag{4.26b}$$

Now, τ_{ij} may be decomposed into a symmetric part τ_{ij}^{S} and an antisymmetric part τ_{ij}^{A}, respectively, which are given by

$$\tau_{ij}^{\mathrm{S}} = 2\rho\frac{\partial y_i}{\partial x_m}\frac{\partial U}{\partial C_{mr}}\frac{\partial y_j}{\partial x_r} + \frac{1}{2}\rho\frac{\partial U}{\partial N_m}\left(\frac{\partial y_i}{\partial x_m}\mu_j + \frac{\partial y_j}{\partial x_m}\mu_i\right), \tag{4.27a}$$

$$\tau_{ij}^{\mathrm{A}} = \frac{1}{2}\rho\frac{\partial U}{\partial N_m}\left(\frac{\partial y_i}{\partial x_m}\mu_j - \frac{\partial y_j}{\partial x_m}\mu_i\right). \tag{4.27b}$$

Substituting from (4.26b) and (2.6) into (4.27b), we obtain

$$\tau_{ij}^{\mathrm{A}} = \tfrac{1}{2}(M_iH_j^{\mathrm{L}} - H_i^{\mathrm{L}}M_j), \tag{4.28}$$

which is identical with Eq. (3.10) of I.

Equation (4.26b) is identical with Eq. (6.35b) of I. Equation (4.27a) is equivalent to Eq. (6.36) of I, as may be verified by employing (6.31) of I.

Thus, we have verified that the variational principle yields the system of equations derived in I subject to the restrictions mentioned.

Note added in proof: In recent direct correspondence I have learned that work of a somewhat similar nature has been done by W. F. Brown, Jr., J. Appl. Phys. Suppl. **36**, 994(1965).

[11] R. A. Toupin, J. Ratl. Mech. Anal. 5, 849 (1956).

Lagrangian Theory for the Second-Rank Tensor Field*

K. J. BARNES
Laboratory of Nuclear Studies,† Cornell University, Ithaca, New York
(Received 26 August 1964)

The second-rank tensor field $\phi^{\mu\nu}$ is decomposed into its various subspaces under the Lorentz group and the appropriate projection operators are exhibited explicitly. The most general local, Hermitian, free-field Lagrangian which can be formed from this field is written down, and the corresponding equations of motion and subsidiary conditions are derived by means of a variational principle. Finally some possible applications of this theory are discussed (in particular spin-2 boson theory), and all the possible couplings of this field to a Dirac particle are listed in full.

1. DECOMPOSITION UNDER THE LORENTZ GROUP

IT is well known that the second-rank tensor $\phi^{\mu\nu}$ can be decomposed under the Lorentz group into a direct sum of subspaces, one of which has five components and is a possible candidate for the description of a spin-2 field.[1] To be precise, the complete decomposition is into the sum of one spin-2, three spin-1, and two spin-0 representations. This decomposition and the appropriate projection operators are now exhibited explicitly.

Firstly, the 16-component $\phi^{\mu\nu}$ may be decomposed into the sum of a 10-component symmetric part and a 6-component antisymmetric part. This can be achieved by the use of the operators S and A defined in terms of the metric tensor $g^{\mu\nu}$ by

$$\begin{aligned} 2S &= g^{\mu\mu'}g^{\nu\nu'} + g^{\mu\nu'}g^{\nu\mu'}, \\ 2A &= g^{\mu\mu'}g^{\nu\nu'} - g^{\mu\nu'}g^{\nu\mu'}, \end{aligned} \tag{1.1}$$

which pick out the symmetric and antisymmetric parts of $\phi^{\mu\nu}$, respectively.

For example,

$$\begin{aligned} S^{\mu\nu\mu'\nu'}\phi^{\mu'\nu'} &= \tfrac{1}{2}[\phi^{\mu\nu} + \phi^{\nu\mu}] \\ &= \text{symmetric part of } \phi^{\mu\nu}. \end{aligned} \tag{1.2}$$

To achieve clarity the indices are frequently omitted in the subsequent work whenever it is felt that no confusion can arise. Hence Eq. (1.2) would read

$$S\phi = \text{symmetric part of } \phi, \tag{1.2a}$$

and the easily verifiable relations which demonstrate that S and A are projection operators take the form

$$\begin{aligned} SA &= 0 = AS, \\ S + A &= \mathbb{1}, \\ SS &= S, \\ AA &= A. \end{aligned} \tag{1.3}$$

The symmetric tensor, $S\phi$, can be further decomposed into a nine-component traceless part and a single-component part which is diagonal, by the use of the operators

$$T = \tfrac{1}{4}g^{\mu\nu}g^{\mu'\nu'} \quad \text{and} \quad R = S - T. \tag{1.4}$$

Again it may be verified that

$$\begin{aligned} RT &= 0 = TR, \\ TT &= T, \\ RR &= R, \\ R + T &= S\,[= \mathbb{1} \text{ for the subspace } (S\phi)], \end{aligned} \tag{1.5}$$

so that these are indeed projection operators.

The single component subspace $(T\phi)$ is a spin-0 representation but both $(R\phi)$ and $(A\phi)$ contain representations of more than one spin value, and can be still further decomposed. For this purpose it has been found most convenient to work in momentum space, and k^μ in the following work is defined as the 4-momentum.

In standard notation, the decomposition

$$\phi = R\phi + T\phi + A\phi \tag{1.6}$$

is equivalent to a decomposition

$$\begin{aligned} D \otimes D = {} & [D(1) \otimes D(1)] + [D(0) \otimes D(0)] \\ & + [D(0) \otimes D(1) + D(1) \otimes D(0)], \end{aligned} \tag{1.7}$$

where $D = D(1) + D(0)$ transforms like a vector, and $D(s)$ is the irreducible representation corresponding to spin s. Hence R should decompose as a direct sum $D(2) + D(1) + D(0)$, and A as $D(1) + D(1)$, giving the final decomposition

$$D \otimes D = D(2) + 3D(1) + 2D(0). \tag{1.8}$$

It is therefore necessary to find the projection operators, P_i $(i = 0, 1, 2)$, which will produce the decomposition of $R\phi$. The most general dimensionless fourth-rank tensor $P_i^{\mu\nu\mu'\nu'}$ which can be constructed from k^μ and $g^{\mu\nu}$, so as to be symmetric

* Supported in part by the Office of Naval Research.
† The greater part of this work was performed at Imperial College, London, and formed part of an unpublished Ph.D. thesis (1963).
[1] C. Fronsdal, Nuovo Cimento Suppl. 9, 416 (1958).

under the exchanges $\mu \leftrightarrow \nu$ and $\mu' \leftrightarrow \nu'$ and to be traceless in the pairs (μ, ν) and (μ', ν'), can be written in the form

$$\begin{aligned} P_i^{\mu\nu\mu'\nu'} &= a_i g^{\mu\nu} g^{\mu'\nu'} + b_i [g^{\mu\mu'} g^{\nu\nu'} + g^{\mu\nu'} g^{\nu\mu'}] \\ &\quad - 2(2a_i + b_i)[g^{\mu\nu} k^{\mu'} k^{\nu'} + g^{\mu'\nu'} k^{\mu} k^{\nu}]/k^2 \\ &\quad + \frac{c_i}{k^2} \begin{bmatrix} g^{\mu\mu'} k^{\nu} k^{\nu'} + g^{\nu\nu'} k^{\mu} k^{\mu'} \\ + g^{\mu\nu'} k^{\nu} k^{\mu'} + g^{\nu\mu'} k^{\mu} k^{\nu'} \end{bmatrix} \\ &\quad + 4[2b_i + 4a_i - c_i] k^{\mu} k^{\nu} k^{\mu'} k^{\nu'}/k^4. \end{aligned} \tag{1.9}$$

At this point it is possible to find P_2, P_1, and P_0, by imposing the projection operator conditions

$$P_i P_j = \delta_{ij} P_i, \tag{1.10}$$

$$\sum_{i=0}^{2} P_i = 1 = \tfrac{1}{2}[g^{\mu\mu'} g^{\nu\nu'} + g^{\mu\nu'} g^{\nu\mu'}] - \tfrac{1}{4} g^{\mu\nu} g^{\mu'\nu'}, \tag{1.11}$$

and identifying the three solutions exactly by finding the number of eigenvalues which should of course be $(2i + 1)$ for P_i. In practice however, it is easier to find P_2 separately by imposing the conditions

$$P_2 P_2 = P_2, \tag{1.12}$$

$$k^{\mu} P_2^{\mu\nu\mu'\nu'} = 0. \tag{1.13}$$

This method was used by Fronsdal[1] who gives general formulas for the projection operator of highest spin in any such decomposition. The result of the ensuing tensor manipulation may be conveniently written in the form:

$$P_2^{\mu\nu\mu'\nu'} = \tfrac{1}{2}[\theta^{\mu\mu'} \theta^{\nu\nu'} + \theta^{\mu\nu'} \theta^{\nu\mu'}] - \tfrac{1}{3} \theta^{\mu\nu} \theta^{\mu'\nu'}, \tag{1.14}$$

where

$$\theta^{\mu\nu} = g^{\mu\nu} - k^{\mu} k^{\nu}/k^2. \tag{1.15}$$

Notice that $P_2^{\mu\nu\mu\nu} = 5 = (2s + 1)$ where $s = 2$ as was required.

As P_2 has been found explicitly, the application of conditions (1.10) and (1.11) to the general form (1.9), to find P_1 and P_0 is now comparatively simple. The results may be taken in the convenient form

$$\begin{aligned} P_0 &= \psi^{\mu\nu} \psi^{\mu'\nu'}/12, \\ P_1 &= \frac{1}{2k^2} \begin{bmatrix} \theta^{\mu\mu'} k^{\nu} k^{\nu'} + \theta^{\nu\nu'} k^{\mu} k^{\mu'} \\ + \theta^{\mu\nu'} k^{\nu} k^{\mu'} + \theta^{\nu\mu'} k^{\mu} k^{\nu'} \end{bmatrix}, \end{aligned} \tag{1.17}$$

where

$$\psi^{\mu\nu} = g^{\mu\nu} - 4k^{\mu} k^{\nu}/k^2. \tag{1.18}$$

It is a simple matter to check that $P_0^{\mu\nu\mu\nu} = 1$ and $P_1^{\mu\nu\mu\nu} = 3$, so that the subscripts are indeed the appropriate spin values.

By similar methods to the above, the six-component subspace $(A\phi)$ decomposes into two three-component subspaces, by means of the projection operators

$$A_I = \frac{1}{2k^2} \begin{bmatrix} g^{\mu\mu'} k^{\nu} k^{\nu'} - g^{\nu\mu'} k^{\mu} k^{\nu'} \\ - g^{\mu\nu'} k^{\nu} k^{\mu'} + g^{\nu\nu'} k^{\mu} k^{\mu'} \end{bmatrix}, \tag{1.19}$$

$$A_1 = A - A_I, \tag{1.20}$$

and it is easy to verify directly that $A_I^{\mu\nu\mu\nu} = 3 = A_1^{\mu\nu\mu\nu}$, so that both subspaces indeed do have three components.

2. LAGRANGIAN THEORY

The purpose of this section is to find the most general local, Hermitian, Lagrangian which can be formed from the second-rank tensor field, and the corresponding equations of motion and subsidiary conditions derivable by a variational principle from this Lagrangian. It is the author's experience, however, that the complexity of the second-rank tensor field and the fourth-rank tensor operators acting on this field tend to produce confusion unless the basic pattern of the method to be used is understood clearly from the start. Therefore, the most general Lagrangian for a vector field will be derived first, and it is hoped that this will provide a concrete example to guide the reader through the later work.

A. Vector Field

(i). The Lagrangian and the Equations of Motion

A neutral spin-1 particle of mass m is most usually described by a four-component Hermitian vector field A^{μ}. This field contains in addition to the spin-1 field, a subsidiary scalar field, and is in fact a direct sum $D(1) + D(0)$, the projection operators being

$$P_1^{\mu\nu} = g^{\mu\nu} - k^{\mu} k^{\nu}/k^2, \tag{2.1}$$

$$P_0 = k^{\mu} k^{\nu}/k^2. \tag{2.2}$$

The most general local free-field Lagrangian, bilinear in this field, and containing no higher than first derivatives, may be written

$$\begin{aligned} L &= am^2 A^{\mu} g^{\mu\nu} A^{\nu} \\ &\quad + \partial^{\mu} A^{\nu} [b g^{\mu\mu'} g^{\nu\nu'} + c g^{\mu\nu'} g^{\nu\mu'} + d g^{\mu\nu} g^{\mu'\nu'}] \partial^{\mu'} A^{\nu'}, \end{aligned} \tag{2.3}$$

where a, b, c, and d are constants. Application of a variational principle to this Lagrangian yields the Euler–Lagrange equations

$$2am^2 A^{\mu} - 2b\partial^{\mu}\partial^{\nu} A^{\nu} - 2c \,\Box^2 A^{\mu} - 2d\partial^{\mu}\partial^{\nu} A^{\nu} = 0,$$

which may be written in momentum space as

$$\Lambda^{\mu\nu} A^{\nu} = 0,$$

where

$$\begin{aligned} \Lambda^{\mu\nu} &= [2am^2 + 2bk^2] P_1 \\ &\quad + [2am^2 + 2k^2(b + c + d)] P_0. \end{aligned} \tag{2.4}$$

Since the field A^μ decomposes into the direct sum $(P_1A)^\mu + (P_0A)^\mu$, Eqs. (2.4) may be separated in the form

$$P_1\Lambda A = 0,$$
$$P_0\Lambda A = 0,$$

or in full

$$[2am^2 + 2bk^2](P_1A)^\mu = 0, \tag{2.4a}$$
$$[2am^2 + 2k^2(b + c + d)](P_0A)^\mu = 0. \tag{2.4b}$$

(ii) Application to Spin-1 Theory of a Neutral Boson

If this is to be a free-field theory for spin-1 particles of mass m, Eq. (2.4a) must give the Klein–Gordon equation, and (2.4b) must give the subsidiary condition that the scalar part of the field vanish. Hence, $b = \frac{1}{2}$, $a = -\frac{1}{2}$, $c + d = -\frac{1}{2}$, and the most general Lagrangian is given by

$$L = -\tfrac{1}{2}m^2A^\mu A^\mu + \tfrac{1}{2}\,\partial^\mu A^\nu \partial^\mu A^\nu + c\partial^\mu A^\nu \partial^\nu A^\mu - (c + \tfrac{1}{2})\partial^\mu A^\mu \partial^\nu A^\nu. \tag{2.5}$$

In the standard treatment of this field (see e.g., Bogoliubov and Shirkov[2]) c is set equal to $-\frac{1}{2}$ and the Lagrangian is given by

$$L = -\tfrac{1}{2}m^2A^\mu A^\mu + \tfrac{1}{4}[\partial^\mu A^\nu - \partial^\nu A^\mu]^2. \tag{2.6}$$

It may then be verified that the energy density given by the application of Noether's theorem to this Lagrangian is positive definite.

Consider now the effect of introducing an interaction, defined by an interaction Lagrangian of the form $J^\mu A^\mu$, between this field and some external field. The equation of motion (2.4) then becomes

$$\Lambda^{\mu\nu}A^\nu = J^\mu \tag{2.7}$$

and separates into

$$(k^2 - m^2)(P_1A)^\mu = (P_1J)^\mu \tag{2.7a}$$
$$-m^2(P_0A)^\mu = (P_0J)^\mu. \tag{2.7b}$$

Hence,

$$(k^2 - m^2)A^\mu = (P_1J)^\mu - (k^2 - m^2)(P_0J)^\mu/m^2 = [g^{\mu\nu} - k^\mu k^\nu/m^2]J^\nu, \tag{2.8}$$

and thus, by Umezawa's method,[3] the numerator of the propogator of the vector field is $d^{\mu\nu} = g^{\mu\nu} - k^\mu k^\nu/m^2$. It is important to notice that although the Lagrangian contains an arbitrary parameter c (if such considerations as the positive definiteness of the energy and the quantization of the theory are not considered). The equations of motion and the propogator are well determined.

B. Second-Rank Tensor Field

(i). The Lagrangian and the Equations of Motion

The first task is to write down the most general Hermitian Lagrangian bilinear in $\phi^{\mu\nu}$ and containing no higher than second derivatives. Making use of the notation developed in Sec. 1, this may be achieved in the convenient form

$$\begin{aligned}L = {} & m^2[a(R\phi)^{\mu\nu}(R\phi)^{\mu\nu} + b(T\phi)^{\mu\nu}(T\phi)^{\mu\nu} + c(A\phi)^{\mu\nu}(A\phi)^{\mu\nu}] \\ & + \partial^\lambda(R\phi)^{\mu\nu}[d\partial^\lambda(R\phi)^{\mu\nu} + e\partial^\mu(R\phi)^{\lambda\nu}] \\ & + \partial^\lambda(A\phi)^{\mu\nu}[p\partial^\lambda(A\phi)^{\mu\nu} + n\partial^\mu(A\phi)^{\lambda\nu}] \\ & + g[\partial^\lambda(R\phi)^{\mu\nu}\partial^\mu(A\phi)^{\lambda\nu} + \partial^\lambda(A\phi)^{\mu\nu}\partial^\mu(R\phi)^{\lambda\nu}] \\ & + r[\partial^\mu(T\phi)^{\mu\nu}\partial^\lambda(T\phi)^{\lambda\nu}] + u[\partial^\mu(A\phi)^{\mu\nu}\partial^\lambda(A\phi)^{\lambda\nu}] \\ & + s[\partial^\mu(T\phi)^{\mu\nu}\partial^\lambda(A\phi)^{\lambda\nu} + \partial^\mu(A\phi)^{\mu\nu}\partial^\lambda(T\phi)^{\lambda\nu}] \\ & + y[\partial^\mu(T\phi)^{\mu\nu}\partial^\lambda(R\phi)^{\lambda\nu} + \partial^\mu(R\phi)^{\mu\nu}\partial^\lambda(T\phi)^{\lambda\nu}] \\ & + z[\partial^\mu(A\phi)^{\mu\nu}\partial^\lambda(R\phi)^{\lambda\nu} + \partial^\mu(R\phi)^{\mu\nu}\partial^\lambda(A\phi)^{\lambda\nu}] \\ & + v[\partial^\mu(R\phi)^{\mu\nu}\partial^\lambda(R\phi)^{\lambda\nu}].\end{aligned} \tag{2.9}$$

The task of applying a variational principle to this Lagrangian to obtain the Euler–Lagrange equations

$$\partial L/\partial\phi^{\mu\nu} - \partial^\lambda \partial L/\partial(\partial^\lambda\phi^{\mu\nu}) = 0 \tag{2.10}$$

is tedious, but can be simplified by inserting the unit operators in each term. Thus for example the term with coefficient g would be written

$$g\left[\partial^\lambda\phi^{\mu\nu}R^{\mu\nu\alpha\beta}A^{\lambda\beta\mu'\nu'}\,\partial^\alpha\phi^{\mu'\nu'} + \partial^\lambda\phi^{\mu\nu}A^{\mu\nu\alpha\beta}R^{\lambda\beta\mu'\nu'}\,\partial^\alpha\phi^{\mu'\nu'}\right].$$

After manipulation the final result of this process may be expressed in the form

$$\Lambda^{\mu\nu\mu'\nu'}\phi^{\mu'\nu'} = 0, \tag{2.11}$$

where

$$\begin{aligned}\Lambda = {} & P_2[2am^2 + 2dk^2] + P_1[2am^2 + k^2(2d + e + v)] \\ & + P_0[2am^2 + \tfrac{1}{2}k^2(4d + 3e + 3v)] + T[2bm^2 + \tfrac{1}{2}rk^2] \\ & + A_1[2cm^2 + 2pk^2] + A_I[2cm^2 + k^2(2p + n + u)] \\ & - (g + z)k^2[\Lambda_{A_IP_1} + \Lambda_{P_1A_I}] \\ & + \tfrac{1}{2}\sqrt{3}yk^2[\Lambda_{P_0T} + \Lambda_{TP_0}],\end{aligned} \tag{2.12}$$

and

$$\Lambda_{P_0T} = \frac{1}{\sqrt{3}}\left[\frac{4k^\mu k^\nu}{k^2} - g^{\mu\nu}\right]\frac{g^{\mu'\nu'}}{4} = -\frac{\psi^{\mu\nu}g^{\mu'\nu'}}{4\sqrt{3}}, \tag{2.13}$$

$$\Lambda_{TP_0} = -g^{\mu\nu}\psi^{\mu'\nu'}/4\sqrt{3}, \tag{2.14}$$

$$\Lambda_{A_IP_1} = \frac{1}{2k^2}\left[g^{\mu\mu'}k^\nu k^{\nu'} - g^{\nu\mu'}k^\mu k^{\nu'} + g^{\mu\nu'}k^\nu k^{\mu'} - g^{\nu\nu'}k^\mu k^{\mu'}\right], \tag{2.15}$$

[2] N. N. Bogoliubov and D. V. Shirkov, *Introduction to the Theory of Quantized Fields* (Interscience Publishers, Inc., New York, 1959).

[3] H. Umezawa, *Quantum Field Theory* (North-Holland Publishing Company, Amsterdam, 1956).

$$\Lambda_{P_1A_1} = \frac{1}{2k^2}\begin{bmatrix} g^{\mu\mu'}k^\nu k^{\nu'} + g^{\nu\mu'}k^\mu k^{\nu'} \\ - g^{\mu\nu'}k^\nu k^{\mu'} - g^{\nu\nu'}k^\mu k^{\mu'} \end{bmatrix}. \tag{2.16}$$

The reasons for the explicit choice of the operators defined by Eqs. (2.13)–(2.16) will be explained shortly, but the significance of their presence should be appreciated immediately. The set of six projection operators

$$0_i = \{P_2, P_1, P_0, T, A_1, A_I\}$$

is not a sufficient basis for the space of operators allowed in Λ. This is to be contrasted with the case of the vector field, where the two projection operators did indeed form a basis. In the present case, it is necessary to introduce four more independent operators to form such a basis. The particular choice defined by Eqs. (2.13)–(2.16) is considered by the author to be the most convenient, and it is hoped that the notation is almost self-explanatory. They have the properties

$$0_i\Lambda_{0_i0_j} = \Lambda_{0_i0_j} = \Lambda_{0_i0_j}0_j, \tag{2.17}$$

$$\Lambda_{0_i0_j}\Lambda_{0_j0_k} = 0_i, \tag{2.18}$$

and their products with any other 0 or Λ are zero.

In view of the decomposition performed in Sec. 1, the Eq. (2.11) may be decomposed into the six equations

$$0_i\Lambda\phi = 0 \qquad (i = 1, \cdots, 6), \tag{2.19}$$

which written out in detail are

$$[2am^2 + 2dk^2](P_2\phi) = 0, \tag{2.19a}$$

$$[2am^2 + k^2(2d + e + v)](P_1\phi) - (g + z)k^2\Lambda_{P_1A_I}(A_I\phi) = 0, \tag{2.19b}$$

$$[2am^2 + \tfrac{1}{2}k^2(4d + 3e + 3v)](P_0\phi) + \tfrac{1}{2}\sqrt{3}k^2y\Lambda_{P_0T}(T\phi) = 0, \tag{2.19c}$$

$$[2bm^2 + \tfrac{1}{2}rk^2](T\phi) + \tfrac{1}{2}\sqrt{3}k^2y\Lambda_{TP_0}(P_0\phi) = 0, \tag{2.19d}$$

$$[2cm^2 + k^2(2p + n + u)](A_I\phi) - (g + z)k^2\Lambda_{A_IP_1}(P_1\phi) = 0, \tag{2.19e}$$

$$[2cm^2 + 2pk^2](A_1\phi) = 0. \tag{2.19f}$$

These equations are the analog of (2.7a) and (2.7b) for the vector case. Note, however, that for the present case certain of the equations are coupled because of the presence of the four additional operators.

(ii). Application to the Theory of a Neutral Boson of Spin 2

If this is to be a free-field theory for spin-2 particles of mass m, then the set of equations (2.19) must give the Klein–Gordon equation

$$(k^2 - m^2)(P_2\phi) = 0, \tag{2.20a}$$

and the subsidiary conditions

$$0_i\phi = 0 \qquad (0_i \neq P_2), \tag{2.20b–f}$$

that the spin-1 and spin-0 parts of the field vanish. At this point a choice of approaches presents itself. The condition $T\phi = 0$ is purely algebraic, as are the pair $A_I\phi = 0$, $A_1\phi = 0$ taken in the form $A\phi = 0$. Thus, these conditions may be imposed on the theory *a priori*, by stipulating that ϕ be symmetric, or traceless, or both. Those cases will now be dealt with separately.

Case 1 ($\phi^{\mu\nu}$ *is taken to be symmetric and traceless*). The set of equations (2.19) reduce to the form

$$[2am^2 + 2dk^2](P_2\phi) = 0, \tag{2.21a}$$

$$[2am^2 + k^2(2d + e + v)](P_1\phi) = 0, \tag{2.21b}$$

$$[2am^2 + \tfrac{1}{2}k^2(4d + 3e + 3v)](P_0\phi) = 0. \tag{2.21c}$$

It is clearly *not* possible to adjust the parameters to obtain Eq. (2.20a) and the subsidiary conditions simultaneously. This result was obtained by Fronsdal[1] in a rather more abstract manner.

Case 2 ($\phi^{\mu\nu}$ *is taken to be symmetric*). The equations (2.19) reduce to the form

$$[2am^2 + 2dk^2](P_2\phi) = 0, \tag{2.22a}$$

$$[2am^2 + k^2(2d + e + v)](P_1\phi) = 0, \tag{2.22b}$$

$$[2am^2 + \tfrac{1}{2}k^2(4d + 3e + 3v)](P_0\phi) + \tfrac{1}{2}\sqrt{3}yk^2\Lambda_{P_0T}(T\phi) = 0, \tag{2.22c}$$

$$[2bm^2 + \tfrac{1}{2}rk^2](T\phi) + \tfrac{1}{2}\sqrt{3}yk^2\Lambda_{TP_0}(P_0\phi) = 0. \tag{2.22d}$$

Choosing $a = -\frac{1}{2}$, $d = \frac{1}{2}$ and $(e + v) = -1$ produces the equation (2.20a) and the subsidiary condition on $(P_1\phi)$. The equations (2.22c) and (2.22d) may then be combined by use of the properties of the 0's and Λ's, to give

$$[(m^2 + \tfrac{1}{2}k^2)(2bm^2 + \tfrac{1}{2}rk^2) + \tfrac{3}{4}k^4y^2](\alpha\phi) = 0, \tag{2.23}$$

where $\alpha = P_0$ or T. Hence, the subsidiary conditions on $(P_0\phi)$ and $(T\phi)$ may finally be obtained by setting $r = -3y^2$; $b = \frac{3}{2}y^2$; $y \neq 0$. There results a two-parameter family of Lagrangians given by Eq. (2.9), where

$$a = -\tfrac{1}{2}, \quad d = \tfrac{1}{2}, \quad e + v = -1, \quad r = -3y^2, \quad 2b = 3y^2 \neq 0, \tag{2.24}$$

and all other parameters are zero, and a single-parameter family of equations of motion specified by Λ given in Eq. (2.12) with the above values

of the parameters (y being the only remaining free parameter).

Case 3 ($\phi^{\mu\nu}$ *is traceless*). The above type of analysis gives, omitting the details,

$$b = r = s = y = 0,$$
$$a = -\tfrac{1}{2}, \quad d = \tfrac{1}{2}, \quad e + v = -\tfrac{2}{3}, \quad p = 0, \qquad (2.25)$$
$$c = 3(n + u) = 9(g + z)^2 \neq 0.$$

There thus results a four-parameter family of Lagrangians specified by e, n, g, and z, and a single-parameter family of equations of motion specified by $(g + z)$.

Case 4 ($\phi^{\mu\nu}$ *is completely general*). Here the results are

$$a = -\tfrac{1}{2}, \qquad d = \tfrac{1}{2}, \qquad p = 0,$$
$$(1 + e + v)(n + u) = (g + z)^2$$
$$= 2c(1 + e + v)^2, \qquad c \neq 0, \qquad (2.26)$$
$$r(2 + 3e + 3v) = 3y^2$$
$$= 8b(1 + \tfrac{3}{2}e + \tfrac{3}{2}v)^2, \qquad b \neq 0.$$

There results a seven-parameter family of Lagrangians; and a three-parameter family of equations of motion specified by $(e + v)$, y, $(g + z)$.

Consider now the effect of introducing an interaction between ϕ and some external field (e.g., a Dirac field), specified by an interaction Lagrangian of the form

$$L' = J^{\mu\nu}\phi^{\mu\nu}, \qquad (2.27)$$

where the tensor current $J^{\mu\nu}$ is independent of $\phi^{\mu\nu}$. The equations of motion are now modified to the form

$$\Lambda\phi = J, \qquad (2.28)$$

where (in analogy with the vector field case) this equation may be decomposed, and the propogator for the field $\phi^{\mu\nu}$ can be derived by Umezawa's method,[3] utilizing the properties of the 0's and Λ's. The results are given below (without the tedious but simple derivations) for each of the above cases 2, 3, and 4.

Case 2 ($\phi^{\mu\nu}$ *is symmetric*).

$$d^{\mu\nu\mu'\nu'} = P_2 + \frac{k^2 - m^2}{m^4}\left[-m^2P_1 + \left(\frac{k^2}{2} - m^2\right)P_0 + \frac{k^2}{2\sqrt{3}y}(\Lambda_{P_0T} + \Lambda_{TP_0}) + \frac{1}{3y^2}\left(m^2 + \frac{k^2}{2}\right)T\right]. \qquad (2.29)$$

Case 3 ($\phi^{\mu\nu}$ *is traceless*).

$$d^{\mu\nu\mu'\nu'} = P_2 + \frac{k^2 - m^2}{m^4} \times \begin{bmatrix} -m^2P_0 - \left(\frac{k^2}{3} + m^2\right)P_1 \\ - \frac{k^2}{9(g+z)^2}(\Lambda_{A_IP_1} + \Lambda_{P_1A_I}) \\ - \left(\frac{k^2}{3} - m^2\right)\frac{A_I}{9(g+z)^2} \\ + \frac{m^2A_1}{9(g+z)^2} \end{bmatrix}. \qquad (2.30)$$

Case 4 ($\phi^{\mu\nu}$ *is completely general*).

$$d = P_2 + \frac{k^2 - m^2}{m^4} \begin{bmatrix} -P_1[m^2 + k^2(1 + e + v)] - P_0[m^2 + \tfrac{1}{2}k^2(2 + 3e + 3v)] \\ + \left(\frac{1 + e + v}{g + z}\right)^2[m^2A_1 + \{m^2 - k^2(1 + e + v)\}A_I] \\ - \frac{k^2(1 + e + v)^2}{(g + z)}[\Lambda_{A_IP_1} + \Lambda_{P_1A_I}] \\ + \frac{2\left(1 + \frac{3e}{2} + \frac{3v}{2}\right)^2}{3y^2}\{[2m^2 - k^2(1 + 3e + 3v)T + \sqrt{3}\, yk^2[\Lambda_{P_0T} + \Lambda_{TP_0}]\} \end{bmatrix}. \qquad (2.31)$$

The author would like to draw the reader's attention to the free parameters which are present in these propogators. These parameters have no counterpart in the usual theories of lower spin (e.g., the spin-1 theory derived in Sec. 2A), and there seems to be no method by which they can be determined or interpreted in this classical theory. (See however Sec. 2Biii). Rivers[4] has now extended this theory and provided a method of quantization. His method of quantization gives strong restrictions on the parameters, leading to the conclusion that only Case 2 of the above theory is possible, and that in this case $y = 1$. Thus, the theory is given by a single-parameter family of Lagrangians (specified by e) leading to unique field equations specified by

$$\Lambda = (k^2 - m^2)P_2 - m^2P_1 - (m^2 + \tfrac{1}{2}k^2)P_0 + 3T(m^2 - \tfrac{1}{2}k^2) + \tfrac{1}{2}\sqrt{3}k^2[\Lambda_{P_0T} + \Lambda_{TP_0}] \qquad (2.32)$$

[4] R. J. Rivers, preprint, Imperial College, London (1964) (to be published).

and a unique propagator specified by

$$d = P_2\{\text{evaluated with } k^2 = m^2\}$$
$$= P_2 + (k^2 - m^2)/m^4$$
$$\times\begin{bmatrix} -m^2P_1 + [\tfrac{1}{2}k^2 - m^2]P_0 + [\tfrac{1}{2}k^2 + m^2]T/3 \\ + (k^2/2\sqrt{3})[\Lambda_{P_0T} + \Lambda_{TP_0}] \end{bmatrix}. \quad (2.33)$$

(iii). Other Possible Applications

It is possible by suitable choice of the parameters in the Lagrangian to utilize the set of equations (2.19) for other purposes than the description of a spin-2 field. To be precise, the particular choice $a \neq 0$, $d = 0$ gives from Eq. (2.19a),

$$P_2\phi = 0 \quad (2.34)$$

so that the remaining five equations in (2.19) form a set of equations of motion and subsidiary conditions on the spin-1 and spin-0 fields. This gives rise to several fascinating possibilities, and in particular, it seems to provide a natural framework in which to describe two spin-1 fields simultaneously.[5] In this context the g, and z terms in the Lagrangian (2.9) could be given the interpretation of interaction terms between the two spin-1 fields giving rise to direct transitions of one field into the other. This concept is rather a mild generalization of the direct transitions between a photon and ρ meson fashionable in recent work[6] on nucleon form factors.

3. THE COUPLING OF THE SECOND-RANK TENSOR FIELD TO DIRAC PARTICLES

The purpose of this section is to write down all possible couplings of the second-rank tensor field to Dirac particles, and to classify these couplings according to their parity, C parity, and the part of the field to which they couple. To define the notation, consider a general vertex at which a particle of integral spin is absorbed by a free Dirac particle of mass μ, and let $P=p'+p$ and $k=p'-p$, where p and p' are the initial and final values of the momentum of the Dirac particle. With this notation, the relevant interaction current j (P, k) may be defined by the equation

$$\langle p's'|\, F(k)\, |ps\rangle = \bar{U}_{s'}(p')j(P, k)U_s(p), \quad (3.1)$$

where F and j carry the requisite number of indices for the spin of the particle emitted or absorbed. All insignificant constants have been absorbed into the $F(k)$ which has the transformation property

$$\mathcal{C}F(k)\mathcal{C}^{-1} = \eta_c F(k), \quad (3.2)$$

where $\mathcal{C}$ is the operation of charge conjugation, and η_c is the charge parity of the particle absorbed or emitted.

Using the approach of Lehmann, Symanzik, and Zimmermann[7] so that

$$|p, s\rangle = \lim_{t_0\to\pm\infty} \int \bar{\psi}(x)Au_s(p)e^{-ip\cdot x}\, d_3x\, |0\rangle, \quad (3.3)$$

the interaction current may be written in the form

$$j(P, k) = -\int d_4x\, d_4y\, e^{i(p'\cdot y - p\cdot x)}\vec{D}_y \times \langle 0|\, T[\psi(y)F(k)\bar{\psi}(x)]\, |0\rangle\, \overleftarrow{D}_x, \quad (3.4)$$

where $\vec{D}_y = (i\gamma\vec{\partial}/\partial y - m)$ and $\overleftarrow{D}_x = (-i\gamma\overleftarrow{\partial}/\partial x - m)$ and the T indicates that the time-ordered product is to be taken. The notation is such that the Dirac equation takes the form $(i\gamma^\mu\partial_\mu - m)\psi(x) = 0$, and adjoint and charge conjugate spinors are defined by the equations

$$\bar{\psi} = \psi^\dagger A, \qquad \psi_c = -C\bar{\psi}^{\mathrm{T}}, \quad (3.5)$$

where the $^{\mathrm{T}}$ and $\dagger$ respectively indicate transposition and Hermitian conjugation, and A and C are defined by the equations

$$\begin{aligned} \gamma_\mu &= A\gamma_\mu A^{-1}, & A^\dagger &= A, \\ \gamma_\mu^{\mathrm{T}} &= -C\gamma_\mu C^{-1}, & C^{\mathrm{T}} &= C, \end{aligned} \quad (3.6)$$

Invariance under the operation of charge conjugation implies that

$$j(P, k) = \mathcal{C}j(P, k)\mathcal{C}^{-1} \quad (3.7)$$

or, using Eqs. (3.2) and (3.3) and transposing

$$j^{\mathrm{T}}(P, k) = \int d_4x\, d_4y\, e^{i(p'\cdot y - p\cdot x)}(-i\gamma^{\mathrm{T}}\vec{\partial}/\partial x - m) \times \langle 0|\, T[\bar{\psi}_c^T(x)\eta_c F\psi_c^T(y)]\, |0\rangle\, (i\gamma^{\mathrm{T}}\overleftarrow{\partial}/\partial y - m). \quad (3.8)$$

When the variables of integration x and y are interchanged, and Eq. (3.5) is used, this leads to the result

$$j^T(P, k) = \eta_c C\int d_4x\, d_4y\, e^{i(p'\cdot x - p\cdot y)} \times \vec{D}_y\, \langle 0|\, T[\psi(y)F(k)\bar{\psi}(x)]\, |0\rangle\, \overleftarrow{D}_x C^{-1} = \eta_c C j(-P, k)C^{-1}$$

or

$$C^{-1}j(P, k)C = \eta_c j^T(-P, k). \quad (3.9)$$

For the present case of a second-rank tensor field, there are six symmetric and four antisymmetric second-rank tensors of positive space parity which

[5] G. Feldman and P. T. Matthews, Phys. Rev. **132**, 823 (1963).

[6] See, e.g., P. T. Matthews, Proc. Conf. Elementary Particles, Aix-en-Provence **2**, 87 (1961).

[7] H. Lehmann, K. Symanzik, and W. Zimmermann, Nuovo Cimento **1**, 205 (1955).

TABLE I. The couplings of a virtual spin-2 boson to free Dirac spinors.

	$\eta_P = +1$		$\eta_P = -1$	
$\eta_c = +1$	$\frac{P^\mu P^\nu}{4\mu^2 - k^2} + \frac{1}{3}\left(\frac{k^\mu k^\nu}{k^2} - g^{\mu\nu}\right)$	(P_2)	$\gamma_5\left[\frac{P^\mu P^\nu}{4\mu^2 - k^2} + \frac{1}{3}\left(\frac{k^\mu k^\nu}{k^2} - g^{\mu\nu}\right)\right]$	(P_2)
	$\frac{P^\mu\gamma^\nu + \gamma^\mu P^\nu}{4\mu} + \frac{1}{3}\left(\frac{k^\mu k^\nu}{k^2} - g^{\mu\nu}\right)$	(P_2)	$\gamma_5\left[\frac{k^\mu\gamma^\nu + \gamma^\mu k^\nu}{4\mu} + \frac{k^\mu k^\nu}{k^2}\right]$	(P_1)
	$\frac{k^\mu k^\nu}{k^2} - \frac{g^{\mu\nu}}{4}$	(P_0)	$\gamma_5\left[\frac{k^\mu k^\nu}{k^2} - \frac{g^{\mu\nu}}{4}\right]$	(P_0)
	$g^{\mu\nu}$	(T)	$\gamma_5 g^{\mu\nu}$	(T)
	$\frac{P^\mu\gamma^\nu - \gamma^\mu P^\nu}{2\mu}$	(A_1)	$\gamma_5\left[\frac{k^\mu\gamma^\nu - \gamma^\mu k^\nu}{2\mu}\right]$	(A_1)
$\eta_c = -1$	$\frac{k^\mu\gamma^\nu + \gamma^\mu k^\nu}{2\mu}$	(P_1)	$\gamma_5\left[\frac{P^\mu\gamma^\nu + \gamma^\mu P^\nu}{2\mu} + \frac{P^\mu k^\nu + k^\mu P^\nu}{k^2}\right]$	(P_2)
	$\frac{P^\mu k^\nu + k^\mu P^\nu}{4\mu^2}$	(P_1)	$\gamma_5\left[\frac{P^\mu k^\nu + k^\mu P^\nu}{4\mu^2}\right]$	(P_1)
	$\frac{P^\mu k^\nu - k^\mu P^\mu}{4\mu^2}$	(A_1)	$\gamma_5\left[\frac{P^\mu\gamma^\nu - \gamma^\mu P^\nu}{2\mu} + \frac{P^\mu k^\nu - k^\mu P^\nu}{k^2}\right]$	(A_1)
	$\frac{k^\mu\gamma^\nu - \gamma^\mu k^\nu}{2\mu}$	(A_1)	$\gamma_5\left[\frac{P^\mu k^\nu - k^\mu P^\nu}{2\mu}\right]$	(A_1)
	$i\sigma^{\mu\nu} + \frac{2\mu}{k^2}(k^\mu\gamma^\nu - \gamma^\mu k^\nu) + \frac{P^\mu k^\nu - k^\mu P^\nu}{k^2}$	(A_1)	$\gamma_5\left[i\sigma^{\mu\nu} + \frac{P^\mu k^\nu - k^\mu P^\nu}{k^2}\right]$	(A_1)

can be constructed from P^μ, k^μ and Dirac matrices, and which are linearly independent. These may conveniently be taken as

$$\left.\begin{array}{l} P^\mu\gamma^\nu + \gamma^\mu P^\nu, \\ P^\mu P^\nu, \\ k^\mu k^\nu, \\ g^{\mu\nu}, \end{array}\right\}\eta_c = +1; \qquad \left.\begin{array}{l} P^\mu k^\nu - k^\mu P^\nu \\ k^\mu\gamma^\nu - \gamma^\mu k^\nu \\ \sigma^{\mu\nu} = \frac{i}{2}[\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu] \end{array}\right\}\eta_c = -1;$$

$$\left.\begin{array}{l} P^\mu k^\nu + k^\mu P^\nu, \\ k^\mu\gamma^\nu + \gamma^\mu k^\nu \end{array}\right\}\eta_c = -1; \quad P^\mu\gamma^\nu - \gamma^\mu P^\nu, \quad \eta_c = +1,$$

where the C-parity assignment has been determined by the application of Eq. (3.9). A similar procedure can be carried through for the tensors of opposite space parity which may be taken to be the above used set multipled by γ_5 from the left, and the corresponding C-parity assignments may again be derived by use of Eq. (3.9).

All that remains is to investigate which combinations of these tensors are projected out by the operators defined in the first part of this paper. The results of this piece of tensor manipulation are summarized in Table I, where the notation is such that if a general one of the expressions is designated by $\omega^{\mu\nu}(o_i)$ then

$$O_i\omega^{\mu\nu}(o_j) = \delta_{ij}\omega^{\mu\nu}(o_i).$$

ACKNOWLEDGMENTS

The author is indebted to Professor P. T. Matthews and Professor A. Salam for constant advice and encouragement, and to many friends, both at Imperial College and Cornell, for valuable discussions.

Exact Bootstrap Solutions in Some Static Models of Meson–Baryon Scattering*

KERSON HUANG AND FRANCIS E. LOW
Department of Physics and Laboratory for Nuclear Science, Massachusetts Institute of Technology, Cambridge, Massachusetts
(Received 22 October 1964)

We study the exact bootstrap solutions to four well-known models of meson–baryon scattering in the nonrecoil, one-meson approximation. The models are the neutral scalar theory, the charged scalar theory, the symmetric scalar theory, and the neutral pseudoscalar theory. A bootstrap solution is defined to be a solution satisfying Levinson's theorem of potential scattering. It is found that the existence of a bootstrap solution depends crucially on the high-energy conditions, which enter the problem through a cutoff function and through subtractions in the dispersion relations. In all the models considered there is no bootstrap solution with no subtraction. With one subtraction there exists more than one bootstrap solution. However, the requirements that (a) the meson–baryon coupling constant should be different from zero, and (b) there should be no inelastic threshold below the elastic threshold, render the bootstrap solution unique. Positions of bound states and their coupling constants depend on two arbitrary parameters, which may be taken to be the cutoff momentum and the subtraction constant.

1. INTRODUCTION AND SUMMARY OF RESULTS

THIS investigation[1] is motivated by a desire to find examples of the bootstrap mechanism.[2] The physical idea behind the bootstrap is that none of the particles observed in nature is elementary, but that all are composite states of one another, so that all masses and coupling constants are determined.

It is generally accepted that the S matrix satisfies certain dispersion relations, which represent a system of coupled nonlinear integral equations for the S-matrix elements of various scattering channels, and the particles are represented by poles in the S-matrix elements in the appropriate channels. The dispersion relations themselves do not seem to possess a unique solution with a unique distribution of poles. Unless they are postulated in an *ad hoc* way, the poles can be determined only by an additional condition on the S matrix. The bootstrap mechanism, being an example of such a condition, is therefore not a property of the equations satisfied by the S matrix, but a property of the solution.

Since it is difficult to consider the complete S matrix, efforts[2] in the practical implementation of the bootstrap idea have concerned themselves with a small number of scattering channels, such as π–π scattering and π–n scattering. The coupling of the selected channels to other channels is either neglected, or taken into account through some cutoff parameters. In so doing, one may have to allow the possibility that bootstrap solutions, whatever they are defined to be, may not be unique, and that the nonuniqueness reflects the insufficient account taken of other channels.

In this investigation we study some simplified models of meson–baryon scattering, with the purpose of testing whether a reasonable, precisely defined bootstrap criterion does limit the choice of a solution. We study five well-known models of meson–baryon scattering in the nonrecoil, one-meson approximation: (a) neutral scalar theory, (b) charged scalar theory, (c) symmetric scalar theory, (d) neutral pseudoscalar theory, (e) special case of symmetric pseudoscalar theory.

Each of these models has a crossing matrix that is appropriate to the intuitive idea of a bootstrap mechanism, i.e., exchange of particle in a channel produces attraction in a crossed channel, and vice versa. [In case (a) the two channels are the same.] The S matrices of these models share with the S matrix of physical processes the general properties of elastic unitarity, crossing symmetry, and analyticity. They have however the simplifying feature that the scattering proceeds via a unique state of orbital angular momentum of the meson: S wave for scalar theories, and P wave for pseudoscalar theories. This feature enables one to obtain the most general solution to the first four models, and a special class of solutions to the pseudoscalar symmetric theory, namely solutions for which the S matrix is a product of the S matrices of the neutral pseudoscalar theory. In all models the unknown interactions at high energies are taken into account by the introduction of a cutoff function, and by

* This work was supported in part through funds provided by the U. S. Atomic Energy Commission under Contract AT (30-1)2098.

[1] A summary of the results of this work has previously been published. K. Huang and F. E. Low, Phys. Rev. Letters **13**, 596 (1964).

[2] G. F. Chew and S. Mandelstam, Nuovo Cimento, **19**, 752 (1961); F. Zachariazen and C. Zemach, Phys. Rev. **128**, 849 (1962).

allowing subtractions in the dispersion relations. The cutoff function is chosen to be $v(\omega) = \kappa^{2c}/(q^2 + \kappa^2)^c$, where ω is the energy of the meson, q is the momentum of the meson, $c = 0, 1, 2, \cdots$, and κ is the cutoff momentum ($\kappa > 1$). There are therefore $2 + n$ arbitrary parameters, where n is the number of subtractions.

The criterion for a bootstrap solution is taken to be Levinson's theorem of potential scattering, which states that the difference between the scattering phase shift at infinite energy and at threshold is $-\pi$ times the number of bound states in the scattering channel in question. The motivation of this criterion is described in the text.

The results are as follows:

1. For all the models considered there exists no bootstrap solution satisfying an unsubstracted dispersion relation.

2. There exists no bootstrap solution for the charged scalar theory, except for a degenerate case which reduces to the neutral scalar theory.

3. There exists a unique bootstrap solution to the neutral scalar theory with one subtraction. The solution is characterized by three parameters: $\kappa > 1$, $c \geq 1$, and the effective subtraction constant. Depending on their values, there may either be no bound state or one bound state, which may or may not be the target baryon.

4. All bootstrap solutions to the symmetric scalar theory with one subtraction must have $c = 1$. There are however several solutions, differing in numbers and locations of bound states. The solution is rendered unique by making the physical requirement that the meson–baryon coupling constant must be different from zero (i.e., the target baryon must appear as a bound state in the $I = \frac{1}{2}$ channel). The unique solution is labeled by two arbitrary parameters: the cutoff momentum $\kappa > 1$, and an effective subtraction constant $\beta_0 \neq 0$. For $\beta_0 < 0$ there is one bound state in each of the two channels with $I = \frac{1}{2}, \frac{3}{2}$. For $\beta_0 > 0$ there is one bound state in the $I = \frac{1}{2}$ channel, and no bound state in the $I = \frac{3}{2}$ channel. The bound state in the $I = \frac{1}{2}$ channel is by construction always the target baryon.

5. All bootstrap solutions to the neutral pseudoscalar theory with one subtraction must have $c = 2$. There are again several solutions differing in numbers and locations of bound states. We can render the solution unique by imposing two physical requirements, namely (a) the meson–baryon coupling constant must be different from zero (i.e., the target baryon must appear as a bound state in the $J = \frac{1}{2}$ channel), and (b) there must be no inelastic threshold below the elastic threshold (i.e., no bound state can have a smaller mass then the target baryon). The unique solution is labeled by two arbitrary parameters: the cutoff momentum $\kappa > 1$, and an effective subtraction constant. For all possible values of these parameters there is one bound state in the $J = \frac{1}{2}$ channel, representing the target baryon, and no bound state in the $J = \frac{3}{2}$ channel. An explicit construction of the solution is easy only for a certain range of the subtraction constant, in which we find that there is a resonance in the $J = \frac{3}{2}$ channel.

6. In the special class of solutions to the symmetric pseudoscalar theory the bootstrap solutions are trivially related to the bootstrap solutions of the neutral pseudoscalar theory. They fail to contain the physically interesting solution that corresponds to low-energy π–n scattering, namely a solution with the target baryon as a bound state in the $I = \frac{1}{2}$, $J = \frac{1}{2}$ channel, and no other bound states in this or other channels.

The general conclusions drawn from these results are as follows.

(A) These models all differentiate sharply between cutoff functions of arbitrary power on the one hand, and subtracted dispersion relations on the other. That is, without subtractions no bootstrap solutions exist, no matter how strong the cutoff. This suggests that the existence and properties of bootstrap solutions of the complete S matrix depend sensitively on the finer details of high-energy phenomena.

(B) If we apply the usual bootstrap philosophy to the solutions described earlier, we would expect that the positions and coupling constants of bound states, virtual states, or resonances to satisfy a certain set of equations. For the case of two bound states with positions ω_1 and ω_2, and squared coupling constants λ_1 and λ_2, and with $\omega_1 = 0$ taken to be the target baryon, these equations should have the forms

$$0 = f_1(\omega_2, \lambda_1, \lambda_2, g, \kappa),$$
$$\omega_2 = f_2(\omega_2, \lambda_1, \lambda_2, g, \kappa),$$
$$\lambda_1 = F_1(\omega_2, \lambda_1, \lambda_2, g, \kappa),$$
$$\lambda_2 = F_2(\omega_2, \lambda_1, \lambda_2, g, \kappa),$$

where g is the subtraction constant, and κ is the cutoff momentum. The mass of the scattered meson serves only as a scaling parameter, and is taken to be unity. Since there are five variables and four equations, we should expect ω_2, λ_1, and λ_2 to be determined up to one arbitrary parameter. Our results disagree with this counting, for there are two arbitrary parameters in all the solutions obtained.

This must mean that one of the equations above is an identity. We may perhaps understand the reason for this by regarding the static models to be the limit of some relativistic model as the mass M of the target baryon tends to infinity. We should then rewrite the equations in the forms

$$M = M + f_1(\omega_2, \lambda_1, \lambda_2, g, \kappa, M),$$
$$M + \omega_2 = M + f_2(\omega_2, \lambda_1, \lambda_2, g, \kappa, M),$$
$$\lambda_1 = F_1(\omega_2, \lambda_1, \lambda_2, g, \kappa, M),$$
$$\lambda_2 = F_2(\omega_2, \lambda_1, \lambda_2, g, \kappa, M).$$

For the static models to be valid, the functions f_i and F_i must be insensitive functions of M, as $M \to \infty$. That is, these equations can be satisfied for whatever M, as long as M is sufficiently large. Consequently, one of them must become an identity when we pass to the limit $M \to \infty$.

We end this section by listing the unique bootstrap solutions referred to earlier.

Neutral Scalar Theory

Let the scattering amplitude be $f(\omega) = v(\omega)h(\omega)$, where $v(\omega)$ is the cutoff function given by $v(\omega) = \kappa^{2c}/(q^2 + \kappa^2)^c$, with ω being the energy of the meson, and $q = (\omega^2 - 1)^{\frac{1}{2}}$. Then

$$h(\omega) = [g - \omega^2 L(\omega^2)]^{-1},$$

where g is a positive constant and

$$L(\omega^2) = \frac{2\kappa^{2c}}{\pi}\int_0^\infty dp\, p^2[(p^2+1)(p^2+1-\omega^2) \times (p^2+\kappa^2)^c]^{-1}.$$

For $g > L(1)$ there is no bound state. For $g < L(1)$ there is one bound at $\omega_0 \neq 0$. For $g = 0$ there is a double pole in $h(\omega)$ at $\omega = 0$, which represents the target baryon. With $g = 0$, the two arbitrary parameters are $c = 1, 2, 3, \cdots$, and $\kappa > 1$.

Symmetric Scalar Theory

There are two scattering channels with $I = \frac{1}{2}, \frac{3}{2}$. The respective S-matrix elements are

$$S_1 = [B/(B-1)][(B-2)/(B+1)]D,$$
$$S_3 = [B/(B-1)]D,$$

where

$$B(\omega) = \tfrac{1}{2} + i[\pi^{-1}\log(\omega+q) - (\omega/q)\beta_0] \quad (\beta_0 \neq 0),$$

$$D(\omega) = \begin{cases} \dfrac{(1-iq)(1-iq/\kappa)(1-iq/s_2)(1+iq/s_1)}{(1+iq)(1+iq/\kappa)(1+iq/s_2)(1-iq/s_1)}, & (\beta_0 < 0) \\ \dfrac{(1-iq)(1-iq/\kappa)(1-iq/s_0)(1+iq/s_1)}{(1+iq)(1+iq/\kappa)(1+iq/s_0)(1-iq/s_1)}, & (\beta_0 > 0) \end{cases}$$

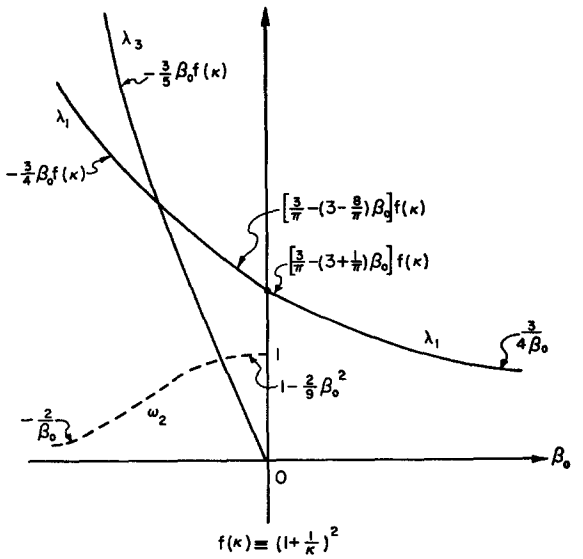


FIG. 1. Coupling constants and bound states in the bootstrap solution of the symmetric scalar theory. The two parameters labeling the solution are the cutoff momentum κ and the effective subtraction constant β_0. The bound state in the $I = \frac{1}{2}$ channel is always at $\omega = 0$, with squared coupling constant λ_1. The bound state in the $I = \frac{3}{2}$ channel is at ω_2, which exists only for $\beta_0 < 0$. Its squared coupling constant is λ_3. The formulas attached to various portions of the curves give the asymptotic forms of the curves.

with $s_\gamma = (1 - \omega_\gamma^2)^{\frac{1}{2}}$, where ω_γ is the unique root of $B(\omega_\gamma) = \gamma$ satisfying the following conditions: For $\gamma = 0$, either $0 < \omega_0 < 1$, or ω_0 is pure imaginary. For $\gamma \geq 1$, $0 < \omega_\gamma < 1$.

For $\beta_0 < 0$, there is a bound state in the $I = \frac{1}{2}$ channel at $\omega = 0$ representing the target baryon, and a bound state in the $I = \frac{3}{2}$ channel at $\omega = \omega_2$. The respective squared coupling constants are

$$\lambda_1 = 3\left(\frac{1}{\pi} - \beta_0\right)\frac{(1-s_1)(1+s_2)}{(1+s_1)(1-s_2)}\left(1+\frac{1}{\kappa}\right)^2,$$

$$\lambda_3 = \frac{2s_2(1+s_2)(s_1-s_2)(\kappa+s_2)}{\omega_2(1-s_2)(s_1+s_2)(\kappa-s_2)}\left(1-\frac{1}{\kappa^2}\right).$$

There are positive because $0 < s_2 < s_1$.

For $\beta_0 > 0$, there is only one bound state, i.e., the target baryon. The meson–baryon squared coupling constant is in this case

$$\lambda_1 = 3\left(\frac{1}{\pi} - \beta_0\right)\frac{(1+s_0)(1-s_1)}{(1-s_0)(1+s_1)}\left(1+\frac{1}{\kappa}\right)^2,$$

It is positive because $0 < s_1 < 1$, and $0 < s_0 < 1$ for $\beta_0 < 1/\pi$, $s_0 > 1$ for $\beta_0 > 1/\pi$.

The two arbitrary parameters are the cutoff momentum $\kappa > 1$, and the effective subtraction constant $\beta_0 \neq 0$. Graphs for λ_1, λ_3, and ω_2 are given in Fig. 1.

Neutral Pseudoscalar Theory

There are two scattering channels with $J = \frac{1}{2}, \frac{3}{2}$. The respective S matrix elements are

$$S_1 = [B/(B-1)][(B-2)/(B+1)]D,$$
$$S_3 = [B/(B-1)]D,$$

where

$$B(\omega) = \tfrac{1}{2} + i[\pi^{-1}\log(\omega+q) - (\omega/q^3)(\beta_0+\beta_1\omega^2)] \qquad (\beta_0+\beta_1 \neq 0).$$

The two constants β_0 and β_1 are related by a threshold condition given below. The function $D(\omega)$ is a real analytic rational function of q, such that $D(1)=1$, $|D(\omega)| = 1$ for $\omega \geq 1$. It is completely determined by the conditions that

$$D(\omega) \text{ have } \begin{cases} \text{double poles at } \pm i(\kappa^2-1)^{\frac{1}{2}}, \\ \text{simple poles at all the roots of } B(\omega)=0, \\ \text{simple zeroes at all the roots of } B(\omega)=-1. \end{cases}$$

The threshold condition is

$$D(\omega) \underset{\omega\to 1}{\longrightarrow} 1 + O(q^3),$$

which is a relation between β_0 and β_1 for given κ. For all β_0, β_1, and κ, there is only the target baryon bound state at $\omega = 0$.

For $\beta_0 > 0$, $\beta_1 > 0$, the roots of $B(\omega) = \gamma$ are located as follows: $B = 0$ at $\omega = 0, \pm x, \pm x^*$, where x is complex; $B = -1$ at $\omega = -\omega_1, y, y^*$, where $0 < \omega_1 < 1$, and y is complex with $\mathrm{Re}\, y > 1$. Letting $s_1 = (1-\omega_1^2)^{\frac{1}{2}}$, $q_0 = (x^2-1)^{\frac{1}{2}}$, $q_1 = (y^2-1)^{\frac{1}{2}}$, we have

$$D(\omega) = \left(\frac{1-iq/\kappa}{1+iq/\kappa}\right)^2 \frac{(1-iq)(1+iq/s_1)(1+q/q_0)(1-q/q_0^*)(1-q/q_1)(1+q/q_1^*)}{(1+iq)(1-iq/s_1)(1+q/q_0^*)(1-q/q_0)(1-q/q_1^*)(1+q/q_1)}.$$

The meson–baryon coupling constant is

$$\lambda_1 = 3\left(\frac{1}{\pi}+\beta_0\right)\frac{1-s_1}{1+s_1}\left(1+\frac{1}{\kappa}\right)^4.$$

The threshold condition reads

$$s_1^{-1} + 2\,\mathrm{Im}\,(q_0^{-1} - q_1^{-1}) = 1 + 2\kappa^{-1}.$$

For $\kappa \sim 1$, this reduces to the approximate relation

$$\beta_1 + \beta_0 \approx \tfrac{3}{2}(1+\kappa^{-1})^{-3}.$$

There is a resonance in the $J = \frac{3}{2}$ channel represented by a pole at $\omega = y$ on the second Riemann sheet. For $\kappa \sim 1$ it is near threshold, and the position and width of the resonance are respectively given by the approximate expressions

$$\mathrm{Re}\, y \approx 1 + \tfrac{1}{4}(1+\kappa^{-1})^{-2},$$
$$\mathrm{Im}\, y \approx \tfrac{1}{4}\sqrt{3}\,(1+\kappa^{-1})^{-2}.$$

The two arbitrary parameters are the cutoff momentum $\kappa > 1$, and the effective subtraction constant β_0 (or β_1).

2. DEFINITION OF THE PROBLEM

A. The Static Models

We state the mathematical problem common to all the models considered in this investigation. Let ω be the energy of the meson, and q its momentum[3]:

$$q = (\omega^2-1)^{\frac{1}{2}}. \tag{1}$$

In the complex ω plane, the branch cuts of q are chosen to run from 1 to ∞, and from -1 to $-\infty$, and q is defined to be real and positive just above the cut $(1, \infty)$. It follows from this definition that on the first Riemann sheet

$$q^*(\omega) = -q(\omega^*), \qquad \mathrm{Im}\, q \geq 0, \tag{2}$$

so that iq is a real analytic function in the cut ω plane. Let $S_\alpha(\omega)$ be the S-matrix element for the channel α (which refers to total spin or isospin depending on the particular model). It is customary to write

$$S_\alpha(\omega) = 1 + 2iq^{2l+1}v(\omega)h_\alpha(\omega), \tag{3}$$

where l is the orbital angular momentum, and $v(\omega)$ is a cutoff function, chosen for definiteness to be

$$v(\omega) = \kappa^{2c}/(q^2+\kappa^2)^c, \qquad \kappa > 1, \quad c = 0, 1, 2, \cdots. \tag{4}$$

The condition $\kappa > 1$ ensures that $v(\omega)$ does not introduce poles of S_α on the real axis. The function $h_\alpha(\omega)$ satisfies a dispersion relation expressing the requirements of elastic unitarity, crossing symmetry, and analyticity, which may be taken as basic postulates, or as consequences of a Lagrangian theory[4,5]:

$$h_\alpha(\omega) = P_\alpha(\omega) + \frac{1}{\pi}\int_1^\infty d\omega'\,(q')^{2l+1}v(\omega') \times \left[\frac{|h_\alpha(\omega')|^2}{\omega'-\omega} + \sum_\beta A_{\alpha\beta}\frac{|h_\beta(\omega')|^2}{\omega'+\omega}\right], \tag{5}$$

[3] We use units in which $\hbar = c = m = 1$, where m is the meson mass.

[4] F. E. Low, Phys. Rev. **97**, 1392 (1955).

[5] G. F. Chew and F. E. Low, Phys. Rev. **101**, 1570 (1956).

where $A_{\alpha\beta}$ is a crossing matrix with the general property

$$\sum_{\gamma} A_{\alpha\gamma}A_{\gamma\beta} = \delta_{\alpha\beta}. \tag{6}$$

It is specified later for the different models. The term P_α is a sum of poles located on the real ω axis between ± 1:

$$P_\alpha(\omega) = \sum_i \left[\frac{\lambda_{i\alpha}}{\omega_i - \omega} + \sum_\beta A_{\alpha\beta}\frac{\lambda_{i\beta}}{\omega_i + \omega}\right], \qquad \lambda_{i\alpha} \geq 0, \quad |\omega_i| < 1. \tag{7}$$

The physical value of $h_\alpha(\omega)$, for real $\omega \geq 1$, is defined to be the limit of $h_\alpha(\omega + i\epsilon)$ as $\epsilon \to 0^+$.

We also consider cases in which (5) is replaced by a subtracted dispersion relation, because subtractions, together with the cutoff function, are the means at our disposal to take into account the high-energy interactions not treated in detail in the static models. For the case of one subtraction, (5) should be replaced by

$$h_\alpha(\omega) = P_\alpha(\omega) + C_\alpha + \frac{\omega}{\pi}\int_1^\infty \frac{d\omega'}{\omega'}(q')^{2l+1}v(\omega') \times \left[\frac{|h_\alpha(\omega')|^2}{\omega' - \omega} - \sum_\beta A_{\alpha\beta}\frac{|h_\beta(\omega')|^2}{\omega' + \omega}\right], \tag{5'}$$

where the subtraction is made at $\omega = 0$, and C_α are the subtraction constants. If the subtraction were made at a different point, the dispersion integral would differ from the above only by a finite constant, which can be absorbed into C_α. By crossing symmetry, $\sum_\beta A_{\alpha\beta}C_\beta = C_\alpha$. Hence a subtraction introduces only one arbitrary parameter.

Each nonvanishing positive value of $\lambda_{i\alpha}$ in (7) gives rise to a bound state in the channel α, with $\lambda_{i\alpha}$ as the squared coupling constant. The bound-state energy ω_i is physically the mass difference between the bound state and the target baryon. Thus, if the target baryon occurs as a bound state in some channel, then the corresponding ω_i is zero. If $\lambda_{i\alpha} < 0$ the pole would be a ghost state, and must be excluded. The first term of (7) displays all the bound states in channel α, while the second term arises from crossing symmetry. In a relativistic theory the crossed poles in P_α appear as short cuts. The only poles of $h_\alpha(\omega)$ in the cut ω plane are those in $P_\alpha(\omega)$. However, $h_\alpha(\omega)$ may have poles on other Riemann sheets. In particular those on the second sheet (reached from the physical sheet by crossing the right-hand cut), have simple physical meaning if they lie close to the real axis. Namely, they represent resonant states if Re $\omega > 1$, virtual states if Re $\omega < 1$.

The physically admissible solutions of (5) must have the proper threshold behavior and high-energy behavior. The former requires $h_\alpha(1)$ to be finite, in order that

$$S_\alpha(\omega) \underset{\omega\to 1}{\to} 1 + O(q^{2l+1}). \tag{8}$$

The high-energy condition depends on the number of subtractions in the dispersion relation. For the case of no subtraction, as expressed by (5), the requirement that the dispersion relation be a consequence of the analyticity of $h_\alpha(\omega)$ and Cauchy's theorem implies

$$h_\alpha(\omega) \underset{\omega\to\infty}{\to} 0. \tag{9}$$

For $\omega \geq 1$, the unitarity condition embodied in the dispersion relation states Im $h_\alpha = q^{2l+1}v\,|h_\alpha|^2$, which implies

$$|h_\alpha(\omega)| \leq [q^{2l+1}v(\omega)]^{-1} \qquad (\omega \geq 1). \tag{10}$$

We rule out an essential singularity at $\omega = \infty$, so that the behavior near $\omega = \infty$ is independent of directions. Therefore (10) either implies (9), or is implied by (9), depending on $v(\omega)$.

When there are subtractions the condition (9) is relaxed. We see from (10) that if $v(\omega) \equiv 1$, then no subtraction is required, and if one were made it could be removed. Hence we need subtractions only if a cutoff is introduced such that $\lim_{\omega\to\infty} q^{2l+1}v(\omega) < 1$.

The mathematical problem defined so far is self-consistent, for there exist solutions to it. However, a solution to the problem is physically reasonable only if two other requirements are met, namely,

(a) the target baryon should be a bound state in the channel with the appropriate quantum numbers,

(b) no bound state should have a smaller mass than the target baryon (i.e., no ω_i should be negative).

If (a) were not true, the meson–baryon coupling constant would be zero. If (b) were not true, then there would be an inelastic threshold below the elastic threshold. In either case the static theory with elastic unitarity would be of even more doubtful validity as a physical model than is usually the case.

B. Criterion for a Bootstrap Solution

For the physical range of energies $\omega \geq 1$, let

$$S_\alpha(\omega) = e^{2i\delta_\alpha(\omega)}. \tag{11}$$

This defines the phase shift $\delta_\alpha(\omega)$ up to an additive multiple of π, which may be fixed by an arbitrary

convention. In potential scattering, Levinson's theorem[6] states that

$$\Delta\delta_\alpha(\omega) \equiv \delta_\alpha(\infty) - \delta_\alpha(1) = -\pi b_\alpha, \qquad (12)$$

where b_α is the number of bound states in the channel α. We take over this statement as the condition for a bootstrap solution. The motivation comes from two sources: analogy to potential scattering, and current ideas on N/D solutions of dispersion relations, as we now explain.

In potential scattering Levinson's theorem implies that all bound states are results of the interaction, for the phase shift becomes identically zero (mod π) when the potential is turned off. Let us extend this interpretation of (12) to the present case. If we define an elementary particle to be a bound state that persists when the interaction is turned off, then (12) states that there are no elementary particles in any of the channels.

To see the connection with current ideas on bootstrapping, let us write $h_\alpha(\omega)$ in the so-called N/D form[7]:

$$h_\alpha(\omega) = N_\alpha(\omega)/D_\alpha(\omega), \qquad (13)$$

where $N_\alpha(\omega)$ has only the left-hand cut, and $D_\alpha(\omega)$ has only the right-hand cut. To specify N_α and D_α uniquely, we require that D_α has no pole, and that

$$\lim_{\omega\to\infty} [\log D(\omega)/\log\omega] = 0. \qquad (14)$$

Levinson's theorem is equivalent to the statement that *every bound state in the channel α is a zero of $D_\alpha(\omega)$, and vice versa.* Thus, all the bound state poles of $h_\alpha(\omega)$, which are contained in the first term in (7), are zeros of $D_\alpha(\omega)$. Furthermore, the "pseudopoles" arising from crossing symmetry, i.e., those contained in the second term in (7), are poles of $N_\alpha(\omega)$. Such an allotment of poles agrees with the view embodied in current approximate bootstrap calculations.[2] They make use of the idea based on analogy with potential scattering, that the "crossed poles" in $N_\alpha(\omega)$ furnish a generalized potential which gives rise to the bound-state zeros in $D_\alpha(\omega)$. It is then hoped that the self-consistency of this procedure would determine the energies and coupling constants of all bound states.

To show the equivalence with Levinson's theorem, let us temporarily adopt the convention $\delta_\alpha(1) = 0$. As is well-known,[7] a function $\mathfrak{D}_\alpha^{-1}$ having the same right-hand cut as $h_\alpha(\omega)$, but having no left-hand cut, is the Omnés function

$$\mathfrak{D}_\alpha(\omega) = \exp\left[-\frac{\omega-\omega_0}{\pi}\int_1^\infty d\omega' \times \frac{\delta_\alpha(\omega')}{(\omega'-\omega_0)(\omega'-\omega)}\right], \qquad (15)$$

where ω_0 is a fixed number less than 1. This function has neither pole nor zero on the physical sheet. As $\omega \to \infty$ we have[8]

$$\lim_{\omega\to\infty} [\log \mathfrak{D}_\alpha(\omega)/\log\omega] = \delta_\alpha(\infty)/\pi. \qquad (16)$$

The condition that $D_\alpha(\omega)$ has b_α zeros and no pole, and the normalization condition (14), then imply that $D_\alpha(\omega)$ is $\mathfrak{D}_\alpha(\omega)$ times a polynomial of degree b_α, with $b_\alpha + \delta_\alpha(\omega)/\pi = 0$, which is Levinson's theorem. Conversely, Levinson's theorem implies that $D_\alpha(\omega)$ has b_α zeros, which, because of (13), must be poles of $h_\alpha(\omega)$. Since b_α is by definition the number of bound states, it is possible to define $D_\alpha(\omega)$ so that its zeros are the bound states.

The mathematical problem is then the following: Given a specific static model, find out whether there exist solutions satisfying (12), for any choice of the high-energy condition, which includes a choice of the cutoff function $v(\omega)$ and the number of subtractions, with the restriction that if $v(\omega) \equiv 1$ then there must be no subtraction.

3. NEUTRAL SCALAR THEORY AND CHARGED SCALAR THEORY

A. General Solution

The neutral scalar theory describes the scattering of a neutral scalar meson by a fixed baryon. There is only one channel, and the crossing matrix is 1. Its solutions have been considered by Lee and Serber,[9] and Castillejo, Dalitz, and Dyson.[10] Mathematically it may be considered a special case of the charged scalar theory.

In the charged scalar theory there are two scalar mesons of opposite charge, which are coupled in a charge-symmetric manner to a charged baryon. There are two channels with $\alpha = \pm 1$, referring respectively to the charge of the meson. The crossing matrix is

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad (17)$$

so that crossing symmetry simply means $h_1(-\omega) =$

[6] N. Levinson, Mat. Fys. Medd. Dan. Vid. Selsk. **25**, No. 9 (1949).

[7] See, for example, G. F. Chew, *S-Matrix Theory of Strong Interactions* (W. A. Benjamin, Inc., New York, 1961), Chap. 10.

[8] M. Sugawara and A. Kanazawa, Phys. Rev. **126**, 2251 (1962), Appendix 2.

[9] T. D. Lee and R. Serber (unpublished). See Ref. 10.

[10] L. Castillejo, R. H. Dalitz, and F. J. Dyson, Phys. Rev. **101**, 453 (1956).

$h_{-1}(\omega)$. Castillejo, Dalitz, and Dyson[10] have shown that the most general solutions are

$$\frac{1}{h_1(\omega)} = -\frac{2\omega^2}{\pi}\int_1^\infty \frac{d\omega'}{\omega'}\frac{q'v(\omega')}{\omega'^2-\omega^2} + \omega R(\omega),$$
$$\frac{1}{h_{-1}(\omega)} = -\frac{2\omega^2}{\pi}\int_1^\infty \frac{d\omega'}{\omega'}\frac{q'v(\omega')}{\omega'^2-\omega^2} - \omega R(-\omega), \tag{18}$$

where $v(\omega)$ is given by (4), and $R(\omega)$ is a meromorphic function with the property $R^*(\omega) = R(\omega^*)$. The dispersion relation for this model implies that Im (ωh_α)/Im ω is positive definite, hence $h_\alpha(\omega)$ has no zero in the complex plane. This means that $R(\omega)$ can have poles only on the real axis. The requirement that $h_\alpha(\omega)$ has no pole in the complex plane is satisfied by requiring Im $(\omega h_\alpha)^{-1}$/Im $\omega \neq 0$ for Im $\omega \neq 0$, which leads to the form

$$R(\omega) = -a\omega + d - \sum_n \frac{g_n}{\epsilon_n - \omega}, \qquad g_n \geq 0, \quad a \geq 0. \tag{19}$$

The poles ϵ_n will be referred to as CDD poles.[11]

It is easily seen that the threshold condition (8) is satisfied. As $\omega \to \infty$, $h_\alpha(\omega)$ approach either 0 or a constant. Hence we need at most one subtraction in the dispersion relation. The high-energy requirements are summarized in the following table:

	Asymp. form	Cutoff	No. subtraction
$a > 0$	ω^{-2}	$c \geq 0$	0
$a = 0, d \neq 0$	ω^{-1}	$c \geq 0$	0
$a = d = 0$	ω^{-1}	$c = 0$	0
	constant	$c \geq 1$	1.

(20)

The $h(\omega)$ for the neutral scalar theory is given by either formula in (18) with $R(\omega)$ chosen to be an odd function. It is therefore a special case that need not be treated separately.

B. Phase Shifts

The phase shifts are obtainable from (18) by using the formula $h_\alpha = (\sin \delta_\alpha)(\exp i\,\delta_\alpha)/qv$, which leads to

$$\cot \delta_1(\omega) = [qv(\omega)]^{-1}[J(\omega) + \omega R(\omega)],$$
$$\cot \delta_{-1}(\omega) = [qv(\omega)]^{-1}[J(\omega) - \omega R(-\omega)], \tag{21}$$

where

$$J(\omega) = -\frac{2\omega^2}{\pi} P\int_1^\infty \frac{d\omega'}{\omega'}\frac{q'v(\omega')}{\omega'^2-\omega^2} = -\frac{2\omega^2\kappa^{2c}}{\pi} P \times \int_0^\infty dp \frac{p^2}{(p^2+1)(p^2+1-\omega^2)(p^2+\kappa^2)^c}, \tag{22}$$

in which (4) has been used. The bootstrap criterion requires $\Delta\delta_\alpha/\pi = -b_\alpha$. Since b_α, the number of bound states, must be a nonnegative integer, we immediately rule out solutions for which $\Delta\delta_\alpha > 0$, or for which $-\Delta\delta_\alpha/\pi$ is not an integer. From (21) we find

$$\cot \delta_1(\omega) \underset{\omega\to 1}{\to} [qh_1(1)]^{-1}, \qquad \cot \delta_1(\omega) \underset{\omega\to\infty}{\to} \omega^{2c-1}(-a\omega^2 + d\omega + Q),$$
$$\cot \delta_{-1}(\omega) \underset{\omega\to 1}{\to} [qh_1(-1)]^{-1}, \qquad \cot \delta_{-1}(\omega) \underset{\omega\to\infty}{\to} \omega^{2c-1}(-a\omega^2 - d\omega + Q), \tag{23}$$

where Q is the positive-definite number

$$Q = \frac{2\kappa^{2c}}{\pi}\int_0^\infty dp \frac{p^2}{(p^2+1)(p^2+\kappa^2)^c} + \sum_n g_n. \tag{24}$$

If there is no cutoff ($c = 0$) we must have $a > 0$ in order that $\Delta\delta_\alpha/\pi$ be an integer. We note that the expression (24) diverges if there is no cutoff. In that event Q in (23) should be replaced by the number -1, which however may be neglected because $a > 0$. We may therefore use (23) as it stands for all c.

According to (21) each CDD pole with $\epsilon_n > 1$ contributes π to $\Delta\delta_1$, and each with $\epsilon_n < -1$ contributes π to $\Delta\delta_{-1}$. Let

$$n_+ = \text{No. of CDD poles with } \epsilon_n > 1,$$
$$n_- = \text{No. of CDD poles with } \epsilon_n < -1. \tag{25}$$

With the help of (23), we construct Tables I and II, which are self-explanatory.

An examination of Tables I and II shows that, with respect to the choice of a and d, there are only two possibilities: either $a > 0$, or $a = 0$ and $d = 0$. To see whether the bootstrap criterion can actually be satisfied for either case we turn to an examination of poles and bound states.

C. Poles and Bound States

The poles of $h_\alpha(\omega)$ are exhibited in the pole term (7), which we rewrite for the charged scalar theory in the form

$$P_1(\omega) = -\frac{\lambda_0}{\omega} - \sum_n \frac{\lambda_n}{\omega-\omega_n} + \sum_m \frac{\lambda'_m}{\omega+\omega'_m},$$
$$P_{-1}(\omega) = \frac{\lambda_0}{\omega} - \sum_m \frac{\lambda'_m}{\omega-\omega'_m} + \sum_n \frac{\lambda_n}{\omega+\omega_n}, \tag{26}$$
$$\lambda_n \geq 0, \quad \lambda'_m \geq 0, \quad 0 < |\omega_n| < 1, \quad 0 < |\omega'_m| < 1.$$

Clearly h_1 and h_{-1} must have the same number of poles. The number of bound states in each

[11] Note that ϵ_n is a zero of $h_\alpha(\omega)$ if and only if $\epsilon_n \neq 0$.

TABLE I. $\Delta\delta_1 \equiv \delta_1(\infty) - \delta_1(1)$, charged scalar theory.

$\cot \delta_1(1)$	$\cot \delta_1(\infty)$	$\Delta\delta_1/\pi$	Conditions	Bootstrap criterion
$+\infty$	$+\infty$	n_+	$h_1(1) > 0$ $a = 0, d \geq 0$	$n_+ = 0, b_1 = 0$
$-\infty$	$+\infty$	$n_+ - 1$	$h_1(1) < 0$ $a = 0, d \geq 0$	either $n_+ = 0, b_1 = 1$ or $n_+ = 1, b_1 = 0$
$+\infty$	$-\infty$	$n_+ + 1$	$h_1(1) > 0$ either $a > 0$ or $a = 0, d < 0$	Impossible
$-\infty$	$-\infty$	n_+	$h_1(1) < 0$ either $a > 0$ or $a = 0, d < 0$	$n_+ = 0, b_1 = 0$

channel can be seen by examining $P_1(\omega)$ alone: a pole of $P_1(\omega)$ at $\omega \neq 0$ is a bound state in channel 1 if its residue is negative; otherwise it is a bound state in channel -1. The pole of $P_1(\omega)$ at $\omega = 0$ follows a different rule: if $\lambda_0 > 0$, it is a bound state either in channel 1, or in both channels; if $\lambda_0 < 0$, it is a bound state either in channel -1, or in both channels. The choice involved can be made arbitrarily. Since this theory does not distinguish between a short cut and a pole, its simple crossing structure makes it impossible to recognize a ghost state with $|\omega| < 1$. This is not true of the symmetric scalar and neutral pseudoscalar theories (discussed later) except for the isolated point $\omega = 0$.

To find the number of poles in $h_\alpha(\omega)$, given $R(\omega)$ and $v(\omega)$, it proves convenient to consider the function

$$H_\alpha(\omega) = \omega h_\alpha(\omega). \tag{28}$$

We first find the number of poles of $H_\alpha(\omega)$ by finding the number of zeros of $1/H_\alpha(\omega)$, through an appropriate application of the formula

$$\frac{1}{2\pi i}\oint d\omega \frac{f'(\omega)}{f(\omega)} = z - p, \tag{29}$$

where z and p are, respectively, the number of zeros and poles enclosed in the contour of integration.[12] It is shown in Appendix A that

$$\text{No. of poles of } H_\alpha = M + \sigma, \tag{30}$$

where

$$M = \text{No. of CDD poles with } |\epsilon_n| \leq 1, \tag{31}$$

$$\sigma = \begin{cases} 1 & \text{if } h_1(1) < 0, \text{ and } h_1(-1) < 0, \\ 0 & \text{if } h_1(1)h_1(-1) < 0, \\ -1 & \text{if } h_1(1) > 0, \text{ and } h_1(-1) > 0. \end{cases} \tag{32}$$

If $H_\alpha(0) \neq 0$, then $h_\alpha(\omega)$ has one more pole than $H_\alpha(\omega)$; if $H_\alpha(0) = 0$, then $h_\alpha(\omega)$ has the same number of poles as $H_\alpha(\omega)$. In the former case the target baryon is a bound state; in the latter case it is not. Thus,

$$\text{No. of poles of } h_\alpha = \begin{cases} M + \sigma + 1 & \text{if } \omega = 0 \text{ is a pole,} \\ M + \sigma & \text{otherwise.} \end{cases} \tag{33}$$

The latter is the case if and only if there is a CDD pole with $\epsilon_n = 0$.

D. Bootstrap Solutions

We refer to Tables I and II, and recall that the bootstrap criterion rules out all but two possibilities: either $a > 0$, or $a = d = 0$. For $a > 0$, we must have $h_1(1) < 0$, $h_1(-1) < 0$; hence by (32) $\sigma = 1$. By (30) $h_\alpha(\omega)$ has at least one pole. This is inconsistent with $b_1 = b_{-1} = 0$. Therefore this case is excluded.

TABLE II. $\Delta\delta_{-1} \equiv \delta_{-1}(\infty) - \delta_{-1}(1)$, charged scalar theory.

$\cot \delta_{-1}(1)$	$\cot \delta_{-1}(\infty)$	$\Delta\delta_{-1}/\pi$	Conditions	Bootstrap criterion
$+\infty$	$+\infty$	n_-	$h_1(-1) > 0$ $a = 0, d \leq 0$	$n_- = 0, b_{-1} = 0$
$-\infty$	$+\infty$	$n_- - 1$	$h_1(-1) < 0$ $a = 0, d \leq 0$	either $n_- = 0, b_{-1} = 1$ or $n_- = 1, b_{-1} = 0$
$+\infty$	$-\infty$	$n_- + 1$	$h_1(-1) > 0$ either $a > 0$ or $a = 0, d > 0$	Impossible
$-\infty$	$-\infty$	n_-	$h_1(-1) < 0$ either $a > 0$ or $a = 0, d > 0$	$n = 0, b_{-1} = 0$

[12] We always count a pole or a zero as often as its order.

We consider from now on the case $a = d = 0$, for which a cutoff is required in order that $-\Delta\delta_\alpha/\pi$ be an integer. The solution (18) becomes

$$\frac{1}{h_1(\omega)} = -\omega^2 L(\omega^2) + \omega \sum_n \frac{g_n}{\omega - \epsilon_n}, \qquad (34)$$

$$\frac{1}{h_{-1}(\omega)} = -\omega^2 L(\omega^2) + \omega \sum_n \frac{g_n}{\omega + \epsilon_n}, \qquad (35)$$

where

$$L(\omega^2) = \frac{2\kappa^{2c}}{\pi} \int_0^\infty dp \, \frac{p^2}{(p^2+1)(p^2+1-\omega^2)(p^2+\kappa^2)^c}. \qquad (36)$$

If the CDD poles are symmetrically placed about $\omega = 0$ then $h_1(\omega) = h_{-1}(\omega)$, and the theory reduces to the neutral scalar theory. According to Tables I and II, there are only four possibilities for the choice of (n_+, n_-), namely (0, 0), (1, 0), (0, 1), (1, 1). Of these (1, 0) and (0, 1) are mathematically equivalent. In Table III we display all the distinct possibilities, with the corresponding requirements for $h_1(\pm 1)$ and $b_{\pm 1}$, which are read off Tables I and II. The number of poles of $h_\alpha(\omega)$ is also listed. They are obtained from (32) and (34) by taking $N = 0$, which is a necessary condition because $b_{\pm 1} \leq 1$.

The last column in Table III gives the bootstrap criterion. The last three cases are clearly impossible because there are more poles than bound states. The other two impossible cases fail to meet the bootstrap criterion for the following reason. The number of poles is $M + 1$ if $\omega = 0$ is a pole; otherwise it is M. If $\omega = 0$ is a pole, then $M = 0$, and from (35) we can show $h_1(1) = h_1(-1) < 0$. If $\omega = 0$ is not a pole, then $M = 1$, and the CDD pole must have $\epsilon_n = 0$. Again we can show from (35) that $h_1(1)h_1(-1) > 0$. This is inconsistent with $h_1(1)h_1(-1) < 0$, as required in Table III.

TABLE III. Bootstrap criterion, charged scalar theory.

n_+	n_-	$h_1(1)$	$h_1(-1)$	No. of Poles [Eq. (33)]	b_1	b_{-1}	Bootstrap criterion
0	0	+	+	$M - 1$	0	0	$M = 1$
		−	−	$M + \begin{cases} 2^a \\ 1 \end{cases}$	1	1	$M = \begin{cases} 0^* \\ 1 \end{cases}$
		+	−	$M + \begin{cases} 1^a \\ 0 \end{cases}$	0	1	
		−	+		1	0	Impossible
1	0	−	+	≥ 2	0	0	Impossible
		−	−	≥ 3	0	1	Impossible
1	1	−	−	≥ 4	0	0	Impossible

[a]Upper number applies if $\omega = 0$ is a pole; otherwise lower number applies.

There now remains the top two rows in Table III, which indeed fulfill the bootstrap criterion. They represent three cases in all, with $M = 0$ if $\omega = 0$ is a pole of $h_\alpha(\omega)$, and $M = 1$ otherwise. In the latter case the CDD pole must have $\epsilon_n = 0$. Thus all three cases reduce to the neutral scalar theory, with

$$h(\omega) = [g - \omega^2 L(\omega^2)]^{-1} \qquad (g \geq 0), \qquad (37)$$

where $L(\omega^2)$ is given by (36). We note that $h(\omega)$ has no zero on the physical sheet. The three cases mentioned above correspond to different values of g:

$$\begin{aligned} g > L(1): &\quad \text{No bound state,} \\ 0 < g < L(1): &\quad \text{Bound state at } \omega_0, \quad 0 < |\omega_0| < 1, \\ g = 0: &\quad \text{Bound state at } \omega = 0. \end{aligned} \qquad (38)$$

The residue of the pole has the correct sign at either $+\omega_0$ or $-\omega_0$. The target baryon is a bound state only in the limiting case $g = 0$, where it is represented by a second-order pole at $\omega = 0$. Presumably the second-order pole is to be viewed physically as a near cancellation between the true pole and the short cut. As g increases from 0, the second-order pole splits into two simple poles, which eventually move to the second Riemann sheet. As $\omega \to \infty$, $h(\omega)$ approaches a constant, and therefore satisfies a once-subtracted dispersion relation. There are three undetermined parameters: $g \geq 0$, $\kappa > 1$, and $c = 1, 2, 3, \cdots$. There are no bootstrap solutions other than (37).

In conclusion, we find that there is no bootstrap solution for the charged scalar theory, except for the degenerate case $h_1 = h_{-1} = h$, where h is a unique family of bootstrap solutions for the neutral scalar theory (Eq. 37).

4. SYMMETRIC SCALAR THEORY

A. General Solution

The symmetric scalar theory describes the scattering of a scalar S-wave meson of isospin 1 from a fixed baryon of isospin $\frac{1}{2}$, with conservation of total isospin. There are two channels labeled by $\alpha = 1, 3$, corresponding respectively to total isospin $\frac{1}{2}$ and $\frac{3}{2}$. The crossing matrix is

$$A = \frac{1}{3}\begin{bmatrix} -1 & 4 \\ 2 & 1 \end{bmatrix}. \qquad (39)$$

Wilson[13] has studied some special solutions to the

[13] K. Wilson, Ph.D. thesis, Physics Department, California Institute of Technology (1961) (unpublished).

model. The most general solution is given by Wanders,[14] who considers the mathematically equivalent problem of a the neutral pseudoscalar theory. We briefly state the Wanders solution for this problem.

The solution is given in terms of the S-matrix element S_α, instead of the function h_α:

$$S_1 = [B/(B-1)][(B-2)/(B+1)]D, \qquad S_3 = [B/(B-1)]D, \tag{40}$$

where

$$B(\omega) = \tfrac{1}{2} + i[\pi^{-1}\log(\omega+q) - (\omega/q)\beta(\omega)] = \pi^{-1}\sin^{-1}\omega - i(\omega/q)\beta(\omega), \tag{41}$$

$$D(\omega) = [1 - iq\alpha(\omega)]/[1 + iq\alpha(\omega)], \tag{42}$$

where $\alpha(\omega)$ and $\beta(\omega)$ are two arbitrary even meromorphic functions of ω ,with $\alpha^*(\omega) = \alpha(\omega^*)$, $\beta^*(\omega) = \beta(\omega^*)$. Thus B is an odd function of ω, and D is an even function of ω. The solution so far is the most general consistent with unitarity, crossing symmetry, and the property $S_\alpha^*(\omega) = S_\alpha(\omega^*)$, which is a consequence of (5). The functions α and β are to be further restricted so as to satisfy other requirements of the model.

It is convenient to rewrite D by noting that it is a meromorphic function of q. We may represent it as a rational function of q, times an entire function with no zero, which may be written as $\exp[iqf(q)]$, with $f^*(q) = f(-q)$. Requiring D to have no essential singularity at $\omega = \infty$ reduces f to a positive constant, which must be zero in order that there be at most a finite number of resonances in the scattering. Thus D is a rational function of q. We see from (42) that $D^*(\omega) = D(\omega^*)$, and that $|D(\omega)| = 1$ for real $\omega \geq 1$. Therefore D may be represented in the form

$$D(\omega) = \prod_m \frac{1 - ir_m q}{1 + ir_m q} \prod_n \frac{(1 - a_n q)(1 + a_n^* q)}{(1 + a_n q)(1 - a_n q)}, \qquad \operatorname{Im} r_m = 0, \quad \operatorname{Re} a_n > 0. \tag{43}$$

The over-all normalization is so chosen that

$$D \underset{q\to 0}{\longrightarrow} 1,$$

which is necessary for the threshold condition

$$S_\alpha \underset{q\to 0}{\longrightarrow} 1.$$

The choice of the arbitrary function α in (42) is now replaced by the choice of the sets of numbers $\{r_m\}$ and $\{a_n\}$. We note that each a_n with $\operatorname{Im} a_n > 0$ (< 0) gives rise to four complex poles (zeros), of $D(\omega)$ that are mirror images of each other with respect to the real and imaginary axes. Each $r_m > 0$ (< 0) gives rise to two poles (zeros) of $D(\omega)$ that are either real or pure imaginary, placed symmetrically about the origin.

The threshold condition is satisfied by requiring $\beta(1) \neq 0$. At high energies

$$B(\omega) \underset{\omega\to\infty}{\longrightarrow} (i/\pi)\log\omega - \beta(\omega), \tag{44}$$

$$D(\omega) \underset{\omega\to\infty}{\longrightarrow} 1 + d_1/\omega + d_2/\omega^2 + \cdots. \tag{45}$$

Therefore

$$h_\alpha = \frac{S_\alpha - 1}{2iqv} \underset{\omega\to\infty}{\longrightarrow} \omega^{2c-1} \times \left[\frac{\mu_\alpha}{B(\omega)} + \frac{d_1}{\omega} + \frac{d_2}{\omega^2} + \cdots\right], \tag{46}$$

where $\mu_1 = -2$, $\mu_3 = 1$, and c is defined in (4). The condition (10) is satisfied for all choices of c and $\beta(\omega)$. We note that since $\beta(\omega)$ is an even meromorphic function,

$$\beta(\omega) \underset{\omega\to\infty}{\longrightarrow} k\omega^{2n} \qquad (n = 0, \pm 1, \pm 2, \cdots). \tag{47}$$

Thus $B(\omega) \approx \log\omega$ if $n \leq 0$, and $B(\omega) \approx \omega^{2n}$ if $n > 0$.

The number of subtractions required in the dispersion relation for h_α is determined by the choice of c, n, and the numbers d_1, d_2, etc. For example, for no subtraction we have the following conditions:

$$\begin{aligned} &c = 0: && \text{No restriction on } \beta(\omega) \text{ and } D(\omega);\\ &c = 1: && n \geq 1, \quad d_1 = 0;\\ &c = 2: && n \geq 2, \quad d_1 = d_2 = d_3 = 0; \end{aligned} \tag{48}$$

For the case of K subtractions, we have the following conditions:

$c = \frac{1}{2}(1 + K)$:

No restriction on $\beta(\omega)$ and $D(\omega)$;

$c > \frac{1}{2}(1 + K)$:

$n \geq c - \frac{1}{2}(1 + K)$, plus condition on $D(\omega)$. (49)

B. Phase Shifts

For $\omega \geq 1$ the phase shifts $\delta_\alpha(\omega)$ are obtained from (40) by setting $S_\alpha(\omega) = \exp[2i\,\delta_\alpha(\omega)]$:

$$\delta_1(\omega) = \phi(\omega) - \theta(\omega) - \psi(\omega), \qquad \delta_3(\omega) = \phi(\omega) - \theta(\omega), \tag{50}$$

where ϕ, θ, and ψ are defined by

$$\frac{B}{B-1} = e^{2i\phi}, \quad \frac{B-2}{B+1} = e^{-2i\psi}, \quad D = e^{-2i\theta}, \tag{51}$$

[14] G. Wanders, Nuovo Cimento **23**, 817 (1962).

and are explicitly given by

$$\cot\phi = -2[\pi^{-1}\log(\omega+q) - (\omega/q)\beta(\omega)],$$
$$\cot\psi = \tfrac{1}{3}\cot\phi, \tag{52}$$
$$\theta = \sum_m \tan^{-1}(r_m q) + \sum_n \tan^{-1}\frac{2q\ \mathrm{Im}\ a_n}{1-q^2|a_n|^2}.$$

Let $\Delta\phi = \phi(\infty) - \phi(1)$, $\Delta\psi = \psi(\infty) - \psi(1)$, and $\Delta\theta = \theta(\infty) - \theta(1)$. From (52) we see that $\cot\phi$ is $\pm\infty$ at $\omega = 1$ and at $\omega = \infty$. Hence $\Delta\phi = \Delta\psi$, and we have

$$\Delta\delta_1 = \delta_1(\infty) - \delta_1(1) = -\Delta\theta,$$
$$\Delta\delta_3 = \delta_3(\infty) - \delta_3(1) = \Delta\phi - \Delta\theta. \tag{53}$$

We can obtain $\Delta\phi$ and $\Delta\theta$ from (52):

$$\Delta\theta/\pi = M_+ - M_- + \tfrac{1}{2}(m_+ - m_-),$$
$$\Delta\phi/\pi = N_+ - N_- + \sigma, \tag{54}$$

where

$$M_\pm = \text{No. of } a_n\text{'s in (43) with Im } a_n \gtrless 0,$$
$$m_\pm = \text{No. of } r_m\text{'s in (43) with } r_m \gtrless 0, \tag{55}$$
$$N_\pm = \text{No. of poles of } \beta(\omega) \text{ on real axis } \omega > 1, \text{ with } \pm \text{ residues},$$

and σ is 0 or ± 1, as given in Table IV. Recalling the remarks after (43), we note that $M_+(M_-)$ is $\frac{1}{4}$ the number of complex poles (zeros) of $D(\omega)$, and that $m_+(m_-)$ is $\frac{1}{2}$ the number of real or pure imaginary poles (zeros) of $D(\omega)$. The $N_+ + N_-$ poles of $\beta(\omega)$ are repeated on the real axis $\omega \leq -1$, because $\beta(\omega)$ is an even function.

C. Poles and Bound States

The poles of $h_\alpha(\omega)$ must conform to (7), from which we can read off the bound states in channel α. We see from (3) that $S_\alpha(\omega)$ has the same poles as $h_\alpha(\omega)$, plus the poles of $v(\omega)$, which are of order c, located at $\omega = \pm i(\kappa^2-1)^{\frac{1}{2}}$. These cutoff poles must be poles of $D(\omega)$, for the only other possible way to introduce a pole in both S_1 and S_3 is to make $B^2 - 1 = 0$; but this equation has no roots on the imaginary ω axis [see (62)].

TABLE IV. σ of Eq. (54).

$n \geq 1$		$n \leq 0$	
$k\beta(1) > 0$[a]	$k\beta(1) < 0$	$\beta(1) > 0$	$\beta(1) < 0$
0	$1 (k > 0)$ $-1 (k < 0)$	1	0

[a] If $\beta(\omega)$ has a pole at $\omega = 1$, the sign of $\beta(1)$ is defined to be that at $\omega = 1 + \epsilon$, $\epsilon \to 0^+$.

Apart from the cutoff poles, all other poles of $S_\alpha(\omega)$ on the physical sheet must lie on the real axis between $\omega = \pm 1$. A zero of $S_\alpha(\omega)$ on the physical sheet, on the other hand, is a pole of $S_\alpha(\omega)$ at the same point on the second sheet, because the unitarity condition may be stated in the form $S_\alpha^{(2)}(\omega) = 1/S_\alpha^{(1)}(\omega)$, where the superscripts identify the Riemann sheets.

From (40) we see that poles of $S_\alpha(\omega)$ can occur only at the poles of D and the roots of $B^2 = 1$. The converse is not true, because a pole of D may be cancelled by a root of $B = 2$, and a root of $B^2 = 1$ may be cancelled by a zero of D. In order that S_α has the poles allowed by (7) and no others, the following rules must be observed in the choice of D:

1. The poles of $D(\omega)$ can occur only at (a) $\pm i(\kappa^2-1)^{\frac{1}{2}}$, (b) roots of $B = 0$, (c) real roots of $B = 2$ lying between $\omega = \pm 1$.

2. There must be zeros of D at all complex roots of $B^2 = 1$. Other zeros of D (called extra zeros) may be freely chosen.

When poles do occur in $S_\alpha(\omega)$, we have to compare the resulting pole structure of $S_\alpha(\omega)$ with (7) to find out whether it is a bound state, and if so in which channel. The residues of bound-state poles in $S_\alpha(\omega)$ have the opposite signs to those in $h_\alpha(\omega)$, because in (3) the factor iq is negative for $|\omega| < 1$. To illustrate the structure of bound-state poles in $S_\alpha(\omega)$, it suffices to give the pole terms $\mathcal{P}_\alpha$ of $S_\alpha(\omega)$ when there is one bound state in each channel, omitting the cutoff poles. Let ω_1 and ω_3 be the respective bound-state energies, and Λ_1, Λ_3 the effective coupling constants, related to the actual squared coupling constants λ_α [as defined in (7)] by

$$\Lambda_\alpha = 2v(\omega_\alpha)(1-\omega_\alpha^2)^{\frac{1}{2}}\lambda_\alpha. \tag{56}$$

Then

$$\mathcal{P}_1 = -\frac{\Lambda_1}{\omega_1-\omega} + \frac{1}{3}\frac{\Lambda_1}{\omega_1+\omega} - \frac{4}{3}\frac{\Lambda_3}{\omega_3+\omega},$$
$$\mathcal{P}_3 = -\frac{2}{3}\frac{\Lambda_1}{\omega_1+\omega} - \frac{\Lambda_3}{\omega_3-\omega} - \frac{1}{3}\frac{\Lambda_3}{\omega_3+\omega}, \tag{57}$$
$$\Lambda_1 \geq 0, \qquad \Lambda_3 \geq 0.$$

The connection between poles and bound states are summarized in Table V.

To find the bound states, we need to know the roots of

$$B(\omega) = \gamma, \tag{58}$$

where γ is real. Since B is an odd function $B(\omega_0) = \gamma$ implies $B(-\omega_0) = -\gamma$. Since $B^*(\omega) = B(\omega^*)$, nonreal roots must occur in complex conjugate pairs.

TABLE V. Bound states in symmetric scalar theory. The constants Λ_α are related to the squared coupling constants λ_α by (56). This Table is also valid for the neutral pseudoscalar theory if we replace Λ_α by $-\Lambda_\alpha$.

Bound-state energy	Which channel	Conditions on B, D	Coupling constants
0	1, 3, or both	$B = 0$, $D = \infty$ BD has simple pole	$\Lambda_1 - \Lambda_3 = \frac{3}{2}$ res (BD)
$0 < \lvert\omega_0\rvert < 1$	ω_0 in 1, and $-\omega_0$ in 3, or vice versa	$B = 0$, $D = \infty$ BD has simple pole	$\Lambda_3 = \frac{1}{2}\Lambda_1 = \begin{cases} 3 \text{ res } (BD) \\ -3 \text{ res } (BD) \end{cases}$
	1	$B + 1 = 0$ (simple zero) $D \neq 0, \infty$	$\Lambda_1 = -\frac{3}{2}$ res $[D/(B+1)]$
	3	$B - 2 = 0$ (simple zero) $D = \infty$ (simple pole)	$\Lambda_3 = 2$ res D

Some properties of the roots are studied in Appendix B, and the results are summarized in the following theorem. Let[12]

$$K_\gamma = \text{No. of roots of Eq. (58).} \tag{59}$$

Let $\beta(\omega)$ be characterized by the five parameters k, n, $N_\pm$, N, defined in the following:

$$\begin{aligned} &\beta(\omega) \underset{\omega\to\infty}{\to} k\omega^{2n}, \qquad (n = 0, \pm1, \pm2, \cdots) \\ &N_\pm = \text{No. of poles of } \beta(\omega) \text{ on real axis} \\ &\qquad \omega > 1, \text{ with } \pm \text{ residues,} \\ &2N = \text{No. of poles of } \beta(\omega) \text{ not on real axis} \\ &\qquad |\omega| \geq 1. \end{aligned} \tag{60}$$

The total number of poles and zeros of $\beta(\omega)$ anywhere on the physical sheet are then respectively given by

$$\begin{aligned} P_\beta &= 2(N + N_+ + N_-), \\ Z_\beta &= 2(n + N + N_+ + N_-). \end{aligned} \tag{61}$$

The theorem states

1. There is no root on the imaginary axis except for $\gamma = 0$,
2. There is no root on the real axis $|\omega| > 1$ for any γ,
3. $K_\gamma = K_{-\gamma}$,
4. K_γ is independent of γ for $|\gamma| > \frac{1}{2}$,
5. $K_0 - K_1 = 2(\Delta\phi/\pi)$,
6. $K_1 = 1 - \nu + \max(Z_\beta, P_\beta)$,

(62)

where $\Delta\phi/\pi$ is given by (55), and

$$\nu = \begin{cases} 1 & \text{if } \beta(\omega) \text{ has a pole at } \omega = 0, \\ 0 & \text{otherwise.} \end{cases} \tag{63}$$

D. Necessary Bootstrap Conditions

Using (53) and (12), we write the bootstrap criterion in the form

$$\begin{aligned} \Delta\theta/\pi &= b_1, \\ \Delta\phi/\pi &= b_1 - b_3. \end{aligned} \tag{64}$$

Table V shows that b_1 can occur only at the roots of $B = 0, -1$, and b_3 can occur only at the roots of $B = 0, 2$. Accordingly we write

$$\begin{aligned} b_1 &= b_{10} + b_{11}, \\ b_3 &= b_{30} + b_{32}, \end{aligned} \tag{65}$$

where b_{10}, b_{11} are respectively the number of bound states in channel 1 at the roots of $B = 0, -1$; b_{30}, b_{32} are respectively the number of bound states in channel 3 at the roots of $B = 0, 2$. Apart from the cutoff poles, the poles of D can occur only at the real roots of $B = 0, 2$ between $\omega = \pm1$. Apart from the "extra zeros," the zeros of D can occur only at the roots of $B = 0, \pm1, 2$, and must occur at the complex roots of $B = \pm1$. Since D is an even function, and B an odd function, $D = 0$ at $B = \gamma$ implies $D = 0$ at $B = -\gamma$. Each zero or pole in D contributes to $\Delta\theta/\pi$ according to (54). Thus we write

$$\Delta\theta/\pi = \tfrac{1}{2}c + (\Delta\theta_0 + \Delta\theta_1 + \Delta\theta_2)/\pi - \tfrac{1}{2}X, \tag{66}$$

where c is defined in (4), and $\Delta\theta_\gamma/\pi$ is the contribution to $\Delta\theta/\pi$ arising from a pole or zero of D at $B = \gamma$, and $\frac{1}{2}X$ is the contribution from extra zeros of D ($X = 0, 1, 2, \cdots$). We now find the relation between $\Delta\theta$ and b_α by using Table V. In doing so we assume that all coupling constants have the correct sign, an assumption requiring verification only if we succeed in finding a bootstrap solution.

Consider first the roots of $B = -1$. Let there be K_{1R} real roots and $2K_{1C}$ complex roots:

$$K_{1R} + 2K_{1C} = K_1. \tag{67}$$

Every complex root must be cancelled, and some or all of the real roots may be cancelled by zeros of D. Suppose K'_{1R} real roots are cancelled. Then

$$\Delta\theta_1/\pi = -K_{1C} - \tfrac{1}{2}K'_{1R}, \quad b_{11} = K_{1R} - K'_{1R}. \tag{68}$$

Using (67) we obtain

$$\Delta\theta_1/\pi = \tfrac{1}{2}(b_{11} - K_1). \tag{69}$$

Next consider the roots of $B = 2$. Let there be K_{2R} real roots and $2K_{2C}$ complex roots:

$$K_{2R} + 2K_{2C} = K_1. \tag{70}$$

Poles of D may be placed at the real roots but not at the complex roots. Suppose K'_{2R} of the real roots are cancelled by poles of D. Then $\Delta\theta_2/\pi = \frac{1}{2}K'_{2R}$ and $b_{32} = K'_{2R}$. Hence

$$\Delta\theta_2/\pi = \tfrac{1}{2}b_{32}. \tag{71}$$

Combining (64), (65), (66), (69), and (71), we find a necessary condition for a bootstrap solution:

$$\Delta\phi/\pi = c - K_1 + 2(\Delta\theta_0/\pi) - (b_{10} + b_{30}) - X. \tag{72}$$

We now consider the roots of $B = 0$. It is obvious from (41) that there is always a root at $\omega = 0$, unless $\beta(\omega)$ has a pole there. In the former case, the number of roots at $\omega = 0$ must be odd, as we can deduce from the fact that K_0 is odd and that the roots at $\omega \neq 0$ must always occur in pairs at $\pm\omega$, $\pm\omega^*$. Accordingly we designate the number of roots as follows:

$$\begin{aligned} &\text{at } \omega = 0: && (1 + 2m_0)\delta_{\nu 0};\\ &\text{at } \omega \neq 0: && 2K_{0R} \text{ real roots},\\ & && 2K_{0I} \text{ pure imaginary roots},\\ & && 4K_{0C} \text{ complex roots}, \end{aligned} \tag{73}$$

where ν is given by (63). Thus

$$(1 + 2m_0)\delta_{\nu 0} + 2(K_{0R} + K_{0I} + 2K_{0C}) = K_0. \tag{74}$$

Let there be a factor $[(1 - iq)/(1 + iq)]^p$ in D to cancel some of the roots of $B = 0$ at $\omega = 0$, and let other poles of D be so chosen as to cancel respectively $2K'_{0R}$, $2K'_{0I}$, $4K'_{0C}$ of the other kinds of roots. Then

$$\Delta\theta_0/\pi = \tfrac{1}{2}p + \tfrac{1}{2}(K'_{0R} + K'_{0I}) + K'_{0C}. \tag{75}$$

We are free to choose these numbers as long as we observe that BD can have at most a simple pole at the real roots, and cannot have a pole at the nonreal roots. This leads to the restrictions

$$\begin{gathered} K'_{0C} \leq K_{0C}, \quad K'_{0I} \leq K_{0I}, \quad K'_{0R} \leq K_{0R} + z,\\ p \leq (1 + m_0)\delta_{\nu 0}, \end{gathered} \tag{76}$$

where the last inequality is obtained by noting that near $\omega = 0$, $BD \approx \omega^s$, where $s = (1 + 2m_0)\delta_{\nu 0} - 2p$. Only for $\nu = 0$ and $p = 1 + m_0$ can there be a bound state at $\omega = 0$. The number z in (76) is the number of pairs of simple poles of BD located at some or all of the $2K_{0R}$ real roots (which must occur in pairs symmetrically placed about the origin). If all the $2K_{0R}$ roots are simple poles, then $0 \leq z \leq K_{0R}$. If some of the roots are of multiple order, then $0 \leq z \leq$ (No. of distinct roots). From Table V we see that

$$b_{10} + b_{30} = 2z + \delta(p, 1 + m_0), \tag{77}$$

where $\delta(x, y)$ is the Kronecker delta.

Combining (74), (75), and all but the last inequalities of (76), we have

$$2(\Delta\theta_0/\pi) \leq \tfrac{1}{2}K_0 + p - (\tfrac{1}{2} + m_0)\delta_{\nu 0} + z. \tag{78}$$

Substituting this into (72), using (77), and using (62) to express K_0 in terms of K_1, we find

$$\begin{aligned} 0 \leq c - \tfrac{1}{2}K_1 - X - z + p\\ - (\tfrac{1}{2} + m_0)\delta_{\nu 0} - \delta(p, 1 + m_0). \end{aligned} \tag{79}$$

Using now the last inequality in (76), and dropping some obvious terms, we reduce this to $0 \leq c - \frac{1}{2}K_1 + \frac{1}{2}\delta_{\nu 0}$. Upon using (62) to express K_1 explicitly we find

$$0 \leq c - (N + N_+ + N_-) - \max(n, 0). \tag{80}$$

To test for a bootstrap solution we first use (80). If it is satisfied we then proceed to the sharper condition (79). If (79) is still satisfied the solution will have to be examined in more detail. We now establish the following theorems.

Theorem 1. There exists no bootstrap solution satisfying an unsubtracted dispersion relation.

Proof. Suppose there is no cutoff ($c = 0$). Then it is clear from (80) that $N = N_+ = N_- = n = 0$, and by (63) and (62) we have $\nu = 0$, $K_1 = 1$. Thus (79) is reduced to

$$0 \leq p - (1 + m_0) - \delta(p, 1 + m_0) - X - z, \tag{81}$$

which is impossible because $p \leq 1 + m_0$ and $x \geq 0$, $z \geq 0$.

Suppose $c \geq 1$. If there is no subtraction, the high-energy condition (49) requires $n \geq c$. Then (80) requires $n = c \geq 1$, and $N = N_+ = N_- = 0$. From (62) we find $K_1 = 1 + 2n$, which again reduces (79) to the impossible condition (81). q.e.d.

Theorem 2. For a once subtracted dispersion relation, a bootstrap solution must have $c = 1$. If, further, the target baryon is required to be a bound state, then $\beta(\omega) = \beta_0$, where β_0 is a nonvanishing constant.

Proof. According to the high-energy condition (49), $c \geq 1$. We first show that $c = 1$.

If $c \geq 2$ then (49) requires $n \geq c$. On the other hand (80) requires $n \leq c$. Hence $n = c$. By (80),

(63), and (62), we must have

$$N = N_+ = N_- = 0,$$
$$\nu = 0, \tag{82}$$
$$K_1 = 1 + 2n.$$

These requirements reduce (79) to (81), which is impossible.

For $c = 1$, (80) requires $N + N_+ + N_- + \max(n, 0) = 0, 1$. Hence there are only the following possibilities:

	N	N_+	N_-	n	ν	K_1
1.	0	0	0	0	0	1
2.	1	0	0	$-1, 0$	1	2
3.	$N + N_+ + N_- = 1$			$-1, 0$	0	3.

(83)

Substituting $\nu = 0$, $K_1 = 3$ into (79) reduces that inequality to (81), which is impossible. Hence case 3 is eliminated. Cases 1 and 2 can satisfy (79), but for case 2 the target baryon cannot be a bound state, for $\nu = 1$. Thus only case 1 remains, which corresponds to $\beta(\omega) =$ constant. We must rule out $\beta(\omega) = 0$ because it violates the threshold condition of the model. q.e.d.

E. Bootstrap Solutions

We consider bootstrap solutions satisfying a once-subtracted dispersion relation, with the target baryon as a bound state. By Theorem 2 we must have $c = 1$, and

$$\beta(\omega) = \beta_0 \neq 0. \tag{84}$$

According to (46) the high-energy behavior is $h_\alpha \approx \omega/\log \omega$, verifying the fact that one subtraction is needed. We may thus regard β_0 as an effective subtraction constant. For given β_0 and cutoff momentum κ the solution is unique, the proof of which is given in Appendix C. We merely quote the unique solution here.

The S-matrix elements are given by (40) with

$$B(\omega) = \tfrac{1}{2} + i[\pi^{-1} \log(\omega + q) - (\omega/q)\beta_0] \tag{85}$$

and $D(\omega)$ is given by separate expressions for $\beta_0 > 0$ and $\beta_0 < 0$.

For $\beta_0 < 0$ we have $K_1 = K_0 = 1$. The roots of $B(\omega) = \gamma$ are all real and are denoted as follows:

$B(\omega)$	ω	$s = (1 - \omega^2)^{\frac{1}{2}}$
0	0	1
± 1	$\mp\omega_1$	s_1
± 2	$\pm\omega_2$	s_2 ,

(86)

where

$$0 < \omega_1 < \omega_2 < 1,$$
$$0 < s_2 < s_1 < 1. \tag{87}$$

We then have

$$D(\omega) = \frac{(1 - iq)(1 - iq/\kappa)(1 - iq/s_2)(1 + iq/s_1)}{(1 + iq)(1 + iq/\kappa)(1 + iq/s_2)(1 - iq/s_1)}. \tag{88}$$

Apart from the cutoff poles, S_1 has only one pole at $\omega = 0$, and S_3 has two poles at $\omega = 0, \omega_2$. In accordance with Table V, we can assign the bound states such that

$$b_1 = 1 \quad \text{(bound state at } \omega = 0),$$
$$b_3 = 1 \quad \text{(bound state at } \omega = \omega_2), \tag{89}$$

with the coupling constants $\Lambda_1 = \frac{3}{2}$ res $(BD)_{\omega=0}$, $\Lambda_3 = 2$ res $(D)_{\omega=\omega_2}$, or

$$\lambda_1 = 3\left(\frac{1}{\pi} - \beta_0\right)\frac{(1 - s_1)(1 + s_2)}{(1 + s_1)(1 - s_2)}\left(1 + \frac{1}{\kappa}\right)^2 > 0,$$
$$\lambda_3 = \frac{2s_2(1 + s_2)(s_1 - s_2)(\kappa + s_2)}{\omega_2(1 - s_2)(s_1 + s_2)(\kappa - s_2)}\left(1 - \frac{1}{\kappa^2}\right) > 0. \tag{90}$$

The position ω_2 of the bound state in channel 3 and the coupling constants λ_1, λ_3, are shown in Fig. 1 as functions of β_0. Using (53) and (54), we find that $\Delta\phi/\pi = 0$, and

$$\Delta\delta_1 = \Delta\delta_3 = -\pi, \tag{91}$$

which fulfills Levinson's theorem. The only zeros of BD on the physical sheet are at the roots of $B^2 = 1$. Therefore S_1 has no zero and S_3 has a zero at $\omega = \omega_1$. This means that on the second Riemann sheet S_1 has no pole, and S_3 has only one pole at $\omega = \omega_1$, which represents a virtual state.

For $\beta_0 > 0$ we have $K_1 = 1$, $K_0 = 3$. The roots of $B = \pm 1$ are real. Of the three roots of $B = 0$, one is at $\omega = 0$, and the other two are symmetrically placed about the origin, either on the real axis between $\omega = \pm 1$ (if $\beta_0 < 1/\pi$), or on the imaginary axis (if $\beta_0 > 1/\pi$). There is also the possibility that all three roots are at $\omega = 0$, but this limiting case need not be considered separately. We denote the roots as follows:

$B(\omega)$	ω	$s = (1 - \omega^2)^{\frac{1}{2}}$
0	0	1
	$\pm\omega_0$	s_0
± 1	$\pm\omega_1$	s_1
± 2	$\mp\omega_2$	s_2 ,

(92)

where

$$\begin{aligned} &\text{for } 0 < \beta_0 < 1/\pi\text{:} \quad 0 < \omega_0 < \omega_1 < \omega_2 < 1, \\ &\qquad\qquad 0 < s_2 < s_1 < \omega_0 < 1, \\ &\text{for } \beta_0 > 1/\pi\text{:} \quad 0 < \omega_1 < \omega_2 < 1, \quad \mathrm{Re}\,\omega_0 = 0, \\ &\qquad\qquad 0 < s_2 < s_1 < 1 < s_0. \end{aligned} \tag{93}$$

We then have

$$D(\omega) = \frac{(1 - iq)(1 - iq/\kappa)(1 - iq/s_0)(1 + iq/s_1)}{(1 + iq)(1 + iq/\kappa)(1 + iq/s_0)(1 - iq/s_1)}. \tag{94}$$

Apart from the cutoff poles, S_1 and S_3 have a pole at $\omega = 0$ and nowhere else. In accordance with Table V, we can make the assignment

$$\begin{aligned} b_1 &= 1 \qquad \text{(bound state at } \omega = 0\text{)}, \\ b_3 &= 0, \end{aligned} \tag{95}$$

with a coupling constant $A_1 = \frac{3}{2}\,(\mathrm{res}\ BD)_{\omega=0}$, or

$$\lambda_1 = 3\left(\frac{1}{\pi} - \beta_0\right)\frac{(1 + s_0)(1 - s_1)}{(1 - s_0)(1 + s_1)}\left(1 + \frac{1}{\kappa}\right)^2 > 0. \tag{96}$$

This is always positive because $(\pi^{-1} - \beta_0)$ has the same sign as $1 - s_0$, as indicated in (93). In Fig. 1, λ_1 is shown as a function of β_0. Using (53) and (54) we find that $\Delta\phi/\pi = 1$, and

$$\Delta\delta_1 = -\pi, \qquad \Delta\delta_3 = 0, \tag{97}$$

which fulfills Levinson's theorem. On the physical sheet S_1 has a zero at $-\omega_2$, and S_3 has a zero at ω_1. These represent poles on the second Riemann sheet, and are therefore virtual states.

There are bootstrap solutions not having the target baryon as a bound state. Although these solutions are to be ruled out because they do not lead to physically reasonable models, they may still be of some mathematical interest. We study them in Appendix C.

5. NEUTRAL PSEUDOSCALAR THEORY

The neutral pseudoscalar theory describes the scattering of a neutral pseudoscalar P-wave meson from a fixed baryon of spin $\frac{1}{2}$, with conservation of total angular momentum. There are two channels labeled by $\alpha = 1, 3$, corresponding respectively to total angular momentum $\frac{1}{2}$ and $\frac{3}{2}$. The crossing matrix is the same as (39). The theory is mathematically equivalent to the symmetric scalar theory, except for the following differences:

1. $h_\alpha = (S_\alpha - 1)/q^3 v$,
2. $S_\alpha \underset{\omega\to 1}{\longrightarrow} 1 + O(q^3)$, (98)
3. Pole terms of S_α have opposite signs to (57).

Under Theorem 1 in the last section we proved that there are no bootstrap solutions of the symmetric scalar theory for $c = 0$. That proof applies without change to the present theory, because it is independent of the conditions (98). Hence we need to consider only $c \geq 1$. Owing to the first condition of (98) the high-energy conditions are the same as for the symmetric scalar theory if statements made about c in that theory are taken to be statements about $c - 1$ of this theory. For example, a solution with no subtraction imposes no restriction on n if $c - 1 = 0$, and requires $n \geq c - 1$ if $c - 1 \geq 1$. For K subtractions we have $c \geq \frac{1}{2}(3 + K)$, with the following requirements:

$$\begin{aligned} &c = \tfrac{1}{2}(3 + K)\text{:} \\ &\quad \text{No restriction on } \beta(\omega),\ D(\omega), \\ &c > \tfrac{1}{2}(3 + K)\text{:} \\ &\quad n > \tfrac{1}{2}(3 + K), \text{ plus conditions on } D(\omega). \end{aligned} \tag{99}$$

The threshold condition 2 of (98) requires

$$D(\omega) \underset{\omega\to 1}{\longrightarrow} 1 + O(q^3), \qquad \beta^{-1}(\omega) \underset{\omega\to 1}{\longrightarrow} O(q^2), \tag{100}$$

which represent extra restrictions on the solution. In particular the locations of poles and zeros of $D(\omega)$ must satisfy the first equation of (100), and $\beta(\omega)$ must have at least a second-order pole at $\omega = 1$, i.e., $N \geq 1$.

The theorem that there exists no bootstrap solution without subtraction also applies to the present case. The proof given earlier requires only the following amendments: (a) The case $c = 0$ is now replaced by the case $c = 1$; (b) The case $n \geq c \geq 1$ is now replaced by the case $n \geq c - 1 \geq 1$; (c) The condition $N \geq 0$ is now replaced by $N \geq 1$. The increase of N by 1 compensates for the decrease of c by 1, and preserves the previous proof.

The bootstrap solutions satisfying a once-subtracted dispersion relation are considered in Appendix D. In this case $c \geq 2$. The requirement that the target baryon be a bound state, and that no bound state have a smaller energy than the target baryon uniquely determines a two-parameter family of solutions with $c = 2$, the parameters being the cutoff momentum κ, and an effective subtraction constant.

The unique family of solutions corresponds to

$$\beta(\omega) = q^{-2}(\beta_0 + \beta_1\omega^2), \tag{101}$$

where β_0 and β_1 are constants restricted by the condition

$$\beta_0 + \beta_1 > 0. \tag{102}$$

They will be further related by the threshold condition, so that only one is an independent parameter. The S-matrix elements are given by (40), with

$$B(\omega) = \tfrac{1}{2} + i[\pi^{-1} \log (\omega + q) - (\omega/q^3)(\beta_0 + \beta_1\omega^2)]. \quad (103)$$

From (98) we verify that

$$h_\alpha \underset{\omega\to\infty}{\longrightarrow} \omega/\log \omega,$$

so that one subtraction is needed.

From Table IV and (62) we find $\sigma = 1$, $\Delta\phi/\pi = 1$, and

$$K_1 = 3, \qquad K_0 = 5. \quad (104)$$

One of the roots of $B = -1$ is always real, while the other two may be real or complex. One of the roots of $B = 0$ is always at $\omega = 0$, while the other four may be all real, all complex, or two of them may be real and the other two pure imaginary. The precise locations of the roots depend on the values of β_0 and β_1. Regardless of how the roots are located, however, the function $D(\omega)$ is completely determined by (43) and the instruction that

$$\begin{aligned} &D(\omega) \text{ have double poles at } \pm i(\kappa^2 - 1)^{\frac{1}{2}}, \\ &\text{simple poles at all the roots of } B = 0, \\ &\text{simple zeros at all the roots of } B = -1. \end{aligned} \quad (105)$$

and that there be no other poles or zeros.

It is clear that on the physical sheet, S_1 and S_3 have their only pole at $\omega = 0$. The values of β_0 and β_1 must be so restricted that (a) $\Lambda_1 = -\frac{3}{2}$ res $(BD)_{\omega=0} > 0$, so that the pole at $\omega = 0$ is a bound state in channel 1; and (b) the threshold condition (100) is satisfied. When this is done, we would have $b_1 = \Delta\theta/\pi = 1$, and $b_3 = b_1 - \Delta\phi/\pi = 0$, thus satisfying Levinson's theorem. The two independent parameters of the solution would then be the cutoff momentum κ and β_0 or β_1. It remains for us to show that the conditions (a) and (b) can be satisfied.

We demonstrate the existence of the family of solutions described above by constructing the subfamily corresponding to

$$\beta_0 > 0, \qquad \beta_1 > 0. \quad (105)$$

By examining the graphical solutions of $B = \gamma$, it can be seen that for $\gamma = 0$ there are four complex roots besides $\omega = 0$, and for $|\gamma| > 1$ there are two complex roots besides a real one. The general properties of the complex roots can be found by a method similar to that discussed by Wanders for the case $\beta_0 > 0$, $\beta_1 = 0$ in the Appendix of Ref. 14. The results are qualitatively the same as for the case discussed by Wanders. We summarize them as follows:

$B(\omega)$	ω		
0	0,	$\pm x$,	$\pm x^*$
-1	$-\omega_1$,	y,	y^*
$+1$	ω_1,	$-y$,	$-y^*$
2	ω_2,	z,	z^*,

(106)

where ω_1 and ω_2 are real, and x, y, z are complex, with

$$\begin{aligned} &0 < \omega_1 < \omega_2 < 1 \\ &\operatorname{Re} x > 1, \qquad \operatorname{Im} x \neq 0, \\ &\operatorname{Re} y > 1, \qquad \operatorname{Im} y \neq 0, \\ &\operatorname{Re} z < 0, \qquad \operatorname{Im} z \neq 0. \end{aligned} \quad (107)$$

Both x and y approach 1 as $\beta_0(1 - \beta_1) \to 0$, and their magnitudes approach ∞ as $\beta_0 \to \infty$. The function $D(\omega)$ is given by

$$D(\omega) = \frac{(1 - iq/\kappa)^2(1 - iq)(1 + iq/s_1)(1 + q/q_0)(1 - q/q_0^*)(1 - q/q_1)(1 + q/q_1^*)}{(1 + iq/\kappa)^2(1 + iq)(1 - iq/s_1)(1 + q/q_0^*)(1 - q/q_0)(1 - q/q_1^*)(1 + q/q_1)}, \quad (108)$$

where

$$\begin{aligned} s_1 &= (1 - \omega_1^2)^{\frac{1}{2}}, \\ q_0 &= (x^2 - 1)^{\frac{1}{2}}, \qquad \operatorname{Im} q_0 > 0, \\ q_1 &= (y^2 - 1)^{\frac{1}{2}}, \qquad \operatorname{Im} q_1 > 0. \end{aligned} \quad (109)$$

The threshold condition (100) requires that

$$\frac{1}{s_1} + 2 \operatorname{Im}\left(\frac{1}{q_0} - \frac{1}{q_1}\right) = 1 + \frac{2}{\kappa}, \quad (110)$$

which, for given κ, is a relation between β_0 and β_1. We only give the result for the case $(\beta_0 + \beta_1) \ll 1$, for which s_1 and q_γ are easily calculable:

$$\begin{aligned} s_1 &\approx [2(\beta_0 + \beta_1)]^{\frac{1}{2}}, \\ q_\gamma &\approx [(\beta_0 + \beta_1)/(\gamma + \tfrac{1}{2})]^{\frac{1}{2}} e^{i\pi/6}. \end{aligned} \quad (111)$$

Then (110) reduces to

$$\beta_1 + \beta_0 \approx \tfrac{3}{2}(1 + \kappa^{-1})^{-3}, \quad (112)$$

which is consistent only if $\kappa \sim 1$.

The only pole of S_1 and S_3 on the physical sheet is $\omega = 0$, with residues consistent with the assign-

ment of a bound state in channel 1, and no bound state in channel 3. The coupling constant is $\Lambda_1 = -\frac{3}{2}$ res $(BD)_{\omega=0}$, or

$$\lambda_1 = 3\left(\frac{1}{\pi} + \beta_0\right)\frac{1 - s_1}{1 + s_1}\left(1 + \frac{1}{\kappa}\right)^4 > 0. \qquad (113)$$

On the physical sheet S_1 has three zeros at the roots of $B = 2$, which all appear to the left of $\omega = 1$. On the second sheet, therefore, S_1 has three poles representing virtual states. For S_3 there are three zeros on the physical sheet at the roots of $B = -1$, i.e., $-\omega_1, y, y^*$. On the second sheet, therefore, there are three poles of S_3. The one at $-\omega_1$ represents a virtual state, and the conjugate pair y, y^*, with Re $y > 1$, represents a resonance of spin $\frac{3}{2}$.

For $(\beta_0 + \beta_1) \ll 1$, (which requires $\kappa \sim 1$), the resonance pole in channel 3 is near threshold, and depends only on κ:

$$y = (1 + q_1^2)^{\frac{1}{2}} \approx 1 + \tfrac{1}{2}(1 + \kappa^{-1})^{-2}e^{i\pi/3}. \qquad (114)$$

The position and width of the resonance are respectively given by

$$\begin{aligned} \text{Position} &= \text{Re } y \approx 1 + \tfrac{1}{4}(1 + \kappa^{-1})^{-2}, \\ \text{Width} &= \text{Im } y \approx \tfrac{1}{4}\sqrt{3}\,(1 + \kappa^{-1})^{-2}. \end{aligned} \qquad (115)$$

6. SYMMETRIC PSEUDOSCALAR THEORY

The symmetric pseudoscalar theory[5] describes the scattering of a pseudoscalar P-wave meson of isospin 1 by a fixed baryon of spin $\frac{1}{2}$ and isospin $\frac{1}{2}$. Among the models we consider this comes closest to describing a physical process, i.e., low-energy pion–nucleon scattering. There are four channels that are best labeled by two indices (α, β), in which α refers to the total angular momentum and β refers to the total isospin of the channel. We take the value $\alpha = (1, 3), \beta = (1, 3)$ to refer to respectively angular momentum and isospin $(\frac{1}{2}, \frac{3}{2})$. The S-matrix elements will accordingly be denoted by $\mathcal{S}_{\alpha\beta}(\omega)$. If we regard $\mathcal{S}$ as a vector in a product space spanned by vectors of the form $\psi_\alpha\phi_\beta$, then the crossing matrix is a direct product

$$\mathcal{C} = A \times A, \qquad (116)$$

where the first A acts on ψ_α, the second on ϕ_β, and A is given by (39), i.e.,

$$\mathcal{C}_{\alpha'\beta',\alpha\beta} = A_{\alpha'\alpha}A_{\beta'\beta}. \qquad (117)$$

In the Lagrangian formulation of this theory one has the restriction $\mathcal{S}_{13} = \mathcal{S}_{31}$, but we do not insist on it here.

The most general solution of the model is so far unknown; but the form of the crossing matrix shows that there exists a special class of solutions for which $\mathcal{S}_{\alpha\beta} = S_\alpha S_\beta$ where S_α is formally a solution of the neutral pseudoscalar theory:

$$\begin{aligned} \mathcal{S}_{11} &= D\frac{B_1}{B_1 - 1}\frac{B_2}{B_2 - 1}\frac{B_1 - 2}{B_1 + 1}\frac{B_2 - 2}{B_2 + 1}, \\ \mathcal{S}_{13} &= D\frac{B_1}{B_1 - 1}\frac{B_2}{B_2 - 1}\frac{B_1 - 2}{B_1 + 1}, \\ \mathcal{S}_{31} &= D\frac{B_1}{B_1 - 1}\frac{B_2}{B_2 - 1}\frac{B_2 - 2}{B_2 + 1}, \\ \mathcal{S}_{33} &= D\frac{B_1}{B_1 - 1}\frac{B_2}{B_2 - 1}, \end{aligned} \qquad (118)$$

where B_1 and B_2 are two functions of the form (41), and D is of the form (43). There are therefore three arbitrary functions. The threshold condition and high-energy conditions of the neutral pseudoscalar theory must be satisfied separately by D and by B_1 and B_2, because the behavior of B_1 and B_2 at threshold and at high energies contribute additively to the behavior of $\mathcal{S}_{\alpha\beta}$. Similarly, a bootstrap solution within this class exists if and only if B_1 and B_2 separately lead to bootstrap solutions of the neutral pseudoscalar theory.

It is not necessary to consider (118) any further, except to point out that (118) does not contain the physically interesting solution in which there is a bound state at $\omega = 0$ in channel 11, a resonance in channel 33, and neither bound state nor resonance in other channels.

APPENDIX A: PROOF OF EQ. (30)

The roots of the equation $H_\alpha^{-1}(\omega) = 0$ can occur only on the real axis between $\omega = \pm 1$. To find their number we apply (29) by taking $f(\omega) = H_\alpha^{-1}(\omega)$, and choosing the contour to be a counterclockwise circle centered at $\omega = 0$, with radius $1 - \epsilon$ $(\epsilon \to 0^+)$.

From (28), (18), and (19), we obtain

$$\frac{1}{H_1(\omega)} = -\frac{2\omega}{\pi}\int_1^\infty \frac{d\omega'}{\omega'}\frac{q'v(\omega')}{\omega'^2 - \omega^2} - \sum_n \frac{g_n}{\epsilon_n - \omega} - a\omega + d, \qquad (A1)$$

and $H_{-1}(\omega) = H_1(-\omega)$. The only poles of $H_\alpha^{-1}(\omega)$ are the CDD poles ϵ_n, which lie on the real axis. We assume $|\epsilon_n| \neq 1$. This involves no loss of generality because, $H_\alpha(\omega)$ being a continuous function of ϵ_n, we may displace ϵ_n slightly. By convention we displace ϵ_n towards $\omega = 0$ by 2ϵ, if originally $|\epsilon_n| = 1$, thus preserving the definition (25). Let

$$H_\alpha^{-1}(\omega) = |H_\alpha(\omega)|^{-1}\, e^{i\eta_\alpha(\omega)}, \qquad (A2)$$

and let $\Delta\eta_\alpha$ denote the net change in η_α when ω goes around the contour once. Noting that the value of the integral in (29) is $\Delta\eta_\alpha/2\pi$, we have

$$\text{No. of poles of } H_\alpha(\omega) = M + \Delta\eta_\alpha/2\pi. \qquad (A3)$$

In the limit $\epsilon \to 0^+$, M is the number of CDD poles with $|\epsilon_n| \leq 1$. To compute $\Delta\eta_\alpha$ we consider

$$\cot \eta_\alpha(\omega) = \text{Re } H_\alpha^{-1}(\omega)/\text{Im } H_\alpha^{-1}(\omega), \quad \text{(A4)}$$

which is single-valued along the contour. Therefore $\Delta\eta_\alpha$ may be deduced from the poles of cot η_α along the contour: Each pole with $\pm$ residue additively contributes $\pm\pi$ to $\Delta\eta_\alpha$.

Since Re $H_\alpha^{-1}(\omega)$ is finite on the contour, and since by construction Im $H_\alpha^{-1}(\omega) \neq 0$ unless Im $\omega = 0$, the only poles of cot η_α are located on the real axis at $\omega = 1 - \epsilon$ and $\omega = -1 + \epsilon$. It is then straightforward to verify (30) by examining the signs of the residues.

APPENDIX B: ROOTS OF $B(\omega) = \gamma$

From the form

$$B(\omega) = \pi^{-1} \sin^{-1} \omega - (i\omega/q)\beta(\omega), \quad \text{(B1)}$$

we see that B is pure imaginary when ω is pure imaginary, because $\beta(\omega^*) = \beta^*(\omega)$, and $\beta(-\omega) = \beta(\omega)$. Therefore statement 1 of (62) is obvious. From the form

$$B(\omega) = \tfrac{1}{2} + i\left[\frac{1}{\pi} \log (\omega + q) - \frac{\omega}{q}\beta(\omega)\right], \quad \text{(B2)}$$

we see that Im $B \neq 0$ if ω is real and $|\omega| > 1$. Hence statement 2 of (62) is obvious. Since $B(\omega) = -B(-\omega)$, statement 3 of (62) follows.

All real roots of $B = \gamma$ must lie between $\omega = \pm 1$. They may be found by a graphical solution of

$$(1/\pi) \sin^{-1} \omega = \omega\beta(\omega)/(1 - \omega^2)^{\frac{1}{2}} + \gamma. \quad \text{(B3)}$$

Recalling that the threshold condition requires $\beta(1) \neq 0$ we see that for any γ there is at least one root on the real axis between $\omega = \pm 1$, if $\beta(0) \neq \infty$.

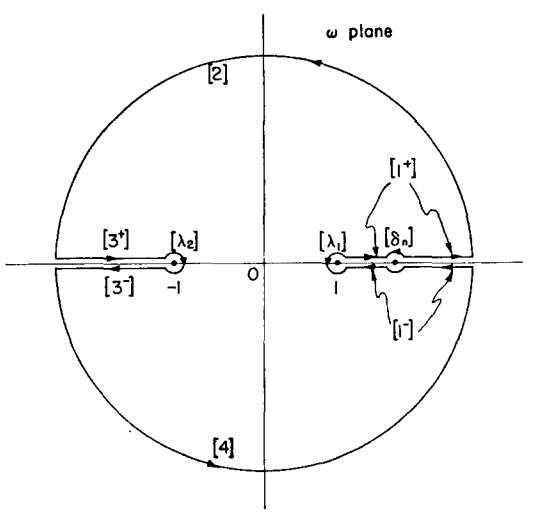


FIG. 2. The contour in the complex ω referred to in the calculation of (B5).

TABLE VI. Contributions to $\Delta\eta$ of Eq. (B5).

From	Conditions			Contribution to
[2] + [4]	$n > 0$	$\lvert\gamma\rvert > \frac{1}{2}$		$2n + 1$
		$\gamma = 0$	$\beta(1)$ k	
			+ +	$2n + 1$
			+ −	$2n + 3$
			− +	$2n - 1$
			− −	$2n + 1$
	$n \leq 0$	$\lvert\gamma\rvert > \frac{1}{2}$		0
		$\lvert\gamma\rvert < \frac{1}{2}$		1
$[\lambda_1] + [\lambda_2]$	$(\frac{1}{2} - \gamma)\beta(1)$	$(\frac{1}{2} + \gamma)\beta(1)$		
	+	+		2
	+	−		1
	−	+		1
	−	−		0
each N_+	$\lvert\gamma\rvert > \frac{1}{2}$			4
	$\gamma = 0$			2
each N_-	$\lvert\gamma\rvert > \frac{1}{2}$			2
	$\gamma = 0$			0

To establish statements 3–6 of (62) we calculate the number of zeros of the function $B(\omega) - \gamma$ by the same method as used in Appendix A. Namely, we write

$$B(\omega) - \gamma = |B(\omega) - \gamma|\, e^{i\eta(\omega)}, \quad \text{(B4)}$$

and let $\Delta\eta$ be the change of $\eta(\omega)$ when ω goes once around the contour shown in Fig. 2, where the detours $[\delta_n]$ are to be made around the poles $\beta(\omega)$. Then

$$K_\gamma = (\Delta\eta/2\pi) + 2N - \nu, \quad \text{(B5)}$$

where N and ν are defined in (60) and (63). To avoid having too many different cases to cope with in the calculation, we assume that $\beta(\omega)$ has no pole at $\omega = \pm 1$. This involves no loss of generality, because without affecting K_γ, we can always displace a pole of $\beta(\omega)$ at $\omega = \pm 1$ by an infinitesimal amount toward $\omega = 0$ along the real axis. Accordingly we include all poles at $\omega = \pm 1$ in N, as the definition (60) indicates. To calculate $\Delta\eta$, we consider

$$\tan \eta(\omega) = \frac{\pi^{-1} \log |\omega + q| - \text{Re } [\omega\beta(\omega)/q]}{\frac{1}{2} - \gamma - \text{phase } (\omega + q) + \text{Im } [\omega\beta(\omega)/q]}. \quad \text{(B6)}$$

It is readily verified that $\tan\eta$ is single-valued and continuous at all junctions of different segments of the contour. The contributions to $\Delta\eta$ therefore come only from the poles of $\tan\eta$ within each segment of the contour: Each pole with $\pm$ residue additively contributes $\mp\pi$ to $\Delta\eta$.

The task of finding the contribution to $\Delta\eta$ from

TABLE VII. $\Delta\eta$ of Eq. (B5).

	$n \geq 1$		$n \leq 0$	
	$k\beta(1) > 0$	$k\beta(1) < 0$	$\beta(1) > 0$	$\beta(1) < 0$
$\gamma = 0$	$2n + 1 + 4N_+$	$2n + 3 + 4N_+$ $(k < 0)$ $2n - 1 + 4N_+$ $(k > 0)$	$3 + 4N_+$	$1 + 4N_+$
$\lvert\gamma\rvert > \frac{1}{2}$	$2n + 1 + 2(N_+ + N_-)$		$1 + 2(N_+ + N_-)$	

the various segments of the contour is straightforward, except perhaps for [2] and [4], for which a helpful rule may be noted: In calculating the denominator of (B6) it is sufficient to take Im $(\omega\beta/q) \approx$ Im β, and to retain only the leading power of β. The reason is that any lower-order terms can affect the result only if they fail to vanish at the endpoints of [2] and [4]; but Im $(\omega\beta/q) = 0$ at the endpoints, which lie on the real axis.

The contributions to $\Delta\eta$ from the different segments of the contour are tabulated in Table VI. The segments $[1^{\pm}]$, $[3^{\pm}]$ do not contribute. Adding these results we obtain $\Delta\eta$ as given in Table VII, from which statements 3–6 of (62) follow.

APPENDIX C: ALL BOOTSTRAP SOLUTIONS IN SYMMETRIC SCALAR THEORY WITH ONE SUBTRACTION

We discuss all bootstrap solutions of the symmetric scalar theory with one subtraction, but no restrictions on the number and positions of bound states. As explained in Sec. 2, a physically acceptable solution must have the target baryon as a bound state, and must have no bound state lighter than the target baryon. It is of some mathematical interest, however, to see the extent to which Levinson's theorem restricts the bound states without extra physical requirements.

Before limiting ourselves to one subtraction, we note the following theorem.

Theorem. For any $c \geq 1$ there always exist bootstrap solutions with no bound state (with an appropriate number of subtractions.)

Proof. The theorem is proved by constructing an example for each value of $c \geq 1$. Let us choose $\beta(\omega)$ such that

$$N_+ = N_- = 0, \qquad n \leq 0, \qquad \beta(1) < 0.$$

From (62) we find $K_1 = 1 - \nu + 2N$, where N still remains arbitrary. By Table IV and (54) we have $\Delta\phi/\pi = \sigma = 0$, which is consistent with $b_1 = b_3 = 0$. We now choose $D(\omega)$ to have zeros only at all the roots of $B = -1$, and to have only the poles of the cutoff function. Obviously this makes $b_1 = b_3 = 0$. The only remaining bootstrap condition is $\Delta\theta/\pi = b_1 = 0$. To satisfy it we note that $\Delta\theta/\pi = \frac{1}{2}(c - K_1) = \frac{1}{2}[c - (1 - \nu + 2N)]$, and put accordingly $c = 1 - \nu + 2N$. For $N = 0$, $\nu = 0$ necessarily, and we have $c = 1$. For $N = 1, 2, 3, \cdots$, $\nu = 0, 1$ independently, and we generate all higher values of c, thus completing the construction. q.e.d.

All possible bootstrap solutions satisfying a once-subtracted dispersion relation have $c = 1$, and are listed in the proof of Theorem 2 in Section 4D, i.e., cases 1 and 2 of (83). These cases exhaust all possible bootstrap solutions with $c = 1$, because (a) any solution with $c = 1$ needs at most one subtraction, as we can see from (46), and (b) a bootstrap solution must have at least one subtraction, which is Theorem 1 of Sec. 4D.

We now construct all the bootstrap solutions contained in cases 1 and 2 of (83). They can both be included by choosing

$$\beta(\omega) = \beta_0(1 - f\omega^{-2}), \qquad \beta_0(1 - f) \neq 0. \tag{C1}$$

Case 1 corresponds to $f = 0$, and case 2 corresponds to $f \neq 0$. All relevant parameters for the two cases are listed in the following table, in which the first four columns are obtained from (62) and (54), the next three columns are requirements imposed by (79), and the last column is a consequence of (77):

	ν	K_1	K_0	$\Delta\phi/\pi$	z	x	p	$b_{10}+b_{30}$
Case 1								
$(f=0)$	0	1	$1+2\sigma$	σ	0	0	$1+m_0$	1
							m_0	0
Case 2								
$(f\neq 0)$	1	2	$2+2\sigma$	σ	0	0	0	0

(C2)

where

$$\sigma = \begin{cases} 1, & \beta(1) > 0 \\ 0, & \beta(1) < 0. \end{cases} \tag{C3}$$

The bootstrap conditions (64) can now be written

$$\begin{aligned} b_1 &= \tfrac{1}{2} + (\Delta\theta_0 + \Delta\theta_1 + \Delta\theta_2)/\pi, \\ b_1 - b_3 &= \sigma, \end{aligned} \tag{C4}$$

where (66) has been used. Furthermore, we have from (72), (69), and (71) respectively,

$$\begin{aligned}\Delta\theta_0/\pi &= \tfrac{1}{2}[\sigma + K_1 - 1 + b_{10} + b_{30}],\\ \Delta\theta_1/\pi &= \tfrac{1}{2}(b_{11} - K_1),\\ \Delta\theta_2/\pi &= \tfrac{1}{2}b_{32}.\end{aligned} \tag{C5}$$

It should be noted that if $b_{10} + b_{30} = 0$, then $b_1 = b_{11}$ and $b_2 = b_{32}$. If $b_{10} + b_{30} = 1$, then there are two possibilities: $(b_{10}, b_{30}) = (1, 0), (0, 1)$. The procedure for constructing bootstrap solutions is as follows. For each sets of possible values of σ and p in (C2), we consider all possible distribution of bound states that can satisfy (C4) and (C5). To do this it is necessary to know where the roots of $B = \gamma$ are located. After this is done, the sufficient condition for a bootstrap solution is that all residues at the poles of S_α have the correct signs (as given in Table V), consistent with the assumed bound-state distribution.

For case 1 we first consider $\beta_0 < 0$, which requires $\sigma = 0$. Then $K_0 = K_1 = 1$. The roots of $B = \gamma$ are as follows

$$\begin{array}{cc} B(\omega) & \omega \\ 0 & 0 \\ -1 & -\omega_1 \\ 2 & \omega_2 \end{array} \qquad (0 < \omega_1 < \omega_2 < 1). \tag{C6}$$

Thus $m_0 = 0$, and $p = 0, 1$. Only the following choices of bound states can satisfy (C4) and (C5):

	σ	p	b_{10}	b_{30}	b_{11}	b_{32}	b_1	b_3	Bootstrap?
case 1a	0	0	0	0	0	0	0	0	yes
1b	0	1	1	0	0	1	1	1	yes
1c	0	1	0	1	1	0	1	1	no
1c′	0	0	0	0	1	1	1	1	no.

(C7)

The information given above uniquely determines $D(\omega)$, and hence the complete solution, in each case. The last column is arrived at by examining the residues at the bound states, which we explain below.

Case 1a is automatically a bootstrap solution, because there are no bound states, and hence no further conditions to satisfy.

Case 1b is the solution given in (88), which was demonstrated to be a bootstrap solution.

Case 1c requires

$$D(\omega) = \frac{(1 - iq)(1 - iq/\kappa)}{(1 + iq)(1 + iq/\kappa)}, \tag{C8}$$

and that the pole of S_α at $\omega = 0$ must be a bound state in channel 3. By Table V, $\Lambda_3 = -\frac{3}{2}\,\mathrm{res}\,(BD)_{\omega=0}$, which we easily show to be negative. Hence the pole represents a ghost state in channel 3, and this solution must be ruled out.

Case 1c′ is ruled out by similar arguments, which show that $\Lambda_1 < 0$.

We consider now case 1 with $\beta_0 > 0$, which requires $\sigma = 1$, $K_1 = 1$, $K_0 = 3$. The roots are located as follows

$$\begin{array}{cc} B(\omega) & \omega \\ 0 & 0, \pm\omega_0 \\ -1 & \omega_1 \\ 2 & -\omega_2 \end{array} \qquad (0 \le \omega_0 < \omega_1 < \omega_2 < 1). \tag{C9}$$

We assume that $\omega_0 \neq 0$, so that $m_0 = 0$. This involves no loss of generality because according to (C2) $z = 0$. Only two subcases can satisfy (C4) and (C5):

	σ	p	b_{10}	b_{30}	b_{11}	b_{32}	b_1	b_3	Bootstrap?
Case 1d	1	1	1	0	0	0	1	0	yes
Case 1e	1	1	1	0	1	1	2	1	no

(C10)

Case 1d is that given in (94), which was demonstrated to be a bootstrap solution.

Case 1e requires

$$D(\omega) = \frac{(1 - iq)(1 - iq/\kappa)(1 - iq/s_0)(1 - iq/s_2)}{(1 + iq)(1 + iq/\kappa)(1 + iq/s_0)(1 + iq/s_2)}, \tag{C11}$$

$$s_\gamma = (1 - \omega_\gamma^2)^{\frac{1}{2}}.$$

For the bound state in channel 1 at $\omega = \omega_1$, Table V requires $\Lambda_1 = -\frac{3}{2}\,\mathrm{res}\,[D/(B + 1)]_{\omega=\omega_1}$, which is easily shown to be negative. Hence it is a ghost state, and this solution must be ruled out.

We turn to case 2 of (C2), for which S_α cannot have a pole at $\omega = 0$. This means that the target baryon cannot be a bound state in this group of solutions. We have

$$\sigma = \begin{cases} 1 & \beta_0(1 - f) > 0 \\ 0 & \beta_0(1 - f) < 0. \end{cases} \tag{C12}$$

Under each alternative above, there are two further alternatives corresponding to $f > 1$ and $f < 1$. To satisfy (C4) and (C5) we must have

$$p = b_{10} = b_{30} = 0, \tag{C13}$$

and the following possible distribution of bound states:

	σ	$b_{11}=b_1$	$b_{32}=b_3$	Bootstrap?	
Case 2a	0	0	0	yes	
2b	0	1	1	yes	$(f<1)$
2c	0	2	2	no	(C14)
2d	1	1	0	yes	$(f>1)$
2e	1	2	1	no.	

To arrive at the last column we need to know about the roots of $B=\gamma$. They are summarized below.

For $\sigma=0, K_1=K_0=2$. The roots are as follows:

$B(\omega)$	ω	
	$(f>1)$	$(f<1)$
0	$\pm i\omega_0$	$\pm\omega_0$
-1	$-\omega_1, -\omega_1'$	$\omega_1, -\omega_1'$
2	ω_2, ω_2'	$-\omega_2, \omega_2'$

(C15)

where, for $f>1$:

$$0<\omega_2<\omega_1<\omega_1'<\omega_2'<1,$$
$$\text{res}\,(B+1)^{-1}>0 \text{ at } -\omega_1'$$
$$<0 \text{ at } -\omega_1; \tag{C16}$$

and, for $f<1$:

$$0<\omega_2<\omega_1<\omega_0<\omega_1'<\omega_2',$$
$$\text{res}\,(B+1)^{-1}>0 \text{ at } \omega_1 \text{ and } -\omega_1'.$$

For $\sigma=1, K_1=2, K_0=4$. The roots are as follows:

$B(\omega)$	ω	
	$(f>1)$	$(f<1)$
0	$\pm\omega_0, \pm i\omega_0'$	$\pm\omega_0, \pm\omega_0'$
		or all complex
-1	$-\omega_1, \omega_1'$	ω_1, ω_1'
2	$\omega_2, -\omega_2'$	$-\omega_2, -\omega_2'$

(C17)

where, for $f>1$:

$$0<\omega_2<\omega_1<\omega_0<\omega_1'<\omega_2'<1,$$
$$\text{res}\,(B+1)^{-1}<0 \text{ at } -\omega_1 \text{ and } \omega_1'; \tag{C18}$$

and, for $f<1$:

$$0<\omega_2<\omega_1<\omega_0<\omega_0'<\omega_1'<\omega_2'<1$$
$$\text{res}\,(B+1)^{-1}>0 \text{ at } \omega_1$$
$$<0 \text{ at } \omega_1'.$$

We now explain the last column of (C14).

TABLE VIII. All bootstrap solutions of symmetric scalar theory with $c=1$. β_0 and f are defined in Eq. (C1). Each entry in the Table gives the case number followed by (b_1, b_3), where b_α = No. of bound states in channel α. Cases 1(b) and 1(d) are the only cases having the nucleon as a bound state.

	$f\neq 0$	
$f=0$	$\beta_0(1-f)>0$	$\beta_0(1-f)<0$
1(a) (0, 0) 1(b) (1, 1) 1(d) (1, 0)	2(d) (1, 0)	2(a) (0, 0) 2(b) (1, 1)

Case 2a is automatically a bootstrap solution because there are no bound states.

Case 2b requires

$$D(\omega)=\frac{(1-iq/\kappa)(1-iq/s_0)(1-iq/s_2)(1+iq/s_1)}{(1+iq/\kappa)(1+iq/s_0)(1+iq/s_2)(1-iq/s_1)}, \tag{C19}$$

$$s_\gamma=(1-\omega_\gamma^2)^{\frac{1}{2}},$$

or we may independently replace S_2 by S_2', S_1 by S_1', thus obtaining four cases in all. For $f>1$, the bound state in channel 3 is always a ghost. For $f<1$ we obtain two bootstrap solutions corresponding to

$$\begin{aligned} &1. \quad b_1 \text{ at } \omega_1, \quad b_3 \text{ at } \omega_2', \\ &2. \quad b_1 \text{ at } -\omega_1', \quad b_3 \text{ at } -\omega_2, \end{aligned} \tag{C20}$$

where b_α means bound state in channel α. The second alternative leads to bound states lighter than the target baryon.

Case 2c leads to a ghost state in channel 3 for $f>1$, and a ghost state in channel 1 for $f<1$.

Case 2d leads to a ghost state in channel 1 for $f<1$. For $f>1$ there are two bootstrap solutions corresponding to

$$\begin{aligned} &1. \quad b_1 \text{ at } \omega_1', \\ &2. \quad b_1 \text{ at } -\omega_1. \end{aligned} \tag{C21}$$

The second alternative leads to a bound state lighter than the target baryon.

Case 2e leads to a ghost state in channel 3 for $f>1$, and a ghost state in channel 1 for $f<1$.

This completes the construction of all bootstrap solutions with $c=1$. The results are summarized in Table VIII.

APPENDIX D: BOOTSTRAP SOLUTIONS IN NEUTRAL PSEUDOSCALAR THEORY

We prove that among all the bootstrap solutions of the neutral pseudoscalar theory with one subtraction, the family given by (103) and (105) is the only one satisfying the physical requirements that

the target baryon be a bound state, and that there be no bound state lighter than the target baryon.

For one subtraction, the high-energy condition (99) requires $c \geq 2$. For $c = 2$ there is no restriction on n, and for $c \geq 3$, $n \geq c - 1$.

We first rule out $c \geq 3$. For this case (80) can be satisfied only if $c = n + 1$ or $c = n + 2$. The latter is immediately ruled out because $n \geq c - 1$. Remembering that $N \geq 1$, we conclude that

$$\begin{aligned} N &= 1 \\ N_+ &= N_- = \nu = 0 \\ K_1 &= 3 + 2n. \end{aligned} \tag{D1}$$

Applying (79), we obtain the further requirements

$$X = z = 0, \qquad p = m_0. \tag{D2}$$

Thus the target baryon cannot be a bound state, because $p \neq 1 + m_0$.

We turn to $c = 2$. Applying (80), we find that there are only five possible cases:

	N	N_+	N_-	n	ν	K_1	
1.	1	0	0	$0, -1$	0	3	
2.	2	0	0	0	1	4	
3.	1	$N_+ + N_- = 1$		$0, -1$	0	5	(D3)
4.	2	0	0	0	0	5	
5.	2	0	0	0	1	4.	

Applying (79), we rule out cases 3, 4, 5, and obtain the following requirements for cases 1 and 2:

	N	N_+	N_-	n	ν	K_1	X	z	1	p	
1.	1	0	0	$0, -1$	0	3	0	0	$1+m_0$		(D4)
2.	2	0	0	0	1	4	0	0	0.		

Only case 1 can have the baryon as a bound state. It corresponds to

$$\beta(\omega) = q^{-2}(\beta_0 + \beta_1\omega^2), \qquad \beta_0 + \beta_1 \neq 0, \tag{D5}$$

where the last condition is necessary to satisfy the threshold requirement (100).

With (D5), we have

$$K_1 = 3, \qquad K_0 = 3 + 2\sigma, \qquad \Delta\phi/\pi = \sigma, \tag{D6}$$
$$\sigma = \begin{cases} 1 & \beta(1) > 0 \\ 0 & \beta(1) < 0. \end{cases}$$

The following relations are obtained respectively from (77), (72), (69), and (71):

$$\begin{aligned} b_{10} + b_{30} &= 1, \\ \Delta\theta_0/\pi &= 1 + \tfrac{1}{2}\sigma, \\ \Delta\theta_1/\pi &= \tfrac{1}{2}(b_{11} - 3), \\ \Delta\theta_2/\pi &= \tfrac{1}{2}b_{32}. \end{aligned} \tag{D7}$$

Using these we reduce the bootstrap condition $b_1 = \Delta\theta/\pi$ to

$$b_{10} + \tfrac{1}{2}(b_{11} - b_{32}) = \tfrac{1}{2}(1 + \sigma). \tag{D8}$$

If this is satisfied, then the condition $b_1 - b_3 = \Delta\phi/\pi$ is automatic. The remaining task is to list all possible distributions of bound states satisfying (D8), and then examine the residues. There are 12 possible cases satisfying (D8), and they are tabulated in Table IX.

Any solution with $b_{10} = 0$ cannot have the target baryon as a bound state. Hence we rule out cases c_1, c_2, c_3, d_1, and d_2. Any solution with $b_{11} = 3$ or $b_{32} = 3$ must have at least one bound state lighter than the target baryon, for a graphical examination of the equation $B(\omega) = \gamma$ ($\gamma = -1, 2$) shows that if all three roots are real, then at least one of them lie in the interval $-1 < \omega < 0$. Hence we rule out cases a_3 and b_4.

A closer examination of the graph for $B(\omega) = \gamma$ further rules out cases a_2 and b_3. In case a_2, $b_{32} = 2$. The graph of $B = 2$ shows that if there are three real roots, then two of them must occur at $\omega < 0$, and the other at $\omega > 0$. Therefore at least one bound state in channel 3 must be lighter than the target baryon. In case a_1, a similar argument shows that at least one bound state in channel 1 must be lighter than the target baryon.

We are now left with only the cases a_1, b_1, and b_2. Explicit construction of $D(\omega)$ for each of these cases show that case a_1 has a ghost state in channel 3, and case b_2 has a ghost state in channel 1. In verifying this, we should remember that for the present theory the coupling constants have opposite signs to that given in Table V.

We have eliminated all cases except case b_1, which is the solution described in Section V.

TABLE IX. Possible cases satisfying (D8).

Case	σ	b_{10}	b_{30}	b_{11}	b_{32}	b_1	b_3
a_1	0	1	0	0	1	1	1
a_2	0	1	0	1	2	2	2
a_3	0	1	0	2	3	3	3
b_1	1	1	0	0	0	1	0
b_2	1	1	0	1	1	2	1
b_3	1	1	0	2	2	3	2
b_4	1	1	0	3	3	4	3
c_1	0	0	1	1	0	1	1
c_2	0	0	1	2	1	2	2
c_3	0	0	1	3	2	3	3
d_1	1	0	1	2	0	2	1
d_2	1	0	1	3	1	3	2

Compactification of Minkowski Space and $SU(3)$ Symmetry*

ROBERT HERMANN
Argonne National Laboratory, Argonne, Illinois and Northwestern University, Evanston, Illinois
(Received 12 October 1964)

Several ideas from the geometric side of Lie group theory are presented that may be relevant to the search for groups containing the internal and Poincaré group symmetries.

1. INTRODUCTION

THERE has recently been interest in investigating symmetry groups that contain both the Poincaré group and the internal symmetry groups. One would think that this choice of larger group is partially determined by geometric considerations. In this paper, we suggest one such geometric property, namely that the larger group act as a transformation group on a compactification of Minkowski space. Certainly, it seems self-evident that the various "asymptotic conditions" that one finds in physics should be related to the geometric behavior of points in Minkowski space near its boundary, in a suitable compactification. In view of the expected covariance with respect to the Poincaré group, it also seems reasonable to require that the compactification be "equivariant" in the sense that the transformations defined by the Poincaré group can be extended to act on the larger compactification. To restrict the number of possibilities, perhaps one should also require that there is a larger group containing the Poincaré group that acts transitively on the compactification. This leads to the following

Definition. Let G, H, G', H' be Lie groups, with H a subgroup of G and H', H' and G subgroups of G'. Suppose that $H' \cap G = H$. Then, the coset space G/H may be considered as a subspace of G'/H' by assigning the coset gH' to the coset gH, for $g \in G$. If in addition G'/H' is compact, we will say that it is a *homogeneous space compactification of* G/H.

The case of interest in physics is that where G is the Poincaré group and H is the homogenous part, i.e., the Lorentz group, so that G/H is Minkowski space. The general theory of these homogeneous space compactifications is being developed in a series of papers.[1,2] We propose here several candidates for G' and H' that may be related to $SU(3)$ symmetry and scattering theory and discuss some of their geometric properties. The general background for this work is Cartan's theory of symmetric spaces.[3]

2. THE CONFORMAL COMPACTIFICATION OF MINKOWSKI SPACE

R^n will denote the real Euclidean space of n dimensions. We start off with R^6, defined by coordinates $y_1, \cdots, y_6$, which are to be considered as functions on R^6. A point of R^6 is a vector $(y_1, \cdots, y_6)$. $P^5(R)$, the real projective space of five dimensions, is obtained by identifying two such vectors that differ only by a scalar, nonzero multiple. Let q be the polynomial $y_1^2 + \cdots + y_4^2 - y_5^2 - y_6^2$. Since q is homogeneous, the relation $q(y) = 0$ passes to the quotient to define a five-dimensional quadric hypersurface Q in $P^5(R)$. Let $GL(6, R)$ be the group of invertable 6×6 real matrices, acting on R^6 in the linear way. It too passes to the quotient to act on $P^5(R)$ as the projective group. Let $GL(6, R, q)$ be the subgroup of $GL(6, R)$ that preserves q.

Let $x_1, \cdots, x_4$ be the real functions defined by

$$x_1 = y_2/(y_6 - y_1), \cdots, x_4 = y_5/(y_6 - y_1).$$

Since they are invariant under scalar multiplication on R^6, they pass to the quotient to define functions everywhere on $P^5(R)$, except on the hyperplane $y_6 - y_1 = 0$, which will be used to identify this subset of $P^5(R)$ with R^4 and to define the Lorentz metric $ds^2 = dx_1^2 + dx_2^2 + dx_3^2 - dx_4^2$. Let us compute this ds^2 in terms of the y's.

$$dx_1 = (y\,dy_2 - y_2\,dy)/y^2, \quad \text{with} \quad y = y_6 - y_1,$$

$$dx_1^2 = (y^2\,dy_2^2 + y_2^2\,dy^2 - 2y_2 y\,dy\,dy_2)/y^4,$$

$$ds^2 = (dy_2^2 + \cdots - dy_5^2)/y^2 + [y(y_6 + y_1)\,dy^2 - y\,dy(dy_6^2) - d(y_1)^2]/y^4.$$

The numerator of the second term is

$$y(y_6 + y_1)\,dy^2 - y\,dy(dy(y_6 + y_1) + d(y_6 + y_1)y) = -y^2(dy_6^2 - dy_1^2),$$

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

[1] R. Hermann, Proc. Nat. Acad. Sci. U. S. 51, 456 (1964).

[2] R. Hermann, "Compactifications of Homogeneous Space," J. Math. and Mech. (to be published).

[3] S. Helgason, *Differential Geometry and Symmetric Spaces* (Academic Press Inc., New York, 1962).

whence

$$y^2\,ds^2 = dy_1^2 + \cdots - dy_5^2 - dy_6^2. \qquad (1)$$

Now, $GL(6, R, q)$ preserves the right-hand side, hence preserves ds^2 up to a scalar factor, i.e., is a conformal transformation of the Lorentz metric. It is well known that every conformal transformation is of this form. (In fact, we have just given the classical Klein construction of the conformal group.) The conformal group acts only "locally" on Minkowski space, but acts globally on the compact space Q in which Minkowski space is imbedded.

The subgroup of $GL(6, R, q)$ corresponding to the Poincaré group can most readily be found by working backwards to the usual definition of the Poincaré group. Consider a linear transformation R^6, and let $y_1', \cdots, y_6'$ be the transformed functions, i.e.,

$$y_i' = \sum_j a_{ij} y_j.$$

Let $x_1', \cdots, x_4'$ be the transforms of the x's,

$$x_i' = \frac{y_i'}{y_6 - y_1}$$

$$= \frac{\sum_j a_{ij} y_j}{\sum_j (a_{6j} - a_{1j}) y_j} = \frac{\sum_j a_{ij} \dfrac{y_2}{y_6 - y_1}}{\sum_j (a_{6j} - a_{1j}) \dfrac{y_2}{y_6 - y_1}}.$$

A transformation of the Poincaré group gives a relation of the form

$$x_i = \sum_{j=2}^{5} b_{ij} x_j + c, \qquad 2 \leq i \leq 5.$$

Comparing coefficients gives

$$\begin{aligned} b_{ij} &= a_{ij} \quad \text{for} \quad 2 \leq i, j \leq 5, \\ a_{ij} &= c_i(a_{61} - a_{11}), \\ a_{i6} &= c_i(a_{66} - a_{16}). \end{aligned} \qquad (2)$$

On the other hand, a group closely related to $SU(3)$ is a subgroup of $GL(6, R, q)$. This can be seen qualitatively as follows: If everything is made complex, the group leaving q invariant is just $SO(6, C)$, the orthogonal group in six variables, i.e., $GL(6, R, q)$ is a "real form" of $SO(6, C)$. $SU(3)$ is known to be a subgroup of $SO(6, R)$, which is a compact real form of $SO(6, C)$. Hence, if one is willing to ignore the distinction between real and complex coefficients (which is more-or-less all right if one is concerned with finite-dimensional representations of the group), $SO(3)$ can also act on Q. More precisely, proceed as follows: Introduce the complex variables

$$z_1 = y_1 + iy_2, \qquad z_2 = y_3 + iy_4, \qquad z_3 = y_5 + iy_6,$$

where $i = (-1)^{\frac{1}{2}}$. This means relabeling R^6 as C^3, the spare of three complex variables. An arbitrary element of $GL(6, R, q)$ permutes the z's and their conjugates. The subgroup that permutes the z's *alone* is a subgroup which is a real form of the complex group whose compact real form is $SU(3)$. Then, we have the possibility of realizing the vectors of the observed representation of $SU(3)$ as "wavefunctions" that "live" on Q, just as in ordinary Lorentz-invariant physics the wavefunctions "live" on Minkowski space.

The reader might ask: Why do we bother to use such an elaborate compactification of R^4, when there is a very simple one available, namely the compactification obtained by tacking on a single "point at infinity." The answer to this is that it is precisely the distinction between timelike and spacelike lines in Minkowski space that forces such an elaboration. For, in the one-point compactification, all lines meet at the same point at infinity. Both for physical and geometric reasons it is desirable to use compactifications which reflect this fundamental distinction. We might remark that if one started with Euclidean R^4, precisely the one-point compactification is obtained: For, $q(z)$ would be $z_1^2 + z_2^2 + \cdots + z_5^2 - z_6^2$, $Q - R^4$ would be determined by the relation $z_1 - z_6 = 0$ and $q(z) = 0$, which is satisfied by precisely one point.

It is very easy to compute the limits of lines in Minkowski space. Suppose $s \to sx = s(x_1, \cdots, x_4)$ is such a line. Along such a line, $z_1 - z_6$ should approach zero; we can suppose without loss of generality that $z_1 - z_6 = 1/s$. Thus, the line approaches, say, the point $(z_1, x_1, x_2, x_3, x_4, z_1)$. It is then seen using relations (2) that the translation subgroup of the Poincaré group maps this point into itself, i.e., acts as the identity on the boundary of R^4. Since the translations from an invariant subgroup, and the quotient of the Poincaré group by it is the Lorentz group, we see that the Lorentz group passes to the quotient to act on the boundary. The isotropy subgroup of the Lorentz group at the boundary point is the Wigner "little group," hence there are esentially three different orbits of the Poincaré group on the boundary, corresponding to whether the vector $x = (x_1, \cdots, x_4)$ is time-, light-, or spacelike. Then, the "wavefunctions" corresponding to the representations of the Poincaré group can be considered as functions on the boundary

of Minkowski space in the conformal compactification, rather than, as customary, as functions on the mass shell.

Now we turn to indicating possible relations to the thory of asymptotic behavior of solutions of equations of the form

$$\Box\psi = V\psi. \tag{3}$$

(This is the mathematical problem inherent in scattering theory, of course.) Now, the D'Alembertian operator $\Box$ is the Laplace–Beltrami operator associated with the Lorentz metric ds^2. If ds'^2 is another metric conformal to ds^2, i.e.,

$$ds^2 = ds'^2/\lambda^2,$$

then

$$\Box\psi = \lambda^2\,\Box'\psi + \cdots,$$

where $\Box'$ is the Laplace–Beltrami operator with respect to the primed metric, and the dots indicate terms of higher order in λ. Since $(y_1, \cdots, y_6)$ are homogeneous coordinates, we can normalize so that, say, $y_6 = 1$. Q is then determined by the relation

$$y_1^2 + \cdots - y_5^2 = 1. \tag{4}$$

Then, if $ds'^2 = dy_1^2 + \cdots - dy_5^2$, we know that

$$ds^2 = ds'^2/(z_1 - 1)^2,$$

so that $\lambda = y_1 - 1$. $\Box'$ is the Laplace–Beltrami operator on the hyperboloid (4) with respect to the metric that is induced from the s-dimensional Lorentz metric on $(y_1, \cdots, y_5)$ space, so remains well behaved near the boundary of R^4 in Q. λ can be written in more familiar terms:

$$\lambda(z_1 + 1) = y_2^2 + \cdots - y_5^2 = \lambda^2(x_2^2 + \cdots - x_4^2),$$

so that $1/\lambda$ grows as the square of the Minkowski distance.

3. COMPACTIFICATION OF MINKOWSKI SPACE IN $P^2(C)$

We now briefly indicate another possible compactification of Minkowski space that is more closely related to $SU(3)$. One of the most remarkable facts about the observed $SU(3)$ symmetry is that only the representations of the adjoint group $SU(3)/Z_3$ seem to have appeared. [Z_3 is the center of $SU(3)$, and is the cyclic group of order three.] The representations of the other groups that appear in physics (e.g., the Poincaré and rotation groups) all occur by decomposing the geometric action of the group on a space possibly with an attached homogeneous vector bundle to account for internal symmetries. If it should happen that $SU(3)$ also enters physics in this way, (which might indicate a "modified" geometry underlying physics, just as general relativity "modifies" the Minkowski geometry), it would be to a crucial question whether the center of $SU(3)$ acts as the identity. Now, the simplest such space is $P^2(C)$, the two (complex)-dimension projective space. It is obtained by starting with C^3, the three-dimensional complex vector space, say with coordinates (z_1, z_2, z_3), and identifying the two vectors that differ by a complex scalar multiple. $SU(3)$ acts on C^3 in its usual linear way, and passes to the quotient to act on $P^2(C)$, with the center acting as the identity on $P^2(C)$. Off the hyperplane at "infinity" determined by $z_3 = 0$, homogeneous coordinates can be introduced, say $z_1/z_3, z_2/z_3$, identifying the complement of the hyperplane with C^2, which can be identified with R^4. Again, the Poincaré group on R^4 would extend to an action on $P^2(C)$, and the same procedure could be used to determine the D'Alembertian near the boundary. However, it would not be conformal to an operator behaving nicely near the boundary, hence the asymptotic behavior would be expected to be more complicated.

Problem of Subtractions in Potential Scattering

D. Bessis
Centre d'Etudes Nucléaires de Saclay, Service de Physique Théorique, Gif-sur-Yvette (Seine-et-Oise), France
(Received 9 November 1964)

We study the problem of subtractions in potential scattering theory. We show that the single spectral functions can be obtained by analytic continuation of the Mellin transform of the double spectral function. The N/D method could thus be avoided, in principle.

I. INTRODUCTION

THE Mandelstam representation in potential scattering can be written down, following Blankenbecler, Goldberger, Khuri, and Treiman[1]:

$$f(s,t) = f_B(t) + \frac{t^{L_0+1}}{\pi^2}\int_0^\infty\int_0^\infty \frac{ds'\,dt'\,\rho(s',t')}{(s'-s)(t'-t)t'^{L_0+1}} + \sum \frac{\Gamma_i(t)}{s+s_i} + \sum_{k=0}^{k=L_0} t^k \frac{1}{\pi}\int_0^\infty \frac{\rho_k(s')}{s'-s}\,ds', \quad \text{Im } s \neq 0. \tag{1}$$

Using only unitarity it is possible to calculate the double spectral function $\rho(s,t)$,[1] from the asymptotic behavior of the amplitude for $s \to \infty$. We would like to be able to calculate the simple spectral functions $\rho_k(s')$, the position of the poles $-s_i$ (bound states) and the value of their residues $\Gamma_i(t)$, without having to come back to the Schrödinger equation or to use partial wave expansions, which are both, in some sense, foreign to the Mandelstam representation.

In this paper, we limit ourselves to the problem of the simple spectral functions. The results for the problem of the poles and their residues are given in the conclusion. Their demonstration can be found in.[2]

II. THE PRINCIPLE OF THE METHOD

The $\rho_p(s)$ are, for $p \leq L_0$ and $s > 0$, the derivatives taken at $t = 0$ of the absorptive part of the amplitude $f_s(s,t)$. More precisely we have

$$\rho_p(s) = \frac{1}{p!}\frac{\partial^p}{\partial t^p} f_s(s,t)|_{t=0}, \qquad p \leq L_0, \tag{2}$$

because of the fact that

$$f_s(s,t) = \frac{t^{L_0+1}}{\pi}\int_0^{+\infty} \frac{\rho(s,t')\,dt'}{t'^{L_0+1}(t'-t)} + \sum_{p=0}^{p=L_0} t^p \rho_p(s). \tag{3}$$

If we evaluate the derivatives at $t = 0$ of $f_s(s,t)$ for $p > L_0$ we get

$$\frac{1}{p!}\frac{\partial^p}{\partial t^p} f_s(s,t)|_{t=0} = \frac{1}{\pi}\int_0^{+\infty} \frac{\rho(s,t')}{t'^{p+1}}\,dt', \quad p > L_0. \tag{4}$$

Equation (4) only has a significance if $p > L_0$, because for $p \leq L_0$ the integral in (4) is divergent. Otherwise the number of subtractions could have been reduced. Let us suppose that a meaning can be given to the integral for $p \leq L_0$. It could then happen that the functions we obtain for $p = 0, 1, \cdots, L_0$ be precisely the $\rho_p(s)$. One possibility is that the analytic continuation (if it exists) of the integral (4) for the values $p = 0, 1, \cdots, L_0$, gives the functions $\rho_p(s)$. We demonstrate that it is so.

The demonstration results from the following steps:

(A) Using the expansion of the amplitude in partial waves, we calculate the derivatives at $t = 0$ of $f_s(s,t)$. The formula we obtain, allows us to define without difficulty, an extrapolation for p complex of the derivatives of $f_s(s,t)$:

$$R_p(s) = \frac{1}{p!}\frac{\partial^p}{\partial t^p} f_s(s,t)|_{t=0}. \tag{5}$$

The extrapolated function obtained in this way, $R(p,s)$ has the following properties:

(1) $R(p,s)$ is a meromorphic function of p for Re $p > -\frac{1}{2}$ $(s > 0)$;

(2) for Re $p > L_0$, $R(p,s)$ is holomorphic in p $(s > 0)$;

(3) for Re $p > L_0$ and $|p| \to \infty$,

$$|R(p,s)| < Ce^{\alpha|\text{Im } p| + \beta \text{ Re } p},$$

where C is a constant independent of p, as well as α and β, and $\alpha < \pi$, and

(4) $$R(p,s) = R_p(s) = \frac{1}{p!}\frac{\partial^p}{\partial t^p} f_s(s,t)|_{t=0}$$

for $p = 0, 1, 2, \cdots, \infty$.

[1] R. Blankenbecler, M. L. Goldberger, N. N. Khuri, and S. B. Treiman, Ann. Phys. (N. Y.) **10**, 62 (1960).
[2] D. Bessis, thesis, Paris, 1965.

(B) The study of the function

$$\bar{R}(p, s) = \frac{1}{\pi}\int_0^\infty \frac{\rho(s, t')}{t'^{p+1}}\, dt' \tag{6}$$

will show us that:

(1) $\bar{R}(p, s)$ is holomorphic in p for Re $p > L_0$, $(s > 0)$;

(2) $\bar{R}(p, s)$ is bounded by: $\bar{C}e^{\bar{\alpha}|\mathrm{Im}\, p| + \bar{\beta}\,\mathrm{Re}\, p}$, $(s > 0)$;

with $\bar{C}$, $\bar{\alpha}$, $\bar{\beta}$, constants independent of p and $\bar{\alpha} < \pi$;

$$(3) \qquad \bar{R}(p, s) = R_p(s) = \frac{1}{p!}\frac{\partial^p}{\partial t^p} f_s(s, t)|_{t=0}$$

$$\text{for} \quad p = L_0 + 1, \quad L_0 + 2, \cdots, \infty.$$

(C) If one applies Carlson's theorem[3] to the function

$$R(p, s) - \bar{R}(p, s)$$

which vanishes for $p = L_0 + 1, L_0 + 2, \cdots, \infty$; which is holomorphic for Re $p > L_0$; which is bounded by $C'e^{\alpha'|\mathrm{Im}\, p| + \beta'\,\mathrm{Re}\, p}$, when $|p| \to \infty$ and Re $p > L_0$, with

$$C' = \sup\,[C, \bar{C}],$$

$$\beta' = \sup\,[\beta, \bar{\beta}],$$

$$\alpha' = \sup\,[\alpha, \bar{\alpha}] < \pi,$$

we see that

$$R(p, s) - \bar{R}(p, s) \equiv 0. \tag{7}$$

(D) It follows that $\bar{R}(p, s)$ actually has an analytic continuation for Re $p \leq L_0$ at least up to Re $p > -\frac{1}{2}$. The single spectral functions are simply equal to

$$\rho_p(s) = \bar{R}(p, s), \qquad p = 0, 1, 2, \cdots, L_0. \tag{8}$$

We now give a rigorous demonstration of the properties (A) and (B).

Remark: To be concise, we shall say that a function is "Carlsonian" in the variable z if (1) it is holomorphic in z for Re $z > N$, N finite; (2) for $|z| \to \infty$ Re $z > N$, it is bounded by $Ce^{\alpha|\mathrm{Im}\, z| + \beta\,\mathrm{Re}\, z}$ with C, α, β finite constants independent of z and $\alpha < \pi$.

III. DEFINITION OF THE FUNCTION $R(p, s)$

We know that the absorptive part of the amplitude can be written

$$f_s(k, \cos\theta) = \frac{1}{k}\sum_{l=0}^{l=\infty}(2l + 1) \times \mathrm{Im}\, f(l, k)P_l(\cos\theta), \qquad k > 0, \tag{9}$$

with $s = k^2$ and $\cos\theta = 1 + t/2s$. If we use Neumann's theorem, we see that this series converges uniformly with respect to $\cos\theta$, inside every ellipse $(\mathcal{E})$ of the $\cos\theta$ plane, having foci -1, $+1$, and in which $f_s(k, \cos\theta)$ is holomorphic in $\cos\theta$. We know from the formula (3) and the fact that $\rho(s, t) = 0$ for $t < 4\mu_0^2 + \mu_0^4/s$,[1] that the greatest ellipse $\mathcal{E}$ passes through the point

$$\cos\theta_0 = 1 + 4\mu_0^2/2s + \mu_0^4/2s^2. \tag{10}$$

If we go to the t plane, we see that the series

$$f(k, t) = \frac{1}{k}\sum_{l=0}^{l=\infty}(2l + 1)\ \mathrm{Im}\, f(l, k)P_l(1 + t/2s) \tag{11}$$

converges uniformly in t, when $t \in \mathcal{E}_s$, where $(\mathcal{E}_s)$ is an ellipse in the t plane having foci 0 and $-4s$ which passes through the point $4\mu_0^2 + \mu_0^4/s$. This ellipse which is s dependent always contains in its interior the circle of center 0 and radius $4\mu_0^2$.

Then, when $|t| < 4\mu_0^2$ Eq. (11) is a series of holomorphic functions of t, which is uniformly convergent with respect to t.

It is known[4] that we can differentiate the series as many times as we like. The series of derivatives is still uniformly convergent and equal to the derivative of the function. So,

$$k\frac{\partial^p}{\partial t^p} f_s(k, t)|_{|t|<4\mu_0^2} = \frac{1}{(2s)^p}\sum_{l=0}^{l=\infty}(2l + 1)\ \mathrm{Im}\, f(l, k)P_l^{(p)}(1 + t/2s).$$

In particular,

$$\frac{1}{p!}\frac{\partial^p}{\partial t^p} f_s(k, t)|_{t=0} = \frac{1}{kp!\,(2k^2)^p}\sum_{l=0}^{l=\infty}(2l + 1) \times \mathrm{Im}\, f(l, k)P_l^{(p)}(1). \tag{12}$$

But we have[5]

$$P_l^{(p)}(1) = \begin{cases}(l + p)!/2^p p!\,(l - p)!, & p \leq l\\ 0, & p > l,\end{cases}$$

and so

$$(1/p!)(\partial^p/\partial t^p)f_s(k, t)|_{t=0} = \frac{2}{[p!]^2 2^{2p+1}k^{2p+1}}\sum_{l=p}^{l=+\infty}(2l + 1)\ \mathrm{Im}\, f(l, k)\,\frac{(l + p)!}{(l - p)!};$$

[3] M. Froissart, *High Energy Properties in Theoretical Physics*, Trieste, Italy, 1962 (International Atomic Agency, Vienna, 1963), p. 38.

[4] G. Valiron, *Théorie des fonctions* (Masson, Paris, 1948), p. 376.

[5] L. Robin, *Fonctions sphériques de Legendre* (Gauthier-Villars, Paris, 1957), Vol. 1, p. 77.

if we put $l = n + p$, we get

$$(1/p!)(\partial^p/\partial t^p)f_s(k, t)|_{t=0} = 2e^{-(2p+1)\log 2k}/[\Gamma(p+1)]^2 \times \sum_{n=0}^{n=\infty}(2n+1+2p) \times \operatorname{Im} f(n+p, k)\Gamma(n+2p+1)/n!. \quad (14)$$

It is then natural to introduce the function $R(p, s)$, defined for complex p by

$$R(p, s) = \frac{2e^{-(2p+1)\log 2k}}{[\Gamma(p+1)]^2}\sum_{n=0}^{\infty}(2n+1+2p) \times f(n+p, k)f^*(n+p^*, k)\Gamma(n+2p+1)/n!, \quad (15)$$

through the use of generalized unitarity:

$$-i[f(l, k) - f^*(l^*, k^*)] = 2f(l, k)f^*(l^*, k^*). \quad (16a)$$

IV. STUDY OF THE FUNCTION $R(p, s)$

The function

$$2e^{-(2p+1)\log 2k}/[\Gamma(p+1)]^2$$

is an integral function of p. Let us look at a term of the series:

$$(2n+1+2p)f(n+p, k)f^*(n+p^*, k) \times \Gamma(n+2p+1)/n!. \quad (16b)$$

We know that $f(l, k)$ is a meromorphic function of l, for $k > 0$ and $\operatorname{Re} l > -\frac{1}{2}$ whose poles are at the left of the line $\operatorname{Re} l = L_0$ and above the line $\operatorname{Im} l = 0$, when k is positive.[6]

As soon as n is greater than L_0, and $\operatorname{Re} p > -\frac{1}{2}$, the term (16b) defines a holomorphic function of p. We split the series into two parts:

$$\Sigma_1(p) = \sum_{n=0}^{n=L_0}(2n+1+2p)f(n+p, k)f^*(n+p^*, k) \times \Gamma(n+2p+1)/n!, \quad (17)$$

$$\Sigma_2(p) = \sum_{n=L_0+1}^{n=\infty}(2n+1+2p)f(n+p, k) \times f^*(n+p^*, k)\Gamma(n+2p+1)/n!. \quad (18)$$

It is clear that $\Sigma_1(p)$ is a meromorphic function of p, the poles of which are very simply related to the Regge poles: to each Regge pole $l_\alpha(k)$, $\alpha = 1, 2, \cdots, \alpha_M(k)$ at the energy k^2 in the l plane corresponds a double series of poles in the p plane:

$$l_\alpha(k) \to l_\alpha(k);\quad l_\alpha(k) - 1;\quad l_\alpha(k) - 2;\ \cdots \quad (19)$$

and

$$l_\alpha(k) \to l_\alpha^*(k);\quad l_\alpha^*(k) - 1;\quad l_\alpha^*(k) - 2;\ \cdots .$$

[6] D. Bessis, Nuovo Cimento **33**, 797 (1964).

We show that $\Sigma_2(p)$ is holomorphic in p, by showing the uniform and absolute convergence of the series (18). For this we need an upper bound on the amplitude $f(l, k)$ which decreases exponentially when $\operatorname{Re} l \to +\infty$.

It is known[7] that for the class of potentials in which we are interested, the amplitude is, for $k > 0$ fixed and $\operatorname{Re} l > L_0$ $|l| \to \infty$, bounded by

$$Ce^{-\alpha(k)\operatorname{Re} l}, \quad (20)$$

where

$$\alpha(k) = 2\log\left[\frac{\mu_0}{2k} + (1 + \mu_0^2/4k^2)^{\frac{1}{2}}\right] > 0.$$

If we take into account the formulas

$$n! > (2\pi n)^{\frac{1}{2}}n^n e^{-n}, \quad n \text{ integer},[8] \quad (21)$$

and

$$|\Gamma(z)| < (2\pi)^{\frac{1}{2}}\exp\{[\operatorname{Re} z - (\tfrac{1}{2})]\log|z| - \operatorname{Im} z \arg z - \operatorname{Re} z + (12\operatorname{Re} z)^{-1}\} \quad (22)$$

for $\operatorname{Re} z > 0$.

We see, using Eqs. (20), (21), and (22), that for n sufficiently large, the general term of the series (18) is less than

$$C^2[2n+1+2|p|]\exp\{-2n\alpha(k) - 2\alpha(k)\operatorname{Re} p + [n + 2\operatorname{Re} p + (\tfrac{1}{2})]\log|n+2p+1|\} \times \exp\{-2\operatorname{Im} p \arg(n+2p+1) - 2\operatorname{Re} p - 1 - (n+\tfrac{1}{2})\log n\}.$$

When p is in any compact set, each term of the series (18) is bounded by the corresponding term of an absolutely convergent series independent of p: the convergence is therefore uniform and Σ_2 is holomorphic in p, for $\operatorname{Re} p > -\frac{1}{2}$.

We see that $R(p, s)$ is an analytic function of p for each real value of s, which is holomorphic in p for $\operatorname{Re} p > L_0$ and meromorphic in p for $-\frac{1}{2} < \operatorname{Re} p \leq L_0$. The poles of $R(p, s)$ are given by Eq. (19).

We now show that $R(p, s)$ is "Carlsonian" for $\operatorname{Re} p > L_0$.

When $\operatorname{Re} z > 0$ and $|\gamma| < \frac{1}{2}\pi$, one can write

$$\Gamma(z) = e^{i\gamma z}\int_0^{+\infty} e^{-ue^{i\gamma}+(z-1)\log u}\,du. \quad (23)$$

From Eq. (23) we get

$$|\Gamma(z)| \leq \exp[-\gamma|\operatorname{Im} z| + \operatorname{Re} z \log(\cos\gamma)^{-1}]\Gamma(\operatorname{Re} z) \quad (24)$$

[7] A. Martin, Nuovo Cimento **31**, 1229 (1964).

[8] E. T. Whittaker and G. N. Watson, *Modern Analysis* (Cambridge University Press, Cambridge, England, 1958), p. 253.

for any γ satisfying $0 \leq \gamma < \frac{1}{2}\pi$ and $\mathrm{Re}\ z > 0$. When $\mathrm{Re}\ p > L_0$, all terms of the series (15) are holomorphic in p, and we have

$$|R(p,s)| \leq \sigma \frac{2C^2}{|\Gamma(p+1)|^2} \times \exp\left\{-[2\ \mathrm{Re}\ (p+1)]\log 2k - 2\alpha(k)\ \mathrm{Re}\ p - 2\gamma\ |\mathrm{Im}\ p| + [(2\ \mathrm{Re}\ (p+1)]\log\frac{1}{\cos\gamma}\right\} \quad (25)$$

with

$$\sigma = \sum_{n=0}^{n=\infty} [2n+1+2\ |p|]\exp\{-n[2\alpha(k) - \log(\cos\gamma)^{-1}]\}\Gamma[(n+2\ \mathrm{Re}\ (p+1)]/n!. \quad (26)$$

If we take into account that ($\rho > 0$, $x > 0$)

$$\sum_{n=0}^{n=\infty} e^{-\rho n}\frac{\Gamma(n+x+1)}{n!} = \frac{\Gamma(x+1)}{[1-e^{-\rho}]^{1+x}},$$

$$\sum_{n=0}^{n=\infty} ne^{-\rho n}\frac{\Gamma(n+x+1)}{n!} = e^{-\rho}\frac{\Gamma(x+2)}{[1-e^{-\rho}]^{x+2}},$$

we see that σ can be written, for

$$\rho = 2\alpha(k) - \log(\cos\gamma)^{-1} > 0,$$

or

$$\gamma < \arccos e^{-2\alpha(k)}, \quad (26)$$

$$\sigma = [1+2\ |p|]\frac{\Gamma[2\ \mathrm{Re}\ (p+1)]}{[1-e^{-\rho}]^{2\ \mathrm{Re}\ (p+1)}} + 2e^{-\rho}\frac{\Gamma[2\ \mathrm{Re}\ (p+2)]}{[1-e^{-\rho}]^{2\ \mathrm{Re}\ (p+2)}};$$

whence

$$\sigma \leq 2[2\ |p|+1]\frac{\Gamma[2\ \mathrm{Re}\ (p+1)]}{[1-e^{-\rho}]^{2\ \mathrm{Re}(p+2)}}, \quad (27)$$

and

$$|R(p,s)| \leq (2C^2/\pi)[2\ |p|+1] \times \exp\{-[2\ \mathrm{Re}\ (p+1)]\log 2k - 2\alpha(k)\ \mathrm{Re}\ p + [2\ \mathrm{Re}\ (p+1)]\log(1/\cos\gamma)\} \times \exp\{2\ \mathrm{Re}\ [(p+1)/2]\log|2\ \mathrm{Re}\ (p+1)| - [2\ \mathrm{Re}\ (p+2)]\log(1-e^{-\rho}) - [2\ \mathrm{Re}\ (p+1)]\log|p+1|\} \times \exp\{2\ \mathrm{Im}\ p\ \arg(p+1) - 2\gamma\ |\mathrm{Im}\ p|\},$$

and finally

$$|R(p,s)| \leq C'e^{\beta\ \mathrm{Re}\ p + \alpha\ |\mathrm{Im}\ p|}, \quad (28)$$

where C', β, α are finite and $\alpha \leq \pi - 2\gamma < \pi$.

V. THE FUNCTION $\bar{R}(p,s)$

The function

$$\bar{R}(p,s) = \frac{1}{\pi}\int_{4\mu_0^2+\mu_0^4/s}^{+\infty}\frac{dt}{t^{p+1}}\rho(s,t) \quad (29)$$

is the Mellin transform of $\rho(s,t)$ with respect to t. To study its analytic properties in p, we need some bound on $\rho(s,t)$.

It has been shown in Ref. 9 that, for potentials $V(r)$, (1) holomorphic in $\mathrm{Re}\ r > 0$, (2) such that $|V(r)| < K/|r|^\rho$ for $|r| \leq 1$, $\rho < 2$; (3) $|V(r)| < K/|r|^\gamma e^{-\mu_0\ \mathrm{Re}\ r}$, for $|r| \leq 1$, $\mu_0 > 0$, $\gamma > \frac{7}{4}$; $\rho(s,t)$ is a continuous function of s and t, bounded by

$$|\rho(s,t)| < C(V)\left\{\frac{1}{s^{\frac{1}{4}}t^{\frac{1}{4}}} + \frac{(2+t)^{L_0}}{s^{\frac{1}{2}}}\right\}, \quad (30)$$

where $C(V)$ is a constant independent of s and t and L_0 is the number of subtractions in t.

When $\mathrm{Re}\ p > L_0 + \eta$ $\eta > 0$, the integral converges uniformly, because of the bound (30) on $\rho(s,t)$ so that the integral defines a holomorphic function of p, for $\mathrm{Re}\ p > L_0$. Since

$$|\bar{R}(p,s)| \leq \frac{1}{\pi}\int_{4\mu_0^2+\mu_0^4/s}^{+\infty}\frac{dt}{t^{\mathrm{Re}\ (p+1)}}|\rho(s,t)| \leq \frac{1}{\pi}\int_{4\mu_0^2+\mu_0^4/s}^{+\infty}\frac{dt}{t^{1+\eta+L_0}}|\rho(s,t)|,$$

and using the bound (30), we find

$$|\bar{R}(p,s)| < C(\eta, L_0, \mu_0), \quad (31)$$

where C does not depend on p or s.

This shows that for $0 \leq s \leq +\infty$: $R(p,s)$ is holomorphic in p for $\mathrm{Re}\ p > L_0$ and "Carlsonian."

VI. THE EQUALITY OF $R(p,s)$ AND $\bar{R}(p,s)$

The function $R(p,s) - \bar{R}(p,s)$ is, for all positive s, (1) holomorphic in p for $\mathrm{Re}\ p > L_0$; (2) "Carlsonian" for $|p| \to \infty$ $\mathrm{Re}\ p > L_0$, because it is the difference of two "Carlsonian" functions; (3) equal to zero for $p = L_0 + 1, L_0 + 2, \cdots, \infty$, by construction.

It then follows from Carlson's theorem that it is zero everywhere. Thus $\bar{R}(p,s)$ has an analytic continuation which is $R(p,s)$. This last function is meromorphic in p for $\mathrm{Re}\ p > -\frac{1}{2}$. The simple spectral functions are given by

$$\rho_K(s) = R(k,s). \quad (32)$$

[9] D. Bessis, J. Math. Phys. 6, 637 (1965).

VII. CONCLUSION

We have shown that the Mellin transform of the double spectral function, defined only for Re $p >$ L_0 = number of subtractions, by

$$\bar{R}(p, s) = \frac{1}{\pi}\int_0^\infty \frac{dt}{t^{p+1}}\,\rho(s, t), \qquad s > 0$$

can be analytically continued up to Re $p > -\frac{1}{2}$, into a meromorphic function of p: $R(p, s)$. The poles of this function $R(p, s)$ are very simply related to the Regge poles: to each Regge pole $l(k)$ in the l plane correspond two series of poles of $R(p, s)$,

$$l(k);\quad l(k)-1;\quad l(k)-2;\quad \cdots,$$
$$l^*(k);\quad l^*(k)-1;\quad l^*(k)-2;\quad \cdots.$$

The simple spectral functions $\rho_p(s)$ are given by

$$\rho_p(s) = R(p, s) \quad \text{for} \quad p = 0, 1, 2, \cdots L_0.$$

To complete our program, we would like to be able to calculate the bound-state energies $-s_i^{(\alpha)}$ and their residues $\Gamma_i^\alpha(t)$, which we know are polynomials in t, directly from the knowledge of $\rho(s, t)$.

If we start from the Mandelstam representation,

$$f(s, t) = f_B(t) + \frac{t^{L_0+1}}{\pi^2}\int_0^{+\infty} ds' \times \int_{4\mu_0^2}^{+\infty} \frac{dt'\rho(s', t')}{(s'-s)(t'-t)t'^{L_0+1}} + \sum_{p=0}^{p=L_0} t^p \frac{1}{\pi}\int_0^{+\infty}\frac{\rho_p(s')}{s'-s}\,ds' + \sum_{l=0}^{l=L_0}\sum_{\alpha=0}^{\alpha=\alpha(l)} (2l+1)\frac{R_l^\alpha}{s+s_l^\alpha}P_l\left(1-\frac{t}{2s_l^\alpha}\right), \quad (33)$$

where s_l^α is the αth bound state of spin l and binding energy equal to $-s_l^\alpha$. As is known[1] the spin of a bound state cannot exceed the number of subtractions L_0, so this can be rewritten

$$f(s, t) = f_B(t) + \frac{t^{L_0+1}}{\pi^2}\int_0^{\infty} ds' \times \int_{4\mu_0^2}^{+\infty} \frac{dt'\rho(s', t')}{(s'-s)(t'-t)t'^{L_0+1}} + \sum_{p=0}^{p=L_0} t^p F_p(s), \quad (34)$$

with

$$F_p(s) = \frac{1}{\pi}\int_0^\infty \frac{ds'}{s'-s}\,\rho_p(s') + \sum_{L=p}^{L=L_0}\sum_{\alpha=0}^{\alpha=\alpha(l)} \frac{\gamma_{lp}^\alpha}{s+s_l^\alpha}. \quad (35)$$

In particular, we have

$$F_0(s) - \frac{1}{\pi}\int_0^\infty \frac{ds'}{(s'-s)}\,\rho_0(s') = \sum_{\substack{\text{all bound}\\ \text{states}}} \frac{(2l+1)R_l^\alpha}{s+s_l^\alpha}. \quad (36)$$

By using a procedure analogous to the one described in this paper, but somewhat more complicated, it is possible to show that the function $\bar{F}(p, s)$, defined by

$$\bar{F}(p, s) = \frac{1}{\pi^2}\int_0^\infty \frac{dt'}{t'^{p+1}}\int_0^\infty \frac{ds'}{(s'-s)}\,\rho(s', t'), \quad (37)$$

has an analytic continuation $F(p, s)$ to the left of Re $p = L_0$ which is meromorphic down to Re $p > -\frac{1}{2}$. The poles are the same as those of $R(p, s)$ except that there are no poles in Im $p < 0$. Moreover, we have

$$F_p(s) = F(p, s) \quad \text{for} \quad p = 0, 1, 2, \cdots, L_0. \quad (38)$$

With this result, we see that Eq. (36) can be rewritten

$$\text{A.C.}\left\{\frac{1}{\pi}\int_0^\infty \frac{dt'}{t'^{p+1}}\frac{1}{\pi}\int_0^\infty \frac{ds'\rho(s', t')}{s'-s}\right\}_{p=0} - \frac{1}{\pi}\int_0^\infty \frac{ds'}{s'-s} \times \text{A.C.}\left\{\frac{1}{\pi}\int_0^\infty \frac{\rho(s', t')}{t'^{p+1}}\,dt'\right\}_{p=0} = \sum_{\substack{\text{all bound}\\ \text{states}}} \frac{(2l+1)R_l^\alpha}{s+s_l^\alpha},$$

where A.C. stands for analytic continuation in p variable.

Equation (39) gives the solution to our problem: to calculate the position of the bound states and their residues knowing only the double spectral function, we only need the commutator of two operations. The first one is the analytic continuation in p. The second one is the integration over s'. This commutator applied to $\rho(s', t')$ gives a rational fraction whose poles are the bound states.

We summarize here the conditions which we have imposed on the potential and under which our results are valid:

(1) $V(r)$ is holomorphic in Re $r > 0$;

(2) $|V(r)| < K/|r|^\rho$, $|r| \le 1$, Re $r > 0$, $\rho < 2$;

(3) $|V(r)| < Ke^{-\mu_0 \text{Re}\, r}/|r|^\gamma$, $|r| \ge 1$, Re $r > 0$, $\mu_0 > 0$, $\gamma > \frac{1}{4}$.

ACKNOWLEDGMENT

It is a pleasure to thank Professor Froissart for discussions essential to this work.

Note on Non-Landau Singularities*

PAUL FEDERBUSH
Department of Physics and the Laboratory of Nuclear Science, Massachusetts Institute of Technology, Cambridge, Massachusetts
(Received 18 August 1964)

Some non-Landau singularities are discussed using the formalism of Fotiadi, Froissart, Lascoux, and Pham. The simple cases of self-energy and vertex diagrams are treated, as well as the sixth-order scattering ladder diagram.

I. INTRODUCTION

THE singularities of Feynman diagrams can be classified into Landau and non-Landau types. If the Feynman amplitude is studied as an integral over the internal momenta (where the path of integration may be distorted to avoid integrand singularities) the non-Landau singularities arise when the contour experiences difficulties at infinity. In the formalism of Ref. 1, the integration region is transformed into a surface in a compact analytic manifold. We choose to study some non-Landau singularities in this formalism.

The vertex and self-energy diagrams have non-Landau singularities of a very special type. These arise because an extra effective denominator appears in the integral when one transforms into the compact analytic manifold. A more interesting case occurs when one has more than one loop in the diagram, as in the sixth-order scattering diagram.

II. SELF-ENERGY AND VERTEX DIAGRAMS

We change from

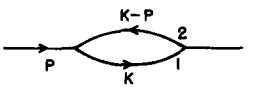
$$k_i \in C^4 \quad \text{to} \quad (x_1, \cdots, x_6) \in W \subset CP^5$$

by

$$\begin{aligned} x_\alpha &= k_\alpha, \qquad \alpha = 1, \cdots, 4, \\ x_5 &= \tfrac{1}{2}(1 - k^2), \qquad x_6 = 1 + k^2, \\ &\left\{\sum_1^5 x_i^2 = \frac{1}{4} x_6^2\right\} = W. \end{aligned} \tag{1}$$

In changing the integral from over k_i to over the x_i variables it is shown in Ref. 1 that one gets an additional effective denominator of the form $(x_6 + 2x_5)$. The non-Landau singularities of the self-energy diagram of Fig. 1 occur when *1*, *2*, and $(x_6 + 2x_5)$ are not in general position in W, and *1* and *2* are in general position. [That is, the zeros of x_6, $x_6 - 2P_\alpha x_\alpha + \frac{1}{2}(x_6 + 2x_5) \times (M_2^2 + P^2 - 1)$, and $x_6 + 2x_5$ are not in general position in W. M_1 has been taken equal one.] In this case the only time the three surfaces are not in general position is when the intersection of these three manifolds is singular. That is, when

FIG. 1. The self-energy diagram.



$$\begin{aligned} &\sum_1^5 x_i^2 = \frac{1}{4} x_6^2, \qquad x_6 + 2x_5 = 0, \qquad x_6 = 0, \\ &x_6 - 2P_\alpha x_\alpha + \tfrac{1}{2}(x_6 + 2x_5)(M_2^2 + P^2 - 1) = 0, \end{aligned} \tag{2}$$

and the four differential forms

$$\begin{aligned} &\sum_2^5 2x_\alpha \, dx_\alpha + 2x_5 \, dx_5 - \frac{1}{2} x_6 \, dx_6, \\ &dx_6 + 2 \, dx_5, \qquad dx_6, \\ &dx_6 - 2P_\alpha \, dx_\alpha + \tfrac{1}{2}(dx_6 + 2 \, dx_5)(M_2^2 + P^2 - 1) \end{aligned} \tag{3}$$

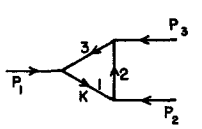
are linearly dependent. [We have assumed, without loss of generality in the conclusion, that $x_1 \neq 0$, $P_1 = 0$.] This occurs as one sees easily when $P^2 = 0$.

In considering the vertex diagram of Fig. 2, the non-Landau singularities arise when surfaces *1*, *2*, *3* and $(x_6 + 2x_5)$ are not in general position in W, but *1*, *2*, and *3* are. If *1*, *2*, and $(x_6 + 2x_5)$ are not in general position we deduce as in the self-energy case that $P_2^2 = 0$. By considering the other two cases corresponding to omitting denominators *1* or *2* we deduce the non-Landau singularity

$$P_1^2 = 0 \quad \text{or} \quad P_2^2 = 0 \quad \text{or} \quad P_3^2 = 0. \tag{4}$$

The final case arises when all these sets of surfaces are in general position but the set *1*, *2*, *3*, and $(x_6 + 2x_5)$ is not.

FIG. 2. The vertex diagram.



* Work supported in part through funds provided by the Atomic Energy Commission under Contract No. AT (30-1)-2098.

[1] D. Fotiadi, M. Froissart, J. Lascoux, and F. Pham (to be published).

Then

$$\sum_1^5 x_i^2 = \frac{1}{4} x_6^2,$$
$$x_6 = 0, \qquad x_6 + 2x_5 = 0,$$
$$x_6 + 2P_{2\alpha}x_\alpha + \tfrac{1}{2}(x_6 + 2x_5)\cdot(M_2^2 + P_2^2 - 1) = 0,$$
$$x_6 - 2P_{1\alpha}x_\alpha + \tfrac{1}{2}(x_6 + 2x_5)\cdot(M_3^2 + P_1^2 - 1) = 0, \tag{5}$$

and the five differentials

$$\sum_2^5 2x_\alpha\, dx_\alpha - \frac{1}{2} x_6\, dx_6,$$
$$dx_6, \qquad dx_6 + 2\, dx_5,$$
$$dx_6 + 2P_{2\alpha}\, dx_\alpha + \tfrac{1}{2}(dx_6 + 2\, dx_5)\cdot(M_2^2 + P_2^2 - 1),$$
$$dx_6 - 2P_{1\alpha}\, dx_\alpha + \tfrac{1}{2}(dx_6 + 2\, dx_5)\cdot(M_3^2 + P_1^2 - 1) \tag{6}$$

are linearly dependent. (Again we assume $x_1 \neq 0$, $P_{11} = P_{21} = 0$.) This occurs when

$$P_1^2 P_2^2 = (P_1 \cdot P_2)^2; \tag{7}$$

Eqs. (4) and (7) comprise the non-Landau singularities of the vertex diagram.

III. THE SIXTH-ORDER LADDER DIAGRAM

The diagram of Fig. 3 is studied in Ref. 2. First one maps k, k^* into $W_1 \times W_2 \subset CP^5 \times CP^5$ similar to what was done in Sec. II. There does not arise in this case any effective denominators. Singularities occur when *1*, *2*, *3*, *4*, *5*, *6*, and *7* are not in general position in $W_1 \times W_2$. In general position all the denominators except *4* are nonsingular manifolds. We stay off the union of *1* and *7* and therefore may take $x_6 = y_6 = 1$. *4* is singular when $x_5 = y_5 = -\frac{1}{2}$, $x_\alpha = -y_\alpha$, $\alpha = 1, \cdots, 4$, $\sum_1^4 x_\alpha^2 = 0$. We call these points *SG*. The set of denominators and their intersections are singular in general position only where they intersect *SG*. In general position a single denominator other than *4* intersects *SG* in a two-sphere —the intersection of two such denominators in a one-sphere—the intersection of three such denominators in a zero sphere. As in Ref. 2, we do not impose momentum conservation at vertex *V*.

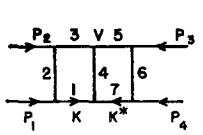


FIG. 3. The sixth-order ladder diagram.

[2] P. Federbush, "Calculation of Some Homology Groups Relevant to Sixth-Order Feynman Diagrams," J. Math. Phys. (to be published).

Remaining off *1* and *7* and exploring the above conditions, we find that there is the following correspondence between surfaces, and some of the non-Landau singularities where they intersect *SG* incorrectly:

$$\begin{aligned}
&6, && P_4^2 = 0;\\
&5, && (P_3 + P_4)^2 = 0;\\
&2, && P_1^2 = 0;\\
&3, && (P_1 + P_2)^2 = 0;\\
&2 \cap 6, && (P_1 + P_4)^2 = 0;\\
&5 \cap 6, && P_3^2 = 0;\\
&2 \cap 3, && P_2^2 = 0;\\
&3 \cap 5, && (P_1 + P_2 + P_3 + P_4)^2 = 0;\\
&3 \cap 6, && (P_1 + P_2 + P_4)^2 = 0;\\
&5 \cap 2, && (P_1 + P_3 + P_4)^2 = 0.
\end{aligned} \tag{8}$$

We consider only one choice of three denominators, the other choices can be obtained by judicious substitution. $2 \cap 5 \cap 6$,

$$P_3^2(P_1 + P_4)^2 - [P_3 \cdot (P_1 + P_4)]^2 = 0. \tag{9}$$

Note added in proof: Other singularities occurring when the intersections with SG are incorrect arise when certain external momenta are linearly dependent (the external momenta are those associated with the surfaces in question). For example, to $2 \cap 6$ is associated the non-Landau singularity with P_1 and P_4 linearly dependent; to $2 \cap 5 \cap 6$ the singularity with P_1, P_3, and P_4 linearly dependent. (I would like to thank J. C. Polkinghorne for a communication on this point.)

IV. CONCLUSION

Non-Landau singularities, which may be calculated by the methods of Ref. 3 or by the present type procedure, are of various types. In more complicated diagrams they will depend on internal masses. The non-Landau singularities that arise from other than the effective denominators are most interesting. In this viewpoint these arise from considering the permanent singularities of certain denominator surfaces or their intersections, in the other method they arise from pinches involving the discriminant. Much remains to be done to understand the significance of these singularities.

[3] D. P. Fairlie, P. V. Landshoff, J. Nuttall, and J. C. Polkinghorne, J. Math. Phys. **3**, 594 (1962).