JOURNAL OF
MATHEMATICAL
PHYSICS

VOLUME 4, NUMBER 10 OCTOBER 1963

Canonical Form of the Covariant Free-Particle Equations

A. CHAKRABARTI
Centre de Physique Théorique de l'Ecole Polytechnique, Paris, France
(Received 21 March 1963)

The canonical form of the covariant equations for free particles of nonzero rest mass is proposed to be taken as $[(p^\nu p_\nu)^{\frac{1}{2}} - \kappa]\,\psi = 0$, instead of $[\gamma^0(p^2 + \kappa^2)^{\frac{1}{2}} - p^0]\,\psi = 0$, as suggested by Foldy. The connection of our representation with the usual forms of the Dirac and the Klein–Gordon (K–G) equations are discussed, each feature being compared with the corresponding one in Foldy's case. The case of the Dirac equation is treated in some detail. A study of the infinitesimal operators of the Poincaré group and the transformation properties of the wavefunction and the polarization operator in our representation lead us to conclude that the choice of operators and the definition of spin states adopted by Iu. M. Shirokov in his study of the Poincaré group corresponds directly to our representation and the canonical form proposed, rather than that proposed by Foldy, as is sometimes supposed. It is also shown that the proposed canonical form corresponds to Wigner's unitary representation of the Poincaré group in terms of the little group of $(\kappa, 0, 0, 0)$ (for $\kappa > 0$). In Appendix A, we give a brief outline of the decomposition of the direct-product representation of the Poincaré group to bring out the special features that arise in our representation. In Appendix B, we compare in detail, for the case of the Dirac equation, our transformation with the well-known Foldy–Wouthuysen transformation. The case of zero rest mass has not been considered. Also the discussion of the position operators in our representation has been left aside and is to be taken up in a following article.

NOTATIONS

WE have taken the metric

$$g^{00} = -g^{11} = -g^{22} = -g^{33} = 1.$$

The γ matrices are choosen such that

$$(\gamma^0)^* = \gamma^0, \qquad (\gamma^i)^* = -\gamma^i \qquad (i = 1, 2, 3),$$

$*$ denoting the complex conjugate of the transposed. As usual $\beta = \gamma^0$ and $\boldsymbol{\alpha} = \gamma^0\boldsymbol{\gamma}$ are all Hermitian.

The dot product of two tensors denotes a contraction over the neighboring indices.

The exterior product of two 4-vectors is defined as

$$(a \wedge b)_{\alpha\beta} = (a_\alpha b_\beta - a_\beta b_\alpha).$$

For the case of 3-vectors, $\mathbf{a} \times \mathbf{b}$ represents the usual vector product (and *not* the 3×3 antisymmetric tensor).

A 4-vector is often denoted as $(a^0, \mathbf{a})$, while a skew tensor S is denoted as

$$S = (\mathbf{A}, \mathbf{B}),$$

meaning $S_{0i} = A_i$, $S_{ij} = \epsilon_{ijk}B_k$, where ϵ_{ijk} is the completely antisymmetric tensor of third rank with components $\pm 1 (\epsilon_{123} = 1)$.

The dual of a skew tensor S is denoted by

$$S^*_{\alpha\beta} = \tfrac{1}{2}\epsilon_{\alpha\beta\rho\lambda}S^{\rho\lambda}$$

(where $\epsilon_{\alpha\beta\rho\lambda}$ being the completely antisymmetric tensor of fourth rank with $\epsilon_{0123} = 1$), leading to

$$S^* = (-\mathbf{B}, \mathbf{A}).$$

I. THE SPIN-$\frac{1}{2}$ CASE AND ITS GENERALIZATION

In his attempt to synthesize the covariant free-particle equations, Foldy[1] proposed the "canonical" form

[1] L. L. Foldy, Phys. Rev. 102, 568 (1956).

$$[i(\partial/\partial t) - \beta\omega]\varphi = 0, \tag{1.1}$$

where $\omega = (\mathbf{p}^2 + \kappa)^{\frac{1}{2}}$, κ being the rest mass. He considered the case of the Dirac, the Klein–Gordon (K–G), and the Proca equations.

First we consider the case of the Dirac equation. The usual Hamiltonian form

$$[i(\partial/\partial t) - H]\varphi = 0, \tag{1.2}$$

where

$$H = \beta\kappa + \boldsymbol{\alpha}\cdot\mathbf{p},$$

is brought to the canonical form by the well-known Foldy–Wouthuysen (F–W) transformation,[2] where

$$\varphi_F = F\varphi, \quad \text{and} \quad F = (\beta H + \omega)/[2\omega(\omega + \kappa)]^{\frac{1}{2}}. \tag{1.3}$$

The infinitesimal operators of the Poincaré group, namely

$$p,\ M = (\mathbf{N}, \mathbf{M}) = (x \wedge p + S)$$

$$[x^\mu \equiv -i(\partial/\partial p_\mu)],$$

where, for the usual representation,

$$(\mathbf{N}, \mathbf{M}) = (x^0\mathbf{p} - \mathbf{x}p^0 + \tfrac{1}{2}i\boldsymbol{\alpha},\ \mathbf{x} \times \mathbf{p} + \tfrac{1}{2}\boldsymbol{\sigma})$$

become, in the new representation (with the subscript F to denote it),

$$p_F = (\beta\omega, \mathbf{p}),$$

$$\mathbf{M}_F = \mathbf{x}\times\mathbf{p} + \tfrac{1}{2}\boldsymbol{\sigma},$$

$$\begin{aligned}\mathbf{N}_F &= x^0\mathbf{p} - \mathbf{x}\beta\omega - [\beta\mathbf{p}\times\boldsymbol{\sigma}/2(\omega + \kappa)] + i\beta\mathbf{p}/2\omega \\ &= x^0\mathbf{p} - \tfrac{1}{2}\beta(\mathbf{x}\omega + \omega\mathbf{x}) \\ &\qquad + [\beta/(\omega + \kappa)]\tfrac{1}{2}\boldsymbol{\sigma}\times\mathbf{p}\end{aligned} \tag{1.4}$$

(In the above we have put $N^1 = M^{01}$ etc . . ., so that Foldy's $\mathbf{K} = -\mathbf{N}$).

For the general case, Foldy proposed $\mathbf{M}$, $\mathbf{N}$ of the above form with $\frac{1}{2}\boldsymbol{\sigma}$ replaced by $\mathbf{S}$, the spin matrix for spin S, in the form

$$\mathbf{S} = \begin{vmatrix} \mathbf{S}' & 0 \\ 0 & \mathbf{S}' \end{vmatrix}, \tag{1.4'}$$

where S' are the $(2s + 1) \times (2s + 1)$ Hermitian matrices satisfying

$$[s_i', s_j'] = i\epsilon_{ijk}s_k'.$$

For the general case, β is defined to be

$$\begin{pmatrix} I' & 0 \\ 0 & -I' \end{pmatrix}, \tag{1.4''}$$

where I' is the $(2s + 1) \times (2s + 1)$ unit matrix.

[2] L. L. Foldy and S. A. Wouthuysen, Phys. Rev. 78, 29 (1950).

Instead of (1.1), we propose the canonical form

$$(\gamma^0 m - \kappa)\varphi = 0, \tag{1.5}$$

where $m \equiv m_{\text{op}} \equiv \epsilon(p^\nu p_\nu)^{\frac{1}{2}}$ and $\epsilon = p^0/|p^0| = \pm 1$ is the sign of the energy, the square root being always taken in the positive sense (see however the argument relating to Eq. (1.11).

To bring the Dirac equation

$$(\gamma\cdot p - \kappa)\varphi = 0 \tag{1.6}$$

to the proposed form we diagonalize, instead of the Hamiltonian (as is the case with F–W transformation), the operator $(\gamma\cdot p)$ through a similarity transformation with the operator[3]

$$Q_{\text{op}} = [\gamma^0(\gamma\cdot p) + m][2m(p^0 + m)]^{-\frac{1}{2}}. \tag{1.7}$$

The infinitesimal operators of the Poincaré group become (the subscript Q denoting the new representation)

$$\begin{aligned} p_Q &= p, \\ \mathbf{M}_Q &= \mathbf{x}\times\mathbf{p} + \tfrac{1}{2}\boldsymbol{\sigma}, \\ \mathbf{N}_Q &= (x^0\mathbf{p} - \mathbf{x}p^0) + \tfrac{1}{2}[\boldsymbol{\sigma}\times\mathbf{p}/(p^0 + m)]. \end{aligned} \tag{1.8}$$

It will be seen at once that the expressions in (1.8) are simpler than those in (1.4) in the sense that β does not occur in our case. Since we have gotten rid of both $\boldsymbol{\alpha}$ and β, replacing $\frac{1}{2}\boldsymbol{\sigma}$ by $\mathbf{S}$, the general spin matrix, we can directly generalize the above formulas for the case of arbitrary spin as

$$\begin{aligned} \mathbf{M} &= \mathbf{x}\times\mathbf{p} + \mathbf{S}, \\ \mathbf{N} &= (x^0\mathbf{p} - \mathbf{x}p^0) + \mathbf{s}\times\mathbf{p}/(p^0 + m), \end{aligned} \tag{1.9}$$

where

$$[s_i, s_j] = i\epsilon_{ijk}S_k.$$

The above form of $\mathbf{M}$ and $\mathbf{N}$ imply that taking $\mathbf{S}$ and β as in (1.4') and (1.4'') amounts simply to treating the upper and lower components of the wavefunction simultaneously, though they transform in an identical manner. So that we might as well

[3] The operator $U(\Lambda)$ for a homogeneous Lorentz transformation can be written as $U(\Lambda) = Q(\Lambda)\ T(\Lambda)$ where

$$(T(\Lambda)\varphi)(p) = \varphi(\Lambda^{-1}\ p),$$

and

$$(Q(\Lambda)\varphi)(p) = Q(p,\ \Lambda)\varphi(p).$$

By their very definition and construction, a transformation by $Q(\Lambda)\ T(\Lambda)$ leaves Eq. (1.6) form-invariant. On the other hand, transforming by $Q(\Lambda)$ alone, since the momentum values are left undisturbed, we get, instead of a change of the frame of reference, a change of the representation, leading to a possible change of the form of the equation.

In particular, the Q of (1.7) leading to Eq. (1.5) can be written as

$$Q(p,\ \Lambda_{k\leftarrow p})_{\text{op}}, \text{ where } \Lambda p_{k\leftarrow p} = K \equiv (\kappa, 0, 0, 0),$$

so that

$$Q_{\text{op}}\ \varphi(p') = Q(p',\ \Lambda_{k\leftarrow p'})\varphi(p').$$

get rid of γ^0 in (1.5) and thus reduce the components by half.

Indeed, since in the momentum space the solutions of (1.5) belong, as may easily be seen, to the eigenvalues ± 1 of γ^0 according as $\epsilon = \pm 1$, we can write, separating the cases of positive and negative energy, and finally take as the canonical form,

$$(\epsilon m - \kappa)\varphi = 0, \tag{1.10}$$

or

$$\{(p^\nu p_\nu)^{\frac{1}{2}} - \kappa\}\varphi = 0, \tag{1.11}$$

the form of the equation being the same for positive and negative energy. After this separation, **S** can directly be taken as the $(2s + 1) \times (2s + 1)$ spin matrix instead of as the direct sum of two such matrices.

In fact it is just because (1.1) involve β, that Foldy,[1] after starting with the canonical form

$$[i(\partial/\partial t) - \omega]\varphi = 0,$$

introduced β in this equation and in the expressions of the corresponding infinitesimal operators [compare his equations (B25–28) and (D61–67)] through the additional hypothesis of invariance under "Pauli" (unitary) time inversion. This was needed to make the correspondence with $\mathbf{N}_F$, $\mathbf{M}_F$ exact. In our case no such artifice is necessary.

Our representation is valid in the general case in the sense that **M**, **N** as given in (1.9), along with p, satisfy the fundamental commutation relations of the Poincaré group, and also the canonical equation is form-invariant under the group transformations (the verification of this last point being trivial in our case since the p^μ's appear in the combination $p^\nu p_\nu$). The transformation (1.7) is intended to show that for $S = \frac{1}{2}$ our scheme is indeed equivalent to the standard Dirac equation. We have confined ourselves to the case $\kappa > 0$, since Q is not well-defined for $\kappa = 0$.

II. TRANSFORMATION OF THE K–G EQUATION

In order to bring the K–G equation to his canonical form Foldy[1] introduced the transformation

$$\chi = \begin{bmatrix} 1/\sqrt{2}\ [\omega^{-\frac{1}{2}}(\partial\varphi/\partial t) - i\omega^{\frac{1}{2}}\varphi] \\ 1/\sqrt{2}\ [\omega^{-\frac{1}{2}}(\partial\varphi/\partial t) + i\omega^{\frac{1}{2}}\varphi] \end{bmatrix}, \tag{2.1}$$

where

$$(\omega^2 - \partial^2/\partial t^2)\varphi = 0,$$

and then verified at some length that, for χ, which satisfies

$$[i(\partial/\partial t) - \beta\omega]\chi = 0,$$

and the infinitesimal operators of the Poincaré group are given by (1.4) where $\frac{1}{2}\delta$ is now to be replaced by *null* 2×2 matrices, corresponding to zero spin.

In our case the problem is very much simpler. Putting

$$\chi = C \begin{vmatrix} (m + \kappa)\varphi \\ (m - \kappa)\varphi \end{vmatrix} \tag{2.2}$$

(where C is a normalization constant), we find that χ satisfies (1.5), and if we take (1.11) as the standard form we have only to put

$$\chi = C'(\epsilon m + \kappa)\varphi. \tag{2.3}$$

In our case, the problem of the verification of the transformation properties reduce to a trivial one.

We will not consider the Proca equation as it is of lesser interest.

III. COMPARISON WITH SHIROKOV'S FORMULATION

It will be noted that, apart from some changes of notations and conventions, the infinitesimal operators (1.9) coincide with those choosen by Iu. M. Shirokov[4] in his study of the Poincaré group. It is sometimes said that Shirokov's formulation corresponds to Foldy's canonical form of the equations. Thus M. I. Shirokov[5] speaks of the Foldy–Iu. Shirokov representation. In our opinion, though Foldy's representation has closely analogous properties, it is more justifiable to associate Iu. Shirokov's choice of operators with the canonical form (1.5) [or (1.11)]. The corresponding transformation, for the Dirac equation is given by (1.7).

The parallel features of the F–W and our transformation will be compared in detail in Appendix B. Here we analyze further the connection between our representation and Iu. Shirokov's choice of operators.

The connection of course, is not far to seek and depends on the construction of the 4-vector Γ (which we will denote by W following Bargmann and Wigner) adopted by Shirokov.[4]

In the rest frame (for $\kappa > 0$) W is supposed to coincide with the 4-vector $m(0, \mathbf{S})$, and for the laboratory frame, W is obtained by the corresponding Lorentz transformation for vectors being applied to $m(0, \mathbf{S})$.

Now it may be verified that (see Appendix B)

$$W_Q = QWQ^{-1} = \Lambda^{-1}_{k\leftarrow p} m(0, \tfrac{1}{2}\delta), \tag{3.1}$$

which is indeed a direct consequence of the fundamental transformation property of the γ's imposed to assure form invariance of the Dirac equation.

From (3.1) we see that defining W after Shirokov[4]

[4] Iu. M. Shirokov, Zh. Eksperim. i. Teor. Fiz., **33**, 1196 (1957) [English transl.: Soviet Phys.—JETP **6**, 919 (1958)].

[5] M. I. Shirokov, Zh. Eksperim. i. Teor. Fiz., **39**, 633 (1960) [English transl.: Soviet Phys.—JETP **12**, 445 (1961)].

is equivalent to adopting our representation and the cononical form of the equation proposed.

Iu. Shirokov derives the explicit expressions for **M**, **N** through those for the operators defined (in a slightly different notation) as

$$W = M^* \cdot p, \qquad g = M \cdot p, \tag{3.2}$$

in terms of which M can be expressed as

$$M = \{(g \wedge p) - (W \wedge p)^*\}(m^2)^{-1}. \tag{3.3}$$

W is expressed in terms of p and **S** in the way described above. As for g, Shirokov simply writes down an expression [Eq. (10), reference 3], stating that it may be verified that such a g satisfies the required commutation relations (which have been evaluated through those satisfied by p and M).

In our representation we have

$$M_Q = (x \wedge p) + [\mathbf{S} \times \mathbf{p}/(p^0 + m), \mathbf{S}]$$
$$(\mathbf{S} = \tfrac{1}{2}\boldsymbol{\sigma} \quad \text{for} \quad S = \tfrac{1}{2}),$$

and hence,

$$W_Q = M_Q^* \cdot p = [\mathbf{S} \cdot \mathbf{p},\ m\mathbf{S} + (\mathbf{S} \cdot \mathbf{p})\mathbf{p}/(p^0 + m)] \tag{3.4}$$

and

$$g_Q = M_Q \cdot p = (x \wedge p) \cdot p + [0,\ m/(p^0 + m)\mathbf{S} \times \mathbf{p}]$$
$$= x(p^2) - (x \cdot p)p + [0,\ m/(p^0 + m)\mathbf{S} \times \mathbf{p}].$$

Thus we see that whereas W_Q coincides with Shirokov's Γ, g_Q does not entirely coincide with g.

Apart from questions of conventions adopted, the essential factor $m/(p^0 + m)$ (which arises in the F–W representation also) is lacking in Shirokov's definition. It seems that there a mistake (or a misprint) here and indeed (3.4) gives the correct form of g required to ensure the proper commutation relations, leading to the correct expressions for **M** and **N**.

IV. TRANSFORMATION PROPERTIES IN THE NEW REPRESENTATION

In this section we discuss the transformation properties of the wavefunction and the polarization operator, in our representations, following closely the notation of Wightman.[6]

Under the inhomogeneous Larentz transformation (a, Λ) the wavefunction transforms as follows:

$$(U(a, \Lambda)\varphi)(p) = \exp ip \cdot a Q(p, \Lambda)\varphi(\Lambda^{-1}p), \tag{4.1}$$

where the matrices Q satisfy

$$Q(p, \Lambda_1)Q(\Lambda_1^{-1}p, \Lambda_2) = Q(p, \Lambda_1\Lambda_2). \tag{4.2}$$

$Q(p, \Lambda)$ can be constructed in terms of the representation of the little group of any one point k of the orbit as follows:

$$Q(p, \Lambda) = Q(k, \Lambda_{p\leftarrow k}^{-1})^{-1} \times Q(k, \Lambda_{p\leftarrow k}^{-1}\Lambda\Lambda_{p'\leftarrow k})Q(k, \Lambda_{p'\leftarrow k}^{-1}) \tag{4.3}$$

(where $\Lambda p' = p$), as may be verified with the help of (4.2) and one of its consequences, namely

$$Q(k, \Lambda_{p\leftarrow k}^{-1})^{-1} = Q(p, \Lambda_{p\leftarrow k}).$$

Since

$$(\Lambda_{p\leftarrow k}^{-1}\Lambda\Lambda_{p'\leftarrow k})k = k,$$

$Q(k, \Lambda_{p\leftarrow k}^{-1}\Lambda\Lambda_{p'\leftarrow k})$ belongs to the representation of the little group of k.

In our case ($\kappa > 0$) we choose $k = (\kappa, 0, 0, 0)$ when

$$Q(k, \Lambda_{p\leftarrow k}^{-1}\Lambda\Lambda_{p'\leftarrow k}) = D^{(s)}(\Lambda_{p\leftarrow k}^{-1}\Lambda\Lambda_{p'\leftarrow k}),$$

$D^{(s)}$ being the standard spin matrix for spin S and $(\Lambda_{p\leftarrow k}^{-1}\Lambda\Lambda_{p'\leftarrow k})$ representing a purely three-dimensional spatial rotation.

Wightman remarks[6] that while (4.3) is a consequence of (4.2), it may as well be taken as the definition of $Q(p, \Lambda)$ [i.e., $Q(\Lambda)$ defined by (4.3) automatically satisfies (4.2) if we choose some suitable continuous function of p for $Q(k, \Lambda_{p\leftarrow k})$].

Wightman puts

$$Q(k, \Lambda_{p\leftarrow k}) = 1 \tag{4.4}$$

and obtains

$$\varphi'(p) \equiv (U(a, \Lambda)\varphi)(p) = \exp ip \cdot a D^{(s)}(\Lambda_{p\leftarrow k}^{-1}\Lambda\Lambda_{p'\leftarrow k})\varphi(p'), \tag{4.5}$$

where

$$\Lambda p' = p.$$

We know, however, that for the usual representation of the Dirac equation, (4.4) does not hold [in fact $Q(k, \Lambda_{p\leftarrow k})$ is the inverse of the matrix (1.7)], and so neither does the simple transformation formula (4.5).

In our representation, however, since

$$\varphi_Q(p'') = Q(p'', \Lambda_{k\leftarrow p''})\varphi(p''),$$

we have

$$\varphi_Q'(p) = \exp ip \cdot a[Q(p, \Lambda_{k\leftarrow p})Q(p, \Lambda) \times Q(p', \Lambda_{k\leftarrow p'})^{-1}]\varphi_Q(p') = \exp ip \cdot a Q(k, \Lambda_{p\leftarrow k}^{-1}\Lambda\Lambda_{p'\leftarrow k})\varphi_Q(p'), \tag{4.6}$$

or

$$\varphi_Q'(p) = \exp ip \cdot a D^{(s)}(\Lambda_{p\leftarrow k}^{-1}\Lambda\Lambda_{p'\leftarrow k})\varphi_Q(p'). \tag{4.7}$$

Thus we see that (for $S = \frac{1}{2}$) passing from the Dirac to the Wigner[7] representation is tantamount to performing the similarity transformation proposed and taking the equation instead of in the usual form

[6] A. S. Wightman, "L'invariance dans la mécanique relativiste" in *Lectures, Ecole d'été de Physique Théorique, Les Houches* (Hermann & Cie, Paris, 1960).

[7] E. P. Wigner, Ann. Math. 40, 149 (1939).

$$(\gamma \cdot p - \kappa)\varphi = 0,$$

in the proposed canonical form

$$(\gamma^0 m - \kappa)\varphi = 0.$$

In fact, as is evident, the argument is quite general. For any spin S, starting from a representation (which may not be unitary) we can always arrive at Wigner's unitary representation through a transformation by $Q(p, \Lambda_{k\leftarrow p})$. This has been noted by Shirokov.[5]

In the case of the K–G equation, though there was no problem about Q, we had to transform φ by the operator $(m + \kappa)$ to obtain the canonical form. But since the p^μ's appear in this operator in a Lorentz-invariant combination, no difficulty is introduced. In fact, the general result is that we can transform by the operator

$$f(p^2, W^2)Q(p, \Lambda_{k\leftarrow p})$$

(f being a function of the invariant p^2, W^2 of the group) to obtain the desired transformation property. Such a factor f, however, affects the position operators which will be discussed separately.

Let us now consider the polarization operator. Putting

$$\mathbf{\Sigma}_Q = Q^{-1}\boldsymbol{\sigma}Q,$$

we have

$$\Lambda_{k\leftarrow p}(W/m) = \tfrac{1}{2}(0, \mathbf{\Sigma}_Q) \qquad (4.8)$$

(see Appendix B),

which implies that the square of the 4-vector $(0, \mathbf{\Sigma})$, i.e. $\mathbf{\Sigma}^2 (= W^2)$, is Lorentz-invariant. This gives a suitable prescription for a relatavistic description of the spin with the help of a 3-vector $\mathbf{\Sigma}$, (which replaces $\boldsymbol{\sigma}$ in the Dirac representation and becomes just $\boldsymbol{\sigma}$ in ours).

Since under a Lorentz transformation

$$W'(p) = \Lambda W(p') \quad \text{where} \quad \Lambda p' = p,$$

we have

$$(0, \mathbf{\Sigma}') = \Lambda_{k\leftarrow p}\Lambda\Lambda^{-1}_{k\leftarrow p'}(0, \mathbf{\Sigma}), \qquad (4.9)$$

or, since Q commutes with Λ's, we have in our representation

$$\boldsymbol{\sigma}' = (\Lambda^{-1}_{p\leftarrow k}\Lambda\Lambda_{p'\leftarrow k})\boldsymbol{\sigma}, \qquad (4.10)$$

i.e., under a Lorentz transformation the spin is rotated through the angle that appears in Wigner's representation in terms of the little group. This is the angle "relativistic rotation of the spin" and has been studied by several authors,[8,9] who have given its explicit expression in terms of p and p'.

The conceptual difference between the treatment of those authors and ours should be noted, though both lead to the same result.

In the case of the authors mentioned[5,8,9,10] the spin state is measured in the reference system in which the particle is at rest, i.e., S is applied, not to φ, but to

$$Q(\Lambda)T(\Lambda)\varphi \qquad (\Lambda \equiv \Lambda_{k\leftarrow p}),$$

and then we come back again to the laboratory system.

In our case we apply S simply to

$$Q(\Lambda)\varphi.$$

Thus in one case there is a change of the reference system in the same representation, and in the other there is a change of representation in the same reference system (cf. reference 3).

V. APPLICATIONS TO SCATTERING

The motivation of the examination of the proper definition and transformation properties of the spin operator by the above mentioned authors[7,8,9] has often been the analysis of the scattering problems of the type

$$a + b \rightarrow c + d.$$

This involves effectively the decomposition into its irreducible components of the direct-product representation of the Poincaré group (leading to a maximum possible diagonalization of the S matrix) and the calculation of the relativistic generalization of the Clebsch–Gordan coefficients.

In Appendix A we give a brief outline of this process using our representation. We have chosen to follow closely the method of Jacob and Wick.[11,12,13] Wick's treatment[12] is valid in any representation, not being dependent on the form of the equation satisfied. The special features of our representation appear only if we try to write the states explicitly, since Wick's notation defining the transformed states conceals the complications that arise in the case of the usual representation of the Dirac equation.

In defining [Eq. (A.2), reference 12]

$$|p, \lambda\rangle = R_{\varphi\vartheta 0}Z\,|p^0, \lambda\rangle,$$

we should note that in our case transformation by Z does not imply any multiplication by a matrix

[8] H. P. Stapp, Phys. Rev. **103**, 425 (1956).
[9] Chou Kuang-Chao and M. I. Shirokov, Zh. Eksperim. i. Teor. Fiz. **34**, 1230 (1958) [English transl.: Soviet Phys.—JETP **7**, 851 (1958)].
[10] Iu. M. Shirokov, Zh. Eksperim. i. Teor. Fiz. **35**, 1005 (1958) [English transl.: Soviet Phys.—JETP **8**, 703 (1959)].
[11] M. Jacob and G. C. Wick, Ann. Phys. (N. Y.) **7**, 404 (1959).
[12] G. C. Wick, Ann. Phys. (N. Y.) **18**, 65 (1962).
[13] J. S. Lomont, J. Math. Phys. **1**, 237 (1960).

as in the Dirac representation. If the states on both sides of Eqs. (4) and (A.4) of reference 12 are written explicitly, the extra matrices corresponding to the $Q(k, \Lambda_{p\leftarrow k}^{-1})$ of our Eq. (4.3) would make their appearance in the Dirac representation and consequently also in Eq. (15) of reference 12 which corresponds to our Eq. (A7).

ACKNOWLEDGMENTS

I am indebted to Professor L. Michel for guidance and encouragement. Also, it is a pleasure to acknowledge the many profitable discussions I had with Dr. H. Bacry. In his thesis (presented to Faculté des Sciences, Marseille) some considerations on the transformation we have proposed for the Dirac equation and the related position operator discussed by us in the following paper, have already appeared (see his Appendix D. 2.—"Transformation de Chakrabarti de l'equation de Dirac").

APPENDIX A

The infinitesimal operators of the product representation are [for component particles (κ', s'), (κ'', s'')]

$$P = P' + P'', \qquad M = M' + M''. \tag{A1}$$

The invariants are, as usual,

$$P^2,\ W^2 \text{ (where } W = M^* \cdot p).$$

We have

$$P^2 = P'^2 + P''^2 + 2P' \cdot P'', \tag{A2}$$

giving for the mass spectrum[13] the range

$$\kappa' + \kappa'' \leq \kappa < \infty,$$

also, since

$$W = (\mathbf{M} \cdot \mathbf{P},\ \mathbf{M}P^0 + \mathbf{N} \times \mathbf{P}), \tag{A3}$$

where

$$\mathbf{M} = \mathbf{M}' + \mathbf{M}'', \qquad \mathbf{N} = \mathbf{N}' + \mathbf{N}''$$

in the reference system in which the total momentum vanishes, i.e.,

$$\begin{aligned} \mathbf{P} = \mathbf{P}' + \mathbf{P}'' = 0 \quad \text{or} \quad \mathbf{P}' = -\mathbf{P}'' = \mathbf{p}_{(0)}, \quad \text{say,} \\ W &= \kappa(0, \mathbf{M}'_{(0)} + \mathbf{M}''_{(0)}) \\ &= \kappa[0, (\mathbf{x}' - \mathbf{x}'') \times \mathbf{p}_{(0)} + \mathbf{s}' + \mathbf{s}''] \\ &= \kappa(0, \mathbf{l} + \mathbf{s}' + \mathbf{s}''), \end{aligned} \tag{A4}$$

where $\mathbf{l}$ is the total orbital angular momentum of the particles about some origin (say the center of mass) in this frame.

Thus the "spin" of the composite system (defined as the intrinsic angular momentum which remains when the total momentum vanishes and whose square coincides with the invariant W^2) has no upper limit due to the term $\mathbf{l}$.

In Wightman's notation,[6] the direct product can be expressed as the direct integral, as

$$[\kappa', s'] \otimes [\kappa'', s''] \cong \int_{\kappa'+\kappa''}^{\oplus} d\kappa \bigoplus_{l=0}^{\infty} \bigoplus_{\xi=|s'-s''|}^{s'+s''} \bigoplus_{J=|\xi-l|}^{\xi+l} [\kappa, J]. \tag{A5}$$

The result follows from the usual rules for composition of angular momenta (for details see reference 6, as also for the modifications necessary when the two particles are identical).

Now labeling the intrinsic angular-momentum states by J and m_i, we can write (in the frame $\mathbf{p} = 0$), the eigenstates as[12] (leaving aside all questions of proper choice of phases and normalizations in our case)

$$\varphi(k, J, m_J; \lambda', \lambda'') \cong \int d\Omega D^{(J)}_{m_i\lambda}(r^{-1}) R \times [\varphi(p'_{(0)}, \lambda') \otimes \varphi(p''_{(0)}, \lambda'')], \tag{A6}$$

where $p'_{(0)}$, $p'_{(0)}$ are the particle momenta [with $\mathbf{p}'_{(0)} = -\mathbf{p}''_{(0)}$ pointing in the Z direction] λ', λ'' their helicities, $\lambda = \lambda' - \lambda''$, $K = (\kappa, 0, 0, 0)$, r is the rotation with Euler angles $\varphi\vartheta\gamma$, $R = R_{\varphi\vartheta\gamma}$ is the corresponding operator, and $d\Omega = \sin\theta\, d\,\theta d\varphi$ can be converted to $dp'\, dp'\, \delta_4(p' + p'' - p')$ with a suitable change of normalization.

The above result of course follows from the standard method of decomposition of the direct-product representations of the rotation group.

Now let us apply a proper Lorentz transformation Λ to both sides according to Wick's convention, i.e., let

$$\Lambda = R_{\varphi\vartheta 0} Z,$$

where Z is a pure Lorentz transformation along the z axis and is followed by a pure rotation R.

This conserves the helicities and we have

$$\varphi(P, J, m_J; \lambda', \lambda'') \cong \int d\Omega D^{(J)}_{m_J\lambda} \sum_{\nu', \nu''} D^{(s')}_{\lambda'\nu'}(r') D^{(s'')}_{\lambda''\nu''}(r'') \times [\varphi(p', \nu') \otimes \varphi(p'', \nu'')], \tag{A7}$$

where

$$\varphi(P) = D^{(J)}(R_{\varphi\vartheta 0})\varphi(\Lambda^{-1}k),$$

and

$$r' = (\Lambda^{-1}_{p_{(0)}'\leftarrow k'}\Lambda'\Lambda_{p'\leftarrow k'}); \quad \Lambda' = \Lambda R, \quad \Lambda' p' = p'_{(0)},$$
$$r'' = (\Lambda^{-1}_{p_{(0)}''\leftarrow k''}\Lambda'\Lambda_{p''\leftarrow k''}); \quad \Lambda' p'' = p''_{(0)}.$$

This gives us the required formula for the wavefunctions providing the carrier spaces of the irreducible components into which the direct-product representation may be decomposed. The product of the three D's on the right (with the proper normalization factor) gives the relativistic Clebsch–Gordan coefficients.

APPENDIX B

In this appendix some features of our transformation (1.7) are compared with the corresponding ones of the well-known F–W transformation.[2]

The transforming operator of F–W is

$$F = (\beta H + \omega)/[2\omega(\omega + \kappa)]^{\frac{1}{2}} \tag{B1}$$

where

$$\omega = (\mathbf{p}^2 + \kappa^2)^{\frac{1}{2}} \quad \text{and} \quad H = \beta x + \boldsymbol{\alpha}\cdot\mathbf{p}.$$

While in our case the transforming operator is just the matrix Q of the Lorentz transformation operator

$$Q(\Lambda_{k\leftarrow p})T(\Lambda_{k\leftarrow p}), \tag{B2}$$

it has been noted by several authors,[14,15] that F is also closely related to Q.

Good and Rose[15] demonstrate that

$$\psi_F = F\psi = (\omega/\kappa)^{\frac{1}{2}}(Q\psi) = (\omega/\kappa)^{\frac{1}{2}}\psi_Q \tag{B3}$$

(in reference 16 $\kappa = 1$), with the help of the following devices:

Writing Q as

$$Q' = [2\omega(\omega + \kappa)]^{-\frac{1}{2}}[\omega + \kappa - \boldsymbol{\alpha}\cdot\mathbf{p}\epsilon], \tag{B4}$$

which involves, as compared to (1.7), replacing the operators p^0 and m by $\epsilon\omega$ and $\epsilon\kappa$, (with $\epsilon = p^0/|p^0| = \pm 1$), they put

$$-\boldsymbol{\alpha}\cdot\mathbf{p}\epsilon = -\boldsymbol{\alpha}\cdot\mathbf{p}\,\frac{p^0}{|p^0|} = -\boldsymbol{\alpha}\cdot\mathbf{p}\,\frac{H}{\omega} = \frac{\mathbf{p}^2 + \kappa\beta\boldsymbol{\alpha}\cdot\mathbf{p}}{\omega}. \tag{B5}$$

The replacement of $p^0[\equiv i(\partial/\partial t)]$ by H is justified by the argument that p^0 is being applied directly on the wavefunction.

This gives

$$Q'\psi \simeq (\kappa/\omega)^{\frac{1}{2}}F\psi. \tag{B6}$$

[We shall use the sign $\simeq$ to distinguish the equalities obtained by replacing certain operators by some other with the help of the equation satisfied by ψ. The above replacement and hence the equivalence $Q' \simeq (\kappa/\omega)^{\frac{1}{2}}F$ is not valid when applied to transform other operators and not ψ.]

We note that

$$\psi_F = (\omega/\kappa)^{\frac{1}{2}}\psi_Q$$

does not have the desirable simple transformation property indicated in (4.7), since the factor $\omega^{\frac{1}{2}}$ is not Lorentz-invariant.

Let us now compare another aspect of the two transformations. We have

$$\begin{aligned}
&\psi_F^*\psi_F = \psi^*\psi; \qquad \bar{\psi}_F\psi_F = \psi^*\frac{H}{\omega}\psi \\
&\qquad\qquad \simeq \psi^*\frac{p^0}{\omega}\psi = \psi_F^*\frac{p^0}{\omega}\psi_F, \\
&\bar{\psi}_Q\psi_Q = \bar{\psi}\psi; \qquad \psi_Q^*\psi_Q = \bar{\psi}\frac{(\gamma\cdot p)}{m}\psi \\
&\qquad\qquad \simeq \bar{\psi}\frac{\kappa}{m}\psi = \bar{\psi}\frac{m}{\kappa}\psi = \bar{\psi}_Q\frac{m}{\kappa}\psi_Q.
\end{aligned} \tag{B7}$$

As is well known,

$$(\gamma\cdot p - \kappa)\psi = 0$$

leads to

$$\psi^*\psi = \bar{\psi}(p^0/\kappa)\psi = \pm(\omega/\kappa)\bar{\psi}\psi.$$

Similarly the equations

$$\bar{\psi}_F\psi_F = \psi_F^*(p^0/\omega)\psi_F = \pm\psi_F^*\psi_F,$$

and

$$\overset{*}{\psi}_Q\psi_Q = \bar{\psi}_Q\psi(m/\kappa)\psi_Q = \pm\bar{\psi}_Q\psi_Q$$

of (B7) also follow from the respective transformed equations.

Thus we find that Q, indeed from its very definition, preserves the form of the Lorentz-invariant scalar product $\bar{\psi}\psi$, which F fails to do.

But on the other hand, as is well-known, Q is not unitary in the ordinary sense (i.e. does not leave invariant the scalar product $\psi^*\psi$ as does F) but is "unitary within the scalar product $\bar{\psi}\psi$".

For an observable whose matrix elements are defined by $\psi^*0\psi$, we have only to take, as the operator corresponding to the observable as $0' = \gamma^0 0$ and write the matrix elements as

$$\bar{\psi}0'\psi \qquad (=\psi^*0\psi),$$

which will be invariant under the transformation

$$0'_{tr} = Q0'Q^{-1},$$

[14] C. G. Bollini and J. J. Girambiagi, Nuovo Cimento **21**, 107 (1961).

[15] R. H. Good, Jr. and M. E. Rose, Nuovo Cimento **22**, 864 (1962).

since

$$\psi_Q^* \gamma^0 = \psi^* \gamma^0 Q^{-1},$$

the matrix y^0 determining the metric being of course left undisturbed. Also $0'$ will be Hermitian within this scalar product, since $0'^* \gamma^0 = \gamma^0 0'$ is a consequence of $0^* = 0$.

Furthermore, since from (B7) and (B8) we have

$$\psi_Q^* \psi_Q \simeq (\kappa/\omega) \psi_F^* \psi_F \qquad [\pm \psi_Q^* \psi_Q \simeq \bar{\psi}_Q \psi_Q = \bar{\psi}\psi \simeq \pm(\kappa/\omega)\psi^*\psi = \pm(\kappa/\omega)\psi_F^* \psi_F], \tag{B9}$$

comparing (B6) and (B9) we conclude that the factor $(\omega/\kappa)^{\frac{1}{2}}$ arises out of the respective behaviors of F and Q with respect to the two equivalent scalar products for the Dirac equation ($\kappa > 0$), namely

$$\psi\bar{\psi} \quad \text{and} \quad (\kappa/p^0)\psi^*\psi.$$

This seems more reasonable than to suggest, as do Good and Rose,[15] that the factor is necessitated by volume contraction on passing from the rest to the laboratory system, since a transformation by Q (just like that by F) does not imply a change of frame of reference (see reference 3).

In the Dirac representation, we have

$$W = M^* \cdot p = \tfrac{1}{2}(\mathfrak{d}\cdot\mathbf{p},\ \mathfrak{d}p^0 + i\boldsymbol{\alpha}\times\mathbf{p});$$

also,

$$\boldsymbol{\Sigma}_F = F^{-1}\mathfrak{d}\mathbf{F} = \frac{\kappa}{\omega}\mathfrak{d} + \frac{i\beta\boldsymbol{\alpha}\times\mathbf{p}}{\omega} + \frac{(\mathfrak{d}\cdot\mathbf{p})\mathbf{p}}{\omega(\omega+\kappa)},$$
$$\boldsymbol{\Sigma}_Q = Q^{-1}\mathfrak{d}Q = \frac{p^0}{m}\mathfrak{d} + \frac{i\boldsymbol{\alpha}\times\mathbf{p}}{m} - \frac{(\mathfrak{d}\cdot\mathbf{p})\mathbf{p}}{m(p^0+m)}. \tag{B10}$$

We find that since

$$\Lambda_{k\leftarrow p} W = [0,\ \mathbf{W} - (W^0/p^0 + m)\mathbf{p}] \tag{B11}$$

[where in applying

$$\Lambda_{k\leftarrow p} = \Lambda_{t\leftarrow u} = 1 - (t+u)\otimes(t+u)/(1+t\cdot u) + 2t\otimes u,$$

with

$$t = (k/\kappa) = (1,\mathbf{0}), \qquad u = p/m,$$

we have used the condition $W\cdot p = 0$], we have identically

$$\Lambda_{k\leftarrow p} W = \tfrac{1}{2}m(0,\ \boldsymbol{\Sigma}_Q),$$
$$Q(\Lambda_{k\leftarrow p}W)Q^{-1} = (\Lambda_{k\leftarrow p}W)_Q = \tfrac{1}{2}m(0,\ \mathfrak{d}), \tag{B12}$$

as has already been noted in (3.1), and this leads to the transformation property of $\mathfrak{d}$ in our representation given by (4.10).

As is well-known, in the F–W representation it is $\beta\mathfrak{d}$, rather than $\mathfrak{d}$ which has a parallel connection with $(\Lambda_{k\leftarrow p}W)$.[16]

But since in both the representations

$$\mathbf{M} = \mathbf{x}\times\mathbf{p} + \tfrac{1}{2}\mathfrak{d},$$

the direct connection of $\mathfrak{d}$ with the Wigner 4-vector polarization operator which we obtain in our case, seems to be more desirable.

It is true, however, that whereas

$$H_F = FHF^{-1} = \beta\omega$$

leads to

$$[H_F,\ \mathfrak{d}]_- = 0,$$
$$H_Q = QHQ^{-1} = (\boldsymbol{\alpha}\cdot\mathbf{p})[1 - \beta\kappa/m] + p^0(\beta\kappa/m) \tag{B13}$$

leads to

$$[H_Q,\ \mathfrak{d}]_- = 2i(\mathfrak{d}\times\mathbf{p})[1 - \beta\kappa/m].$$

But since

$$(\beta m - \kappa)\psi_Q = 0,$$

we have

$$[H_Q,\ \mathfrak{d}]_-\psi_Q = 0. \tag{B14}$$

Thus, the matrix elements of $\mathfrak{d}$ vanish in our representation also, and we may after all consider our $\boldsymbol{\delta}$ as a suitable polarization operator in this respect also, with

$$[H_Q,\ \mathfrak{d}] \simeq 0. \tag{B15}$$

Another appealing feature of our transformation is that since it leaves γ^5 invariant, separation of ψ into two chirality states by means of the projection operators $\frac{1}{2}(1 \pm i\gamma^5)$ is unaffected. The F–W transformation does not have this property.

Finally we just mention that the simple nonlinear spinor equation at first proposed by Heisenberg[17] is brought by our transformation to the form

$$[\gamma^0 m - (\bar{\psi}\psi)]\psi = 0, \tag{B16}$$

which again can be separated, as usual, into two equations, thus eliminating all γ matrices.

The discussion of the position operators will be taken up elsewhere.

[16] D. M. Fradkin and R. H. Good, Jr., Rev. Mod. Phys. **33**, 343 (1961).

[17] W. Heisenberg, Z. Naturforsch. **9a**, 292 (1954).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 4, NUMBER 10 OCTOBER 1963

Relativistic Position Operator for Free Particles

A. CHAKRABARTI
Centre de Physique Théorique de l'Ecole Polytechnique, Paris, France
(Received 21 March 1963)

Certain definitions and derivations for the relativistic center of mass are considered and are related to the position operator corresponding to the representation proposed in a previous paper. It is shown that the classical definition of Pryce has indeed a more direct correspondence with our representation than the Foldy–Wouthuysen representation, in spite of the fact that Pryce's particular method of symmetrization leads to the latter. In the process of derivation, the points of view of Synge, Pryce, and Shirokov are compared and the expressions for the velocity and the intrinsic angular momentum corresponding to the different definitions are given. In the Appendix, our operator is compared with that of Newton and Wigner for the case of spin $\frac{1}{2}$.

1. INTRODUCTION

In a previous paper[1] we considered a certain transformation of the Dirac equation leading to the form

$$(\gamma^0 m - \kappa)\varphi = 0, \tag{1.1}$$

where κ is the rest mass and $m = (p^0/|p^0|)(p^r p_r)^{\frac{1}{2}}$. The transformation properties of the wavefunction and of the corresponding polarization operator were discussed in the momentum space. Here, we intend to discuss the corresponding position operator.

In what follows we will give a general expression, valid for arbitrary spin, which reduces for spin $\frac{1}{2}$ to the operator

$$X = Q^{-1}xQ \tag{1.2}$$

(Q having the meaning ascribed to it in reference 1), which is transformed to x in our representation.

It is known from the discussion of the definition of the relativistic "center of mass" by various authors,[2,3] that when the conception of a particle is generalized by attributing to them "spin" (intrinsic angular momentum), a suitable definition of the centroid leads, even in the case of a single particle, to an expression different from the usual x. We thus arrive at a new definition of the position of a particle. This relatively complicated definition is preferable in the sense that the expression for velocity and that for the "spin part" of the total angular momentum tensor corresponding to the decomposition

$$M = -p \wedge X + S_x$$

has certain desirable properties.

Pryce[2] derived a classical expression for the "center" from his definitions and showed that when the passage to quantum mechanics is made in a certain way (i.e., adopting certain processes of symmetrization and so on) we arrive at a position operator which, for the spin-$\frac{1}{2}$ case, corresponds to a change of representation of the usual Dirac equation. This operator is nowadays usually associated with the Foldy–Wouthuysen (F–W) transformation.[4] The "self-reciprocal" transformation operator actually given by Pryce is, however, somewhat different. It gives the same result as the F–W transformation for any operator commuting with β, but not for H (the Dirac Hamiltonian) for example.

We will show that our transformation[1] corresponds even more directly to the classical formula of Pryce. We will also rederive it in a somewhat different manner along with an alternative form.

Our method will be to start with explicitly accepting the existence of the spin part of the total angular momentum tensor even for the classical case, as has been done, for example, by Synge.[3]

The terms will be ordered in such a way that no difficulty will arise when we will consider them as quantum mechanical operators. For spin $\frac{1}{2}$, for example, we will be able to consider the case of a Dirac particle merely by substituting the well-known operators for the spin tensor and, of course, regarding x and p in the operator sense. The commutation difficulties will not induce any further change.

In the following we use the same tensorial notation as in reference 1.

2. DERIVATION OF THE EXPRESSIONS FOR THE "CENTERS"

Synge[3] introduces the center (which we will denote by q') by imposing the condition that Sq', given by

[1] A. Chakrabarti, J. Math. Phys. **4**, 1215 (1963).
[2] M. H. L. Pryce, Proc. Roy. Soc. (London), **A195**, 62 (1948).
[3] J. L. Synge, *Special Relativity* (North-Holland Publishing Company, Amsterdam, 1956).
[4] L. L. Foldy and S. A. Wouthuysen, Phys. Rev. **78**, 29 (1950).

$$M = -p \wedge q' + S_{q'} \tag{2.1}$$

should satisfy

$$p \cdot S_{q'} = 0.$$

He derives an expression (noncovariant) for q' by considering the rest frame only and shows that this coincides with that derived by Pryce,[2] starting from his definition

$$\mathbf{q}_{\mathrm{Pr}} = \int T^{00}\mathbf{x}\, d\tau / \int T^{00}\, d\tau,$$

($T^{\mu\nu}$) being the energy–momentum tensor.

Starting, however, with (2.1) we can immediately derive a covariant expression for q'.

We have

$$M = (\mathbf{N}, \mathbf{M}) = (1/p^2)[(p \wedge \omega)^* - (p \wedge g)], \tag{2.2}[5]$$

where

$$\omega = -p \cdot M^*, \qquad g = -p \cdot M.$$

Hence, putting

$$q' = (1/p^2)g, \tag{2.3}$$

$$S_{q'} = (1/p^2)(p \wedge \omega)^*$$

$$= (1/p^2)(-\mathbf{p} \times \boldsymbol{\omega}, p^0\boldsymbol{\omega} - \mathbf{p}\omega^0), \tag{2.4}$$

satisfies

$$p \cdot S_{q'} = 0.$$

This is the definition of "center" adopted by Shirokov.[6]

In the above,

$$\begin{aligned} \omega &= (\mathbf{p}\cdot\mathbf{M}, p^0\mathbf{M} - \mathbf{p}\times\mathbf{N}), \\ g &= -(\mathbf{p}\cdot\mathbf{N}, p^0\mathbf{N} + \mathbf{p}\times\mathbf{M}). \end{aligned} \tag{2.5}$$

This makes the expression for q' quite complicated and even for spin zero (though Sq' is zero), q' is quite different from the usual x.

For a single particle there is, however, another simpler covariant solution of (2.1).

According to the usual decomposition,

$$M = -p \wedge x + S, \tag{2.6}$$

where

$$S = (\mathbf{n}, \boldsymbol{\zeta}), \quad \text{say},$$

is independent of x.

If we define

$$\lambda = -p \cdot S = -(\mathbf{p}\cdot\mathbf{n}, p^0\mathbf{n} + \mathbf{p}\times\boldsymbol{\zeta}), \tag{2.7}$$

and

$$q = x + (1/p^2)\lambda, \tag{2.8}$$

since

$$\omega = -p \cdot M^* = -p \cdot S^* = (\mathbf{p}\cdot\boldsymbol{\zeta}, p^0\boldsymbol{\zeta} - \mathbf{p}\times\mathbf{n}),$$

and

$$S = (1/p^2)\{(p \wedge \omega)^* - (p \wedge \lambda)\}, \tag{2.9}$$

we see that evidently $S_q = S'_q$ and now, moreover, for spin zero (i.e., $S = 0$), $q = x$.

Writing

$$S_q = (\mathbf{n}_q, \boldsymbol{\zeta}_q),$$

we note that

$$\mathbf{n}_q = -(\mathbf{p}/p^0) \times \boldsymbol{\zeta}_q, \tag{2.10}$$

which is a general feature of any S satisfying (2.1),[3] leading to

$$\mathbf{p}\cdot\mathbf{n} = 0 = \boldsymbol{\zeta}\cdot\mathbf{n} \tag{2.11}$$

and

$$S{:}S^* = 0.$$

Pryce,[2] having obtained a noncovariant expression for q_{Pr}, deduced another (X_{Pr}) supposed to be valid for any frame and coinciding with q_{Pr} in the rest frame. In the language of Bohm, Vigier, and Halbwachs,[7] q_{Pr} and X_{Pr} correspond to the "pseudocenter" and the "center," respectively.

In our case we have the covariant center from the start. To obtain the pseudocenter we transform $q(q')$ to the rest frame.

Since λ, ω transform as 4-vectors and $p\cdot\lambda = 0 = p\cdot\omega$, denoting by the subscript (0) the corresponding expressions transformed to rest frame,[8]

$$\begin{aligned} \omega_{(0)} &= [0, \boldsymbol{\omega} - (\mathbf{p}/p^0 + m)\omega^0], \\ \lambda_{(0)} &= [0, \boldsymbol{\lambda} - (\mathbf{p}/p^0 + m)\lambda^0], \end{aligned} \tag{2.12}$$

or, in terms of $\mathbf{n}$, $\boldsymbol{\zeta}$,

$$\begin{aligned} \boldsymbol{\omega}_{(0)} &= m\left[\boldsymbol{\zeta} - \frac{\mathbf{p}\times(\mathbf{p}\times\boldsymbol{\zeta})}{m(p^0+m)} - \frac{\mathbf{p}\times\mathbf{n}}{m}\right] \\ &= m\boldsymbol{\zeta}_{(0)}, \\ \boldsymbol{\lambda}_{(0)} &= -m\left[\mathbf{n} - \frac{\mathbf{p}\times(\mathbf{p}\times\mathbf{n})}{m(p^0+m)} + \frac{\mathbf{p}\times\boldsymbol{\zeta}}{m}\right] \\ &= -m\mathbf{n}_{(0)}, \end{aligned} \tag{2.13}$$

[5] In defining such operators we will adopt the convention of keeping p systematically to the left. This will later lead to no ambiguities involving noncommuting operators, and no symmetrization will be necessary. However, it is for "g" only that the precaution is essential.

[6] Iu. M. Shirokov, Zh. Eksperim. i Teor. Fiz. **35**, 1005 (1958) [English transl.: Soviet Phys.—JETP **8**, 703 (1959)].

[7] F. Halbwachs, *Théorie relativiste des fluides à spin* (Gauthier-Villars, Paris, 1960).

[8] Any 4-vector A is transformed to

$$A_{(0)} = LA,$$

where

$$L = 1 - \frac{(\mathrm{K}+p)\otimes(\mathrm{K}+p)}{(\mathrm{K}+p)\cdot p} + \frac{2\mathrm{K}\otimes p}{p^2},$$

and $\mathrm{K} = (m, \mathbf{o})$.

where $(\mathbf{n}_{(0)}, \boldsymbol{\zeta}_{(0)})$ is just the transformed tensor $(\mathbf{n}, \boldsymbol{\zeta})$.[9]

Substituting, we find that

$$(-p \wedge q + S_q)_{(0)} = -p_{(0)} \wedge q_{(0)} + [1/p^0_{(0)}](\mathbf{0}, \boldsymbol{\omega}_{(0)}), \tag{2.14}$$

where

$$q_{(0)} = L[x + (1/p^1)\lambda] \quad \text{and} \quad p_{(0)} = (p^0_{(0)}, \mathbf{0}).$$

Thus we can define in any frame, the pseudocenter ξ as

$$M = -p \wedge \xi + (1/p^0)(\mathbf{0}, \boldsymbol{\omega}) = -p \wedge \xi + S_\xi, \quad \text{say.} \tag{2.15}$$

To obtain an explicit expression for ξ we note that the definition

$$M = -p \wedge q + S_q$$

(where Sq is given) does not determine q completely. We can always add a term proportional to p, say fp, where, if we want q to be covariant, f must be a scalar. Similarly, the equations

$$-p \wedge q + S_q = -p \wedge \xi + S_\xi,$$

i.e.,

$$p^0(\boldsymbol{\xi} - \mathbf{q}) - \mathbf{p}(\xi^0 - q^0) = \mathbf{n}_\xi - \boldsymbol{\zeta}_\xi$$

and

$$\mathbf{p} \times (\boldsymbol{\xi} - \mathbf{q}) = \boldsymbol{\zeta}_\xi - \boldsymbol{\zeta}_q, \tag{2.16}$$

or

$$p \cdot (S_\xi - S_q) = -p[p \cdot (\xi - q)] + p^2(\xi - q)$$

leave us free to impose one additional condition on ξ. For example, choosing $\xi^0 = q^0$ [when finally also $p(\xi - q) = 0$], we have

$$\xi^0 = x^0 - (1/p^2)(\mathbf{p} \cdot \mathbf{n}), \qquad \boldsymbol{\xi} = \mathbf{x} - (1/p^0)\left[\mathbf{n} + \frac{1}{p^2}\,\mathbf{p}(\mathbf{p} \cdot \mathbf{n})\right]. \tag{2.17}$$

Exactly similarly, proceeding in terms of q', N, M instead of q, $\mathbf{n}$, $\boldsymbol{\zeta}$, we can derive a pseudocenter corresponding to q' as

$$\xi'^0 = -(1/p^2)(\mathbf{p} \cdot \mathbf{N}), \qquad \boldsymbol{\xi}' = -(1/p^0)[\mathbf{N} + (1/p^2)\mathbf{p}(\mathbf{p} \cdot \mathbf{N})]. \tag{2.18}$$

Taking account of the fact that $\mathbf{N}$ of Pryce[2] (which we will denote as $\mathbf{N}_{\mathrm{Pr}}$) is related to our $\mathbf{N}$ as

$$\mathbf{N}_{\mathrm{Pr}} = -\mathbf{N} + \mathbf{p}x^0,$$

our q' and ξ' are seen to differ, respectively, from X_{Pr} and q_{Pr} by the term

$$(p/p^0)[x^0 + (1/p^2)\mathbf{p} \cdot \mathbf{N}] \tag{2.19}$$

[where Pryce moreover replaces p^2, p^0 by κ^2 and $E = (\mathbf{p}^2 + \kappa^2)^{\frac{1}{2}}$, respectively].[10] This, of course, leaves the spin part unaffected. As was noted by Pryce,[2] neither the center nor the pseudocenter has components with the usual commutation relations (or Poisson brackets in the classical case) and, to satisfy this requirement, he defined a "weighted mean"

$$\tilde{\mathbf{q}}_{\mathrm{Pr}} = (E\mathbf{q}_{\mathrm{Pr}} + \kappa\mathbf{X}_{\mathrm{Pr}})/(E + \kappa).$$

In our case we can likewise define

$$X = (mq + p^0\xi)/(m + p^0), \tag{2.20}$$

giving

$$X^0 = x^0 - \frac{1}{m^2}(\mathbf{p} \cdot \mathbf{n}), \qquad \mathbf{X} = \mathbf{x} - \left[\frac{1}{m}\,\mathbf{n} + \frac{1}{m^2(p^0 + m)}\,\mathbf{p}(\mathbf{p} \cdot \mathbf{n}) + \frac{1}{m(p^0 + m)}\,\mathbf{p} \times \boldsymbol{\zeta}\right]. \tag{2.21}$$

For spin $\frac{1}{2}$, substituting

$$\mathbf{n} = \tfrac{1}{2}i\boldsymbol{\alpha}, \qquad \boldsymbol{\zeta} = \tfrac{1}{2}\boldsymbol{\vartheta},$$

we find that X is just the operator

$$Q^{-1}xQ,$$

where

$$Q = \frac{[\gamma^0(\gamma \cdot p) + m]}{[2m(p^0 + m)]^{\frac{1}{2}}},$$

and

$$p^0 \equiv i(\partial/\partial x^0), \qquad m = \pm(p^\nu p_\nu)^{\frac{1}{2}}.$$

(Evidently, in the general case, X has the proper commutation properties with the above definitions of p^0 and m.)

It will be noticed that q, ξ, and X, as defined above, all give a certain nonlocality in time (as well as in space) as is evident from the expression for $q^0 (= \xi^0 = X^0)$. (This nonlocality, however, disappears in the rest frame while the spatial part does not.)

If we want to avoid this, without affecting the respective spin tensors, we can add the term

[9] C. Möller, *Theory of Relativity* (Clarendon Press, Oxford, England, 1955).

[10] Since our q' is manifestly covariant, Pryce's statement that X_{pr} is covariant "in spite of its appearance" depends on the term $(1/p^0)[x^0 + (1/p^2)\ \mathbf{p} \cdot \mathbf{N}]$ being scalar. This seems to be true only for purely spatial notation.

$$(p/p^0)(1/m^2)(\mathbf{p}\cdot\mathbf{n}) \tag{2.22}$$

to each. Denoting the operators thus obtained by q'', ξ'', and X'', respectively, we have in particular

$$X''^0 = x^0, \tag{2.23}$$

$$\mathbf{X}'' = \mathbf{x} - \left[\frac{1}{m}\mathbf{n} - \frac{\mathbf{p}(\mathbf{p}\cdot\mathbf{n})}{p^0 m(p^0+m)} + \frac{\mathbf{p}\times\boldsymbol{\zeta}}{m(p^0+m)}\right] = \mathbf{x} - \frac{1}{p^0}\left[\mathbf{n} + \frac{1}{m(p^0+m)}\,\mathbf{p}\times\boldsymbol{\omega}\right].$$

For the case of a Dirac particle, exactly this formula is obtained, if in $Q^{-1}xQ$, p^0 is replaced by $\pm(p^2+x^2)^{\frac{1}{2}}$. ($p^0 = \pm E$ when $m = \pm\kappa$.) Hence evidently, with this convention, X'' has also the proper commutation relations.[11]

Proceeding similarly, adding the term (2.19) to q' and ξ', we get, of course (in a slightly different notation), just

$$\tilde{\mathbf{q}}_{\mathrm{Pr}} = -\left[\frac{1}{m}\mathbf{N} - \frac{1}{p^0 m(p^0+m)}\,\mathbf{p}(\mathbf{p}\cdot\mathbf{N}) + \frac{1}{m(p^0+m)}\,\mathbf{p}\times\mathbf{M}\right] = -\frac{1}{p^0}\left[\mathbf{N} + \frac{1}{m(p^0+m)}\,\mathbf{p}\times\boldsymbol{\omega}\right]. \tag{2.24}$$

Moreover, in quantum mechanics, when p^0 is replaced by $\pm(\mathbf{p}^2+\kappa^2)^{\frac{1}{2}}$, to preserve the commutation relations we should put

$$\mathbf{N} = -p^0\mathbf{x} + \mathbf{n}$$

(rather than $\mathbf{N} = \mathbf{p}x^0 - p^0\mathbf{x} + \mathbf{n}$). Thus, since $\boldsymbol{\omega}$ is already independent of the orbital part, we find that

$$\tilde{\mathbf{q}}_{\mathrm{Pr}} = \mathbf{x} - \frac{1}{p^0}\left[\mathbf{n} + \frac{1}{m(p^0+m)}\,\mathbf{p}\times\boldsymbol{\omega}\right] = \mathbf{X}''. \tag{2.25}$$

Thus we find that Pryce's initial definition of the "center" (having the proper Poisson brackets) corresponds directly to our transformation (for the spin-$\frac{1}{2}$ case) when we pass over to quantum mechanics—we have just to substitute for $\mathbf{n}$, $\boldsymbol{\zeta}$; nothing else is required. The quantum-mechanical position operator finally obtained by Pryce (for a Dirac particle) differs from (2.25) in the denominators and through the appearance of β. Also, Pryce's method of passing over to quantum mechanics is not wholly consistent—whereas for $\mathbf{N}$ the spin part ($\mathbf{n} = \frac{1}{2}i\boldsymbol{\alpha}$) is obtained by symmetrization, for $\mathbf{M}$ its existence is taken for granted.

It may be noted however that just as

$$\mathbf{q}_{\mathrm{Pr}}\ (=E^{-1}\mathbf{N}_{\mathrm{Pr}}) = \tfrac{1}{2}[H/E^2, \tfrac{1}{2}[\mathbf{x}, H]_+]_+$$

where $H = \beta\kappa + \boldsymbol{\alpha}\cdot\mathbf{p}$, we can obtain our covariant q as

$$q = \tfrac{1}{2}[(\boldsymbol{\gamma}\cdot p/p^2), \tfrac{1}{2}[x, \boldsymbol{\gamma}\cdot p]_+]_+$$

which corresponds to the fact that in our case it is $\boldsymbol{\gamma}\cdot p$ that is diagonalized, rather than H.

In our case (as discussed in reference 1) the unitarity of the transformation and the Hermiticity of the relevant operators hold, of course, with respect to the metric γ^0 for spin $\frac{1}{2}$, and, in general, a metric given by Ω, such that

$$[\mathbf{n}, \Omega]_+ = 0, \qquad [\xi, \Omega]_- = 0. \tag{2.26}$$

3. PROPERTIES OF THE CENTERS DEFINED

We have already mentioned that even for a single particle the utility of these position operators lies in the fact that the corresponding velocities and intrinsic angular momenta have certain desirable properties.

As for the spin tensors, we have, respectively, for q, ξ, and X (the same expressions holding also for q', ξ', X' and q'', ξ'', X'', where X' is obtained as the weighted mean of q', ξ'),

$$\begin{aligned} S_q \equiv (\mathbf{n}_q, \boldsymbol{\zeta}_q) &= \frac{1}{m^2}(p \wedge \omega)^* \\ &= \frac{1}{m^2}(-\mathbf{p}\times\boldsymbol{\omega}, p^0\boldsymbol{\omega} - \mathbf{p}\omega^0), \\ S_\xi \equiv (\mathbf{n}_\xi, \boldsymbol{\zeta}_\xi) &= \frac{1}{p^0}(\mathbf{0}, \boldsymbol{\omega}), \\ S_x \equiv (\mathbf{n}_x, \boldsymbol{\zeta}_x) &= \frac{1}{m}\left(\frac{\mathbf{p}\times\boldsymbol{\omega}}{p^0+m}, \boldsymbol{\omega} - \frac{\mathbf{p}}{p^0+m}\omega^0\right) \\ &= \frac{1}{m}\left(\frac{-1}{p^0+m}\,\mathbf{p}\times\boldsymbol{\omega}_{(0)}, \boldsymbol{\omega}_{(0)}\right). \end{aligned} \tag{3.1}$$

We note that in all the cases

$$S{:}S^* = 0$$

(though only for q, $p\cdot S_q = 0$).

The significance of the fact that $\boldsymbol{\zeta}_X$ (or $\boldsymbol{\zeta}_{X''}$) satisfies

$$\boldsymbol{\zeta}_X = (1/m)\boldsymbol{\omega}_{(0)}$$

has been discussed in detail elsewhere.[1] It was also shown that the Dirac Hamiltonian commutes

[11] This form of the transformation corresponds to taking the Dirac equation in the time-independent form

$$(\pm\gamma^0 E - \boldsymbol{\gamma}\cdot\mathbf{p} - \kappa)\,\varphi = 0,$$

when the transformed equation is simply (since $m = \pm x$)

$$(\gamma^0 \mp 1)\,\varphi = 0,$$

accordingly as the energy is positive or negative. That is, it is separated into two 2-component equations, each spinor, satisfying the K–G equation.

effectively with the polarization operator, so that it may be considered as a constant of motion.

In considering the velocities associated with the different definitions of the position operator, in absence of a suitable general Hamiltonian formulation we will consider the case of a Dirac particle only. Since

$$q = x - (1/2m^2)(i\boldsymbol{\alpha}\cdot\mathbf{p},\, i\boldsymbol{\alpha}p^0 + \mathbf{p}\times\boldsymbol{\sigma}),$$

$$\dot{\mathbf{q}} = i[H, \mathbf{q}] = (\boldsymbol{\alpha}/m^2)p^0(p^0 - H) + (\mathbf{p}/m^2)$$

$$\times [p^0 - (\boldsymbol{\alpha}\cdot\mathbf{p})] \approx (\mathbf{p}/m^2)\beta\kappa \approx \beta(\mathbf{p}/\kappa) \qquad (3.2)$$

[since $(p^0 - H)\varphi = (p^0 - \boldsymbol{\alpha}\cdot\mathbf{p} - \beta\kappa)\varphi = 0$, and $A \approx B$ signifies[1] $(A - B)\varphi = 0$].

Proceeding similarly, we can show that

$$\dot{\mathbf{q}} \approx \dot{\boldsymbol{\xi}} \approx \dot{\mathbf{X}} \approx \beta\mathbf{p}/\kappa. \qquad (3.3)$$

Again since

$$i\left[H, \frac{1}{2m^2}(\boldsymbol{\alpha}\cdot\mathbf{p})\frac{\mathbf{p}}{p^0}\right] = -\frac{\beta\mathbf{p}}{\kappa}\left(\frac{H - \kappa\beta}{p^0}\right)$$

$$\approx -\frac{\beta\mathbf{p}}{\kappa} + \frac{\mathbf{p}}{p^0}, \qquad (3.4)$$

we find directly from (2.22) and (3.3)

$$\dot{\mathbf{q}}'' \approx \dot{\boldsymbol{\xi}}'' \approx \dot{\mathbf{X}}'' \approx \mathbf{p}/p^0 \qquad (3.5)$$

[where $p^0 = \pm(\mathbf{p}^2 + \kappa^2)^{\frac{1}{2}}$].

Thus we see that the operator X'', obtained by transforming with $p^0 = \pm E$, has exactly the required relativistic expression for velocity.

On the other hand, X [which we obtain with $p^0 \equiv i(\partial/\partial x^0)$ in Q] has for its time derivatives, so to say, the spatial components of the 4-velocity. (Since β anticommutes with the charge conjugation operator we may say that, for the same $\mathbf{p}$, the two eigenvalues ± 1 of β correspond to the particle and the antiparticle, respectively.) Since β does not commute with H, we get an accelerated motion. It will be noticed, however, that the squares of the velocity and acceleration are constants of motion. This indicates a circular motion over a region of nonlocality, somewhat analogous to the motion of the "center of matter" of a "classical Dirac particle" considered by Halbwachs.[7]

The third typical feature of such a position operator is that it implies a certain nonlocality, i.e., it implies that the particle is spread over a certain small region. As is well-known, what appears as the region covered by the zitterbewegung in the Dirac representation appears as the region of nonlocality (of the order of one Compton wavelength) in the F–W representation, this nonlocality being related to the uncertainly principle. It may be verified (for example by a calculation similar to that of Rose[12]) that the same phenomenon appears in connection with our representation also.

4. CONCLUSION

We have derived two expressions for the position operator corresponding to our representation accordingly as we put $p^0 \equiv i(\partial/\partial x^0)$, or $p^0 \equiv \pm E$. The former implies, in an eigenstate of energy, a closed circular internal motion, and also a nonlocality in time, vanishing in the rest system. The latter operator, though differing in form from the F–W position operator, has effectively the same expression for velocity. Also we have shown that this operator has a more direct correspondence with the classical definition of Pryce than the F–W operator. We have also connected the operator (X, X'') with the other definitions of "Center" and "pseudocenter" (q, ξ and q'', ξ'', respectively).

From the remarks on the form of the spin tensor, it is evident that $\mathbf{N}_Q$, $\mathbf{M}_Q$ considered in reference 1 does not change form whether or not we replace p^0 by $\pm E$.

It will be remembered however,[1] that our transformations, and hence our position operators, are valid for nonzero rest mass only.

APPENDIX

It is often remarked that the Newton–Wigner (N–W) position operator[13] coincides (for spin $\frac{1}{2}$) with the F–W one. There is, however, some difference which should be noted.

The N–W operator (denoted as q_N) is, in their notation,

$$q_N^k = E\prod_{\alpha=1}^{2S}(1 + \gamma_\alpha^0)\frac{p_0^{2S+\frac{1}{2}}}{(p^0 + \mu)^S}$$

$$\times\left(-i\frac{\partial}{\partial p_k}\right)\frac{p_0^{-\frac{1}{2}}}{(p_0 + \mu)^{\frac{1}{2}}}E, \qquad (A1)$$

where E is the energy projection operator. For $S = \frac{1}{2}$ this reduces to

$$\mathbf{q}_N = E(1 + \gamma^0)\frac{p_0^{\frac{1}{2}}}{(p^0 + \mu)^{\frac{1}{2}}}\times\frac{p_0^{-\frac{1}{2}}}{(p^0 + \mu)^{\frac{1}{2}}}E, \qquad (A2)$$

where

$$E = \frac{\boldsymbol{\gamma}\cdot p + \mu}{2p_0}\gamma^0 = \frac{H + p_0}{2p_0}.$$

[12] M. E. Rose, *Relativistic Electron Theory* (John Wiley & Sons, Inc., New York, 1961).

[13] T. D. Newton and E. P. Wigner, Rev. Mod. Phys. **21**, 400 (1949).

Since

$$E(1+\beta)E = \frac{p^0 + H\beta}{2p^0}(1+\beta)\frac{p^0+\beta H}{2p^0}$$

$$= F^{-1}[2p_0(p_0+\mu)]^{\frac{1}{2}}(1+\beta)[2p_0(p_0+\mu)]^{\frac{1}{2}}F,$$

where F is the F–W transformation operator, we have

$$\mathbf{q}_N = F^{-1}[\tfrac{1}{2}(1+\beta)p_0\mathbf{x}p_0^{-1}]F,$$

or

$$F\mathbf{q}_N F^{-1} = [\mathbf{x} + i\mathbf{p}/p_0^2]\tfrac{1}{2}(1+\beta), \qquad \text{(A3)}$$

instead of just $\mathbf{x}$, as would have been the case for $F^{-1}\mathbf{x}F$. The additional terms are, however, easy to interpret. $\frac{1}{2}(1+\beta)$ is just the energy projection operator in the F–W representation (since $H_F = \beta p_0$) and the term $i\mathbf{p}/p_0$ is necessary to make the operator Hermitian in presence of the term p_0^{-2} in the volume element.

To compare our representation (and $\mathbf{X}''$) with the above, we note that if we write

$$\mathbf{q}'_N = E'(1+\gamma^0)\frac{p_0^{\frac{3}{2}}}{(p_0+\mu)^{\frac{3}{2}}}\left(\frac{p_0}{\mu}\right)^{-\frac{1}{2}}\mathbf{x}$$

$$\times\left(\frac{p_0}{\mu}\right)^{-\frac{1}{2}}\frac{p_0^{-\frac{1}{2}}}{(p_0+\mu)^{\frac{1}{2}}}E, \qquad \text{(A4)}$$

where $E' = (\gamma\cdot p + \mu)/2\mu$ is the other form of the projector, then

$$Q\mathbf{q}'_N Q^{-1} = [\mathbf{x} + (i\mathbf{p}/p^0)]\tfrac{1}{2}(1+\gamma^0). \qquad \text{(A5)}$$

In our case the additional factor $(p_0/\mu)^{-\frac{1}{2}}$ arises due to the fact that we have taken the scalar product as[1] $\bar{\psi}\psi$ rather than $\psi^*\psi$, and for nonzero rest mass,

$$\bar{\psi}\psi = (\mu/p^0)\psi^*\psi.$$

Invariant Solutions of the Exact Bethe–Salpeter Equation in the General-Mass Case*

NOBORU NAKANISHI†
Applied Mathematics Division, Brookhaven National Laboratory, Upton, New York
(Received 9 August 1962)

The invariant solutions of the exact Bethe–Salpeter equation are found in the case of nonzero exchanged-meson mass. They are expressed as series of "truncated vertex functions," which are certain combinations of the usual vertex functions. The eigenvalues of the coupling constant are determined by a transcendental equation, which is not solved explicitly.

I. INTRODUCTION

IT has been a long-standing problem to solve the Bethe–Salpeter equation[1] for bound states of a relativistic two-body system. The first attempt of solving it relativistically was made by Goldstein.[2] He dealt with the spinor–spinor system in which both the total energy–momentum and the exchanged-meson mass are zero. He could not, however, find any discrete spectrum without a cutoff. An essential progress for this problem was made by Wick[3] and Cutkosky.[4] They obtained a complete set of solutions for the scalar–scalar system in which the exchanged-meson mass is zero. The scalar–spinor case was investigated by Sugano and Munakata.[5] As for the case of nonzero exchanged-meson mass, the present author[6] solved the Bethe–Salpeter equation for the scalar–scalar system in which the bound-state mass is zero by making use of a double dispersion representation. All these works were done in ladder approximation. No attempt of solving the exact Bethe–Salpeter equation has been made so far.

The purpose of the present paper is to solve the exact Bethe–Salpeter equation for the scalar–scalar system in general-mass case. We shall assume that the exchanged-meson mass is nonzero, and consider invariant solutions only. The method employed in the following is an elegant generalization of our previous work.[6] In Sec. II we shall analytically continue the Bethe–Salpeter amplitude to the region where the squared bound-state mass is negative. Then a double dispersion representation becomes always possible for the amplitude. We shall convert the Bethe–Salpeter equation into a system of inhomogeneous integral equations for spectral functions, which can be solved by an iteration method. Next, we shall carry out the spectral integrations of the solutions term by term, and prove that each term can be written in terms of the "truncated vertex function" which will be introduced in Sec. III. Finally, we shall analytically continue the solutions to the physical value of the squared bound-state mass. In Sec. IV some remarks will be made on our invariant solutions of the Bethe–Salpeter equation.

II. SOLUTIONS IN THE DOUBLE DISPERSION REPRESENTATION

The exact Bethe–Salpeter equation for two scalar particles having masses m and μ is

$$\tilde{f}(p, k) = [m^2 - (p + k)^2 - i\epsilon]^{-1} \times [\mu^2 - (p - k)^2 - i\epsilon]^{-1} \int d^4q\, H(p, q, k)\tilde{f}(q, k). \tag{2.1}$$

Here $2k$ is the total 4-momentum of the system, p being the relative 4-momentum of the two particles; $\tilde{f}(p, k)$ is the Bethe–Salpeter amplitude, and the kernel $H(p, q, k)$ corresponds to Fig. 1 (irreducible graphs only). The radiative corrections Δ_F'/Δ_F should be understood to be included in $H(p, q, k)$. We assume $\kappa \neq 0$, where κ denotes the mass of the lightest particle which interacts with the two particles m and μ.

In ladder approximation we have

$$H(p, q, k) = [\lambda/\pi^2 i][\kappa^2 - (p - q)^2 - i\epsilon]^{-1}, \tag{2.2}$$

where λ is the squared coupling constant. This expression will not be used in the following consideration, because we are interested in a more general case.

Hereafter we consider invariant solutions (S-wave solutions) only. Then $\tilde{f}(p, k)$ becomes a function

* This work was performed under the auspices of the U. S. Atomic Energy Commission.
† On leave of absence from the Institute for Advanced Study, Princeton, New Jersey.
[1] E. E. Salpeter and H. A. Bethe, Phys. Rev. 84, 1232 (1951). M. Gell-Mann and F. Low, Phys. Rev. 84, 350 (1951). H. Kita, Progr. Theoret. Phys. (Kyoto) 7, 217 (1952).
[2] J. S. Goldstein, Phys. Rev. 91, 1516 (1953).
[3] G. C. Wick, Phys. Rev. 96, 1124 (1954).
[4] R. E. Cutkosky, Phys. Rev. 96, 1135 (1954).
[5] R. Sugano and Y. Munakata, Progr. Theoret. Phys. (Kyoto) 16, 532 (1956).
[6] N. Nakanishi, Progr. Theoret. Phys. (Kyoto) 24, 1275 (1960).

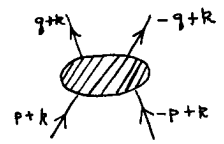


FIG. 1. Graphical explanation of the kernel $H(p, q, k)$.

of $(p + k)^2$, $(p - k)^2$, and $(2k)^2$ only. Since the squared bound-state mass $u \equiv (2k)^2$ can be regarded as a parameter, we may write

$$\tilde{f}(p, k) \equiv f(s, t), \tag{2.3}$$

where $s \equiv (p + k)^2$ and $t \equiv (p - k)^2$.

Now, we shall investigate (2.1) in an unphysical region $u < 0$. As was proved by Yamamoto[7] in axiomatic theory, $f(s, t)$ can be represented as

$$f(s, t) = \int_{m^2}^{\infty} ds' \int_{\mu^2}^{\infty} dt' \frac{\varphi(s', t')}{(s' - s - i\epsilon)(t' - t - i\epsilon)}. \tag{2.4}$$

Substituting (2.4) in the q-integral of (2.1) we get

$$\int d^4q\, H(p, q, k)\tilde{f}(q, k) = \int_{m^2}^{\infty} ds'' \int_{\mu^2}^{\infty} dt''\, M(s, t; s'', t'')\varphi(s'', t''), \tag{2.5}$$

where the function

$$M(s, t; s'', t'') \equiv \int d^4q \times \frac{H(p, q, k)}{[s'' - (q + k)^2 - i\epsilon][t'' - (q - k)^2 - i\epsilon]} \tag{2.6}$$

is the vertex function shown in Fig. 2. Using Yamamoto's result[7] again, we have

$$M(s, t; s'', t'') = \int_0^{\infty} ds' \int_0^{\infty} dt' \times \frac{K_0(s', t'; s'', t'')}{(s' - s - i\epsilon)(t' - t - i\epsilon)}, \tag{2.7}$$

where $K_0(s', t'; s'', t'')$ vanishes unless

$$s' \geq (m + \kappa)^2, \qquad t' \geq (\mu + \kappa)^2 \tag{2.8}$$

for $s'' \geq m^2$ and $t'' \geq \mu^2$. Here, for simplicity, we have assumed that the particles m and μ can not be converted into lighter particles.

In ladder approximation the concrete expression for K_0 is

$$\lambda^{-1}K_0(s', t'; s'', t'') = [(s' - t')^2 - 2u(s' + t') + u^2]^{-\frac{1}{2}} \times \theta(s' - [s''^{\frac{1}{2}} + \kappa]^2)\cdot\theta(t' - [t''^{\frac{1}{2}} + \kappa]^2)\theta(D), \tag{2.9}$$

[7] K. Yamamoto, Progr. Theoret. Phys. (Kyoto) **25**, 720 (1961).

with

$$D \equiv \begin{vmatrix} 2s'' & s'' + t'' - u & s'' + \kappa^2 - s' \\ s'' + t'' - u & 2t'' & t'' + \kappa^2 - t' \\ s'' + \kappa^2 - s' & t'' + \kappa^2 - t' & 2\kappa^2 \end{vmatrix}, \tag{2.10}$$

but this expression will not be used in the following calculation.

The substitution of (2.7) in (2.5) leads to

$$\int d^4q\, H(p, q, k)\tilde{f}(q, k) = \int ds' \int dt' \times \frac{\Phi(s', t')}{(s' - s - i\epsilon)(t' - t - i\epsilon)}, \tag{2.11}$$

with

$$\Phi(s', t') \equiv \int ds'' \int dt''\, K_0(s', t'; s'', t'')\varphi(s'', t''). \tag{2.12}$$

By using an identity

$$(m^2 - s - i\epsilon)^{-1}(s' - s - i\epsilon)^{-1} = (s' - m^2 - i\epsilon)^{-1} \times [(s' - s - i\epsilon)^{-1} - (m^2 - s - i\epsilon)^{-1}], \tag{2.13}$$

the right-hand side of (2.1) can be written as

$$\int ds' \int dt' \frac{\Psi(s', t')}{(s' - s - i\epsilon)(t' - t - i\epsilon)}, \tag{2.14}$$

with

$$\begin{aligned}\Psi(s', t') \equiv{}& \delta(s' - m^2)\delta(t' - \mu^2) \\ &\times \int dv \int dw \frac{\Phi(v, w)}{(v - m^2 - i\epsilon)(w - \mu^2 - i\epsilon)} \\ &- \delta(t' - \mu^2)\frac{1}{s' - m^2 - i\epsilon}\int dw \frac{\Phi(s', w)}{w - \mu^2 - i\epsilon} \\ &- \delta(s' - m^2)\frac{1}{t' - \mu^2 - i\epsilon}\int dv \frac{\Phi(v, t')}{v - m^2 - i\epsilon} \\ &+ \frac{\Phi(s', t')}{(s' - m^2 - i\epsilon)(t' - \mu^2 - i\epsilon)}.\end{aligned} \tag{2.15}$$

Substituting (2.4) and (2.14) in (2.1), we obtain

$$\varphi(s', t') = \Psi(s', t'). \tag{2.16}$$

Hence, according to (2.15) together with (2.12),

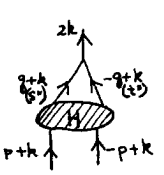


FIG. 2. The graph corresponding to $M(s, t; s'', t'')$.

(2.8), and $\kappa > 0$, $\varphi(s', t')$ must have the following spectrum:

$$\varphi(s', t') = \delta(s' - m^2)\delta(t' - \mu^2) + \varphi_1(s')\delta(t' - \mu^2) + \varphi_2(t')\delta(s' - m^2) + \varphi_{12}(s', t'), \quad (2.17)$$

where

$$\varphi_1(s') = \varphi_1(s')\theta(s' - [m + \kappa]^2),$$
$$\varphi_2(t') = \varphi_2(t')\theta(t' - [\mu + \kappa]^2), \quad (2.18)$$
$$\varphi_{12}(s', t') = \varphi_{12}(s', t')\theta(s' - [m + \kappa]^2) \times \theta(t' - [\mu + \kappa]^2).$$

For simplicity, we normalize the coefficient of the first term of (2.17) to unity (i.e., $a = 1$), provided that $a \neq 0$.

From (2.16), (2.17), and (2.15), we obtain

$$1 = \int dv \int dw \frac{\Phi(v, w)}{(v - m^2)(w - \mu^2)}, \quad (2.19)$$

$$\varphi_1(s') = -\frac{1}{s' - m^2} \int dw \frac{\Phi(s', w)}{w - \mu^2}, \quad (2.20)$$

$$\varphi_2(t') = -\frac{1}{t' - \mu^2} \int dv \frac{\Phi(v, t')}{v - m^2}, \quad (2.21)$$

$$\varphi_{12}(s', t') = \frac{\Phi(s', t')}{(s' - m^2)(t' - \mu^2)}. \quad (2.22)$$

The substitution of (2.12) and (2.17) in (2.22) gives an equation for $\varphi_{12}(s', t')$:

$$\varphi_{12}(s', t') = \eta_{12}(s', t') + \int ds'' \int dt'' K(s', t'; s'', t'')\varphi_{12}(s'', t''), \quad (2.23)$$

with

$$\eta_{12}(s', t') \equiv K(s', t'; m^2, \mu^2) + \int ds'' K(s', t'; s'', \mu^2)\varphi_1(s'') + \int dt'' K(s', t'; m^2, t'')\varphi_2(t''), \quad (2.24)$$

$$K(s', t'; s'', t'') \equiv (s' - m^2)^{-1} \times (t' - \mu^2)^{-1} K_0(s', t'; s'', t''). \quad (2.25)$$

Hereafter we shall use the following "matrix" notation:

$$\int dv \int dw\, A(s', t'; v, w)B(v, w; s'', t'') \equiv (AB)(s', t'; s'', t''). \quad (2.26)$$

Equation (2.23) is an inhomogeneous integral equation, so that it can be solved by an iteration method. Namely, the resolvent R is given by the Neumann series

$$R(s', t'; s'', t'') = \sum_{n=1}^{\infty} K^n(s', t'; s'', t''), \quad (2.27)$$

where K^n is defined according to (2.26). Thus the solution of (2.23) is

$$\varphi_{12}(s', t') = \eta_{12}(s', t') + \int ds'' \int dt'' \times R(s', t'; s'', t'')\eta_{12}(s'', t'') = R(s', t'; m^2, \mu^2) + \int ds'' R(s', t'; s'', \mu^2)\varphi_1(s'') + \int dt'' R(s', t'; m^2, t'')\varphi_2(t''), \quad (2.28)$$

where we have used a property of the resolvent

$$R = K + KR = K + RK. \quad (2.29)$$

Next, substituting (2.12), (2.17), and (2.28) in (2.20) and (2.21), and using (2.29) again, we get

$$\varphi_1(s') = -\int dt' R(s', t'; m^2, \mu^2) - \int dt' \int ds'' R(s', t'; s'', \mu^2)\varphi_1(s'') - \int dt' \int dt'' R(s', t'; m^2, t'')\varphi_2(t''), \quad (2.30)$$

$$\varphi_2(t') = -\int ds' R(s', t'; m^2, \mu^2) - \int ds' \int ds'' R(s', t'; s'', \mu^2)\varphi_1(s'') - \int ds' \int dt'' R(s', t'; m^2, t'')\varphi_2(t''). \quad (2.31)$$

These equations are simultaneous inhomogeneous integral equations, which can be solved by the iteration method again. Substituting the solutions of (2.30) and (2.31) in (2.28), we obtain

$$\varphi_{12}(s', t') = R(s', t'; m^2, \mu^2) + \sum_{n=1}^{\infty} (S^n R)(s', t'; m^2, \mu^2), \quad (2.32)$$

where

$$S(s', t'; s'', t'') \equiv -R(s', t'; s'', \mu^2) - R(s', t'; m^2, t''), \quad (2.33)$$

and

$$\varphi_1(s') = -\int dt' \varphi_{12}(s', t'), \quad (2.34)$$

$$\varphi_2(t') = -\int ds' \varphi_{12}(s', t'). \quad (2.35)$$

The rearrangement of (2.32) in powers of K leads to

$$\varphi_{12}(s', t') = K(s', t'; m^2, \mu^2) + \sum_{n=1}^{\infty} (T^n K)(s', t'; m^2, \mu^2), \quad (2.36)$$

where

$$T(s', t'; s'', t'') \equiv K(s', t'; s'', t'') - K(s', t'; s'', \mu^2) - K(s', t'; m^2, t''). \quad (2.37)$$

Finally, from (2.19) and (2.22), we have

$$\int ds' \int dt' \, \varphi_{12}(s', t') = 1, \quad (2.38)$$

which determines the eigenvalues.

III. SOLUTIONS IN TERMS OF THE TRUNCATED VERTEX FUNCTIONS

We define the following quantities:

$$\psi(p, q, k) \equiv \frac{H(p, q, k)}{[m^2 - (q+k)^2 - i\epsilon][\mu^2 - (q-k)^2 - i\epsilon]}, \quad (3.1)$$

$$M^{(1)}(s, t) \equiv \int d^4q \, \psi(p, q, k) = M(s, t; m^2, \mu^2), \quad (3.2)$$

$$M^{(n+1)}(s, t) \equiv \int d^4q \, \psi(p, q, k) \times M^{(n)}([q+k]^2, [q-k]^2). \quad (3.3)$$

$M^{(n)}(s, t)$ is the usual vertex function containing $H(p, q, k)$ n times. Likewise we define

$$L^{(1)}(s, t) \equiv M^{(1)}(s, t), \quad (3.4)$$

$$L^{(n+1)}(s, t) \equiv \int d^4q \, \psi(p, q, k) \times \{L^{(n)}([q+k]^2, [q-k]^2) - L^{(n)}(m^2, \mu^2)\}, \quad (3.5)$$

and we call $L^{(n)}(s, t)$ the "truncated vertex function." It can be easily shown by mathematical induction that

$$L^{(n)}(s, t) = \sum_{l=1}^{n} (-1)^{l-1} \sum_{P_l} M^{(n_1)}(s, t) \times M^{(n_2)}(m^2, \mu^2) \cdots M^{(n_l)}(m^2, \mu^2), \quad (3.6)$$

where P_l stands for a permutation $(n_1, n_2, \cdots, n_l)$ satisfying $n_1 + n_2 + \cdots + n_l = n$, ($n_i$ is a natural number). If one solves (3.6) with respect to $M^{(n)}(s, t)$, one obtains

$$M^{(n)}(s, t) = \sum_{l=1}^{n} \sum_{P_l} L^{(n_1)}(s, t) \times L^{(n_2)}(m^2, \mu^2) \cdots L^{(n_l)}(m^2, \mu^2). \quad (3.7)$$

This formula is closely analogous to Haag's definition[8] of the truncated Wightman function.

Now, the following theorem gives the relationship between the truncated vertex functions and the solutions of the Bethe–Salpeter equation obtained in the preceding section.

Theorem. For $n \geq 1$,

$$L^{(n+1)}(s, t) = \int ds' \int dt' \int dv' \int dw' \times (s' - s - i\epsilon)^{-1}(t' - t - i\epsilon)^{-1} \times [K_0(s', t'; v', w') - K_0(s', t'; v', \mu^2) - K_0(s', t'; m^2, w')](T^{n-1}K)(v', w'; m^2, \mu^2). \quad (3.8)$$

Proof: We first calculate the following quantity:

$$\tilde{L}(s, t; a, b) \equiv \int ds' \int dt' \int dv' \int dw' \times (s' - s - i\epsilon)^{-1}(t' - t - i\epsilon)^{-1} \times [K_0(s', t'; v', w') - K_0(s', t'; v', \mu^2) - K_0(s', t'; m^2, w')]K(v', w'; a, b). \quad (3.9)$$

From (2.7) we have

$$\tilde{L}(s, t; a, b) \equiv \int dv' \int dw' \times [M(s, t; v', w') - M(s, t; v', \mu^2) - M(s, t; m^2, w')]K(v', w'; a, b). \quad (3.10)$$

The substitution of (2.6) and (2.25) in (3.10) leads to

$$\tilde{L}(s, t; a, b) = \int dv' \int dw' \int dv \int dw \times [(v' - v - i\epsilon)^{-1}(w' - w - i\epsilon)^{-1} - (v' - v - i\epsilon)^{-1}(\mu^2 - w - i\epsilon)^{-1} - (m^2 - v - i\epsilon)^{-1}(w' - w - i\epsilon)^{-1}] \times \int d^4q \, \delta(v - [q+k]^2) \, \delta(w - [q-k]^2) \times H(p, q, k)(v' - m^2)^{-1}(w' - \mu^2)^{-1} \times K_0(v', w'; a, b) = \int dv' \int dw' \int dv \int dw \times \int d^4q \, \delta(v - [q+k]^2) \, \delta(w - [q-k]^2) \times H(p, q, k)(m^2 - v - i\epsilon)^{-1}(\mu^2 - w - i\epsilon)^{-1} \times \{[(v' - v - i\epsilon)^{-1} - (v' - m^2)^{-1}] \times [(w' - w - i\epsilon)^{-1} - (w' - \mu^2)^{-1}]$$

[8] R. Haag, Phys. Rev. **112**, 669 (1958).

$$+ [(v' - v - i\epsilon)^{-1} - (v' - m^2)^{-1}](w' - \mu^2)^{-1}$$
$$+ [(w' - w - i\epsilon)^{-1} - (w' - \mu^2)^{-1}](v' - m^2)^{-1}\}$$
$$\times K_0(v', w'; a, b) = \int dv' \int dw' \int d^4q\psi(p, q, k)$$
$$\times \{[v' - (q + k)^2 - i\epsilon]^{-1}[w' - (q - k)^2 - i\epsilon]^{-1}$$
$$- (v' - m^2)^{-1}(w' - \mu^2)^{-1}\}K_0(v', w'; a, b). \quad (3.11)$$

Now, let us prove the theorem by mathematical induction. We will write the right-hand side of (3.8) as $\tilde{L}^{(n+1)}(s, t)$. First, for $n = 1$, from (3.8), (3.9), and (3.11), we have

$$\tilde{L}^{(2)}(s, t) = \int dv' \int dw' \int d^4q\psi(p, q, k)$$
$$\times \{[v' - (q + k)^2 - i\epsilon]^{-1}[w' - (q - k)^2 - i\epsilon]^{-1}$$
$$- (v' - m^2)^{-1}(w' - \mu^2)^{-1}\}K_0(v', w'; m^2, \mu^2). \quad (3.12)$$

On account of (2.7) and (3.2), (3.12) is rewritten as

$$\tilde{L}^{(2)}(s, t) = \int d^4q\psi(p, q, k)$$
$$\times [M^{(1)}([q + k]^2, [q - k]^2) - M^{(1)}(m^2, \mu^2)]. \quad (3.13)$$

Thus, by the definition (3.5) with (3.4), we obtain $\tilde{L}^{(2)}(s, t) = L^{(2)}(s, t)$. Similarly, in general case, (3.8), (2.37), (3.9), and (3.11) yield

$$\tilde{L}^{(n+1)}(s, t) = \int dv' \int dw' \int d^4q\psi(p, q, k)$$
$$\times \{[v' - (q + k)^2 - i\epsilon]^{-1}[w' - (q - k)^2 - i\epsilon]^{-1}$$
$$- (v' - m^2)^{-1}(w' - \mu^2)^{-1}\} \int dv'' \int dw''$$
$$\times [K_0(v', w'; v'', w'') - K_0(v', w'; v'', \mu^2)$$
$$- K_0(v', w'; m^2, w'')](T^{n-2}K)(v'', w''; m^2, \mu^2). \quad (3.14)$$

The assumption of induction leads to

$$\tilde{L}^{(n+1)}(s, t) = \int d^4q\psi(p, q, k)$$
$$\times \{L^{(n)}([q + k]^2, [q - k]^2) - L^{(n)}(m^2, \mu^2)\}. \quad (3.15)$$

Thus we have $\tilde{L}^{(n+1)}(s, t) = L^{(n+1)}(s, t)$ because of (3.5). Q.E.D.

Now, from (2.37) and (2.25) we obtain

$$\int ds' \int dt' \frac{(T^nK)(s', t'; m^2, \mu^2)}{(s' - s - i\epsilon)(t' - t - i\epsilon)}$$
$$= \int ds' \int dt' \int dv' \int dw' \frac{K_0(s', t'; v', w') - K_0(s', t'; v', \mu^2) - K_0(s', t'; m^2, w')}{(s' - s - i\epsilon)(t' - t - i\epsilon)(s' - m^2)(t' - \mu^2)} (T^{n-1}K)(v', w'; m^2, \mu^2)$$
$$= (m^2 - s - i\epsilon)^{-1}(\mu^2 - t - i\epsilon)^{-1} \int ds' \int dt' \int dv' \int dw'[(s' - s - i\epsilon)^{-1} - (s' - m^2)^{-1}]$$
$$\times [(t' - t - i\epsilon)^{-1} - (t' - \mu^2)^{-1}][K_0(s', t'; v', w') - K_0(s', t'; v', \mu^2) - K_0(s', t'; m^2, w')]$$
$$\times (T^{n-1}K)(v', w'; m^2, \mu^2) = \frac{L^{(n+1)}(s, t) - L^{(n+1)}(s, \mu^2) - L^{(n+1)}(m^2, t) + L^{(n+1)}(m^2, \mu^2)}{(m^2 - s - i\epsilon)(\mu^2 - t - i\epsilon)}, \quad (3.16)$$

where we have used (3.8) in the last step. It is easy to show

$$\int ds' \int dt' \frac{K(s', t'; m^2, \mu^2)}{(s' - s - i\epsilon)(t' - t - i\epsilon)} = \frac{L^{(1)}(s, t) - L^{(1)}(s, \mu^2) - L^{(1)}(m^2, t) + L^{(1)}(m^2, \mu^2)}{(m^2 - s - i\epsilon)(\mu^2 - t - i\epsilon)} \quad (3.17)$$

by using (2.25), (2.7), (3.2), and (3.4). Thus, from (2.36), (3.16), and (3.17), we obtain

$$\int ds' \int dt' \frac{\varphi_{12}(s', t')}{(s' - s - i\epsilon)(t' - t - i\epsilon)}$$
$$= (m^2 - s - i\epsilon)^{-1}(\mu^2 - t - i\epsilon)^{-1}$$
$$\times \sum_{n=1}^{\infty} [L^{(n)}(s, t) - L^{(n)}(s, \mu^2)$$
$$- L^{(n)}(m^2, t) + L^{(n)}(m^2, \mu^2)]. \quad (3.18)$$

Similarly, the use of (3.8) leads to

$$\int ds' \int dt' \frac{\varphi_1(s')\delta(t' - \mu^2)}{(s' - s - i\epsilon)(t' - t - i\epsilon)}$$
$$= \frac{\sum_{n=1}^{\infty} [L^{(n)}(s, \mu^2) - L^{(n)}(m^2, \mu^2)]}{(m^2 - s - i\epsilon)(\mu^2 - t - i\epsilon)}, \quad (3.19)$$

$$\int ds' \int dt' \frac{\varphi_2(t')\delta(s' - m^2)}{(s' - s - i\epsilon)(t' - t - i\epsilon)}$$
$$= \frac{\sum_{n=1}^{\infty} [L^{(n)}(m^2, t) - L^{(n)}(m^2, \mu^2)]}{(m^2 - s - i\epsilon)(\mu^2 - t - i\epsilon)}. \quad (3.20)$$

The eigenvalue equation (2.38) now becomes

$$1 = \sum_{n=1}^{\infty} L^{(n)}(m^2, \mu^2). \tag{3.21}$$

From (2.4), (2.17), and (3.18) $\sim$ (3.21), the invariant solutions are written as

$$f(s, t) = (m^2 - s - i\epsilon)^{-1} \times (\mu^2 - t - i\epsilon)^{-1} \sum_{n=1}^{\infty} L^{(n)}(s, t). \tag{3.22}$$

Finally, we must analytically continue the results (3.22) and (3.21) to $u = M^2$, where M stands for the physical mass of the bound state. As is seen from (3.6), the truncated vertex function has the same analyticity as the usual vertex function. Therefore, according to our previous result,[9] $\sum_{n=1}^{\infty} L^{(n)}(s, t)$ can be represented as

$$\int_0^1 \int_0^1 \int_0^1 \delta\left(1 - \sum_{i=1}^{3} z_i\right) dz_1\, dz_2\, dz_3 \times \int_0^{\infty} d\alpha \frac{\rho(\alpha; z_1, z_2, z_3)}{\alpha - z_1 s - z_2 t - z_3 u - i\epsilon}. \tag{3.23}$$

Here $\rho(\alpha; z_1 z_2 z_3)$ vanishes unless

$$\alpha \geq \text{Max}\,[M_s^2 z_1 + M_t^2 z_2 + (M_s - M_t)^2 z_3, \quad M_s^2 z_1 + (M_s - M_u)^2 z_2 + M_u^2 z_3, \quad (M_t - M_u)^2 z_1 + M_t^2 z_2 + M_u^2 z_3], \tag{3.24}$$

where M_s, M_t, and M_u are the lowest intermediate-state masses (or their lower bounds) in the respective channels, and it is assumed that they satisfy the triangular inequalities. In the present case, since

$$M_s = m + \kappa, \quad M_t = \mu + \kappa, \quad M < M_u \leq m + \mu, \tag{3.25}$$

(3.23) is analytic for

$$u < M_u^2 \tag{3.26}$$

if s and t satisfy

$$zs + (1 - z)t < \text{Max}\,[(m + \kappa)^2 z + (m + \kappa - M_u)^2(1 - z), \quad (\mu + \kappa - M_u)^2 z + (\mu + \kappa)^2(1 - z)], \tag{3.27}$$

for any z satisfying $0 \leq z \leq 1$. We can therefore continue $f(s, t)$ to $u = M^2$ for such s and t. For other values of s and t, we should continue it after the above continuation with respect to u. In this case, of course, $f(s, t)$ may have an absorptive part.

[9] N. Nakanishi, Progr. Theoret. Phys. (Kyoto), Suppl. **18**, 1 (1961), Sec. 16.

Thus we obtain the invariant solutions of the Bethe–Salpeter equation for the physical mass.

IV. REMARKS

In the above argument we have been indifferent to the convergence of the series. If we replace K by xK where x is a parameter, these series will be convergent and the rearrangement in (2.36) will be justified when $|x|$ is small. The correct result will then be obtained from (3.22) and (3.21) by continuing them to $x = 1$. In ladder approximation, (2.27) is absolutely convergent for all λ because of the θ functions in (2.9).

The series of the truncated vertex functions appearing in (3.22) is formally factorized as

$$\sum_{n=1}^{\infty} L^{(n)}(s, t) = \left[\sum_{n=1}^{\infty} M^{(n)}(s, t)\right] \times \left[1 - \sum_{n=1}^{\infty} L^{(n)}(m^2, \mu^2)\right]. \tag{4.1}$$

Hence $f(s, t)$ would vanish identically because of (3.21). However, this is naturally not true, because $f(s, t)$ necessarily contains a nonvanishing pole term as is seen from (2.4) and (2.17). Indeed, if (4.1) is true for $s = m^2$ and $t = \mu^2$, it cannot be consistent with (3.21) unless $\sum_{n=1}^{\infty} M^{(n)}(m^2, \mu^2)$ is infinite. Thus the series $\sum_{n=1}^{\infty} M^{(n)}(s, t)$ should be divergent for the values of λ satisfying (3.21).[10] It will be quite natural that the usual perturbation series becomes divergent when the coupling of the particles becomes so strong that they can be bound.

Blankenbecler and Cook[11] derived an equation determining the coupling constant from an unsubtracted dispersion relation. Their equation can be understood as the lowest-order approximation of our eigenvalue equation (3.21).

In our approach, the assumption $\kappa \neq 0$ is very essential. For $\kappa \to 0$, our solutions do not smoothly tend to the solutions for $\kappa = 0$. To see this, consider a special case of the ladder approximation in which $m = \mu$ and $k = 0$ (i.e., $s = t$). When $\kappa = 0$, Wick's simplest solution[3] reads

$$f(s, s) = (m^2 - s - i\epsilon)^{-3}, \tag{4.2}$$

with $\lambda = 2m^2$. Thus this solution has a triple pole, while our solutions for $\kappa \neq 0$ have only a double

[10] Namely, (4.1) shows that $(m^2 - s)(\mu^2 - t)f(s, t)$ is essentially nothing but the residue of the usual vertex function at a pole in the λ plane.

[11] R. Blankenbecler and L. F. Cook, Jr., Phys. Rev. **119**, 1745 (1960).

pole $(m^2 - s - i\epsilon)^{-2}$. It has been pointed out recently by Schroer[12] that a similar situation does occur also in the case of one-body propagators.

The stability condition for the bound state played an important role in Wick's method.[3] In our approach we used it in the final step. If $M > m + \mu$, we could not analytically continue the result to $u = M^2$.

Our formalism may equally apply to the case of the spinor–spinor system. But in this case the integral of $L^{(n)}(s, t)$ is divergent. This difficulty corresponds to Goldstein's one.[2] Extension of our approach to noninvariant solutions will be an important next task.

[12] B. Schroer (to be published).

ACKNOWLEDGMENTS

The author would like to express his sincere thanks to Dr. J. W. Calkin and Dr. Y. Shimamoto for their kind hospitality extended to him at the Brookhaven National Laboratory.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 4, NUMBER 10 OCTOBER 1963

Remarks on the Double Dispersion Approach to the Bethe–Salpeter Equation*

NOBORU NAKANISHI†
Applied Mathematics Division, Brookhaven National Laboratory, Upton, New York
(Received 25 September 1962)

The following remarks are made on the applicability of the double dispersion approach to the Bethe–Salpeter equation introduced previously. (1) Any invariant solution of the Bethe–Salpeter equation in ladder approximation satisfies the double dispersion representation when the total energy–momentum is spacelike. (2) There are some exceptional invariant solutions which are not given by the previous method in the equal-mass case, but the existence of such solutions is very unlikely in the unequal-mass case. (3) In the case of the general separated kernel the previous results give the correct solutions even if the kernel does not reproduce the double dispersion representation.

I. INTRODUCTION

In a previous work,[1] we have obtained the invariant solutions of the general Bethe–Salpeter equation for bound states of two scalar particles. In this work it was essential to use the double dispersion representation of the Bethe–Salpeter amplitude for $u < 0$, where $u \equiv 4(k_0^2 - \mathbf{k}^2)$ is the square of the total energy–momentum. The purpose of the present paper is to investigate to what extent the double dispersion approach is applicable.

In Sec. II it will be shown by using Ida and Maki's method[2] that all (normal and abnormal) invariant solutions of the Bethe–Salpeter equation in ladder approximation satisfy the double dispersion representation for $u < 0$. In Sec. III we shall discuss some exceptional solutions which were not dealt with in I. In the final section it will be demonstrated that in the case of the general separated kernel our previous solutions are correct even if the kernel does not reproduce the double dispersion representation.

Throughout the present paper, the notation in I will be used unless otherwise indicated.

II. DOUBLE DISPERSION REPRESENTATION

In this section it will be investigated in ladder approximation whether or not the possibility of the double dispersion representation for $u < 0$ is an additional requirement to the invariant solutions.

Theorem 1. In a Bethe–Salpeter equation

$$\tilde{f}(p, k) = [m^2 - (p + k)^2 - i\epsilon]^{-1}[\mu^2 - (p - k)^2 - i\epsilon]^{-1}(\lambda/\pi^2 i) \int d^4q \frac{\tilde{f}(q, k)}{\kappa^2 - (p - q)^2 - i\epsilon}, \quad (2.1)$$

the eigenvalues are discrete and positive definite when $u \equiv 4k^2 < 0$.

Proof: Since $u < 0$, k is spacelike. Hence we can

* This work done under the auspices of the U. S. Atomic Energy Commission.
† On leave of absence from the Institute for Advanced Study, Princeton, N. J.
[1] N. Nakanishi, J. Math. Phys. 4, 1229 (1963). This paper is referred to as I.
[2] M. Ida and K. Maki, Progr. Theoret. Phys. (Kyoto) 26, 470 (1961).

pole $(m^2 - s - i\epsilon)^{-2}$. It has been pointed out recently by Schroer[12] that a similar situation does occur also in the case of one-body propagators.

The stability condition for the bound state played an important role in Wick's method.[3] In our approach we used it in the final step. If $M > m + \mu$, we could not analytically continue the result to $u = M^2$.

Our formalism may equally apply to the case of the spinor–spinor system. But in this case the integral of $L^{(n)}(s, t)$ is divergent. This difficulty corresponds to Goldstein's one.[2] Extension of our approach to noninvariant solutions will be an important next task.

[12] B. Schroer (to be published).

ACKNOWLEDGMENTS

The author would like to express his sincere thanks to Dr. J. W. Calkin and Dr. Y. Shimamoto for their kind hospitality extended to him at the Brookhaven National Laboratory.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 4, NUMBER 10 OCTOBER 1963

Remarks on the Double Dispersion Approach to the Bethe–Salpeter Equation*

NOBORU NAKANISHI†
Applied Mathematics Division, Brookhaven National Laboratory, Upton, New York
(Received 25 September 1962)

The following remarks are made on the applicability of the double dispersion approach to the Bethe–Salpeter equation introduced previously. (1) Any invariant solution of the Bethe–Salpeter equation in ladder approximation satisfies the double dispersion representation when the total energy–momentum is spacelike. (2) There are some exceptional invariant solutions which are not given by the previous method in the equal-mass case, but the existence of such solutions is very unlikely in the unequal-mass case. (3) In the case of the general separated kernel the previous results give the correct solutions even if the kernel does not reproduce the double dispersion representation.

I. INTRODUCTION

In a previous work,[1] we have obtained the invariant solutions of the general Bethe–Salpeter equation for bound states of two scalar particles. In this work it was essential to use the double dispersion representation of the Bethe–Salpeter amplitude for $u < 0$, where $u \equiv 4(k_0^2 - \mathbf{k}^2)$ is the square of the total energy–momentum. The purpose of the present paper is to investigate to what extent the double dispersion approach is applicable.

In Sec. II it will be shown by using Ida and Maki's method[2] that all (normal and abnormal) invariant solutions of the Bethe–Salpeter equation in ladder approximation satisfy the double dispersion representation for $u < 0$. In Sec. III we shall discuss some exceptional solutions which were not dealt with in I. In the final section it will be demonstrated that in the case of the general separated kernel our previous solutions are correct even if the kernel does not reproduce the double dispersion representation.

Throughout the present paper, the notation in I will be used unless otherwise indicated.

II. DOUBLE DISPERSION REPRESENTATION

In this section it will be investigated in ladder approximation whether or not the possibility of the double dispersion representation for $u < 0$ is an additional requirement to the invariant solutions.

Theorem 1. In a Bethe–Salpeter equation

$$\tilde{f}(p, k) = [m^2 - (p + k)^2 - i\epsilon]^{-1}[\mu^2 - (p - k)^2 - i\epsilon]^{-1}(\lambda/\pi^2 i) \int d^4q \frac{\tilde{f}(q, k)}{\kappa^2 - (p - q)^2 - i\epsilon}, \quad (2.1)$$

the eigenvalues are discrete and positive definite when $u \equiv 4k^2 < 0$.

Proof: Since $u < 0$, k is spacelike. Hence we can

* This work done under the auspices of the U. S. Atomic Energy Commission.
† On leave of absence from the Institute for Advanced Study, Princeton, N. J.
[1] N. Nakanishi, J. Math. Phys. 4, 1229 (1963). This paper is referred to as I.
[2] M. Ida and K. Maki, Progr. Theoret. Phys. (Kyoto) 26, 470 (1961).

take a frame of $k_0 = 0$. Following Wick,[3] we analytically continue (2.1) as $p_0 \to ip_4$ and $q_0 \to iq_4$. Then

$$\tilde{f}(ip_4, \mathbf{p}, 0, \mathbf{k}) = [m^2 + (\mathbf{p}+\mathbf{k})^2 + p_4^2]^{-1}[\mu^2 + (\mathbf{p}-\mathbf{k})^2 + p_4^2]^{-1}(\lambda/\pi^2)\int d\mathbf{q}\, dq_4 \times \frac{\tilde{f}(iq_4, \mathbf{q}, 0, \mathbf{k})}{\kappa^2 + (\mathbf{p}-\mathbf{q})^2 + (p_4 - q_4)^2}. \quad (2.2)$$

By a transformation

$$g(p_4, \mathbf{p}, \mathbf{k}) \equiv [m^2 + (\mathbf{p}+\mathbf{k})^2 + p_4^2]^{\frac{1}{2}}[\mu^2 + (\mathbf{p}-\mathbf{k})^2 + p_4^2]^{\frac{1}{2}}\tilde{f}(ip_4, \mathbf{p}, 0, \mathbf{k}), \quad (2.3)$$

the kernel becomes symmetric and of positive type.[4] It is easy to make the integration range a finite one.[3] Therefore the eigenvalues are discrete and positive definite. Q.E.D.

For $u > 0$, these properties of the eigenvalues are proved only in the case $m = \mu$.[3,2]

Hereafter we will consider invariant (S-wave) solutions only. Namely, we consider the case in which $\tilde{f}(p, k)$ can be written as $f(s, t)$, where $s \equiv (p+k)^2$ and $t \equiv (p-k)^2$. Then the kernel becomes

$$[m^2 - s - i\epsilon]^{-1}[\mu^2 - t - i\epsilon]^{-1}(\pi^2 i)^{-1}\int d^4q \times \frac{\delta(s' - [q+k]^2)\delta(t' - [q-k]^2)}{\kappa^2 - (p-q)^2 - i\epsilon}. \quad (2.4)$$

The eigenvalues of (2.4) are, of course, discrete and positive definite for $u < 0$, because they are necessarily eigenvalues of (2.1).

Theorem 2. The trace of (2.4) is equal to $\bar{M}^{(1)}(m^2, \mu^2)$, where $\bar{M}^{(1)}(s, t)$ stands for the third-order vertex function, namely[5]

$$\bar{M}^{(1)}(s, t) \equiv (\pi^2 i)^{-1}\int d^4q\, [m^2 - (q+k)^2 - i\epsilon]^{-1} \times [\mu^2 - (q-k)^2 - i\epsilon]^{-1}[\kappa^2 - (p-q)^2 - i\epsilon]^{-1}. \quad (2.5)$$

A proof of this theorem will be given in Appendix A.

Now, we will use Ida and Maki's argument.[2] Let λ_i be the eigenvalues of λ corresponding to the invariant solutions. Then the total sum of λ_i^{-1} is equal to the trace of (2.4) according to Mercer's theorem, namely

$$\sum_i \lambda_i^{-1} = \bar{M}^{(1)}(m^2, \mu^2). \quad (2.6)$$

The totality of the eigenvalues corresponding to the solutions which satisfy the double dispersion representation is naturally a subset of $\{\lambda_i\}$. Since $\lambda_i > 0$ and the trace is invariant under similarity transformations, all the solutions satisfy the double dispersion representation for $u < 0$ if, and only if, the trace of the kernel $\mathbf{K}(s', t'; s'', t'')$ in the equation for the double spectral function is equal to (2.6).

From I, the kernel $\mathbf{K}(s', t'; s'', t'')$ reads

$$\begin{aligned}\mathbf{K}(s', t'; s'', t'') &\equiv \delta(s' - m^2)\delta(t' - \mu^2)\int dv \int dw \frac{\bar{K}_0(v, w; s'', t'')}{(v - m^2)(w - \mu^2)} \\ &- \delta(t' - \mu^2)(s' - m^2)^{-1}\int dw \frac{\bar{K}_0(s', w; s'', t'')}{w - \mu^2} \\ &- \delta(s' - m^2)(t' - \mu^2)^{-1}\int dv \frac{\bar{K}_0(v, t'; s'', t'')}{v - m^2} \\ &+ (s' - m^2)^{-1}(t' - \mu^2)^{-1}\bar{K}_0(s', t'; s'', t''). \quad (2.7)\end{aligned}$$

Here in ladder approximation $\bar{K}_0(s', t'; s'', t'')$ vanishes unless

$$s'^{\frac{1}{2}} \geq s''^{\frac{1}{2}} + \kappa \quad \text{and} \quad t'^{\frac{1}{2}} \geq t''^{\frac{1}{2}} + \kappa; \text{ and}$$

$$\bar{M}^{(1)}(s, t) = \int ds' \int dt' \frac{\bar{K}_0(s', t'; m^2, \mu^2)}{(s' - s - i\epsilon)(t' - t - i\epsilon)}. \quad (2.8)$$

Therefore the trace of $\mathbf{K}(s', t'; s'', t'')$ is

$$\begin{aligned}\int ds' \int dt'\, \mathbf{K}(s', t'; s', t') &= \int dv \int dw \frac{\bar{K}_0(v, w; m^2, \mu^2)}{(v - m^2)(w - \mu^2)} \\ &= \bar{M}^{(1)}(m^2, \mu^2). \quad (2.9)\end{aligned}$$

From (2.6) and (2.9) we can conclude that all the invariant solutions of (2.1) satisfy the double dispersion representation for $u < 0$.

The above reasoning is not adequate for the case $\kappa = 0$, because $\bar{M}^{(1)}(m^2, \mu^2)$ then contains an infrared divergence. In Appendix B, using Cutkosky's solutions[6] we shall prove that the same is true also for $\kappa = 0$.

III. EXCEPTIONAL SOLUTIONS

Because of (2.7) the double spectral function $\varphi(s', t')$ must have the following spectrum:

[3] G. C. Wick, Phys. Rev. **96**, 1124 (1954).

[4] A kernel $K(p, q)$ is said to be of positive type if $\int dp \int dq\, \varphi(p)\, K(p, q)\varphi(q) \geq 0$ for any real function $\varphi(p)$. Hence, in the present case, it is sufficient to show that $[\kappa^2 + (\mathbf{p}-\mathbf{q})^2 + (p_4 - q_4)^2]^{-1}$ is of positive type. According to Bochner's theorem, this statement is equivalent to the positive definiteness of the Fourier transform of $(\kappa^2 + \mathbf{p}^2 + p_4^2)^{-1}$, which is given by $\kappa r^{-1}K_1(\kappa r) > 0$ with $r \equiv (\mathbf{x}^2 + x_4^2)^{\frac{1}{2}}$, where K_1 denotes a modified Bessel function.

[5] The bar means the excursion of the factor λ.

[6] R. E. Cutkosky, Phys. Rev. **96**, 1135 (1954).

$$\varphi(s', t') = a\delta(s' - m^2)\delta(t' - \mu^2) + \varphi_1(s')\delta(t' - \mu^2) + \varphi_2(t')\delta(s' - m^2) + \varphi_{12}(s', t'), \quad (3.1)$$

where $\varphi_1(s')$, $\varphi_2(t')$, and $\varphi_{12}(s', t')$ vanish unless the respective explicit variables satisfy $s' \geq (m + \kappa)^2$ and $t' \geq (\mu + \kappa)^2$. In I, we have found the solutions in the case $a \neq 0$. If $a = 0$, the equations for the spectral functions become

$$0 = \int ds' \int dt' \, \varphi_{12}(s', t'), \quad (3.2)$$

$$\varphi_{12}(s', t') = \int ds'' \, K(s', t'; s'', \mu^2)\varphi_1(s'') + \int dt'' \, K(s', t'; m^2, t'')\varphi_2(t'') + \int ds'' \int dt'' \, K(s', t'; s'', t'') \times \varphi_{12}(s'', t''), \quad (3.3)$$

$$\varphi_1(s') = -\int dt' \int ds'' \, R(s', t'; s'', \mu^2)\varphi_1(s'') - \int dt' \int dt'' \, R(s', t'; m^2, t'')\varphi_2(t''), \quad (3.4)$$

$$\varphi_2(t') = -\int ds' \int ds'' \, R(s', t'; s'', \mu^2)\varphi_1(s'') - \int ds' \int dt'' \, R(s', t'; m^2, t'')\varphi_2(t''), \quad (3.5)$$

where

$$K(s', t'; s'', t'') \equiv (s' - m^2)^{-1}(t' - \mu^2)^{-1}K_0(s', t'; s'', t''), \quad (3.6)$$

and

$$R \equiv \sum_{n=1}^{\infty} K^n. \quad (3.7)$$

While (3.3) is an inhomogeneous (Volterra) integral equation, (3.4) and (3.5) are homogeneous Fredholm integral equations, which cannot easily be solved. These equations are, of course, satisfied only for some particular values of λ.

Now, let us consider the case in which the two particles m and μ are identical ($m = \mu$). Since $\mathbf{K}(s', t'; s'', t'') = \mathbf{K}(t', s'; t'', s'')$ in this case, the solution $\varphi(s', t')$ is either symmetric or antisymmetric with respect to s' and t'. We define the antisymmetric kernel by

$$\mathbf{K}^{(-)}(s', t'; s'', t'') \equiv \tfrac{1}{2}[\mathbf{K}(s', t'; s'', t'') - \mathbf{K}(s', t'; t'', s'')]. \quad (3.8)$$

Then on account of (2.7) its trace is

$$\int ds' \int dt' \, \mathbf{K}^{(-)}(s', t'; s', t') = \int ds' \int dt' \, \frac{\bar{K}_0(s', t'; m^2, s')}{(s' - m^2)(t' - m^2)}, \quad (3.9)$$

which is nonvanishing, at least in ladder approximation. Thus antisymmetric solutions must exist in ladder approximation.[2] They automatically satisfy (3.2), and hence belong to the case $a = 0$. Since $\varphi_1(s') = -\varphi_2(s')$, (3.4) and (3.5) become identical. Thus the eigenvalues are determined by a homogeneous equation

$$\varphi_1(s') = \int ds'' \, A(s', s'')\varphi_1(s''), \quad (3.10)$$

where

$$A(s', s'') \equiv -\int dt' \, R(s', t'; s'', m^2) + \int dt' \, R(s', t'; m^2, s''). \quad (3.11)$$

On the other hand, if $m \neq \mu$, (3.2) becomes an additional requirement which will be independent of the condition on the eigenvalues imposed by (3.4) and (3.5). It is, therefore, very unlikely that there would exist the solutions for $a = 0$ in the unequal-mass case. Hence the antisymmetric solutions for $m = \mu$ may be obtained from the unequal-mass case by a limiting procedure $\mu \to m$.[7]

Finally, we remark that an extremely exceptional situation happens if

$$K(s', t'; m^2, \mu^2) \equiv 0. \quad (3.12)$$

Since in this case the inhomogeneous term vanishes even for $a \neq 0$, the equations for the spectral functions become (3.3), (3.4), and (3.5). However, since we have no restriction corresponding to (3.2), there will exist some solutions in general. Such an example will be given in the next section. In practical uses, this exceptional case should be avoided by a small variation of m^2 or μ^2.

IV. SEPARATED KERNEL

According to the result of I, the invariant solutions ($a \neq 0$) of the general Bethe–Salpeter equation,

$$\tilde{f}(p, k) = [m^2 - (p + k)^2 - i\epsilon]^{-1}[\mu^2 - (p - k)^2 - i\epsilon]^{-1} \times \int d^4q \, H(p, q, k)\tilde{f}(q, k), \quad (4.1)$$

[7] A similar situation occurs also in Cutkosky's solutions when $m = \mu$ and $k_\mu = 0$.

are given by

$$\bar{f}(p, k) \equiv f(s, t) = a[m^2 - s - i\epsilon]^{-1}[\mu^2 - t - i\epsilon]^{-1} \times \sum_{n=1}^{\infty} L^{(n)}(s, t), \quad (4.2)$$

with

$$1 = \sum_{n=1}^{\infty} L^{(n)}(m^2, \mu^2). \quad (4.3)$$

Here $L^{(n)}(s, t)$ is a truncated vertex function, which was introduced in I.

Now, as is well known, (4.1) is solved very easily if $H(p, q, k)$ is a separated kernel, i.e.,

$$H(p, q, k) \equiv \sum_{i=1}^{N} I_i(s, t) J_i(q, k), \quad (4.4)$$

where $I_1(s, t), \cdots, I_N(s, t)$ are assumed to be linearly independent. The solutions are

$$f(s, t) = \frac{\sum_i c_i I_i(s, t)}{(m^2 - s - i\epsilon)(\mu^2 - t - i\epsilon)}, \quad (4.5)$$

where

$$\sum_i (\delta_{ij} - A_{ij}) c_i = 0, \quad (4.6)$$

with

$$A_{ij} \equiv \int d^4q \times \frac{J_i(q, k) I_j([q + k]^2, [q - k]^2)}{[m^2 - (q + k)^2 - i\epsilon][\mu^2 - (q - k)^2 - i\epsilon]}. \quad (4.7)$$

In order that (4.6) has nontrivial c_i, we must have an equation

$$\det(\delta_{ij} - A_{ij}) = 0, \quad (4.8)$$

which determines the eigenvalues.

It will be instructive to calculate these solutions by using (4.2) together with (4.3). For this purpose we introduce the following matrix and vector notations:

$$\begin{aligned} E &\equiv (\delta_{ij}), \\ A &\equiv (A_{ij}), \\ \mathbf{I}(s, t) &\equiv (I_1(s, t), \cdots, I_N(s, t)), \\ \mathbf{I} &\equiv \mathbf{I}(m^2, \mu^2), \\ \mathbf{b} &\equiv (b_1, b_2, \cdots, b_N)^{\mathrm{T}}, \end{aligned} \quad (4.9)$$

where

$$b_i \equiv \int d^4q \frac{J_i(q, k)}{[m^2 - (q + k)^2 - i\epsilon][\mu^2 - (q - k)^2 - i\epsilon]}, \quad (4.10)$$

and the superscript T means "transposed". We assume $\mathbf{b} \neq 0$. From Eqs. (3.2)–(3.5) of I we have

$$M^{(n)}(s, t) = \mathbf{I}(s, t) A^{n-1} \mathbf{b}, \quad (4.11)$$

$$L^{(n)}(s, t) = \mathbf{I}(s, t)(A - \mathbf{b}\cdot\mathbf{I})^{n-1}\mathbf{b}, \quad (4.12)$$

where $\mathbf{b}\cdot\mathbf{I}$ is naturally an $N \times N$ matrix. Hence (4.2) is written as

$$f(s, t) = \frac{a\mathbf{I}(s, t)\left[\sum_{n=0}^{\infty} (A - \mathbf{b}\cdot\mathbf{I})^n\right]\mathbf{b}}{(m^2 - s - i\epsilon)(\mu^2 - t - i\epsilon)}. \quad (4.13)$$

The series will be convergent at least when the coupling constant is small. For its large values, we should analytically continue $f(s, t)$ from its small values. Then we obtain

$$f(s, t) = \mathbf{I}(s, t)\mathbf{c}/[(m^2 - s - i\epsilon)(\mu^2 - t - i\epsilon)], \quad (4.14)$$

with

$$\mathbf{c} \equiv a(E - A + \mathbf{b}\cdot\mathbf{I})^{-1}\mathbf{b}. \quad (4.15)$$

Likewise (4.3) leads to

$$1 = \mathbf{I}(E - A + \mathbf{b}\cdot\mathbf{I})^{-1}\mathbf{b}. \quad (4.16)$$

The use of (4.16) leads to

$$(E - A)\mathbf{c} = 0, \quad (4.17)$$

because

$$\begin{aligned} (E - A)\mathbf{c} &= a(E - A + \mathbf{b}\cdot\mathbf{I} - \mathbf{b}\cdot\mathbf{I}) \\ &\quad \times (E - A + \mathbf{b}\cdot\mathbf{I})^{-1}\mathbf{b} \\ &= a\{\mathbf{b} - \mathbf{b}[\mathbf{I}(E - A + \mathbf{b}\cdot\mathbf{I})^{-1}\mathbf{b}]\} \\ &= 0. \end{aligned} \quad (4.18)$$

The assumption $\mathbf{b} \neq 0$ implies $\mathbf{c} \neq 0$. Hence from (4.17) we have

$$\det(E - A) = 0. \quad (4.19)$$

Since (4.16) is an algebraic equation of nth degree with respect to A_{ij}, it must be equivalent to (4.19). Thus the final results (4.14), (4.17), and (4.19) completely coincide with (4.5), (4.6), and (4.8), respectively.

In the above reasoning, we did not assume the possibility of the double dispersion representations for $I_i(s, t)$. This suggests that our solutions (4.2) with (4.3) will have very wide applicability. Indeed, it is easily checked that (4.2) with (4.3) always satisfies (4.1) at least formally.

From (4.11), the series of the usual vertex functions formally becomes

$$\sum_{n=1}^{\infty} M^{(n)}(s, t) = \mathbf{I}(s, t)(E - A)^{-1}\mathbf{b}, \quad (4.20)$$

but $(E - A)^{-1}$ does not exist because of (4.19). Thus this series is indeed divergent[8] for bound states as was noticed in I.

If $\mathbf{b} = 0$, all $L^{(n)}(s, t)$ vanish, and hence (4.2) with (4.3) no longer gives correct solutions. This case is the extremely exceptional case stated in the end of the preceding section. Indeed, if

$$\mathrm{I}(s, t) = \int ds' \int dt' \frac{\delta(s', t')}{(s' - s - i\epsilon)(t' - t - i\epsilon)}, \tag{4.21}$$

we have

$$K(s', t', m^2, \mu^2) = (s' - m^2)^{-1}(t' - \mu^2)^{-1}\delta(s', t')\mathbf{b} = 0. \tag{4.22}$$

ACKNOWLEDGMENTS

The author would like to express his sincere thanks to Dr. J. W. Calkin and Dr. Y. Shimamoto for their kind hospitality extended to him at the Brookhaven National Laboratory.

APPENDIX A: TRACE OF THE S-WAVE KERNEL

We shall calculate the trace of (2.4). We take a special frame of $k = (0, k_1, 0, 0)$ in which $u = 4k^2 = -4k_1^2$. After Wick's analytic continuation,[3] we adopt the following variables:

$$\begin{aligned} p_1,\ r &\equiv (p_2^2 + p_3^2 + p_4^2)^{\frac{1}{2}}, \\ q_1,\ r' &\equiv (q_2^2 + q_3^2 + q_4^2)^{\frac{1}{2}}, \end{aligned} \tag{A1}$$

instead of the invariant squares s, t, s', t'. Then from (2.2) the S-wave kernel is given by[2]

$$\pi^{-1}[m^2 + (p_1 + k_1)^2 + r^2]^{-1}[\mu^2 + (p_1 - k_1)^2 + r^2]^{-1} \times \log \frac{(p_1 - q_1)^2 + (r + r')^2 + \kappa^2}{(p_1 - q_1)^2 + (r - r')^2 + \kappa^2}, \tag{A2}$$

which can be obtained also from (2.4) by transforming the variables. Hence the trace is

$$\begin{aligned} \sigma &\equiv \int_0^\infty dr \int_{-\infty}^\infty dp_1 \\ &\quad \times \frac{\pi^{-1} \log [(4r^2 + \kappa^2)/\kappa^2]}{[m^2 + r^2 + (p_1 + k_1)^2][\mu^2 + r^2 + (p_1 - k_1)^2]} \\ &= \int_0^1 dy \int_0^\infty dr \int_{-\infty}^\infty dp_1 \\ &\quad \times \frac{\pi^{-1} \log [(4r^2 + \kappa^2)/\kappa^2]}{[r^2 + p_1^2 + k_1^2 + 2(2y - 1)p_1 k_1 + ym^2 + (1 - y)\mu^2]^2}. \end{aligned} \tag{A3}$$

Transforming p_1 into $p_1 - (2y - 1)k_1$, we obtain

$$\sigma = \int_0^1 dy \int_0^\infty dr \int_{-\infty}^\infty dp_1 \times \frac{\pi^{-1} \log [(4r^2 + \kappa^2)/\kappa^2]}{[r^2 + p_1^2 + ym^2 + (1 - y)\mu^2 - y(1 - y)u]^2}. \tag{A4}$$

Ida and Maki's identity,[2]

$$\int_0^\infty dr \int_{-\infty}^\infty dp_1 \frac{\pi^{-1} \log [(4r^2 + \kappa^2)/\kappa^2]}{[r^2 + p_1^2 + \alpha]^2} = \int_0^1 dx \frac{1 - x}{x\kappa^2 + (1 - x)^2\alpha}, \qquad (\alpha > 0), \tag{A5}$$

leads to

$$\sigma = \int_0^1 (1 - x)\, dx \int_0^1 dy \{x\kappa^2 + (1 - x)^2 \times [ym^2 + (1 - y)\mu^2 - y(1 - y)u]\}^{-1}. \tag{A6}$$

Putting $x_1 = (1 - x)y$, $x_2 = (1 - x)(1 - y)$, and $x_3 = x$, we have

$$\sigma = \int_0^1 dx_1 \int_0^1 dx_2 \int_0^1 dx_3\, \delta(1 - x_1 - x_2 - x_3) \times [x_1 m^2 + x_2\mu^2 + x_3\kappa^2 - x_1x_3m^2 - x_2x_3\mu^2 - x_1x_2u]^{-1}, \tag{A7}$$

which is nothing but the Feynman parametric representation of $\bar{M}^{(1)}(m^2, \mu^2)$.

APPENDIX B: DOUBLE DISPERSION REPRESENTATIONS OF CUTKOSKY'S SOLUTIONS

We shall prove that all Cutkosky's invariant solutions satisfy the double dispersion representation when $u < (m - \mu)^2$. They are given by certain linear combinations of

$$f_n^k(s, t) = \int_{-1}^1 \frac{g_n^k(z)dz}{[\frac{1}{2}(1 + z)(m^2 - s) + \frac{1}{2}(1 - z)(\mu^2 - t) - i\epsilon]^{n+2}}, \qquad (n \geq 1, k < n). \tag{B1}$$

The functions g_n^k ($k = 1, 2, \cdots, n - 1$) are determined from $g_n^0 \equiv g_n$ by inhomogeneous integral equations. The functions g_n satisfy differential equations

$$g_n'' + 2(n - 1)z(1 - z^2)^{-1}g_n' + (1 - z^2)^{-1} \times \{-n(n - 1) + \lambda[\tfrac{1}{2}(1 + z)m^2 + \tfrac{1}{2}(1 - z)\mu^2 - \tfrac{1}{4}(1 - z^2)u]^{-1}\}g_n = 0, \tag{B2}$$

with $g_n(\pm 1) = 0$. Since (B2) is linear, the locations of singularities of any solution coincide with those of the coefficients. If $u < (m - \mu)^2$, therefore, the singularities of $g_n(z)$ are located only in $z \geq 1$ and

[8] It should be remarked that this infinity corresponds to the pole of the scattering amplitude for the particles m and μ.

$z \leq -1$. Since $g_n(z)$ is real and analytic in $-1 < z < 1$, branch cuts must lie in these regions because of the reflection principle. Since it is easily demonstrated that $z = \infty$ is a regular singular point, $g_n(z)$ satisfies a dispersion relation

$$g_n(z) = \sum_{\nu=1}^{N-1} a_\nu z^\nu + z^N \left[\int_{-\infty}^{-1} dz' + \int_1^\infty dz'\right] \frac{\rho(z')}{z'^N(z' - z)}. \tag{B3}$$

Substituting (B3) in the inhomogeneous equations for g_n^k, we can inductively show that g_n^k satisfies such a dispersion relation as (B3). Inserting the dispersion relation for g_n^k into (B1), and deforming the contour of z, we can easily see that $f_n^k(s, t)$ is analytic except for $s \geq m^2$ or $t \geq \mu^2$. The double dispersion representation may be obtained in the following way.

The following formulas are easily verified by direct calculations:

$$\int_{-1}^{1} dz\, [\alpha - \tfrac{1}{2}(1 + z)v - \tfrac{1}{2}(1 - z)w - i\epsilon]^{-1}$$

$$= 2 \int_\alpha^\infty dv' \int_\alpha^\infty dw' \frac{\delta(v' - w')}{(v' - v - i\epsilon)(w' - w - i\epsilon)}, \tag{B4}$$

$$\int_{-1}^{1} dz(z' - z)^{-1}[\alpha - \tfrac{1}{2}(1 + z)v - \tfrac{1}{2}(1 - z)w - i\epsilon]^{-1}$$

$$= \frac{z'}{|z'|} \int_\alpha^\infty dv' \int_\alpha^\infty dw' \times \frac{\delta(\alpha - \tfrac{1}{2}[1 + z']v' - \tfrac{1}{2}[1 - z']w')}{(v' - v - i\epsilon)(w' - w - i\epsilon)},$$

$(|z'| > 1)$. (B5)

In order to multiply the integrands of the left-hand sides by z^ν, we make

$$\left[\int_\alpha^\infty d\alpha \left(\frac{\partial}{\partial v} - \frac{\partial}{\partial w}\right)\right]^\nu \tag{B6}$$

operate on (B4) and (B5). After differentiating these formulas by α $(n + 1)$ times, we put $\alpha = 0$, $v = s - m^2$, $w = t - \mu^2$. An appropriate linear combination of the resulting formulas gives the double dispersion representation of (B1). Of course, no subtraction terms appear.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 4, NUMBER 10 OCTOBER 1963

Local Commutativity and the Analytic Continuation of the Wightman Function

YUKIO TOMOZAWA
Istituto Nazionale di Fisica Nucleare, Sezione di Pisa, Pisa, Italy
(Received 10 April 1963)

It is proved that the analytic continuation of the Wightman function, the vacuum expectation value of the product of field operators, due to local commutativity is single-valued in the union of the extended tubes which correspond to the Wightman functions obtained by permuting the order of the field operators in the product, and that the extended tubes, the union of them and the intersection of any two are simply connected.

1. INTRODUCTION

IN the systematic analysis of the frame of quantum field theory, the investigation of the analytic property of the vacuum expectation value, called the Wightman function (denoted W function hereafter), of product of field operators turned out to be important: Wightman has shown[1] that a set of analytic functions with certain properties, such as suitable invariance properties and boundedness, is equivalent to quantum field theory with certain axioms (see below), identifying the boundary values of these analytic functions with the W functions of the theory.

[1] A. S. Wightman, Phys. Rev. **101**, 860 (1956).

We take the following axioms[2] as the basis of the theory:

[2] For detailed discussion of these axioms and for further references, see A. S. Wightman in *Les Problèmes Mathématiques de la Théorie Quantique des Champs* (Colloque Internationaux du CNRS, Lille, 1957), p. 1; Cours de la Faculté des Sciences de l'Université de Paris (1957–58); Nuovo Cimento Suppl. **14**, 192 (1959); J. Indian Math. Soc. **24**, 625 (1960); R. Haag, Nuovo Cimento Suppl. **14**, 131 (1959); G. Källen in *Relations de Dispersion et Particules Elémentaires* (Hermann & Cie., Paris, 1960), p. 387; R. Jost in *Theoretical Physics in the Twentieth Century* (Interscience Publishers, Inc., New York, 1960), p. 107; in *Lectures on Field Theory and the Many Body Problem* (Academic Press Inc., New York, 1961), p. 127; H. Araki, Suppl. Progr. Theoret. Phys. (Kyoto) **18**, 83 (1961); R. Haag and B. Schroer, J. Math. Phys. **3**, 248 (1962); W. Schmidt and K. Baumann, Nuovo Cimento **4**, 860 (1956).

$z \leq -1$. Since $g_n(z)$ is real and analytic in $-1 < z < 1$, branch cuts must lie in these regions because of the reflection principle. Since it is easily demonstrated that $z = \infty$ is a regular singular point, $g_n(z)$ satisfies a dispersion relation

$$g_n(z) = \sum_{\nu=1}^{N-1} a_\nu z^\nu + z^N \left[\int_{-\infty}^{-1} dz' + \int_1^\infty dz'\right] \frac{\rho(z')}{z'^N(z' - z)}. \tag{B3}$$

Substituting (B3) in the inhomogeneous equations for g_n^k, we can inductively show that g_n^k satisfies such a dispersion relation as (B3). Inserting the dispersion relation for g_n^k into (B1), and deforming the contour of z, we can easily see that $f_n^k(s, t)$ is analytic except for $s \geq m^2$ or $t \geq \mu^2$. The double dispersion representation may be obtained in the following way.

The following formulas are easily verified by direct calculations:

$$\int_{-1}^{1} dz\, [\alpha - \tfrac{1}{2}(1 + z)v - \tfrac{1}{2}(1 - z)w - i\epsilon]^{-1}$$

$$= 2 \int_\alpha^\infty dv' \int_\alpha^\infty dw' \frac{\delta(v' - w')}{(v' - v - i\epsilon)(w' - w - i\epsilon)}, \tag{B4}$$

$$\int_{-1}^{1} dz(z' - z)^{-1}[\alpha - \tfrac{1}{2}(1 + z)v - \tfrac{1}{2}(1 - z)w - i\epsilon]^{-1}$$

$$= \frac{z'}{|z'|} \int_\alpha^\infty dv' \int_\alpha^\infty dw' \times \frac{\delta(\alpha - \tfrac{1}{2}[1 + z']v' - \tfrac{1}{2}[1 - z']w')}{(v' - v - i\epsilon)(w' - w - i\epsilon)},$$

$(|z'| > 1)$. (B5)

In order to multiply the integrands of the left-hand sides by z^ν, we make

$$\left[\int_\alpha^\infty d\alpha \left(\frac{\partial}{\partial v} - \frac{\partial}{\partial w}\right)\right]^\nu \tag{B6}$$

operate on (B4) and (B5). After differentiating these formulas by α $(n + 1)$ times, we put $\alpha = 0$, $v = s - m^2$, $w = t - \mu^2$. An appropriate linear combination of the resulting formulas gives the double dispersion representation of (B1). Of course, no subtraction terms appear.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 4, NUMBER 10 OCTOBER 1963

Local Commutativity and the Analytic Continuation of the Wightman Function

YUKIO TOMOZAWA
Istituto Nazionale di Fisica Nucleare, Sezione di Pisa, Pisa, Italy
(Received 10 April 1963)

It is proved that the analytic continuation of the Wightman function, the vacuum expectation value of the product of field operators, due to local commutativity is single-valued in the union of the extended tubes which correspond to the Wightman functions obtained by permuting the order of the field operators in the product, and that the extended tubes, the union of them and the intersection of any two are simply connected.

1. INTRODUCTION

IN the systematic analysis of the frame of quantum field theory, the investigation of the analytic property of the vacuum expectation value, called the Wightman function (denoted W function hereafter), of product of field operators turned out to be important: Wightman has shown[1] that a set of analytic functions with certain properties, such as suitable invariance properties and boundedness, is equivalent to quantum field theory with certain axioms (see below), identifying the boundary values of these analytic functions with the W functions of the theory.

[1] A. S. Wightman, Phys. Rev. **101**, 860 (1956).

We take the following axioms[2] as the basis of the theory:

[2] For detailed discussion of these axioms and for further references, see A. S. Wightman in *Les Problèmes Mathématiques de la Théorie Quantique des Champs* (Colloque Internationaux du CNRS, Lille, 1957), p. 1; Cours de la Faculté des Sciences de l'Université de Paris (1957–58); Nuovo Cimento Suppl. **14**, 192 (1959); J. Indian Math. Soc. **24**, 625 (1960); R. Haag, Nuovo Cimento Suppl. **14**, 131 (1959); G. Källen in *Relations de Dispersion et Particules Elémentaires* (Hermann & Cie., Paris, 1960), p. 387; R. Jost in *Theoretical Physics in the Twentieth Century* (Interscience Publishers, Inc., New York, 1960), p. 107; in *Lectures on Field Theory and the Many Body Problem* (Academic Press Inc., New York, 1961), p. 127; H. Araki, Suppl. Progr. Theoret. Phys. (Kyoto) **18**, 83 (1961); R. Haag and B. Schroer, J. Math. Phys. **3**, 248 (1962); W. Schmidt and K. Baumann, Nuovo Cimento **4**, 860 (1956).

(I) Invariance under the proper[3] inhomogeneous Lorentz group,

(II) Spectral condition, i.e., existence of the Hilbert space spanned by the physical state vectors, nonnegativity of the energy spectrum of these states, and the existence of the vacuum as the lowest energy state,

(III) Existence of field operators as temperate distribution operators,

(IV) Local commutativity, i.e., field operators commute or anticommute for spacelike separation.

From axioms (I)–(III) it follows that the W function

$$W^N_\nu(\xi_1, \cdots, \xi_N) = \langle 0 | \psi^{(0)}_{\nu_0}(x_0)\psi^{(1)}_{\nu_1}(x_1) \cdots \psi^{(N)}_{\nu_N}(x_N) | 0 \rangle, \quad (1)$$

where $\xi_j = x_j - x_{j-1}$ $(j = 1, 2, \cdots, N)$ $[\psi^{(j)}_{\nu_j}(x_j)$ are field operators, ν_j are the spin indices and ν stands for the set $\nu_0, \nu_1, \cdots, \nu_N]$, is a temperate distribution which is the boundary value of a function analytic in the forward tube $\mathfrak{I}_N = \{\{\zeta_j\}; \operatorname{Im} \zeta_j \in V_+\}$, where V_+ is the forward cone. The Bargmann–Hall–Wightman Theorem[4] enables us to enlarge the analyticity domain of the W function: $W^N_\nu(\zeta_1, \cdots, \zeta_N)$ is a single valued analytic function in the extended tube $\mathfrak{I}'_N = \{\{\zeta_j\}; \zeta_j = L_+(C)\zeta'_j, \{\zeta'_j\} \in \mathfrak{I}_N\}$, where $L_+(C)$ is the totality of the proper homogeneous complex Lorentz transformations (with determinant $+1$).

Local commutativity (axiom IV) then relates the W functions which correspond to various permutations of the field operators in the product, and gives us analytic continuation in the union

$$\bigcup_{g \in S_{N+1}} P(g)\,\mathfrak{I}'_N$$

of the extended tubes $P(g)\mathfrak{I}'_N$. Here we adopt the following notation:

$$g = \begin{pmatrix} 0, & 1, & \cdots, & N \\ i_0, & i_1, & \cdots, & i_N \end{pmatrix} \downarrow$$

is an element of the symmetric group of degree $N + 1$, S_{N+1}; the set $\{\tilde{\xi}_j\} = P(g)\{\xi_j\} = \{P(g)\xi_j\}$ is the set of the transformed variables of $\{\xi_j\}$ induced by the permutation g operating on the suffix of $(x_0, x_1, \cdots, x_N)$, i.e.,

$$\tilde{\xi}_j = P(g)\xi_j = x_{i_j} - x_{i_{j-1}} = \sum_{k=1}^{N} p_{jk}(g)\xi_k = \begin{cases} \xi_{i_j} + \xi_{i_j-1} + \cdots + \xi_{i_{j-1}+1}, & i_j > i_{j-1}, \\ -(\xi_{i_{j-1}} + \xi_{i_{j-1}-1} + \cdots + \xi_{i_j+1}), & i_j < i_{j-1}. \end{cases} \quad (2)$$

Similarly we write $\{\tilde{\zeta}_j\} = P(g)\{\zeta_j\} = \{P(g)\zeta_j\}$; the permuted forward tube $P(g)\mathfrak{I}_N$ and the permuted extended tube $P(g)\mathfrak{I}'_N$ are defined as follows, writing the variables $\{\zeta_j\}$ explicitly,

$$P(g)\,\mathfrak{I}_N(\{\zeta_j\}) = \mathfrak{I}_N(P(g)\{\zeta_j\}) = \{\{\zeta_j\}; \operatorname{Im}(P(g)\zeta_j) \in V_+\}, \quad (3a)$$

and

$$P(g)\,\mathfrak{I}'_N(\{\zeta_j\}) = [P(g)\,\mathfrak{I}_N]'(\{\zeta_j\}) = \mathfrak{I}'_N(P(g)\{\zeta_j\}) = \{P(g)\{\zeta_j\}; P(g)\zeta_j = L_+(C)P(g)\zeta'_j, \operatorname{Im}(P(g)\zeta'_j) \in V_+\} = \{\{\zeta_j\}; \zeta_j = L_+(C)\zeta'_j, \{\zeta'_j\} \in P(g)\,\mathfrak{I}_N\}, \quad (3b)$$

where $L_+(C)P(g)\zeta_j = P(g)L_+(C)\zeta_j$ [if we write $\zeta_j = \zeta^\mu_j$, μ denoting the component of the 4-vector $(\mu = 0, 1, 2, 3)$, $P(g)$ operates only on j, and $L_+(C)$ operates only on μ].

The aim of this article is to prove that the analytic continuation of the W function due to local commutativity, mentioned above, is single valued in

$$\bigcup_{g \in S \subseteq S_{N+1}} P(g)\,\mathfrak{I}'_N$$

(Sec. 2), where S is an arbitrary subset of S_{N+1}, and that the domain

$$\bigcup_{g \in S \subseteq S_{N+1}} P(g)\mathfrak{I}'_N$$

is simply connected (Sec. 3). In previous papers (see reference 2 and the discussions of Sec. 5), these statements were often assumed but no proof was given. (In the special case $N = 2$, they are, however, easy consequences of the results of Källen and Wightman.[5]) It is also proved that the inter-

[3] Here "proper" means "connected component with unit element." Sometimes it is called "restricted" or "proper orthochronous."

[4] D. Hall and A. S. Wightman, Kgl. Danske Videnskab. Selskab. Mat-fys. Medd. 31, n. 5 (1957); R. Jost, in *Theoretical Physics in the Twentieth Century* (Interscience Publishers, Inc., New York, 1960), p. 107.

[5] It follows from the Dyson Theorem[6] that the analytic continuation of the W function due to local commutativity is single-valued [at least in the small neighborhood of the totally spacelike points $S = \{\{\xi_j\}; \xi_j = x_j - x_{j-1}, x_k - x_l \in V_s, (k, l = 0, 1, 2, \cdots, N; k \neq l)\}$]. Araki has concluded[7] the connectedness of

$$\bigcup_{g \in S_{N+1}} P(g)\mathfrak{I}'_N.$$

by showing that

$$\bigcup_{g \in S_{N+1}} P(g)J_N$$

is connected, $P(g)J_N \subset P(g)\mathfrak{I}'_N$ being the set of Jost points

section $P(g_1)\mathfrak{I}_N' \cap P(g_2)\mathfrak{I}_N'$ is simply connected (Sec. 4).

2. SINGLE VALUEDNESS OF THE ANALYTIC CONTINUATION OF THE W FUNCTION

The set J_N of the real points $\{\xi_i\}$, called the Jost points, of the extended tube $\mathfrak{I}_N'$ is characterized by the following condition:[9,10]

$$\left(\sum_{j=1}^{N} \lambda_j \xi_j\right)^2 > 0, \tag{4}$$

for

$$\sum_{j=1}^{N} \lambda_j = 1, \qquad \lambda_j \geq 0. \tag{5}$$

The set $P(g)J_N$ of the Jost points of the permuted extended tube $P(g)\mathfrak{I}_N'$ is given by

$$P(g)J_N(\{\xi_i\}) = J_N(P(g)\{\xi_i\}) = \left\{\{\xi_i\}; \left(\sum_{j=1}^{N} \lambda_j P(g)\xi_j\right)^2 > 0, \sum_{j=1}^{N} \lambda_j = 1, \lambda_j \geq 0\right\}. \tag{6}$$

Lemma 1. $\bigcap_{\alpha=1,2,3} P(g_\alpha)J_N$ is nonempty for $\forall g_\alpha \in S_{N+1}, (\alpha = 1, 2, 3)$.

Proof: Take a set $Q(g_1, g_2, g_3)$ of points $\{\xi_i\}$, such that

$$Q(g_1, g_2, g_3) = \{\{\xi_i\}; \xi_i^0 = 0, P(g_\alpha)\xi_i^\alpha > 0 \text{ for } \forall_i, \text{ or } < 0 \text{ for } \forall_i, (\alpha = 1, 2, 3)\} \tag{7}$$

which is nonempty because $P(g)\{\xi_i\}$ affords a representation[11] of S_{N+1}, and thus, Eq. (2) is solvable in terms of $\{\xi_i\}$ [or see Eq. (2') after Lemma 3]. Since

$$\left(\sum_{j=1}^{N} \lambda_j P(g_\alpha)\xi_j\right)^2 = \sum_{\beta=1}^{3} \left(\sum_{j=1}^{N} \lambda_j P(g_\alpha)\xi_j^\beta\right)^2 > 0 \tag{8}$$

for a point $\{\xi_i\} \in Q(g_1, g_2, g_3)$, for $\alpha = 1, 2, 3$, and for $\{\lambda_i\}$ satisfying (5), we get

$$Q(g_1, g_2, g_3) \subset \bigcup_{\alpha=1,2,3} P(g_\alpha)J_N. \tag{9}$$

(Q.E.D.)

Now, $P(g)J_N \subset P(g)\mathfrak{I}_N'$, and so we have the following theorem:

Theorem 1. The intersection of any three of the permuted extended tubes $P(g)\mathfrak{I}_N'$ is nonempty.

Lemma 2. Any arbitrary point belonging to $P(g_1)\mathfrak{I}_N' \cap P(g_2)\mathfrak{I}_N'$ is connected by a path inside $P(g_1)\mathfrak{I}_N' \cap P(g_2)\mathfrak{I}_N'$ to a point belonging to $P(g_1)\mathfrak{I}_N' \cap P(g_2)\mathfrak{I}_N$ and to a point belonging to $P(g_1)\mathfrak{I}_N \cap P(g_2)\mathfrak{I}_N'$.

Proof: Take a point $\{\zeta_i'\} \in P(g_1)\mathfrak{I}_N' \cap P(g_2)\mathfrak{I}_N'$. By the definition of the extended tube, we can find complex Lorentz transformations $\Lambda_1, \Lambda_2 \in L_+(C)$ such that

$$\{\Lambda_\alpha \zeta_i'\} = \{(\zeta_i)_\alpha\} \in P(g_\alpha)\mathfrak{I}_N, \quad (\alpha = 1, 2). \tag{10}$$

Since $L_+(C)$ is a connected set, we can find continuous paths $\Lambda_\alpha(t)$ such that

$$\Lambda_\alpha(t) \in L_+(C), \qquad 0 \leq t \leq 1, \qquad \Lambda_\alpha(0) = 1, \quad \Lambda_\alpha(1) = \Lambda_\alpha \qquad (\alpha = 1, 2). \tag{11}$$

From the invariance of $P(g)\mathfrak{I}_N'$ under the operation of $L_+(C)$, and from Eqs. (10) and (11) it follows that the continuous curves

$$\{(\zeta_i(t))_\alpha\} = \{\Lambda_\alpha(t)\zeta_i'\} \subset P(g_1)\mathfrak{I}_N' \cap P(g_2)\mathfrak{I}_N', \qquad 0 \leq t \leq 1, \qquad \alpha = 1, 2 \tag{12}$$

give the required paths to connect the points of the lemma. (Q.E.D.)

According to this lemma, the question of the connectedness of $P(g_1)\mathfrak{I}_N' \cap P(g_2)\mathfrak{I}_N'$ was reduced to that of $P(g_1)\mathfrak{I}_N' \cap P(g_2)\mathfrak{I}_N$. Clearly it is sufficient

---

(for definition, see Sec. 2) which is connected. The fact that $\bigcup_{g \in S_{N+1}} P(g)J_N$ is connected follows from Araki's Lemma,[7] which states that any point of the connected set $K = \{\{\xi_i\}; \xi_i^0 = 0, \xi_j = \mathbf{x}_j - \mathbf{x}_{j-1}$ are real, $(\mathbf{x}_0, \mathbf{x}_1, \cdots, \mathbf{x}_N)$ are all distinct$\}$ is contained in $\bigcup_{g \in S_{N+1}} P(g)J_N$ and from the fact that $K \cap P(g)J_N$ is nonempty for $\forall g \in S_{N+1}$. [It also follows from the existence[8] of the nonempty intersection $J_N \cap P((k-1, k))J_N$, $g = (k-1, k)$ being a neighboring transposition belonging to S_{N+1}, which seems to be not a direct consequence of that lemma contrary to Araki's statement]. Then Araki has stated that the analytic continuation of W function due to local commutativity is single-valued in a small neighborhood of $\bigcup_{g \in S_{N+1}} P(g)J_N$, which is the consequence of the connectedness of $\bigcup_{g \in S_{N+1}} P(g)J_N$. This is, of course, a weaker conclusion than that of the Dyson Theorem, since $\bigcup_{g \in S_{N+1}} P(g)J_N \subset S$.

For the analysis of the analyticity domain of the three-point function ($N = 2$), see G. Källen and A. S. Wightman, Kgl. Danske Videnskab. Selskab. Mat. Fys. Skr. **1**, no. 6 (1958); D. Ruelle, Helv. Phys. Acta **34**, 587 (1961).

[6] F. J. Dyson, Phys. Rev. **110**, 579 (1958).

[7] H. Araki, Progr. Theoret. Phys. (Kyoto) Suppl. **18**, 83 (1961), Sec. 5; Ann. Phys. **11**, 260 (1960), Lemma 1.

[8] A. S. Wightman, J. Indian Math. Soc. **24**, 625 (1960); Y. Tomozawa, Nuovo Cimento **27**, 543 (1963) Lemma 2.

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

[10] We use the metric given by $x^2 = -x_0^2 + \mathbf{x}^2$.

[11] Actually it is the irreducible representation corresponding to the partition $(\lambda) = (N, 1)$. (See Appendix B).

to discuss the case with $g_1 = 1$, and $g_2 = g$ an arbitrary permutation of S_{N+1}, since

$$P(g_1)\,\mathfrak{I}'_N(\{\zeta_i\}) \cap P(g_2)\,\mathfrak{I}_N(\{\zeta_i\}) = \mathfrak{I}'_N(P(g_1)\{\zeta_i\}) \cap P(g_2 g_1^{-1})\,\mathfrak{I}_N(P(g_1)\{\zeta_i\}), \quad (13)$$

and we can take $P(g_1)\{\zeta_i\}$ as the set of new variables. Here we have used the definition (3) of the permuted forward and extended tubes, and the relation

$$P(g_2 g_1)\,\mathfrak{I}_N = P(g_2)P(g_1)\,\mathfrak{I}_N \quad (14)$$

or

$$P(g_2 g_1)\,\mathfrak{I}'_N = P(g_2)P(g_1)\,\mathfrak{I}'_N, \quad (14')$$

which follows from the fact that $P(g)\{\zeta_i\}$ affords a representation[11] of S_{N+1}.

Lemma 3. (Jost[9]) $\mathfrak{I}'_N = P(g_I)\mathfrak{I}'_N$, where

$$g_I = g_I^{-1} = \begin{pmatrix} 0, & 1, \cdots, & N \\ N, N-1, & \cdots, & 0 \end{pmatrix} \downarrow .$$

Proof: g_I induces the transformation $\{\zeta_i\} \to P(g_I)\{\zeta_i\} = \{-\zeta_i\}$. Take

$$\Lambda(-1) = \begin{pmatrix} -1 & & & \\ & -1 & & 0 \\ & & -1 & \\ & 0 & & -1 \end{pmatrix} \in L_+(C).$$

Using $L_+(C)\Lambda(-1) = L_+(C)$, and Eq. (3), we have

$$\begin{aligned} P(g_I)\,\mathfrak{I}'_N &= \{\{\zeta_i\};\ \zeta_i = L_+(C)\zeta'_i,\ \mathrm{Im}\,(-\zeta'_i) \in V_+\} \\ &= \{\{\zeta_i\};\ \zeta_i = L_+(C)(\Lambda(-1)\zeta'_i), \\ &\qquad \mathrm{Im}\,(\Lambda(-1)\zeta'_i) \in V_+\} \\ &= \mathfrak{I}'_N. \qquad \text{(Q.E.D.)} \end{aligned}$$

According to this lemma we need to discuss the connectedness of $\mathfrak{I}'_N \cap P(g)\mathfrak{I}_N$ only for the cases $g \neq 1, g_I$.

A point $\{\zeta_i\} \in \mathfrak{I}'_N \cap P(g)\mathfrak{I}_N$ has the following properties:

(a) $\exists\, \Lambda \in L_+(C)$ such that $\mathrm{Im}\,(\Lambda\zeta_i) \in V_+$,

(b) $\mathrm{Im}\,\xi_i \in V_+$, where

$$\xi_i = P(g)\zeta_i = \sum_{k=1}^{N} p_{ik}(g)\zeta_k$$

[see Eq. (2)].

Using

$$g = \begin{pmatrix} 0, & 1, & \cdots, & N \\ i_0, & i_1, & \cdots, & i_N \end{pmatrix} \downarrow = \begin{pmatrix} l_0, & l_1, & \cdots, & l_N \\ 0, & 1, & \cdots, & N \end{pmatrix} \downarrow ,$$

the inverse of Eq. (2) (with $\xi_i \to \zeta_i$) is given by

$$\zeta_i = P(g^{-1})\xi_i = \sum_{k=1}^{N} p_{ik}(g^{-1})\xi_k = \begin{cases} \xi_{l_i} + \xi_{l_i-1} + \cdots + \xi_{l_{i-1}+1}, & l_i > l_{i-1}, \\ -(\xi_{l_{i-1}} + \xi_{l_{i-1}-1} + \cdots + \xi_{l_i+1}), & l_i < l_{i-1}. \end{cases} \quad (2')$$

According to the property (b) and the relation (2'), we can classify j, the suffix of the component 4-vector ζ_i of the point $\{\zeta_i\} \in P(g)\mathfrak{I}_N$ into two classes $Z_\pm(g)$ as follows:

$$j \in Z_+(g) \quad \text{or} \quad Z_-(g), \quad \text{if} \quad p_{ik}(g^{-1}) \geq 0 \quad \text{or} \quad \leq 0 \quad (k = 1, 2, \cdots, N). \quad (15)$$

Thus, a necessary condition of the property (b) is that

$$\mathrm{Im}\,\zeta_j \in V_\pm \quad \text{for} \quad j \in Z_\pm(g), \quad (16)$$

according to Eq. (2') and the fact that the $V_\pm$ are convex sets. Incidentally both sets $Z_\pm(g)$ are nonempty unless $g = 1, g_I$.

A complex Lorentz transformation $\Lambda \in L_+(C)$ can be expressed in the normal form

$$\Lambda = L_1 M L_2, \quad (17)$$

where $L_1, L_2 \in L_+^\uparrow$ ($L_+^\uparrow$ being the proper homogeneous real Lorentz group) and $M \in L_+(C)$ has one of two possible forms[4]:

$$M_1(\varphi, \chi) = \begin{pmatrix} \cos\varphi & i\sin\varphi & 0 & 0 \\ i\sin\varphi & \cos\varphi & 0 & 0 \\ 0 & 0 & \cosh\chi & i\sinh\chi \\ 0 & 0 & -i\sinh\chi & \cosh\chi \end{pmatrix}, \quad \varphi, \chi \text{ real}, \quad (18a)$$

or

$$M_2^*(\tau) = \pm \begin{pmatrix} 1 & 0 & \tau & i\tau \\ 0 & 1 & \tau & i\tau \\ \tau & -\tau & 1 & 0 \\ i\tau & -i\tau & 0 & 1 \end{pmatrix}, \quad \tau \text{ real}. \quad (18b)$$

Since $L_+^\uparrow$ is connected and leaves $P(g)\mathfrak{I}_N$ and $P(g)\mathfrak{I}'_N$ invariant, we can ignore L_1 and L_2. [There exists a continuous curve which connects $\{\zeta_i\}$ and $\{L_2\zeta_i\}$ inside $\mathfrak{I}'_N \cap P(g)\mathfrak{I}_N$. We write $\{L_2\zeta_i\}$ as $\{\zeta_i\}$ for simplicity. As for L_1, if $\{\Lambda\zeta_i\} \in \mathfrak{I}_N$, then $\{L_1^{-1}\Lambda\zeta_i\} \in \mathfrak{I}_N$.]

Lemma 4. The second normal form $M_2^*(\tau)$ cannot transform any 4-vector, ζ_j, with $\mathrm{Im}\,\zeta_j \in V_-$ into $\mathfrak{I}_1$, i.e., into a 4-vector, ζ'_j, with $\mathrm{Im}\,\zeta'_j \in V_+$.

Proof: Take the case of $M_2^+(\tau)$ and $\mathrm{Im}\,\zeta_i = \eta_i \in V_-$. We can readily get[4]

$$\zeta_i(\tau) = M_2^+(\tau)\zeta_i = \xi_i(\tau) + i\eta_i(\tau), \tag{19}$$

$$\eta_i^0(\tau) = \eta_i^0 + \tau(\eta_i^2 + \xi_i^3), \tag{20a}$$

$$\begin{aligned} -[\eta_i(\tau)]^2 &= [\eta_i^0(\tau)]^2 - [\mathbf{n}_i(\tau)]^2 \\ &= -(\eta_i)^2 + 2\tau[\xi_i^3(\eta_i^0 - \eta_i^1) - \eta_i^3(\xi_i^0 - \xi_i^1)] \\ &\quad - \tau^2[(\eta_i^0 - \eta_i^1)^2 + (\xi_i^0 - \xi_i^1)^2]. \end{aligned} \tag{20b}$$

Since $\eta_i^0(0) = \eta_i^0 < 0$, the condition $\eta_i^0(\tau) > 0$ gives the range $\tau > \tau_0 > 0$ or $\tau < \tau_0 < 0$, according to whether $\tau_0 > 0$ or $\tau_0 < 0$, where $\eta_i^0(\tau_0) = 0$. On the other hand $-[\eta_i(\tau)]^2 \leq 0$ for $\tau > \tau_0 > 0$ or $\tau < \tau_0 < 0$, respectively, since $-[\eta_i(\tau)]^2$ is at most a quadratic function of τ in which the coefficient of τ^2 is ≤ 0, and $-[\eta_i(0)]^2 > 0$ and $-[\eta_i(\tau_0)]^2 \leq 0$. Thus, $\eta_i(\tau) \notin V_+$, for $\forall\tau$ real and $\forall\eta_i \in V_-$.

The case of $M_2^-(\tau)$ and $\eta_i \in V_+$ can be proved quite similarly. (Q.E.D.)

Corollary. The second normal form $M_2^\pm(\tau)$ cannot transform a point $\{\zeta_i\} \in P(g)\mathfrak{I}_N$ into $\mathfrak{I}_N$ for $g \neq 1, g_I$.

Proof: For $g \neq 1, g_I$ the classes $Z_\pm(g)$ are non-empty. According to Lemma 4 and the definition (16), ζ_i, where $j \in Z_\pm(g)$, cannot be transformed into $\mathfrak{I}_1$ by $M_2^\mp(\tau)$. This establishes the statement. (Q.E.D.)

Thus, a point $\{\zeta_i\} \in \mathfrak{I}_N' \cap P(g)\mathfrak{I}_N$, $(g \neq 1, g_I)$ must be transformed into $\mathfrak{I}_N$ by the first normal form[12] $M_1(\varphi, \chi)$.

Lemma 5. Assume that $\{M_1(\varphi, \chi)\zeta_i\} \in \mathfrak{I}_N$ for a point $\{\zeta_i\} \in \mathfrak{I}_N' \cap P(g)\mathfrak{I}_N$ and for a particular (φ, χ).

Define

$$\zeta_i(\rho) = \begin{bmatrix} \rho\xi_i^0 \\ \xi_i^1 \\ \rho\xi_i^2 \\ \rho\xi_i^3 \end{bmatrix} + i\begin{bmatrix} \eta_i^0 \\ \rho\eta_i^1 \\ \rho\eta_i^2 \\ \rho\eta_i^3 \end{bmatrix}. \tag{21}$$

Then

$$\left.\begin{aligned} &\{\zeta_i(\rho)\} \in \mathfrak{I}_N' \cap P(g)\,\mathfrak{I}_N \\ &\text{and} \\ &\{M_1(\varphi, \chi)\zeta_i(\rho)\} \in \mathfrak{I}_N \\ &\text{for} \\ &-1 \leq \rho \leq 1 \end{aligned}\right\}. \tag{22}$$

Proof: Since

$$P(g)\zeta_i(\rho) = \xi_i(\rho) = \xi_i(\rho) + i\tilde{\eta}_i(\rho), \tag{23a}$$

and since

$$\tilde{\eta}_i \in V_+ \Rightarrow \tilde{\eta}_i(\rho) \in V_+ \quad \text{for} \quad -1 \leq \rho \leq 1, \tag{23b}$$

we get

$$\{\zeta_i\} \in P(g)\,\mathfrak{I}_N \Rightarrow \{\zeta_i(\rho)\} \in P(g)\,\mathfrak{I}_N \tag{24}$$

for $-1 \leq \rho \leq 1$.

The following formulas and definitions are self-explanatory:

$$\eta_i(\varphi, \chi) = \mathrm{Im}\,\zeta_i(\varphi, \chi) = \mathrm{Im}\,(M_1(\varphi, \chi)\zeta_i) = \begin{bmatrix} \eta_i^0\cos\varphi + \xi_i^1\sin\varphi \\ \eta_i^1\cos\varphi + \xi_i^0\sin\varphi \\ \eta_i^2\cosh\chi + \xi_i^3\sinh\chi \\ \eta_i^3\cosh\chi - \xi_i^2\sinh\chi \end{bmatrix}, \tag{25}$$

$$\zeta_i(\varphi, \chi; \rho) = M_1(\varphi, \chi)\zeta_i(\rho) = \xi_i(\varphi, \chi; \rho) + i\eta_i(\varphi, \chi; \rho), \tag{26}$$

$$\zeta_i(\varphi, \chi; 1) = \zeta_i(\varphi, \chi), \tag{27}$$

$$\eta_i^0(\varphi, \chi; \rho) = \eta_i^0(\varphi, \chi) = \eta_i^0\cos\varphi + \xi_i^1\sin\varphi, \tag{28}$$

and

$$\begin{aligned} -[\eta_i(\varphi, \chi; \rho)]^2 &= -[\eta_i(\varphi, \chi)]^2 + (1 - \rho^2)[\mathbf{n}(\varphi, \chi)]^2 \\ &\geq -[\eta_i(\varphi, \chi)]^2, \\ &\qquad \text{for} \quad -1 \leq \rho \leq 1. \end{aligned} \tag{29}$$

Then, using Eqs. (28) and (29) it follows from the condition $\eta_i(\varphi, \chi) \in V_+$ that $\eta_i(\varphi, \chi; \rho) \in V_+$ for $-1 \leq \rho \leq 1$. That is to say,

$$\{M_1(\varphi, \chi)\zeta_i\} \in \mathfrak{I}_N \Rightarrow \{M_1(\varphi, \chi)\zeta_i(\rho)\} \in \mathfrak{I}_N \quad \text{for} \quad -1 \leq \rho \leq 1. \tag{30}$$

Equations (24) and (30) establish the lemma. (Q.E.D.)

Lemma 6. The set C of the points $\{\zeta_i\}$, which are of the form

$$\zeta_i = \begin{bmatrix} 0 \\ \xi_i^1 \\ 0 \\ 0 \end{bmatrix} + i\begin{bmatrix} \eta_i^0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{31}$$

and have the properties

$$\mathrm{Im}\,(P(g)\zeta_i) = \tilde{\eta}_i^0 > 0 \tag{32a}$$

and

$$\xi_i^1 > 0, \quad \text{for} \quad \forall j, \tag{32b}$$

is a connected subset of $\mathfrak{I}_N' \cap P(g)\mathfrak{I}_N$.

[12] Apart from the real Lorentz transformation belonging to $L_+^\uparrow$. (See the discussion given before Lemma 4).

Proof: Clearly C is connected (actually it is convex), and it follows from (32a) that $C \subset P(g)\mathfrak{I}_N$. To prove that $C \subset \mathfrak{I}_N'$ we operate with $M_1(\varphi, \chi)$ on $\{\zeta_i\}$ getting

$$M_1(\varphi,\chi)\zeta_i = \zeta_i(\varphi,\chi) = \begin{pmatrix} 0 \\ \xi_i^1 \cos\varphi - \eta_i^0 \sin\varphi \\ 0 \\ 0 \end{pmatrix} + i\begin{pmatrix} \eta_i^0 \cos\varphi + \xi_i^1 \sin\varphi \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad (33)$$

which does not depend on χ. According to (15), (16), and (32a), we have $\eta_i^0 \gtrless 0$ for $j \in Z_{\pm}(g)$. Defining φ_i such that

$$0 < \varphi_i = \tan^{-1}(-\eta_i^0/\xi_i^1) < \pi, \quad (34)$$

we have that

$$\tfrac{1}{2}\pi < \varphi_i < \pi \quad \text{for} \quad j \in Z_+(g), \quad (35a)$$

and

$$0 < \varphi_i < \tfrac{1}{2}\pi \quad \text{for} \quad j \in Z_-(g), \quad (35b)$$

according to (32b). Since

$$\eta_i^0(\varphi,\chi) = \eta_i^0 \cos\varphi + \xi_i^1 \sin\varphi > 0 \quad (36a)$$

when

$$\varphi_i - \pi < \varphi < \varphi_i \quad \text{for} \quad j \in Z_+(g), \quad (36b)$$

and

$$0 < \varphi_i < \varphi \quad \text{and} \quad -\pi < \varphi < \varphi_i - \pi < 0 \quad \text{for} \quad j \in Z_-(g), \quad (36c)$$

we conclude that

$$\operatorname{Im}[M_1(\varphi,\chi)\zeta_i] = \eta_i(\varphi,\chi) \in V_+ \quad (37)$$

for

$$\max_{j\in Z_-(g)} \varphi_i < \varphi < \min_{j\in Z_+(g)} \varphi_i, \quad \chi \text{ arbitrary}. \quad (38)$$

We note that the domain of (φ, χ) given by Eq. (38) is nonempty according to Eq. (35). This establishes that $C \subset \mathfrak{I}_N'$ and so the lemma is proved. (Q.E.D.)

Lemma 7. A point $\{\zeta_i\} \in \mathfrak{I}_N' \cap P(g)\mathfrak{I}_N$ is connected, inside $\mathfrak{I}_N' \cap P(g)\mathfrak{I}_N$, to the set C, C being defined in Lemma 6.

Proof: It is sufficient to consider the point $\{\zeta_i\} \in \mathfrak{I}_N' \cap P(g)\mathfrak{I}_N$ which satisfies $\{M_1(\varphi,\chi)\zeta_i\} \in \mathfrak{I}_N$ for a (φ, χ) according to the explanation given just before Lemma 4, and the corollary of Lemma 4. (The lemma is trivial for $g = 1$ and g_I, since $C \subset P(g)\mathfrak{I}_N \subset \mathfrak{I}_N'$ for $g = 1$ and g_I. Assuming $g \neq 1, g_I$, the above statement is correct).

Thus we can apply Lemma 5: For $\{\zeta_i(\rho)\}$, which is defined by Eq. (21), Eq. (22) is valid. Changing ρ of $\{\zeta_i(\rho)\}$ from 1 to 0, we get a continuous curve which connects $\{\zeta_i\} = \{\zeta_i(1)\}$ and $\{\zeta_i'\} = \{\zeta_i(0)\}$ inside $\mathfrak{I}_N' \cap P(g)\mathfrak{I}_N$, where $\zeta_i' = \xi_i' + i\eta_i'$ is of the form (31) and satisfies (32a), and

$$\{M_1(\varphi,\chi)\zeta_i'\} \in \mathfrak{I}_N,$$

i.e.,

$$\eta_i^0 \cos\varphi + \xi_i^1 \sin\varphi > 0 \quad \text{for } \forall j \text{ and for a } \varphi. \quad (39)$$

The allowed domain of φ of Eq. (39) can be either

$$0 < \max_{j\in Z_-(g)} \varphi_i < \varphi < \min_{j\in Z_+(g)} \varphi_i < \pi, \quad (40a)$$

or

$$-\pi < \max_{j\in Z_+(g)} (\varphi_i - \pi) < \varphi < \min_{j\in Z_-(g)} (\varphi_i - \pi) < 0, \quad (40b)$$

depending on whether

$$\max_{j\in Z_-(g)} \varphi_i < \min_{j\in Z_+(g)} \varphi_i, \quad (41a)$$

or

$$\max_{j\in Z_+(g)} \varphi_i < \min_{j\in Z_-(g)} \varphi_i, \quad (41b)$$

where φ_i is defined by (34). However, the case (41b) can be reduced to the case (41a): Operating with the space rotation $R_3(\pi)$ of angle π, around the third axis, on $\{\zeta_i^1\}$, the sign of all ξ_i' is inverted (the new point being denoted by $\{\bar{\zeta}_i\}$) and φ_i is changed into $\bar{\varphi}_i = \pi - \varphi_i$ which satisfies Eq. (41a). Since the space rotation R leaves $\mathfrak{I}_N' \cap P(g)\mathfrak{I}_N$ invariant, and R is connected, $\{\zeta_i'\}$ and $\{\bar{\zeta}\}$ are connected inside $\mathfrak{I}_N' \cap P(g)\mathfrak{I}_N$. Thus, we need to consider only the case (41a).

Now $\varphi_i = \tan^{-1}(-\eta_i^0/\xi_i^1)$ is an increasing or decreasing function of ξ_i^1 for $j \in Z_+(g)$ or $j \in Z_-(g)$, respectively. Thus for any $\xi_i^{1\prime} > \xi_j^1$, we have

$$0 < \max_{j\in Z_-(g)} \tan^{-1}(-\eta_i^0/\xi_i^{1\prime}) < \max_{j\in Z_-(g)} \varphi_i < \min_{j\in Z_+(g)} \varphi_i < \min_{j\in Z_+(g)} \tan^{-1}(-\eta_i^0/\xi_i^{1\prime}) < \pi. \quad (42)$$

This means that the increasing ξ_i^1 in $\{\zeta_i'\}$ of the form (31), in the case (41a), does not change the property that $\{\zeta_i'\} \in \mathfrak{I}_N'$. The condition $\{\zeta_i'\} \in P(g)\mathfrak{I}_N$ is

invariant under the change of Re ζ_i' since the above condition is relevant only to Im ζ_i'. Therefore, the continuous curve which is given by increasing ξ_i^1 can connect $\{\zeta_i'\}$ with the point $\{\zeta_i''\} \in C$, where Re $\zeta_i^{1''} >$ max (Re $\zeta_i^{1'} = \xi_i^1$, 0) and Im $\zeta_i^{0''} =$ Im $\zeta_i^{0'} = \eta_i^0$, and is inside $\mathfrak{T}_N' \cap P(g)\mathfrak{T}_N$. This establishes the lemma. (Q.E.D.)

Corollary. $P(g_1)\mathfrak{T}_N' \cap P(g_2)\mathfrak{T}_N$ is connected.

Proof: Lemmas 6 and 7 show that any point $\{\zeta_i\} \in \mathfrak{T}_N' \cap P(g)\mathfrak{T}_N$ is connected to the connected set $C \subset \mathfrak{T}_N' \cap P(g)\mathfrak{T}_N$, inside $\mathfrak{T}_N' \cap P(g)\mathfrak{T}_N$, which proves the connectedness of $\mathfrak{T}_N' \cap P(g)\mathfrak{T}_N$. Equation (13) generalizes it accordingly. (Q.E.D.)

Theorem 2. $P(g_1)\mathfrak{T}_N' \cap P(g_2)\mathfrak{T}_N'$ is connected.

Proof: It follows from Lemma 2 and the last corollary.

Theorem 3. The analytic continuation of W function, due to local comutativity, is single valued in the domain

$$\bigcup_{g \in S \subseteq S_{N+1}} P(g)\,\mathfrak{T}_N',$$

where S is an arbitrary subset of the symmetric group S_{N+1}.

Proof: From Theorems 1, 2, and Lemma 1, it follows that

$$P(g_1)\,\mathfrak{T}_N' \cap P(g_2)\,\mathfrak{T}_N', \quad \text{for } \forall g_1, g_2 \in S_{N+1}$$

is a nonempty connected domain, and contains $P(g_1)J_N \cap P(g_2)J_N$, which is nonempty. The latter forms a real environment,[13] since the set of Jost points is an open set in the real $4N$-dimensional Minkowski space. Thus, local commutativity equates (up to a sign) the W functions at $P(g_1)J_N \cap P(g_2)J_N$, and gives the single-valued analytic continuation in $P(g_1)\mathfrak{T}_N' \cup P(g_2)\mathfrak{T}_N'$ since $P(g_1)\mathfrak{T}_N' \cap P(g_2)\mathfrak{T}_N'$ is connected. Applying this process to all the pairs of $\{P(g)\mathfrak{T}_N'\}_{g \in S \subseteq S_{N+1}}$, the statement of the theorem follows.

3. SIMPLE CONNECTEDNESS OF THE UNION OF THE EXTENDED TUBES

First we prove the simple connectedness of an extended tube itself. To this end it is convenient to use the covering group (the universal covering group) $\bar{L}_+(C)$ of $L_+(C)$. [$\bar{L}_+(C)$ is isomorphic to $SL(2, C) \otimes SL(2, C)$ and is simply connected, where $SL(2, C)$ is the complex unimodular group of two dimensions. Clearly $\bar{L}_+(C)\mathfrak{T}_N = \mathfrak{T}_N' = L_+(C)\mathfrak{T}_N$, since $\bar{L}_+(C)\mathfrak{T}_N$ covers $\mathfrak{T}_N'$ twice, as does $L_+(C)\mathfrak{T}_N$.]

Lemma 8. If a point $\{\zeta_i'\} \in \mathfrak{T}_N'$ can be expressed by

$$\zeta_i' = \bar{\Lambda}_0(\zeta_i)_0 = \bar{\Lambda}_1(\zeta_i)_1, \tag{43a}$$

where

$$\bar{\Lambda}_\alpha \in \bar{L}_+(C) \text{ and } \{(\zeta_i)_\alpha\} \in \mathfrak{T}_N, \quad (\alpha = 0, 1), \tag{43b}$$

then there exist continuous curves

$$\bar{\Lambda}(t) \subset \bar{L}_+(C) \text{ and } \{\zeta_i(t)\} \subset \mathfrak{T}_N, \quad 0 \leq t \leq 1, \tag{44a}$$

with

$$\bar{\Lambda}(\alpha) = \bar{\Lambda}_\alpha \text{ and } \zeta_i(\alpha) = (\zeta_i)_\alpha, \quad (\alpha = 0, 1), \tag{44b}$$

such that

$$\zeta_i' = \bar{\Lambda}(t)\zeta_i(t) \quad \text{for} \quad 0 \leq t \leq 1. \tag{44c}$$

Proof: From the conditions $(\zeta_i)_1 = (\bar{\Lambda}_1)^{-1}\bar{\Lambda}_0(\zeta_i)_0$, $\{(\zeta_i)_0\}$, $\{(\zeta_i)_1\} \in \mathfrak{T}_N$, and $(\bar{\Lambda}_1)^{-1}\bar{\Lambda}_0 \in \bar{L}_+(C)$, it follows that there exist curves

$$\bar{\Lambda}'(t) \subset \bar{L}_+(C) \text{ and } \{\zeta_i(t)\} \subset \mathfrak{T}_N, \quad 0 \leq t \leq 1, \tag{45a}$$

with

$$\begin{aligned} &\bar{\Lambda}'(0) = 1, \qquad \bar{\Lambda}'(1) = (\bar{\Lambda}_1)^{-1}\bar{\Lambda}_0, \\ &\zeta_i(0) = (\zeta_i)_0, \qquad \zeta_i(1) = (\zeta_i)_1, \end{aligned} \tag{45b}$$

such that

$$\zeta_i(t) = \bar{\Lambda}'(t)(\zeta_i)_0, \qquad 0 \leq t \leq 1, \tag{45c}$$

according to a lemma of Hall and Wightman included in Lemma 1 of reference 4.

Then we have

$$\zeta_i' = \bar{\Lambda}_0(\zeta_i)_0 = \bar{\Lambda}(t)\zeta_i(t), \qquad 0 \leq t \leq 1, \tag{46}$$

where we put

$$\bar{\Lambda}(t) = \bar{\Lambda}_0(\bar{\Lambda}'(t))^{-1}. \tag{47}$$

It is easy to see that $\bar{\Lambda}(t) \subset \bar{L}_+(C)$ and $\{\zeta_i(t)\} \subset \mathfrak{T}_N$, $(0 \leq t \leq 1)$, as defined by (47) and (45c), respectively, are the required curves. (Q.E.D.)

Lemma 9. For a continuous closed curve

$$\{\zeta_i'(t)\} \subset \mathfrak{T}_N', \qquad 0 \leq t \leq 1, \tag{48a}$$

with

$$\zeta_i'(0) = \zeta_i'(1), \tag{48b}$$

we can find continuous closed curves

$$\bar{\Lambda}(s) \subset \bar{L}_+(C) \text{ and } \{\zeta_i(s)\} \subset \mathfrak{T}_N, \quad 0 \leq s \leq 1, \tag{49a}$$

with

$$\bar{\Lambda}(0) = \bar{\Lambda}(1) \quad \text{and} \quad \zeta_i(0) = \zeta_i(1), \tag{49b}$$

[13] S. Bochner and W. T. Martin, *Several Complex Variables* (Princeton University Press, Princeton, New Jersey, 1948), Chap. II, Sec. 2.

such that

$$\zeta_i'(t(s)) = \zeta_i''(s) = \bar{\Lambda}(s)\zeta_i(s), \qquad 0 \leq s \leq 1, \tag{50}$$

with a suitable parametrization by s, where $t(s)$ is a nondecreasing continuous function of s with $t(0) = 0$ and $t(1) = 1$.

Proof: From the facts that

$$\mathfrak{Z}_N' = \bigcup_{\bar{\Lambda} \in \bar{L}_+(C)} \bar{\Lambda}\mathfrak{Z}_N,$$

that $\bar{\Lambda}\mathfrak{Z}_N$ is an open set, and that the curve (48) is an image of the compact set [0, 1] due to a continuous mapping T, it follows[14] that there exists a $\delta > 0$ such that the image of the δ neighborhood of each point $t \in [0, 1]$ due to T is found in a suitably chosen $\bar{\Lambda}\mathfrak{Z}_N$. Therefore, we can find a finite number of open sets $\bar{\Lambda}_\alpha\mathfrak{Z}_N$, $(\alpha = 1, 2, \cdots, n)$, which covers the curve (48) entirely in such a way that for a suitable division $0 = t_0 < t_1 < \cdots < t_n = 1$ of [0, 1] and images $Q_\alpha' = \{\zeta_i'(t_\alpha)\}$ of t_α, $(\alpha = 0, 1, \cdots, n)$, the arc $Q_{\alpha-1}'Q_\alpha'$ of the curve is contained in $\bar{\Lambda}_\alpha\mathfrak{Z}_N$, $(\alpha = 1, 2, \cdots, n)$. Since $\bar{\Lambda}_\alpha \in \bar{L}_+(C)$ is a homeomorphic mapping of $\mathfrak{Z}_N$ onto $\bar{\Lambda}_\alpha\mathfrak{Z}_N$, the mapping $(\bar{\Lambda}_\alpha)^{-1}$ of arc $Q_{\alpha-1}'Q_\alpha' \subset \mathfrak{Z}_N'$ gives a continuous curve

$$Q_{\alpha-1}^{(\alpha)}Q_\alpha^{(\alpha)} = \{\{\zeta_i^{(\alpha)}(t)\}; t_{\alpha-1} \leq t \leq t_\alpha\} \subset \mathfrak{Z}_N, \tag{51a}$$

with

$$Q_\alpha' = \bar{\Lambda}_\alpha Q_\alpha^{(\alpha)} = \bar{\Lambda}_{\alpha+1}Q_\alpha^{(\alpha+1)}, \qquad (\alpha = 1, 2, \cdots, n), \tag{51b}$$

where we put $\bar{\Lambda}_{n+1} = \bar{\Lambda}_1$ and $Q_n^{(n+1)} = Q_0^{(1)}$. The relation (51) enables us to apply Lemma 8 to find curves

$$\bar{\Lambda}_{(\alpha,\alpha+1)}(\tau_\alpha) \subset \bar{L}_+(C) \quad \text{and} \quad Q^{(\alpha,\alpha+1)}(\tau_\alpha) \subset \mathfrak{Z}_N, \qquad 0 \leq \tau_\alpha \leq 1, \tag{52a}$$

with

$$\begin{aligned} &\bar{\Lambda}_{(\alpha,\alpha+1)}(0) = \bar{\Lambda}_\alpha, \qquad \bar{\Lambda}_{(\alpha,\alpha+1)}(1) = \bar{\Lambda}_{\alpha+1}, \\ &Q^{(\alpha,\alpha+1)}(0) = Q_\alpha^{(\alpha)}, \quad Q^{(\alpha,\alpha+1)}(1) = Q_\alpha^{(\alpha+1)}, \end{aligned} \tag{52b}$$

such that

$$Q_\alpha' = \bar{\Lambda}_{(\alpha,\alpha+1)}(\tau_\alpha)Q^{(\alpha,\alpha+1)}(\tau_\alpha), \quad 0 \leq \tau_\alpha \leq 1, \tag{52c}$$

for $\alpha = 1, 2, \cdots, n$. Now we introduce the parameter s, $0 \leq s \leq 1$, and with the partition

$$0 = s^0 < s_1 < s^1 < s_2 < s^2 < \cdots < s^{n-1} < s_n < s^n = 1, \tag{53a}$$

we put

$$\left.\begin{aligned} &t_{\alpha-1} \leq t(s) \leq t_\alpha \quad \text{for} \quad s^{\alpha-1} \leq s \leq s_\alpha, \\ &t(s) = t_\alpha \quad \text{and} \\ &0 \leq \tau_\alpha(s) \leq 1 \end{aligned}\right\} \quad \text{for} \quad s_\alpha \leq s \leq s^\alpha, \qquad \alpha = 1, 2, \cdots, n, \tag{53b}$$

where the parameters $t(s)$ and $\tau_\alpha(s)$ are nondecreasing continuous functions of s. The continuous closed curves $\{\zeta_i'(t(s))\} = \{\zeta_i''(s)\} \subset \mathfrak{Z}_N'$ connecting $Q_0', Q_1', \cdots, Q_n' = Q_0'$; $\bar{\Lambda}(s) \subset \bar{L}_+(C)$ connecting $\bar{\Lambda}_1, \bar{\Lambda}_2, \cdots, \bar{\Lambda}_n, \bar{\Lambda}_{n+1} = \bar{\Lambda}_1$ (putting $\bar{\Lambda}(s) = \bar{\Lambda}_\alpha$ for $s^{\alpha-1} \leq s \leq s_\alpha$); and $\{\zeta_i(s)\} \subset \mathfrak{Z}_N$ connecting $Q_0^{(1)}, Q_1^{(1)}, Q_1^{(2)} \cdots, Q_n^{(n)}, Q_n^{(n+1)} = Q_0^{(1)}$, which were described above, give the required relation (50). (Q.E.D.)

Corollary. For a continuous curve (48a), we can find continuous curves (49a) such that (50) is satisfied with a suitable parametrization by s.

Proof: This follows from Lemma 9 and from the fact that any continuous curve can be considered as a part of a continuous closed curve.

Theorem 4. The extended tube $\mathfrak{Z}_N'$ is simply connected.

Proof: $\mathfrak{Z}_N'$ is connected since $\bar{L}_+(C)$ and $\mathfrak{Z}_N$ are connected, and $\mathfrak{Z}_N'$ is a continuous image of $\bar{L}_+(C) \otimes \mathfrak{Z}_N$. According to Lemma 9, any continuous closed curve (48) belonging to $\mathfrak{Z}_N'$ can be expressed by Eq. (50) in terms of the continuous closed curves (49) belonging to $\bar{L}_+(C)$ and $\mathfrak{Z}_N$. Since $\bar{L}_+(C)$ and $\mathfrak{Z}_N$ are simply connected (note that $\mathfrak{Z}_N$ is convex), we can let the curves (49) shrink to points inside each domain. Therefore, the curve (48) shrinks to a point inside $\mathfrak{Z}_N'$. Thus the theorem is established. (Q.E.D.)

For proving the simple connectedness of $\bigcup P(g)\mathfrak{Z}_N'$ we prove the following lemmas:

Lemma 10. If the simply connected domains $D_1, D_2, \cdots, D_n$ have a nonempty common intersection

$$\bigcap_{i=1}^{n} D_i,$$

and the intersection $D_k \cap D_l$ $(k, l, = 1, \cdots, n)$, of any two is connected, then the union

$$\bigcup_{i=1}^{n} D_i$$

is simply connected.

[14] H. Seifert and W. Threlfall, *Lehrbuch der Topologie* (B. G. Teubner, Leipzig, 1934), Chap. 2, Sec. 7, Satz IV; or F. Hausdorff, *Set Theory* (Chelsea Publishing Company, New York, 1957), Chap. VI, p. 26, Theorem III (the Borel Covering Theorem for separable spaces).

Proof:

$$\bigcup_{j=1}^{n} D_j$$

is connected, since any point of it is connected to

$$\bigcap_{j=1}^{n} D_j.$$

Take a curve belonging to

$$\bigcup_{j=1}^{n} D_j$$

which runs through domains $D_{i_1} \to D_{i_2} \to \cdots \to D_{i_m}$, where $(i_1, i_2, \cdots, i_m)$ is a set of integers taken from the set of integers $(1, 2, \cdots, n)$ with repetition allowed, but $i_k \neq i_{k+1}$ $(k = 1, 2, \cdots, m-1)$. Since the curve necessarily must pass through $D_{i_k} \cap D_{i_{k+1}}$ before leaving D_{i_k}, we can choose a set of points, $Q_{1,2}, \cdots, Q_{m-1,m}$ on the given curve Q_1Q_m such that $Q_1 \in D_{i_1}$, $Q_{k,k+1} \in D_{i_k} \cap D_{i_{k+1}}$ $(k = 1, \cdots, m-1)$, $Q_m \in D_{i_m}$, and arc $Q_1Q_{1,2} \subset D_{i_1}$, arc $Q_{k-1,k}Q_{k,k+1} \subset D_{i_k}$ $(k = 2, \cdots, m-1)$, arc $Q_{m-1,m}Q_m \subset D_{i_m}$, and where arc $Q_{k-1,k}Q_{k,k+1}$ is the portion of the curve between $Q_{k-1,k}$ and $Q_{k,k+1}$. Taking a point

$$O \in \bigcap_{j=1}^{n} D_j,$$

we can draw the continuous curves joining O and Q's in such a way that arc $OQ_1 \subset D_{i_1}$, arc $OQ_{k,k+1} \subset D_{i_k} \cap D_{i_{k+1}}$ $(k = 1, \cdots, m-1)$, and arc $OQ_m \subset D_{i_m}$, since $D_k \cap D_l$ $(k, l = 1 \cdots, n)$, is connected. Thus, all the curves $OQ_{k-1,k}$, $Q_{k-1,k}Q_{k,k+1}$, and $OQ_{k,k+1}$ are inside D_{i_k}. Using the terminology of equivalence[15] (denoted by $\sim$) for the case where two curves with the same ends can be deformed continuously to each other, and multiplication (denoted by $\cdot$) for joining two curves when the end point of the first is the starting point of the second, we have the following equivalence relations, since D_j is simply connected:

$$\begin{aligned} Q_1Q_{1,2} &\sim Q_1O \cdot OQ_{1,2}, \\ Q_{1,2}Q_{2,3} &\sim Q_{1,2}O \cdot OQ_{2,3}, \\ &\cdots \\ Q_{m-1,m}Q_m &\sim Q_{m-1,m}O \cdot OQ_m. \end{aligned} \tag{54}$$

By multiplying these successively we get

$$Q_1Q_m \sim Q_1O \cdot OQ_m,$$

since[15] $AB \cdot BA \sim E$ and $AB \cdot E \sim AB$. (Here BA represents the same curve as AB but with with opposite direction). Thus, we have proved that any curves joining Q_1 and Q_m are equivalent to $Q_1O \cdot OQ_m$, and therefore, are equivalent to each other. This proves the statement.

Lemma 11. Let $D_1, D_2, \cdots, D_n$ be simply connected domains. If any three of them have a non-empty common intersection, and the intersection of any two of them is connected, then the union

$$\bigcup_{j=1}^{n} D_j$$

is simply connected.

Proof:

$$\bigcup_{j=1}^{n} D_j$$

is connected since any two points $Q_\alpha \in D_{i_\alpha}$, $(\alpha = 1, 2)$, are connected to each other through $D_{i_1} \cap D_{i_2}$. Take a continuous curve contained in

$$\bigcup_{j=1}^{n} D_j$$

which runs through the domains $D_{i_1} \to D_{i_2} \to \cdots \to D_{i_m}$, and take the points $Q_1, Q_{1,2}, \cdots, Q_{m-1,m}, Q_m$ on the curve, defined in the same way as in the proof of Lemma 10. For the case $n \leq 3$ and the case $n > 3$ and $m \leq 3$, the statement is true, according to Lemma 10. Assume $n > 3$ and $m > 3$. Since D_{i_1}, D_{i_2}, and D_{i_3} satisfy the conditions of Lemma 10, the curve $Q_1Q_{1,2} \cdot Q_{1,2}Q_{2,3}$ is equivalent to $Q_1Q_{1,2,3} \cdot Q_{1,2,3}Q_{2,3}$ where $Q_{1,2,3} \in D_{i_1} \cap D_{i_2} \cap D_{i_3}$ and $Q_1Q_{1,2,3} \subset D_{i_1}$, $Q_{1,2,3}Q_{2,3} \subset D_{i_2}$. Then the curve Q_1Q_m is equivalent to the curve

$$Q_1Q_{1,2,3} \cdot Q_{1,2,3}Q_{2,3} \cdot Q_{2,3}Q_m$$

which runs through $D_{i_1} \to D_{i_3} \to \cdots \to D_{i_m}$, where the number of domains is reduced by one from that of Q_1Q_m. Thus, by induction, we arrive at the statement of the lemma. (Q.E.D.)

Theorem 5. The union

$$\bigcup_{g \in S \subseteq S_{N+1}} P(g)\, \mathfrak{T}'_N$$

of any number of extended tubes is simply connected.

Proof: According to Theorems 1, 2, and 4, any subset of $\{P(g)\mathfrak{T}'_N\}_{g \in S_{N+1}}$ satisfies the conditions of Lemma 11. Thus the theorem is established.

4. SIMPLE CONNECTEDNESS OF THE INTERSECTION OF TWO EXTENDED TUBES

Lemma 12: $P(g_1)\mathfrak{T}'_N \cap P(g_2)\mathfrak{T}_N$ is simply connected.

Proof: It is sufficient to prove the case $g_1 = 1$ and $g_2 = g$ arbitrary, but $g \neq 1$, g_I, as mentioned

[15] L. Pontrjagin, *Topological Groups* (Princeton University Press, Princeton, New Jersey, 1958), Chap. VIII, Sec. 46.

before [see Eq. (13) and Lemma 3]. We can prove the lemma by continuously deforming a continuous closed curve of $\mathfrak{I}_N' \cap P(g)\mathfrak{I}_N$ into a continuous closed curve of the convex set $C \subset \mathfrak{I}_N' \cap P(g)\mathfrak{I}_N$, C being defined in Lemma 6. [C is convex because of the definition (31) and (32), and thus is simply connected].

For a continuous closed curve

$$\{\zeta_i'(t)\} \subset \mathfrak{I}_N' \cap P(g)\mathfrak{I}_N, \qquad 0 \leq t \leq 1,$$

we can find continuous closed curves $\bar{\Lambda}(s) \subset \bar{L}_+(C)$ and $\{\zeta_i(s)\} \subset \mathfrak{I}_N$ such that

$$\zeta_i'(t(s)) = \zeta_i'(s) = \bar{\Lambda}(s)\zeta_i(s), \qquad 0 \leq s \leq 1, \tag{55}$$

with a suitable parametrization by s, according to Lemma 9. A corresponding expression to Eqs. (17), (18), and the corollary of Lemma 4 for the case of the covering group $\bar{L}_+(C)$ enables us to express $\bar{\Lambda}(s)$ by[4]

$$(\bar{\Lambda}(s))^{-1} = \bar{L}_1(s)M_1(\varphi(s), \chi(s))\bar{L}_2(s), \quad 0 \leq s \leq 1, \tag{56}$$

where $\bar{L}_1(s)$, $\bar{L}_2(s) \subset \bar{L}_+^{\uparrow}$, $\bar{L}_+^{\uparrow}$ being the covering group of $L_+^{\uparrow}$; $M_1(\varphi(s), \chi(s)) \subset \bar{L}_+(C)$ is given by Eq. (18a), and $\bar{L}_1(s)$, $\bar{L}_2(s)$, and $M_1(\varphi(s), \chi(s))$ $(0 \leq s \leq 1)$, are continuous closed curves. Without loss of generality, we can ignore the $\bar{L}_1(s)$ and $\bar{L}_2(s)$ by a reason[16] similar to that described before Lemma 4.

For $\zeta_i'(s) = \xi_i'(s) + i\eta_i'(s)$, define a set of continuous closed curves

$$\zeta_i'(s;\rho) = \begin{pmatrix} \rho\xi_i^{0'}(s) \\ \xi_i^{1'}(s) \\ \rho\xi_i^{2'}(s) \\ \rho\xi_i^{3'}(s) \end{pmatrix} + i \begin{pmatrix} \eta_i^{0'}(s) \\ \rho\eta_i^{1'}(s) \\ \rho\eta_i^{2'}(s) \\ \rho\eta_i^{3'}(s) \end{pmatrix}, \quad 0 \leq s \leq 1, \quad \rho \text{ real.} \tag{57}$$

Since $\{\zeta_i'(s)\} = \{\zeta_i'(s;1)\} \subset \mathfrak{I}_N' \cap P(g)\mathfrak{I}_N$ and $\{M_1(\varphi(s), \chi(s))\zeta_i'(s)\} \subset \mathfrak{I}_N$, we get

$$\{\zeta_i'(s;\rho)\} \subset \mathfrak{I}_N' \cap P(g)\,\mathfrak{I}_N, \tag{58a}$$

and

$$\{M_1(\varphi(s), \lambda(s))\zeta_i'(s;\rho)\} \subset \mathfrak{I}_N, \tag{58b}$$

for

$$-1 \leq \rho \leq 1 \quad \text{and} \quad 0 \leq s \leq 1, \tag{58c}$$

according to Lemma 5. Thus, a continuous change of ρ from 1 to 0 affords a continuous deformation of $\{\zeta_i'(s)\}$ into the continuous closed curve $\{\zeta_i'(s;0)\} \subset \mathfrak{I}_N' \cap P(g)\mathfrak{I}_N$, inside $\mathfrak{I}_N' \cap P(g)\mathfrak{I}_N$. The latter curve is of the form of Eq. (31), and satisfies the relation $\{M_1(\varphi(s), \chi(s))\zeta_i'(s;0)\} \subset \mathfrak{I}_N$, $0 \leq s \leq 1$. Finally, the procedure described in the proof of Lemma 7 affords a continuous deformation of $\{\zeta_i'(s;0)\}$ into a continuous closed curve contained in the convex set $C \subset \mathfrak{I}_N' \cap P(g)\mathfrak{I}_N$, inside $\mathfrak{I}_N' \cap P(g)\mathfrak{I}_N$. [A space rotation and an increase of the same amount of the real part ξ_i^1 of each of the component 4-vectors of $\{\zeta_i'(s;0)\}$ as described in Lemma 7 keep the continuous closed curve inside $\mathfrak{I}_N' \cap P(g)\mathfrak{I}_N$.] Since the set C is simply connected, we can deform the derived continuous closed curve into a point in $C \subset \mathfrak{I}_N' \cap P(g)\mathfrak{I}_N$. Thus the lemma is established. (Q.E.D.)

Theorem 6. $P(g_1)\mathfrak{I}_N' \cap P(g_2)\mathfrak{I}_N'$ is simply connected.

Proof: For a continuous closed curve $\{\zeta_i''(t)\} \subset P(g_1)\mathfrak{I}_N' \cap P(g_2)\mathfrak{I}_N'$, $0 \leq t \leq 1$, we can find continuous closed curves $\bar{\Lambda}(s) \subset \bar{L}_+(C)$ and $\{\zeta_i(s)\} \subset P(g_1)\mathfrak{I}_N' \cap P(g_2)\mathfrak{I}_N$, such that the Eq. (55) is satisfied with a suitable parametrization by s, according to Lemma 9. From the fact that $\bar{L}_+(C)$ and $P(g_1)\mathfrak{I}_N' \cap P(g_2)\mathfrak{I}_N$ are simply connected (Lemma 12), we can conclude the theorem.

5. DISCUSSION

(1) Theorem 3 follows from Theorem 5, since the analytically continued function is connected. We note, however, that the former has been derived by a smaller number of pieces of knowledge compared with the case of the latter, as seen in the proofs in the text. Similarly, Theorem 4 can be considered as a stronger statement than that of the Bargmann–Hall–Wightman Theorem.[4]

(2) According to Theorem 3, the Ruelle Theorem,[17] which states that the holomorphy envelope of

$$\bigcup_{g \in S_{N+1}} P(g)\,\mathfrak{I}_N'$$

contains the totally spacelike points[5] S, turns out to be applicable to the quantum field theory which is based on axioms (I)–(IV). The difference of the contents of the Ruelle Theorem and the Dyson Theorem[6] lies in that the former is global while the latter is local in character.

(3) A continuous mapping[1,18]

[16] The continuous closed curve $\{\zeta_i'(s)\} \subset \mathfrak{I}_N' \cap P(g)\mathfrak{I}_N$ can be deformed continuously into the continuous closed curve $\{\bar{L}_2(s)\zeta_i'(s)\} \subset \mathfrak{I}_N' \cap P(g)\mathfrak{I}_N$, inside $\mathfrak{I}_N' \cap P(g)\mathfrak{I}_N$, since $\bar{L}_+^{\uparrow}$ leaves $\mathfrak{I}_N' \cap P(g)\mathfrak{I}_N$ invariant, and since it follows from the simple-connectedness of $\bar{L}_+^{\uparrow}$ that we can deform continuously the continuous closed curve $\bar{L}_2(s) \subset \bar{L}_+^{\uparrow}$ into the unit element of $\bar{L}_+^{\uparrow}$, inside $\bar{L}_+^{\uparrow}$. For the continuous closed curve $\{\zeta_i(s)\} \subset \mathfrak{I}_N$, the continuous closed curve $\{\bar{L}_1^{-1}(s)\zeta_i(s)\}$ is contained in $\mathfrak{I}_N$.

[17] D. Ruelle, Helv. Phys. Acta **32**, 135 (1959).

[18] A. S. Wightman, J. Indian Math. Soc. **24**, 625 (1960).

$$\{\zeta_i\} \to \{\zeta_i \cdot \zeta_k\}, \qquad j, k = 1, 2, \cdots, N \tag{59}$$

maps the domain $\mathfrak{Z}'_N$ (or $\mathfrak{Z}_N$) into a space composed of a symmetric complex $N \times N$ matrix of rank ≤ 4, the image of the mapping being denoted by $\mathfrak{M}_N$. Since the mapping (59) is such that an inside point of $\mathfrak{Z}'_N$ is mapped to an inside point of $\mathfrak{M}_N$, and a boundary point to a boundary point, all the Theorems 1–6 of this article are valid if we replace $\mathfrak{Z}'_N$ by $\mathfrak{M}_N$.

(4) The results of this article seem to clarify the statements about local commutativity given in the paper of reference 1. This is due to the reason that, in constructing a quantum field theory from a set of analytic functions, following Wightman, we need knowledge of what the domains of analyticity of these analytic functions are.

(5) Streater[19] has extended the discussion of the analytic properties of the W function to that of an arbitrary matrix element of the product of field operators, getting the same analyticity domain as that of the W function. Our results are then equally applicable to the case of Streater's treatment.

ACKNOWLEDGMENTS

The author wishes to thank Professor L. A. Radicati for kind hospitality and constant encouragement. He is indebted to Dr. H. Matsumura for instruction in mathematics, and for criticism and collaboration in Sec. 3. Thanks are also due to Professor F. Gürsey, Professor E. Fabri, Dr. A. Di Giacomo, Dr. L. Lovitch, Dr. P. Menotti, and Dr. L. E. Picasso for useful comments.

APPENDIX A: SOME PROPERTIES OF THE JOST POINTS AND THE EXTENDED TUBE[20]

Define the sets of the real points, $K_N^{[\alpha]}$ and $K_N^{[\alpha]*}$ as follows:

$$K_N^{[\alpha]} = \{\{\xi_i\}; \xi_i^\mu = 0 \quad \text{for} \quad \mu \neq \alpha\}, \qquad (\alpha = 1, 2, 3), \tag{A1}$$

and

$$K_N^{[\alpha]*} = \{\{\xi_i\}; \xi_i^\alpha \gtrless 0, \xi_i^\mu = 0 \quad \text{for} \quad \mu \neq \alpha\}, \qquad (\alpha = 1, 2, 3). \tag{A2}$$

Clearly $K_N^{[\alpha]}$ is invariant under the operation $P(g)$,

$$P(g)K_N^{[\alpha]}(\{\xi_i\}) = K_N^{[\alpha]}(P(g)\{\xi_i\}) = K_N^{[\alpha]}(\{\xi_i\}), \tag{A3}$$

and

$$P(g)K_N^{[\alpha]*}(\{\xi_i\}) = K_N^{[\alpha]*}(P(g)\{\xi_i\}) = \{\{\xi_i\}; P(g)\xi_i^\alpha = \xi_i^\alpha \gtrless 0, \xi_i^\mu = 0 \text{ for } \mu \neq \alpha\}. \tag{A4}$$

If we write as

$$K_N^{[\alpha]}(\{\xi_i\}) = K_1^{[\alpha]}(\xi_1) \otimes K_1^{[\alpha]}(\xi_2) \otimes \cdots \otimes K_1^{[\alpha]}(\xi_N), \tag{A5}$$

$K_1^{[\alpha]}(\xi)$ is the αth coordinate axis ($\alpha = 1, 2, 3$) of the space component in the real Monkowski space. The similarly defined $K_1^{[\alpha]+}(\xi)$ and $K_1^{[\alpha]-}(\xi)$ stand for the αth positive and negative coordinate axes. The set $Q(g_1, g_2, g_3)$ defined by (7) in the text can be expressed in terms of $K_N^{[\alpha]*}$ as follows:

$$Q(g_1, g_2, g_3) = \bigotimes_{\alpha=1,2,3} (P(g_\alpha)K_N^{[\alpha]+} + P(g_\alpha)K_N^{[\alpha]-}). \tag{A6}$$

Lemma A1.

$$K_N^{[\alpha]} \cap P(g)J_N = P(g)K_N^{[\alpha]+} + P(g)K_N^{[\alpha]-}. \tag{A7}$$

Proof: It is sufficient to prove the case $g = 1$, i.e.,

$$K_N^{[\alpha]} \cap J_N = K_N^{[\alpha]+} + K_N^{[\alpha]-}, \tag{A7'}$$

since (A7) follows from (A7') by operating with $P(g)$ and the property (A3). First, we have

$$K_N^{[\alpha]+} + K_N^{[\alpha]-} \subset J_N,$$

since

$$\left(\sum_{i=1}^N \lambda_i \xi_i\right)^2 = \left(\sum_{i=1}^N \lambda_i \xi_i^\alpha\right)^2 > 0$$

for a point $\{\xi_i\} \in K_N^{[\alpha]*}$ and for $\{\lambda_i\}$ satisfying Eq. (5). For a point

$$\{\xi_i\} \in (K_N^{[\alpha]} - \sum_{s=*} K_N^{[\alpha]s}),$$

either at least one component 4-vector $\xi_i = 0$, or one such pair $(\xi_{i_1}^\alpha, \xi_{i_2}^\alpha)$ have opposite signs. Thus, for both cases, we can find a $\{\lambda_i\}$ satisfying (5) which gives

$$\sum_{i=1}^N \lambda_i \xi_i = 0,$$

and so

$$(K_N^{[\alpha]} - \sum_{s=*} K_N^{[\alpha]s}) \cap J_N = \phi.$$

This completes the proof. (Q.E.D.)

Define

$$K_N^{[\alpha,\beta]} = K_N^{[\alpha]} \otimes K_N^{[\beta]} = \{\{\xi_i\}; \xi_i^\mu = 0 \text{ for } \mu \neq \alpha, \beta\} \qquad (\alpha \neq \beta), \tag{A8}$$

[19] R. F. Streater, J. Math. Phys. **3**, 256 (1962).

[20] These properties were not used in the text, but it might help in discussing the structure of the extended tube. For systematic analysis of the boundary, see A. S. Wightman, reference 18.

where $K_1^{[\alpha,\beta]}$ is the (α, β) coordinate plane in the space part of the Minkowski space.

Lemma A2.

$$K_N^{[\alpha,\beta]} \subset P(g)\bar{J}_N, \tag{A9}$$

where the right-hand side stands for the closure of the set $P(g)J_N$.

Proof: Take the case $\alpha = 1$, $\beta = 2$ for simplicity. Clearly

$$\{\{\xi_i\}; \xi_i^0 = 0, P(g)\xi_i^3 > 0 \text{ for } \forall j \text{ or } < 0 \text{ for } \forall j\} \subset P(g)J_N, \tag{A10}$$

according to Lemma A1, and such a point can be found in any neighborhood of any point of $K_N^{[1,2]}$.

Lemma A3. The extended tube is concave at any point belonging to $\partial K_N^{[\alpha]*}$, the boundary of $K_N^{[\alpha]*}$, which is contained in $\partial \mathfrak{I}_N' \cap \partial \mathfrak{I}_N$.

Proof: Clearly $\partial K_N^{[\alpha]*} \subset \partial \mathfrak{I}_N' \cap \partial \mathfrak{I}_N$, according to Lemma A1. First we prove the concavity of $\mathfrak{I}_N'$ at the origin $\{O\} \in \partial K_N^{[\alpha]*}$. Take a hyperplane passing through the origin, which is expressed as

$$\sum_{i=1}^{N} \sum_{\nu=0}^{3} (a_i^\nu \xi_i^\nu + b_i^\nu \eta_i^\nu) = 0, \tag{A11}$$

where a_i^ν and b_i^ν are real. Then it can be shown that for any choice of a_i^ν and b_i^ν, we can find a point belonging to $J_N \subset \mathfrak{I}_N'$, which satisfies Eq. (A11), in any neighborhood of the origin. To prove this, take a real point $\{\xi_i\} \in J_N$ such that

$\xi_i^0 = 0$, $\xi_i^\alpha > 0$ (or < 0) for $\forall j$ and a fixed α

(a) either when $\exists a_i^\beta \neq 0$ $(\beta \neq \alpha)$

$$(\alpha, \beta = 1, 2, 3), \tag{A12}$$

or

(b) when $a_i^\beta = 0$

$$(\beta = 1, 2, 3 \text{ and for } \forall j).$$

The point (A12) easily satisfies Eq. (A11) by adjusting ξ_i^β for the case (a), and evidently for the case (b). Moreover, if a $\{\xi_i\}$ of (A12) satisfies Eq. (A11), then all $\{l\xi_i\}$ do, l being arbitrary real number. This proves the above statement.

Next we consider a point $\{\bar{\xi}_i\} \in \partial K_N^{[\alpha]*}$. For simplicity take the case $\alpha = 1$, $+$ sign in the rhs, i.e., $\bar{\xi}_i$ satisfies the conditions $\bar{\xi}_i^1 \geq 0$ and $\bar{\xi}_i^\mu = 0$ $(\mu \neq 1)$. The equation for a hyperplane passing through the point $\{\bar{\xi}_i\}$ is

$$\sum_{i=1}^{N} \sum_{\nu=0}^{3} \{a_i^\nu(\xi_i^\nu - \bar{\xi}_i^\nu) + b_i^\nu \eta_i^\nu\} = 0. \tag{A11'}$$

Take $\xi_i = \bar{\xi}_i$ for j such that $\bar{\xi}_i^1 > 0$, and ξ_i of the type (A12) (the case >0) for j such that $\bar{\xi}_i = 0$. It is easy to see that this is a Jost point and that we can find a solution of (A11') from such points in any neighborhood of the point $\{\bar{\xi}_i\}$. Thus the lemma is established. (Q.E.D.)

APPENDIX B: PROOF OF THE IRREDUCIBILITY OF THE REPRESENTATION $P(g)$.

Lemma B. The set $\{\xi_i\}$, $(\xi_i = x_i - x_{i-1}, j = 1, 2, \cdots, N)$ forms a basis of the irreducible representation of the symmetric group S_{N+1} operating on the suffix of $(x_0, x_1, \cdots, x_N)$, which corresponds to the partition $(\lambda) = (N, 1)$.

Proof: Permutation

$$g = \begin{bmatrix} 0, & 1, & \cdots, & N \\ i_0, & i_1, & \cdots, & i_N \end{bmatrix} \downarrow$$

induces the transformation

$$\begin{pmatrix} \bar{x}_0 \\ \bar{x}_1 \\ \vdots \\ \bar{x}_N \end{pmatrix} = A(g) \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_N \end{pmatrix},$$

or for short, $\{\bar{x}_i\} = A(g)\{x_i\}$,

$$j = 0, 1, \cdots, N, \tag{B1}$$

where $A(g)$ is a representation of S_{N+1}, and thus, $\det(A(g)) \neq 0$. However, $A(g)$ is not irreducible, since

$$\xi_0 = \frac{1}{N+1} \sum_{i=0}^{N} x_i$$

is invariant under S_{N+1}. Make the following change of variables:

$$\{\xi_i\} = B\{x_i\}, \quad j = 0, 1, \cdots, N, \tag{B2}$$

where

$$B = \begin{bmatrix} \frac{1}{N+1} & \frac{1}{N+1} & \frac{1}{N+1} & \cdots & \frac{1}{N+1} \\ -1 & 1 & 0 & \cdots & 0 \\ 0 & -1 & 1 & \cdots & 0 \\ & & \cdots & & \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix},$$

$$\det B = 1. \tag{B3}$$

Writing

$$B^{-1} = \begin{bmatrix} 1 & b_{01} & \cdots & b_{0N} \\ 1 & b_{11} & \cdots & b_{1N} \\ & & \cdots & \\ 1 & b_{N1} & \cdots & b_{NN} \end{bmatrix}, \tag{B4}$$

we have the relations

$$\sum_{i=0}^{N} b_{ik} = 0, \qquad (k = 1, \cdots, N), \tag{B5}$$

from $BB^{-1} = 1$. Now the change of basis (B2) leads to

$$\{\xi_i\} = A'(g)\{\xi_i\}, \qquad j = 0, 1, \cdots, N, \tag{B6}$$

where, using Eqs. (B1)–(B5), we get

$$A'(g) = BA(g)B^{-1}$$
$$= B\begin{bmatrix} 1 & b_{i_0 1} & \cdots & b_{i_0 N} \\ 1 & b_{i_1 1} & \cdots & b_{i_1 N} \\ & \cdots & & \\ 1 & b_{i_N 1} & \cdots & b_{i_N N} \end{bmatrix} = \left[\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & P(g) & \\ 0 & & & \end{array}\right]. \tag{B7}$$

Thus, we get a representation $P(g)$ of S_{N+1}, whose basis is $\{\xi_i\}$, $(j = 1, \cdots, N)$, and $\det(P(g)) \neq 0$. Using the relation $\mathrm{tr}(A'(g)) = \mathrm{tr}(A(g))$ the number of x_i which are not changed by the permutation $g \in S_{N+1}$, we can calculate the character $\chi^P_{(1^\alpha, 2^\beta, \cdots)}$, for the representation $P(g)$, of the class corresponding to a partition in cycles $(1^\alpha, 2^\beta, \cdots)$ (where notation is the same as that used by Hamermesh[21]). Since $\chi^{A'}_{(1^\alpha, 2^\beta, \cdots)} = \alpha$, we have

$$\chi^P_{(1^\alpha, 2^\beta, \cdots)} = \alpha - 1 = \chi^{(N,1)}_{(1^\alpha, 2^\beta, \cdots)}, \tag{B8}$$

where $\chi^{(N,1)}_{(1^\alpha, 2^\beta, \cdots)}$ is the character of the class $(1^\alpha, 2^\beta, \cdots)$ for the irreducible representation corresponding to the partition $(N, 1)$, and the second equality can be readily obtained by the graphical method.[22] Therefore, $P(g) = [(p_{ik}(g))]$ is the irreducible representation which corresponds to the partition $(N, 1)$. Incidentally the dimension n^P of the representation is

$$n^P = \chi^P_{(1^{N+1})} = N,$$

which is, of course, consistent with the number of of $\{\xi_i\}$. (Q.E.D.)

Note added in proof: (a) Professor R. Jost has kindly communicated to the author that Theorem 3 of the text, i.e., the single-valuedness of the analytic continuation of the W function into the union of the extended tubes, has been proved by Dr. D. Ruelle with an essentially similar method. This work has not been published.

(b) Closedness of the continuous curves of the right-hand side of Eq. (56) is not trival, but can be proved for the case where the continuous closed curve $\bar{\Lambda}(s) \subset \bar{L}_+(C)$ $(0 \leq s \leq 1)$ transforms the points of $\mathfrak{I}'_N \cap P(g)\mathfrak{I}_N$ $(g \neq 1, g_I)$ into $\mathfrak{I}_N$.

The normal form, Eq. (18), has been derived by the decomposition,[4]

$$\{A, B\} = \{C^{-1}, C^{-1}\}\{C\,AB^{-1}C^{-1}, 1\}\{CB, CB\},$$

where $\{A, B\} \in \bar{L}_+(C)$ and $A, B, C \in SL(2, C)$. The Lorentz transformation associated with $\{A, B\}$ is

$$Z \to AZB^+ = \bar{\Lambda}\begin{bmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{bmatrix}$$

for

$$Z = \begin{bmatrix} x^0 + x^1 & x^2 - ix^3 \\ x^2 + ix^3 & x^0 - x^1 \end{bmatrix}, \quad \text{and} \quad \{A, A\} \in \bar{L}^\uparrow_+.$$

Equation (18a) or (56) corresponds to the case

$$CAB^{-1}C^{-1} = \begin{bmatrix} \alpha & 0 \\ 0 & \alpha^{-1} \end{bmatrix},$$

where

$$\alpha = e^{-\chi + i\varphi} \qquad (\chi, \varphi \text{ real})$$

is a continuous function of the matrix elements of A and B being an eigenvalue of AB^{-1}.

It is easy to see that a Lorentz transformation $\bar{\Lambda}$ which corresponds to a real value of α, i.e., $\varphi = n\pi$ (n integer), cannot transform the points belonging to $\mathfrak{I}'_N \cap P(g)\mathfrak{I}_N$ $(g \neq 1, g_I)$ into $\mathfrak{I}_N$. The reason is that to such a $\bar{\Lambda}$ corresponds the normal form

$$M_1(\varphi = n\pi, \chi) = M^*_1(\chi)$$
$$= \begin{bmatrix} \pm 1 & 0 & 0 & 0 \\ 0 & \pm 1 & 0 & 0 \\ 0 & 0 & \cosh\chi & i\sinh\chi \\ 0 & 0 & -i\sinh\chi & \cosh\chi \end{bmatrix},$$

so that $M^*_1(\chi)$, and thus $\bar{\Lambda}$, cannot transform a 4-vector ζ_j into $\mathfrak{I}_1$ for $j \in Z_*(g)$.

Since $\alpha(s)$ cannot be real, and since, therefore, α cannot move to α^{-1} continuously, $\alpha(s)$ $(0 \leq s \leq 1)$ is a continuous *closed* curve in the complex α plane. It is easy to construct a continuous closed curve $C(s)$ of the matrix which transforms the matrix $A(s)B(s)^{-1}$ into diagonal form. From this, Eq. (56) follows.

[21] H. Hamermesh, *Group Theory and its Applications to Physical Problems* (Addison-Wesley Publishing Company, Inc., Reading, Massachusetts, 1962), Chap. 7.

[22] H. Hamermesh, *Group Theory and its Applications to Physical Problems* (Addison-Wesley Publishing Company, Inc., Reading, Massachusetts, 1962), p. 206, Problem (3a).

Complex Singularities in Production Amplitudes

D. BESSIS AND F. PHAM
Centre d'Etudes Nucléaires de Saclay, Service de Physique Mathématique, Gif-sur-Yvette (Seine et Oise), France
(Received 5 April 1963)

It is shown for a simple model of production that Watson's scheme for final-state interaction may be good even though the amplitude has complex singularities. The method used consists in continuing analytically partial-wave dispersion relations. Stress is put upon the necessity of correctly choosing the cuts starting at the complex singularities. The safest choice in this approach is to take the cut along the line on which the series of partial waves diverges. A simple geometrical recipe is given to see the evolution of this line, the appearance of an anomalous threshold and a complex singularity.

1. INTRODUCTION

IT is well-known that, for production processes, the Mandelstam amplitude has complex singularities.[1] This is the case, for example, in π production:

$$N + \pi \to N + \pi + \pi.$$

We would like to study this process in some detail, and see the role of complex singularities. We completely neglect spin, and isotopic spin, the singularities being completely independent of these quantities. We shall therefore consider only one scalar amplitude.

Among the numerous graphs that can be considered, let us study more particularly the following one, previously studied,[2] and which will play an important role in our discussion. (See Fig. 1).

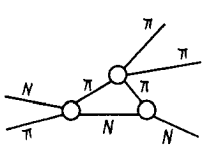


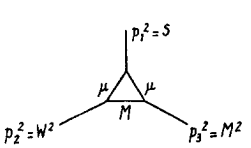
FIG. 1. Contractions of more general graphs leading to the vertex approximation.

The blobs correspond to "contractions" made on more general graphs. We let W be the "mass" of the $(N - \pi)$ system and s the mass of the $(\pi\pi)$ system. We must then look for the singularities of the graph. (Fig. 2).

Setting

$$x = (2\mu^2 - S)/2\mu^2,$$
$$y = (M^2 + \mu^2 - W^2)/2M\mu,$$
$$z = \mu/2M < 1,$$

FIG. 2. The vertex approximation.



we know[3] that the singularities are at

$$x = yz \pm [(y^2 - 1)(z^2 - 1)]^{\frac{1}{2}}, \tag{1.1}$$

or

$$S_\pm = (1/2M^2)(\mu^2[3M^2 + W^2 - \mu^2] \pm \mu[4M^2 - \mu^2]^{\frac{1}{2}} \times \{[W^2 - (M - \mu)^2][(M + \mu)^2 - W^2]\}^{\frac{1}{2}}). \tag{1.2}$$

A more careful analysis shows that only S_+ is on the physical sheet, S_- being on the unphysical sheet.

If we use a representation slightly different from the one used in Bonnevay's paper,[4] we can follow the singularities as W varies.

Let us represent the couple of two complex numbers (z_1, z_2) by a point (Re z_1, Re z_2), and a vector coming out from the point with components (Im z_1; Im z_2) in the plane. In this way, we get a very useful representation of the 4-dimensional space, that we shall use constantly in the future. (Fig. 3).

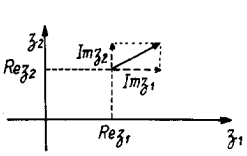


FIG. 3. Vector representation of complex points.

The necessary condition for the singularity to be on the physical sheet is that the origin of the representative vector in the (x, y), $C_2 \otimes C_2$ plane should be in the shadowed region. (Fig. 4).

The Landau surface for $z = \mu/2M$ is the set of complex points of an ellipse. The singularities are on the intersection of the ellipse with $y = y_0$. When W^2 grows from M^2 (point A^+) to $M^2 + 2\mu^2$ (point B^+), x first follows AB, and then BC on the ellipse, arriving at C^+ for $W^2 = (M + \mu)^2$. Then the intersection becomes complex, and the real part of

[1] L. F. Cook, Jr. and J. Tarski, J. Math. Phys. **3**, 1 (1962).
[2] V. V. Anisovich, A. A. Anselm, V. N. Gribov, and I. Diatlov, Zh. Eksperim i Teor. Fiz. **43**, 906 (1962) [English transl.: Soviet Phys.—JETP **16**, 643 (1963)].
[3] R. J. Eden, *Lectures in Theoretical Physics, Brandeis Summer Institute, 1961* (W. A. Benjamin, Inc., New York).
[4] G. Bonnevay, I. J. R. Aitchison, and J. S. Dowker, Nuovo Cimento **21**, 1001 (1961).

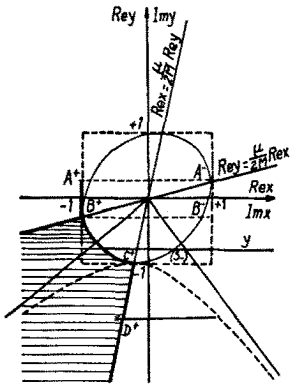


FIG. 4. Landau surface (complex points of the ellipse).

S_- follows C^+D^+. For $W^2 \geq M^2 + 2\mu^2$, it is necessary to give W^2 an imaginary part $W^2 = W^2 + i\epsilon$, because the singularity is on the boundary of the shadowed domain. It is easy to show then, that the real part of one of the singularities is inside the shadowed domain.

We show here the trajectory of S_- and S_+ in complex plane S, as a function of $W^2 + i\epsilon$. (Fig. 5).

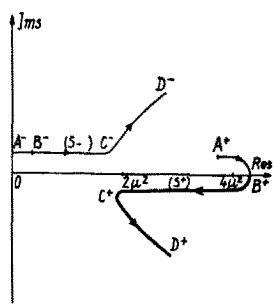


FIG. 5. Trajectories of the singularities in the complex s plane. Only the thick line (trajectory of the true singularity) is in the physical sheet.

One sees that there is an anomalous threshold from $W^2 = M^2 + 2\mu^2$ to $W^2 = (M + \mu)_0^2$, and then a complex singularity.

Another way to make the singularities of the amplitude appear, is to solve the scattering for $W = M$, for which there is no anomalous threshold, and then to continue analytically the solution in the W variable to its physical value $W > M + 2\mu$, following a procedure already used by Mandelstam.[5] This is the method we shall employ.

2. KINEMATICS FOR THE REACTION

We are treating the initial $(N - \pi)$ system as one unstable particle with mass W. M is the nucleon mass, and μ the pion mass. The four-momenta involved are shown on Fig. 6.

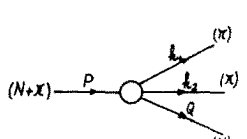


FIG. 6. Four-momenta involved in the reaction.

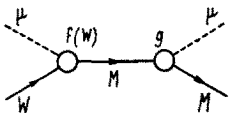


FIG. 7. The Born term in t and u channels.

Let us call $\mathbf{p}$ and $\mathbf{q}$, respectively, the momenta of the outgoing nucleon and pions in the center-of-mass system of the outgoing pions. The invariants

$$s = (k_1 + k_2)^2,\ t = (P - k_1)^2,\ \mu = (P - k_2)^2 \quad (2.1)$$

are thus related to the variables

$$p = |\mathbf{p}| = \{[s - (W - M)^2][s - (W + M)^2]/4s\}^{\frac{1}{2}},$$
$$q = |\mathbf{q}| = (s/4 - \mu^2)^{\frac{1}{2}}, \quad (2.2)$$
$$\theta = \text{angle between } \mathbf{p} \text{ and } \mathbf{q},$$

through the formulas

$$s = 4(\mu^2 + q^2),$$
$$t = (W^2 + M^2 - s)/2 + \mu^2 + 2pq\cos\theta, \quad (2.3)$$
$$\mu = (W^2 + M^2 - s)/2 + \mu^2 - 2pq\cos\theta.$$

3. CALCULATION OF THE AMPLITUDE FOR $W = M$

Let us assume that the amplitude $A(s, t, u, W)$ is an analytic function of W, and that it satisfied Mandelstam's representation for $W = M$, with a cut starting at $s = 4\mu^2$ due to unitarity in the channel $W + \bar{N} = \pi + \pi$. It is shown in reference 5 why these assumptions are reasonable.

In the s channel, the unitarity condition for the first partial wave is

$$\text{Im } A_l(s, W = M) = A_l^*(s, W = M)e^{i\delta_l} \times \sin\delta_l \quad \text{for} \quad s > 4\mu^2, \quad (3.1)$$

where δ_l is the $(\pi - \pi)$ phase shift.

Let us emphasize that $A_l(s, W = M)$ in formula (3.1) represents the l-wave projection of the analytic continuation to $W = M$ of the physical amplitude $A(s, t, u, W)$, and that it does not necessarily coincide with the analytic continuation to $W = M$ of the physical partial wave $A_l(s, W)$, because, as will be seen later, during the process of analytic continuation, the series of partial waves may diverge:

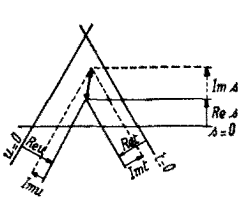
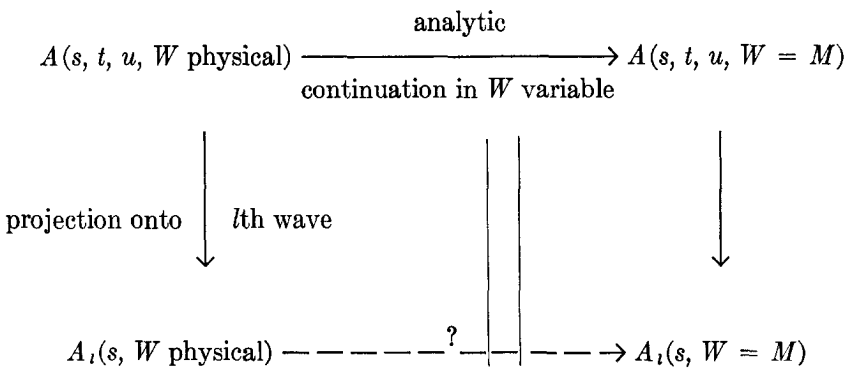


FIG. 8. Vector representation of complex points in trilinear coordinates.

[5] S. Mandelstam, Phys. Rev. Letters 4, 84 (1960).



(3.2)

series of partial waves may diverge.

The amplitude has other singularities—for example, the pole coming from the Born term, cuts coming from unitarity in the crossed channels, etc. For the sake of simplicity we will keep only the Born term (Fig. 7)

$$B(s, t, u) = f(W)g[1/(t - M^2) + 1/(u - M^2)].$$

In fact, we could be more general and consider the case of any isolated singularity at $t = M'^2$ or $u = M'^2$ (with M' an arbitrary mass). Thus,

$$A_l(s, W = M) = B_l(s, W = M) + F_l(s, W = M), \qquad (3.3)$$

where F_l (the final-state interaction term) has only the right-hand cut $(4\mu^2, \infty)$ with the phase δ_l. The solution of this problem is well-known.[6]

It is, up to an arbitrary real polynomial,

$$F_l(s, W = M) = \frac{1}{D_l(s)} \int_{4\mu^2}^{\infty} 4q'(s')^{-\frac{1}{2}} \times \frac{N_l(s')B_l(s', W = M)}{s' - s}\, ds', \qquad (3.4)$$

where

$$D_l(s) = \exp\left[-\frac{1}{\pi}\int_{4\mu^2}^{\infty} \frac{\delta_l(s')}{s' - s}\, ds'\right], \qquad (3.5)$$

and

$$N_l(s) = (s^{\frac{1}{2}}/4\pi q)D_l(s)e^{i\delta_l(s)} \sin \delta_l(s). \qquad (3.6)$$

The only remaining task is the analytic continuation of the expression (3.3) to the physical value of W. However, in all our continuations, we must be careful not to cross values of S for which a singularity touches on the line between -1 and $+1$ in the $\cos\theta$ plane, so as to be sure that the continued A_l really represents the integral

$$\frac{1}{2}\int_{-1}^{+1} A(s, t, u)P_l(\cos\theta)d\cos\theta,$$

and not the same integral over a deformed contour.

[6] R. Omnès, Nuovo Cimento 8, 316 (1958); N. I. Muskhelishvili, *Singular Integral Equations* (P. Noordhoff Ltd., Gronigen, The Netherlands, 1963).

4. ANALYTIC CONTINUATION

For fixed W, let us draw the curve $\cos\theta = \pm 1$ (Kibble's cubic) in trilinear coordinates (Re s, Re t, Re u). It will be convenient to represent any complex point on the same diagram as a vector as shown on Fig. 8.

With this representation, all complex points which lie on a straight line with real coefficients will appear as vectors along the real part of this straight line, and we shall say that they are carried by this line.

Let us begin at values of W near M, where we know there are no anomalous thresholds, and increase W, following the evolution of the singularities (which, in the simple model considered here, are only the cut s real $\geq 4\mu^2$ and the poles $t = M'^2$ and $u = M'^2$). What happens when one projects onto a partial wave? To see that, let us consider any line $s = s_0$ complex constant. It intersects Kibble's cubic at infinity, and at two other points M, N symmetric with respect to the exchange $t \leftrightarrow u$. These points M and N correspond to $\cos\theta = \pm 1$. As t is a linear function of $\cos\theta$ for s fixed, the points corresponding to the segment $-1 < \cos\theta < +1$ will belong to the linear interpolation between M and N. One such point (P) is represented on Fig. 9. One sees that a pole will fall on the segment $-1 < \cos\theta < +1$ if and only if the real straight lines carrying the vectors M and N intersect at the point α where the pole lines $t = M'^2$ and $u = M'^2$ meet, and if these pole lines are between M and N, as shown on Fig. 9.

Thus, to get the "forbidden" points (cuts) in the s complex plane, one need only intersect Kibble's cubic with all such real straight lines and read the values of s obtained. In particular, the end-point

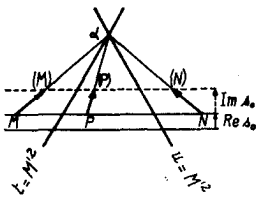


FIG. 9. Intersection of a straight line $s = s_0$ with Kibble's cubic (points M and N) showing a case when poles ($t = M'^2$ and $u = M'^2$) fall on the segment $-s < \cos\theta < +1$.

singularities correspond to the intersection of the cubic with the pole lines. Their analytic expression in the case $M' = M$ is

$$s_\pm = (\tfrac{1}{2}M^2)\{\mu^2(W^2 + 3M^2 - \mu^2) \pm i\mu(4M^2 - \mu^2)^{\frac{1}{2}} \times [W^2 - (M + \mu)^2]^{\frac{1}{2}}[W^2 - (M - \mu)^2]^{\frac{1}{2}}\}, \quad (4.1)$$

coinciding with the expression given by Landau's method. The different configurations are shown on Fig. 10, in the case $M' = M$, together with the corresponding cuts in the s plane (thick lines). One sees that the integral (3.4) giving $F_l(s, W)$ will have an end-point singularity for $W^2 = M^2 + 2\mu^2$ due to the presence of B_l in the integrand. To avoid this we must give W^2 a small positive imaginary part, corresponding to the fact that W actually is the center-of-mass energy of the incident system $N + \pi$. Then the cuts in the s plane move as shown on Fig. 10 (dashed lines). One might wonder whether the amplitude A_l has also kinematic cuts coming from the $d(\cos\theta) = (1/2pq)\,dt$ term in the expression

$$\frac{1}{2}\int_{-1}^{+1} A(s, t, u)P_l(\cos\theta)d(\cos\theta).$$

In fact, this is not the case because changing the sign of pq is the same thing as changing $\cos\theta$ into $-\cos\theta$, which means exchanging the outgoing pions, and $A(s, t, u)$ is symmetrical with respect to that exchange.

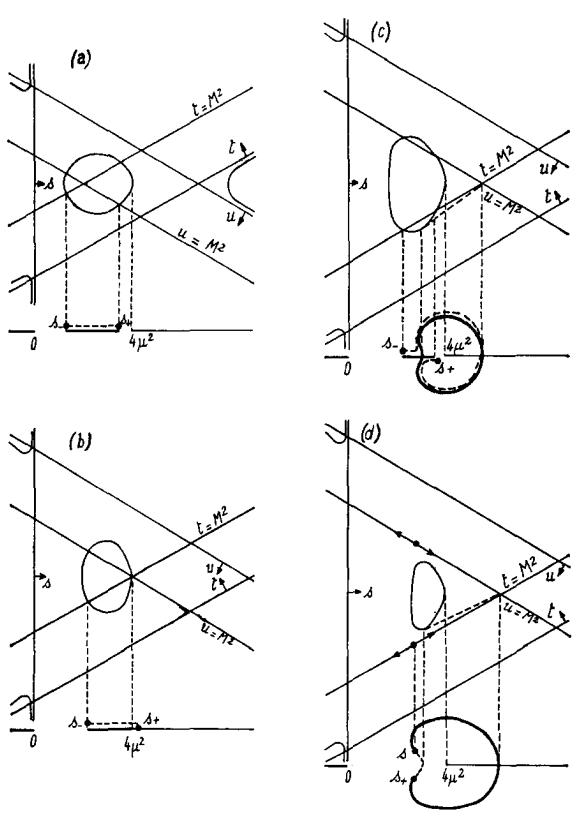


FIG. 10. Mandelstam's diagrams for increasing values of W^2. (a) $W^2 < M^2 + 2\mu^2$; (b) $W^2 = M^2 + 2\mu^2$; (c) $M^2 + 2\mu^2 < W^2 < (M + \mu)^2$; (d) $(M + \mu)^2 < W^2 < (M + 2\mu)^2$. Under each diagram are shown the lines in the complex s plane where the series of partial waves diverges (thick lines). Dashed lines are obtained by giving W^2 a small positive imaginary part.

It is easily seen on Fig. 10 that when W^2 becomes greater than $M^2 + 2\mu^2$, we must deform the contour of the integral (3.4) in the lower half-plane. Thus (3.4) becomes

$$F_l(s, W) = \frac{-1}{D_l(s)}\int_{C_W} 4q's'^{-\frac{1}{2}}\frac{N_l(s')B_l(s', W)}{s' - s}\,ds', \quad (4.2)$$

where C_W is the contour shown on Fig. 11 in the case $W^2 > (M + \mu)^2$, and the minus sign arises from the way we have cut the q function (the cuts chosen for p and q are also shown on Fig. 11).

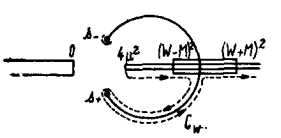


FIG. 11. The contour C_W, with the cuts for the functions p and q.

One sees the importance of cutting the amplitude as we did. If we had chosen the cut, let us say, along the trajectories of s_+ and s_-, we would have calculated B_l on another sheet and obtained surprising results like an infinite value of the amplitude at $s = 4\mu^2$. This means that one must not simply continue analytically in W for s fixed, but continue both in W and s, and the safest way to do this is to chose cuts in the s plane according to our prescription; that is, to move only in regions where the series of Legendre polynomials converges, i.e., where partial waves have a meaning.

5. SIMPLIFIED FORM OF THE SOLUTION

We want to evaluate the integral

$$I(s) = \int_{C_W} 4q's'^{-\frac{1}{2}}\frac{N(s')B(s', W)}{s' - s}\,ds' \quad (5.1)$$

(here and in the following, we drop the index l). Remembering that $q(s) = -q^*(s^*)$, we get

$$I^*(s^*) = -\int_{C_{W^*}} \text{same integrand}, \quad (5.2)$$

where C_W^* is the complex conjugate of the contour C_W. Thus,

$$I(s) + I^*(s^*) = \int_{C_W - C_{W^*}} \text{same integrand}, \qquad (5.3)$$

and by Cauchy's theorem,

$$I(s) + I^*(s^*) = -2i\pi(4q)s^{-\frac{1}{2}}N(s)B(s, W) + \int_{C'} 4q's'^{-\frac{1}{2}} \frac{N(s')B(s', W)}{s' - s}\, ds', \qquad (5.4)$$

where C' is the contour $(-\infty + i\epsilon, 0, -\infty - i\epsilon)$. Analogously,

$$I(s) - I^*(s^*) = \int_{C_W + C_{W^*}} \text{same integrand}. \qquad (5.5)$$

But if we set

$p_1 = p$; $q_1 = q$ in the upper half-plane,

$p_1 = -p$; $q_1 = -q$ in the lower half-plane

(the cuts of the functions p_1 and q_1 are shown on Fig. 12), we get

$$\begin{aligned} I(s) - I^*(s^*) &= \int_{C_{W^*} - C_W} 4q_1's'^{-\frac{1}{2}} \frac{N(s')B(s', W)}{s' - s}\, ds' \\ &= \int_K 4q_1's'^{-\frac{1}{2}} \frac{N(s')B(s', W)}{s' - s}\, ds' \\ &= f(W)g(-4i\pi) \int_{s_+}^{s_-} \frac{N(s')}{p_1's'^{\frac{1}{2}}(s' - s)}\, ds' \end{aligned} \qquad (5.6)$$

(K is the contour shown on Fig. 12).

Setting (5.4) and (5.6) together, we get

$$\begin{aligned} I(s) = &-4i\pi qs^{-\frac{1}{2}}N(s)B(s, W) \\ &+ \int_{C'} 2q's'^{-\frac{1}{2}} \frac{N(s')B(s', W)}{s' - s}\, ds' \\ &- 2i\pi f(W)g \int_{s_+}^{s_-} \frac{N(s')}{p_1's'^{\frac{1}{2}}(s' - s)}\, ds'. \end{aligned} \qquad (5.7)$$

Hence, substituting (5.7) in (3.3), we find

$$\begin{aligned} A(s, W) &= B(s, W) - I(s, W)/D(s) \\ &= B(s, W)e^{i\delta} \cos\delta + [1/D(s)] \\ &\quad \times \left[-\int_{C'} 2q's'^{-\frac{1}{2}} \frac{N(s')B(s', W)}{s' - s}\, ds' \right. \\ &\quad \left. + 2i\pi f(W)g \int_{s_+}^{s_-} \frac{N(s')}{p_1's'^{\frac{1}{2}}(s' - s)}\, ds' \right]. \end{aligned} \qquad (5.8)$$

This form can be further simplified if we suppose that all nearby singularities of the $(\pi - \pi)$ amplitude are on the second Riemann sheet (corresponding either to resonances or to "virtual" bound states). In that case, $N(s)$ has none of these singularities and is a slowly varying function in the physical region. One thus reads at once from equation (3.6) that

$$1/D(s) = \text{slowly varying function} \times e^{i\delta} \sin\delta/q.$$

On the other hand, the integrals in (5.8) are slowly varying in the interesting region provided the singularities of the $(\pi - \pi)$ amplitude are much nearer $4\mu^2$ than the contours C' and (s_+s_-). In that case $B(s, W)$ also is slowly varying, so that we may write

$$A(s, W) \simeq B(W)e^{i\delta}[\cos\delta + (q_0/q)\sin\delta], \qquad (5.9)$$

where q_0 is a complex constant whose real part arises from the integral over C' and imaginary part from the integral over (s_+s_-).

One thus gets the usual approximation[7] for final-state interaction.

FIG. 12. The contour K, with the cuts for the functions p_1 and q_1.

6. DISCUSSION AND CONCLUSION

Our model shows that the usual approximation for final-state interaction is valid even in the case where there are complex singularities, provided the intermediary particles giving these singularities are stable enough [that is, provided the contour (s_+s_-) is far enough from the physical region, or equivalently, the pole lines on Mandelstam's diagram are greatly separated from the physical region]. This result is not surprising because if the intermediary particles in the t and u channels were near the physical region, we should have to deal with multiple final-state interaction and should not hope to get simple results by projecting onto the partial waves of one channel. In that case, $A(s, \cos\theta)$ would be strongly dependent upon $\cos\theta$, higher-order partial waves would be important, and the problem of summing up the series of Legendre polynomials could not be avoided. This shows the importance of our choice of cuts—although it may seem a complicated one in the case of multiple final-state interaction—if one wants to take into account unitarity in the crossed channels. Perhaps in that case it would be better to avoid considering partial waves altogether when using unitarity.

ACKNOWLEDGMENTS

We would like to thank Professor M. Froissart and Professor R. Stora for fruitful and illuminating discussions.

[7] K. M. Watson, Phys. Rev. **88**, 1163 (1952).

Regge Poles in High-Energy Potential Scattering

S. TANI*
Department of Physics, New York University, University Heights, Bronx, New York
(Received 6 May 1963)

High-energy potential scattering based on the Schrödinger equation has been investigated using Regge poles. When the Watson–Sommerfeld (W–S) integral is directly applied in the left half-plane, it is not useful at all in order to derive the scattering amplitude but gives a condition which governs the behavior of Regge poles. They are distributed with residues such that their contributions cancel against each other in the nonforward direction, provided that the integral at infinity is ignored. When Mandelstam's technique is used, the Legendre function of the first kind and the secant in the original W–S integral are replaced by the Legendre function of the second kind and the cosecant, respectively. This enables one to sum over partial waves more easily than with the conventional series. In high-energy potential scattering one cannot single out an individual Regge pole. The same conclusion follows when Khuri's representation is used.

I

THE Regge trajectory[1] for nonrelativistic potential scattering has been investigated by Newton.[2] In several examples it has been studied in detail by Singh,[3] and by Barut and Calogero.[4] The high-energy limit has been explored for a certain class of potential by Bethe and Kinoshita.[5] Since a Regge pole is moving to the left along the real axis in the angular momentum plane with decreasing negative energy, there is a certain negative energy below which all Regge poles are shifted to the left half-plane (Re $l < -\frac{1}{2}$). Admitting that the limit of infinite energy is realized independently of the direction in the complex energy plane for the class of potential described below, one expects that all Regge poles are distributed in the left half-plane at high energies. The analytical continuation of the scattering amplitude into the left half-plane has been studied by Mandelstam[6] and more recently by Khuri.[7]

In this note the high-energy potential scattering will be studied off the forward direction for a *physical* value of scattering angle. This problem has been studied extensively using conventional methods. For some common class of potential, which includes the Yukawa potential, the asymptotic expansions of both the phase shift and the scattering amplitude have been made by Rosendorff and Tani.[8] In that paper the Klein-Gordon equation has been studied, but with relatively minor changes one obtains the results for the Schrödinger equation. The potential is assumed to be expansible in the form

$$V(r) = V_{-1}r^{-1} + V_0 + V_1 r + V_2 r^2 + \cdots \qquad (1)$$

around the origin, where some one of the coefficients of an odd order does not vanish; it is also assumed that it is infinitely differentiable everywhere. A Regge pole for the same class of potential has been investigated by Bethe and Kinoshita.[5] It is our purpose to review the results of I using Regge poles. In conclusion it will be seen that Regge poles are aids to summing over partial waves, but a particular Regge pole (in the left half-plane) cannot be singled out as giving a dominant contribution, in sharp contrast to a Regge pole (in the right half-plane) for a hard core.[9]

For the class of potential, (1), both the scattering amplitude and the phase shift can be well approximated by their first Born approximations. The asymptotic expansion of the first-order phase shift takes the form[10]

* Supported by the National Science Foundation.
[1] T. Regge, Nuovo Cimento **14**, 951 (1959); **18**, 947 (1960); A. Bottino, A. M. Longoni, and T. Regge, Nuovo Cimento **23**, 954 (1962).
[2] R. G. Newton, J. Math. Phys. **3**, 867 (1962).
[3] V. Singh, Phys. Rev. **127**, 632 (1962).
[4] A. O. Barut and F. Calogero, Phys. Rev. **128**, 1383 (1962).
[5] H. A. Bethe and T. Kinoshita, Phys. Rev. **128**, 1418 (1962).
[6] S. Mandelstam, Ann. Phys. (N. Y.) **19**, 254 (1962).
[7] N. N. Khuri, Phys. Rev. **130**, 429 (1963). The author would like to express his gratitude to Professor B. Zumino for calling his attention to this work.
[8] S. Rosendorff and S. Tani, Phys. Rev. **128**, 457 (1962); this paper will be referred to as I. Further references on high-energy potential scattering can be found in this paper.
[9] A. J. W. Sommerfeld, *Partial Differential Equations in Physics* (Academic Press Inc., New York, 1949), p. 279; B. R. Levy and J. B. Keller, Communs. Pure Appl. Math. **12**, 159 (1959).
[10] This is a theorem which can be established after some algebra using Gegenbauer's addition theorem for Bessel functions. The use of an auxiliary function given in I enables one to obtain the expansion up to a desired order much more easily.

$$\delta(k, l) = A_0(k) + l(l+1)A_1(k) + (l-1)l(l+1)(l+2)A_2(k) + \cdots + \psi(l+1)[B_0(k) + l(l+1)B_1(k) + (l-1)l(l+1)(l+2)B_2(k) + \cdots], \quad (2)$$

where $\psi(l+1)$ is the logarithmic derivative of the gamma function, and A's and B's are functions of momentum which becomes small at high energies,

$$A_0(k) = -V_{-1}(\log k/2k) - \lim_{\epsilon\to 0}\left[V_{-1}\log\epsilon + \int_\epsilon^\infty V(r)\,dr\right]\Big/ 2k, \quad \text{etc.},$$

$$B_0(k) = V_{-1}(1/2k),$$

$$B_1(k) = V_1(1/4k^3),$$

$$B_2(k) = V_3(3/16k^5), \quad \text{etc.}$$

When the phase shift given by (2) is substituted in the sum for the amplitude one obtains

$$f(k, z) = \frac{1}{2ik}\sum_{l=0}^{\infty}(2l+1)\{\exp[2i\delta(k, l)] - 1\}P_l(z),$$

where z is the cosine of the scattering angle. Only a term involving $\psi(l+1)$ turns out to be essential in order to derive an asymptotic expansion, which decreases as only some inverse power of k off the forward direction.[11] The sum for

$$\sum_{l=0}(2l+1)(l-n)\cdots l(l+1)\cdots(l+n+1)P_l(z) \quad (4a)$$

or

$$\sum_{l=0}^{\infty}(2l+1)(l-n)\cdots l(l+1)\cdots(l+n+1)\times\psi(l+1)P_l(z) \quad (4b)$$

can be evaluated by using the well-known formula

$$\sum_{l=0}^{\infty}x^lP_l(z) = (1-2xz+x^2)^{-\frac{1}{2}}, \quad (5)$$

or some other formula derived from it.

II

As usual the sum over partial waves, (3), is converted into the Watson-Sommerfeld (W-S) integral

$$f(k, z) = (1/2k)\int_C d\lambda\,\lambda[S(k, \lambda-\tfrac{1}{2}) - 1] \times \sec\pi\lambda P_{\lambda-\frac{1}{2}}(-z), \quad (6)$$

where $\lambda = l + \frac{1}{2}$. If it is understood that there is no Regge pole in the right half-plane (Re $\lambda > 0$), the contour C may be taken downwards along the imaginary axis. When the contour is shifted into the left half-plane, there will be contributions from negative half-integers as well as from Regge poles. Using the relation

$$P_{-\nu}(z) = P_{\nu+1}(z) \quad (7)$$

and the so-called Mandelstam symmetry[12] for half-integral λ,

$$S(k, -\lambda-\tfrac{1}{2}) = -S(k, \lambda-\tfrac{1}{2}), \quad (8)$$

the contribution from negative half-integers turns out to be

$$f_N(k, z) = \frac{1}{2ik}\sum_{l=0}^{\infty}(2l+1)[S(k, l)+1]P_l(z), \quad (9)$$

if it is allowed to shift the contour C to the left without limit. We have to study the S matrix as an analytic function of λ for a given class of potential in order to justify the last statement. We do not make such an analysis here, but the reader can be referred to Froissart's theory[13] which establishes that the S matrix for the potential considered here is meromorphic in the whole λ plane, except at infinity. On the right side of (9) the contribution of the incident wave to the forward amplitude appears to have reversed its sign as compared to the original sum, (3). Therefore, the difference between them,

$$f_N(k, z) - f(k, z) = \frac{1}{ik}\sum_{l=0}^{\infty}(2l+1)P_l(z), \quad (10)$$

must be compensated by the contribution from Regge poles and the integral at infinity, if the latter cannot be ignored. Denoting the contribution from Regge poles by $f_R(k, z)$, we then find that

$$f_R(k, z) = \frac{i}{k}\sum_{l=0}^{\infty}(2l+1)P_l(z) - I(\infty), \quad (11)$$

where $I(\infty)$ is the W–S integral along the contour going around counterclockwise with an infinite radius.

Suppose the nth Regge pole is located at $\lambda = \alpha_n(K)$, and set

$$\lim_{\lambda\to\alpha_n}(\lambda-\alpha_n)\lambda S(k, \lambda-\tfrac{1}{2}) = -\beta_n(k). \quad (12)$$

[11] The terms involving A's consist of polynomials of an even order in $\lambda = l + \frac{1}{2}$. If all B's are set equal to zero, the amplitude is given by a sum of terms of the form (4a). Each term, (4a), vanishes for $z < 1$.

[12] See reference 6; this is proved in reference 2 for potentials with finite first and second moments. For the Coulomb potential, which does not belong to this class, the symmetries, (8) and (18), follow directly from the S matrix, (23).

[13] M. Froissart, J. Math. Phys. 3, 922 (1962); see also reference 2.

Writing $f_R(k, z)$ explicitly in terms of α's and β's, Eq. (11) reads

$$\sum_n \frac{i}{k} \frac{\pi\beta_n}{\cos \pi\alpha_n} P_{\alpha_n - \frac{1}{2}}(-z) = \frac{i}{k} \sum_{l=0}^{\infty} (2l+1) P_l(z) - I(\infty). \quad (13)$$

This means that, if the W-S integral is directly applied to the left plane, the Regge poles are not useful at all in order to derive the scattering amplitude. Instead a condition, Eq. (13), is obtained concerning the high-energy behavior of Regge poles for potential scattering. [*Note added in proof*: Equation (13) is a general result for the sum over all Regge poles when the S matrix is meromorphic in the entire λ plane. The author would like to express his gratitude to Professor M. A. Ruderman for calling his attention to this point.]

As will be shown in Sec. IV, the integral may be ignored in a nonforward direction in the perturbation calculation with Coulomb potential. If this is the case, we may set

$$\sum_n \frac{i}{k} \frac{\pi\beta_n(k)}{\cos \pi\alpha_n(k)} P_{\alpha_n - \frac{1}{2}}(-z) = 0. \quad (14)$$

The residues and the location of Regge poles at high energies are such that their contributions cancel against each other. Obviously, one cannot single out an individual Regge pole in this calculation.

III

The contribution from negative half-integers can be dropped if we use Mandelstam's technique.[6] Using the identity

$$\frac{Q_{-\lambda-\frac{1}{2}}(z)}{\pi \sin \pi\lambda} = \frac{P_{\lambda-\frac{1}{2}}(z)}{\cos \pi\lambda} + \frac{Q_{\lambda-\frac{1}{2}}(z)}{\pi \sin \pi\lambda}, \quad (15)$$

Eq. (6) can be replaced by

$$f(k, z) = \frac{1}{2\pi k} \int_c d\lambda\, \lambda [S(k, \lambda - \tfrac{1}{2}) - 1] Q_{-\lambda-\frac{1}{2}}(z) \csc \pi\lambda + \frac{1}{i\pi k} \sum_{n=1}^{\infty} n [S(k, n - \tfrac{1}{2}) - 1](-1)^n Q_{n-\frac{1}{2}}(-z). \quad (16)$$

The Legendre function of the second kind for the physical value of z is defined by

$$Q_\lambda(z) = \tfrac{1}{2} \lim_{\epsilon \to 0} [Q_\lambda(z + i\epsilon) + Q_\lambda(z - i\epsilon)], \quad \text{real } z, \quad |z| < 1. \quad (17)$$

The contour C in Eq. (16) runs downwards along the imaginary axis, as in Eq. (6). If the contour is shifted into the left half-plane, the contribution from negative integers cancels the sum in (16) because of the Mandelstam symmetry for integral λ[14]

$$S(k, -n - \tfrac{1}{2}) = S(k, n - \tfrac{1}{2}), \quad \text{integral } n. \quad (18)$$

Therefore, if the remaining integral is ignored, we have only the contribution from Regge poles left,

$$f(k, z) = \frac{i}{k} \sum_n \frac{\beta_n(k)}{\sin \pi\alpha_n(k)} Q_{-\alpha_n+\frac{1}{2}}(-z). \quad (19)$$

If one recalls the asymptotic form of the Legendre function, only the Regge pole which is nearest to the origin may be retained for large $|z|$. However, it is not justified to interpret each term as a representative of a resonance for a physical value of z. Each term on the right side of (19) can be expanded into Legendre polynomials using the formula

$$\int_{-1}^{1} Q_\sigma(z) P_l(z)\, dz = \frac{1 - \cos \pi(l - \sigma)}{(l - \sigma)(l + \sigma + 1)}, \quad \text{integral } l. \quad (20)$$

There is no term which dominates for $l = \text{Re } \sigma$ in this expansion. In fact, for a physical value of z, all Regge poles make a collective contribution to the first Born approximation of the amplitude, the leading term at high energies. In the example of Coulomb potential, Mandelstam's technique is quite useful because the summation over partial waves can be made more easily by starting with (19) than with (3).

An ingenious representation of the amplitude has been developed by Khuri[7] which is most convenient to locate the singularities. Using this method one obtains

$$f(k, z) = \frac{i}{2\sqrt{2}k} \sum_n \frac{\beta_n}{\alpha_n} \int_\xi^\infty \frac{\exp(\alpha_n x) \sinh x}{(\cosh x - z)^{\frac{1}{2}}} dx, \quad (21)$$

where

$$\xi = \cosh^{-1}(1 + 1/2k^2).$$

Expanding the right side into Legendre polynomials it assumes the form

$$f(k, z) = \frac{1}{2ik} \sum_n \sum_{l=0}^{\infty} \frac{\beta_n}{\alpha_n(\alpha_n - l - \frac{1}{2})} \times \exp(\alpha_n - l - \tfrac{1}{2})\xi (2l+1) P_l(z). \quad (22)$$

Again each term does not stand for a resonance because all Regge poles are in the left half-plane (Re $\alpha_n < 0$); one cannot single out an individual

[14] Specifically for integral λ (or half-integral l), no new information is gained by changing from positive λ to negative λ. The potential-free wavefunction is given in terms of a Bessel function of an integral order, which is invariant (except for a sign) to the change of the sign of its order. Thus, the same S matrix follows for a negative integral λ as for the corresponding positive integral λ, if the potential is of short range decreasing faster than r^{-2}.

Regge pole in order to obtain the first Born approximation for the amplitude.

IV

We shall show how the analyses mentioned above apply to the example of Coulomb potential. The S matrix as a function of λ is[3]

$$S(k, \lambda - \tfrac{1}{2}) = \frac{\Gamma(\lambda + \frac{1}{2} - i(e^2/2k))}{(\Gamma\lambda + \frac{1}{2} + i(e^2/2k))}. \quad (23)$$

The location of the nth Regge pole is

$$\alpha_n(k) = -n - \tfrac{1}{2} + i\, e^2/2k, \qquad n = 0, 1, 2, \cdots, \quad (24)$$

while the Born expansion of the residue is given by

$$\beta_n(k) = i(e^2/2k)(2n + 1) + \tfrac{1}{2}(e^2/k)^2[1 + (2n + 1) \times \psi(n + 1)] + O(e^6). \quad (25)$$

Using (7), the left side of (14), the amplitude which comes from Regge poles, reads

$$f_R(k, z) = \frac{i\pi}{k \cdot \sin(i\pi e^2/2k)} \sum_{n=0}^{\infty} (-1)^n \beta_n(k) P_{n - i(e^2/2k)}(-z). \quad (26)$$

Using the formula for integral n and nonintegral δ,

$$P_{n+\delta}(z) = \frac{\sin \pi\delta}{\pi} \times \sum_{m=0}^{\infty} \frac{(-1)^{m+n}(2m + 1)}{(n - m + \delta)(n + m + 1 + \delta)} P_m(z), \quad (27)$$

the Born expansion of $P_{n - i(e^2/2k)}$ is given by

$$P_{n - i(e^2/2k)}(z) = P_n(z) + i\frac{e^2}{2k}\left[\frac{1}{2n + 1} P_n(z) + \sum_{m=0}^{\infty}{}' (-1)^{m+n}\left(\frac{1}{m - n} + \frac{1}{n + m + 1}\right) P_m(z)\right] + O(e^4). \quad (28)$$

Let us recall

$$\psi(l + 1) = \sum_{m=1}^{l} \frac{1}{m} + \lim_{N\to\infty}\left(\log N - \sum_{n=1}^{N} \frac{1}{n}\right), \quad (29)$$

and substitute (25) and (28) into (26). After some algebra we find

$$f_R(k, z) = \frac{i}{k} \sum_{l=0}^{\infty} (2l + 1) P_l(z) + \frac{e^2}{k^2} \lim_{N\to\infty} (\log N) \sum_{l=0}^{\infty} (2l + 1) P_l(z) + \cdots, \quad (30)$$

which confirms Eq. (14) up to the first order in e^2. The first term on the right of (30) is the same as the first term on the right of (13), while the second term of (30) should come from the integral $I(\infty)$ in (13).

When Mandelstam's technique is used, we have the amplitude $f(k, z)$, according to (19),

$$f(k, z) = \frac{-i}{k \cos(ie^2/2k)} \sum_{n=0}^{\infty} (-1)^n \beta_n(k) Q_{n - i(e^2/2k)}(-z). \quad (31)$$

Using (25), the first term of the Born expansion is given by

$$f_1(k, z) = \frac{e^2}{2k^2} \sum_{l=0}^{\infty} (-1)^l (2l + 1) Q_l(-z). \quad (32)$$

When one starts with the formula

$$\sum_{n=0}^{\infty} x^n Q_n(z) = (1 - 2xz + x^2)^{-\frac{1}{2}} \times \log\{[z - x + (1 - 2xz + x^2)^{\frac{1}{2}}](1 - z^2)^{-\frac{1}{2}}\}, \quad (33)$$

one can easily derive

$$\sum_{n=0}^{\infty} x^n (2n + 1) Q_n(z) = \left(2x\frac{d}{dx} + 1\right) \sum_{n=0}^{\infty} x^n Q_n(z) = -2x(1 - 2zx + x^2)^{-1} + \frac{1 - x^2}{(1 - 2zx + x^2)^{\frac{3}{2}}} \times \log\left[\frac{z - x + (1 - 2xz + x^2)^{\frac{1}{2}}}{(1 - z^2)^{\frac{1}{2}}}\right]. \quad (34)$$

Then it is straightforward to obtain the sum in (32) by putting $x = -1$ and $z \to -z$ in (34). The result is well-known:

$$f_1(k, z) = (e^2/2k^2)[1/(1 - z)]. \quad (35)$$

The above calculation is much simpler than the conventional sum over partial waves, (3). The Born expansion of the S matrix, (23), leads to

$$S(k, l) = 1 - i(e^2/k)\psi(l + 1) + \cdots, \quad \text{integral } l. \quad (36)$$

Therefore, the conventional sum over partial waves starts with

$$f(k, z) = -\frac{e^2}{2k^2} \sum_{l=0}^{\infty} (2l + 1)\psi(l + 1) P_l(z) + \cdots. \quad (37)$$

The formula (5) plays an essential role in summing the right side of (37). The calculation is straightforward, but fairly tedious; the reader is referred to, for instance, Appendix C of I. We have seen here that Mandelstam's technique is very useful in the calculation of the amplitude.

When Khuri's representation is used, we may substitute (24) and (25) into (22). Then we obtain, using (29), the right side of (37) and a term similar to the second term in (30), so that this method does not seem to be useful in order to derive the Born approximations for the amplitude. It is necessary to scrutinize the limit of various integrals as the contour recedes to infinity in order to analyze the nature of the latter term, but we do not attempt such an analysis in this paper.

In conclusion it has been shown for the class of potential described in (1) that:

(i) If the W-S integral is applied directly to the left half-plane, the Regge poles are not useful in order to derive the amplitude, but one obtains a condition, (13) or (14), which governs their behavior;

(ii) by using Mandelstam's technique the amplitude can be derived more easily than with the conventional sum over partial waves; and

(iii) an individual Regge pole cannot be singled out in order to derive the amplitude for the physical value of z. The conclusion remains the same when analyzed using Khuri's representation.

ACKNOWLEDGMENT

The author would like to express his sincere gratitude to Professor S. Borowitz and Dr. S. Rosendorff for stimulating conversations.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 4, NUMBER 10 OCTOBER 1963

Computation of Matrix Elements of Integrable Nuclear Two-Particle Interaction Operators*

R. K. NESBET†
Nuclear Engineering Department, Brookhaven National Laboratory, Upton, Long Island, New York
(Received 26 April 1963)

A general procedure is described for computing matrix elements of operators that occur in a realistic nuclear Schrödinger Hamiltonian, in a basis of orbitals with radial factors of the form $N_a r^{2n_a+l_a} \exp(-\gamma_a r^2)$, where n_a is an arbitrary nonnegative integer and γ_a is an arbitrary positive number. The method is suitable for efficient large-scale computation of these matrix elements, needed when orbitals of physical interest (such as Hartree–Fock orbitals) are expressed as linear combinations of basis functions of the kind indicated. The analysis of matrix elements of a tensor operator provides a new method of reduction to linearly independent reduced matrix elements, the number of which is smaller than it is in the usual analysis. Thus, the number of independent parameters in the corresponding empirical theory of complex nuclear spectra is reduced. The specific operator forms considered here are those present in the asymptotic one-pion exchange potential, except that the functional forms of the potentials $V(r)$ are not prescribed beyond the requirement that they should be integrable.

I. INTRODUCTION

In recent years, with the advent of the electronic digital computer, it has become possible to work towards the actual solution of the many-particle Schrödinger equation for atoms and small molecules.[1] The technique which appears to be most fruitful at present is to obtain first a good approximate solution to the Hartree–Fock equations, using the matrix Hartree–Fock method proposed by Roothaan,[2] which can easily be generalized for use with open-shell configurations. Following this, the effects of interparticle correlation are evaluated by means of perturbation theory, which may vary in complexity depending on the nature of the problem at hand from the simple second-order formalism to the more sophisticated theory of Brueckner, based on the Bethe–Goldstone equation, and to the consideration of collective excitations. The beginning step in analysis of this kind is the evaluation of matrix elements of the one- and two-particle operators in the Schrödinger Hamiltonian of the many-particle system, in a basis of arbitrary functions chosen for ease of evaluation of these matrix elements and for their qualitative resemblance to the Hartree–Fock orbitals of the system.

* Work performed under the auspices of the U. S. Atomic Energy Commission.
† Present address: IBM Research Laboratory, San Jose, California.
[1] R. K. Nesbet, Rev. Mod. Phys. **33**, 28 (1961).
[2] C. C. J. Roothaan, Rev. Mod. Phys. **23**, 69 (1951).

When Khuri's representation is used, we may substitute (24) and (25) into (22). Then we obtain, using (29), the right side of (37) and a term similar to the second term in (30), so that this method does not seem to be useful in order to derive the Born approximations for the amplitude. It is necessary to scrutinize the limit of various integrals as the contour recedes to infinity in order to analyze the nature of the latter term, but we do not attempt such an analysis in this paper.

In conclusion it has been shown for the class of potential described in (1) that:

(i) If the W-S integral is applied directly to the left half-plane, the Regge poles are not useful in order to derive the amplitude, but one obtains a condition, (13) or (14), which governs their behavior;

(ii) by using Mandelstam's technique the amplitude can be derived more easily than with the conventional sum over partial waves; and

(iii) an individual Regge pole cannot be singled out in order to derive the amplitude for the physical value of z. The conclusion remains the same when analyzed using Khuri's representation.

ACKNOWLEDGMENT

The author would like to express his sincere gratitude to Professor S. Borowitz and Dr. S. Rosendorff for stimulating conversations.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 4, NUMBER 10 OCTOBER 1963

Computation of Matrix Elements of Integrable Nuclear Two-Particle Interaction Operators*

R. K. NESBET†
Nuclear Engineering Department, Brookhaven National Laboratory, Upton, Long Island, New York
(Received 26 April 1963)

A general procedure is described for computing matrix elements of operators that occur in a realistic nuclear Schrödinger Hamiltonian, in a basis of orbitals with radial factors of the form $N_a r^{2n_a+l_a} \exp(-\gamma_a r^2)$, where n_a is an arbitrary nonnegative integer and γ_a is an arbitrary positive number. The method is suitable for efficient large-scale computation of these matrix elements, needed when orbitals of physical interest (such as Hartree–Fock orbitals) are expressed as linear combinations of basis functions of the kind indicated. The analysis of matrix elements of a tensor operator provides a new method of reduction to linearly independent reduced matrix elements, the number of which is smaller than it is in the usual analysis. Thus, the number of independent parameters in the corresponding empirical theory of complex nuclear spectra is reduced. The specific operator forms considered here are those present in the asymptotic one-pion exchange potential, except that the functional forms of the potentials $V(r)$ are not prescribed beyond the requirement that they should be integrable.

I. INTRODUCTION

In recent years, with the advent of the electronic digital computer, it has become possible to work towards the actual solution of the many-particle Schrödinger equation for atoms and small molecules.[1] The technique which appears to be most fruitful at present is to obtain first a good approximate solution to the Hartree–Fock equations, using the matrix Hartree–Fock method proposed by Roothaan,[2] which can easily be generalized for use with open-shell configurations. Following this, the effects of interparticle correlation are evaluated by means of perturbation theory, which may vary in complexity depending on the nature of the problem at hand from the simple second-order formalism to the more sophisticated theory of Brueckner, based on the Bethe–Goldstone equation, and to the consideration of collective excitations. The beginning step in analysis of this kind is the evaluation of matrix elements of the one- and two-particle operators in the Schrödinger Hamiltonian of the many-particle system, in a basis of arbitrary functions chosen for ease of evaluation of these matrix elements and for their qualitative resemblance to the Hartree–Fock orbitals of the system.

* Work performed under the auspices of the U. S. Atomic Energy Commission.
† Present address: IBM Research Laboratory, San Jose, California.
[1] R. K. Nesbet, Rev. Mod. Phys. **33**, 28 (1961).
[2] C. C. J. Roothaan, Rev. Mod. Phys. **23**, 69 (1951).

The present paper describes a method of evaluating matrix elements of the operators in a physically realistic nuclear Hamiltonian, for a basis set of sufficient generality that one can expect to fit nuclear Hartree–Fock orbitals with linear combinations of a relatively small number of basis orbitals. The purpose of this work is to make a start towards quantitative computation of nuclear energy levels, transition probabilities, and other properties of nuclear wavefunctions.

The structure of the two-body interaction that must be included in a nuclear Hamiltonian is still a subject for speculation. The asymptotic internucleon interaction coming from meson theory, the one-pion exchange potential, is now fairly well established, but the inner region is still indefinitely determined either by empirical data or by meson theory.[3] The present paper will treat only the forms of interaction that occur in the one-pion exchange potential: integrable velocity-independent scalar and tensor interactions. The functional form of the potential is left indefinite except for the requirement of integrability, which excludes the extreme form of hard-core interaction (infinite in a finite region of configuration space) but includes many functional forms that are large and positive at short range, and possibly singular at an isolated point ($r_{12} = 0$) in configuration space. An example of such a potential is that used by Goldhammer[4] for calculations of the binding energy of O^{16} and of the doublet splitting in the $1p$ shell due to the tensor interaction. Goldhammer's results are in reasonable agreement with experiment.

The analysis of the tensor interaction given here is a new method of reduction of matrix elements to independent components. It can be shown that the usual formulation[5] in terms of quantum numbers L', L for the initial and final two-particle states leads to reduced matrix elements $(L \parallel L')$ that are not, in general, linearly independent. The present analysis, by coupling the angular momenta in each one-particle space to a definite quantum number, leads to linearly independent reduced matrix elements $I^{kk'}$. An extension of the empirical theory of complex atomic spectra to the nuclear case should be based on these parameters and not on the $(L \parallel L')$ reduced matrix elements.

Computational technique is stressed in the present paper, to the extent that it should be possible to write a fairly efficient digital computer program from the information given here. Such a program has been written for the IBM 7090 by the present author, and tested with the nuclear Hamiltonian used by Goldhammer. Some of the apparently arbitrary procedures described here are justified in that they simplify the resulting program and make it easier to write and to check.

Formulas for one-particle integrals (overlap and kinetic energy) are included, for completeness, in Sec. V, together with formulas for matrix elements of the two-particle operator $(\hbar^2/M)(\nabla_1 \cdot \nabla_2)$.

II. MATRIX ELEMENTS OF A SCALAR OPERATOR

Let $V(r)$ be an arbitrary function of the interparticle distance $r = r_{12}$. It is well known that all matrix elements $(ab \mid V \mid cd)$ of $V(r)$, in a basis of single-particle wavefunctions (orbitals)

$$\eta_a = R_a(r)Y(l_a, m_{la}; \theta, \phi)\chi(m_{sa}, m_{ta}), \tag{1}$$

can be expressed as linear combinations of the linearly independent Slater parameters

$$F^k(ac \mid db) = \int_0^\infty r_1^2\, dr_1 \int_0^\infty r_2^2\, dr_2 R_a^*(r_1) \times R_c(r_1) f_k(r_1, r_2) R_d(r_2) R_b^*(r_2), \tag{2}$$

where k is restricted to values obtainable by coupling together both pairs (l_a, l_c) and (l_d, l_b) to a common value k of the total orbital angular momentum.

Here

$$V(r) = \sum_{k=0}^{\infty} f_k(r_1, r_2) P_k(\cos\omega), \tag{3}$$

or conversely,

$$f_k(r_1, r_2) = (k + \tfrac{1}{2}) \int_0^\pi V(r) P_k(\cos\omega) \sin\omega\, d\omega, \tag{4}$$

where ω is the angle from $\mathbf{r}_1$ to $\mathbf{r}_2$.

If $V(r)$ is multiplied by an operator depending on spin or isotopic spin, the spatial matrix elements are multiplied by factors that depend only on the corresponding quantum numbers m_{sa}, m_{ta}. Since this is irrelevant to the present discussion, it will be assumed that $m_{sa} = m_{sc}$, $m_{ta} = m_{tc}$, $m_{sd} = m_{sb}$, $m_{td} = m_{tb}$, and the spin-isotopic spin functions χ will be ignored. The spin quantum numbers must of course be taken into account in discussing the tensor interaction.

It has been shown by Talmi[6] that when the orbitals

[3] See, for example, N. K. Glendenning and G. Kramer, Phys. Rev. **126**, 2159 (1962). These authors give earlier references.

[4] P. Goldhammer, Phys. Rev. **116**, 676 (1959); **125**, 660 (1962).

[5] J. P. Elliott, Proc. Roy. Soc. (London) **A218**, 345 (1953).

[6] I. Talmi, Helv. Phys. Acta **25**, 185 (1952); E. H. Kronheimer, Phys. Rev. **90**, 1003 (1953).

η_a are of the form of harmonic oscillator eigenfunctions it is convenient to write the Slater parameters as triple integrals,

$$F^k(ac \mid db) = \int_0^\infty r_1^2\, dr_1 \int_0^\infty r_2^2\, dr_2 \int_0^\pi \sin\omega\, d\omega \times (k + \tfrac{1}{2}) V(r) P_k(\cos\omega) R_a^*(r_1) R_c(r_1) R_d(r_2) R_b^*(r_2), \quad (5)$$

and then to transform to a coordinate system in which r occurs as a variable of integration. The other coordinates are chosen to simplify the integral of the product of the orbitals and the angular factor. When $V(r)$ is a general functional form, the functions $f_k(r_1, r_2)$ are of little use as intermediates in a calculation. The coordinate transformation proposed by Talmi makes two of the three integrations trivial, and leaves only a final one-dimensional quadrature over r which can be performed by numerical methods if necessary. The function $V(r)$ affects the calculation only in the final quadrature. This has an obvious practical advantage in that $V(r)$ is incompletely determined by present empirical knowledge. General computer programs based on Talmi's analysis are committed to a definite form of $V(r)$ only through the single subroutine that carries out the quadrature over r.

In order to generalize Talmi's analysis to the case of the tensor interaction it is convenient to consider the reduction of angular integrals from a somewhat more general point of view than that usually used in deriving Eq. (5). The matrix element to be evaluated for a scalar potential is

$$(ab \mid V \mid cd) = \iint d\tau_1\, d\tau_2 R_a^*(r_1) \times R_c(r_1) V(r) R_d(r_2) R_b^*(r_2) Y^*(l_a, m_{la}; 1) \times Y(l_c, m_{lc}; 1) Y(l_d, m_{ld}; 2) Y^*(l_b, m_{lb}; 2). \quad (6)$$

The region of integration and volume element is

$$\iint d\tau_1\, d\tau_2 = \int_0^\infty r_1^2\, dr_1 \int_0^\infty r_2^2\, dr_2 \int_0^\pi \sin\omega\, d\omega \times \int_0^{2\pi} d\phi \int_0^\pi \sin\theta\, d\theta \int_0^{2\pi} d\psi, \quad (7)$$

where two geometrical interpretations of the angular variables are possible. The first is that (r_1, θ, ϕ) and (r_2, ω, ψ) are, respectively, the spherical polar coordinates of particles 1 and 2 in a fixed external frame of reference. The second is that the angles ω and ψ specify the orientation of vector $\mathbf{r}_2$ with respect to $\mathbf{r}_1$ in the sense that ω is the angle from $\mathbf{r}_1$ to $\mathbf{r}_2$ and ψ is the third Euler angle which together with θ and ϕ specifies the orientation of the $\mathbf{r}_1\mathbf{r}_2$ plane in a fixed external frame of reference. The volume element and limits of integration are the same in both cases.

When the second interpretation is used, it is clear that the matrix element is most easily evaluated by integrating first over the Euler angles (θ, ϕ, ψ) to obtain a function of (r_1, r_2, ω) defined by

$$\iiint Y_a^*(1) Y_c(1) Y_d(2) Y_b^*(2)\, d\Omega, \quad (8)$$

where $d\Omega$ denotes $d\phi \sin\theta\, d\theta\, d\psi$.

To carry out the integration of Eq. (8), it is convenient first to combine the spherical harmonics in each of the single-particle spaces. Thus,[7]

$$Y^*(l_a, m_{la}; 1) Y(l_c, m_{lc}; 1) = \sum_k (-1)^{m_{lc}} \times [(2l_a + 1)(2l_c + 1)(2k + 1)(4\pi)^{-1}]^{\frac{1}{2}} \times \begin{pmatrix} l_a & l_c & k \\ -m_{la} & m_{lc} & m_{la} - m_{lc} \end{pmatrix} \begin{pmatrix} l_a & l_c & k \\ 0 & 0 & 0 \end{pmatrix} \times Y(k, m_{lc} - m_{la}; 1) = \sum_k [(2k + 1)(4\pi)^{-1}]^{\frac{1}{2}} \times C^k(l_a, m_{la}; l_c, m_{lc}) Y^*(k, m_{la} - m_{lc}; 1). \quad (9)$$

The spherical harmonics are those defined by Condon and Shortley,[8] who have tabulated the Gaunt coefficients $C^k(l, m; l'm')$. The bracket symbols are Wigner 3-j symbols as usually defined.[7] In particle space 2,

$$Y(l_d, m_{ld}; 2) Y^*(l_b, m_{lb}; 2) = \sum_{k'} [(2k' + 1)(4\pi)^{-1}]^{\frac{1}{2}} \times C^{k'}(l_d, m_{ld}; l_b, m_{lb}) Y(k', m_{ld} - m_{lb}; 2). \quad (10)$$

The spherical harmonics in Eqs. (6) and (8) are defined in a fixed external reference frame P. They are related to spherical harmonics in a "body-fixed" reference frame Q, defined in a constant relationship to the $\mathbf{r}_1, \mathbf{r}_2$ plane, by the transformation[7]

$$\begin{aligned} Y^*(k, m; \theta_1, \phi_1) &= \sum_s Y^*(k, s; \theta_{1Q}, \phi_{1Q}) \times \mathfrak{D}_{sm}^{(k)*}(\phi, \theta, \psi), \\ Y(k', m'; \theta_2, \phi_2) &= \sum_{s'} Y(k', s'; \theta_{2Q}, \phi_{2Q}) \times \mathfrak{D}_{s'm'}^{(k')}(\phi, \theta, \psi), \end{aligned} \quad (11)$$

where (ϕ, θ, ψ) are the Euler angles of the coordinate transformation and the matrices $\mathfrak{D}$ are the irreducible representations of the rotation group. Since

[7] A. R. Edmonds, *Angular Momentum in Quantum Mechanics* (Princeton University Press, Princeton, New Jersey, 1957).

[8] E. U. Condon and G. H. Shortley, *The Theory of Atomic Spectra* (Cambridge University Press, New York, 1953).

the specific values θ_{1Q}, ϕ_{1Q}, etc., are functions of (r_1, r_2, ω) only, Eq. (11) provides an explicit expansion of the integrand of Eq. (8) in the variables of Eq. (7), with the variables (ϕ, θ, ψ) appearing only in the irreducible representation matrices $\mathfrak{D}$. Obviously this technique can be used to obtain an explicit expansion of any expression involving irreducible tensor functions or operators.

With these remarks, the integral of Eq. (8) becomes

$$\sum_k \sum_{k'} \frac{[(2k+1)(2k'+1)]^{\frac{1}{2}}}{4\pi} \times C^k(l_a, m_{la}; l_c, m_{lc}) C^{k'}(l_d, m_{ld}; l_b, m_{lb}) \times \sum_s \sum_{s'} Y^*(k, s; 1_Q) Y(k', s'; 2_Q) \times \iiint \mathfrak{D}^{(k)}_{sm}{}^*(\phi, \theta, \psi) \mathfrak{D}^{(k')}_{s'm'}(\phi, \theta, \psi)\, d\Omega, \quad (12)$$

where $m = m_{la} - m_{lc}$ and $m' = m_{ld} - m_{lb}$.

It is shown by Edmonds[7] that

$$\iiint \mathfrak{D}^{(k)}_{sm}{}^*(\phi, \theta, \psi) \mathfrak{D}^{(k')}_{s'm'}(\phi, \theta, \psi)\, d\Omega = \frac{8\pi^2}{2k+1} \delta_{kk'} \delta_{ss'} \delta_{mm'}. \quad (13)$$

Hence, Eq. (12) reduces to

$$\sum_k C^k(l_a m_{la}; l_c m_{lc}) C^k(l_d m_{ld}; l_b m_{lb}) \times [2\pi \sum_s Y^*(k, s; 1_Q) Y(k, s; 2_Q)]. \quad (14)$$

From the spherical harmonic addition theorem,[7] the factor in square brackets is just $(k + \frac{1}{2}) P_k(\cos \omega)$, where ω is the same as in Eq. (7). With F^k defined by Eq. (5), these results can be combined to give the well known formula[8]

$$(ab \mid V \mid cd) = \sum_k C^k(a; c) C^k(d; b) F^k(ac \mid db), \quad (15)$$

where

$$F^k(ac \mid db) = \int_0^\infty r_1^2\, dr_1 \int_0^\infty r_2^2\, dr_2 \int_0^\pi \sin \omega\, d\omega \times R^*_a(r_1) R_c(r_1) V(r) R_d(r_2) R^*_b(r_2) \times [2\pi \sum_s Y^*(k, s; 1_Q) Y(k, s; 2_Q)]. \quad (16)$$

III. MATRIX ELEMENTS OF A TENSOR OPERATOR

Consider a tensor operator of the usual form[5]

$$V(r) S_{12} = V(r) \sum_M (-1)^M \mathcal{S}(2, -M) \mathcal{R}(2, M) = V(r)[(\sigma_1 \cdot \mathbf{e})(\sigma_2 \cdot \mathbf{e}) - \tfrac{1}{3}(\sigma_1 \cdot \sigma_2)], \quad (17)$$

where

$$\mathcal{S}(2, -M) = \sqrt{5}\,(-1)^M \sum_{\nu+\nu'=-M} \begin{bmatrix} 1 & 2 & 1 \\ \nu & M & \nu' \end{bmatrix} \sigma_\nu(1) \sigma_{\nu'}(2), \quad (18)$$

and

$$\mathcal{R}(2, M) = \sqrt{5}\,(-1)^M \sum_{\nu+\nu'=M} \begin{bmatrix} 1 & 2 & 1 \\ \nu & -M & \nu' \end{bmatrix} e_\nu e_{\nu'}. \quad (19)$$

Here the indices ν refer to the irreducible tensor form of vectors[7] and $\mathbf{e}$ is a unit vector in the direction $\mathbf{r} = \mathbf{r}_2 - \mathbf{r}_1$.

Thus,

$$\mathcal{S}(2, 0) = (\tfrac{3}{2})^{\frac{1}{2}} (\sigma_z(1) \sigma_z(1) - \tfrac{1}{3}) \quad (20)$$

and

$$\mathcal{R}(2, 0) = (\tfrac{3}{2})^{\frac{1}{2}} (e_z^2 - \tfrac{1}{3}). \quad (21)$$

In the "body-fixed" frame of reference of Sec. II, if the polar axis is taken parallel to vector $\mathbf{r}$ and the reference plane $(y = 0)$ for polar coordinates is taken to be the $(\mathbf{r}_1, \mathbf{r}_2)$ plane, the tensor $\mathcal{R}$ reduces to

$$\mathcal{R}_Q(2, M) = (\tfrac{2}{3})^{\frac{1}{2}} \delta_{M0} \quad (22)$$

since in this frame of reference

$$\mathbf{e}_Q = (0, 0, 1). \quad (23)$$

It follows from the definition of $\mathcal{R}$ as an irreducible tensor that equations analogous to Eqs. (11) hold for it, so that in the external frame of reference

$$\mathcal{R}(2, M) = \sum_{M'} \mathcal{R}_Q(2, M') \mathfrak{D}^{(2)}_{M'M}(\Omega) = (\tfrac{2}{3})^{\frac{1}{2}} \mathfrak{D}^{(2)}_{0M}(\Omega). \quad (24)$$

The matrix of the spin operator $\mathcal{S}(2, -M)$ can easily be evaluated. When the quantum numbers $(m_{sa}, m_{sb}, m_{sc}, m_{sd})$ are specified, the matrix elements of $\mathcal{S}$ vanish unless

$$-M = m_{sa} + m_{sb} - m_{sc} - m_{sd}. \quad (25)$$

Similarly, the matrix elements of $\mathcal{R}(2, M)$ vanish unless

$$M = m_{la} + m_{lb} - m_{lc} - m_{ld}. \quad (26)$$

Thus, a unique value of M is determined for any nonvanishing matrix element of the tensor operator. Coefficients combining the constant factor of Eq. (24) with the appropriate matrix elements of $\mathcal{S}(2, -M)$,

$$\{ac \mid db\} = (-1)^M (\tfrac{2}{3})^{\frac{1}{2}} (ab \mid \mathcal{S}(2, -M) \mid cd), \quad (27)$$

are listed in Table I.

TABLE I. Coefficients $\{ac|db\} = (-1)^M(\frac{2}{3})^{\frac{1}{2}}(ab|S(2,-M)|cd)$.

m_{sa}, m_{sc}	m_{sd}, m_{sb} $\frac{1}{2}, \frac{1}{2}$	$\frac{1}{2}, -\frac{1}{2}$	$-\frac{1}{2}, \frac{1}{2}$	$-\frac{1}{2}, -\frac{1}{2}$
$\frac{1}{2}, \frac{1}{2}$	$\frac{2}{3}$	$-(\frac{2}{3})^{\frac{1}{2}}$	$(\frac{2}{3})^{\frac{1}{2}}$	$-\frac{2}{3}$
$\frac{1}{2}, -\frac{1}{2}$	$(\frac{2}{3})^{\frac{1}{2}}$	$-\frac{2}{3}$	$(\frac{8}{9})^{\frac{1}{2}}$	$-(\frac{2}{3})^{\frac{1}{2}}$
$-\frac{1}{2}, \frac{1}{2}$	$-(\frac{2}{3})^{\frac{1}{2}}$	$(\frac{8}{9})^{\frac{1}{2}}$	$-\frac{2}{3}$	$(\frac{2}{3})^{\frac{1}{2}}$
$-\frac{1}{2}, -\frac{1}{2}$	$-\frac{2}{3}$	$(\frac{2}{3})^{\frac{1}{2}}$	$-(\frac{2}{3})^{\frac{1}{2}}$	$\frac{2}{3}$

These coefficients multiply an integral of the same form as Eq. (6), except that the angular factor in the integrand is

$$Y^*(l_a, m_{la}; 1)Y(l_c, m_{lc}; 1)\mathfrak{D}^{(2)}_{0M}(\Omega) \times Y(l_d, m_{ld}; 2)Y^*(l_b, m_{lb}; 2). \quad (28)$$

Again from Eqs. (9), (10), and (11), the angular integral analogous to Eq. (12) is

$$\sum_k \sum_{k'} \frac{[(2k+1)(2k'+1)]^{\frac{1}{2}}}{4\pi} \times C^k(l_a, m_{la}; l_c, m_{lc})C^{k'}(l_d, m_{ld}; l_b, m_{lb}) \times \sum_s \sum_{s'} Y^*(k, s; 1_Q)Y(k', s'; 2_Q) \times \iiint \mathfrak{D}^{(k)*}_{sm}(\Omega)\mathfrak{D}^{(2)}_{0M}(\Omega)\mathfrak{D}^{(k')}_{s'm'}(\Omega)\, d\Omega. \quad (29)$$

Two formulas given by Edmonds[7] are used to evaluate the integral over Euler angles. They are

$$\mathfrak{D}^{(k)*}_{sm}(\Omega) = (-1)^{s-m}\mathfrak{D}^{(k)}_{-s,-m}(\Omega), \quad (30)$$

and

$$\iiint \mathfrak{D}^{(k)}_{-s,-m}(\Omega)\mathfrak{D}^{(2)}_{0M}(\Omega)\mathfrak{D}^{(k')}_{s'm'}(\Omega)\, d\Omega = 8\pi^2 \begin{pmatrix} k & 2 & k' \\ -s & 0 & s' \end{pmatrix} \begin{pmatrix} k & 2 & k' \\ -m & M & m' \end{pmatrix}, \quad (31)$$

where, as before, $m = m_{la} - m_{lc}$ and $m' = m_{ld} - m_{lb}$. The 3-j symbols vanish unless Eq. (26) is satisfied and unless $s' = s$. It is useful to introduce a third Gaunt coefficient

$$C^2(k, m; k', m') = (-1)^{-m}\,[(2k+1)(2k'+1)]^{\frac{1}{2}} \times \begin{pmatrix} k & 2 & k' \\ -m & M & m' \end{pmatrix} \begin{pmatrix} k & 2 & k' \\ 0 & 0 & 0 \end{pmatrix}. \quad (32)$$

If these equations are combined, Eq. (29) reduces to

$$\sum_k \sum_{k'} C^k(l_a, m_{la}; l_c, m_{lc}) \times C^{k'}(l_d, m_{ld}; l_b, m_{lb})C^2(k, m; k', m') \times \left[2\pi \sum_s \frac{(-1)^s \begin{pmatrix} k & 2 & k' \\ -s & 0 & s \end{pmatrix} Y^*(k, s; 1_Q)Y(k', s; 2_Q)}{\begin{pmatrix} k & 2 & k' \\ 0 & 0 & 0 \end{pmatrix}} \right]. \quad (33)$$

The factor in square brackets is an explicit function of (r_1, r_2, ω).

In terms of the spin matrix elements $\{ac \mid db\}$ defined above the general matrix element of the tensor operator is

$$(ab \mid VS \mid cd) = \{ac \mid db\} \sum_k \sum_{k'} C^k(a; c)C^{k'}(d; b) \times C^2(k, m_{la}-m_{lc}; k', m_{ld}-m_{lb})I^{kk'}(ac \mid db) \quad (34)$$

where

$$I^{kk'}(ac \mid db) = \int_0^\infty r_1^2\, dr_1 \int_0^\infty r_2^2\, dr_2 \times \int_0^\pi \sin\omega\, d\omega\, R^*_a(r_1)R_c(r_1)V(r)R_d(r_2)R^*_b(r_2) \times \left[2\pi \sum_s (-1)^s \frac{\begin{pmatrix} k & 2 & k' \\ -s & 0 & s \end{pmatrix}}{\begin{pmatrix} k & 2 & k' \\ 0 & 0 & 0 \end{pmatrix}} \times Y^*(k, s; 1_Q)Y(k', s; 2_Q) \right]. \quad (35)$$

These formulas constitute a direct generalization of the Slater formulas, Eqs. (15) and (16). The quantum numbers k and k' must satisfy the usual vector-coupling conditions for nonvanishing Gaunt coefficients, i.e.,

$$|l_a - l_c| \le k \le l_a + l_c;$$
$$|l_d - l_b| \le k' \le l_d + l_b;$$
$$|k - k'| \le 2 \le k + k';$$
$$k + l_a + l_c,$$
$$k' + l_d + l_b, \quad k + k' + 2 \quad \text{even}. \quad (36)$$

The quantum numbers m_s and m_l must satisfy the conservation rule [from Eqs. (25) and (26)]

$$m_{la} + m_{sa} + m_{lb} + m_{sb} = m_{lc} + m_{sc} + m_{ld} + m_{sd}. \quad (37)$$

Except for constant factors, the quantities $I^{kk'}$

are reduced matrix elements, in the sense of Racah,[9] of the tensor operator, in a vector coupling scheme in which the angular momenta in each variable space are coupled to definite orbital angular momenta, k and k'. This is different from the usual scheme,[5] in which l_a and l_b are coupled to total L and l_c and l_d coupled to total L'. The (L, L') reduced matrix elements are not, in general, all linearly independent.

IV. COMPUTATION OF REDUCED MATRIX ELEMENTS

The evaluation of the reduced matrix elements F^k or $I^{kk'}$ is greatly simplified if the radial factors of the basis orbitals are of the form

$$R_a(r) = N_a r^{2n_a+l_a} \exp(-\gamma_a r^2). \tag{38}$$

Complete sets of functions can be made up from this class of basis orbitals in many different ways. Orbitals of physical interest (such as nuclear Hartree–Fock functions) can be expressed as linear combinations of basis orbitals of this class. Hence, all matrix elements of integrable operators in a nuclear Hamiltonian can be expressed as linear combinations of matrix elements with basis orbitals of this kind.

The exponential factor in the integrand of Eq. (16) or Eq. (35), for four different orbitals of the form of Eq. (38), is

$$\exp(-\gamma_1 r_1^2 - \gamma_2 r_2^2), \tag{39}$$

where

$$\gamma_1 = \gamma_a + \gamma_c, \qquad \gamma_2 = \gamma_d + \gamma_b. \tag{40}$$

Following Talmi,[6] it is convenient to use coordinates (r, R, α), instead of (r_1, r_2, ω), where r is the magnitude of the relative vector

$$\mathbf{r} = \mathbf{r}_2 - \mathbf{r}_1, \tag{41}$$

R is the magnitude of the "center-of-mass" vector

$$\mathbf{R} = (\gamma_1\mathbf{r}_1 + \gamma_2\mathbf{r}_2)/(\gamma_1 + \gamma_2), \tag{42}$$

and α is the angle from $\mathbf{R}$ to $\mathbf{r}$. Thus,

$$\mathbf{R}\cdot\mathbf{r} = rR\cos\alpha, \tag{43}$$

$$|\mathbf{R}\times\mathbf{r}| = rR\sin\alpha. \tag{44}$$

As a result of this coordinate change, the exponential function of Eq. (39) factorizes into

$$\exp(-\gamma r^2 - \lambda R^2), \tag{45}$$

where

$$\gamma = \gamma_1\gamma_2/(\gamma_1 + \gamma_2), \tag{46}$$

[9] G. Racah, Phys. Rev. 62, 438 (1942).

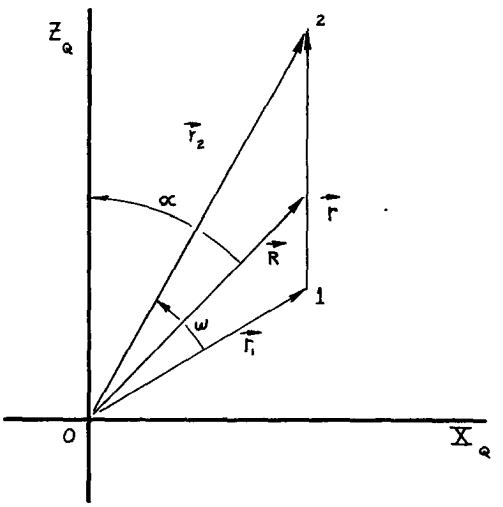


FIG. 1. Vectors in plane $y_Q = 0$.

and

$$\lambda = \gamma_1 + \gamma_2. \tag{47}$$

The remaining factors in the integrand for the reduced matrix elements are, except for the arbitrary function $V(r)$, polynomials in r, R, and $\cos\alpha$. This remark is trivial in the case of even powers of r_1 and r_2, since

$$\begin{aligned} r_1^2 &= R^2 - 2\gamma_2 rR\cos\alpha/(\gamma_1+\gamma_2) + \gamma_2^2 r^2/(\gamma_1+\gamma_2)^2, \\ r_2^2 &= R^2 + 2\gamma_1 rR\cos\alpha/(\gamma_1+\gamma_2) + \gamma_1^2 r^2/(\gamma_1+\gamma_2)^2. \end{aligned} \tag{48}$$

To analyze the spherical harmonic factors, it is convenient to use Cartesian coordinates

$$\begin{aligned} \mathbf{r}_1 &= \mathbf{R} - \gamma_2\mathbf{r}/(\gamma_1+\gamma_2), \\ \mathbf{r}_2 &= \mathbf{R} + \gamma_1\mathbf{r}/(\gamma_1+\gamma_2). \end{aligned} \tag{49}$$

In the special coordinate frame Q, $\mathbf{r}$ is parallel to the polar axis and $\mathbf{R}$ lies in the reference plane $y_Q = 0$, as indicated in Fig. 1. Hence,

$$\begin{aligned} \mathbf{r}_Q &= (0, 0, r), \\ \mathbf{R}_Q &= (R\sin\alpha, 0, R\cos\alpha), \end{aligned} \tag{50}$$

and

$$\begin{aligned} \mathbf{r}_{1Q} &= (R\sin\alpha, 0, R\cos\alpha - \gamma_2 r/(\gamma_1+\gamma_2)), \\ \mathbf{r}_{2Q} &= (R\sin\alpha, 0, R\cos\alpha + \gamma_1 r/(\gamma_1+\gamma_2)). \end{aligned} \tag{51}$$

Since the functions $r^k Y(k, s; \theta, \phi)$ are homogeneous polynomials of degree k in the Cartesian coordinates, it follows that

$$r_{1Q}^k Y^*(k, s; \theta_{1Q}, \phi_{1Q}) r_{2Q}^{k'} Y(k', s; \theta_{2Q}, \phi_{2Q}) \tag{52}$$

is a homogeneous polynomial of degree $k + k'$ in (r, R). It will be shown below that only even powers of $\sin \alpha$ occur. Since by Eqs. (36) $l_a + l_c - k$ and $l_d + l_b - k'$ are both even, it follows that the integrand is a polynomial in $\cos \alpha$.

The region of integration and volume element associated with variables (r, R, α) is easily found to be

$$\int_0^\infty r^2\,dr \int_0^\infty R^2\,dR \int_0^\pi \sin\alpha\,d\alpha, \tag{53}$$

since the Jacobian of Eqs. (41) and (42) is unity.

Since the integrand is of the form

$$N_a N_c N_d N_b V(r) \exp(-\gamma r^2 - \lambda R^2) \text{ Polynomial } (r, R, \cos\alpha), \tag{54}$$

the integrations over α and R are trivial. The final quadrature in r will, in general, be carried out numerically. The chief problem in evaluating the reduced matrix elements is the determination of the coefficients of the polynomial in Eq. (54).

The remainder of this section describes a method of evaluating these coefficients designed for an automatic computer program. For this purpose it is important to specify the order in which operations are to be carried out as well as to state the formulas used.

The normalized spherical harmonics as defined by Condon and Shortley[8] lead to the homogeneous polynomials

$$r^l Y(l, m; \theta, \phi) = \left[\frac{2l+1}{4\pi}\frac{(l+m)!}{(l-m)!}\right]^{\frac{1}{2}} \frac{(-1)^m}{2^l l!} \times (x + iy)^m r^{l-m} \Pi_{lm}(\cos\theta), \tag{55}$$

where

$$\Pi_{lm}(\cos\theta) = (\cos^2\theta - 1)^{-m}\left(\frac{d}{d\cos\theta}\right)^{l-m}(\cos^2\theta - 1)^l \tag{56}$$

$$= \sum_{b=0,1}^{l-m} p_{lm}(b)(\cos\theta)^b(1 - \cos^2\theta)^{\frac{1}{2}(l-m-b)}. \tag{57}$$

The functions Π_{lm} are evaluated by a recurrence formula obtained from their definition, Eq. (56),

$$\Pi_{ll} = 1, \tag{58}$$

$$\Pi_{lm} = \left\{2(m+1)\cos\theta - (1 - \cos^2\theta)\frac{d}{d\cos\theta}\right\}\Pi_{l,m+1}. \tag{59}$$

In terms of the coefficients $p_{lm}(b)$ this is

$$p_{ll}(b) = \delta_{b0}, \tag{60}$$

$$p_{lm}(b) = (l + m + 2 - b)p_{l,m+1}(b - 1) - (b + 1)p_{l,m+1}(b + 1). \tag{61}$$

The computer program for Eq. (61) carries out the following sequence of operations on coefficients stored in a linear array with index b. This operation replaces the array $p_{l,m+1}$ by the array p_{lm}. An arrow will be used to denote "replaces".

(1) $(l + m + 1 - b)p(b) \rightarrow p(b + 1)$
(2) $p(b - 1) - bp(b) \rightarrow p(b - 1)$
(3) zero $\rightarrow p(b)$
(4) $b + 2 \rightarrow b$
(5) Repeat (1–4) until $b > l - m$. (62)

In the Cartesian coordinates of Eq. (51), x_{1Q} and x_{2Q} are both equal to $X_Q = R \sin\alpha$, and $y_{1Q} = y_{2Q} = 0$. Since $r_1^2(1 - \cos^2\theta_{1Q}) = r_2^2(1 - \cos^2\theta_{2Q}) = X_Q^2$, the bracketed expressions in Eqs. (16) and (35), when multiplied by $r_1^k r_2^{k'}$, are both of the form

$$\sum_{s=0}^{\text{Min}(k,k')} C(k, s)X_Q^{2s} \sum_a \sum_b p_{ks}(a)p_{k's}(b) \times z_{1Q}^a z_{2Q}^b X_Q^{k+k'-2s-a-b}. \tag{63}$$

Since a and b are either both even or both odd, and since $k + k'$ is even, only even powers of X_Q occur in this formula. Thus, as remarked above, the integrand in F^k or $I^{kk'}$ is a polynomial in $\cos\alpha$. The coefficients $C(k, s)$ combine the normalization constants of the spherical harmonics with the coefficients in the bracketed expressions of Eqs. (16) and (35). Since the terms in s and $-s$ are identical, a factor of two is included in $C(k, s)$ for $s > 0$, and only nonnegative values of s are included in the summation in Eq. (63).

The four different recurrence formulas for $C(ks)$ needed for the scalar and tensor operators are:

Scalar $k' = k$

$$C(k, k) = (2k + 1)!!/(2k)!!$$

$$C(k, s) = \frac{C(k, s+1)}{(k - s)(k + s + 1)}(1 - \tfrac{1}{2}\delta_{s0}); \tag{64}$$

Tensor $k' = k - 2$

$$C(k, k-2) = \frac{2[(2k+1)(2k-3)]^{\frac{1}{2}}(2k-1)!!}{k(2k-1)(2k)!!}$$

$$C(k, s) = \frac{C(k, s+1)}{(k - s - 2)(k + s + 1)}(1 - \tfrac{1}{2}\delta_{s0}); \tag{65}$$

Tensor $k' = k$

$$C(k, k) = \frac{-2(2k-1)(2k+1)!!}{(2k+2)!!}$$

$$C(k, s) = \frac{[k(k+1) - 3s^2]C(k, s+1)}{[k(k+1) - 3(s+1)^2](k-s)(k+s+1)} \times (1 - \tfrac{1}{2}\delta_{s0}); \quad (66)$$

Tensor $k' = k + 2$

$$C(k, k) = \frac{2[(2k+1)(2k+5)]^{\frac{1}{2}}(2k+3)!!}{(k+2)(2k+3)(2k+4)!!}$$

$$C(k, s) = \frac{C(k, s+1)}{(k-s)(k+s+3)}(1 - \tfrac{1}{2}\delta_{s0}). \quad (67)$$

Eq. (66) must be modified by treating the factor $[k(k+1) - 3s^2]$ separately when $C(k, s)$ vanishes, to avoid division by zero.

Eq. (63) is homogeneous of degree $(k + k')$ in the three coordinates (z_{1Q}, z_{2Q}, X_Q). For this reason, it is convenient to represent it as a rectangular array of coefficients of $z_{1Q}^a z_{2Q}^b$,

$$D'(a, b) = \sum_s C(k, s)p_{ks}(a)p_{k's}(b). \quad (68)$$

These coefficients are obtained by combining the appropriate one of Eqs. (64)–(67) with the recurrence procedure of Eq. (62), recurring downward on s from its maximum value to zero.

This rectangular array of coefficients $D'(a, b)$ is stored in a rectangular block of locations of dimension $l_a + l_c + 1$ by $l_d + l_b + 1$, to accommodate the additional powers of r_1^2 and r_2^2, respectively, that are inherent in the definition, Eq. (38), of basis orbitals corresponding to given spherical harmonics. Since $r_1^k r_2^{k'}$ is already included in Eq. (63), these additional powers are $r_1^{l_a+l_c-k}$ and $r_2^{l_d+l_b-k'}$, even in both cases. In terms of the Cartesian coordinates, the additional factor to Eq. (63) is

$$(z_{1Q}^2 + X_Q^2)^{\frac{1}{2}(l_a+l_c-k)}(z_{2Q}^2 + X_Q^2)^{\frac{1}{2}(l_d+l_b-k')}. \quad (69)$$

This factor is obviously a homogeneous polynomial in (z_{1Q}, z_{2Q}, X_Q) and preserves the structure of Eq. (63). The program to multiply Eqs. (63) and (69) is very simple. Each factor $(z_{1Q}^2 + X_Q^2)$ is multiplied in by the procedure

(1) $D(a, b) + D(a+2, b) \rightarrow D(a+2, b)$,
(2) $a - 2 \rightarrow a$,
(3) Repeat (1–2) until $a < 0$. (70)

Each factor $(z_{2Q}^2 + X_Q^2)$ is entered by the procedure

(1) $D(a, b) + D(a, b+2) \rightarrow D(a, b+2)$,
(2) $b - 2 \rightarrow b$,
(3) Repeat (1–2) until $b < 0$. (71)

The count in both cases is downward from the current maximum value of a or b, respectively. The first process is repeated $\frac{1}{2}(l_a + l_c - k)$ times and the second is repeated $\frac{1}{2}(l_d + l_b - k')$ times, to produce the array of coefficients $D(a, b)$, where $r_1^{l_a+l_c} r_2^{l_d+l_b}$ times the bracketed expression in Eqs. (16) or (35) is equal to

$$\sum_a \sum_b D(a, b) z_{1Q}^a z_{2Q}^b X_Q^{l_a+l_c+l_d+l_b-a-b}. \quad (72)$$

These arrays $D(a, b)$ depend only on the quantum numbers $(l_a, l_c, l_d, l_b, k, k')$, with different arrays for the scalar and tensor operators. They can be computed before computing the individual integrals in the block of matrix elements with orbitals having these angular quantum numbers. Thus, the computation of these arrays occupies a negligible fraction of the total computing time when there is a reasonably large number of independent basis orbitals with the same angular momentum quantum numbers.

Given these arrays, for each individual matrix element $F^k(ac \mid db)$ or $I^{kk'}(ac \mid db)$ it is necessary to convert Eq. (72) to the form of a polynomial in $(r, R, \cos\alpha)$, to multiply in the additional factors $r_1^{2n_a+2n_c}$, $r_2^{2n_d+2n_b}$, and then to carry out the three-dimensional quadrature. The coefficients of the polynomial in $(r, R, \cos\alpha)$ obviously depend on the exponents γ_1 and γ_2 which occur in Eqs. (51).

If $l_1 = l_a + l_c$, $l_2 = l_d + l_b$, $\lambda_1 = -\gamma_2/(\gamma_1 + \gamma_2)$, and $\lambda_2 = \gamma_1/(\gamma_1 + \gamma_2)$, Eq. (72) becomes

$$\sum_a \sum_b D(a, b)(R\cos\alpha + \lambda_1 r)^a \times (R\cos\alpha + \lambda_2 r)^b (R\sin\alpha)^{l_1+l_2-a-b} = \sum_u \sum_v G(u, v) R^{l_1+l_2-2u+v} r^{2u-v} \times (\sin\alpha)^{l_1+l_2-2u}(\cos\alpha)^v. \quad (73)$$

This transformation is carried out by collecting the terms in Eq. (73) with the same power of $\sin\alpha$, for which $a + b = 2u$. The corresponding coefficients form an antidiagonal linear array in the rectangular array $D(a, b)$. For a given value of u, the transformation is defined by

$$\sum_{a+b=2u} D(a, b)(R\cos\alpha + \lambda_1 r)^a (R\cos\alpha + \lambda_2 r)^b = \sum_u G(u, v) R^v r^{2u-v} (\cos\alpha)^v. \quad (74)$$

Given u, the coefficients $G(u, v)$ form a linear array in index v, which satisfies $0 \leq v \leq 2u$. Thus, the full array $G(u, v)$ is triangular and can itself be treated as a linear array, in increasing order of the index u with range $0 \leq 2u \leq l_1 + l_2$. There are $2u + 1$ elements $G(u, v)$ for each value of u.

The set of coefficients $G(u, v)$ with given index u is obtained by the following procedure, which is repeated for values of u in increasing order:

(1) Zero $\rightarrow G(u, v)$ for $0 \leq v \leq 2u$; call this array $G(v)$; zero $\rightarrow G'(v)$, same number of locations.
(2) Max $(2u - l_2, 0) \rightarrow a$; $2u - a \rightarrow b$.
(3) $D(a, b) \rightarrow G'(0)$.
(4) Do (5–9) a times.
(5) Zero $\rightarrow v$.
(6) $G'(v) + G'(v + 2) \rightarrow G'(v + 2)$.
(7) $\lambda_1 G'(v) \rightarrow G'(v)$.
(8) $v + 1 \rightarrow v$.
(9) Repeat (6–8) until $v > 2u$, then return to 5.
(10) Do (11–15) b times.
(11) Zero $\rightarrow v$.
(12) $G'(v) + G'(v + 2) \rightarrow G'(v + 2)$.
(13) $\lambda_2 G'(v) \rightarrow G'(v)$.
(14) $v + 1 \rightarrow v$.
(15) Repeat (12–14) until $v > 2u$, then return to 11.
(16) Array $[G'(v) + G(v)] \rightarrow$ array $[G(v)]$.
(17) $a + 1 \rightarrow a$; $2u - a \rightarrow b$.
(18) Repeat (3–17) until $a > l_1$, or $b < 0$. (75)

The triangular array $G(u, v)$ is not yet complete, since an additional factor $r_1^{2n_1} r_2^{2n_2}$ occurs in each integrand. Here $n_1 = n_a + n_c$ and $n_2 = n_d + n_b$. This factor is of the same form as Eq. (73), since

$$r_1^2 = R^2 \sin^2\alpha + \lambda_1^2 r^2 + 2\lambda_1 Rr \cos\alpha + R^2 \cos^2\alpha,$$
$$r_2^2 = R^2 \sin^2\alpha + \lambda_2^2 r^2 + 2\lambda_2 Rr \cos\alpha + R^2 \cos^2\alpha. \quad (76)$$

For example, the triangular array that represents r_1^2, written as a linear array, is

$$(1, \lambda_1^2, 2\lambda_1, 1). \quad (77)$$

The product of two such arrays is an array of the same kind. The factor $r_1^{2n_1} r_2^{2n_2}$ can be put in by successive multiplications of arrays like Eq. (77) into the array $G(u, v)$. This process can be carried out as follows, counting backward through all the elements of $G(u, v)$:

(1) $G(u, v) + G(u + 1, v + 2) \rightarrow G(u + 1, v + 2)$.
(2) $2\lambda_1 G(u, v) + G(u + 1, v + 1) \rightarrow G(u + 1, v + 1)$.
(3) $\lambda_1^2 G(u, v) + G(u + 1, v) \rightarrow G(u + 1, v)$. (78)

Obviously $u_{\max}$ is increased by one for each factor r_1^2 or r_2^2. The final array, including all factors, will be denoted by $F(u, v)$, with $2u_{\max} = 2U = l_1 + l_2 + 2n_1 + 2n_2$. This is the polynomial indicated in Eq. (54), expressed in the form of Eq. (73).

To carry out the integration over α and R it is convenient to use the following formulas, easily obtained by integration by parts:

$$\frac{1}{2}\int_0^\pi (\sin\alpha)^{2p+1}(\cos\alpha)^{2q}\, d\alpha = \frac{(2p)!!\,(2q-1)!!}{(2p+2q+1)!!}, \quad (79)$$

$$2\int_0^\infty R^{2p+2q+2} e^{-\lambda R^2}\, dR = \frac{(2p+2q+1)!!}{(2\lambda)^{p+q+1}}\left(\frac{\pi}{\lambda}\right)^{\frac{1}{2}}. \quad (80)$$

The final integration over r takes the form

$$I = C_I \int_0^\infty e^{-\gamma r^2} V(r) \sum_{t=0}^{U} G_{2t} r^{2t+2}\, dr, \quad (81)$$

where

$$G_{2t} = \sum_{p+q=U-t} F(U - p, q)\, \frac{(2p)!!\,(2q-1)!!}{(2\lambda)^{p+q+1}}, \quad (82)$$

and

$$C_I = (\pi/\lambda)^{\frac{1}{2}} N_a N_c N_d N_b, \quad (83)$$

with γ and λ given by Eqs. (46) and (47), respectively.

In general, there will be several different independent potential functions $V(r)$ for both the scalar and tensor parts of a nuclear two-body interaction. Different methods of evaluating Eq. (81) are appropriate for different potential functions. If $V(r)$ has a well defined range, then $V(r)$ exp (σr^2) will be more easily approximated by a polynomial than is $V(r)$ itself, for an appropriate choice of the range parameter σ. Depending on whether this polynomial is more nearly even or odd, there are two methods of numerical quadrature that may be generally useful.

Eq. (81) can be written in the form

$$I = C_I \int_0^\infty e^{-(\gamma+\sigma)r^2} f(r)\, dr, \quad (84)$$

where

$$f(r) = V(r) e^{\sigma r^2} \sum_{t=0}^{U} G_{2t} r^{2t+2}. \quad (85)$$

If x_i and α_i are, respectively, the roots and weight factors for Gauss–Laguerre quadrature, as tabulated and defined by Salzer and Zucker,[10] then if

$$h_i = \alpha_i; \qquad r_i^2 = x_i/(\gamma + \sigma), \quad (86)$$

[10] H. E. Salzer and R. Zucker, Bull. Am. Math. Soc. 55, 1004 (1949).

a simple coordinate transformation leads to

$$I \cong \frac{C_I}{2(\gamma+\sigma)} \sum h_i \left[\frac{1}{r_i} f(r_i)\right]. \quad (87)$$

This formula is exact if $f(r)$ is an odd polynomial of order less than four times that of the Laguerre polynomial whose roots are x_i.

If x_i and α_i are, respectively, the positive roots and weight factors for even Gauss–Hermite quadrature, as tabulated and defined by Salzer, Zucker, and Capuano,[11] then if

$$h_i = 2\alpha_i x_i^2; \qquad r_i^2 = x_i^2/(\gamma+\sigma), \quad (88)$$

the integral is approximated by

$$I = \frac{C_I}{2(\gamma+\sigma)^{\frac{1}{2}}} \sum h_i \left[\frac{1}{r_i^2} f(r_i)\right]. \quad (89)$$

This formula is exact if $f(r)$ is an even polynomial of order less than twice that of the Hermite polynomial whose roots are x_i. Since only half of the roots are used, both formulas are of comparable accuracy

V. ONE-PARTICLE MATRIX ELEMENTS

With basis orbitals whose radial factors are given by Eq. (38), the orthogonality and kinetic-energy matrix elements can be expressed in terms of the auxiliary function

$$U(x, y) = \int_0^\infty r^{2x} e^{-yr^2}\, dr = \frac{(2x-1)!!}{(2y)^x}\left(\frac{\pi}{4y}\right)^{\frac{1}{2}}. \quad (90)$$

This is most easily computed by the recurrence formula

$$U(0, y) = (\pi/4y)^{\frac{1}{2}},$$
$$U(x+1, y) = \frac{2x+1}{2y} U(x, y). \quad (91)$$

The orthogonality or "overlap" matrix elements are

$$(a \mid c) = N_a N_c \int_0^\infty r^{2n_a+2n_c+2l+2} e^{-(\gamma_a+\gamma_c)r^2}\, dr$$
$$= N_a N_c U(n_a+n_c+l+1, \gamma_a+\gamma_c), \quad (92)$$

if $l_a = l_c = l$. The kinetic-energy matrix elements are, again with $l_a = l_c = l$,

$$\left(a \left| -\frac{\hbar^2}{2M}\nabla^2 \right| c\right) = \frac{\hbar^2}{2M} N_a N_c$$
$$\times \int_0^\infty \left[\left\{r\frac{d}{dr}(r^{2n_a+l}e^{-\gamma_a r^2})\right\}\right.$$
$$\times \left\{r\frac{d}{dr}(r^{2n_c+l}e^{-\gamma_c r^2})\right\}$$
$$\left. + l(l+1) r^{2n_a+2n_c+2l} e^{-(\gamma_a+\gamma_c)r^2}\right] dr \quad (93)$$
$$= \frac{\hbar^2}{2M} N_a N_c [4n_a n_c U(n_a+n_c+l, \gamma_a+\gamma_c)$$
$$- 4(n_a\gamma_c + n_c\gamma_a) U(n_a+n_c+l+1, \gamma_a+\gamma_c)$$
$$+ 4\gamma_a\gamma_c U(n_a+n_c+l+2, \gamma_a+\gamma_c)]. \quad (94)$$

The normalization constants are

$$N_a = [U(2n_a+l+1, 2\gamma_a)]^{-\frac{1}{2}}. \quad (95)$$

Another type of matrix element occurs in correcting for the center-of-mass motion of a finite nucleus. In terms of the Gaunt coefficients,

$$\left(ab \left| \frac{\hbar^2}{M}\nabla_1\cdot\nabla_2 \right| cd\right) = C^1(l_a, m_{la}; l_c, m_{lc})$$
$$\times C^1(l_d, m_{ld}; l_b, m_{lb}) P^1(ac \mid db), \quad (96)$$

where the reduced matrix element is

$$P^1(ac \mid db) = \frac{\hbar^2}{M} P(ac) P(db), \quad (97)$$

with

$$P(ac) = \int_0^\infty r^2 R_a \left(\frac{d}{dr} - \frac{l_c}{r}\right) R_c\, dr, \quad (98)$$

if $l_a \geq l_c$ and $l_d \geq l_b$. An interchange of one pair of quantum numbers l changes the sign of the matrix element. The integral of Eq. (98) is

$$P(ac) = N_a N_c \int_0^\infty r^{2n_a+2n_c+l_1+l_2+2} e^{-(\gamma_a+\gamma_c)r^2}$$
$$\times \left\{\frac{2n_c}{r} - 2\gamma_c r\right\} dr = N_a N_c$$
$$\times \{2n_c U(n_a+n_c+l_a, \gamma_a+\gamma_c)$$
$$- 2\gamma_c U(n_a+n_c+l_a+1, \gamma_a+\gamma_c)\}, \quad (99)$$

where $l_a = l_c + 1$ only.

If M is the neutron mass, the constant in the kinetic-energy integrals is

$$\hbar^2/2M = 20.720 \text{ meV} \times 10^{-26} \text{ cm}^2. \quad (100)$$

ACKNOWLEDGMENTS

The author is indebted to Dr. C. E. Porter and to Dr. J. R. Stehn for making it possible to carry out this work at Brookhaven, and for their hospitality. He is also indebted to Dr. N. C. Francis for many helpful discussions and for encouraging this work, and to Professor E. L. Lomon for advice on the present status of the nuclear two-body interaction.

[11] Salzer, Zucker, and Capuano, J. Res. Natl. Bur. Std. **48**, 111 (1952).

Reduced Width Amplitude Distributions and Random Sign Rules in R-Matrix Theory*

T. J. KRIEGER AND C. E. PORTER
Brookhaven National Laboratory, Upton, Long Island, New York
(Received 5 April 1963; revised 12 June 1963)

A multivariate reduced width amplitude distribution is derived from quite general assumptions of level independence and of functional form invariance of the distribution. The multivariate distribution reduces to the well-known Gaussian singlet distribution (one degree of freedom) and moreover allows for the possibility of channel–channel reduced width amplitude correlations. Such correlations, which cannot be represented by a singlet distribution, may be especially relevant to partial fission reduced width amplitudes. In particular, it is noted that in statistical theory complete correlation or anticorrelation between two random variables implies a linear relation between them. The multivariate distribution is then used as a basis for a more precise statement of the "random sign" rules of R-matrix theory than hitherto given. It is pointed out that these rules imply a linear averaging over the multivariate reduced width amplitude distribution (in addition to the usual linear energy average) and are not merely consequences of the equiprobability of positive and negative signs for the reduced width amplitudes. The influence of the energy eigenvalue distribution and of variable reduced width amplitude statistics is discussed. It is emphasized that the multivariate reduced width amplitude distribution is relevant to any reaction theory in which level widths are defined even though it is phrased here in the language of R-matrix theory.

I. INTRODUCTION

THE statistical distribution of single-channel reduced width amplitudes ($\gamma_{\lambda c}$) in the R-matrix theory of nuclear reactions is the well-known Gaussian distribution.[1] The original derivation of this distribution was based on physical arguments leading to the use of the central limit theorem of statistical theory. Since that time, it has been pointed out that the distribution can be thought of as a necessary consequence of a random matrix hypothesis.[2] The random matrix hypothesis also has the unifying feature that the eigenvalue statistics are derived from it as well.

The purpose of the present work is to present a more general approach to the problem of the reduced width amplitude distribution based on the principles of level independence and of invariance with respect to orthogonal transformations of the basic state vectors of the compound nucleus. The method enables us to derive without further assumptions, not only the single-channel Gaussian distribution mentioned above, but also the form of the *multivariate* reduced width amplitude distribution. The latter can then be used to establish in a direct manner the rules governing the application of the so-called "random sign approximation" in nuclear reaction theory.

* Research supported by the United States Atomic Energy Commission.

[1] R. G. Thomas and C. E. Porter, Phys. Rev. **104**, 483 (1956).

[2] S. Blumberg and C. E. Porter, Phys. Rev. **110**, 786 (1958). See also C. E. Porter and N. Rosenzweig, Ann. Acad. Sci. Fennicae, Ser. A VI, **44** (1960).

II. MULTIVARIATE DISTRIBUTION

Consider a set of levels $\lambda = 1, 2 \cdots$ and channels $c, c', \cdots$. The levels are assumed to be contained within an interval K, which is small compared to an energy region over which the strength function of any channel varies significantly. We imagine the channel labels c to specify the usual coupling scheme $\{\alpha slJM\}$ where α denotes the target and projectile internal state, s the channel spin, l the orbital angular momentum, J the total angular momentum, and M the z projection of J.[3] Then we can write the joint distribution function of the $\gamma_{\lambda c}$ as

$$P(\{\gamma_{\lambda c}\}) \equiv P(\gamma_{1c}, \gamma_{1c'}, \cdots, \gamma_{2c}, \gamma_{2c'}, \cdots). \quad (1)$$

We now make the following two fundamental assumptions concerning the form of P:

(i) P is *independent* with respect to levels, i.e., P is factorizable into (individually normalized) functions $f_\lambda(\gamma_{\lambda c}, \gamma_{\lambda c'}, \cdots)$ corresponding to the various levels λ of the compound nucleus. (For the sake of simplicity, in the remainder of this section we shall consider levels λ corresponding to the same symmetry labels $JM\Pi$.)

(ii) P is *form invariant* with respect to a proper (i.e., determinant $= +1$) orthogonal transformation among the state functions X_λ of the same symmetry. Since $\gamma_{\lambda c} \equiv (\hbar^2/2M_c a_c)^{\frac{1}{2}} \int \varphi_c^* X_\lambda \, dS$, where M_c is the reduced mass, a_c is the channel radius, φ_c is the channel function, and S is the channel surface, an orthogonal transformation of the X_λ is equivalent

[3] A. M. Lane and R. G. Thomas, Rev. Mod. Phys. **30**, 257 (1958).

to the same orthogonal transformation of the $\gamma_{\lambda c}$ with the same channel index c.

The statement of independence [assumption (i)] may be regarded as a consequence of the random matrix hypothesis mentioned earlier in the limit of large matrix dimension for levels of the same symmetry (spin and parity). The invariance assumption is a general feature of the derivation of the Gaussian matrix ensemble.[2]

From assumption (i) we may write for channels $c, c', \cdots$ of the same symmetry

$$P_{JM\Pi}(\{\gamma_{\lambda c}\}) = f_1(\gamma_{1c}, \gamma_{1c'}, \cdots) f_2(\gamma_{2c}, \gamma_{2c'}, \cdots) \cdots . \tag{2}$$

Channels of different symmetry are not coupled. We wish to develop first the usual single-channel distribution. To do this we form

$$P_c(\gamma_{1c}, \gamma_{2c}, \cdots) \equiv \iint \cdots \int \cdots \int P_{JM\Pi}(\{\gamma_{\lambda c}\})\, d\gamma_{1c'}\, d\gamma_{2c'} \cdots d\gamma_{1c''}\, d\gamma_{2c''} \cdots , \tag{3}$$

which gives on substitution from (2)

$$P_c(\gamma_{1c}, \gamma_{2c}, \cdots) = f_1(\gamma_{1c}) f_2(\gamma_{2c}) \cdots , \tag{4}$$

where the f's are obvious integrals over the remaining reduced width amplitudes.

In order to deduce the form of P_c, we introduce assumption (ii) above. It is convenient to consider only a pair of levels at a time, say $\lambda = 1, 2$; the relevant transformation is

$$\begin{aligned} \gamma_{1c} &= \gamma'_{1c} \cos\theta + \gamma'_{2c} \sin\theta, \\ \gamma_{2c} &= -\gamma'_{1c} \sin\theta + \gamma'_{2c} \cos\theta. \end{aligned} \tag{5}$$

We insist that (invariance)

$$d(\ln P_c)/d\theta = 0, \tag{6}$$

or equivalently that

$$\frac{f_1'(\gamma_{1c})}{f_1(\gamma_{1c})} \gamma_{2c} - \frac{f_2'(\gamma_{2c})}{f_2(\gamma_{2c})} \gamma_{1c} = 0, \tag{7}$$

from which it follows at once that for either f_1 or f_2 and hence for any f_λ,

$$f_\lambda(\gamma_{\lambda c}) = (2\pi\overline{\gamma_c^2})^{-\frac{1}{2}} \exp(-\gamma_{\lambda c}^2/2\overline{\gamma_c^2}), \tag{8}$$

with

$$\overline{\gamma_c^2} \equiv \overline{\gamma_{\lambda c}^2} = \int_{-\infty}^{+\infty} \gamma_{\lambda c}^2 f(\gamma_{\lambda c})\, d\gamma_{\lambda c}. \tag{9}$$

The joint distribution of reduced width amplitudes for channel c is found from (4) and (8) to be

$$P_c(\gamma_{1c}, \gamma_{2c}, \cdots) = \text{const} \times \exp[-(\gamma_{1c}^2 + \gamma_{2c}^2 + \cdots)/2\overline{\gamma_c^2}]. \tag{10}$$

This result is consistent with the usual arguments.[1]

In a similar manner, we may define the joint distribution of reduced width amplitudes corresponding to the two channels c and c' as

$$P_{cc'}(\gamma_{1c}, \gamma_{2c}, \cdots, \gamma_{1c'}, \gamma_{2c'}, \cdots) = \iint \cdots \int P(\{\gamma_{\lambda c}\})\, d\gamma_{1c''}\, d\gamma_{2c''} \cdots , \tag{11}$$

which leads to

$$P_{cc'} = f_1(\gamma_{1c}, \gamma_{1c'}) f_2(\gamma_{2c}, \gamma_{2c'}) \cdots . \tag{12}$$

We now apply our invariance assumption (ii) to $P_{cc'}$. We again insist that

$$d(\ln P_c)/d\theta = 0, \tag{13}$$

upon introducing a pair of transformations (5) corresponding to c and c', respectively. Now (following a suggestion of Wick[4]) we write

$$\begin{aligned} \ln P_{cc'} &= \sum_\lambda \ln f_\lambda(\gamma_{\lambda c}, \gamma_{\lambda c'}) \\ &\equiv \sum_\lambda g_\lambda(\gamma_{\lambda c}, \gamma_{\lambda c'}). \end{aligned} \tag{14}$$

Noting that (for a pairwise level transformation)

$$d\gamma_{1c}/d\theta = +\gamma_{2c}, \qquad d\gamma_{2c}/d\theta = -\gamma_{1c}, \tag{15}$$

and corresponding equations for channel c', we have from (13), (14), and (15)

$$\frac{\partial g_1}{\partial \gamma_{1c}} \gamma_{2c} + \frac{\partial g_1}{\partial \gamma_{1c'}} \gamma_{2c'} = \frac{\partial g_2}{\partial \gamma_{2c}} \gamma_{1c} + \frac{\partial g_2}{\partial \gamma_{2c'}} \gamma_{1c'}. \tag{16}$$

Since $\gamma_{2c'}$ is arbitrary, we may set it equal to zero. This leads to the conclusion that $\partial g_1/\partial\gamma_{1c}$ is a *linear* function of γ_{1c} and $\gamma_{1c'}$. A similar procedure leads to the same conclusion for $\partial g_1/\partial\gamma_{1c'}$ with the consequence that we can write

$$g_1(\gamma_{1c}, \gamma_{1c'}) = a\gamma_{1c}^2 + b\gamma_{1c'}^2 + c\gamma_{1c}\gamma_{1c'}, \tag{17}$$

with a similar conclusion for g_2, and hence for any g_λ. Note that a, b, and c do *not* depend on λ, as can be seen by substitution in (16). This yields the result[5]

$$f_\lambda(\gamma_{\lambda c}, \gamma_{\lambda c'}) = \frac{|M|^{\frac{1}{2}}}{2\pi} \exp[-\tfrac{1}{2}(\gamma_\lambda, M\gamma_\lambda)], \tag{18}$$

where in this case γ_λ is a two-component vector, M is a two-by-two real symmetric matrix, and $|M|$ is the determinant of M. From (12) and (18) we have

$$P_{cc'} = \prod_\lambda \frac{|M|^{\frac{1}{2}}}{2\pi} \exp[-\tfrac{1}{2}(\gamma_\lambda, M\gamma_\lambda)]. \tag{19}$$

Note that, unless M is diagonal, $P_{cc'}$ is *not* an *inde-*

[4] G. C. Wick (private communication).

[5] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis* (John Wiley & Sons, Inc., New York 1958), p. 14.

pendent distribution with respect to channels. Thus, in the present approach, correlations may exist between reduced width amplitudes corresponding to different channels. This result is consistent with the observed fluctuating character of the total fission width since such fluctuation is not compatible with widely fluctuating *independent* partial fission widths, for example. We also note that this channel correlation is manifested in the joint distribution (19) but not in the singlet distribution (8).

The generalization of (19) to an arbitrary number m of channels is straightforward, yielding

$$P_{cc'c''\cdots} = \prod_{\lambda} \frac{|M|^{\frac{1}{2}}}{(2\pi)^{m/2}} \exp\left[-\tfrac{1}{2}(\gamma_\lambda, M\gamma_\lambda)\right], \tag{20}$$

with M an $m \times m$ real symmetric positive-definite matrix. It is easy to show from (20) that

$$\overline{\gamma_{\lambda c}^2} = \overline{\gamma_c^2} = \frac{1}{|M|}\frac{\partial}{\partial M_{cc}}|M| \tag{21}$$

is independent of λ, and that

$$\overline{\gamma_{\lambda c}\gamma_{\mu c'}} = \delta_{\lambda\mu}\frac{1}{2|M|}\frac{\partial}{\partial M_{cc'}}|M|, \tag{22}$$

where $\delta_{\lambda\mu}$ is the Kronecker delta. Of course,

$$\overline{\gamma_{\lambda c}} = 0. \tag{23}$$

For only two channels, c and c', we have

$$\begin{aligned}\overline{\gamma_c^2} &= \frac{M_{c'c'}}{M_{cc}M_{c'c'} - M_{cc'}^2},\\ \overline{\gamma_{\lambda c}\gamma_{\lambda c'}} &= \frac{-M_{cc'}}{M_{cc}M_{c'c'} - M_{cc'}^2},\end{aligned} \tag{24}$$

so that the channel correlation coefficient $C_{cc'}$ becomes

$$\begin{aligned}C_{cc'} &\equiv \overline{(\gamma_{\lambda c}\gamma_{\lambda c'})}/(\overline{\gamma_c^2}\,\overline{\gamma_{c'}^2})^{\frac{1}{2}}\\ &= -M_{cc'}/(M_{cc}M_{c'c'})^{\frac{1}{2}}.\end{aligned} \tag{25}$$

The reduced width correlation coefficient is easily found to be

$$\frac{\overline{\gamma_{\lambda c}^2\gamma_{\lambda c'}^2} - \overline{\gamma_{\lambda c}^2}\,\overline{\gamma_{\lambda c'}^2}}{\{(\overline{\gamma_{\lambda c}^4} - [\overline{\gamma_{\lambda c}^2}]^2)(\overline{\gamma_{\lambda c'}^4} - [\overline{\gamma_{\lambda c'}^2}]^2)\}^{\frac{1}{2}}} = (C_{cc'})^2, \tag{26}$$

implying that only a *positive* width correlation is possible. Note that (20) is inversion invariant so that moments of an odd character vanish.

The matrix M can of course be diagonalized by an orthogonal transformation. This means that in some intermediate state designation scheme there is no correlation. Such a situation may be relevant to, for example, fission for which there are probably correlations in the fragment-pair designated channel scheme, but for which there may be no correlations in the intermediate state designation scheme (see the Appendix of reference 1).

In order to see how fission fragment correlations of a realistic nature might be described, we examine (20) for the two-channel case. It should be pointed out that any multichannel form of (20) reduces to the two-channel form of (20) after all but two of the reduced width amplitudes have been integrated out and after proper identification of the relevant two-by-two matrix M has been made. We have

$$\begin{aligned}P_{cc'} = \prod_{\lambda} \frac{(M_{cc}M_{c'c'} - M_{cc'}^2)^{\frac{1}{2}}}{2\pi} \exp\left[-\tfrac{1}{2}(M_{cc}\gamma_{\lambda c}^2\right.\\ \left. + M_{c'c'}\gamma_{\lambda c'}^2 + 2M_{cc'}\gamma_{\lambda c}\gamma_{\lambda c'})\right].\end{aligned} \tag{27}$$

In order to see how $P_{cc'}$ behaves if correlations are large, we convert to units of the dispersion as given in (24):

$$\begin{aligned}P_{cc'} = \prod_{\lambda} \frac{1}{2\pi\overline{\gamma_{\lambda c}^2}\,\overline{\gamma_{\lambda c'}^2}(1-C^2)^{\frac{1}{2}}} \exp\left[-\frac{1}{2}\frac{1}{1-C^2}\right.\\ \left.\times\left(\frac{\gamma_{\lambda c}^2}{\overline{\gamma_{\lambda c}^2}} + \frac{\gamma_{\lambda c'}^2}{\overline{\gamma_{\lambda c'}^2}} - 2C\frac{\gamma_{\lambda c}\gamma_{\lambda c'}}{(\overline{\gamma_{\lambda c}^2}\,\overline{\gamma_{\lambda c'}^2})^{\frac{1}{2}}}\right)\right].\end{aligned} \tag{28}$$

From (28), it is easy to see that as C approaches ± 1, the behavior of $P_{cc'}$ has a delta function character with the singularity located on the line

$$\frac{\gamma_{\lambda c}}{[\overline{\gamma_{\lambda c}^2}]^{\frac{1}{2}}} = \pm\frac{\gamma_{\lambda c'}}{[\overline{\gamma_{\lambda c'}^2}]^{\frac{1}{2}}}, \tag{29}$$

implying a linear relationship between $\gamma_{\lambda c}$ and $\gamma_{\lambda c'}$ which would explain the correlation between partial fission widths that has been conjectured previously.

We note briefly that (21) and (22) are most easily expressed in terms of the inverse $M^{-1} \equiv \Sigma$ of M. The matrix Σ is called the covariance matrix. The appropriate result is

$$\overline{\gamma_\lambda \times \gamma_\lambda} = \Sigma. \tag{30}$$

It is perhaps relevant to point out that a linear connection between two completely correlated (or completely anticorrelated) random variables is a general result not tied to our particular distribution. The Schwartz inequality for two functions f and g is

$$(f, g)^2 \leq (f, f)(g, g), \tag{31}$$

where we have in mind that f and g are

$$\begin{aligned}f(x, y) &= x,\\ g(x, y) &= y,\end{aligned} \tag{32}$$

with x and y random variables (zero mean) distributed according to a distribution $P(x, y)$ and the

inner product is an integral with respect to the weight function P (probability distribution):

$$(f, g) = \int_{-\infty}^{+\infty} dx \int_{-\infty}^{+\infty} dy\, P(x, y) f(x, y) g(x, y), \quad (33)$$

so that the correlation coefficient C is given as

$$C^2 = \frac{(f, g)^2}{(f, f)(g, g)}. \quad (34)$$

If $C^2 = 1$, then the Schwartz inequality becomes an equality with the solution

$$f = \lambda g, \quad (35)$$

where λ is a constant. Thus if there is complete correlation or complete anti-correlation, we see that x and y are linearly related; hence the generality of the result.

III. RANDOM SIGN RULES

A typical procedure in the calculation of quantities of interest in nuclear reaction theory, such as the collision matrix, cross sections, transmission coefficients, etc., is to expand the collision matrix in terms of the reduced partial width amplitudes $\gamma_{\lambda c}$, say following the method of Thomas.[6] One then obtains series of terms which, aside from numerical factors, are of the general form

$$\frac{\gamma_{\lambda c}\gamma_{\lambda c'}\gamma_{\mu c'}\gamma_{\mu c''} \cdots P_c^{\frac{1}{2}} P_{c'}, P_{c''}^{\frac{1}{2}} \cdots}{\epsilon_\lambda \epsilon_\mu \cdots}, \quad (36)$$

where $\epsilon_\lambda \equiv E_\lambda - E - i\Gamma_\lambda/2$ (the level shifts having been caused to vanish by an appropriate choice of boundary conditions at the channel radii), and P_c is the usual channel penetration factor. At this point it is customary to simplify the calculation by invoking the statistical properties of the $\gamma_{\lambda c}$ in somewhat the following way. One assumes that the $\gamma_{\lambda c}$ are given by a probability law which makes positive and negative signs for each $\gamma_{\lambda c}$ equally likely. As a result, groups of terms formed by summing (36) over the great many levels in the interval K will tend to vanish because of mutual cancellation, provided (1) at least one $\gamma_{\lambda c}$ appears linearly (or to an odd power) in each term of the group, and (2) an average over energy E is taken. Proviso (1) ensures that the signs of the terms will truly fluctuate, while proviso (2) prevents one or two levels near E from dominating the group and thereby rendering the application of any level width statistics invalid. One difficulty with the above argument, often called a "random sign" approximation or rule, is that it refers merely to the equiprobability of positive and negative signs for any single $\gamma_{\lambda c}$, which is a property of the *singlet* distribution. It is clear, however, that the sign of a term of the form (36) is in general a property of the multivariate distribution. In the present section we shall discuss random sign rules from the point of view of the more general multivariate reduced width amplitude distribution derived above.

We postulate at the outset that a random sign rule is simply the application of the $\gamma_{\lambda c}$ statistics, i.e., the multivariate distribution, to the calculation of any quantity of interest Q. If Q is first averaged over energy E in order to prevent a small number of nearby resonances from dominating the contributions to Q, the application of $\gamma_{\lambda c}$ statistics means simply the averaging of Q over the multivariate distribution. In statistical language, we may say that the effect of the energy average is to reduce the dispersion of Q with respect to the multivariate distribution. Let I denote the energy interval over which Q is averaged, and let E_0 denote its midpoint. It will be convenient to choose $I \ll K$, where K is the large interval centered at E_0 over which the $\gamma_{\lambda c}$ statistics may be assumed constant. Since the interval I must, of course, contain many levels, we must have $D \ll I \ll K$. In practice, $K \sim 0.1$ MeV, $I \sim 1$ keV, and $D \sim 1$ eV. If we wish to include the effects of very distant levels, i.e., those outside K, it is necessary to allow for the possibility of variable statistics over the entire energy range. The case of variable statistics will be discussed in Sec. IV.

If we denote by $\bar{Q}$ the result of an energy average and a multivariate $\gamma_{\lambda c}$ average, the basic random sign rule may be written symbolically as

$$\bar{Q}(E_0) = \int d(\{\gamma_{\lambda c}\})\, P(\{\gamma_{\lambda c}\}) \int dE' \times R_I(E_0 - E') Q(E', \{\gamma_{\lambda c}\}), \quad (37)$$

where $R_I(E - E')$ is the energy resolution or weight function characterized by the width I. We do not expect our results to be sensitive to the form of R_I which may be taken in the typical case to be Lorentzian or rectangular in shape. From (37) we note that if $F(Q)$ is some function of Q, in general,

$$\overline{F(Q)} \neq F(\bar{Q}). \quad (38)$$

However, considering the collision matrix

$$U(E_0 + iI/2)$$

evaluated at the complex energy $E_0 + iI/2$ to be a function $U(R(E_0 + iI/2))$ of the R matrix similarly evaluated, it can be shown that $\bar{U}(R(E_0 + iI/2)) =$

[6] R. G. Thomas, Phys. Rev. 97, 224 (1955).

$U(\bar{R}(E_0 + iI/2))$, where here the bar denotes *only* the $\gamma_{\lambda c}$ average. This may appear, in a sense, to be an exception to (38).

To illustrate the somewhat abstract arguments presented above, we shall apply the multivariate distribution to the calculation of the energy average of the collision matrix U. In terms of the $\gamma_{\lambda c}$'s, the matrix element $U_{cc'}$ may be written following Thomas[6] as

$$U_{cc'} = e^{-i(\varphi_c+\varphi_{c'})}\Bigg[\delta_{cc'} + 2i\sum_{\lambda}\frac{\gamma_{\lambda c}\gamma_{\lambda c'}P_c^{\frac{1}{2}}P_{c'}^{\frac{1}{2}}}{\epsilon_\lambda} - 2\sum_{\lambda\neq\mu}\sum_{\lambda}\sum_{c''}\frac{\gamma_{\lambda c}\gamma_{\lambda c''}\gamma_{\mu c''}\gamma_{\mu c'}P_c^{\frac{1}{2}}P_{c''}P_{c'}^{\frac{1}{2}}}{\epsilon_\lambda\epsilon_\mu} + \cdots\Bigg], \tag{39}$$

where φ_c and $\varphi_{c'}$ are the hard-sphere phase shifts, assumed constant (along with P_c) over the interval I. For simplicity, we shall neglect all levels outside a very large interval K centered about E_0. Applying (37), the multivariate width distribution (20), and a Lorentzian resolution function $R_I(E' - E)$, we find

$$\overline{U_{cc'}} = e^{-i(\varphi_c+\varphi_{c'})}\Bigg[\delta_{cc'} + 2i\sum_{\lambda}\frac{\overline{\gamma_{\lambda c}\gamma_{\lambda c'}}P_c^{\frac{1}{2}}P_{c'}^{\frac{1}{2}}}{\epsilon'_\lambda} - 2\sum_{c''}\sum_{\lambda\neq\mu}\sum\frac{\overline{\gamma_{\lambda c}\gamma_{\lambda c''}\,\gamma_{\mu c''}\gamma_{\mu c'}}P_c^{\frac{1}{2}}P_{c''}P_{c'}^{\frac{1}{2}}}{\epsilon'_\lambda\epsilon'_\mu} + \cdots\Bigg], \tag{40}$$

where

$$\epsilon'_\lambda = E_\lambda - E_0 - i(\Gamma_\lambda + I)/2 \cong E_\lambda - E_0 - iI/2. \tag{41}$$

Introducing the symmetric channel matrix Σ defined by (30), and the diagonal matrices P and Ω defined by $P_{cc'} = \delta_{cc'}P_c(E_0)$ and $\Omega_{cc'} = \delta_{cc'}e^{-i\varphi_c}$ respectively, we may rewrite (40) as

$$\bar{U} = \Omega\bar{W}\Omega, \tag{42}$$

$$\bar{W} \equiv 1 + 2i\Big(\sum_{\lambda}\frac{1}{\epsilon'_\lambda}\Big)P^{\frac{1}{2}}\Sigma P^{\frac{1}{2}} - 2\Big(\sum_{\lambda\neq\mu}\sum\frac{1}{\epsilon'_\lambda\epsilon'_\mu}\Big)P^{\frac{1}{2}}\Sigma P\Sigma P^{\frac{1}{2}} + \cdots. \tag{43}$$

To obtain $\bar{W}$, it is therefore necessary to evaluate level sums like

$$\sum_{\lambda}\frac{1}{\epsilon'_\lambda}, \qquad \sum_{\lambda\neq\mu}\sum\frac{1}{\epsilon'_\lambda\epsilon'_\mu}, \quad \text{etc.}$$

These sums depend clearly on the set of values E_λ, which are themselves given by a certain statistical distribution. The question of the level (or eigenvalue) distribution will be discussed in the next section. For present purposes, we shall suppose that the levels are randomly and independently distributed with mean spacing D.

Considering first the sum $\sum_\lambda (1/\epsilon'_\lambda)$, we note that because of its large imaginary part, ϵ'_λ varies slowly as we pass from level to level. The sum may therefore be replaced by an integral, yielding in the limit of large K

$$\sum_{\lambda}\frac{1}{\epsilon'_\lambda} = \frac{1}{D}\int_K\frac{dE_\lambda}{\epsilon'_\lambda} = \frac{i\pi}{D}. \tag{44}$$

Similarly,

$$\sum_{\lambda\neq\mu}\sum\frac{1}{\epsilon'_\lambda\epsilon'_\mu} = \Big(\sum_{\lambda}\frac{1}{\epsilon'_\lambda}\Big)^2 - \Big(\sum_{\lambda}\frac{1}{\epsilon'^2_\lambda}\Big), = -\pi^2/D^2, \tag{45}$$

as follows from (44) and the vanishing of $\sum_\lambda (1/\epsilon'^2_\lambda)$ in the limit of large K. Inserting (44), (45), and similar results for higher-order sums into (43), we have

$$\bar{W} = 1 - \frac{2\pi}{D}P^{\frac{1}{2}}\Sigma P^{\frac{1}{2}} + \frac{2\pi^2}{D^2}P^{\frac{1}{2}}\Sigma P\Sigma P^{\frac{1}{2}} + \cdots = P^{\frac{1}{2}}\Big[1 + \frac{\pi}{D}\Sigma P\Big]^{-1}\Big[1 - \frac{\pi}{D}\Sigma P\Big]P^{-\frac{1}{2}}. \tag{46}$$

In the special case of diagonal Σ, i.e., no channel-channel correlations, we have from (46) and (42)

$$\bar{U}_{cc'} = e^{-2i\varphi_c}\frac{1 - \pi s_c P_c}{1 + \pi s_c P_c}\delta_{cc'}, \tag{47}$$

where $s_c \equiv \overline{\gamma^2_{\lambda c}}/D$ is the usual (reduced) pole strength function for channel c in the interval K. The expression (47) differs from the one given by Feshbach, Porter, and Weisskopf,[7] by Feshbach[8] and rederived by Moldauer,[9] but agrees with that of Thomas[6] for the "picket-fence" model which assumes a constant level separation and constant $\gamma_{\lambda c}$. Recently Moldauer[10] has modified his earlier results[9] and has obtained an expression for $\bar{U}_{cc}$ which, in the special case of constant strength function, agrees with (47). Moldauer[10] has also treated the case of correlations between two channels c and c' in some detail. However, his general method is immediately applicable to the case of correlations between any number of channels; and a combination of his equations yields a result in agreement with (46).

IV. EIGENVALUE DISTRIBUTIONS AND VARIABLE STATISTICS

Thus far, it has been assumed that the energy eigenvalues E_λ are randomly and independently dis-

[7] H. Feshbach, C. E. Porter, and V. F. Weisskopf, Phys. Rev. **96**, 448 (1954).
[8] H. Feshbach, "The Complex Potential Model," *Nuclear Spectroscopy (Part B)*, edited by F. Ajzenberg-Selove (Academic Press, Inc., New York, 1960), Chap. VI. D.
[9] P. A. Moldauer, Phys. Rev. **123**, 968 (1961).
[10] P. A. Moldauer, Phys. Rev. **129**, 754 (1963).

tributed with mean spacing D. We now consider the possible influence on $\bar{Q}$ of the details of the eigenvalue distribution, e.g., of the level-repulsion effect. Let $P(\{E_\lambda\})$ denote the distribution of the N levels in the interval K. The definition of $\bar{Q}$ given in (37) can now be generalized to include, in addition to the energy and reduced width amplitude average, an average over $P(\{E_\lambda\})$:

$$\bar{Q}(E_0) \equiv \int d\{E_\lambda\}\, P(\{E_\lambda\}) \int d\{\gamma_{\lambda c}\}\, P(\{\gamma_{\lambda c}\}) \times \int dE'\, R_I(E_0 - E') Q(E', \{\gamma_{\lambda c}\}, \{E_\lambda\}). \quad (48)$$

In (48), the dependence of Q, the quantity of interest, on the set of eigenvalues $\{E_\lambda\}$ has been included explicitly. The eigenvalue distribution $P(\{E_\lambda\})$, which can be shown to be statistically independent of the $\gamma_{\lambda c}$ distribution, is for the Gaussian orthogonal ensemble[2]

$$P(\{E_\lambda\}) = A[\prod_{\lambda<\mu} |E_\lambda - E_\mu|] \times \exp\,[-\sum_\lambda (E_\lambda - E_0)^2/4\sigma^2], \quad (49)$$

where A is a normalization constant, E_0 is the mean of the eigenvalues and σ is a dispersion constant related to D by

$$D = \pi\sigma/N^{\frac{1}{2}}. \quad (50)$$

[The distribution (49) refers to levels E_λ of the same symmetry quantum numbers J and Π. Since sets of levels corresponding to different symmetries are distributed independently of each other, the most general form of $P(\{E_\lambda\})$ is a product of factors of the form (49), one for each symmetry. In most cases of interest, Q is expressible as a sum of terms each of which is of a single symmetry or at most of a product of symmetries. Averaging over the general $P(\{E_\lambda\})$ is then accomplished with each factor of the form (49) separately. If Q happens to be symmetrical in all the E_λ, and several symmetries are involved, the generalized $P(\{E_\lambda\})$ may be symmetrized without affecting the value of $\bar{Q}$. The symmetrized form of $P(\{E_\lambda\})$ is then closely approximated by a "random" or Poisson distribution, with a mean spacing simply related to the mean spacings of the individual component distributions.]

The distribution of a single eigenvalue E_λ is obtained from (49) by integrating out all variables except E_λ. The resulting distribution is the well-known semicircle law of Wigner

$$P(E_\lambda) = \frac{1}{\pi N^{\frac{1}{2}}\sigma}\left[1 - \frac{(E_\lambda - E_0)^2}{4N\sigma^2}\right]^{\frac{1}{2}}. \quad (51)$$

If all but two variables, E_λ and E_μ, are integrated out of (49), the pair distribution $P(E_\lambda, E_\mu)$ law is obtained. Dyson[12] has derived the form of $P(E_\lambda, E_\mu)$ for the circular orthogonal ensemble in the limit $N \to \infty$. The repulsion effect is manifested in a factor $|E_\lambda - E_\mu|$ when $|E_\lambda - E_\mu| \lesssim D$, while for $|E_\lambda - E_\mu| \gg D$, $P(E_\lambda, E_\mu)$ is essentially a product of two independent single-level distributions $P(E_\lambda)P(E_\mu)$.

As typical examples of the inclusion of the eigenvalue distribution in the averaging procedure, as defined by (48), we consider the sums $\sum_\lambda (1/\epsilon_\lambda')$ and $\sum\sum_{\lambda\neq\mu} (1/\epsilon_\lambda'\epsilon_\mu')$, which arise in the calculation of $\bar{W}$ (Eq. 43). These sums were evaluated under the assumption of uncorrelated levels in (44) and (45), respectively. We shall now see that these results are essentially unchanged by the introduction of the eigenvalue distribution $P(\{E_\lambda\})$. Considering first the single sum $\sum_\lambda (1/\epsilon_\lambda')$, we observe that its average over $P(\{E_\lambda\})$ requires the use of only the single-level law, (51). We have, in fact,

$$\overline{\sum_{\lambda=1}^{N} \frac{1}{\epsilon_\lambda'}} = N\,\overline{\frac{1}{\epsilon_\lambda'}} = N\int_{-\infty}^{\infty} dE_\lambda\, P(E_\lambda)\,\frac{1}{\epsilon_\lambda'} = \frac{i\pi}{D}, \quad (52)$$

in agreement with (44). We also have, in the limit with $\alpha = 2, 3, \cdots N \to \infty$,

$$\overline{\sum_{\lambda=1}^{N} \frac{1}{\epsilon_\lambda'^{\alpha}}} = N\,\overline{\frac{1}{\epsilon_\lambda'^{\alpha}}} = N\int_{-\infty}^{\infty} dE_\lambda\, P(E_\lambda)\,\frac{1}{\epsilon_\lambda'^{\alpha}} = 0. \quad (53)$$

We consider next the average of $\sum\sum_{\lambda\neq\mu} (1/\epsilon_\lambda'\epsilon_\mu')$ over $P(\{E_\lambda\})$. This is more difficult to evaluate, owing to the complicated form of $P(E_\lambda, E_\mu)$. However, for present purposes, an approximate description of $P(E_\lambda, E_\mu)$ which exhibits the essential features of the repulsion effect will be adequate. We therefore take for $P(E_\lambda, E_\mu)$ the separated form $P(E_\lambda)P(E_\mu)$ over the entire E_λ, E_μ plane except for a strip of width $\sim D$ about the line $E_\lambda = E_\mu$. Within this strip, $P(E_\lambda, E_\mu)$ will decrease as the line $E_\lambda = E_\mu$ is approached and will vanish as $|E_\lambda - E_\mu|$ in the neighborhood of the line. It is then readily verified that, except for corrections of order D/I and I/K, and in the limit of large N,

$$\overline{\sum_{\lambda\neq\mu}\sum \frac{1}{\epsilon_\lambda'\epsilon_\mu'}} = \left[\overline{\sum_\lambda \frac{1}{\epsilon_\lambda'}}\right]^2 - \overline{\sum_\lambda \frac{1}{\epsilon_\lambda'^2}} = -\pi^2/D^2, \quad (54)$$

in agreement with (45). Similar arguments show that the higher order sums of the above type appearing in the expansion of $\bar{W}$ are also unaffected by averaging over $P(\{E_\lambda\})$, at least to order D/I and I/K.

In view of the above discussion, we conclude that

$\bar{W}$, and hence also $\bar{W}^\dagger$, are independent of the details of the eigenvalue distribution except, of course, for the mean spacing D. We shall not consider here the influence of the eigenvalue distribution on quantities such as $\overline{|W_{cc'}|^2}$ which have singularities in both upper and lower halves of the complex E plane.

We comment finally on the inclusion of very distant levels, i.e., those outside K, into the present scheme of evaluating $\bar{Q}$. The difficulty which arises is that the reduced width amplitude and eigenvalue statistics may no longer be considered constant over the entire energy range. For example, the mean level spacing and strength functions are actually weakly energy dependent, although we have thus far assumed them to be constant. The slow variation in the underlying statistical distributions over energy intervals larger than K may be accounted for by allowing a correspondingly slow variation with λ in the statistical parameters, i.e., D and Σ. With this understanding, the definition of the basic average (48) is still applicable, λ being no longer restricted to those levels within K. This modification results in some complications in performing the averages. Thus, sums like $\sum_\lambda (1/\epsilon'_\lambda)$ and $\sum\sum_{\lambda\neq\mu} (1/\epsilon'_\lambda\epsilon'_\mu)$ do not factor out as in (43), since Σ is no longer constant. The simple results of (44) and (45) are thereby rendered inapplicable. In place of (44), for example, we need to consider a sum like $\sum_\lambda f(E_\lambda)/\epsilon'_\lambda$ where $f(E_\lambda)$ is some slowly varying real function of E_λ. We then have

$$\sum_\lambda \frac{f(E_\lambda)}{\epsilon'_\lambda} = \int \frac{f(E_\lambda)D^{-1}(E_\lambda)}{\epsilon'_\lambda}\, dE_\lambda$$

$$= P\int \frac{f(E_\lambda)D^{-1}(E_\lambda)}{E_\lambda - E_0}\, dE_\lambda + i\pi \frac{f(E_0)}{D(E_0)}. \qquad (55)$$

The second integral in (55) may be regarded not only as the contribution from distant levels but also as the correction for the variation of the statistics over the interval K, e.g., neglecting levels outside K, a first approximation to the real part of (55) is $K[(d/dE_\lambda)\ f(E_\lambda)D^{-1}(E_\lambda)]_{E_\lambda=E_0}$ which is just the change in $f(E_\lambda)D^{-1}(E_\lambda)$ over the interval K.

V. CONCLUDING REMARKS

We have shown, at least in principle, how to apply the concept of channel–channel correlations to R-matrix theory where random sign rules are needed. On the other hand, Feshbach's approach[8] appears not to require random sign rules, although he does make use of the singlet distribution to simplify[11] the resonant part of the transition matrix in the case of many overlapping levels. Thus channel–channel correlation may be easier to handle in a Hamiltonian-based theory. Certainly, from a statistical point of view, it is expected that if randomness assumptions are needed in the development of a theory, they will enter more naturally through the Hamiltonian matrix elements than through a separate treatment of the reduced width amplitudes and the eigenvalues, since both the width *and* the eigenvalue statistics are implied by statistical hypotheses concerning the Hamiltonian.

As in all problems analyzed with R-matrix theory, a question arises here concerning the independence of the results on the particular choice of the set of channel radii a_c, parameters in the theory whose values may be chosen rather arbitrarily above a certain set of minimum values. This is a difficult question not easily resolved. We merely remark here that the assumptions (i) and (ii) of Sec. II, which lead to the multivariate distribution (20) are not necessarily valid at all values of the a_c. This means that the form of the $\gamma_{\lambda c}$ distribution may actually depend on the choice of the a_c. Presumably, and for similar reasons, the form of the eigenvalue distribution may also depend on the a_c. The results obtained by averaging Q over the distributions should, however, still be independent of the a_c, since the parameters of the distributions can change in a compensating fashion. It seems most likely that the minimum values of a_c, i.e., those which correspond to the sum of the radii of the interacting nuclei, will be those for which assumptions (i) and (ii) and also for which the assumptions leading to the eigenvalue distribution (49) are most nearly valid, since it is for this choice that maximum "complexity density" of the compound wavefunction is achieved. As a_c is increased above its minimum permissible value, one would expect the distributions of both $\{\gamma_{\lambda c}\}$ and $\{E_\lambda\}$ to become increasingly narrow and rigid, with the parameters also changing in order to maintain a constant value of $\bar{Q}$.

It is to be emphasized that the multivariate reduced width amplitude distribution developed in this paper, although presented in the language of R-matrix theory, is not tied to that approach to reaction theory. The definitions of the reduced width amplitudes $\gamma_{\lambda c}$ are easily transferred directly to any other theory whether or not it is necessary to develop

[11] H. Feshbach, "The Compound Nucleus," *Nuclear Spectroscopy (Part B)*, edited by F. Ajzenberg-Selove (Academic Press, Inc., New York, 1960), Chap. V.A.

random sign rules for an expansion of the collision matrix.

We remark briefly that in connection with assumption (i) all considerations could have been stated mathematically in terms of channel, rather than level, independence. Such a theory would have led to identically equal strength functions in all channels which is physically unrealistic.

It is also important to note that we have used only orthogonal invariance to obtain (20). In the presence of a strong external magnetic field, for example, time reversal invariance is violated so that unitary invariance must be required.[12] If an odd half-integral spin system (atom) is placed in an external crystalline multipole field, then symplectic invariance is appropriate since angular momentum conservation is no longer valid. These considerations are discussed by Dyson.[13]

[12] See N. Ullah, J. Math. Phys. 4, 1279 (1963).

ACKNOWLEDGMENTS

We have benefited from a conversation with Dr. N. Rosenzweig relating to the derivation of the reduced width amplitude distribution in Sec. I. The remarks in Sec. V concerning the dependence of the results on the choice of channel radii evolved from discussions with Dr. E. P. Wigner. Communications from Dr. P. A. Moldauer have also been useful in clarifying the connection between the present work and references 9 and 10.

[13] F. J. Dyson, J. Math. Phys. 3, 140 (1962); 3, 157 (1962); 3, 166 (1962); 3, 1191 (1962); 3, 1199 (1962); F. J. Dyson and M. L. Mehta, J. Math. Phys. 4, 701 (1963); M. L. Mehta and F. J. Dyson, J. Math. Phys. 4, 713 (1963).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 4, NUMBER 10 OCTOBER 1963

Reduced Width Amplitude Distributions for the Unitary Ensemble*

NAZAKAT ULLAH†
Brookhaven National Laboratory, Upton, Long Island, New York
(Received 22 April 1963)

The reduced width amplitude distributions for a system not invariant under time reversal (unitary ensemble) are derived. The form of the joint distribution function is compared with that of the orthogonal ensemble. The calculation of the reduced width channel–channel correlation coefficient shows that only a positive width correlation is possible.

I. INTRODUCTION

THE reduced width amplitude distributions for a system invariant under an orthogonal transformation have been derived recently by Krieger and Porter.[1] This derivation is extended here for a unitary ensemble. The unitary ensemble has been defined by Dyson[2] and applies when time-reversal invariance is violated. Such a situation occurs when an atom or a nucleus is placed in an external magnetic field. For such systems the Hamiltonian is an arbitrary complex Hermitian matrix.

We shall derive the reduced width amplitude distributions using the principle of level independence and of invariance with respect to unitary transformations of the basic-state vectors of the compound system. These distributions will be used to calculate the correlation coefficients.

II. DERIVATION OF DISTRIBUTION FUNCTION

The reduced-width amplitudes $\gamma_{\lambda c}$, where λ denotes a particular level and c a particular channel, are complex for a unitary ensemble. Let us denote the real and imaginary parts of $\gamma_{\lambda c}$ by $\gamma^r_{\lambda c}$, $\gamma^i_{\lambda c}$, respectively. The joint distribution function of $\gamma_{\lambda c}$'s is written as

$$P(\{\gamma_{\lambda c}\}) = P(\gamma_{1c}, \gamma^*_{1c}, \gamma_{1c'}, \gamma^*_{1c'}, \cdots \gamma_{2c}, \gamma^*_{2c}, \gamma_{2c'}, \gamma^*_{2c'}, \cdots). \quad (1)$$

Using the assumption of level independence, P can

* Research supported by the U. S. Atomic Energy Commission.

† Present address: Tata Institute of Fundamental Research, Colaba, Bombay, India.

[1] T. J. Krieger and C. E. Porter, J. Math. Phys. 4, 1272 (1963).

[2] F. J. Dyson, J. Math. Phys. 3, 140 (1962).

random sign rules for an expansion of the collision matrix.

We remark briefly that in connection with assumption (i) all considerations could have been stated mathematically in terms of channel, rather than level, independence. Such a theory would have led to identically equal strength functions in all channels which is physically unrealistic.

It is also important to note that we have used only orthogonal invariance to obtain (20). In the presence of a strong external magnetic field, for example, time reversal invariance is violated so that unitary invariance must be required.[12] If an odd half-integral spin system (atom) is placed in an external crystalline multipole field, then symplectic invariance is appropriate since angular momentum conservation is no longer valid. These considerations are discussed by Dyson.[13]

[12] See N. Ullah, J. Math. Phys. 4, 1279 (1963).

ACKNOWLEDGMENTS

We have benefited from a conversation with Dr. N. Rosenzweig relating to the derivation of the reduced width amplitude distribution in Sec. I. The remarks in Sec. V concerning the dependence of the results on the choice of channel radii evolved from discussions with Dr. E. P. Wigner. Communications from Dr. P. A. Moldauer have also been useful in clarifying the connection between the present work and references 9 and 10.

[13] F. J. Dyson, J. Math. Phys. 3, 140 (1962); 3, 157 (1962); 3, 166 (1962); 3, 1191 (1962); 3, 1199 (1962); F. J. Dyson and M. L. Mehta, J. Math. Phys. 4, 701 (1963); M. L. Mehta and F. J. Dyson, J. Math. Phys. 4, 713 (1963).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 4, NUMBER 10 OCTOBER 1963

Reduced Width Amplitude Distributions for the Unitary Ensemble*

NAZAKAT ULLAH†
Brookhaven National Laboratory, Upton, Long Island, New York
(Received 22 April 1963)

The reduced width amplitude distributions for a system not invariant under time reversal (unitary ensemble) are derived. The form of the joint distribution function is compared with that of the orthogonal ensemble. The calculation of the reduced width channel–channel correlation coefficient shows that only a positive width correlation is possible.

I. INTRODUCTION

THE reduced width amplitude distributions for a system invariant under an orthogonal transformation have been derived recently by Krieger and Porter.[1] This derivation is extended here for a unitary ensemble. The unitary ensemble has been defined by Dyson[2] and applies when time-reversal invariance is violated. Such a situation occurs when an atom or a nucleus is placed in an external magnetic field. For such systems the Hamiltonian is an arbitrary complex Hermitian matrix.

We shall derive the reduced width amplitude distributions using the principle of level independence and of invariance with respect to unitary transformations of the basic-state vectors of the compound system. These distributions will be used to calculate the correlation coefficients.

II. DERIVATION OF DISTRIBUTION FUNCTION

The reduced-width amplitudes $\gamma_{\lambda c}$, where λ denotes a particular level and c a particular channel, are complex for a unitary ensemble. Let us denote the real and imaginary parts of $\gamma_{\lambda c}$ by $\gamma^r_{\lambda c}$, $\gamma^i_{\lambda c}$, respectively. The joint distribution function of $\gamma_{\lambda c}$'s is written as

$$P(\{\gamma_{\lambda c}\}) = P(\gamma_{1c}, \gamma^*_{1c}, \gamma_{1c'}, \gamma^*_{1c'}, \cdots \gamma_{2c}, \gamma^*_{2c}, \gamma_{2c'}, \gamma^*_{2c'}, \cdots). \quad (1)$$

Using the assumption of level independence, P can

* Research supported by the U. S. Atomic Energy Commission.

† Present address: Tata Institute of Fundamental Research, Colaba, Bombay, India.

[1] T. J. Krieger and C. E. Porter, J. Math. Phys. 4, 1272 (1963).

[2] F. J. Dyson, J. Math. Phys. 3, 140 (1962).

be written as a product of functions $f_\lambda(\gamma_{\lambda c}, \gamma_{\lambda c}^* \cdots)$, i.e.,

$$P(\{\gamma_{\lambda c}\}) = f_1(\gamma_{1c}, \gamma_{1c}^*, \gamma_{1c'}, \gamma_{1c'}^* \cdots) \times f_2(\gamma_{2c}, \gamma_{2c}^*, \gamma_{2c'}, \gamma_{2c'}^*, \cdots) \cdots . \quad (2)$$

By integrating out over all the other channels but c in (1), we get the single-channel distribution

$$P_c(\gamma_{1c}, \gamma_{1c}^*, \gamma_{2c}, \gamma_{2c}^* \cdots) = \int_{-\infty}^{+\infty} \cdots \int P(\{\gamma_{\lambda c}\})\, d\gamma_{1c}^r\, d\gamma_{1c}^i\, d\gamma_{2c}^r\, d\gamma_{2c}^i \cdots . \quad (3)$$

Using (2), it can be written in the form

$$P_c(\gamma_{1c}, \gamma_{1c}^*, \gamma_{2c}, \gamma_{2c}^* \cdots) = f_1(\gamma_{1c}, \gamma_{1c}^*) f_2(\gamma_{2c}, \gamma_{2c}^*) \cdots , \quad (4)$$

where the functions f are obtained by integrating over all the other channels. We now introduce the second assumption that P is form-invariant with respect to a unitary transformation. The most general 2×2 unitary transformation having the determinant $+1$ is

$$U = \begin{pmatrix} e^{i\xi} \cos\omega & e^{i\eta} \sin\omega \\ -e^{-i\eta} \sin\omega & e^{-i\xi} \cos\omega \end{pmatrix}. \quad (5)$$

Let us apply this unitary transformation to a pair of levels $\lambda = 1, 2$; then

$$\begin{pmatrix} \gamma_{1c} \\ \gamma_{2c} \end{pmatrix} = U \begin{pmatrix} \gamma_{1c}' \\ \gamma_{2c}' \end{pmatrix}, \quad (6)$$

where a prime on $\gamma_{\lambda c}$'s denotes the new $\gamma_{\lambda c}$'s. From (4) we can write for a pair of levels

$$\ln P_c = \ln f_1(\gamma_{1c}, \gamma_{1c}^*) + \ln f_2(\gamma_{2c}, \gamma_{2c}^*).$$

The invariance hypothesis gives $(\partial/\partial t) \ln P_c = 0$, where t can be ξ, η, or ω. Therefore we have

$$\frac{1}{f_1(\gamma_{1c}, \gamma_{1c}^*)} \left[\frac{\partial f_1(\gamma_{1c}, \gamma_{1c}^*)}{\partial \gamma_{1c}} \frac{\partial \gamma_{1c}}{\partial t} + \frac{\partial f_1(\gamma_{1c}, \gamma_{1c}^*)}{\partial \gamma_{1c}^*} \frac{\partial \gamma_{1c}^*}{\partial t} \right] + \frac{1}{f_2(\gamma_{2c}, \gamma_{2c}^*)} \left[\frac{\partial f_2(\gamma_{2c}, \gamma_{2c}^*)}{\partial \gamma_{2c}} \frac{\partial \gamma_{2c}}{\partial t} + \frac{\partial f_2(\gamma_{2c}, \gamma_{2c}^*)}{\partial \gamma_{2c}^*} \frac{\partial \gamma_{2c}^*}{\partial t} \right] = 0. \quad (7)$$

Putting in the values of $\partial\gamma_{\lambda c}/\partial t$ from (6), and equating to zero the expressions multiplying $e^{i(\xi+\eta)}$ and $e^{-i(\xi+\eta)}$, and after some simplification, it can be shown that

$$f_1(\gamma_{1c}, \gamma_{1c}^*) = A \exp[-\mu |\gamma_{1c}|^2], \quad (8)$$

where A and μ are fixed using the relations

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} d\gamma_{\lambda c}^r\, d\gamma_{\lambda c}^i f_\lambda(\gamma_{\lambda c}, \gamma_{\lambda c}^*) = 1$$

(normalization condition),

$$\overline{|\gamma_c|^2} = \overline{|\gamma_{\lambda c}|^2} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} |\gamma_{\lambda c}|^2 f_\lambda(\gamma_{\lambda c}, \gamma_{\lambda c}^*)\, d\gamma_{\lambda c}^r\, d\gamma_{\lambda c}^i. \quad (10)$$

Expression (8) can then be written as

$$f_\lambda(\gamma_{\lambda c}, \gamma_{\lambda c}^*) = (\pi\, \overline{|\gamma_c|^2})^{-1} \exp[-\gamma_{\lambda c}\gamma_{\lambda c}^*/\overline{|\gamma_c|^2}]. \quad (11)$$

Using (4) and (11) we can write the joint distribution for a single channel as

$$P_c(\gamma_{1c}, \gamma_{1c}^*, \gamma_{2c}, \gamma_{2c}^*, \cdots) = (\text{const}) \times \exp[-(|\gamma_{1c}|^2 + |\gamma_{2c}|^2 + \cdots)/\overline{|\gamma_c|^2}]. \quad (12)$$

We shall now derive the single-channel distribution of partial width $\Gamma = |\gamma|^2$. We write $\gamma_{\lambda c} = |\gamma_{\lambda c}|\, e^{i\theta}$, then $d\gamma_{\lambda c}^r\, d\gamma_{\lambda c}^i = |\gamma_{\lambda c}|\, d\,|\gamma_{\lambda c}|\, d\theta = \frac{1}{2} d\Gamma_{\lambda c}\, d\theta$. Therefore using (11) we can write

$$f_\lambda(\gamma_{\lambda c}, \gamma_{\lambda c}^*)\, d\gamma_{\lambda c}^r\, d\gamma_{\lambda c}^i = (\pi\, \overline{|\gamma_c|^2})^{-1} \times \exp[-\Gamma_{\lambda c}/\overline{|\gamma_c|^2}] \tfrac{1}{2}\, d\Gamma_{\lambda c}\, d\theta.$$

Integrating over θ, we get the single-channel distribution for partial width $\Gamma_{\lambda c}$ as[3]

$$P(\Gamma_{\lambda c}) = (\bar{\Gamma}_c)^{-1} \exp[-\Gamma_{\lambda c}/\bar{\Gamma}_c]. \quad (13)$$

It is interesting to note that for real $\gamma_{\lambda c}$'s, i.e., for a system invariant under time reversal, the single-channel distribution for partial width $\Gamma_{\lambda c}$ is[4]

$$P(\Gamma_{\lambda c}) = (2\pi\bar{\Gamma}_c)^{-\frac{1}{2}} \exp[-\Gamma_{\lambda c}/2\bar{\Gamma}_c]/\Gamma_{\lambda c}^{\frac{1}{2}}. \quad (14)$$

Thus the single-channel distribution of partial width for the system invariant under time reversal differs from that which violates time-reversal invariance in having a factor of $\Gamma_{\lambda c}^{\frac{1}{2}}$ in the denominator and a somewhat different dependence on $\bar{\Gamma}_c$.

We shall now derive the joint distribution of reduced-width amplitudes for the multichannel case. Just as in the case of single channel, the joint distribution function for two channels c, c′ is obtained by integrating (1) over the rest of the channels c″, c‴, $\cdots$,

$$P_{cc'}(\gamma_{1c}, \gamma_{1c}^*, \gamma_{2c}, \gamma_{2c}^*, \cdots \gamma_{1c'}, \gamma_{1c'}^*, \cdots) = \int \cdots \int P(\{\gamma_{\lambda c}\})\, d\gamma_{1c''}^r\, d\gamma_{1c''}^i \cdots . \quad (15)$$

Using (2) it can be written as

[3] J. M. C. Scott, Phil. Mag. 45, 1322 (1954).
[4] R. G. Thomas and C. E. Porter, Phys. Rev. 104, 483 (1956).

$$P_{cc'} = f_1(\gamma_{1c}, \gamma_{1c}^*, \gamma_{1c'}, \gamma_{1c'}^*) \times f_2(\gamma_{2c}, \gamma_{2c}^*, \gamma_{2c'}, \gamma_{2c'}^*) \cdots . \quad (16)$$

Therefore,

$$\ln P_{cc'} = \ln f_1(\gamma_{1c}, \gamma_{1c}^*, \gamma_{1c'}, \gamma_{1c'}^*) + \ln f_2(\gamma_{2c}, \gamma_{2c}^*, \gamma_{2c'}, \gamma_{2c'}^*) + \cdots = \sum_\lambda g_\lambda(\gamma_{\lambda c}, \gamma_{\lambda c}^*, \gamma_{\lambda c'}, \gamma_{\lambda c'}^*), \quad (17)$$

where we have now introduced the new functions g_λ's which are easier to work with. The invariance hypothesis $(\partial/\partial t) \ln P_{cc'} = 0$ gives for a pair of levels

$$\frac{\partial g_1}{\partial \gamma_{1c}}\frac{\partial \gamma_{1c}}{\partial t} + \frac{\partial g_1}{\partial \gamma_{1c}^*}\frac{\partial \gamma_{1c}^*}{\partial t} + \frac{\partial g_1}{\partial \gamma_{1c'}}\frac{\partial \gamma_{1c'}}{\partial t} + \frac{\partial g_1}{\partial \gamma_{1c'}^*}\frac{\partial \gamma_{1c'}^*}{\partial t} + \frac{\partial g_2}{\partial \gamma_{2c}}\frac{\partial \gamma_{2c}}{\partial t} + \frac{\partial g_2}{\partial \gamma_{2c}^*}\frac{\partial \gamma_{2c}^*}{\partial t} + \frac{\partial g_2}{\partial \gamma_{2c'}}\frac{\partial \gamma_{2c'}}{\partial t} + \frac{\partial g_2}{\partial \gamma_{2c'}^*}\frac{\partial \gamma_{2c'}^*}{\partial t} = 0. \quad (18)$$

As before, using (6) and simplifying, we get

$$\frac{\partial g_1}{\partial \gamma_{1c}}\gamma_{2c} + \frac{\partial g_1}{\partial \gamma_{1c'}}\gamma_{2c'} = \frac{\partial g_2}{\partial \gamma_{2c}}\gamma_{1c}^* + \frac{\partial g_2}{\partial \gamma_{2c'}^*}\gamma_{1c'}^*, \quad (19)$$

$$\frac{\partial g_1}{\partial \gamma_{1c}^*}\gamma_{2c}^* + \frac{\partial g_1}{\partial \gamma_{1c'}^*}\gamma_{2c'}^* = \frac{\partial g_2}{\partial \gamma_{2c}}\gamma_{1c} + \frac{\partial g_2}{\partial \gamma_{2c'}}\gamma_{1c'}. \quad (20)$$

Since $\gamma_{2c}, \gamma_{2c'}$ are arbitrary, by putting $\gamma_{2c'} = 0$, we find that $\partial g_1/\partial \gamma_{1c}$ is a linear function of $\gamma_{1c}^*, \gamma_{1c'}^*$; similarly by putting $\gamma_{2c} = 0$ we find that $\partial g_1/\partial \gamma_{1c'}$ is a linear function of $\gamma_{1c}^*, \gamma_{1c'}^*$; therefore

$$g_1(\gamma_{1c}, \gamma_{1c}^*, \gamma_{1c'}, \gamma_{1c'}^*) = a\gamma_{1c}\gamma_{1c}^* + b\gamma_{1c}\gamma_{1c'}^* + c\gamma_{1c}^*\gamma_{1c'} + d\gamma_{1c}\gamma_{1c'}^*. \quad (21)$$

Using the relations

$$\left(\frac{\partial g_1}{\partial \gamma_{1c}}\right)^* = \frac{\partial g_1}{\partial \gamma_{1c}^*}, \qquad \left(\frac{\partial g_1}{\partial \gamma_{1c'}}\right)^* = \frac{\partial g_1}{\partial \gamma_{1c'}^*},$$

we get $a^* = a$, $d^* = c$, $b^* = b$; therefore $\begin{pmatrix} a & c \\ c^* & b \end{pmatrix}$ is a 2×2 Hermitian matrix, and so

$$f_\lambda(\gamma_{\lambda c}, \gamma_{\lambda c}^*, \gamma_{\lambda c'}, \gamma_{\lambda c'}^*) = A \exp[-(\gamma_\lambda, M\gamma_\lambda)], \quad (22)$$

where

$$(\gamma_\lambda, M\gamma_\lambda) = M_{cc}|\gamma_{1c}|^2 + M_{c'c'}|\gamma_{1c'}|^2 + M_{cc'}\gamma_{1c}^*\gamma_{1c'} + M_{cc'}^*\gamma_{1c}\gamma_{1c'}^*. \quad (23)$$

The normalization condition gives

$$A = \det M/\pi^2 = |M|/\pi^2.$$

Therefore,

$$f_\lambda(\gamma_{\lambda c}, \gamma_{\lambda c}^*, \gamma_{\lambda c'}, \gamma_{\lambda c'}^*) = (|M|/\pi^2) \exp[-(\gamma_\lambda, M\gamma_\lambda)]. \quad (24)$$

Using (16) and (24), we get

$$P_{cc'} = \prod_\lambda \frac{|M|}{\pi^2} \exp[(-\gamma_\lambda, M\gamma_\lambda)]. \quad (25)$$

The generalization of (25) gives

$$P_{cc'c'', \cdots} = \prod_\lambda \frac{|M|}{\pi^{2m}} \exp[-(\gamma_\lambda, M\gamma_\lambda)], \quad (26)$$

where M is now a $m \times m$ Hermitian matrix corresponding to m channels.

We now show that the multichannel distribution (26) reduces to the two-channel distribution. The Hermitian matrix M can be diagonalized by a unitary transformation.

$$\text{Let} \quad \gamma_\lambda = V\gamma_\lambda', \quad (27)$$

where V is a $m \times m$ unitary matrix; then

$$(\gamma_\lambda, M\gamma_\lambda) = (V\gamma_\lambda', MV\gamma_\lambda'), = (\gamma_\lambda', V^{-1}MV\gamma_\lambda'), = (\gamma_\lambda', D\gamma_\lambda'), \quad (28)$$

where D is the diagonal matrix $V^{-1}MV$.

The two-channel distributions are obtained from (26) by integrating over all the other channels but c, c':

$$P_{cc'} = \int P_{cc'c'', \cdots} d\gamma_{\lambda c''}^r, d\gamma_{\lambda c''}^i, \cdots . \quad (29)$$

Now

$$d\gamma_{\lambda c''}^r, d\gamma_{\lambda c''}^i, \cdots = J\, d\gamma_{\lambda c''}^{r\prime}, d\gamma_{\lambda c''}^{i\prime}, \cdots , \quad (30)$$

where J is the Jacobian of transformation and can be taken outside the integral in (29), since it does not depend on the integration variables. Using (26), (28), (29), and (30), we get

$$P_{cc'} = J \int \prod_\lambda \frac{|M|}{\pi^{2m}} \times \exp[-(\gamma_\lambda', D\gamma_\lambda')]\, d\gamma_{\lambda c''}^r, d\gamma_{\lambda c''}^i, \cdots = \prod_\lambda (\text{const}) \exp[-(\gamma_\lambda', d\gamma_\lambda')], \quad (31)$$

where d is a 2×2 diagonal matrix. Equation (31) can be written as

$$P_{cc'} = \prod_\lambda (\text{const}) \exp[-(\gamma_\lambda, M'\gamma_\lambda)], \quad (32)$$

where M' is a 2×2 Hermitian matrix.

This derivation shows that we can use the two-channel distribution function to calculate various

quantities of interest instead of the multichannel distribution function.

We shall now show that

$$\overline{\gamma_\lambda \times \gamma_\lambda^*} = M^{-1} = \Sigma, \tag{33}$$

where Σ is called the covariance matrix.

It is easy to show that

$$\overline{|\gamma_{\lambda c}|^2} = \frac{1}{|M|}\frac{\partial}{\partial M_{cc}}|M|,$$

$$\overline{|\gamma_{\lambda c'}|^2} = \frac{1}{|M|}\frac{\partial}{\partial M_{c'c'}}|M|,$$

$$\overline{\gamma_{\lambda c}^*\gamma_{\lambda c'}} = \frac{1}{|M|}\frac{\partial}{\partial M_{cc'}}|M|,$$

$$\overline{\gamma_{\lambda c}\gamma_{\lambda c'}^*} = \frac{1}{|M|}\frac{\partial}{\partial M_{cc'}^*}|M|;$$

hence $\overline{\gamma_\lambda \times \gamma_\lambda^*} = \Sigma$.

III. CORRELATION COEFFICIENTS

We shall now calculate the various channel correlation coefficients and compare the results with those of the orthogonal ensemble. The channel correlation coefficients $C^*_{\gamma_c,\gamma_{c'}}$, $C_{\gamma_{c'},\gamma_c^*}$ are given by

$$C^*_{\gamma_c,\gamma_{c'}} = \overline{\gamma_{\lambda c}\gamma_{\lambda c'}^*}(|\gamma_c|^2\,|\gamma_{c'}|^2)^{-\frac{1}{2}}, \tag{34}$$

$$C^*_{\gamma_{c'},\gamma_c} = \overline{\gamma_{\lambda_c}\gamma_{\lambda c}^*}(|\gamma_c|^2\,\overline{|\gamma_{c'}|^2})^{-\frac{1}{2}}. \tag{35}$$

For two channels c, c′, we have

$$\overline{|\gamma_c|^2} = \frac{M_{c'c'}}{|M|}, \qquad \overline{|\gamma_{c'}|^2} = \frac{M_{cc}}{|M|},$$

$$\overline{\gamma_{\lambda c}^*\gamma_{\lambda c'}} = -\frac{M_{cc'}^*}{|M|}, \qquad \overline{\gamma_{\lambda c}\gamma_{\lambda c'}^*} = -\frac{M_{cc'}}{|M|}.$$

Putting in these values in (34) and (35), we get

$$C^*_{\gamma_c,\gamma_{c'}} = -\frac{M_{cc'}}{(M_{cc}M_{c'c'})^{\frac{1}{2}}}, \tag{36}$$

$$C^*_{\gamma_{c'},\gamma_c} = -\frac{M_{cc'}^*}{(M_{cc}M_{c'c'})^{\frac{1}{2}}},$$

$$= [C^*_{\gamma_c,\gamma_{c'}}]^*. \tag{37}$$

The reduced-width correlation coefficient is given by

$$C_{|\gamma_c|^2,|\gamma_{c'}|^2} = \overline{|\gamma_{\lambda c}|^2\,|\gamma_{\lambda c'}|^2} - \overline{|\gamma_{\lambda c}|^2}\;\overline{|\gamma_{\lambda c'}|^2}$$
$$\times [\{\overline{|\gamma_{\lambda c}|^4} - (\overline{|\gamma_{\lambda c}|^2})^2\}\{\overline{|\gamma_{\lambda c'}|^4} - (\overline{|\gamma_{\lambda c'}|^2})^2\}]^{-\frac{1}{2}}. \tag{38}$$

Putting in the various values we get

$$C_{|\gamma_c|^2,|\gamma_{c'}|^2} = \frac{|M_{cc'}|^2}{M_{cc}M_{c'c'}} \tag{39}$$

$$= |C^*_{\gamma_c,\gamma_{c'}}|^2 > 0. \tag{40}$$

Therefore the width correlation is always positive, a result found also for the orthogonal ensemble.[1]

ACKNOWLEDGMENT

The author would like to thank Dr. C. E. Porter for a number of helpful suggestions and criticisms.

Invariant Operators of the Unitary Unimodular Group in n Dimensions*

ABRAHAM KLEIN†
University of Pennsylvania, Philadelphia, Pennsylvania
(Received 27 March 1963)

An elementary derivation is given of Biedenharn's construction of a complete set of independent invariants for the group $SU(n)$. The basic tool is the mapping of the adjoint representation onto the linear space of generators in the defining representation. The trace of any algebraic function of the matrix thus associated is seen to constitute an invariant of the adjoint representation and yields by substitution an invariant operator. The independent invariants are recognized by their isomorphy to the invariant forms under the permutation group.

IN a recent paper, Biedenharn[1] has given an explicit construction of the independent invariants for the unimodular unitary group in n dimensions, $SU(n)$, to serve as the starting point for the laudable program of determining the irreducible representations of these groups in a form suitable for physical application. The purpose of this note is to report a derivation of Biedenharn's main results—the form of the invariants and their relation to a symmetric vector coupling coefficient—which appears to us both completely elementary and bearing a more integral relationship to the fundamental properties of these groups.[2]

We work with the simplest possible notation. The generators X_i, $i = 1, \cdots, n^2 - 1$, of $SU(n)$, which satisfy

$$[X_i, X_j] = c_{ijk}X_k, \tag{1}$$

are taken as an Hermitian set. Throughout most of this section we mean Eq. (1) to refer to a choice of $n \times n$ traceless matrices in the *fundamental* representation. Since in this representation the X_i plus the unit matrix span the space of Hermitian $n \times n$ matrices, we also have for the anticommutator

$$\tfrac{1}{2}\{X_i, X_j\} = c\delta_{ij}1 + d_{ijk}X_k. \tag{2}$$

Whereas the c_{ijk} are antisymmetric in all indices, the d_{ijk}, obviously symmetric in i and j, can be shown from the properties of the trace to be symmetric in all pairs of indices. Similarly it can be seen that the first term on the right-hand side of (2) is necessarily a multiple of the Kronecker delta if we normalize the X_i so that the square of each has the same trace,

$$\text{tr } X_i^2 = m, \qquad \text{for each } i. \tag{3}$$

We introduce, in addition to the fundamental representation of the generators, an $(n^2 - 1)$-dimensional vector space spanned by vectors $\mathbf{\Phi} = (\Phi_1 \cdots \Phi_N)$, $N = n^2 - 1$, transforming under the adjoint representation. With the given Hermitian choice of the X_i, we can, without loss of generality, choose the vectors $\mathbf{\Phi}$ to be real; the representation is then by orthogonal matrices $\mathfrak{U}$,

$$\begin{aligned} \Phi_i' &= \mathfrak{U}_{ij}\Phi_j, \\ \mathfrak{U}^* &= \mathfrak{U}, \qquad \tilde{\mathfrak{U}}\mathfrak{U} = 1. \end{aligned} \tag{4}$$

By definition of the adjoint representation, moreover, if U is a transformation of the defining representation, since U induces $\mathfrak{U}$, $\mathfrak{U} = \mathfrak{U}(U)$, we have

$$UX_iU^{-1} = \mathfrak{U}_{ij}X_j, \tag{5}$$

since this is the integrated form of Eq. (1).

With these preliminaries out of the way we outline the procedure to be followed:

(i) We note that (4) defines necessarily a subgroup of the orthogonal group in $n^2 - 1$ dimensions. We show that this subgroup is characterized by $n - 1$ algebraically independent algebraic invariants constructed from a single arbitrary vector $\mathbf{\Phi}$, and that these are forms of degree 2 to n in $\mathbf{\Phi}$. This is done by explicit construction of the invariants in a form isomorphic to the corresponding independent invariants of the permutation group of n objects. It is also seen that $n - 1$ is the maximum number of independent invariants, this being the rank of the original group.

(ii) It is then observed that every such algebraic

* Supported in part by the U. S. Atomic Energy Commission.
† Alfred P. Sloan Foundation Fellow.
[1] L. C. Biedenharn, J. Math. Phys. 4, 436 (1963).
[2] It has been called to the writer's attention that U. Fano suggested a similar proof (Eastern Theoretical Physics Conference, 1962) which is incorporated in subsequent work by Biedenharn. The invariants derived by V. Bargmann and M. Moshinsky, [Nucl. Phys. 18, 697 (1960); 23, 177 (1961), *et seq.*] also employ essentially the same technique.

invariant yields, by the substitution $\Phi_i \rightarrow X_i$, an invariant operator and conversely. The form of these invariants is prescribed in terms of the numerical coefficients d_{ijk} of Eq. (2).

(iii) We pick up incidentally the proof that the d_{ijk} constitute symmetric vector coupling coefficients.

We turn to the demonstrations. The fundamental tool is the Hermitian $n \times n$ matrix

$$X_\Phi = \Phi_i X_i, \tag{6}$$

which defines a mapping of the linear vector space Φ onto the linear vector space of matrix generators X_i. According to (4) and (5), we have

$$U^{-1} X_\Phi U = \Phi_i' X_i, \tag{7}$$

where

$$\Phi' = \mathfrak{U}\Phi. \tag{8}$$

We choose U such that X_Φ is brought to diagonal form. It is then a linear combination of the subset of $n - 1$ mutually commuting members of the set X_i which we denote by Y_a, $a = 1 \cdots n - 1$. Thus, under a suitable U,

$$X_\Phi \rightarrow \Phi_a' Y_a. \tag{9}$$

This shows that under the transformations $\mathfrak{U}$, an arbitrary Φ will, in any coordinate system, have no fewer than $n - 1$ components. As we shall see, this puts an upper bound of $n - 1$ on the number of independent invariants.

The most important remark of this paper is that as a consequence of (7) and (8), the trace of any power of X_Φ is a form in Φ, invariant under the representation $\mathfrak{U}$,

$$\begin{aligned} \operatorname{tr} (X_\Phi)^p &= s_{i_1 \cdots i_p} \Phi_{i_1} \cdots \Phi_{i_p} \\ &= s_{i_1 \cdots i_p} \Phi'_{i_1} \cdots \Phi'_{i_p}, \end{aligned} \tag{10}$$

or $s_{i_1 \cdots i_p}$, which will be specified in detail later, is a symmetric numerical tensor invariant under the transformation $\mathfrak{U}$,

$$s_{i_1 \cdots i_p} = \mathfrak{U}_{i_1 j_1} \cdots \mathfrak{U}_{i_p j_p} s_{j_1 \cdots j_p}. \tag{11}$$

We now show that for $p = 2, \cdots n$ these forms are algebraically independent (for $p = 1$ the form is zero), whereas for $p > n$, they are dependent upon the previous forms.

For this purpose we change the basis of the Y_a to a set of idempotents P_α, $\alpha = 1 \cdots n$,

$$P_\alpha P_\beta = \delta_{\alpha\beta} P_\alpha, \qquad \operatorname{tr} P_\alpha = \operatorname{tr} (P_\alpha)^p = 1. \tag{12}$$

We may rewrite (9) as

$$\Phi_a' Y_a = Z_a P_a \tag{13}$$

with the subsidiary condition

$$\sum_\alpha Z_\alpha = 0. \tag{14}$$

We now have immediately from (12)

$$J_p \equiv \operatorname{tr} (X_\Phi)^p = \sum_{\alpha=1}^{n} (Z_\alpha)^p. \tag{15}$$

But for $p = 1, \cdots n$, these forms constitute an algebraic basic for the invariants of the symmetric group of n objects,[3] and since the form for $p = 1$ is identically zero, we have found $n - 1$ algebraically independent invariants for the group of transformations $\mathfrak{U}$.

That there are no further independent invariants can be seen as follows: From the form (10) we conclude that the sets

$$\Psi_i^{(p)} = s_{ii_2 \cdots i_p} \Phi_{i_2} \cdots \Phi_{i_p}, \qquad \Psi_i^{(1)} = \Phi_i \tag{16}$$

are vectors under $\mathfrak{U}$. But any n of these vectors are linearly related as can be seen by remembering that there is a coordinate system in which a general Φ has only $n - 1$ nonvanishing components. From the linear dependence of any n of the vectors $\Psi^{(p)}$ follows the algebraic dependence of any n invariants.

The form taken by the independent invariants (15) when expressed in terms of Φ is readily obtained by the help of (2). It turns out that the J_p are not a convenient set. Instead we define the set

$$\begin{aligned} I_2 &= \delta_{i_1 i_2} \Phi_{i_1} \Phi_{i_2}, \\ I_3 &= d_{i_1 i_2 i_3} \Phi_{i_1} \Phi_{i_2} \Phi_{i_3}, \\ &\vdots \\ I_p &= d_{i_1 i_2 j_1} d_{j_1 j_2 i_3} \cdots d_{i_{p-4} i_{p-3} i_{p-2}} \\ &\qquad \times d_{i_{p-3} i_{p-2} i_{p-1}} d_{i_{p-2} i_{p-1} i_p} \Phi_{i_1} \cdots \Phi_{i_p}. \end{aligned} \tag{17}$$

We then find that for $q \leq n$ the J_q are for $p > 3$ simple algebraic functions of the I_p including J_p, whereas according to the theorem, for $p > n$, no new I_p intervene.

The role of the d_{ijk} as symmetric vector coupling coefficients is manifest from the second of Eqs. (17).

Our demonstration is now completed by noticing that according to Eqs. (7), (8), and (11), we obtain an invariant under the group for every I_p by substituting therein $\Phi_i \rightarrow X_i$. For $p > 3$, the use of the symmetrized form of the invariant numerical tensors is implied.

ACKNOWLEDGMENT

The author is indebted to Professor S. Bludman and Professor B. W. Lee for stimulating discussions.

[3] H. Weyl, *The Classical Groups* (Princeton University Press, Princeton, New Jersey, 1946), Chap. II.

Energy-Dependent Boltzmann Equation in Plane Geometry

ROMAN J. BEDNARZ AND JANUSZ R. MIKA
Institute of Nuclear Research, Warsaw, Poland
(Received 9 May 1963)

The paper presents the general procedure of solving the energy-dependent Boltzmann equation in plane geometry. The particular solutions are found and then it is proved that the general solutions can be formed by superposition of particular solutions. As an illustration, the fully degenerate kernel is considered in detail and a solution in a closed form is obtained.

1. INTRODUCTION

THIS paper is devoted to the presentation of the method of solving the energy-dependent Boltzmann equation in plane geometry.

The method was introduced to the transport theory by Case in the case of one-velocity Boltzmann equation with isotropic scattering in plane geometry.[1] A number of works have been devoted to the extension of the method to anisotropic scattering[2,3] and multigroup approximation[4,5] as well as to the thermalization in heavy gas.[6]

We deal with the Boltzmann equation with the kernel dependent on the angle and energy, the total mean free path being a function of energy. It is shown that in general case, the particular solutions to the Boltzmann equation can be found and the spectrum of the permissible values of the separational constant is determined. The general solution is then formed as a superposition of particular solutions. It is proved that such an extension is uniquely determined from the boundary condition.

The possible applications of the method for degenerate kernels and in multigroup approximation is discussed. As an illustration, the case when the scattering function can be written in the form

$$K(\mu, E; \mu', E') = g(E)h(E'), \tag{1.1}$$

is considered in detail. It is shown that the problem is reduced to the equation which is essentially the same as in the one-velocity case.

2. PARTICULAR SOLUTIONS TO THE BOLTZMANN EQUATION

We consider the Boltzmann equation of the following form:

$$l(E)\mu \frac{\partial \tilde{\psi}}{\partial x} + \tilde{\psi}(x, \mu, E) = \int_{E_0}^{E_1} dE' \times \int_{-1}^{+1} d\mu' \tilde{K}(\mu, E; \mu', E')\tilde{\psi}(x, \mu', E'), \tag{2.1}$$

where $\psi(x, \mu, E)$ is the collision density of neutrons, $l(E)$ is the total mean-free path of neutrons, and μ is the cosine of the angle between the neutron velocity and the x axis. The kernel $\tilde{K}(\mu, E; \mu', E')$ gives the number of secondary neutrons per one collision. E_0 and E_1 are the lower and the upper limits of the energy interval, respectively. The specification of those limits will depend on the particular problem considered.

It is convenient to introduce new variables

$$\begin{aligned} v &= \mu l(E), \\ v' &= \mu' l(E'), \end{aligned} \tag{2.2}$$

and with them Eq. (2.1) can be written in the form

$$v \frac{\partial \psi}{\partial x} + \psi(x, v, E) = \int_{E_0}^{E_1} dE' \times \int_{-l(E')}^{l(E')} dv' K(v, E; v', E')\psi(x, v', E'), \tag{2.3}$$

where $\psi(x, v, E)$ and $K(v, E; v', E')$ are in an obvious way connected with $\tilde{\psi}(x, \mu, E)$ and $\tilde{K}(\mu, E; \mu', E')$.

Introduce the following notation:

$$E \in M(v) \quad \text{if} \quad l(E) \geq |v| \tag{2.4}$$

for

$$v \in (-1, +1).$$

For v not belonging to the interval $(-1, +1)$ on the real axis, the set $M(v)$ is empty.

With the definition (2.4), Eq. (2.3) can be written in the alternative form

$$v \frac{\partial \psi}{\partial x} + \psi(x, v, E) = \int_{-1}^{+1} dv' \times \int_{M(v')} dE' K(v, E; v', E')\psi(x, v', E'), \tag{2.5}$$

where $E \in M(v)$.

[1] K. M. Case, Ann. Phys. (N. Y.) 9, 1 (1960).
[2] R. Żelazny, A. Kuszell, and J. Mika, Ann. Phys. (N. Y.) 16, 69 (1961).
[3] J. Mika, Nucl. Sci. Eng. 11, 415 (1961).
[4] R. Żelazny and A. Kuszell, Ann. Phys. (N. Y.) 16, 81 (1961).
[5] R. Żelazny and A. Kuszell, Phys. Fast and Intermediate Reactors 55, (1962).
[6] R. Żelazny, Proc. Conf. Neutr. Therm. Brookhaven, National Laboratories, Upton, Long Island, New York, 1962.

We shall separate the space variable by an *ansatz*:

$$\psi(x, v, E) = e^{-x/t}\,\phi(t, v, E), \tag{2.6}$$

where the separational constant t is an arbitrary complex number to be determined.

Putting (2.6) into Eq. (2.5), we get

$$(t - v)\phi(t, v, E) = t\int_{-1}^{+1} dv' \times \int_{M(v')} dE' K(v, E; v', E')\phi(t, v', E'). \tag{2.7}$$

Denote

$$\int_{-1}^{+1} dv' \int dE' K(v, E; v', E')\phi(t, v', E') = H(t, v, E). \tag{2.8}$$

Then, solving Eq. (2.7) for $\phi(t, v, E)$, we have

$$\phi(t, v, E) = \frac{tH(t, v, E)}{t - v} + \lambda(t, E)\delta(t - v), \tag{2.9}$$

where $\lambda(t, E)$ is an arbitrary function which is connected with $H(t, v, E)$. Multiplying Eq. (29) by $K(v, E; v', E')$, integrating over all possible values of v' and E', and using the relation (2.8), we obtain

$$H(t, v, E) + t\int_{-1}^{+1} \frac{dv'}{v' - t} \times \int_{M(v')} dE' K(v, E; v', E')H(t, v', E') = \int_{M(t)} dE' K(v, E; t, E')\lambda(t, E'). \tag{2.10}$$

The relationship between $H(t, v, E)$ and $\lambda(t, E)$ given by Eq. (2.10) plays a substantial role in the theory. It will be seen that it enables us to reduce the problem of finding the function $H(t, v, E)$ to solving the integral equation of the singular type. A similar procedure has been used by Żelazny.[6]

In further considerations, the important role will be played by the integral operator

$$\Omega_{v,E}(z)f(z, v, E) = f(z, v, E) + z\int_{-1}^{+1} \frac{dv'}{v' - z} \times \int_{M(v')} dE' K(v, E; v', E')f(z, v', E'). \tag{2.11}$$

In the following we shall assume that $K(v, E; v', E')$ satisfies the H^* condition with respect to v and v'.

The function $\varphi(v)$ is said to satisfy the H^* condition in the interval (α, β) if, for v_1 and v_2 belonging to any closed interval contained in the open interval (α, β),

$$|\varphi(v_1) - \varphi(v_2)| \leq A\,|v_1 - v_2|^{\gamma}, \tag{2.12}$$

where A and γ are some positive constants. Moreover, near the ends α and β, $\varphi(v)$ behaves as

$$\varphi(v) = \varphi^*(v)/|v - c|^{\delta}, \qquad 0 \leq \delta < 1, \tag{2.13}$$

where $\varphi^*(v)$ satisfies the condition (2.12) in the closed interval (α, β), and c stands for α and β, respectively.

The sectionally holomorphic function is a function analytic in the cut plane. The boundary values on the cut from above and from below exist and satisfy the H^* condition there.

Now it is seen that the operator $\Omega_{v,E}(z)$, operating on any function $f(z, v, E)$ sectionally holomorphic in z and satisfying the H^* condition in v, will give back the sectionally holomorphic function. It will be said that $\Omega_{v,E}(z)$ is a sectionally holomorphic operator with a cut on the real axis $(-1, +1)$.

Using the Plemelj's formulas,[1] the boundary values of $\Omega_{v,E}(z)$ can be calculated:

$$\Omega^{\pm}_{v,E}(s)f^{\pm}(p, v, E) = P\Omega_{v,E}(s)f^{\pm}(p, v, E) \pm i\pi s w_{v,E}(s, v)f^{\pm}(p, v, E), \tag{2.14}$$

where

$$P\Omega_{v,E}(s)f^{\pm}(p, v, E) = f^{\pm}(p, v, E) + s\int_{-1}^{+1} \frac{dv'}{v' - s} \times \int_{M(v')} dE' K(v, E; v', E')f^{\pm}(p, v', E'). \tag{2.15}$$

Here the integration over v' is understood in the Cauchy principal-value sense.

$$w_{v,E}(s, v)f^{\pm}(p, v, E) = \int_{-1}^{+1} dv'\,\delta(v' - s) \times \int_{M(v')} dE' K(v, E; v', E')f(p, v', E') = \int_{M(s)} dE' K(v, E; s, E')f(p, s, E'). \tag{2.16}$$

In particular, p can be put equal to s in Eqs. (2.14), (2.15), and (2.16).

The right-hand side of Eq. (2.10) vanishes for t which do not belong to the interval $(-1, +1)$ on the real axis. Therefore, one gets the discrete solutions to the Boltzmann equation from the eigenfunctions of the operator $\Omega_{v,E}(z)$ for z outside the cut. Denote the numerable set of eigenvalues of $\Omega_{v,E}(z)$ by L_i. Then we have

$$\Omega_{v,E}(L_i)H(L_i, v, E) = 0, \qquad i = 1, 2, 3, \cdots. \tag{2.17}$$

The corresponding solutions of the Boltzmann equa-

tion will be

$$\phi(L_i, v, E) = L_i H(L_i, v, E)/(L_i - v). \quad (2.18)$$

It may happen that for some $L_k \in (-1, +1)$ there are eigenfunctions of $\Omega^{*}_{v,E}(L_k)$. Denote those eigenfunctions by $H^{*}(L_k, v, E)$. The corresponding solutions to the Boltzmann equation are

$$\phi^{*}(L_k, v, E) = L_k H^{*}(L_k, v, E)/(L_k - v). \quad (2.19)$$

As the integral equation for $\phi(t, v, E)$ (2.7) involves only the real functions, it is seen that $\phi^{*}(L_k, v, E)$ and $H^{*}(L_k, v, E)$ have to be real functions. Then it is immediately seen that

$$H^{+}(L_k, v, E) = H^{-}(L_k, v, E) = H(L_k, v, E), \quad (2.20)$$

and $H(L_k, v, E)$ is the simultaneous eigenfunction of the operators $P\Omega_{v,E}(L_k)$ and $w_{v,E}(L_k)$, that is,

$$H(L_k, v, E) + L_k \int_{-1}^{+1} \frac{dv'}{v' - L_k} \times \int_{M(v')} dE' K(v, E; v', E') H(L_k, v, E') = 0, \quad (2.21)$$

and

$$\int_{M(L_k)} dE' K(v, E; L_k, E') H(L_k, L_k, E') = 0. \quad (2.22)$$

For any representation of the kernel $K(v, E; v', E')$ (finite or infinite), all elements of $H(L_k, L_k, E)$ vanish by virtue of Eq. (2.22) and hence,

$$H(L_k, L_k, E) = 0, \quad (2.23)$$

which means that the corresponding solution of the Boltzmann equation,

$$\phi(L_k, vE) = L_k H(L_k, v, E)/(L_k - v), \quad (2.24)$$

is a bounded function.

The continuous spectrum of the solutions of the Boltzmann equation consist of all $s \in (-1, +1)$. The corresponding solutions are

$$\phi(s, v, E) = sH(s, v, E)/(s - v) + \lambda(s, E)\delta(s - v), \quad (2.25)$$

with $H(s, v, E)$ and $\lambda(s, E)$ related by Eq. (2.10).

Finally we conclude that the spectrum of the Boltzmann equation consists of all $s \in (-1, +1)$, and discrete values $L_i \notin (-1, +1)$ and $L_k \in (-1, +1)$.

3. ORTHOGONALITY RELATIONS

Consider the equation

$$\left(1 - \frac{v}{t}\right)\bar{\phi}(t', v, E) = \int_{-1}^{+1} dv' \times \int_{M(v')} dE' K(v', E'; v, E)\bar{\phi}(t', v', E'). \quad (3.1)$$

It will be proved that the solutions of Eq. (3.1) are orthogonal to the solutions of Eq. (2.7).

Rewrite Eq. (2.7) in the form

$$\left(1 - \frac{v}{t}\right)\phi(t, v, E) = \int_{-1}^{+1} dv' \times \int_{M(v')} dE' K(v, E; v', E')\phi(t, v, E). \quad (3.2)$$

Multiplying Eq. (3.1) by $\phi(t, v, E)$ and Eq. (3.2) by $\bar{\phi}(t', v, E)$, respectively, intergrating over all v and E and subtracting, we get the orthogonality relation

$$\left(\frac{1}{t} - \frac{1}{t'}\right) \int_{-1}^{+1} dv \int_{M(v)} dE v\phi(t, v, E) \times \bar{\phi}(t', v, E) = 0. \quad (3.3)$$

The solution of Eq. (3.1) can be obtained in the same way as that of Eq. (2.7). We have

$$\bar{\phi}(t, v, E) = t\bar{H}(t, v, E)/(t - v) + \bar{\lambda}(t, E)\delta(t - v), \quad (3.4)$$

where

$$\bar{H}(t, v, E) = \int_{-1}^{+1} dv' \times \int_{M(v')} dE' K(v', E'; v, E)\bar{\phi}(t, v', E'), \quad (3.5)$$

and $\bar{\lambda}(t, E)$ satisfies the relation

$$\bar{H}(t, v, E) + t \int_{-1}^{+1} \frac{dv'}{v' - t} \times \int_{M(v')} dE' K(v', E'; v, E)\bar{H}(t, v', E') = \int_{M(t)} dE' K(t, E'; v, E)\bar{\lambda}(t, E'). \quad (3.6)$$

Similarly we can introduce an operator

$$\bar{\Omega}_{v,E}(z)f(z, v, E) = f(z, v, E) + z \int_{-1}^{+1} \frac{dv'}{v' - z} \times \int_{M(v')} dE' K(v', E'; v, E)f(z, v', E'). \quad (3.7)$$

It is evident that the spectra of the operators $\Omega_{v,E}(z)$ and $\bar{\Omega}_{v,E}(z)$ are identical. Thus the permissible values of t for Eq. (3.1) are the same as those for Eq. (2.7)

From now on we will use normalized discrete solutions, that is, such that

$$\int_{-1}^{+1} dv \int_{M(v)} dE v\phi(t, v, E)\bar{\phi}(t, v, E) = 1. \quad (3.8)$$

It may be noted that if the kernel $K(v, E; v', E')$ is symmetric with respect to v and E, the functions $\phi(t, v, E)$ and $\bar{\phi}(t, v, E)$ are identical.

4. COMPLETENESS THEOREM

So far we have obtained formally the particular solutions of Eq. (2.1) of the form (2.6). Moreover, we have found the spectrum of permissible values of the parameter t.

We will prove that the general solution of Eq. (2.1) can be expressed by the superposition of particular solutions (2.6).

Assume that for some $x = x_0$ the energy–angular distribution of neutrons is given. We can always choose such a coordinate system that $x_0 = 0$. We will express the boundary neutron distribution in variables (2.2). The problem can be formulated as follows: Given the function $\psi(v, E)$ satisfying the H^* condition in v, find the functions $H(s, v, E)$ and $\lambda(s, E)$, and the coefficients A_i and A_k so that the expansion

$$\psi(v, E) = \int_{-1}^{+1} \frac{sH(s, v, E)}{s - v} ds + \lambda(v, E) + \sum_i A_i \frac{L_i H(L_i, v, E)}{L_i - v} + \sum_k A_k \frac{L_k H(L_k, v, E)}{L_k - v} \quad (4.1)$$

is unique.

Multiply Eq. (4.1) by $K(p, E; v, E')$ and intergrate over $M(v)$. Using Eq. (2.10), we get

$$\psi(v; p, E) = \int_{M(v)} dE' K(p, E; v, E') \times \int_{-1}^{+1} \frac{sH(s, v, E')\, ds}{s - v} + H(v, p, E) + v \int_{-1}^{+1} \frac{dv'}{v' - v} \times \int_{M(v')} dE' K(p, E; v, E') H(v, v', E'), \quad (4.2)$$

where

$$\psi(v; p, E) = \int_{M(v)} dE' K(p, E; v, E') \times \left[\psi(v, E') - \sum_i A_i \frac{L_i H(L_i, v, E')}{L_i - v} - \sum_k A_k \frac{L_k H(L_k, v, E')}{L_k - v} \right]. \quad (4.3)$$

Introduce a sectionally holomorphic function

$$N(z; v, E) = \frac{1}{2\pi i} \int_{-1}^{+1} \frac{sH(s, v, E)\, ds}{s - z}. \quad (4.4)$$

By the Plemelj formulas, we have

$$N^*(p; v, E) = \frac{1}{2\pi i} \int_{-1}^{+1} \frac{sH(s, v, E)\, ds}{s - z} \pm \tfrac{1}{2} pH(p, v, E) \quad (4.5)$$

for $p \in (-1, +1)$.

From Eq. (4.5) we get

$$\int_{-1}^{+1} \frac{sH(s, v, E)\, ds}{s - v} = i\pi[N^+(v; v, E) + N^-(v; v, E)],$$

$$vH(v, p, E) = N^+(v; p, E) - N^-(v; p, E). \quad (4.6)$$

Using Eqs. (2.14) and (4.6), we can rewrite Eq. (4.2) in the form

$$\Omega^+_{p,E}(v)N^+(v; p, E) - \Omega^-_{p,E}N^-(v; p, E) = v\psi(v; p, E). \quad (4.7)$$

As the operator $\Omega_{p,E}(z)$ and the function $N(z; p, E)$ are sectionally holomorphic in the cut plane, the function

$$F(z; p, E) = \Omega_{p,E}(z)N(z; p, E) \quad (4.8)$$

is sectionally holomorphic in z with the boundary values satisfying the condition

$$F^+(v; p, E) - F^-(v; p, E) = v\psi(v; p, E). \quad (4.9)$$

From this, using Plemelj's formulas, we get

$$F(z; p, E) = \frac{1}{2\pi i} \int_{-1}^{+1} \frac{v\psi(v; p, E)\, dv}{v - z}. \quad (4.10)$$

Thus we see that $F(z; p, E)$ is actually the holomorphic function of z in the cut plane and does not vanish at any point besides the point at infinity.

Equation (4.4) gives the relationship between the sectionally holomorphic function $N(z; v\ E)$ and the density $H(s, v, E)$. Knowing $H(s, v, E)$ we can uniquely determine $N(z; v, E)$, and vice versa. Our theorem will be proved if we show that the function $N(z; v, E)$ can be uniquely determined from the equation

$$\Omega_{v,E}(z)N(z; v, E) = F(z; v, E), \quad (4.11)$$

with $F(z; v, E)$ given by Eq. (4.10). The condition of the solubility of the integral equation (4.11) is that the inhomogeneous term be orthogonal to the eigenfunctions of the adjoint homogeneous equation

$$Q\ ; p, E) + \frac{z}{p - z} \int_{-1}^{+1} dv' \int_{M(v')} dE' K(v', E'; p, E) \times Q(z; p, E') = 0. \quad (4.12)$$

By multiplying both sides of this equation by $p - z$ we see that it is identical with Eq. (3.1). Therefore the eigenfunctions of Eq. (4.14) are all discrete solutions of Eq. (3.1), namely, the functions $\bar{\phi}(L_i, p, E)$ and $\bar{\phi}(L_k, p, E.)$

Finally we have shown that the solution of Eq. (4.12) exists if and only if $F(z; p, E)$ is orthogonal

to all discrete solutions of Eq. (3.1), that is,

$$\int_{-1}^{+1} dp \int_{M(p)} dEpF(L_i; p, E)\bar{\phi}(L_i; p, E) = 0,$$
$$\int_{+1}^{+1} dp \int_{M(p)} dEpF(L_k; p, E)\bar{\phi}(L_k; p, E) = 0. \tag{4.13}$$

We see that Eqs. (4.15) give exactly as many conditions as it is necessary to evaluate the constants A_i and A_k. In fact, using solutions normalized according to Eq. (3.8), we have from (4.13),

$$A_i = \int_{-1}^{+1} dv \int_{M(v)} dEv\psi(v, E)\bar{\phi}(L_i, v, E),$$
$$A_k = \int_{-1}^{+1} dv \int_{M(v)} dEv\psi(v, E)\phi(L_k, v, E). \tag{4.14}$$

The same result could have been obtained by utilizing the orthogonality properties (3.3).

Finally we have proved that all functions and coefficients in the expansion (4.1) can be uniquely determined. The function $N(z; p, E)$ can be obtained by solving the integral equation (4.11) and $H(s; p, E)$ from the second equation (4.6). The function $\lambda(v, E)$ can be evaluated directly from Eq. (4.1).

The proof of existence of the general solution to Eq. (2.1) given in this chapter has a constructive character only when the boundary value of the neutron distribution is given for a total interval $\mu \in (-1, +1)$.

In practice one often meets problems when the neutron distribution for some value of x is given only in half-range interval in μ. The typical example of such problems is the Milne problem.

Consider a semi-infinite medium occupying the region $x > 0$. The condition at infinity states that the neutron distribution increases not faster than e^{-x/L_m}, where L_m is the largest negative value of L_i. From this condition we get immediately

$$H(s, v, E) = 0, \quad s < 0,$$
$$H(L_k, v, E) = 0, \quad L_k < 0, \quad (L_i \neq L_m). \tag{4.15}$$
$$H(L_i, v, E) = 0, \quad L_i < 0,$$

The procedure employed previously can be used also in this case. As the result one gets the equation similar to Eq. (4.2) except that the integration over s involves now the interval (0, 1). We introduce a sectionally holomorphic function

$$\tilde{N}(z; v, E) = \frac{1}{2\pi i} \int_0^1 \frac{sH(s, v, E)\, ds}{s - z}. \tag{4.16}$$

This leads formally to the equation identical to Eq. (4.7). However, now the function $\tilde{N}(z; v, E)$ is holomorphic in the whole plane with a cut (0, 1), whereas $\Omega(z)$ is a sectionally holomorphic operator with a cut $(-1, +1)$.

This fact makes impossible a simple construction of the function $\tilde{N}(z; p, E)$. In such a case the approximate methods have to be employed. They will be considered in the next chapter.

5. DEGENERATE KERNELS AND MULTIGROUP APPROXIMATION

The general procedure developed in the previous chapters can be hardly used in any practical calculations. However, in practice one usually has to do with the degenerate kernels. This means that

$$K(v, E; v', E') = \sum_{KLmn} g_k(E)h_l(E')P_m(v)Q_n(v'), \tag{5.1}$$

where the summation runs over finite number of terms.

In such a case the function $H(t, v, E)$ can be represented by a vector consisting of the finite number of components and the integral operator $\Omega_{v,E}(z)$ by a finite matrix. The theory for the operators having finite matrix representation does not formally differ from the general one given in previous chapter, so we will not repeat the considerations.

It is worthwhile to note, however, that when the mean free path of neutrons depends arbitrarily on energy, one cannot obtain the recurrence relations between the components of the vector representing $H(t, v, E)$. As the result one has to solve the system of singular equations.

The situation becomes much simpler if the mean free path can be expended into the same functions as the kernel. This permits us to obtain the recurrence relations and reduce the problem to the single singular integral equation which can be dealt with by the standard methods.[7]

Another possible application of the method is the multigroup approximation.

It is well known that there are two assumptions upon which the multigroup approximation is based.

Divide the total energy interval into N subintervals. The first assumption is that the total mean free path is constant in each subinterval:

$$l(E) = l_n, \qquad E \in (n), \tag{5.2}$$

where (n) denotes the nth subinterval. Second, the integral of the kernel over the nth interval is a step function of E', that is,

$$\int_{(n)} dEK(\mu, E; \mu', E') = f_{nm}(\mu, \mu') \quad E' \in (m). \tag{5.3}$$

[7] N. J. Muskhelishwili, *Singular Integral Equations* (P. Nordhoff Ltd., Groningen, The Netherlands, 1952).

Denoting

$$\psi_n(x, \mu) = \int_{(n)} \psi(x, \mu, E)\, dE, \tag{5.4}$$

and taking into account the two assumptions (5.2) and (5.3), we can easily get from Eq. (2.1) the following system of multigroup equations:

$$l_n\mu \frac{\partial \psi_n}{\partial x} + \psi_n(x, \mu) = \sum_{m-1}^{N} \int_{-1}^{+1} f_{nm}(\mu, \mu')\psi_m(x, \mu'). \tag{5.5}$$

As previously, we will assume that $l_{\max} = 1$. Introduce new variables:

$$\begin{aligned} v &= \mu l_n, \\ v' &= \mu' l_m. \end{aligned} \tag{5.6}$$

We will say that

$$n \in M(v) \quad \text{if} \quad l_n \leq |v|. \tag{5.7}$$

If v does not belong to the interval $(-1, +1)$ on the real axis, the set $M(v)$ is empty.

Using (5.6) and (5.7) we can rewrite Eq. (5.5) in the form

$$v \frac{\partial \psi_n}{\partial x} + \psi_n(x. v) = \int_{-1}^{+1} dv' \sum_{m \in M(v')} f_{nm}(v, v')\psi_m(x, v'). \tag{5.8}$$

The further procedure is obviously identical with that used in the general case. The only difference is that the intergration over energy is now substituted by the summation over energy groups. All the theorems proved in previous chapters are valid also in this case.

The solutions to the multigroup equations for isotropic scattering were given also by Żelazny and Kuszell.[5] It seems, however, that this method would lead to the less complicated results.

6. ISOTROPIC DEGENERATE KERNEL

By way of illustration, let us consider the simple case when the kernel in Eq. (2.3) can be written as

$$K(v, E; v', E') = g(E)h(E'). \tag{6.1}$$

The corresponding Boltzmann equation in variables v and v' is of the form

$$v \frac{\partial \psi}{\partial x} + \psi(x, v, E) = g(E) \int_{-1}^{+1} dv' \int_{M(v')} dE' h(E')\psi(x, v', E'), \tag{6.2}$$

Such an equation is valid, for example, for infinitely heavy fissionable nuclei.

Using the separational *ansatz* (2.6) we get

$$(t - v)\phi(t, v, E) = tg(E) \int_{-1}^{+1} dv \int_{M(v')} h(E')\phi(t, v', E'). \tag{6.3}$$

Introduce

$$H(t) = \int_{-1}^{+1} dv \int_{M(v')} h(E')\phi(t, v', E')\, dE'; \tag{6.4}$$

then

$$\phi(t, v, E) = \frac{tg(E)H(t)}{t - v} + \lambda(t, E)\delta(t - v). \tag{6.5}$$

From the definition of $H(t)$ we get the following relation:

$$H(t)\left[1 + t \int_{-1}^{+1} \frac{dv'}{v' - t} \int_{M(v')} g(E')h(E')\, dE'\right] = \int_{M(t)} dE'\lambda(t, E')h(E'). \tag{6.6}$$

Introduce the function

$$c(v) = \int_{M(v)} g(E)h(E)\, dE. \tag{6.7}$$

Then Eq. (6.6) can be written as

$$H(t)\left[1 + t \int_{-1}^{+1} \frac{c(v')\, dv'}{v' - t}\right] = \int_{M(t)} dE'h(E')\lambda(t, E'). \tag{6.8}$$

Introduce the sectionally holomorphic function

$$\Lambda(z) = 1 + z \int_{-1}^{+1} \frac{c(v')\, dv'}{v' - z}. \tag{6.9}$$

Obviously the discrete solutions to Eq. (6.2) exist for zeros of the function $\Lambda(z)$.

The boundary values of $\Lambda(z)$ at the cut are

$$\Lambda^{\pm}(s) = 1 + s \int_{-1}^{+1} \frac{c(v')\, dv'}{v' - s} \pm i\pi sc(s). \tag{6.10}$$

As $c(s)$ is a positive function of s vanishing only at $s = \pm 1$ it is seen that

$$\Lambda^{\pm}(s) \neq 0 \tag{6.11}$$

for all values of s, except, possibly, $s = \pm 1$, This means that the discrete solutions will exist only for values of t outside the cut or, possibly, at the points ± 1.

The number of zeros of $\Lambda(z)$ can be calculated by the argument principle. Noting that

$$\lim_{z \to \infty} \Lambda(z) = 1 - \int_{-1}^{+1} c(v')\, dv' = \text{const}, \tag{6.12}$$

we have the following expression for the number of zeros of $\Omega(z)$ outside the cut

$$2r = \frac{1}{2\pi i} \oint \log \Lambda(z)\, dz \tag{6.13}$$

where the contour of integration encircles the cut $(-1, +1)$. As $\Lambda(z)$ is an even function of z, the number of zeros is even.

Using the symmetry properties of $\Lambda^{*}(s)$ and the fact that

$$\arg \Lambda^{*}(0) = 0, \tag{6.14}$$

we can rewrite Eq. (6.13) as follows:

$$r = (1/\pi) \arg \Lambda^{+}(1). \tag{6.15}$$

Consider the integral

$$\int_{-1}^{+1} \frac{c(v')\, dv'}{v' - s} \quad \text{near} \quad s = 1.$$

If $c(1) \neq 0$ which means that the mean free path is identically equal to 1, that is, to its maximum value for some finite interval of E, then the integral is divergent and always

$$\arg \Lambda^{+}(1) = \pi.$$

Therefore $r = 1$ and there are two roots outside the cut.

If $c(1) = 0$, the integral is convergent and three cases are possible:

(a) $$1 + \int_{-1}^{+1} \frac{c(v')\, dv'}{v' - 1} < 0, \qquad r = 1.$$

Both zeros of $\Lambda(z)$ are outside the cut.

(b) $$1 + \int_{-1}^{+1} \frac{c(v')\, dv'}{v' - 1} = 0, \qquad r = 1.$$

There are simultaneous zeros of $\Lambda(z)$ and $\Lambda^{*}(s)$ at $z = \pm 1$ or $s = \pm 1$, respectively.

(c) $$1 + \int_{-1}^{+1} \frac{c(v')\, dv'}{v' - 1} > 0, \qquad r = 0.$$

There are no discrete eigenvalues and no asymptotic solution.

Suppose now that for some value of $x = x_0$ there is given energy–angular distribution of neutrons. For simplicity, put $x_0 = 0$; then the boundary condition written in terms of v is

$$\psi(v, E) = g(E) \int_{-1}^{+1} \frac{sH(s)\, dt}{s - v} + \lambda(v, E) + g(E) \sum_i \frac{L_i H(L_i)}{L_i - v}, \tag{6.16}$$

where $\psi(v, E)$ is a given function and last term is a sum over all discrete solutions.

Multiply through both sides of Eq. (6.16) by $h(E)$ and integrate over $M(v)$. We get

$$\bar{\psi}(v) = c(v) \int_{-1}^{+1} \frac{sH(s)\, ds}{s - v} + \lambda(v) H(v) + \sum_i \frac{c(v) L_i H(L_i)}{L_i - v}, \tag{6.17}$$

where

$$\bar{\psi}(v) = \int_{M(v)} dE h(E) \psi(v, E), \tag{6.18}$$

and

$$\lambda(v) = 1 + v \int_{-1}^{+1} \frac{c(v')\, dv'}{v - v'}. \tag{6.19}$$

We see that Eq. (6.17) does not essentially differ from the one-velocity equation. The only difference is that in the one-velocity theory, $c(v)$ is constant. Equation (6.15) can be easily reduced to the non-homogeneous Hilbert problem which can be solved by the standard methods.

In cases (a) and (b), the index of the Hilbert problem is equal to $(-2r)$, that is, to (-2) or to zero and the number of additional conditions is exactly equal to the number of arbitrary constants $H(L_i)$.

In case (c), the original singular integral equation is not of a regular type. However, the procedure developed in reference 3 can be employed and the discrete solutions on the cut at $s = \pm 1$ can be accounted for.

Knowing the constants $H(L_i)$ and the function $H(t)$, we can find $\lambda(v, E)$ from the original Eq. (6.14).

It will be noted that in this case the analytic solutions may be obtained also with half-range boundary conditions, e.g., for the Milne's problem, that is, a medium occupying the region $x > 0$. We will assume that the characteristic equation has two roots lying outside the cut at $\pm L$.

From the boundary condition at infinity we have that for a semi-infinite medium

$$H(s) = 0, \qquad s < 0. \tag{6.20}$$

There are no neutrons coming from vacuum to the medium. This leads to the following equation:

$$c(v) \int_0^1 \frac{sH(s)\, ds}{s - v} + \lambda(v) H(v) + H_{+} \frac{Lc(v)}{L - v} + H_{-} \frac{Lc(v)}{L + v} = 0. \tag{6.21}$$

The asymptotic part of the solution is expressed by constants H_+ and H_-.

$$\psi_{as}(x, \mu, E) = H_+ \frac{Lc(\mu l(E))}{L - \mu l(E)} l^{-x/L} + H_- \frac{Lc(\mu l(E))}{L + \mu l(E)} e^{x/L}. \quad (6.22)$$

The asymptotic behavior of neutrons is therefore described by the ratio H_+/H_-. The latter can be easily found from Eq. (6.21).

Introduce a sectionally holomorphic function

$$N(z) = \frac{1}{2\pi i} \int_0^1 \frac{sH(s)\, ds}{s - z}. \quad (6.23)$$

Then Eq. (6.21) is reduced to the Hilbert problem

$$\Lambda^+(v)N^+(v) - \Lambda^-(v)N^-(v) = f(v), \quad (6.24)$$

where

$$f(v) = -Lc(v)v\left[\frac{H_+}{L - v} + \frac{H_-}{L + v}\right]. \quad (6.25)$$

The solution to the homogeneous Hilbert problem corresponding to (6.24) is

$$X(z) = (1 - z)e^{-\Gamma(z)}, \quad (6.26)$$

where

$$\Gamma(z) = \int_0^1 \frac{\theta(t)\, dt}{t - z}, \quad (6.27)$$

and

$$\theta(t) = \frac{1}{\pi} \arg \Lambda^+(t). \quad (6.28)$$

Knowing $X(z)$ we can find the solution to problem (6.24) and then to Eq. (6.21). In fact,

$$N(z) = \frac{X(z)}{2\pi i} \int_0^1 \frac{f(v)\, dv}{\Lambda^+(v)X^+(v)(v - z)} \quad (6.29)$$

provided that

$$\int_0^1 \frac{f(v)\, dv}{\Lambda^+(v)X^+(v)} = 0. \quad (6.30)$$

The last condition gives

$$\frac{H_+}{H_-} = \frac{\int_0^1 \frac{vc(v)\, dv}{\Lambda^+(v)X^+(v)(v + L)}}{\int_0^1 \frac{vc(v)\, dv}{\Lambda^+(v)X^+(v)(v - L)}}. \quad (6.31)$$

We have an obvious identity:

$$\frac{1}{X(z)} = \frac{1}{2\pi i} \int_c \frac{dz'}{X(z')(z' - z)}, \quad (6.32)$$

where the contour encircles the cut. Using the equation satisfied by $X(z)$, namely

$$X^+(v)\Lambda^+(v) = x^-(v)\Lambda^-(v),$$

we have

$$\frac{1}{X(z)} = \frac{1}{2\pi i} \int_0^1 \frac{[\Lambda^+(v) - \Lambda^-(v)]\, dv}{X^+(v)\Lambda^+(v)(v - z)} = \int_0^1 \frac{vc(v)\, dv}{X^+(v)\Lambda^+(v)(v - z)}. \quad (6.33)$$

From this we get the final result:

$$\frac{H_+}{H_-} = \frac{X(+L)}{X(-L)}. \quad (6.34)$$

Formally, this is identical to the result obtained in one-velocity theory. From (6.34) one can easily calculate the extrapolated end point.

Integrating the expression (6.22) over μ we get the asymptotic neutron density

$$\rho_{as}(x, E) = H_+ f(E)e^{-x/L} + H_- f(E)e^{x/L}, \quad (6.35)$$

where

$$f(E) = L \int_{-l(E)}^{l(E)} \frac{c(v)\, dv}{L \mp v}. \quad (6.36)$$

Now the extrapolated end point defined by the condition

$$\rho_{as}(z_0) = 0 \quad (6.37)$$

is given by

$$z_0 = \frac{L}{2} \log\left[-\frac{X(-L)}{X(L)}\right]. \quad (6.38)$$

7. CONCLUSIONS

The main objective of this paper has been to present the general procedure of solving the energy-dependent Boltzmann equation. The authors believe that it can be used extensively in many problems of neutron transport theory.

It is fairly obvious that unless the kernel is specified, one cannot expect to obtain the simple results. It is hoped, however, that the results obtained in this paper can help in better understanding of the structure of the Boltzmann equations, which in turn might be used in dealing with specific problems.

By way of illustration, the fully degenerate kernel is considered in detail. It has been shown that in this case the problem is reduced to the equation very similar to that obtained in one-velocity theory.

Computer Studies of Energy Sharing and Ergodicity for Nonlinear Oscillator Systems*

JOSEPH FORD AND JOHN WATERS
School of Physics, Georgia Institute of Technology, Atlanta, Georgia
(Received 7 January 1963)

Weakly coupled systems of N oscillators are investigated using Hamiltonians of the form

$$H = \tfrac{1}{2}\sum_{k=1}^{N}(p_k^2 + \omega_k^2 q_k^2) + \alpha \sum_{i,k,l=1}^{N} A_{ikl} q_i q_k q_l,$$

where the A_{ikl} are constants and where α is chosen to be sufficiently small that the coupling energy never exceeds some small fraction of the total energy. Starting from selected initial conditions, a computer is used to provide exact solutions to the equations of motion for systems of 2, 3, 5, and 15 oscillators. Various perturbation schemes are used to predict, interpret, and extend these computer results. In particular, it is demonstrated that these systems can share energy only if the uncoupled frequencies ω_k satisfy resonance conditions of the form

$$\sum n_k \omega_k \lesssim \alpha$$

for certain integers n_k determined by the particular coupling. It is shown that these systems have N normal modes, where a normal mode is defined as motion for which each oscillator moves with essentially constant amplitude and at a given frequency or some harmonic of this frequency. These systems are shown to have, at least, one constant of the motion, analytic in q, p, and α, other than the total energy. Finally, it is demonstrated that the single-oscillator energy distribution density for a 5-oscillator linear and nonlinear system has the Boltzmann form predicted by statistical mechanics. Thus, these nonlinear systems are shown to have many features in common with linear systems. In particular, it is unlikely that they are ergodic. From the standpoint of statistical mechanics, it is argued that this lack of ergodicity may be a welcome feature.

1. INTRODUCTION

THERE is a long standing desire to solemnize the union of particle mechanics and statistical mechanics by exhibiting a mechanical system whose exact solution can be shown to yield an approach to equilibrium. Due to the work of Poincare,[1] Fermi,[2] and others,[3] physicists tend to believe that weakly coupled nonlinear systems will exhibit that ergodic behavior deemed necessary[3] for an approach to equilibrium. In accepting ergodicity, however, one gives up almost all hope of analytic solution. The solutions to the equations of motion are expected to be incredibly complex, and the constants of the motion are assumed to be pathological monstrosities.

As a first step in circumventing this difficulty, Fermi, Pasta, and Ulam[4] used a computer to investigate the question of equipartition of energy for weakly coupled oscillator systems using Hamiltonians of the form

$$H = \frac{1}{2}\sum_{k=1}^{N}(p_k^2 + \omega_k^2 q_k^2) + \alpha \sum_{i,k,l=1}^{N} A_{ikl} q_i q_k q_l, \qquad (1)$$

where the A_{ikl} are constants defined later and where α is chosen sufficiently small that the coupling energy never exceeds some small fraction of the total energy. To their surprise, FPU found that these nonlinear oscillator systems yielded very little energy sharing at all. Certainly, the systems considered did not exhibit ergodic behavior. Since nonergodic systems might be susceptible to analysis, one of us[5] was led to investigate these nonlinear oscillator systems using perturbation theory. This heuristic investigation indicated that the FPU systems failed to show energy sharing because of an unfortuitous choice of uncoupled frequencies ω_k. Moreover, the available perturbation theories indicated that even energy sharing oscillator systems would not be ergodic.

If one believes that ergodicity is necessary for an approach to equilibrium, then apparently one must abandon these oscillator systems. But the work of Peierls,[6] for example, on lattice thermal conductivity makes one reluctant to believe that these systems do not exhibit an approach to equilibrium. The immediate suggestion then is

* Work supported by the National Science Foundation.
[1] H. Poincare, *Methodes Nouvelles de la Mechanique Celeste*, (Dover Publications, Inc., New York, 1957), Vol. 1, p. 233.
[2] E. Fermi, Z. Physik **24**, 261 (1923).
[3] For references see D. ter Haar, *Elements of Statistical Mechanics* (Rinehart and Company, New York, 1954), Appendix I.
[4] E. Fermi, J. Pasta, and S. Ulam, Los Alamos Scientific Laboratory Report. LA-1940 (1955), hereafter referred to as FPU.
[5] J. Ford, J. Math. Phys. **2**, 387 (1961).
[6] R. E. Peierls, Ann. Physik **3**, 1055 (1929).

that one should give up ergodicity. In addition to the data we shall present in the following sections, there is a growing amount of evidence that many physically interesting systems are not ergodic. Resibois and Prigogine[7] have discovered constants of the motion for gas systems. Kolmogorov[8] claims to have shown that nonlinear systems, more general than the systems considered here, are not ergodic. Statistical mechanics itself yields indirect evidence on this question. Numerous authors[9] have pointed out that only the reduced probability densities f_n, $n \ll N$, are needed for most physical calculations. Thus, the N-particle probability density f_N is not required to approach equilibrium and there is no need for ergodicity.

Perhaps the most important consequence of abandoning ergodicity, however, is that it opens the door for an analytic solution of the approach to equilibrium. It might then be expected that the improvement wrought by nonlinear systems over linear ones in the study of irreversible behavior is more modest than popularly supposed. Thus, for example, the only effect of a weak nonlinearity superimposed on a linear system might be the bringing to equilibrium of certain low-order correlation functions known[10] not to approach equilibrium for the linear system. In what follows, we attempt to show that the weakly coupled oscillator systems of Hamiltonian (1) indeed have that simple behavior expected of nonergodic systems while retaining the capacity for an approach to the equilibrium. It is our hope that these investigations will lay the groundwork for a more general study of irreversibility. We now outline the material to be covered in the sections which follow.

The question of energy sharing for the nonlinear oscillator systems of Hamiltonian (1) is investigated in Sec. II. The criteria for energy sharing are derived using perturbation theory. Their validity is demonstrated using the exact solutions to the equations of motion obtained on a computer for systems of 2, 3, 5, and 15 oscillators. Equipartition of energy is then demonstrated to occur for the 5-oscillator system.

In Sec. III, the energy-sharing oscillator systems of Sec. II are shown to have normal modes, where a normal mode is defined as motion in which each oscillator moves at essentially constant amplitude (energy) and at a given frequency or some harmonic of this frequency. The amplitudes of these normal modes are predicted using perturbation theory; the validity of the calculations is then demonstrated using a computer for systems of 3 and 5 oscillators. The computer solutions indicate that these normal modes are stable in the sense that a system started near one of its normal modes remains near it. Attention is then turned to constants of the motion. Perturbation theory yields an expression for a constant of the motion, analytic in q, p, and α, distinct from the total energy. The validity of the perturbation calculation for the case of the 5-oscillator system is verified. Thus, these nonlinear systems are shown to have many properties in common with linear systems. In particular, they are nonergodic and are susceptible to analysis.

Section IV compares the computer solution of Hamiltonian (1) for $N = 5$ with the exact solution of the linear Hamiltonian

$$H = \frac{1}{2}\sum_{k=1}^{5}(p_k^2 + q_k^2) - \alpha\sum_{k=1}^{4} q_{k+1}q_k. \qquad (2)$$

For both cases, all energy was initially given to one oscillator. The single-oscillator energy distribution density is calculated from these curves, i.e., the probability density that any single oscillator will have energy between E and $E + dE$, and a Boltzmann distribution is obtained for both linear and nonlinear systems. Thus for both systems, four particles evidently form an adequate heat bath for the fifth particle.

Section V reviews our conclusions and offers suggestions for future investigations.

We conclude this introduction on a note of caution. We have used a computer to provide exact, but limited, information on just a few of the systems described by Hamiltonian (1), and the perturbation theories we use are presented without a hint of rigor. Nonetheless, it is our belief that the necessary rigor can be supplied and these results have a rather wide range of validity. Our optimism, however, is certainly open to question; and we offer our results in the hope of spurring further study of these systems.

II. ENERGY SHARING

In this section we investigate the question of energy sharing for systems of 2, 3, 5, and 15 oscilla-

[7] P. Resibois and I. Prigogine, Bull. Acad. Roy. Belg. Class. Sci. **46**, 53 (1960).

[8] A. N. Kolmogorov, *Proceedings of the International Congress of Mathematicians, Amsterdam*, (North-Holland Publishing Co., Amsterdam, 1957), Vol. 1, pp. 315–333. Kolmogorov's article is in Russian. There is a detailed English review by J. Moser, Math. Rev. **20**, 675 (1959).

[9] See, for example, P. Mazur and E. Montroll, J. Math. Phys. **1**, 70 (1960).

[10] G. Klein and I. Prigogine, Physica **19**, 1053 (1953).

tors. Using perturbation theory, we obtain the criteria for energy sharing. The validity of these criteria is then tested by numerically solving the equations of motion on a computer. In order that the main result not be lost in the following welter of mathematical detail, let us now state that the oscillators of Hamiltonian (1) share energy only if the uncoupled frequencies ω_k satisfy resonance conditions of the form

$$\sum n_k\omega_k \lesssim \alpha, \tag{3}$$

where the n_k are integers determined by the particular coupling used. This sharpens a result previously stated by one of the authors.[5]

We are interested in continuing the investigations begun by FPU and thus we examine systems which have the same cubic coupling that they used. Now the Hamiltonian for the systems they considered may be written in the form

$$H = \frac{1}{2}\sum_{k=1}^{N} P_k^2 + \frac{1}{2}\sum_{k=0}^{N}(X_{k+1} - X_k)^2 + \frac{\alpha}{3}\sum_{k=0}^{N}(X_{k+1} - X_k)^3, \tag{4}$$

where $X_0 \equiv X_{N+1} \equiv 0$. But is is well-known that a linear transformation will diagonalize the quadratic terms of Hamiltonian (4), thus bringing it to the form of Hamiltonian (1). Hence, we refer to Eq. (4) as the FPU particle Hamiltonian and to Eq. (1), obtained by substituting the necessary linear transformation into Eq. (4), as the FPU oscillator Hamiltonian.

Since all the essential features of energy sharing appear in a 2-oscillator system, we develop the theory for this simplest case. The FPU oscillator Hamiltonian for two oscillators has the form

$$H = \tfrac{1}{2}(p_1^2 + p_2^2 + \omega_1^2q_1^2 + \omega_2^2q_2^2) + 2^{-\frac{1}{2}}\alpha(q_1^2q_2 - q_2^3). \tag{5}$$

The linear transformation which converts Hamiltonian (4) for $N = 2$ into Hamiltonian (5) determines the ω_k; i.e., $\omega_1 = 1.0$ and $\omega_2 = \sqrt{3}$. However, we wish to determine if there are any ω_1, ω_2 pairs which yield energy sharing using the same cubic coupling; thus, we treat ω_1, ω_2 as free parameters. The equations of motion which follow from Hamiltonian (5) are

$$\ddot{q}_1 = -\omega_1^2q_1 - \sqrt{2}\,\alpha q_1q_2,$$
$$\ddot{q}_2 = -\omega_2^2q_2 - 2^{-\frac{1}{2}}\alpha(q_1^2 - 3q_2^2). \tag{6}$$

In Appendix I, we use Wigner–Brillouin perturbation theory[11] and find that, to order α^2, Eq. (6) has the formal solution

$$\begin{aligned} q_1 = {} & A_1\cos\tau_1 - \alpha\left\{\frac{A_1A_2\cos(\tau_1+\tau_2)}{\sqrt{2}\,[\omega_1^2 - (\Omega_1+\Omega_2)^2]} + \frac{A_1A_2\cos(\tau_1-\tau_2)}{\sqrt{2}\,[\omega_1^2 - (\Omega_1-\Omega_2)^2]}\right\} \\ & + \alpha^2\left\{\frac{A_1^3\cos 3\tau_1}{4[\omega_2^2 - 4\Omega_1^2][\omega_1^2 - 9\Omega_1^2]} - \frac{3A_1A_2^2\cos(\tau_1+2\tau_2)}{4[\omega_2^2 - 4\Omega_2^2][\omega_1^2 - (\Omega_1+2\Omega_2)^2]}\right. \\ & - \frac{3A_1A_2^2\cos(\tau_1-2\tau_2)}{4[\omega_2^2 - 4\Omega_2^2][\omega_1^2 - (\Omega_1-2\Omega_2)^2]} + \frac{A_1A_2^2\cos(\tau_1+2\tau_2)}{2[\omega_1^2 - (\Omega_1+\Omega_2)^2][\omega_1^2 - (\Omega_1+2\Omega_2)^2]} \\ & \left. + \frac{A_1A_2^2\cos(\tau_1-2\tau_2)}{2[\omega_1^2 - (\Omega_1-\Omega_2)^2][\omega_1^2 - (\Omega_1-2\Omega_2)^2]}\right\}, \end{aligned} \tag{7a}$$

$$\begin{aligned} q_2 = {} & A_2\cos\tau_2 - \alpha\left\{\frac{(A_1^2 - 3A_2^2)}{2\sqrt{2}\,\omega_2^2} + \frac{A_1^2\cos 2\tau_1}{2\sqrt{2}\,[\omega_2^2 - 4\Omega_1^2]} - \frac{3A_2^2\cos 2\tau_2}{2\sqrt{2}\,[\omega_2^2 - 4\Omega_2^2]}\right\} \\ & + \alpha^2\left\{\frac{A_1^2A_2\cos(2\tau_1+\tau_2)}{2[\omega_1^2 - (\Omega_1+\Omega_2)^2][\omega_2^2 - (2\Omega_1+\Omega_2)^2]} + \frac{A_1^2A_2\cos(2\tau_1-\tau_2)}{2[\omega_1^2 - (\Omega_1-\Omega_2)^2][\omega_2^2 - (2\Omega_1-\Omega_2)^2]}\right. \\ & - \frac{3A_1^2\cos(2\tau_1-\tau_2)}{4[\omega_2^2 - 4\Omega_1^2][\omega_2^2 - (2\Omega_1-\Omega_2)^2]} - \frac{3A_1^2A_2\cos(2\tau_1+\tau_2)}{4[\omega_2^2 - 4\Omega_1^2][\omega_2^2 - (2\Omega_1+\Omega_2)^2]} \\ & \left. + \frac{9A_2^3\cos 3\tau_2}{4[\omega_2^2 - 4\Omega_2^2][\omega_2^2 - 9\Omega_2^2]}\right\}, \end{aligned} \tag{7b}$$

$$\Omega_1^2 = \omega_1^2 - \alpha^2\left\{\frac{A_1^2 - 3A_2^2}{2\omega_2^2} + \frac{A_1^2}{4[\omega_2^2 - 4\Omega_1^2]} + \frac{A_2^2}{2[\omega_1^2 - (\Omega_1+\Omega_2)^2]} + \frac{A_2^2}{2[\omega_1^2 - (\Omega_1-\Omega_2)^2]}\right\}, \tag{7c}$$

[11] The application of Wigner–Brillouin perturbation methods to oscillator systems is discussed in a paper by E. Atlee Jackson, J. Math. Phys. 4, 551 (1963). We first heard of the W–B method through him, and we are indebted to him for allowing us to read a preprint of this paper.

$$\Omega_2^2 = \omega_2^2 - \alpha^2\Big\{\frac{A_1^2}{2[\omega_1^2 - (\Omega_1 + \Omega_2)^2]} + \frac{A_1^2}{2[\omega_1^2 - (\Omega_1 - \Omega_2)^2]} - \frac{3(A_1^2 - 3A_2^2)}{2\omega_2^2} + \frac{9A_2^2}{4[\omega_2^2 - 4\Omega_2^2]}\Big\}. \quad (7d)$$

Here $\tau_1 = \Omega_1 t + \theta_1$, and $\tau_2 = \Omega_2 t + \theta_2$. The constants A_1, A_2, θ_1, and θ_2 are determined by initial conditions. Thus, the Wigner–Brillouin perturbation scheme converts the problem of solving differential Eq. (6) to the problem of solving the algebraic Eqs. (7c) and (7d) for the perturbed frequencies Ω_1 and Ω_2. Fortunately, our task is even simpler since we may use Eq. (7) to discuss energy sharing without obtaining an exact solution of Eqs. (7c) and (7d).

An approximate solution to Eqs. (7) is obtained by setting $\Omega_1 = \omega_1$ and $\Omega_2 = \omega_2$ in all the denominators. Provided that none of these denominators is zero or of the order of α, we see from Eqs. (7) that, for small α, $q_1 = A_1 \cos \tau_1$ and $q_2 = A_2 \cos \tau_2$. In this case the oscillators move almost as if uncoupled and there is no appreciable energy sharing. In terms of the uncoupled frequencies, no energy sharing occurs provided

$$n_1\omega_1 + n_2\omega_2 \gg \alpha, \quad (8)$$

where n_1 and n_2 range over the small integral values which appear in Eqs. (7). Examining Eq. (7c) in this approximation,

$$\Omega_1^2 = \omega_1^2 - \alpha^2\Big\{\frac{A_1^2 - 3A_2^2}{2\omega_2^2} + \frac{A_1^2}{4(\omega_2^2 - 4\omega_1^2)} + \frac{A_2^2}{2[\omega_1^2 - (\omega_1 + \omega_2)^2]} + \frac{A_2^2}{2[\omega_1^2 - (\omega_1 - \omega_2)^2]}\Big\}, \quad (9)$$

we see that a small denominator will occur only if $(2\omega_1 - \omega_2) \approx \alpha$. The same conclusion is reached examining Eq. (7d). Thus, when $(2\omega_1 - \omega_2) \approx \alpha$, we would anticipate that the exact solution of Eqs. (7c) and (7d) would have the form

$$\Omega_1 = \omega_1 + K_1\alpha + \cdots,$$
$$\Omega_2 = \omega_2 + K_2\alpha + \cdots, \quad (10)$$

where K_1 and K_2 are constants. As a consequence, for example, one of the terms of order α in Eq. (7a) may be written

$$-\frac{\alpha A_1 A_2 \cos(\tau_1 - \tau_2)}{\sqrt{2}\,[\omega_1^2 - (\Omega_1 - \Omega_2)^2]} \approx -KA_1A_2 \cos(\tau_1 - \tau_2), \quad (11)$$

where K is some constant; i.e., what had been a first-order term becomes a zeroth-order term. But this modification of our zeroth-order term means that, even for α arbitarily small, energy sharing will occur.

When $2\omega_1$ is not on the order of ω_2, we see from Eq. (9) that $\Omega_1 = \omega_1 + K_3\alpha^2$ and $\Omega_2 = \omega_2 + K_4\alpha^2$. In this case our only hope for energy sharing lies with terms of order α^2 in Eq (7a). Examining the denominators of these second-order terms, we see that none of these denominators will be of order α^2 unless $(\omega_1 - \omega_2) \approx \alpha^2$. If $\omega_1 \approx \omega_2$, then some of the second-order terms become zeroth order, and energy sharing occurs. As might be expected, the amount of energy sharing for the case $\omega_1 = \omega_2$ is much less than that for the case $2\omega_1 = \omega_2$. This point is verified by the computer solutions discussed below and by the theory presented in Section IV.

We have now exhausted the possibilities of energy sharing for Eq. (6). Unless $2\omega_1 \approx \omega_2$ or $\omega_1 \approx \omega_2$, we have seen that there is no energy sharing through second order. In nth order, $n > 2$, a frequency relation such as $3\omega_1 - 5\omega_2 = 0$ might, at worst, make some denominator become of the order of α^2; but the α^n in the numerator would prevent the term from contributing to zeroth order.

The generalization of these arguments for a larger number of oscillators or for a different type coupling is straightforward. For the cubic coupling of Eq. (1), the conditions for energy sharing are summarized by Eq. (3). We might mention in passing that while the Wigner–Brillouin method is useful for discussing energy sharing, its usefulness for a quantitative solution at resonance, e.g., $2\omega_1 = \omega_2$, is probably limited due to the fact that corrections to the zeroth-order term arise in most of the higher-order terms. Nonetheless, this point needs investigation.

We now turn to a discussion of the verification of the preceding theory by exact solution of the differential equations on a computer. These exact solutions were used to compute the energy E_k of each oscillator as given by

$$E_k = \tfrac{1}{2}(p_k^2 + \omega_k^2 q_k^2). \quad (12)$$

The numerical data were then plotted to give E_k as a function of time. In all cases $\alpha = 0.1$ and the coupling energy in the cubic terms seldom exceeds 10% of the total energy; thus it is legitimate to speak of individual oscillator energies.

In Figs. 1 and 2 we plot E_1 and E_2 as determined from the computer solution of Eq. (6) for various ω_k near the $2\omega_1 = \omega_2$ resonance. Our object here is

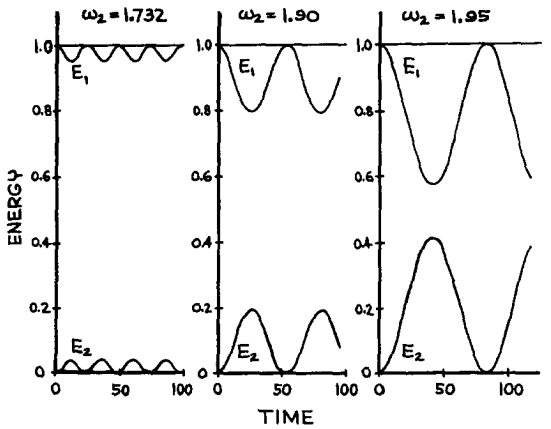


FIG. 1. A plot of the single-oscillator energies for the 2-oscillator system showing the dependence of energy sharing on frequency near the $\omega_2 = 2\omega_1$ resonance. Here $\omega_1 = 1.0$, and ω_2 is varied, starting from its FPU value of $\sqrt{3}$. For all curves, initially one unit of kinetic energy was given to oscillator 1. Here $T_1 = (2\pi/\omega_1) = 6.28$ time units.

not only to show that energy sharing occurs, but also to demonstrate the sharpness of the resonance. For computational definiteness, we set $\omega_1 = 1.0$ and varied ω_2. First, we set $\omega_2 = \sqrt{3}$, the FPU frequency. Then ω_2 is set equal to 1.90, 1.95, 1.99, and 2.00. The initial conditions for all these curves were the same; all positions and momenta were zero except $p_1 = \sqrt{2}$, making the total energy $E_T = 1.0$. A measure of the time scale may be realized by noting that the uncoupled period of oscillator one is $T_1 = (2\pi/\omega_1) = 6.28$.

Figures 3 and 4 plot E_1 and E_2 for Eq. (6) at the resonance $\omega_1 = \omega_2 = 1.0$. In Fig. 3, initially only $p_1 = \sqrt{2}$ was nonzero, making $E_T = 1.0$. For Fig. 4, initially only $p_1 = 1.0$ and $p_2 = 1.0$ were nonzero, again making $E_T = 1.0$. As previously noted, not much energy sharing occurs for this

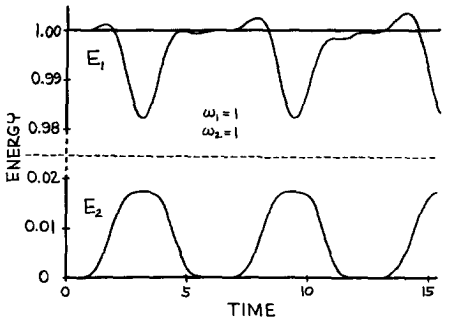


FIG. 3. A plot of the E_k for the weak resonance at $\omega_1 = \omega_2$ for the same system as in Figs. 1 and 2. Initially, one unit of kinetic energy was given to oscillator 1. Oscillator 2 never receives more than 0.02 units of energy.

resonance; but it should be noted that the energy sharing which does occur will continue to occur no matter how small the value of α, excluding $\alpha = 0$ of course.

Figures 5 and 6 plot E_1, E_2, and E_3 for the three oscillator system whose oscillator Hamiltonian is

$$H = \tfrac{1}{2}\sum_{k=1}^{3}(p_k^2 + \omega_k^2 q_k^2) + \alpha\{(1 - 2^{-\frac{1}{2}})q_1^2 q_2 - (1 + 2^{-\frac{1}{2}})q_2 q_3^2 + \sqrt{2}\, q_1 q_2 q_3\}. \quad (13)$$

Hamiltonian (13) was obtained, of course, by substituting the necessary linear transformation into Hamiltonian (4). Figure 5 plots the E_k found using the FPU frequencies $\omega_1 = (2 - \sqrt{2})^{\frac{1}{2}} = 0.765$, $\omega_2 = \sqrt{2} = 1.41$, and $\omega_3 = (2 + \sqrt{2})^{\frac{1}{2}} = 1.85$. Here only $p_1 = \sqrt{6}$ is nonzero, making $E_T = 3.0$. The shortest uncoupled period is $T_3 = (2\pi/\omega_3) = 3.4$; thus, this system has a recurrence time of about 14 T_3. Note that E_3 gets almost no energy. Figure 6 plots the E_k using the frequencies $\omega_1 = 0.6$, $\omega_2 = 1.2$, $\omega_3 = 1.8$. This is the resonance $\omega_1 = \frac{1}{2}\omega_2 = \frac{1}{3}\omega_3$

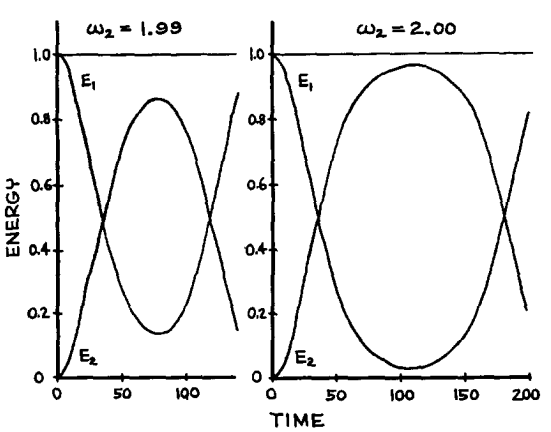


FIG. 2. This figure is a continuation of Fig. 1 and plots the E_k for $\omega_2 = 1.99$ and 2.00.

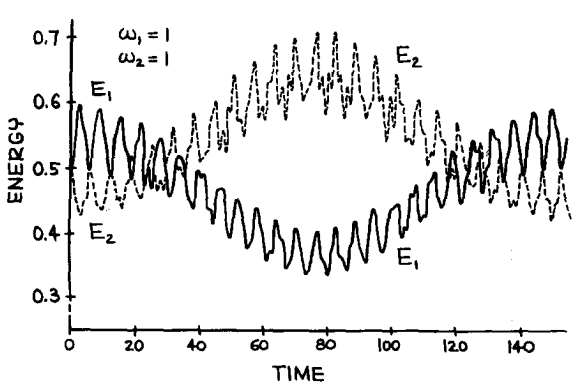


FIG. 4. Plots the E_k for different initial conditions for the same system as in Fig. 3. Here initially one unit of kinetic energy was equally divided between the oscillators. Again, there is very little energy sharing.

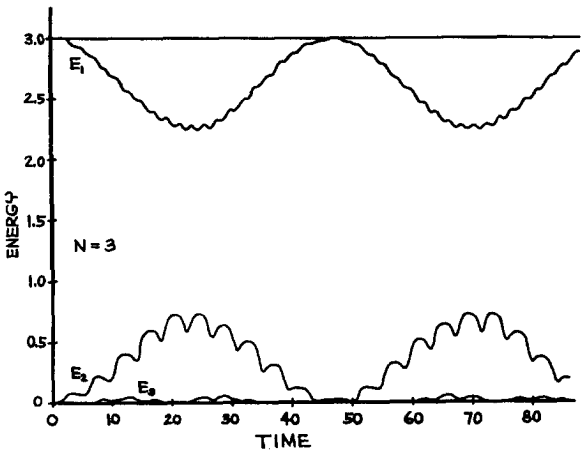


FIG. 5. The E_k for the 3-oscillator system using the FPU frequencies $\omega_1 = 0.765$, $\omega_2 = 1.41$, and $\omega_3 = 1.85$. Initially, three units of kinetic energy were given to oscillator 1. The recurrence time for this system is about $14T_3$, where $T_3 = (2\pi/\omega_3)$.

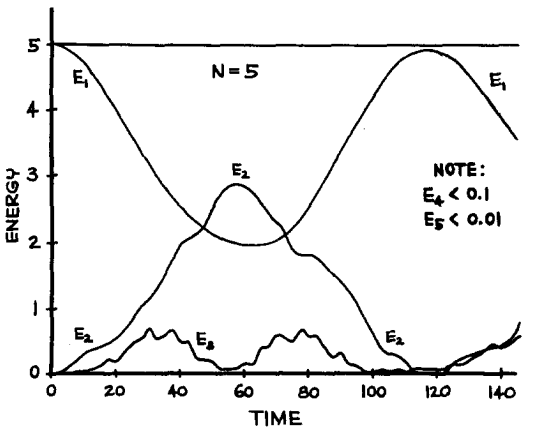


FIG. 7. The E_k for the 5-oscillator system using the FPU frequencies $\omega_k = 2\sin(k\pi/12)$. Initially, five units of kinetic energy were given to oscillator 1. The system has a recurrence time of about $35T_5$, and oscillators 4 and 5 receive little energy.

determined from the Wigner–Brillouin series. Note that these resonant frequencies represent only a slight shift from the FPU frequencies, again indicating the sharpness of the resonance. The initial conditions for Fig. 6 are the same as those for Fig. 5.

Figures 7 and 8 plot the E_k for the oscillator Hamiltonian

$$H = \tfrac{1}{2}\sum_{k=1}^{5}(p_k^2 + \omega_k^2 q_k^2) + \alpha\{0.078q_1^2q_2 + 0.5q_2^2q_4 - 1.078q_2q_5^2 - 0.5q_4^3 + 0.422q_1q_2q_3 + 0.732q_1q_3q_4 + q_1q_4q_5 + 1.578q_2q_3q_5 - 2.732q_3q_4q_5\}. \quad (14)$$

For both figures, only $p_1 = \sqrt{10}$ is initially nonzero, making $E_T = 5.0$. Figure 7 plots the E_k using the FPU frequencies $\omega_1 = (2 - \sqrt{3})^{\frac{1}{2}}$, $\omega_2 = 1.0$, $\omega_3 = \sqrt{2}$, $\omega_4 = \sqrt{3}$, $\omega_5 = (2 + \sqrt{3})^{\frac{1}{2}}$. Note that E_4 and E_5 are negligible on the scale shown and that the recurrence time is about $35T_5$, where $T_5 = (2\pi/\omega_5) = 3.25$. The Wigner–Brillouin series indicate that there should be a resonance at $\omega_1 = \frac{1}{2}\omega_2 = \frac{1}{3}\omega_3 = \frac{1}{4}\omega_4 = \frac{1}{5}\omega_5$. The frequencies used for Fig. 8 were $\omega_1 = 0.4$, $\omega_2 = 0.8$, $\omega_3 = 1.2$, $\omega_4 = 1.6$, and $\omega_5 = 2.0$. As before, these frequencies are shifted only slightly from the FPU frequencies. The increase in energy sharing, however, is dramatic. The total energy given to this resonant system is 5 units; thus if equipartition of energy occurs, each oscillator should have unit energy on the time

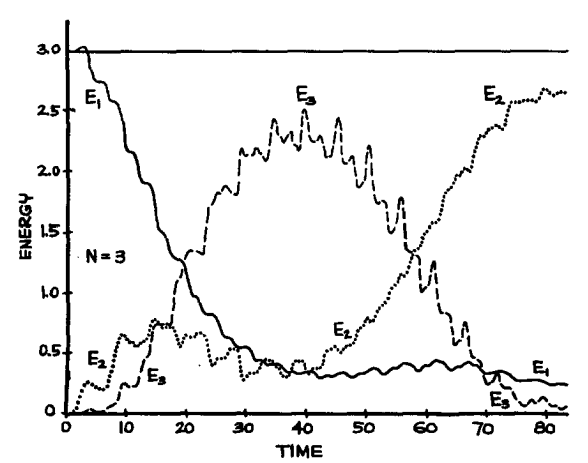


FIG. 6. The E_k for the same system as in Fig. 5 but using the resonant frequencies $\omega_1 = 0.6$, $\omega_2 = 1.2$, $\omega_3 = 1.8$. Note the increase in energy sharing and the disappearance of the short-term recurrence.

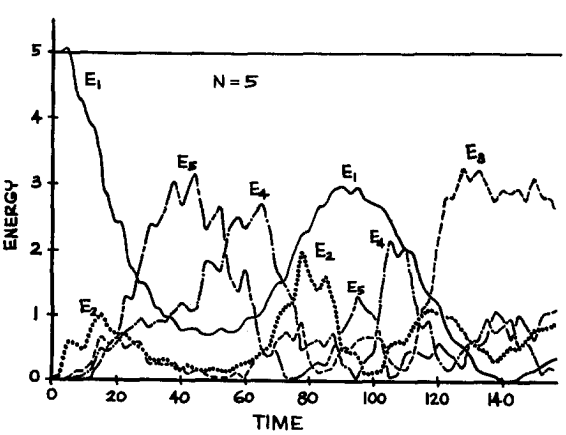


FIG. 8. The E_k using the resonant frequencies $\omega_k = (2k/5)$ for the same system as in Fig. 7. The frequencies are shifted only slightly from their FPU values; however, the increase in energy sharing is dramatic.

average. Table I shows the extent to which equipartition occurs, computed by finding the time average E_k over various time intervals. It perhaps should be noted that the solution for this 5-oscillator system was obtained from $t = 0$, to $t = 1000$. Thus, Fig. 8 and Table I consider only part of the total solution available.

Finally, Fig. 9 plots three of the E_k for the oscillator Hamiltonian for 15 oscillators. The nonresonant FPU case is not plotted since it would look quite similar to the curves appearing in the FPU paper. The resonant curves of Fig. 9 were obtained using the frequencies $\omega_k = (2k/15)$. Only $p_1 = \sqrt{10}$ was nonzero, making $E_T = 5.0$. We have plotted only E_2, E_7, and E_{15} in order to show, without the clutter of 15 interlaced curves, that energy spreads throughout the system.

From the foregoing, we may conclude that the Wigner–Brillouin series is a reasonable method of predicting energy sharing for weakly coupled oscillator systems. It may even be capable of giving quantitative data, although we have not fully investigated this point.

III. NORMAL MODES

In the preceding section we have seen that resonant nonlinear systems share energy, starting from certain initial conditions. Now it is well-known that linear resonant systems share energy, starting from some initial conditions; but if a system of oscillators weakly coupled by linear forces is initially excited in one of its normal modes, the energy of each oscillator remains constant. Thus, not all initial conditions lead to energy sharing. We thus are led to ask whether the same situation holds for weak nonlinear coupling.

Recall that we find linear normal modes by assuming a solution $q_k = A_k \cos \Omega t$, i.e., by assuming that each oscillator can move at the same frequency. The constant amplitudes A_k, as well as the allowed

TABLE I. Time-average single-oscillator energies for the resonant nonlinear 5-oscillator system.

Time interval ΔT	Average $\langle E_1 \rangle$	Average $\langle E_2 \rangle$	Average $\langle E_3 \rangle$	Average $\langle E_4 \rangle$	Average $\langle E_5 \rangle$
0 → 100	2.05	0.59	0.34	0.98	1.14
100 → 200	1.12	0.91	1.48	0.91	0.69
200 → 300	0.70	1.41	0.85	0.70	1.43
300 → 400	1.09	0.76	1.46	0.81	0.99
0 → 400	1.24	0.92	1.03	0.85	1.06
100 → 400	0.97	1.03	1.26	0.81	1.04
200 → 400	0.90	1.09	1.16	0.75	1.21

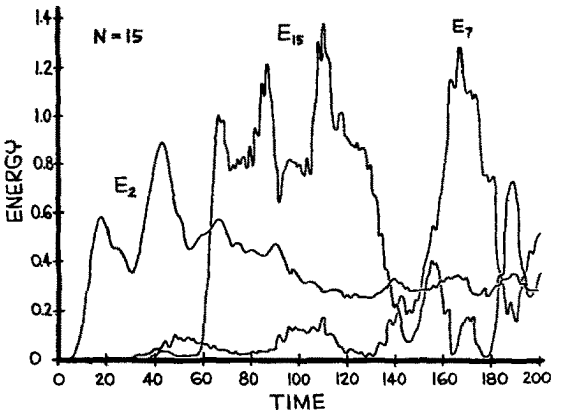


FIG. 9. Three of the E_k for the resonant 15-oscillator system using the frequencies $\omega_k = (2k/15)$. The curves plotted are typical of all 15 curves. Initially, five units of kinetic energy were given to oscillator 1. The comparison FPU curves are not plotted since they would strongly resemble the curves of Figs. 5 and 7.

Ω, are then determined from the equations of motion. We now apply the same technique to the systems discussed in Sec. II. Because these systems are nonlinear, we slightly generalize our assumed solution and write $q_k = A_k \cos n\Omega t$; i.e., we allow for the possibility that the kth oscillator moves at some harmonic of Ω. We now illustrate the method for a resonant 3-oscillator system; we then turn to the computer verification of the theory.

The equations of motion for the 3-oscillator system which follow from Hamiltonian (13) are

$$\begin{aligned}
\ddot{q}_1 + \omega_1^2 q_1 &= \alpha\{(\sqrt{2} - 2)q_1 q_2 - \sqrt{2}\, q_2 q_3\}, \\
\ddot{q}_2 + \omega_2^2 q_2 &= \alpha\{(1 + 2^{-\frac{1}{2}})q_3^2 \\
&\qquad - \sqrt{2}\, q_1 q_3 + (-1 + 2^{-\frac{1}{2}})q_1^2\}, \\
\ddot{q}_3 + \omega_3^2 q_3 &= \alpha\{(\sqrt{2} + 2)q_2 q_3 - \sqrt{2}\, q_1 q_3\}. \qquad (15)
\end{aligned}$$

If the ω_k are such that Eqs. (15) describe nonresonant motion, then, for small α, the oscillators move as if independent. The three normal modes are then easily seen to be

$$\begin{aligned}
&(1) \quad q_1 = A_1 \cos \omega_1 t, \quad q_2 = 0, \quad q_3 = 0; \\
&(2) \quad q_1 = 0, \quad q_2 = A_2 \cos \omega_2 t, \quad q_3 = 0; \qquad (16) \\
&(3) \quad q_1 = 0, \quad q_2 = 0, \quad q_3 = A_3 \cos \omega_3 t.
\end{aligned}$$

However, the existence of normal modes for the nonresonant motion is not particularly interesting; the real interest lies in determining if normal modes continue to exist in the presence of resonance.

Thus, we now consider the case $\omega_1 = \frac{1}{2}\omega_2 = \frac{1}{3}\omega_3$,

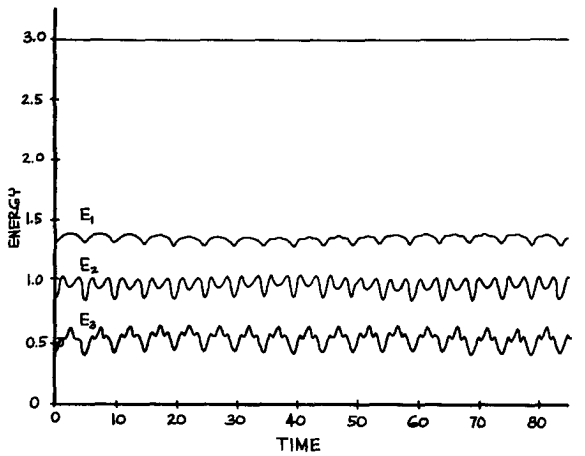


FIG. 10. A plot of the E_k for the first normal mode for the resonant 3-oscillator system. A total energy of three units was chosen for comparison with the energy-sharing curves of Fig. 6.

and we look for a solution of the form

$$\begin{aligned} q_1 &= A_1 \cos \Omega t + \alpha q_{11} + \cdots, \\ q_2 &= A_2 \cos 2\Omega t + \alpha q_{21} + \cdots, \\ q_3 &= A_3 \cos 3\Omega t + \alpha q_{31} + \cdots, \end{aligned} \tag{17}$$

where we shall require that the first-order terms be negligible as $\alpha \to 0$. Putting Eq. (17) into Eq. (15), we find, to order α,

$$\begin{aligned} &(-\Omega^2 + \omega_1^2)A_1 \cos \Omega t + \alpha\{\ddot{q}_{11} + \omega_1^2 q_{11}\} \\ &\quad = \alpha\{-0.293A_1A_2 \cos \Omega t - 0.293A_1A_2 \cos 3\Omega t \\ &\quad - 0.707A_2A_3 \cos \Omega t - 0.707A_2A_3 \cos 5\Omega t\}, \end{aligned} \tag{18a}$$

$$\begin{aligned} &(-\Omega^2 + \omega_1^2)4A_2 \cos 2\Omega t + \alpha\{\ddot{q}_{21} + 4\omega_1^2 q_{21}\} \\ &\quad = \alpha\{1.71A_3^2 + 1.71A_3^2 \cos 6\Omega t \\ &\quad - 0.707A_1A_3 \cos 2\Omega t - 0.707A_1A_3 \cos 4\Omega t \\ &\quad - 0.147A_1^2 - 0.147A_1^2 \cos 2\Omega t\}, \end{aligned} \tag{18b}$$

$$\begin{aligned} &(-\Omega^2 + \omega_1^2)9A_3 \cos 3\Omega t + \alpha\{\ddot{q}_{31} + 9\omega_1^2 q_{31}\} \\ &\quad = \alpha\{1.71A_2A_3 \cos \Omega t + 1.71A_2A_3 \cos 5\Omega t \\ &\quad - 0.707 \cos \Omega t - 0.707A_1A_2 \cos 3\Omega t\}. \end{aligned} \tag{18c}$$

In order to be able to solve for q_{11}, q_{21}, and q_{31} without obtaining small denominators, we must get rid of the cos Ωt terms in Eq. (18a), the cos $2\Omega t$ terms in Eq. (18b), and the cos $3\Omega t$ terms in Eq. (18c). These terms will be eliminated, allowing us to obtain a solution of the desired form in which higher-order terms are negligible as $\alpha \to 0$, provided

$$\begin{aligned} &(\Omega^2 - \omega_1^2)A_1 \\ &\quad - 0.293\alpha A_1A_2 - 0.707\alpha A_2A_3 = 0, \\ &-0.147\alpha A_1^2 \\ &\quad + 4(\Omega^2 - \omega_1^2)A_2 - 0.707\alpha A_1A_3 = 0, \\ &\quad -0.707\alpha A_1A_2 + 9(\Omega^2 - \omega_1^2)A_3 = 0. \end{aligned} \tag{19}$$

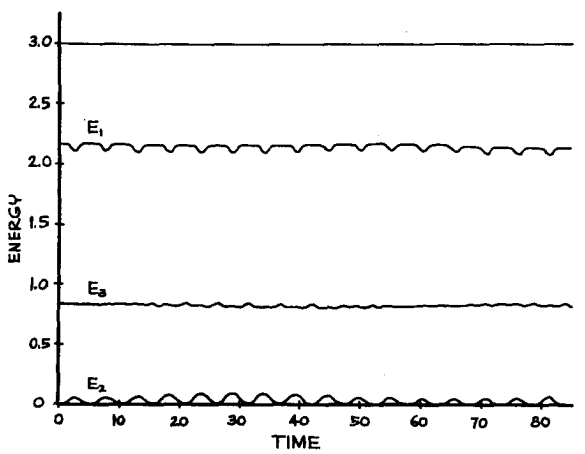


FIG. 11. The second normal mode for three oscillators.

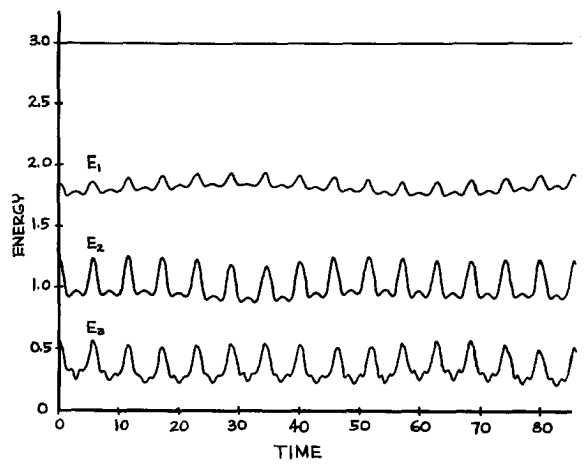


FIG. 12. The third normal mode for three oscillators. Here the initial conditions were not so well chosen and there is a long-term oscillation about the accurate values.

These, of course, are just nonlinear versions of the linear secular equations which arise in finding linear normal modes. Solving Eq. (19), we find the three normal modes:

$$\begin{aligned} &(1)\quad q_1 = 5.4A_3 \cos \Omega t, \quad q_2 = 2.2A_3 \cos 2\Omega t, \\ &\qquad q_3 = A_3 \cos 3\Omega t, \quad \text{and} \quad \Omega^2 = \omega_1^2 + 0.933\alpha A_3; \\ &(2)\quad q_1 = -4.8A_3 \cos \Omega t, \quad q_2 = 0, \\ &\qquad q_3 = A_3 \cos 3\Omega t, \quad \text{and} \quad \Omega^2 = \omega_1^2; \\ &(3)\quad q_1 = 5.4A_3 \cos \Omega t, \quad q_2 = -2.2A_3 \cos 2\Omega t, \\ &\qquad q_3 = A_3 \cos 3\Omega t, \quad \text{and} \quad \Omega^2 = \omega_1^2 - 0.933\alpha A_3. \end{aligned} \tag{20}$$

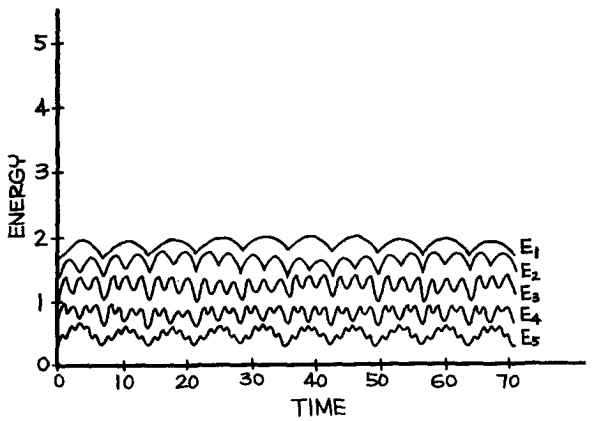


FIG. 13. First normal mode of the 5-oscillator resonant system.

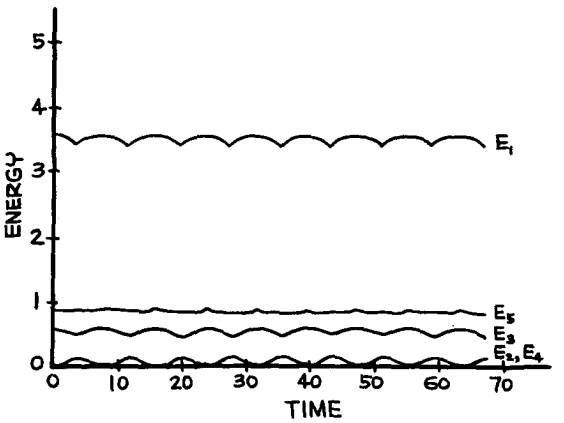


FIG. 15. Third normal mode for five oscillators.

Note that these nonlinear normal modes have two things in common with linear normal modes. First, mode 1 has no nodes, mode 2 has one node, mode 3 has two nodes; second, the frequency splitting is symmetric about ω_1. These same similarities were also observed in the calculations for the resonant 5-oscillator system. Now, because of nonlinearity, the complete solution of Eq. (15) is not just a sum of these three normal modes. However, the computer solutions, to be discussed shortly, indicate that these modes are stable in the sense that a system, started near one of these modes, remains near it. This, at least, suggests that the breakdown of superposition may not be complete. At the present time, we are unable to shed further light on this point.

We now turn to the computer check on the previous theory. Figures 10–12 plot the $E_k = \frac{1}{2}\omega_k^2 A_k^2$ for normal modes 1, 2, and 3, respectively, of Eq. (20). In Eq. (20), A_3 is arbitrary; A_3 is chosen such that the total energy $E_T = 3.0$. A close look at these curves reveals that there is a rapid and a slow oscillation of each E_k. The rapid oscillation is due to higher order terms neglected in Eq. (20). The slow variation is due to the error in calculating the A_k from Eq. (20); i.e., the initial conditions used were slightly off their precise normal values. When the initial E_k deviate from their precise normal mode values, the E_k then oscillate about these precise values. This effect is most noticeable in Fig. 12. It is this oscillation about normal mode values, observed for these and other curves, which leads us to believe that these systems have stable normal modes.

Using the same mathematical technique as for three oscillators, we calculated the normal modes for the resonant 5-oscillator system obeying Hamiltonian (14). Figures 13–17 plot the no-node, one-node, two-node, three-node, and four-node normal modes, respectively. Due to algebraic difficulties, the initial conditions used for Figs. 14 and 16 are rather crude approximations to the accurate normal mode values. No attempt was made to improve

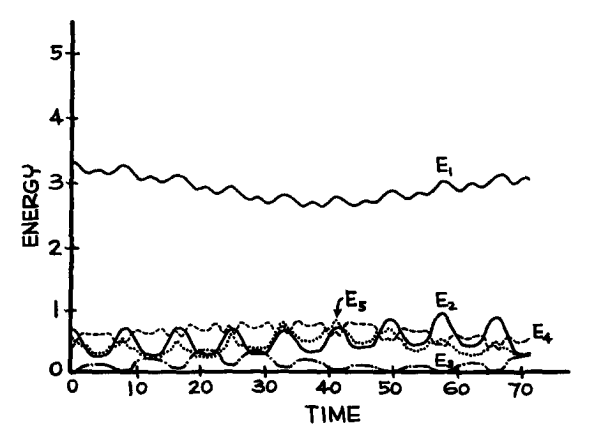


FIG. 14. Second normal mode for five oscillators. Note the long-term oscillation.

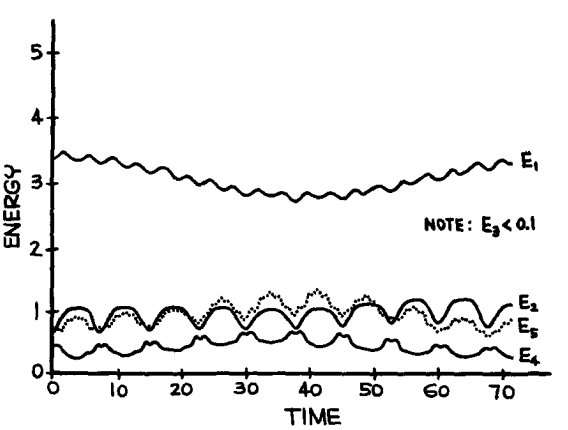


FIG. 16. Fourth normal mode for five oscillators. Again, note the long-term oscillation.

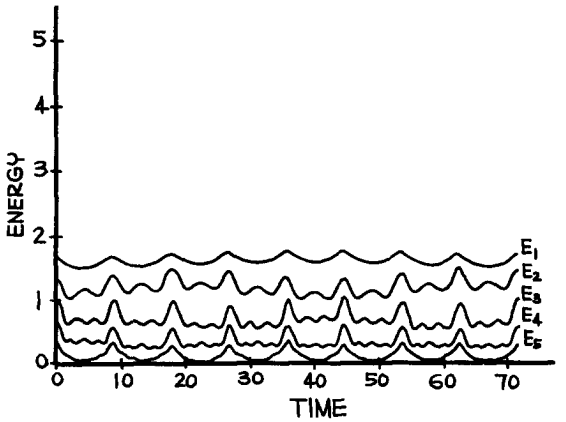


FIG. 17. Fifth normal mode for five oscillators. Note the almost mirror symmetry of Fig. 17 and Fig. 13. Mode one is the completely "in-phase" mode and mode five is the completely "out-of-phase" mode. There are a number of such similarities between linear and nonlinear normal modes.

the approximation, since we wanted to demonstrate the stability of these modes.

Thus, we have shown that the same resonant systems discussed in Sec. II which shared energy for certain initial conditions do not share energy for other initial conditions. We have also established that one may have a small amount of quantitative confidence in these perturbation schemes.

If one accepts the stability of these normal modes, then he concludes that these systems are not ergodic. A system started in a given state may cover a large section of the energy surface, but regions of nonzero measure will be left uncovered. In all the current attempts to derive a master equation for the approach to equilibrium, certain assumptions about initial conditions have been made. We suggest that these initial-condition assumptions are not just mathematical conveniences. They are needed in order to minimize the importance of "normal mode" solutions which form a set of nonzero measure in the totality of solutions. The physical justification for eliminating these normal modes rests with the fact that it is easy to excite one end of a system, but it is hard to excite just one normal mode.

Secs. II and III have now shown that the nonlinear systems obeying Hamiltonian (1) have certain striking similarities to linear systems. We conclude this section by asking whether or not Hamiltonian (1) has, in common with linear systems, N constants of the motion analytic in q, p, and α. Since the weakly coupled oscillators for a nonresonant system do not share very much energy, we see that for a nonresonant system the E_k of Eq. (12) are N constants of the motion; i.e., they are the zeroth-order term in a power series expansion of the constants of the motion.[12] In view of the fact that normal modes exist even in the presence of resonance, one would be almost surprised if N analytic constants of the motion did not also continue to exist in the presence of resonance. The question of analytic constants for 2-oscillator systems has been considered by Whittaker,[12] and he succeeds in finding a formal, analytic integral distinct from the total energy. Cherry[13] considers the case of N oscillators. He proves that there is always at least one analytic integral distinct from the energy. Thus far, we have not been able to improve on these results other than to remark that their proof goes most easily using the canonical formalism of Birkhoff.[14] However, using a computer we can observe whether or not this second, formal constant of the motion is really constant. Before discussing the computer result, we illustrate a simple method for finding this second constant of the motion.

Consider the case of the resonant 2-oscillator Hamiltonian

$$H = \tfrac{1}{2}(p_1^2 + p_2^2 + q_1^2 + 4q_2^2) + 2^{-\frac{1}{2}}\alpha(q_1^2 q_2 - q_2^3). \quad (21)$$

Introduce the canonical transformation

$$q_1 = (2J_1)^{\frac{1}{2}} \cos\theta_1, \qquad q_2 = (J_2)^{\frac{1}{2}} \cos\theta_2,$$
$$p_1 = -(2J_1)^{\frac{1}{2}} \sin\theta_1, \quad p_2 = -2(J_2)^{\frac{1}{2}} \sin\theta_2. \quad (22)$$

Hamiltonian (21) then becomes

$$H = J_1 + 2J_2 + 2^{-\frac{1}{2}}\alpha[J_1 J_2^{\frac{1}{2}} \cos\theta_2 + \tfrac{1}{2}J_1 J_2^{\frac{1}{2}} \times \cos(\theta_2 + 2\theta_1) + \tfrac{1}{2}J_1 J_2^{\frac{1}{2}} \cos(\theta_2 - 2\theta_1) - (\tfrac{3}{4}J_2^{\frac{1}{2}}) \cos\theta_2 - (\tfrac{1}{4}J_2^{\frac{1}{2}}) \cos 3\theta_2]. \quad (23)$$

In zeroth order we expect $\theta_1 = t$, $\theta_2 = 2t$; thus among the first-order terms in Eq. (23), all will be rapidly varying except the term involving $\cos(\theta_2 - 2\theta_1)$. A quick check reveals that this slowly varying term has zero Poisson bracket with the term $(J_1 + 2J_2)$. Thus, we anticipate that one can find a constant of the motion Φ of the form

$$\Phi = J_1 J_2^{\frac{1}{2}} \cos(\theta_2 - 2\theta_1) + \alpha\Phi_1 + \cdots. \quad (24)$$

[12] E. T. Whittaker, *Analytic Dynamics* (Dover Publications, Inc., New York, 1944), Chap. 16.
[13] T. M. Cherry, Proc. Cambridge Phil. Soc. **22**, 510 (1925).
[14] G. D. Birkhoff, *Dynamical Systems* (American Mathematical Society Colloquium Publications, New York, 1927), p. 82. The generalization of Birkhoff's theorem needed for this purpose is stated as Lemma 4 in a paper by J. Moser, Commun. Pure Appl. Math. **11**, 81 (1958).

Transforming back to the q, p variables, we find

$$\Phi = \tfrac{1}{2}(q_1^2 q_2 - q_2 p_1^2 + q_1 p_1 p_2) + \alpha\Phi_1 + \cdots, \quad (25)$$

the desired analytic constant of the motion. In passing, we note that the two constants of the motion which can be determined for a 2-oscillator system allow one to determine[12] the complete solution. Using this method one can obtain quantitative agreement with the curves plotted in Figs. 1 through 4. In particular, one can show that the $\omega_1 = \omega_2$ resonance yields little energy sharing.

Using the method outlined above, we found the slowly varying first-order terms in Hamiltonian (14) for the $\omega_1 = \frac{1}{2}\omega_2 = \frac{1}{3}\omega_3 = \frac{1}{4}\omega_4 = \frac{1}{5}\omega_5$ resonance. The sum of these terms, expressed in the q, p coordinates, is just the desired constant of the motion Φ. Using the same initial conditions that led to the complete energy sharing of Fig. 8, the computer calculated this Φ. The results are plotted in Fig. 18; the plot of Φ is labeled "resonant coupling energy" since it is these slowly varying terms in the coupling which give rise to the small denominators in the Wigner–Brillouin series. For comparison, the rapidly varying total coupling energy is plotted. Note that the total coupling energy seldom exceeds 10% of the total energy $E_T = 5.0$. Figure 18 plots these two curves for 80 time units, where $T_5 = (2\pi/\omega_5) = 3.14$. We have plotted these curves out to $t = 1000$, and the segment presented is typical.

IV. SINGLE-OSCILLATOR ENERGY DISTRIBUTION FUNCTION

In this section we compare the behavior of a resonant 5-oscillator linear system with that of the resonant 5-oscillator nonlinear system obeying Hamiltonian (14) with $\omega_1 = \frac{1}{2}\omega_2 = \frac{1}{3}\omega_3 = \frac{1}{4}\omega_4 = \frac{1}{5}\omega_5 = 0.4$. The linear Hamiltonian used was

$$H = \frac{1}{2}\sum_{k=1}^{5}(p_k^2 + q_k^2) - \alpha(q_1 q_2 + q_2 q_3 + q_3 q_4 + q_4 q_5). \quad (26)$$

For both systems, five units of kinetic energy were initially given to particle number one. A plot of E_1 for both systems appears in Fig. 19, and the plot of E_5 for both systems appears in Fig. 20. For the

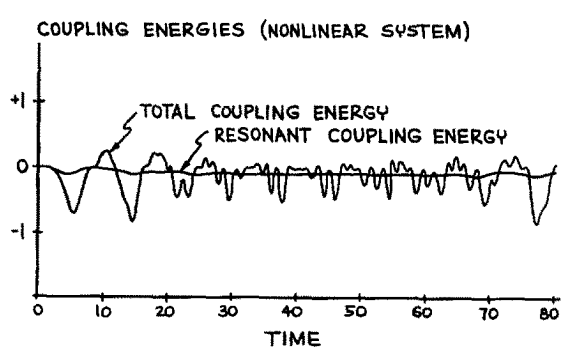


FIG. 18. The curve labeled "resonant coupling energy" is a plot of a constant of the motion distinct from the total energy for the same 5-oscillator system as in Fig. 8. This constant was predicted by perturbation theory and is essentially Whittaker's Adelphic integral. For comparison the total coupling energy is plotted. Note that the total coupling energy seldom exceeds 10% of the total energy of five units.

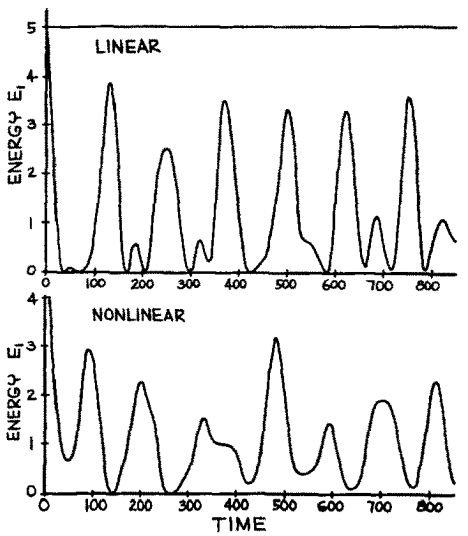


FIG. 19. A comparison of the E_1 curves for a resonant 5-oscillator linear and nonlinear system. For both systems initially, five units of kinetic energy were given to oscillator 1. The nonlinear curve is a continuation of the curve in Fig. 8.

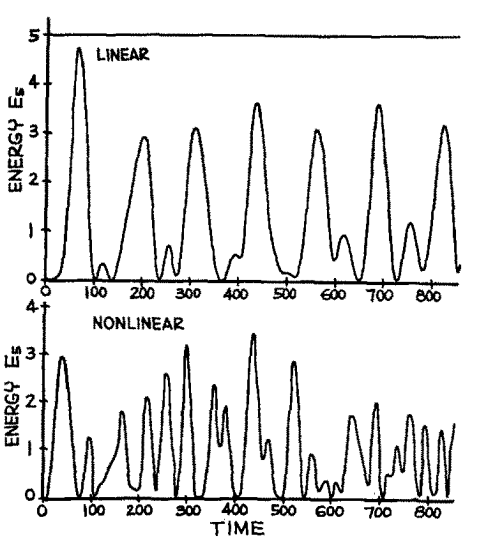


FIG. 20. A comparison of the E_5 curves for the same system as in Fig. 19.

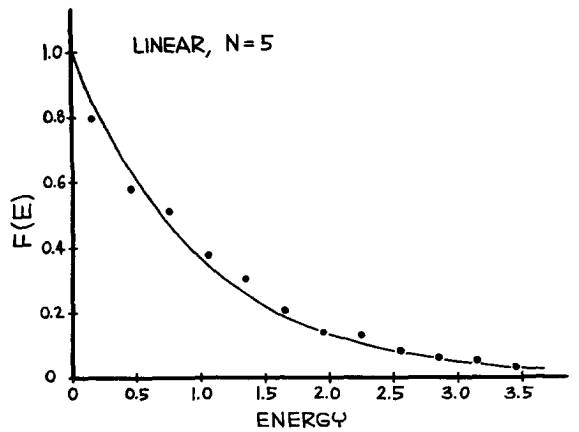


FIG. 21. The dots represent the $f(E)$ data points for the linear system computed from the E_k curves. Figures 19 and 20 present two of these curves. The solid curve is a plot of $f(E) = e^{-E}$ predicted by statistical mechanics.

linear system all the uncoupled periods are equal to 6.28. For the nonlinear system, $T_1 = (2\pi/\omega_1) = 15.7$, and $T_5 = (2\pi/\omega_5) = 3.14$. These curves present at least two interesting features. First, both systems appear to reach a steady state after the elapse of only a small number of uncoupled periods; and second, these E_k curves show a rather striking similarity. Certainly there are differences, but bearing in mind the totally different coupling involved, the similarities are more impressive than the differences.

In view of this similarity, it is interesting to compare the single-oscillator energy distribution densities $f(E)$ for these two systems. Here we define $f(E)$ to be the fractional amount of time a single oscillator spends in the energy interval $(E, E + dE)$. First, we computed $f(E)$ data points using $\Delta E = 0.3$ for each oscillator. Since we found that all the $f(E)$ plots were more or less the same, for presentation we took the average of the linear $f(E)$ data and plotted the data points as the dots shown in Fig. 21. The composite $f(E)$ data points for the nonlinear system are shown as dots in Fig. 22.

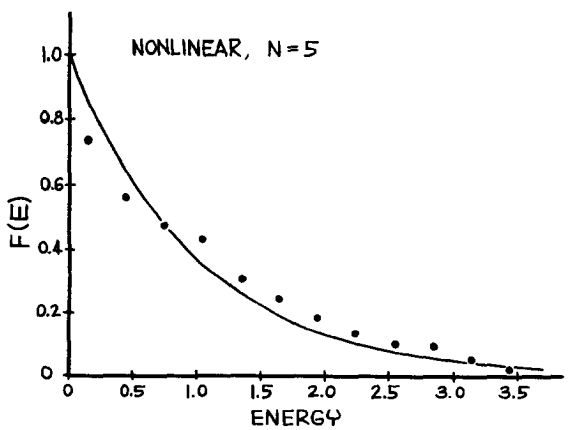


FIG. 22. Plot of the same information as Fig. 21 for the resonant nonlinear oscillator system.

According to equilibrium statistical mechanics, the single oscillator energy distribution density for a system at temperature T is

$$f(E) = (kT)^{-1}e^{-E/kT}, \tag{27}$$

where $kT = \langle E \rangle$ and $\langle E \rangle$ equals the average energy of a single oscillator. For these 5-oscillator systems, $kT = \langle E \rangle = E_T/5 = 1.0$. Thus according to statistical mechanics, $f(E) = e^{-E}$. The solid curve in Figs. 21 and 22 is a plot of e^{-E}.

Thus, the data points represent the time average $f(E)$ while the solid curve represents the phase average $f(E)$. It is interesting to note the agreement of these two averages for both these systems. It thus appears that for these systems, four oscillators form an adequate heat bath for the fifth oscillator.

V. CONCLUSIONS

It is believed that a thorough understanding of oscillator systems obeying Hamiltonians of the form

$$H = \tfrac{1}{2}\sum_{k=1}^{N}(p_k^2 + \omega_k^2 q_k^2) + \alpha \sum_{i,k,l=1}^{N} A_{ikl} q_i q_k q_l, \tag{28}$$

might shed light on some of the unresolved questions of statistical mechanics. This paper has investigated those properties of these systems which, it is hoped, are basic to a complete understanding. Our work relies heavily on numerical solutions to the equations of motion. Clearly, a computer can provide these solutions only for selected initial conditions and for relatively short time intervals. The analytical methods we have used to imply the long-time behavior are open to serious convergence questions, and it remains to be shown whether or not they provide an adequate treatment of the small-denominator problem. Nonetheless, whenever we have been provided with a check, we have found agreement between analytical and numerical solution. Consequently, it is with guarded optimism that we offer our conclusions.

Oscillator systems weakly coupled by nonlinear forces can share energy only if there is resonance. For weak coupling, resonance will occur provided

$$\sum n_k \omega_k \lesssim \alpha \tag{29}$$

for certain values of the n_k determined by the particular coupling. Starting from certain initial conditions, some of the resonances predicted using

Eq. (29) will lead to equipartition of energy. It is interesting to note that the time required to achieve a more or less uniform spread of energy is surprisingly short. In this paper, we have primarily investigated resonances leading to complete energy sharing. Further study of the weak resonances is needed.

Resonant oscillator systems with N oscillators have N normal modes. The computer solutions indicate that these normal modes are stable. The normal-mode frequency spectrum and the normal-mode node count is quite similar to that for linear systems. We have verified that these oscillator systems have at least one constant of the motion analytic in q, p, and α other than the total energy. We suspect that there are N analytic constants, but we have not been able to find all of them. However, one could easily explain our failure since it is possible that these constants are analytic in q and p but nonanalytic in α.[15] In any event, it is unlikely that these oscillator systems are ergodic. Thus, weakly coupled oscillator systems are seen to share many of the properties of linear systems. This observation has been reinforced by showing that the single-oscillator energy distribution densities for a particular choice of linear and nonlinear system are quite similar. Finally, we have shown that the single-oscillator energy distribution for both systems has the Boltzmann form predicted by statistical mechanics.

Our efforts in this paper have been directed toward laying the groundwork for further study of irreversible processes. We have been able to demonstrate that certain simple phase functions approach equilibrium. It remains to be shown that a greater variety of phase functions approach equilibrium. The possibility of providing an understanding of lattice thermal conductivity using these models is a particularly interesting prospect. There are a host of unanswered mathematical questions concerning these systems. The need for a more adequate perturbation scheme is especially pressing. The partial success of the perturbation methods we have used indicates that these systems may eventually succumb completely to analysis; although a computer is likely to remain a useful adjunct to analysis for some time to come. In conclusion, it is our belief that knowledge of these simple nonlinear systems will measurably increase our understanding of irreversibility and that the problems these systems present, though difficult, are on the verge of solution.

[15] T. M. Cherry, Proc. Cambridge Phil. Soc. **22**, 287 (1925).

APPENDIX I

In this appendix we derive the Wigner–Brillouin solution through order α^2 of the equations

$$\begin{aligned} \ddot{q}_1 + \omega_1^2 q_1 &= -\sqrt{2}\,\alpha q_1 q_2, \\ \ddot{q}_2 + \omega_2^2 q_2 &= -2^{-\frac{1}{2}}\alpha(q_1^2 - 3q_2^2). \end{aligned} \tag{30}$$

We assume a solution of the form

$$\begin{aligned} q_1 &= A_1 \cos\tau_1 + \alpha q_{11} + \alpha^2 q_{12} + \cdots, \\ q_2 &= A_2 \cos\tau_2 + \alpha q_{21} + \alpha^2 q_{22} + \cdots, \end{aligned} \tag{31}$$

where $\tau_1 = \Omega_1 t + \theta_1$ and $\tau_2 = \Omega_2 t + \theta_2$ and where Ω_1 and Ω_2 are frequencies to be determined. Substituting Eq. (31) into Eq. (30), we obtain

$$(-\Omega_1^2 + \omega_1^2)A_1 \cos\tau_1 + \alpha\{\ddot{q}_{11} + \omega_1^2 q_{11} + \sqrt{2}\,q_{10}q_{20}\} + \alpha^2\{\ddot{q}_{12} + \omega_1^2 q_{12} + \sqrt{2}(q_{10}q_{21} + q_{11}q_{20})\} = 0, \tag{32a}$$

$$\begin{aligned} &(-\Omega_2^2 + \omega_2^2)A_2 \cos\tau_2 \\ &\quad + \alpha\{\ddot{q}_{21} + \omega_2^2 q_{21} + 2^{-\frac{1}{2}}(q_{10}^2 - 3q_{20}^2)\} \\ &\quad + \alpha^2\{\ddot{q}_{22} + \omega_2^2 q_{21} \\ &\quad + \sqrt{2}\,(q_{10}q_{11} - 3q_{20}q_{21})\} = 0, \end{aligned} \tag{32b}$$

where $q_{10} = A_1 \cos\tau_1$ and $q_{20} = A_2 \cos\tau_2$.

Setting the coefficient of α in Eq. (32) to zero yields

$$\begin{aligned} \ddot{q}_{11} + \omega_1^2 q_{11} &= -2^{-\frac{1}{2}}[A_1 A_2 \cos(\tau_1 + \tau_2) \\ &\qquad + A_1 A_2 \cos(\tau_1 - \tau_2)], \\ \ddot{q}_{21} + \omega_2^2 q_{21} &= -2^{-\frac{1}{2}}[(A_1^2 - 3A_2^2) \\ &\qquad + A_1^2 \cos 2\tau_1 - 3A_2^2 \cos 2\tau_2]. \end{aligned} \tag{33}$$

The solution of Eq. (33) may be written

$$\begin{aligned} q_{11} &= -\frac{A_1 A_2 \cos(\tau_1 + \tau_2)}{\sqrt{2}\,[\omega_1^2 - (\Omega_1 + \Omega_2)^2]} \\ &\qquad - \frac{A_1 A_2 \cos(\tau_1 - \tau_2)}{\sqrt{2}\,[\omega_1^2 - (\Omega_1 - \Omega_2)^2]}, \\ q_{21} &= -\frac{(A_1^2 - 3A_2^2)}{2\sqrt{2}\,\omega_2^2} - \frac{A_1^2 \cos 2\tau_1}{2\sqrt{2}\,[\omega_2^2 - 4\Omega_1^2]} \\ &\qquad + \frac{3A_2^2 \cos 2\tau_2}{2\sqrt{2}\,[\omega_2^2 - 4\Omega_2^2]}. \end{aligned} \tag{34}$$

Setting the coefficient of α^2 in Eq. (32) equal to zero yields

$$\ddot{q}_{12} + \omega_1^2 q_{12} = \left\{\frac{(A_1^2 - 3A_2^2)A_1 \cos\tau_1}{2\omega_2^2} + \frac{A_1^3 \cos\tau_1}{4[\omega_2^2 - 4\Omega_1^2]} + \frac{A_1^3 \cos 3\tau_1}{4[\omega_2^2 - 4\Omega_1^2]}\right.$$

$$- \frac{3A_1A_2^2 \cos(\tau_1 + 2\tau_2)}{4[\omega_2^2 - 4\Omega_2^2]} - \frac{3A_1A_2^2 \cos(\tau_1 - 2\tau_2)}{4[\omega_2^2 - 4\Omega_2^2]}$$

$$+ \frac{A_1A_2^2 \cos \tau_1}{2[\omega_1^2 - (\Omega_1 + \Omega_2)^2]} + \frac{A_1A_2^2 \cos(\tau_1 + 2\tau_2)}{2[\omega_1^2 - (\Omega_1 + \Omega_2)^2]}$$

$$+ \frac{A_1A_2^2 \cos \tau_1}{2[\omega_1^2 - (\Omega_1 - \Omega_2)^2]} + \frac{A_1A_2^2 \cos(\tau_1 - 2\tau_2)}{2[\omega_1^2 - (\Omega_1 - \Omega_2)^2]} \Bigg\}, \tag{35a}$$

$$\ddot{q}_{22} + \omega_2^2 q_{22} = \Bigg\{ + \frac{A_1^2A_2 \cos \tau_2}{2[\omega_1^2 - (\Omega_1 + \Omega_2)^2]}$$

$$+ \frac{A_1^2A_2 \cos(2\tau_1 + \tau_2)}{2[\omega_1^2 - (\Omega_1 + \Omega_2)^2]} + \frac{A_1^2A_2 \cos \tau_2}{2[\omega_1^2 - (\Omega_1 - \Omega_2)^2]}$$

$$+ \frac{A_1^2A_2 \cos(2\tau_1 - \tau_2)}{2[\omega_1^2 - (\Omega_1 - \Omega_2)^2]} - \frac{3(A_1^2 - 3A_2^2)A_2 \cos \tau_2}{2\omega_2^2}$$

$$- \frac{3A_1^2A_2 \cos(2\tau_1 - \tau_2)}{4[\omega_2^2 - 4\Omega_1^2]} - \frac{3A_1^2A_2 \cos(2\tau_1 + \tau_2)}{4[\omega_2^2 - 4\Omega_1^2]}$$

$$+ \frac{9A_2^3 \cos \tau_2}{4[\omega_2^2 - 4\Omega_2^2]} + \frac{9A_2^3 \cos 3\tau_2}{4[\omega_2^2 - 4\Omega_2^2]} \Bigg\}. \tag{35b}$$

The terms involving $\cos \tau_1$ in Eq. (35a) and the terms involving $\cos \tau_2$ in Eq. (35b) must be eliminated since they will lead to small denominators of the type $(\omega_1^2 - \Omega_1^2)$ and $(\omega_2^2 - \Omega_2^2)$. The zeroth-order terms in Eq. (32) are set up for the specific purpose of eliminating these terms. The $\cos \tau_1$ in Eq (35a) are grouped with the zeroth-order term in Eq. (32a); the $\cos \tau_2$ terms of Eq. (35b) are grouped with the zeroth-order terms of Eq. (32b). Ω_1 and Ω_2 are then chosen such that they eliminate these terms. Carrying out this procedure leads to Eqs. (7c) and (7d) of Sec. II. Having eliminated the dangerous terms from Eqs. (35a) and (35b), we may solve for q_{12} and q_{22} just as we solved Eq. (33). Carrying out these steps and putting the results into Eq. (31) leads to Eqs. (7a) and (7b) of Sec. II. Clearly, this procedure can be continued to arbitrary order in α. We eliminate dangerous $\cos \tau_1$ and $\cos \tau_2$ terms in any order by grouping them with the zeroth-order terms, and then proceed to solve for the q_{1n} and q_{2n} in the usual fashion.

Representation Theory for Nonunitary Groups

JOHN O. DIMMOCK*
Raytheon Research Division, Waltham, Massachusetts
(Received 6 May 1963)

The representation theory for nonunitary groups is formulated following the same development used in the case of unitary groups, and the orthogonality relations for the corepresentation matrices are obtained. The general considerations of Wigner are used to show how one may obtain the irreducible corepresentations of a nonunitary group from the irreducible representations of its unitary subgroup.

I. INTRODUCTION

THE purpose of this work is to formulate the representation theory of nonunitary groups, that is, of groups containing both unitary and antiunitary operators, following the same development used in the case of unitary groups. We begin with a discussion of the structure of nonunitary groups and of some of the properties of unitary and antiunitary operators.[1] In Sec. II the formulation is developed, and in Sec. III we discuss the relations between the irreducible corepresentations of nonunitary groups and the irreducible representations of their unitary subgroups. Consider the group G which contains both unitary and antiunitary operators. These operators will be denoted by $\mathbf{u}$ and $\mathbf{a}$, respectively. Further, it is convenient to express the antiunitary operators as $\mathbf{a} = \mathbf{v}\theta$ where $\mathbf{v}$ is unitary and θ is antiunitary. No loss of generality results from the identification of θ with the operation of time reversal. It can be shown that the product of two unitary operators is unitary, the product of two antiunitary operators is also unitary and the product of an antiunitary operator and a unitary operator is antiunitary. Consequently, nonunitary groups contain equal numbers of unitary and antiunitary operators, and the unitary operators form an invariant subgroup H of index two. The antiunitary operators form the coset of H such that by choosing an arbitrary one of these $\mathbf{a}_0$ we can write $G = H + H\cdot\mathbf{a}_0$. The coset is generated by the products of the members of H with $\mathbf{a}_0$, and G can be specified by H and $\mathbf{a}_0$.

Before proceeding let us recall what is meant by a unitary and an antiunitary operator. Consider a system which is invariant under the operations of G. If Φ and Ψ are state functions of the system, $\mathbf{u}$ is unitary if

$$(\mathbf{u}\Psi, \mathbf{u}\Phi) = (\Psi, \Phi), \tag{1}$$

and $\mathbf{a}$ is antiunitary if

$$(\mathbf{a}\Psi, \mathbf{a}\Phi) = (\Phi, \Psi) = (\Psi, \Phi)^*. \tag{2}$$

If Φ is expanded in terms of the eigenstates of the system,

$$\Phi = \sum_\alpha a_\alpha \phi_\alpha,$$

then one can show[1] that

$$\mathbf{u}\Phi = \sum_\alpha a_\alpha \mathbf{u}\phi_\alpha, \tag{3}$$

and

$$\mathbf{a}\Phi = \sum_\alpha a_\alpha^* \mathbf{a}\phi_\alpha. \tag{4}$$

That is, unitary operators are linear and antiunitary operators are antilinear.

II. ORTHOGONALITY RELATIONS FOR IRREDUCIBLE COREPRESENTATIONS OF NONUNITARY GROUPS

Consider that $\phi_\alpha^{(i)}$ is the αth basis function of the ith irreducible corepresentation of G. The corepresentation matrices may be defined by the following[1]:

$$\begin{aligned} \mathbf{u}\phi_\alpha^{(i)} &= D^{(i)}(\mathbf{u})_{\beta\alpha}\phi_\beta^{(i)}, \\ \mathbf{a}\phi_\alpha^{(i)} &= D^{(i)}(\mathbf{a})_{\beta\alpha}\phi_\beta^{(i)}, \end{aligned} \tag{5}$$

adopting the usual convention of an understood summation over repeated indices. Using Eqs. (3), (4), and (5) one finds that the corepresentation matrices multiply as follows:

$$\begin{aligned} \mathbf{D}^{(i)}(\mathbf{u}_1)\,\mathbf{D}^{(i)}(\mathbf{u}_2) &= \mathbf{D}^{(i)}(\mathbf{u}_1\mathbf{u}_2), \\ \mathbf{D}^{(i)}(\mathbf{u})\,\mathbf{D}^{(i)}(\mathbf{a}) &= \mathbf{D}^{(i)}(\mathbf{ua}), \\ \mathbf{D}^{(i)}(\mathbf{a})\,\mathbf{D}^{(i)}(\mathbf{u})^* &= \mathbf{D}^{(i)}(\mathbf{au}), \\ \mathbf{D}^{(i)}(\mathbf{a}_1)\,\mathbf{D}^{(i)}(\mathbf{a}_2)^* &= \mathbf{D}^{(i)}(\mathbf{a}_1\mathbf{a}_2). \end{aligned} \tag{6}$$

* Present address M.I.T. Lincoln Laboratory, Lexington, Massachusetts.

[1] E. P. Wigner, *Group Theory* (Academic Press Inc., New York, 1959), Chap. 26. The first part of Sec. III summarizes some of the results of this chapter. See also E. P. Wigner, J. Math. Phys. 1, 409, 414 (1960).

The **D** matrices do not multiply exactly as do the operators because of the complex-conjugate signs in these relations, and for this reason are referred to as corepresentation matrices.

Next it is important to see how the **D** matrices transform under a transformation of basis functions. Consider a new set of basis functions $\psi_\alpha^{(i)}$ related to the $\phi_\alpha^{(i)}$ by some transformation matrix **V** such that

$$\psi_\alpha^{(i)} = V_{\beta\alpha}\phi_\beta^{(i)}.$$

Then the corepresentation matrices taken with respect to the $\psi_\alpha^{(i)}$ are related to the **D** matrices by the transformation

$$\begin{aligned} \mathbf{B}^{(i)}(\mathbf{u}) &= \mathbf{V}^{-1}\mathbf{D}^{(i)}(\mathbf{u})\mathbf{V} \\ \mathbf{B}^{(i)}(\mathbf{a}) &= \mathbf{V}^{-1}\mathbf{D}^{(i)}(\mathbf{a})\mathbf{V}^* \end{aligned} \quad . \tag{7}$$

This is therefore the form taken by a similarity transformation of corepresentation matrices of nonunitary groups and the two sets of matrices **D** and **B** are considered to be equivalent.

We are now in a position to develop the corepresentation theory of nonunitary groups following the same arguments used in the case of unitary groups.[2]

Theorem I: Any corepresentation by matrices with nonvanishing determinants can be transformed into a corepresentation by unitary matrices through a similarity transformation.

The corepresentation matrices of nonunitary groups can thus be chosen to be unitary just as in the case of unitary groups. The proof of Theorem I is essentially the same as in the case of unitary groups and consequently will not be given here.[2] It is, however, easy to see that the corepresentation matrices will be unitary if the basis functions form an orthonormal set, and since any linearly independent set of functions can be transformed into an orthonormal set, we can choose unitary representation matrices regardless of whether the group in question is unitary or not.

Theorem II: If there exists a Hermitian matrix **M** such that

$$\begin{aligned} \mathbf{M}\mathbf{D}^{(i)}(\mathbf{u}) &= \mathbf{D}^{(i)}(\mathbf{u})\mathbf{M}, \\ \mathbf{M}\mathbf{D}^{(i)}(\mathbf{a}) &= \mathbf{D}^{(i)}(\mathbf{a})\mathbf{M}^*, \end{aligned} \tag{8}$$

for all the matrices of an irreducible corepresentation, then **M** is a constant matrix. If there exists a nonconstant Hermitian matrix which satisfies these relations for all the matrices of a corepresentation then the corepresentation is reducible. If there exists none then it is irreducible.

This theorem differs from the one which is proven for unitary groups in the presence of the complex conjugate, which arises because of Eq. (7), and in that **M** is restricted to Hermitian matrices. The restriction on **M** is important in the derivation of the orthogonality relations between corepresentations and we therefore show here how it comes about. Taking the Hermitian adjoint of Eq. (8) we have

$$\mathbf{M}^\dagger\mathbf{D}^{(i)}(\mathbf{u})^\dagger = \mathbf{D}^{(i)}(\mathbf{u})^\dagger\mathbf{M}^\dagger.$$

Multiplying from the left and right by $\mathbf{D}^{(i)}(\mathbf{u})$, and using the fact that these matrices are unitary, we have

$$\mathbf{M}^\dagger\mathbf{D}^{(i)}(\mathbf{u}) = \mathbf{D}^{(i)}(\mathbf{u})\mathbf{M}^\dagger.$$

It then follows that the Hermitian matrices $\mathbf{H}_1 = \frac{1}{2}(\mathbf{M} + \mathbf{M}^\dagger)$ and $\mathbf{H}_2 = \frac{1}{2}i(\mathbf{M} - \mathbf{M}^\dagger)$ commute with $\mathbf{D}^{(i)}(\mathbf{u})$. In the case of unitary groups it is then sufficient to show that every Hermitian matrix which commutes with $\mathbf{D}^{(i)}(\mathbf{u})$ is a constant matrix, since if $\mathbf{H}_1$ and $\mathbf{H}_2$ are constant matrices, so is $\mathbf{M} = \mathbf{H}_1 - i\mathbf{H}_2$. For nonunitary groups this still holds for the $\mathbf{D}^{(i)}(\mathbf{u})$ matrices. However returning to Eq. (8) we can obtain

$$\mathbf{M}^\dagger\mathbf{D}^{(i)}(\mathbf{a}) = \mathbf{D}^{(i)}(\mathbf{a})\mathbf{M}^{*\dagger}$$

as above, and also

$$\mathbf{H}_1\mathbf{D}^{(i)}(\mathbf{a}) = \mathbf{D}^{(i)}(\mathbf{a})\mathbf{H}_1^*;$$

but

$$\mathbf{H}_2\mathbf{D}^{(i)}(\mathbf{a}) = -\mathbf{D}^{(i)}(\mathbf{a})\mathbf{H}_2^*.$$

The proof of Theorem II for **M** Hermitian is essentially the same as in the case of unitary groups,[2] but whereas in the case of unitary groups this proof suffices to establish the theorem in general, this is no longer the case for nonunitary groups and **M** is in this case required to be Hermitian.

Theorem III: Consider two irreducible corepresentations of the same group $\mathbf{D}^{(i)}$ and $\mathbf{D}^{(j)}$ of dimension l_i and l_j, respectively. If there exists a matrix **M** of l_i rows and l_j columns such that

$$\begin{aligned} \mathbf{M}\mathbf{D}^{(i)}(\mathbf{u}) &= \mathbf{D}^{(j)}(\mathbf{u})\mathbf{M}, \\ \mathbf{M}\mathbf{D}^{(i)}(\mathbf{a}) &= \mathbf{D}^{(j)}(\mathbf{a})\mathbf{M}^*, \end{aligned} \tag{9}$$

then for $l_i \neq l_j$, **M** is a null matrix; for $l_i = l_j$, **M** is either a null matrix or it has a nonvanishing de-

[2] See for example E. P. Wigner, *Group Theory*. (reference 1), Chap. 9. The development of Sec. II follows that given here for unitary groups.

terminant. In the latter case $\mathbf{M}$ has an inverse and the two representations are equivalent.

The proof follows that given for unitary groups.[2]

Theorem IV: If $\mathbf{D}^{(i)}$ and $\mathbf{D}^{(j)}$ are two inequivalent, irreducible, unitary corepresentations of G then

$$\sum_k D^{(i)}(\mathbf{u}_k)_{\alpha\mu} D^{(j)}(\mathbf{u}_k)^*_{\beta\nu} = 0, \qquad \sum_k D^{(i)}(\mathbf{a}_k)_{\alpha\mu} D^{(j)}(\mathbf{a}_k)^*_{\beta\nu} = 0. \tag{10}$$

For the elements of a single, irreducible, unitary corepresentation, we have

$$\sum_k \{D^{(i)}(\mathbf{u}_k)_{\alpha\mu} D^{(i)}(\mathbf{u}_k)^*_{\beta\nu} + D^{(i)}(\mathbf{a}_k)_{\alpha\nu} D^{(i)}(\mathbf{a}_k)^*_{\beta\mu}\} = \frac{n}{l_i}\,\delta_{\alpha\beta}\delta_{\mu\nu}, \tag{11}$$

where n is the number of operators in G and l_i is the dimension of $\mathbf{D}^{(i)}$. In the sum k goes from 1 to $\frac{1}{2}n$, $\mathbf{u}_k$ runs over the members of the unitary subgroup and $\mathbf{a}_k = \mathbf{u}_k \cdot \mathbf{a}_0$ runs over the members of the antiunitary coset. Notice in Eq. (11) that the indices μ and ν are interchanged in the two parts of the sum.

Proof: The matrix

$$\mathbf{M} = \sum_k \{\mathbf{D}^{(i)}(\mathbf{u}_k)\mathbf{X}\mathbf{D}^{(j)}(\mathbf{u}_k)^\dagger + \mathbf{D}^{(i)}(\mathbf{a}_k)\mathbf{X}^*\mathbf{D}^{(j)}(\mathbf{a}_k)^\dagger\} \tag{12}$$

satisfies Eq. (9) for any arbitrary matrix $\mathbf{X}$ and consequently is a null matrix if $\mathbf{D}^{(i)}$ and $\mathbf{D}^{(j)}$ are inequivalent. We can set $X_{\mu\nu} = 1$ and all other elements zero to obtain

$$\sum_k \{D^{(i)}(\mathbf{u}_k)_{\alpha\mu} D^{(j)}(\mathbf{u}_k)^*_{\beta\nu} + D^{(i)}(\mathbf{a}_k)_{\alpha\mu} D^{(j)}(\mathbf{a}_k)^*_{\beta\nu}\} = 0. \tag{13}$$

We can also choose $X_{\mu\nu} = i$ and all other elements zero to obtain

$$\sum_k \{D^{(i)}(\mathbf{u}_k)_{\alpha\mu} D^{(j)}(\mathbf{u}_k)^*_{\beta\nu} - D^{(i)}(\mathbf{a}_k)_{\alpha\mu} D^{(j)}(\mathbf{a}_k)^*_{\beta\nu}\} = 0. \tag{14}$$

Adding and subtracting Eqs. (13) and (14) we get Eq. (10) and establish the first half of Theorem IV. Consider now

$$\mathbf{M} = \sum_k \{\mathbf{D}^{(i)}(\mathbf{u}_k)\mathbf{X}\mathbf{D}^{(i)}(\mathbf{u}_k)^\dagger + \mathbf{D}^{(i)}(\mathbf{a}_k)\mathbf{X}^*\mathbf{D}^{(i)}(\mathbf{a}_k)^\dagger\} \tag{15}$$

which satisfies Eq. (8). If $\mathbf{X}$ is Hermitian then $\mathbf{M}$ is Hermitian and a multiple of the unit matrix. Let us first set $X_{\mu\nu} = X_{\nu\mu} = 1$ for $\mu \neq \nu$ and all other elements zero. Then

$$\sum_k \{D^{(i)}(\mathbf{u}_k)_{\alpha\mu} D^{(i)}(\mathbf{u}_k)^*_{\beta\nu} + D^{(i)}(\mathbf{u}_k)_{\alpha\nu} D^{(i)}(\mathbf{u}_k)^*_{\beta\mu} + D^{(i)}(\mathbf{a}_k)_{\alpha\mu} D^{(i)}(\mathbf{a}_k)^*_{\beta\nu} + D^{(i)}(\mathbf{a}_k)_{\alpha\nu} D^{(i)}(\mathbf{a}_k)^*_{\beta\mu}\} = c_{\mu\nu}\delta_{\alpha\beta}, \tag{16}$$

where $c_{\mu\nu}$ is independent of α and β. To determine $c_{\mu\nu}$, set $\alpha = \beta$ and sum from $\beta = 0$ to l_i. This yields

$$2n\delta_{\mu\nu} = c_{\mu\nu} l_i. \tag{17}$$

If $\mu = \nu$ then we can obtain

$$\sum_k \{D^{(i)}(\mathbf{u}_k)_{\alpha\mu} D^{(i)}(\mathbf{u}_k)^*_{\beta\mu} + D^{(i)}(\mathbf{a}_k)_{\alpha\mu} D^{(i)}(\mathbf{a}_k)^*_{\beta\mu}\} = \frac{n}{l_i}\,\delta_{\alpha\beta}. \tag{18}$$

If $\mu \neq \nu$ we can set $X_{\mu\nu} = -X_{\nu\mu} = i$ and all other elements zero. Then

$$\sum_k \{D^{(i)}(\mathbf{u}_k)_{\alpha\mu} D^{(i)}(\mathbf{u}_k)^*_{\beta\nu} - D^{(i)}(\mathbf{u}_k)_{\alpha\nu} D^{(i)}(\mathbf{u}_k)^*_{\beta\mu} - D^{(i)}(\mathbf{a}_k)_{\alpha\mu} D^{(i)}(\mathbf{a}_k)^*_{\beta\nu} + D^{(i)}(\mathbf{a}_k)_{\alpha\nu} D^{(i)}(\mathbf{a}_k)^*_{\beta\mu}\} = -iy_{\mu\nu}\delta_{\alpha\beta}. \tag{19}$$

Proceeding as above we find that $d_{\mu\nu} = 0$. Therefore, adding Eqs. (16) and (19) and using Eqs. (17) and (18), we get Eq. (11) and establish the second half of Theorem IV.

Using these results we can readily find relations between the characters of the corepresentations. For $\mathbf{D}^{(i)}$ and $\mathbf{D}^{(j)}$ inequivalent corepresentations set $\alpha = \mu$ and $\beta = \nu$ in Eq. (10) and sum over α and β to obtain

$$\sum_k \chi\{\mathbf{D}^{(i)}(\mathbf{u}_k)\}\chi\{\mathbf{D}^{(j)}(\mathbf{u}_k)\}^* = 0, \qquad \sum_k \chi\{\mathbf{D}^{(i)}(\mathbf{a}_k)\}\chi\{\mathbf{D}^{(j)}(\mathbf{a}_k)\}^* = 0, \tag{20}$$

where $\chi\{\mathbf{D}^{(i)}(\mathbf{u}_k)\} = \sum_\beta D^{(i)}(\mathbf{u}_k)_{\beta\beta}$ is the character of $\mathbf{u}_k$ in the ith corepresentation. Similarly for $i = j$ we can set $\alpha = \mu$ and $\beta = \nu$ in Eq. (11) and sum over α and β to obtain

$$\sum_k \{|\chi\{\mathbf{D}^{(i)}(\mathbf{u}_k)\}|^2 + \chi\{\mathbf{D}^{(i)}(\mathbf{a}_k^2)\}\} = n. \tag{21}$$

The orthogonality relations, Eqs. (10), (11), (20), and (21), have an inconvenient form and further development of the representation theory of nonunitary groups along these lines has so far proven untenable. In the following section we shift our attention to another approach which has been used by Wigner. In this approach one finds expressions for the $\mathbf{D}$ matrices in terms of the representation matrices of the unitary subgroup H. These latter

matrices satisfy the usual orthogonality relations and for this reason it is more convenient to discuss the irreducible corepresentations of G in terms of the irreducible representations of the unitary subgroup H.

III. DETERMINATION OF THE COREPRESENTATIONS FROM THE REPRESENTATIONS OF THE UNITARY SUBGROUP

To determine the co-representation matrices $\mathbf{D}^{(i)}$ let us select, following Wigner,[1] a set of l_i' functions which forms a basis for an irreducible representation $\mathbf{\Delta}^{(i)}(\mathbf{u})$ of the unitary subgroup H. That is,

$$\mathbf{u}\phi_\alpha^{(i)} = \Delta^{(i)}(\mathbf{u})_{\beta\alpha}\phi_\beta^{(i)} \tag{22}$$

for all $\mathbf{u}$ in H. The dimension of $\mathbf{\Delta}^{(i)}(\mathbf{u})$ is l_i'. Recall that the nonunitary group G is formed by supplementing H with a coset of operators $\mathbf{ua}_0$. The corepresentation of G corresponding to $\mathbf{\Delta}^{(i)}(\mathbf{u})$ of H is constructed in one of three ways depending on which of the following three cases is realized.

Case (a): $\mathbf{a}_0\phi_\alpha^{(i)}$ reproduces the set of functions $\phi_\alpha^{(i)}$. The irreducible corepresentation $\mathbf{D}^{(i)}$ of G corresponds to a single irreducible representation $\mathbf{\Delta}^{(i)}(\mathbf{u})$ of H and has the same dimension. In this case no new degeneracy is introduced by the coset $\mathbf{ua}_0$.

Case (b): $\mathbf{a}_0\phi_\alpha^{(i)}$ produces the set of functions $\psi_\alpha^{(i)}$ which is independent of the set $\phi_\alpha^{(i)}$ but which also forms a basis for $\mathbf{\Delta}^{(i)}(\mathbf{u})$ of H. $\mathbf{D}^{(i)}$ again corresponds to a single irreducible representation of H but has twice its dimension. In this case the degeneracy of $\mathbf{\Delta}^{(i)}(\mathbf{u})$ is doubled.

Case (c): $\mathbf{a}_0\phi_\alpha^{(i)}$ produces a set of functions $\phi_\alpha^{(j)}$ which forms a basis for the irreducible representation $\mathbf{\Delta}^{(j)}(\mathbf{u})$ of H which is inequivalent to $\mathbf{\Delta}^{(i)}$. $\mathbf{D}^{(i)}$ corresponds to two inequivalent irreducible representations of H, $\mathbf{\Delta}^{(i)}(\mathbf{u})$ and $\mathbf{\Delta}^{(j)}(\mathbf{u})$, such that in this case the antiunitary operators cause $\mathbf{\Delta}^{(i)}(\mathbf{u})$ and $\mathbf{\Delta}^{(j)}(\mathbf{u})$ to become degenerate.

Wigner[1] has obtained the irreducible corepresentations of G explicitly in terms of the irreducible representations of H in each of the three cases, and it remains only to find a method by which one can decide between the three cases given the representation $\mathbf{\Delta}^{(i)}(\mathbf{u})$ of H and one of the antiunitary operators. The following result can be obtained by using Wigner's explicit forms for the corepresentations:

$$\begin{aligned} \sum_k \chi\{\mathbf{\Delta}^{(i)}(\mathbf{a}_k^2)\} &= +n', && \text{Case (a)},\\ &= -n', && \text{Case (b)},\\ &= 0, && \text{Case (c)}, \end{aligned} \tag{23}$$

where $n' = \frac{1}{2}n$ is the number of elements in H. [A calculation similar to that which yields Eq. (23) has been performed previously.[3] In that work we found a result (Eq. 20 therein) which is different from that obtained here. Both results are correct although the present one is more convenient.] Finally it can be shown that each irreducible corepresentation of G is in a one-to-one correspondence with an irreducible representation of H except in Case (c) where the correspondence is one-to-two. Therefore knowing all the inequivalent irreducible representations of H one can find all the inequivalent irreducible corepresentations of G.

Since the antiunitary operators can be written as the product of unitary operators $\mathbf{v}$ with an antiunitary operator $\boldsymbol{\theta}$, which we can take to be the time-reversal operator, we can use the fact that

$$\boldsymbol{\theta}^2 = \omega\mathbf{E}, \tag{24}$$

where $\mathbf{E}$ is the identity operator, and

$$\omega = (-1)^N, \tag{25}$$

where N is the number of fermions in the system to further develop Eq. (23). If $\mathbf{u}$ and $\mathbf{v}$ are spatial operators they commute with $\boldsymbol{\theta}$ such that Eq. (23) becomes

$$\begin{aligned} \sum_k \chi\{\mathbf{\Delta}^{(i)}(\mathbf{u}_k\mathbf{v}_0\mathbf{u}_k\mathbf{v}_0)\} &= \omega n', && \text{Case (a)},\\ &= -\omega n', && \text{Case (b)},\\ &= 0, && \text{Case (c)}, \end{aligned} \tag{26}$$

where $\mathbf{a}_k = \mathbf{u}_k\mathbf{a}_0$ and $\mathbf{a}_0 = \mathbf{v}_0\boldsymbol{\theta}$. If $\boldsymbol{\theta}$ is a member of G we can choose $\mathbf{v}_0 = \mathbf{E}$ and Eq. (26) becomes

$$\begin{aligned} \sum_k \chi\{\mathbf{\Delta}^{(i)}(\mathbf{u}_k^2)\} &= \omega n', && \text{Case (a)},\\ &= -\omega n', && \text{Case (b)},\\ &= 0, && \text{Case (c)}, \end{aligned} \tag{27}$$

which was originally found by Frobenius and Schur.[4]

The general results derived above can be applied to space groups.[3] Identify the group G with the group of the wave vector $G_\mathbf{k}$ for some point in the Brillouin zone defined by the vector $\mathbf{k}$. Then $\mathbf{a}_0$ is an operator of the group of the wave vector and is written as

$$\mathbf{a}_0 = \mathbf{v}_0\cdot\boldsymbol{\theta} = (\boldsymbol{\varrho}_0 \mid \boldsymbol{\tau}_0)\cdot\boldsymbol{\theta}, \tag{28}$$

where $\boldsymbol{\varrho}_0$ is a point operator and $\boldsymbol{\tau}_0$ is the smallest translation operator associated with $\boldsymbol{\varrho}_0$, and since $\boldsymbol{\theta}\mathbf{k} = -\mathbf{k}$,

[3] J. O. Dimmock and R. G. Wheeler, J. Phys. Chem. Solids **23**, 729 (1962). (Hereafter referred to as I.)

[4] G. Frobenius and I. Schur, S. B. Deut. Akad. Wiss. **49**, 186 (1906).

$$\vartheta_0\mathbf{k} = -\mathbf{k} + \mathbf{K}_q, \tag{29}$$

where $\mathbf{K}_q$ is a primitive translation of the reciprocal lattice. The operators $\mathbf{u}$ are members of $G_\mathbf{k}$ and may be written as $\mathbf{u} = (\mathbf{E} \mid \mathbf{R}_n)(\vartheta \mid \tau)$, where $\mathbf{R}_n$ is a primitive translation of the direct lattice and τ is the smallest translation associated with the point operation ϑ. For all ϑ we have

$$\vartheta\mathbf{k} = \mathbf{k} + \mathbf{K}_q. \tag{30}$$

Equation (26) may be shown to reduce to

$$\sum_\vartheta \chi\{\Delta^{(i)}[(\vartheta \mid \tau)\mathbf{v}_0(\vartheta \mid \tau)\mathbf{v}_0]\} = \begin{cases} \omega M, & \text{Case (a)}, \\ -\omega M, & \text{Case (b)}, \\ 0, & \text{Case (c)}, \end{cases} \tag{31}$$

where M is equal to the number of distinct unitary point operators in $G_\mathbf{k}$, that is, to the number of terms in the sum, and with each ϑ we associate only the smallest τ. Equation (31) is equivalent to the result obtained by Herring[5] if θ is a member of the space group. [In I we obtained this result only after placing a restriction on the space group.[6] In fact Eq. (31) is obtained directly from Eq. (26) without any additional assumptions by a calculation similar to that which led to Eq. (29) of I.]

In conclusion we remark that although one can develop the representation theory for nonunitary groups, the resultant orthogonality relations between corepresentation matrices make further development of the theory untenable. However, following another approach which has been used by Wigner, we have seen that the irreducible corepresentations of a nonunitary group may readily be obtained from the irreducible representations of its unitary subgroup, which in turn may be obtained by standard means.

ACKNOWLEDGMENT

The author would like to thank Dr. G. F. Koster for suggesting this calculation and for his continued interest and advice.

[5] C. Herring, Phys. Rev. **52**, 361 (1937).

[6] Equation (30) of I was said to follow from the restriction that the time-reversal operator itself be a member of the space group. We are indebted to R. Guccione (private communication) for pointing out to us that this is incorrect. Equation (33) of that work [Eq. (31) of the present work] is, however, correct regardless of the error, and indeed is independent of the restriction. See also J. O. Dimmock and R. G. Wheeler, Phys. Rev. **127**, 391 (1962).

Convergence of Fugacity Expansions for Fluids and Lattice Gases*

OLIVER PENROSE†
Belfer Graduate School of Science, Yeshiva University, New York, New York

Upper and lower bounds are obtained for $R(V)$, the radius of convergence of the Mayer expansion $V\sum_l b_l(V)z^l$ expressing the logarithm of the classical grand partition function for a finite volume V as a power series in the fugacity z. The particles in V interact only through two-body forces whose potential $\varphi(\mathbf{r})$ satisfies $s^{-1}\sum_{i<j\leq s}\varphi(\mathbf{x}_i - \mathbf{x}_j) \geq \text{const} \equiv -\Phi$ for all s, $\mathbf{x}_1 \cdots \mathbf{x}_s$. The bounds are

$$[e^{l+2\Phi/kT}\int |e^{-\varphi(\mathbf{r})/kT} - 1|\, d^3\mathbf{r}]^{-1} \leq R(V) \leq |e^{\Phi/kT}l/(l-1)\, b_l(V)|^{1/(l-1)}$$

for any $l \geq 2$. For lattice gases the integral becomes a sum. The upper bounds, obtained from the theory of entire functions, include a subsequence converging to $R(V)$ as $l \to \infty$. The lower bound is obtained by using the Kirkwood–Salsburg integral equation to calculate upper bounds on the $b_l(V)$'s and the coefficients in the fugacity expansions of the s-particle distribution functions. For hard-core potentials some of these bounds can be strengthened. For nonnegative potentials, $1/2|b_2(V)|$ is an extra upper bound on $R(V)$. The radius of convergence of the infinite-volume series $\sum b_l z^l$ is shown to be at least $\lim_{V\to\infty} R(V)$, with equality for nonnegative potentials.

1. INTRODUCTION

MANY formal power series expansions are used in the theories of classical fluids[1] and Ising lattices[2] but there are few theorems about their convergence properties. Rigorous results regarding the convergence of these series would be of interest for the theory of phase transitions. For example, let R be the radius of convergence of the Mayer expansion for the thermodynamic pressure p of a fluid in powers of fugacity z:

$$p(z)/kT = \sum_{l=1}^{\infty} b_l z^l; \tag{1.1}$$

and suppose that only a finite number of the b_l's are negative, so that the function $p(z)$ has a singularity at $z = R$. Then if we can show[3] that (1.1) correctly gives the thermodynamic pressure for small z, we can deduce[4] that there is a phase transition either at $z = R$ or at some smaller value of z where (1.1) ceases to give the thermodynamic pressure correctly.

For particles interacting by additive central forces with a nonnegative potential $\varphi(r)$, Groeneveld[5] has obtained upper and lower bounds for R by studying the Mayer graphs of the coefficients b_l. His formulas imply that the same lower bound also holds for $R(V)$, the radius of convergence of the finite-volume analogue of (1.1),

$$\mathcal{V}^{-1}\log \Xi(z, V) = \sum_l b_l(V)z^l, \tag{1.2}$$

where the $b_l(V)$'s are the Mayer cluster integrals for a finite region V, whose volume is $\mathcal{V}$, and $\Xi(z, V)$ is the grand partition function for this region. Moreover, he shows that (1.1) correctly represents the pressure if z is between 0 and the lower bound on R, so that there can be no phase transition for z within this range. Ruelle[6] has generalized Groeneveld's lower bounds to all $\varphi(r)$ with the property

$$\sum_{i<j\leq s}\varphi(|\mathbf{x}_i - \mathbf{x}_j|) \geq -s\Phi \text{ for all } s, \mathbf{x}_1 \cdots \mathbf{x}_s, \tag{1.3}$$

where Φ is a constant, and $\sum_{i<j\leq s}$ means $\sum_{i=1}^{s-1}\sum_{j=i+1}^{s}$ for $s \geq 2$, and 0 for $s = 1$. He used the Kirkwood–Salsburg integral equation[7] and the theory of linear operators in Banach space.

In the present paper we shall obtain a new set of upper bounds on $R(V)$, forming an infinite sequence which converges from above to the limit $R(V)$. We shall also rederive Ruelle's lower bound on $R(V)$ using a method based, like his, on the Kirkwood–Salsburg integral equation, but not

* Work supported by the U. S. Air Force Office of Scientific Research under grant No. 62-64.

† On leave of absence from Imperial College, London, England.

[1] See the review article by G. W. Ford and G. E. Uhlenbeck in *Studies in Statistical Mechanics*, edited by J. de Boer and G. E. Uhlenbeck (North-Holland Publishing Company, Amsterdam, 1962), Vol. I.

[2] See the review article by C. Domb, Phil. Mag. Suppl. **9**, 149 (1960).

[3] A proof is needed since the derivation of (1.1) involves an interchange of the limiting processes $\lim_{V\to\infty}$ and $\sum_{l=1}^{\infty}$. See C. N. Yang and T. D. Lee, Phys. Rev. **87**, 404 (1952).

[4] I owe this argument to Dr. J. Lebowitz (private communication).

[5] J. Groeneveld, Phys. Letters **3**, 50 (1962).

[6] D. Ruelle, "Correlation Functions of Classical Gases," Institute for Advanced Study, Princeton, New Jersey, 1963 [to be published in Ann. Phys. (N. Y.)].

[7] J. G. Kirkwood and Z. W. Salsburg, Discussions Faraday Soc. **15**, 28 (1953); T. L. Hill, *Statistical Mechanics* (McGraw-Hill Book Company, Inc., New York, 1956), p. 251. The equation is a special case of one due to J. E. Mayer [J. Chem Phys. **15**, 187 (1947), Eq. (54′)].

requiring the theory of Banach spaces. Our method incidentally yields upper bounds on the $b_l(V)$'s and on the coefficients in the fugacity expansions of the s-particle distribution functions ($s = 1, 2, 3, \cdots$). These results apply to classes of potentials φ a little wider than those treated by Groeneveld and Ruelle, enough so to make the results applicable to lattice systems as well as to fluids. For potentials with hard cores, very similar methods yield additional upper and lower bounds which are sometimes stronger than those given by the more general methods.

2. PROPERTIES OF THE INTERACTION POTENTIAL

The results of this paper apply to systems of interacting particles whose positions $\mathbf{x}_1, \mathbf{x}_2, \cdots$ are either continuously variable or confined to a lattice (Ising model). Their potential energy is assumed to have the form[8] $\sum_{i<j} \varphi(\mathbf{x}_i - \mathbf{x}_j)$, where $\varphi(\mathbf{r})$ is symmetric under space inversion, though not necessarily spherically symmetric, and falls into one or more of the following classes:

I. *nonnegative potentials* satisfying

$$\varphi(\mathbf{r}) \geq 0 \quad \text{for all} \quad \mathbf{r}; \tag{2.1}$$

II. *hard-core potentials* satisfying

$$\varphi(\mathbf{r}) = +\infty \quad \text{if} \quad r < a, \tag{2.2}$$

$$|\varphi(\mathbf{r})| < Ar^{-\nu-\epsilon} \quad \text{if} \quad r \geq a, \tag{2.3}$$

where r is the length of $\mathbf{r}$, and a, A, and ϵ are positive constants, and ν is the number of space dimensions;

III. *potentials with finite binding energy per particle* satisfying

$$\sum_{i<j\leq s} \varphi(\mathbf{x}_i - \mathbf{x}_j) \geq -s\Phi \text{ for all } s, \mathbf{x}_1 \cdots \mathbf{x}_s, \tag{2.4}$$

where Φ is a (finite) constant depending on the form of $\varphi(\mathbf{r})$ only.

The conditions defining categories I and III are very much like the ones used by Groeneveld[5] and by Ruelle[6]. However, we do not require spherical symmetry of $\varphi(\mathbf{r})$, nor do we require absolute convergence of the integral (or sum) for b_2. Thus our results apply to lattice systems as well as to fluids, and also to long-range potentials such as $\varphi(\mathbf{r}) = \text{const}/r$. The condition (2.1) is equivalent to (2.4) with $\Phi = 0$, so that potentials in the class I are also in the class III. The remainder of this section is devoted to the proof of a theorem and corollary showing that the class II is also included in the class III.

Theorem. For any hard-core or nonnegative potential, there exists a constant Φ' such that

$$\sum_{i\leq t} \varphi(\mathbf{x}_i - \mathbf{y}) \geq -2\Phi' \tag{2.5}$$

for all t, $\mathbf{x}_1 \cdots \mathbf{x}_t$, $\mathbf{y}$ satisfying

$$\sum_{i<j\leq t} \varphi(\mathbf{x}_i - \mathbf{x}_j) < \infty. \tag{2.6}$$

Proof. For nonnegative potentials the result is trivial, with $\Phi' = 0$. For hard-core potentials, (2.2) and (2.6) imply that t spheres of radius $\frac{1}{2}a$ centered on the points $\mathbf{x}_1 \cdots \mathbf{x}_t$ do not intersect. An arbitrary sphere of radius $(r + a/2)$ has volume $(1 + 2r/a)^\nu$ times that of one of the t small spheres, and therefore cannot completely enclose more than $(1 + 2r/a)^\nu$ of them. The function $F(r)$, defined as the number of points from the set $\mathbf{x}_1 \cdots \mathbf{x}_t$ which lie inside a sphere of radius r centered on the point $\mathbf{y}$, is therefore bounded above and below by

$$0 \leq F(r) \leq (1 + 2r/a)^\nu. \tag{2.7}$$

Moreover, we may assume that

$$F(r) = 0 \quad \text{if} \quad r \leq a \tag{2.8}$$

since otherwise (2.5) is trivially satisfied, the left side being infinite. The left side of (2.5) with reversed sign has, by (2.3), an upper bound which may be written as a Stieltjes integral and transformed by partial integration:

$$\begin{aligned} -\sum_i \varphi(\mathbf{x}_i - \mathbf{y}) &< A \sum_i |\mathbf{x}_i - \mathbf{y}|^{-\nu-\epsilon} \\ &= A \int_a^\infty r^{-\nu-\epsilon}\, dF(r) = A[r^{-\nu-\epsilon}F(r)]_a^\infty \\ &\quad + (\nu + \epsilon)A \int_a^\infty r^{-\nu-\epsilon-1}F(r)\, dr. \end{aligned} \tag{2.9}$$

The integrated part vanishes at $r = a$ by (2.8), and as $r \to \infty$ by (2.7); and the integral converges, again by (2.7). Therefore, (2.5) can be satisfied, for example by taking

$$\Phi' = \tfrac{1}{2}(\nu + \epsilon)A \int_a^\infty r^{-\nu-\epsilon-1}(1 + 2r/a)^\nu\, dr < \infty,$$

Q.E.D.

Corollary. In proving (2.4) we may assume that all terms on the left are finite; otherwise (2.4) is satisfied trivially. Transforming the left side of (2.4) and using (2.5) with some changes of notation, we obtain[9]

[8] The results of Sec. (3), which do not depend on integral equations, are easily extended to systems whose potential energy cannot be written as a sum of pair potentials.

[9] I owe this step to Dr. D. Ruelle.

$$\sum_{i\le j<s} \varphi(\mathbf{x}_i - \mathbf{x}_j) = \tfrac{1}{2} \sum_{i=1}^{s} \sum_{j}{}' \varphi(\mathbf{x}_i - \mathbf{x}_j)$$

$$\geq \tfrac{1}{2} \sum_{i=1}^{s} (-2\Phi') = -s\Phi', \qquad (2.10)$$

where $\sum_j'$ means a sum from 1 to s without the $j = i$ term. Thus (2.4) can be satisfied by taking $\Phi = \Phi'$, and the best value of Φ is at most Φ'. This completes the proof that hard-core potentials (class II) have a finite binding energy per particle (class III).

3. UPPER BOUNDS ON $R(V)$

The grand partition function for a ν-dimensional region V at fugacity z is defined by

$$\Xi(z, V) = 1 + \sum_N (z^N/N!)$$

$$\times \int_V \cdots \int_V e_N(\mathbf{x}_1 \cdots \mathbf{x}_N)\, d^\nu\mathbf{x}_1 \cdots d^\nu\mathbf{x}_N, \qquad (3.1)$$

where the sum is from 1 to ∞ and

$$e_N(\mathbf{x}_1 \cdots \mathbf{x}_N) \equiv \exp\{-\sum_{i<j\le N} \varphi(\mathbf{x}_i - \mathbf{x}_j)/kT\}. \qquad (3.2)$$

We interpret $e_1(\mathbf{x}_1)$ as 1. For a lattice system we interpret the integrations over V in (3.1), and throughout this paper, as sums over all lattice sites contained within V.

For hard-core potentials, an inequality due to Yang and Lee[10] yields upper bounds to $R(V)$. Let M be the largest number of nonoverlapping hard spheres of diameter a whose centers can be fitted into the region V. Then (3.1) reduces to a polynomial of degree M, which may be factorized in the form

$$\Xi(z, V) = \prod_\alpha (1 - z/z_\alpha), \qquad (3.3)$$

where $z_1, z_2 \cdots$ are the zeros of $\Xi(z, V)$.

Expanding the logarithm of (3.3) in powers of z and comparing with (1.2) yields Yang and Lee's formula for the $b_l(V)$'s:

$$\mathcal{V}lb_l(V) = -\sum_\alpha z_\alpha^{-l} \qquad (l = 1, 2, \cdots), \qquad (3.4)$$

where

$$\mathcal{V} \equiv \int_V d^\nu\mathbf{x} \qquad (3.5)$$

is the volume of the region V, or the number of lattice sites within V. Equation (3.4) implies

$$|\mathcal{V}lb_l(V)| \le \sum_\alpha |z_\alpha|^{-l} \le M/R(V)^l \qquad (l = 1, 2, \cdots), \qquad (3.6)$$

since $\Xi(z, V)$ has just M zeros, none of them inside the circle of convergence of (1.2). Rearranging (3.6) gives a set of upper bounds[8] on $R(V)$ for hard-core potentials:

$$R(V) \le |M/\mathcal{V}lb_l(V)|^{1/l} \qquad (l = 1, 2, \cdots). \qquad (3.7)$$

By Cauchy's nth root convergence test,[11] the smallest limit point of the right side of (3.7) is precisely $R(V)$; therefore, the sequence of upper bounds includes a subsequence converging from above to the limit $R(V)$.

Even if the potential does not have a hard core, a similar argument can be used. By virtue of (2.4), the series (3.1) can be majorized by an exponential series, so that for any real or complex value of z,

$$|\Xi(z, V)| \le \exp(|z|\, e^{\Phi/kT}\mathcal{V}). \qquad (3.8)$$

This inequality shows that $\Xi(z, V)$, regarded as a function of the complex variable z, is[12] an entire function (integral function) of order ≤ 1. Therefore it may be factorized using Hadamard's theorem[13,14]

$$\Xi(z, V) = e^{G(z)} \prod_\alpha \{(1 - z/z_\alpha)e^{z/z_\alpha}\}, \qquad (3.9)$$

where $G(z)$ is a polynomial of degree 1. Expanding the logarithm of (3.9) in powers of z and comparing with (1.2) yields a set of formulas differing only slightly from (3.4):

$$G(z) = \mathcal{V}b_1(V)z = \mathcal{V}z, \qquad (3.10)$$

$$\mathcal{V}lb_l(V) = -\sum_\alpha z_\alpha^{-l} \qquad (l = 2, 3, \cdots). \qquad (3.11)$$

The upper bound to $b_l(V)$ obtained from (3.11) may be written as a Stieltjes integral

$$|\mathcal{V}lb_l(V)| \le \sum_\alpha |z_\alpha|^{-l} = \int_0^\infty r^{-l}\, d\mathfrak{N}(r), \qquad (3.12)$$

where $l \ge 2$ and $\mathfrak{N}(r)$ is defined as the number of zeros of $\Xi(z, V)$ in $|z| < r$. In particular, since there are no zeros of $\Xi(z, V)$ inside the circle of convergence of (1.2), we have

$$\mathfrak{N}(r) = 0 \quad \text{if} \quad r \le R(V). \qquad (3.13)$$

[10] C. N. Yang and T. D. Lee, Phys. Rev. **87**, 404 (1952).

[11] E. T. Copson, *Theory of Functions of a Complex Variable*, (Oxford University Press, London, 1935), Sec. 3.33; E. C. Titchmarsh, *The Theory of Functions* (Oxford University Press, London, 1939), Section 7.1.

[12] E. T. Copson, reference 11, Sec. 7.4; E. C. Titchmarsh, reference 11, Sec. 8.2.

[13] E. T. Copson, Reference 11, Sec. 7.6; E. C. Titchmarsh, Reference 11, Sec. 8.24.

[14] If the order is less than 1, Hadamard's factorization is $\Pi_\alpha(1 - z/z_\alpha)$, but since $\sum |z_\alpha|^{-1}$ converges in this case the factorization (3.9) is also valid [E. Borel, *Leçons sur les Fonctions Entières* (Gauthier-Villars, Paris, 1921), p. 32].

D. Ruelle [Helv. Phys. Acta **36**, 183 (1963)] shows that for most potentials the order is actually zero.

Further information about $\mathfrak{N}(r)$ is provided by Jensen's inequality[15] which shows, because of (3.8), that

$$\int_0^r \mathfrak{N}(q)\, dq/q \leq re^{\Phi/kT}\mathfrak{v}. \tag{3.14}$$

This has the corollary

$$\mathfrak{N}(r) = \mathfrak{N}(r)\int_r^{er} dq/q \leq \int_0^{er} \mathfrak{N}(q)\, dq/q \leq ere^{\Phi/kT}\mathfrak{v}. \tag{3.15}$$

Equation (3.13) justifies replacing $\int_0^\infty$ by $\int_{R(V)}^\infty$ in (3.12), and then two partial integrations give, if $l \geq 2$,

$$|\mathfrak{v}lb_l(V)| \leq l\int_{R(V)}^\infty \mathfrak{N}(r)r^{-l-1}\, dr = l^2\int_{R(V)}^\infty r^{-l-1}\, dr \int_{R(V)}^r \mathfrak{N}(q)\, dq/q, \tag{3.16}$$

where the integrated parts vanish by (3.13)–(3.15). The inequalities (3.14) and (3.16) combine to give

$$|\mathfrak{v}lb_l(V)| \leq \mathfrak{v}e^{\Phi/kT}l^2/(l-1)R(V)^{l-1} \qquad (l = 2, 3, \cdots), \tag{3.17}$$

which leads to the upper bounds on $R(V)$,

$$R(V) \leq |e^{\Phi/kT}l/(l-1)b_l(V)|^{1/(l-1)} \qquad (l = 2, 3, \cdots). \tag{3.18}$$

Like the upper bounds in (3.7), these include a subsequence converging from above to the limit $R(V)$. They are valid[8] for all potentials with finite binding energy, including hard-core potentials.

4. RELATION BETWEEN R AND $\lim_{V\to\infty} R(V)$

The inequality (3.18) also furnishes a comparison between R and $\lim_{V\to\infty} R(V)$. Let $V \to \infty$ on both sides of (3.18); this gives[16]

$$\lim_{V\to\infty} R(V) \leq |e^{\Phi/kT}l/(l-1)b_l|^{1/(l-1)} \quad \text{if} \quad l \geq 2, \tag{4.1}$$

since $b_l = \lim_{V\to\infty} b_l(V)$. Now let $l \to \infty$ in (4.1); this gives[11]

$$\lim_{V\to\infty} R(V) \leq R. \tag{4.2}$$

This result can also be deduced from Yang and Lee's results[10] on the uniform convergence of the sequence of functions (1.2).

For nonnegative potentials, Groeneveld[5] has shown that $b_l(V) \leq b_l$, which implies $R(V) \geq R$. Combined with (4.2), this yields the result, due to J. Lebowitz,[4]

$$R = \lim_{V\to\infty} R(V) \quad \text{for nonnegative potentials.} \tag{4.3}$$

5. DISTRIBUTION FUNCTIONS

The s-particle distribution function at fugacity z is defined,[17] for any positive integer s, as

$$n_s(\mathbf{x}_1 \cdots \mathbf{x}_s \mid z) \equiv \frac{\sum_{N=s}^\infty \frac{z^N}{(N-s)!}\int_V \cdots \int_V e_N(\mathbf{x}_1 \cdots \mathbf{x}_N)\, d^\nu\mathbf{x}_{s+1} \cdots d^\nu\mathbf{x}_N}{\Xi(z, V)}. \tag{5.1}$$

For hard-core potentials the series in the numerator, like the series (3.1) for the denominator, is[10] a polynomial of degree M in z, so that n_s is a rational function of z. Even if the potential has no hard core, the numerator and denominator are still entire functions, so that n_s is still a meromorphic function of z.

The radius of convergence of the series expansion of this meromorphic function in powers of z is at least $R(V)$, since the denominator has no zeros in $|z| < R(V)$. This series expansion may be written

$$n_s(\mathbf{x}_1 \cdots \mathbf{x}_s \mid z) = \sum_{l=0}^\infty n_{s,l}(\mathbf{x}_1 \cdots \mathbf{x}_s)z^{s+l}. \tag{5.2}$$

The coefficients in the series for $s = 1$ are related to the Mayer coefficients $b_l(V)$; for (5.1), (3.1), and (1.2) imply

$$\int_V n_1(\mathbf{x} \mid z)\, d^\nu\mathbf{x} = z\frac{d(\log \Xi)}{dz} = \mathfrak{v}\sum lb_l(V)z^l, \tag{5.3}$$

whence

$$\mathfrak{v}lb_l(V) = \int_V n_{1,l-1}(\mathbf{x})\, d^\nu\mathbf{x} \quad (l = 1, 2, \cdots). \tag{5.4}$$

We shall use this formula to estimate $R(V)$ from information about the coefficient functions $n_{s,l}(\mathbf{x}_1 \cdots \mathbf{x}_s)$.

[15] E. T. Copson, reference 11, Sec. 7.42; E. C. Titchmarsh, reference 11, Sec. 8.21.

[16] *Note added in proof*: To allow for the possibility that $\lim_{V\to\infty} R(V)$ may not exist, it would be better to write lim sup for lim in Eqs. (4.1), (4.2), and (9.2).

[17] T. L. Hill, reference 7, Sec. 37. To shorten the formulas a little, the fact that n_s also depends on the form of the region V is not shown explicitly in our notation.

The first term in the series (5.2) is easily obtained by division from (5.1) and (3.1). It is

$$n_{s,0}(\mathbf{x}_1 \cdots \mathbf{x}_s) = e_s(\mathbf{x}_1 \cdots \mathbf{x}_s). \tag{5.5}$$

Recurrence relations for the higher coefficients are conveniently obtained by substituting (5.2) either into the Kirkwood–Salsburg integral equation[7] or the Mayer–Montroll integral equation.[18]

The Kirkwood–Salsburg equation is[7]

$$\begin{aligned} n_s(\mathbf{x}_1 \cdots \mathbf{x}_s \mid z) &= z\left\{\prod_{i=2}^{s} e_2(\mathbf{x}_1, \mathbf{x}_i)\right\} \\ &\times \sum_{k=0}^{\infty} \frac{1}{k!} \int \cdots \int n_{s-1+k}(\mathbf{x}_2 \cdots \mathbf{x}_{s+k} \mid z) \\ &\times \prod_{j=s+1}^{s+k} \{f_1(\mathbf{x}_1; \mathbf{x}_j)\, d^\nu \mathbf{x}_j\}, \end{aligned} \tag{5.6}$$

for $s = 1, 2, \cdots$, where

$$f_1(\mathbf{x}; \mathbf{y}) \equiv e_2(\mathbf{x}, \mathbf{y}) - 1, \tag{5.7}$$

and

$$n_0 \equiv 1. \tag{5.8}$$

Substituting (5.2) into (5.6) and equating coefficients of like powers of z gives the recurrence relation

$$\begin{aligned} n_{s,l}(\mathbf{x}_1 \cdots \mathbf{x}_s) &= \left\{\prod_{i=2}^{s} e_2(\mathbf{x}_1, \mathbf{x}_i)\right\} \\ &\times \sum_{k=0}^{l} \frac{1}{k!} \int \cdots \int n_{s-1+k,\,l-k}(\mathbf{x}_2 \cdots \mathbf{x}_{s+k}) \\ &\times \prod_{j=s+1}^{s+k} \{f_1(\mathbf{x}_1, \mathbf{x}_j)\, d^\nu \mathbf{x}_j\}, \end{aligned} \tag{5.9}$$

where

$$n_{0,l} \equiv \delta_{0,l}. \tag{5.10}$$

Equation (5.9) holds for all $s \geq 1$, $l \geq 0$.

The Mayer–Montroll equation is[18]

$$\begin{aligned} n_s(\mathbf{x}_1 \cdots \mathbf{x}_s \mid z) &= z^s e_s(\mathbf{x}_1 \cdots \mathbf{x}_s) \\ &\times \sum_{k=0}^{\infty} \frac{1}{k!} \int \cdots \int n_k(\mathbf{x}_{s+1} \cdots \mathbf{x}_{s+k} \mid z) \\ &\times \prod_{j=s+1}^{s+k} \{f_s(\mathbf{x}_1 \cdots \mathbf{x}_s; \mathbf{x}_j)\, d^\nu \mathbf{x}_j\}, \end{aligned} \tag{5.11}$$

for $s = 1, 2, \cdots$, where

$$f_s(\mathbf{x}_1 \cdots \mathbf{x}_s; \mathbf{y}) \equiv \left\{\prod_{i=1}^{s} e_2(\mathbf{x}_i, \mathbf{y})\right\} - 1, \tag{5.12}$$

and $n_0 \equiv 1$ as before. Equation (5.11) leads to the recurrence relation

$$\begin{aligned} n_{s,l}(\mathbf{x}_1 \cdots \mathbf{x}_s) &= e_s(\mathbf{x}_1 \cdots \mathbf{x}_s) \\ &\times \sum_{k=1}^{l} \frac{1}{k!} \int \cdots \int n_{k,\,l-k}(\mathbf{x}_{s+1} \cdots \mathbf{x}_{s+k}) \\ &\times \prod_{j=s+1}^{s+k} \{f_s(\mathbf{x}_1 \cdots \mathbf{x}_s; \mathbf{x}_j)\, d^\nu \mathbf{x}_j\}, \end{aligned} \tag{5.13}$$

which holds for all $s \geq 1$, $l \geq 1$.

6. LOWER BOUND ON $R(V)$

One way to obtain a lower bound on $R(V)$ is to calculate upper bounds on the coefficient functions $n_{s,l}(\mathbf{x}_1 \cdots \mathbf{x}_s)$ from the recurrence relation (5.9), which depends on the Kirkwood–Salsburg integral equation. These upper bounds may be written

$$|n_{s,m-s}(\mathbf{x}_1 \cdots \mathbf{x}_s)| \leq K_{s,m-s} \quad (s = 1, 2 \cdots m). \tag{6.1}$$

In accordance with (5.5), we can satisfy this inequality for $m = 1$ by taking

$$K_{1,0} = 1. \tag{6.2}$$

Values of $K_{s,m-s}$ satisfying (6.1) for $m > 1$ may now be calculated by induction. If (6.1) is true for $m = M - 1$, where $M - 1$ is any positive integer, then (5.9) and (5.10) imply

$$\begin{aligned} |n_{s,M-s}(\mathbf{x}_1 \cdots \mathbf{x}_s)| &\leq \left\{\prod_{i=2}^{s} e_2(\mathbf{x}_1, \mathbf{x}_i)\right\} \\ &\times \sum_{k=0}^{M-s} \frac{1}{k!} K_{s-1+k,\,M-s-k} B(V)^k, \end{aligned} \tag{6.3}$$

for $s = 1, 2, \cdots M$, where

$$B(V) \equiv \operatorname*{Max}_{\mathbf{x}} \int_V |f_1(\mathbf{x}; \mathbf{y})|\, d^\nu \mathbf{y} \tag{6.4}$$

and

$$K_{0,m} \equiv \delta_{0,m}. \tag{6.5}$$

For nonnegative potentials, the product in braces in (6.3) has the upper bound 1 and may therefore be omitted. Thus, if (6.1) holds for $m = M - 1$, it also holds for $m = M$ provided

$$K_{s,M-s} \geq \sum_{k=0}^{M-s} \frac{1}{k!} K_{s-1+k,\,M-s-k} B(V)^k \quad (s = 1, 2, \cdots M). \tag{6.6}$$

It follows by induction that (6.1) holds for all $m \geq 1$, if $K_{s,l}$ satisfies (6.5) for $s = 0$, (6.2) for $s = 1$, $l = 0$, and (6.6) for all other nonnegative values of s and l. These conditions can be satisfied by taking $K_{s,l} = s(s + l)^{l-1} B(V)^l / l!$; in fact, this is their "best possible" solution, which makes both sides of (6.6) equal. Substituting this formula

[18] J. E. Mayer and E. Montroll, J. Chem. Phys. 9, 2 (1941).

for $K_{s,l}$ into (6.1) we obtain for nonnegative potentials

$$|n_{s,l}(\mathbf{x}_1 \cdots \mathbf{x}_s)| \leq s(s+l)^{l-1}B(V)^l/l! \qquad (6.7)$$

for all $s \geq 1$, $l \geq 0$.

This result may be generalized to other potentials with finite binding energy per particle, using a resource due to Ruelle.[6] Since $n_{s,M-s}(\mathbf{x}_1 \cdots \mathbf{x}_s)$ is symmetric in $\mathbf{x}_1 \cdots \mathbf{x}_s$, it satisfies not only (6.3) but $s-1$ further inequalities obtained by cyclic permutations of the indices:

$$|n_{s,M-s}(\mathbf{x}_1 \cdots \mathbf{x}_s)| \leq \{\prod_i{}' e_2(\mathbf{x}_J, \mathbf{x}_i)\} \times \sum_k K_{s-1+k,M-s-k}B(V)^k/k! \qquad (6.8)$$

for $J = 1, 2 \cdots s$, where $\prod_i'$ means a product over $i = 1, 2, \cdots J-1, J+1, \cdots s$. Therefore $n_{s,M-s}(\mathbf{x}_1 \cdots \mathbf{x}_s)$ is also less than the geometric mean of the s different right-hand sides of (6.8). By (3.2), this geometric mean is a similar expression with $\exp[(-2/skT)\sum_{i<j\leq s}\varphi(\mathbf{x}_i - \mathbf{x}_j)]$ replacing the product in the braces. By (2.4), this exponential is at most $e^{2\Phi/kT}$; therefore (6.1) is satisfied for $m = M$ provided

$$K_{s,M-s} \geq e^{2\Phi/kT}\sum_{k=0}^{M-s}\frac{1}{k!}K_{s-1+k,M-s-k}B(V)^k \qquad (s = 1, \cdots M). \qquad (6.9)$$

It follows, by a discussion parallel to that of (6.6), that

$$|n_{s,l}(\mathbf{x}_1 \cdots \mathbf{x}_s)| \leq K_{s,l} = e^{2(s+l-1)\Phi/kT}s(s+l)^{l-1}B(V)^l/l! \qquad (6.10)$$

for all $s \geq 1$ and $l \geq 0$.

A stronger upper bound on $n_{1,l}(\mathbf{x}_1)$ can be obtained by setting $s = 1$ in (6.3) and substituting on the right from (6.10). Since the product over i in (6.3) now degenerates to 1, we obtain

$$|n_{1,l}(\mathbf{x}_1)| \leq e^{2(l-1)\Phi/kT}(1+l)^{l-1}B(V)^l/l! \qquad (6.11)$$

for all $l \geq 1$. Substituting this into (5.4) gives

$$|lb_l(V)| \leq e^{2(l-2)\Phi/kT}[lB(V)]^{l-1}/l! \qquad (6.12)$$

for all $l \geq 2$. The radius of convergence of the series (1.2) is therefore

$$R(V) = \liminf_{l\to\infty}|b_l(V)|^{-1/l} \geq 1/B(V)e^{1+2\Phi/kT}. \qquad (6.13)$$

If $\Phi = 0$ and $V \to \infty$, (6.12) reduces to one of Groeneveld's results;[5] if $V \to \infty$, (6.13) reduces to one of Ruelle's.[6]

7. UPPER BOUNDS ON $n_{s,l}(x_1 \cdots x_s)$

For nonnegative and hard-core potentials the calculation in the previous section can be supplemented by a similar calculation based on the recurrence relation (5.13) which depends on the Mayer–Montroll equation.[18] The resulting inequalities have the advantage of indicating the dependence of $n_{s,l}$ on $\mathbf{x}_1 \cdots \mathbf{x}_s$. They also lead to a new lower bound on $R(V)$.

The form of Eq. (5.13) suggests looking for bounds of the form

$$|n_{s,l}(\mathbf{x}_1 \cdots \mathbf{x}_s)| \leq e_s(\mathbf{x}_1 \cdots \mathbf{x}_s)M_{s,l} \qquad (s = 1, 2, \cdots). \qquad (7.1)$$

Eq. (5.5) shows that we may take

$$M_{s,0} = 1 \qquad (s = 1, 2, \cdots). \qquad (7.2)$$

As an inductive hypothesis, suppose that (7.1) holds for $l = 0, 1, \cdots L-1$ where L is a positive integer. Then (5.13), (3.2), and (2.4) imply

$$|n_{s,L}(\mathbf{x}_1 \cdots \mathbf{x}_s)| \leq e_s(\mathbf{x}_1 \cdots \mathbf{x}_s) \times \sum_{k=1}^{L}\frac{1}{k!}M_{k,L-k}e^{k\Phi/kT} \times \left\{\int |f_s(\mathbf{x}_1 \cdots \mathbf{x}_s; \mathbf{y})|\, d^\nu\mathbf{y}\right\}^k \qquad (7.3)$$

for all $s \geq 1$.

For nonnegative potentials, Φ vanishes and $|f_s|$ has the upper bound used by Groeneveld[5],

$$|f_s(\mathbf{x}_1 \cdots \mathbf{x}_s; \mathbf{y})| \leq \sum_{i=1}^{s}|f_1(\mathbf{x}_i; \mathbf{y})|; \qquad (7.4)$$

therefore if (7.1) holds for $l = 0, 1, \cdots L-1$ it also holds for $l = L$, provided

$$M_{s,L} \geq \sum_{k=1}^{L}M_{k,L-k}[sB(V)]^k/k! \qquad (7.5)$$

for all $s \geq 1$. It follows by induction that (7.1) holds for all $l \geq 0$ provided the numbers $M_{s,l}$ satisfy (7.2) for $l = 0$ and (7.5) for all $l \geq 1$. The "best possible" solution of these conditions is $M_{s,l} = s(s+l)^{l-1}B(V)^l/l!$. Substituted into (7.1), this gives (nonnegative potentials only)

$$|n_{s,l}(\mathbf{x}_1 \cdots \mathbf{x}_s)| \leq s(s+l)^{l-1}B(V)^l e_s(\mathbf{x}_1 \cdots \mathbf{x}_s)/l! \qquad (7.6)$$

for all $s \geq 1$, $l \geq 0$, which is at least as strong as the corresponding inequality (6.7) obtained from the Kirkwood–Salsburg equation.

For hard-core potentials, a generalization of (7.4) is required. In obtaining this we may assume that

$e_s(\mathbf{x}_1 \cdots \mathbf{x}_s) > 0$, since if the factor e_s vanishes in (7.3) the value of the other factor is immaterial. Consequently we may write, using (5.12) and (2.5),

$$1 + f_s(\mathbf{x}_1 \cdots \mathbf{x}_s; \mathbf{y}) = \prod_{i=1}^{s} (1 + g_i) \leq e^{2\Phi'/kT}, \quad (7.7)$$

where

$$g_i \equiv \exp[-\varphi(\mathbf{x}_i - \mathbf{y})/kT] - 1. \quad (7.8)$$

Using the inequality in (7.7), and the fact that $g_i \geq -1$, it can be proved by induction on s that

$$1 - \sum_{i=1}^{s} g_{i-} \leq \prod_{i=1}^{s} (1 + g_i) \leq 1 + e^{2\Phi'/kT} \sum_{i=1}^{s} g_{i+}/(1 + g_{i+}), \quad (7.9)$$

where

$$g_{i\pm} \equiv \text{Max}\,(0, \pm g_i) \geq 0. \quad (7.10)$$

Combining the equality in (7.7) with (7.9) gives the required generalization of (7.4):

$$|f_s(\mathbf{x}_1 \cdots \mathbf{x}_s; \mathbf{y})| \leq \sum_{i=1}^{s} \text{Max}\,\{g_{i-}, e^{2\Phi'/kT} g_{i+}/(1 + g_{i+})\}. \quad (7.11)$$

Using (7.11) in (7.3) we find that if (7.1) holds for $l = 0, 1, \cdots L - 1$, it also holds for $l = L$ provided

$$M_{s,L} \geq \sum_{k=1}^{L} M_{k,L-k} e^{k\Phi/kT} [sB'(V)]^k/k! \quad (7.12)$$

for all $s \geq 1$, where

$$B'(V) \equiv \text{Max}_{\mathbf{x}} \int_V \text{Max}\,\{(1 - e^{-\varphi(\mathbf{x}-\mathbf{y})/kT}), \times e^{2\Phi'/kT}(1 - e^{\varphi(\mathbf{x}-\mathbf{y})/kT})\}\, d^\nu\mathbf{y}. \quad (7.13)$$

It follows, by a discussion parallel to that following (7.5), that for hard-core potentials

$$|n_{s,l}(\mathbf{x}_1 \cdots \mathbf{x}_s)| \leq s(s + l)^{l-1} \times [B'(V)e^{\Phi/kT}]^l e_s(\mathbf{x}_1 \cdots \mathbf{x}_s)/l! \quad (7.14)$$

for all $s \geq 1$, $l \geq 0$. By (7.6), this inequality also holds for nonnegative potentials. Substituted into (5.4) it gives

$$|lb_l(V)| \leq [lB'(V)e^{\Phi/kT}]^{l-1}/l!, \quad (7.15)$$

and the new radius of convergence estimate

$$R(V) \geq 1/B'(V)e^{1+\Phi/kT}. \quad (7.16)$$

This inequality is stronger than (6.13) for some potentials and temperatures, but weaker for others; the lattice gas example considered in Sec. 9 illustrates both possibilities.

8. SPECIAL UPPER BOUNDS ON R AND $R(V)$

For nonnegative potentials, Groeneveld[5] showed that R has the upper bound $1/2|b_2|$. By virtue of (4.3), this is also an upper bound on $\lim_{V\to\infty} R(V)$. This bound is stronger than the first upper bound $2/|b_2|$ deducible from (3.18), presumably because the result (3.18) is not restricted to two-body interactions.[8] The method of integral equations provides an alternative proof of Groeneveld's upper bound on R and also yields the corresponding result for finite volumes.

For nonnegative potentials, the definition (5.12) implies $f_s(\mathbf{x}_1, \cdots \mathbf{x}_s; \mathbf{y}) \leq 0$; it follows inductively from (5.5) and (5.13) that

$$(-)^l n_{s,l}(\mathbf{x}_1 \cdots \mathbf{x}_s) \geq 0 \quad (8.1)$$

for all $s \geq 1$, $l \geq 0$, and that all terms in the sum (5.13) have the same sign. Consequently the $k = 1$ term in (5.13) gives a lower bound to $|n_{s,l}|$ if $l \geq 1$. For $s = 1$, this bound is

$$|n_{1,l}(\mathbf{x})| \geq \int_V |f_1(\mathbf{x}; \mathbf{y})|\, |n_{1,l-1}(\mathbf{y})|\, d^\nu\mathbf{y} \quad (8.2)$$

for $l \geq 1$, with equality when $l = 1$. Since $n_{1,0}(\mathbf{x}) = 1$ by (5.5), it follows that

$$|n_{1,l}(\mathbf{x})| \geq \int_V K^{(l)}(\mathbf{x}, \mathbf{y})\, d^\nu\mathbf{y}, \quad (8.3)$$

where $K^{(l)}(\mathbf{x}, \mathbf{y})$ is the lth iterate[19] of the symmetric kernel $K^{(1)}(\mathbf{x}, \mathbf{y}) \equiv |f_1(\mathbf{x}; \mathbf{y})|$. Expanding[18] $K^{(l)}(\mathbf{x},\mathbf{y})$ in (8.3) in terms of its (real) eigenvalues $\lambda_0, \lambda_1, \cdots$ and eigenfunctions $\psi_0(\mathbf{x}), \psi_1(\mathbf{x}), \cdots$, and substituting into (5.4), we obtain

$$|\mathcal{V}(l + 1)b_{l+1}(V)| \geq \int_V \int_V K^{(l)}(\mathbf{x}, \mathbf{y})\, d^\nu\mathbf{x}\, d^\nu\mathbf{y} = \sum_p \left\{\int \psi_p(\mathbf{x})\, d^\nu\mathbf{x}\right\}^2/\lambda_p^l. \quad (8.4)$$

Therefore $R(V)$ has the lower bound

$$R(V) = \liminf_{l\to\infty} |b_l(V)|^{-1/l} \leq |\lambda_0|, \quad (8.5)$$

where λ_0 is the eigenvalue of $K(\mathbf{x}, \mathbf{y})$ with smallest absolute value. Since (8.4) is an equality for $l = 1$, it also implies

$$2\mathcal{V}\,|b_2(V)| \leq \sum_p \left\{\int \psi_p(\mathbf{x})\, d^\nu\mathbf{x}\right\}^2/|\lambda_0| \leq \mathcal{V}/|\lambda_0|, \quad (8.6)$$

[19] R. Courant and D. Hilbert, *Methods of Mathematical Physics* (Interscience Publishers, Inc., New York, 1953) Chap. 3, Sec. 5.

by virtue of Bessel's inequality.[20] Combining (8.5) with (8.6) gives the main result of this section,

$$R(V) \leq 1/2|b_2(V)|. \tag{8.7}$$

Taking the limit $V \to \infty$ and using (4.3), we recover Groeneveld's result,[5]

$$R \leq 1/2|b_2|, \tag{8.8}$$

which can also be proved more simply by letting $V \to \infty$ at the beginning of the calculation [Eq. (8.2)] instead of the end.

Unfortunately, it seems to be impossible to extend this result to more general potentials.

9. CONCLUSIONS

Our basic inequalities (6.13), (3.18), and (4.2) for the radii of convergence R and $R(V)$ of the series (1.1) and (1.2) may be summarized in the form[16]

$$1/B(V)e^{1+2\Phi/kT} \leq R(V)$$

$$\leq |e^{\Phi/kT}l/(l-1)b_l(V)|^{1/(l-1)} \quad (l = 2, 3, \cdots), \tag{9.1}$$

$$\lim_{V\to\infty} R(V) \leq R. \tag{9.2}$$

These are valid for any pair potential[8] with finite binding energy per particle. Hard-core potentials also satisfy (7.16) and (3.7),

$$1/B'(V)e^{1+\Phi/kT} \leq R(V) \leq |M/\mathcal{V}lb_l(V)|^{1/l} \quad (l = 1, 2, \cdots), \tag{9.3}$$

and if the potential is nonnegative, the relations

$$R \leq R(V) \leq |1/2b_2(V)|, \tag{9.4}$$

$$\lim_{V\to\infty} R(V) = R \tag{9.5}$$

supplement (9.1) and (9.2), and (9.3) if it applies.

The quantities $B(V)$, Φ, $B'(V)$, and Φ' are defined in (6.4), (2.4), (7.13), and (2.5); M is defined just before (3.3). By virtue of the inequality

$$B(V) \leq B \equiv \lim_{V\to\infty} B(V)$$

$$= \int_{\text{all space}} |e^{-\varphi(\mathbf{r})/kT} - 1|\, d^3\mathbf{r}, \tag{9.6}$$

we may replace $B(V)$ in (9.1) by B, whose definition is neater. This makes the lower bound in (9.1) identical with that given by Ruelle.[6] Similarly we may replace $B'(V)$ in (9.3) by its upper bound

$$B' \equiv \int_{\text{all space}} \text{Max}\,\{(1 - e^{-\varphi(\mathbf{r})/kT}), e^{2\Phi'/kT} \times (1 - e^{\varphi(\mathbf{r})/kT})\}\, d^3\mathbf{r}. \tag{9.7}$$

These simplifications destroy hardly any information unless V is small or B infinite. A case where the original form of (9.1) does give more information is the Coulomb potential $\varphi(\mathbf{r}) = \text{const}/r$, for which (9.1) shows that $R(V) = O(\mathcal{V}^{-1})$.

A system for which R can be calculated exactly is the 1-dimensional gas of hard rods of length a, for which $\varphi(\mathbf{r}) = +\infty$ for $r < a$ and $= 0$ for $r \geq a$. The equation of state is[21] $p = \rho kT/(1 - a\rho)$ from which it follows that $z = (p/kT)\exp(ap/kT)$ and hence, by (1.1) and Lagrange's inversion formula,[22] that

$$b_l = (-al)^{l-1}/l!. \tag{9.8}$$

The function $p(z) = kT\sum b_l z^l$ is analytic on the positive real axis, but it has a branch point at $z = -1/ea$, so that $R = 1/ea = 0.368a^{-1}$. This system shows no phase transition, and the divergence of $\sum b_l z^l$ for real $z > R$ has no physical significance.

Equation (9.5) shows that $\lim_{V\to\infty} R(V) = R = 0.368a^{-1}$. The lower bound on $\lim_{V\to\infty} R(V)$ computed from (9.1) is $1/eB = 1/2e|b_2| = 0.184a^{-1}$, just half the correct value. The upper bound on $\lim_{V\to\infty} R(V)$ given by (9.4) is $1/2a = 0.500a^{-1}$. The convergent sequence of upper bounds given by (9.1) starts $2.000a^{-1}$, $1.000a^{-1}$, $0.795a^{-1}$, $0.700a^{-1}$, $0.644a^{-1}$, etc., while the one given by (9.3) starts a^{-1}, $0.707a^{-1}$, $0.606a^{-1}$, $0.553a^{-1}$, $0.521a^{-1}$, $0.499a^{-1}$, $0.483a^{-1}$, etc. The convergence to the limit $0.368a^{-1}$ is a little faster in the second sequence, whose derivation takes advantage of the information that the potential has a hard core.

For an Ising lattice gas with coordination number q and negative interaction energy -2ϵ between nearest neighbours, Lee and Yang have shown[23] that $\lim_{V\to\infty} R(V) = 1/w^q$ where $w \equiv e^{\epsilon/kT}$. This is a hard-core potential, for which $\Phi' = \Phi = q\epsilon$, and our lower bounds on $\lim R(V)$ are

$$1/Be^{1+2\Phi/kT} = 1/[1 + q(w^2 - 1)]ew^{2q}, \tag{9.9}$$

$$1/B'e^{1+\Phi/kT} = 1/[1 + w^{2q}q(1 - w^{-2})]ew^q. \tag{9.10}$$

For a given lattice (fixed $q \geq 3$), the first of these lower bounds is the stronger when $w \gg 1$ (low temperatures) and the second when w is close to 1 (high temperatures). Our simplest upper bound on $\lim R(V)$ is 1 [from (9.3) with $l = 1$], which is actually attained when $\epsilon = 0$.

The physical significance of the lower bounds (9.1)

[20] Reference 18, Chap. 2, Sec. 1.

[21] L. Tonks, Phys. Rev. **50**, 955 (1936).

[22] E. T. Copson, reference 11, Sec. 6.23.

[23] T. D. Lee and C. N. Yang, Phys. Rev. **87**, 410 (1952). In the limit $V \to \infty$ their z becomes our $z\exp(q\epsilon/kT)$.

and (9.3) is[5,6] that they imply there can be no phase transition for fugacities less than $1/Be^{1+2\Phi/kT}$ (or $1/B'e^{1+\Phi/kT}$ for hard-core potentials) since the convergent power series (1.1) correctly gives the pressure for these fugacities. This is proved[5] most directly by using (6.12) and (9.6) [or (7.15) and (9.7)] together with Weierstrass' theorem[24] to show that the series (1.2) converges uniformly in the limit $V \to \infty$ for any z inside the disk $|z| < 1/Be^{1+2\Phi/kT}$ (or $|z| < 1/B'e^{1+\Phi'/kT}$). Alternatively, one may apply Yang and Lee's theory[10,25], since (6.13) and (9.6) [or 7.16 and (9.7)] show that this disk contains no zeros of $\Xi(z, V)$ for any V. Unfortunately the upper bounds on R and $R(V)$ cannot be interpreted so easily, since the singularity of (1.1) or (1.2) nearest to the origin need not be on the positive real axis; for nonnegative potentials this singularity is in fact on the negative real axis.[5] An example is the one-dimensional gas of hard rods, which has no phase transition although the fugacity series diverges for $z > 1/ea$.

[24] E. C. Titchmarsh, reference 11, 1.11 and 1.14.
[25] D. Ruelle, Helv. Phys. Acta **36**, 183 (1963) extends Yang and Lee's theorems to a more general class of potentials.

10. ACKNOWLEDGMENTS

I am indebted to J. Groeneveld, J. Lebowitz, E. Lieb, J. Percus, D. Ruelle, and G. Stell for helpful correspondence and discussions.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 4, NUMBER 10 OCTOBER 1963

Remarks on the Combinatorial Approach to the Ising Problem*

P. N. BURGOYNE
Department of Mathematics, University of California, Berkeley, California
(Received 30 April 1963)

A new proof is given of a certain conjecture due to Feynman. This conjecture relates graphs and paths on a lattice and was first proved by Sherman. It is the key step in a particular method of obtaining Onsager's formula for the partition function of the two-dimensional Ising model.

I. INTRODUCTION

IMAGINE a lattice in which each site α has two possible states $\sigma_\alpha = +1$ or $\sigma_\alpha = -1$. For example σ_α might specify whether an atom at α had spin up or spin down. Imagine further that the energy of the system is proportional to $\sum \sigma_\alpha\sigma_\beta$ where the sum is taken over all pairs of neighboring sites (see Fig. 1). This is the Ising model.[1]

This model is of interest for many reasons, both mathematical and physical, not the least important being that it shows (in 2 and 3 dimensions) a definite phase transition. To rigorously derive this fact, and to help understand the Ising model, it is necessary to have a closed formula for its partition function. This is the Ising problem.

For a one-dimensional lattice, the solution to the problem is quite straightforward. For a two-dimensional lattice it has been done but is not exactly straightforward. The three-dimensional problem has not been solved. We might mention, however, that for the latter case, there are powerful approximate techniques.[2]

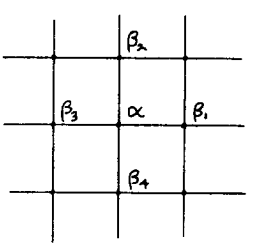


FIG. 1. The neighboring sites of α for a two-dimensional square lattice.

A closed formula for the partition function in the two-dimensional case was first obtained by Onsager.[3] His approach is purely algebraic. In 1952 Kac and

* Research supported by National Science Foundation Grant GP-2.
[1] E. Ising, Z. Physik **31**, 253 (1925). There are two good review articles; one by G. F. Newell and E. W. Montroll, Rev. Mod. Phys. **25**, 353 (1953); the other by C. Domb, Advan. Phys. **9**, 151 (1960).
[2] See for example G. A. Baker Jr., Phys. Rev. **124**, 768 (1961).
[3] L. Onsager, Phys. Rev. **65**, 117 (1944); see also B. Kaufman, Phys. Rev. **76**, 1232 (1949).

and (9.3) is[5,6] that they imply there can be no phase transition for fugacities less than $1/Be^{1+2\Phi/kT}$ (or $1/B'e^{1+\Phi/kT}$ for hard-core potentials) since the convergent power series (1.1) correctly gives the pressure for these fugacities. This is proved[5] most directly by using (6.12) and (9.6) [or (7.15) and (9.7)] together with Weierstrass' theorem[24] to show that the series (1.2) converges uniformly in the limit $V \to \infty$ for any z inside the disk $|z| < 1/Be^{1+2\Phi/kT}$ (or $|z| < 1/B'e^{1+\Phi'/kT}$). Alternatively, one may apply Yang and Lee's theory[10,25], since (6.13) and (9.6) [or 7.16 and (9.7)] show that this disk contains no zeros of $\Xi(z, V)$ for any V. Unfortunately the upper bounds on R and $R(V)$ cannot be interpreted so easily, since the singularity of (1.1) or (1.2) nearest to the origin need not be on the positive real axis; for nonnegative potentials this singularity is in fact on the negative real axis.[5] An example is the one-dimensional gas of hard rods, which has no phase transition although the fugacity series diverges for $z > 1/ea$.

[24] E. C. Titchmarsh, reference 11, 1.11 and 1.14.
[25] D. Ruelle, Helv. Phys. Acta **36**, 183 (1963) extends Yang and Lee's theorems to a more general class of potentials.

10. ACKNOWLEDGMENTS

I am indebted to J. Groeneveld, J. Lebowitz, E. Lieb, J. Percus, D. Ruelle, and G. Stell for helpful correspondence and discussions.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 4, NUMBER 10 OCTOBER 1963

Remarks on the Combinatorial Approach to the Ising Problem*

P. N. BURGOYNE
Department of Mathematics, University of California, Berkeley, California
(Received 30 April 1963)

A new proof is given of a certain conjecture due to Feynman. This conjecture relates graphs and paths on a lattice and was first proved by Sherman. It is the key step in a particular method of obtaining Onsager's formula for the partition function of the two-dimensional Ising model.

I. INTRODUCTION

IMAGINE a lattice in which each site α has two possible states $\sigma_\alpha = +1$ or $\sigma_\alpha = -1$. For example σ_α might specify whether an atom at α had spin up or spin down. Imagine further that the energy of the system is proportional to $\sum \sigma_\alpha\sigma_\beta$ where the sum is taken over all pairs of neighboring sites (see Fig. 1). This is the Ising model.[1]

This model is of interest for many reasons, both mathematical and physical, not the least important being that it shows (in 2 and 3 dimensions) a definite phase transition. To rigorously derive this fact, and to help understand the Ising model, it is necessary to have a closed formula for its partition function. This is the Ising problem.

For a one-dimensional lattice, the solution to the problem is quite straightforward. For a two-dimensional lattice it has been done but is not exactly straightforward. The three-dimensional problem has not been solved. We might mention, however, that for the latter case, there are powerful approximate techniques.[2]

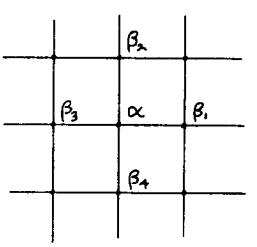
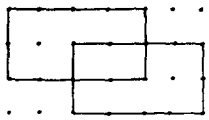


FIG. 1. The neighboring sites of α for a two-dimensional square lattice.

A closed formula for the partition function in the two-dimensional case was first obtained by Onsager.[3] His approach is purely algebraic. In 1952 Kac and

* Research supported by National Science Foundation Grant GP-2.
[1] E. Ising, Z. Physik **31**, 253 (1925). There are two good review articles; one by G. F. Newell and E. W. Montroll, Rev. Mod. Phys. **25**, 353 (1953); the other by C. Domb, Advan. Phys. **9**, 151 (1960).
[2] See for example G. A. Baker Jr., Phys. Rev. **124**, 768 (1961).
[3] L. Onsager, Phys. Rev. **65**, 117 (1944); see also B. Kaufman, Phys. Rev. **76**, 1232 (1949).

FIG. 2. An admissible graph with two crossings.



Ward[4] gave a brief discussion of a quite different, combinatorial method of obtaining Onsager's formula. Though they made their method seem extremely plausible, they did not intend to give a rigorous proof. Feynman, in some unpublished lecture notes, developed a variation of the Kac–Ward approach based on a conjectured identity [see Eq. (1)] between graphs and paths on a lattice. This identity was proved by Sherman,[5] who not only generalized it but also gave a number of further applications. This paper was the first correct treatment of the combinatorial method. Indpendently, Hurst and Green,[6] using pfaffians, have given another, very concise, justification of the Kac–Ward approach.

Sherman's proof is quite involved and at one place incomplete.[7] The purpose of this paper is to provide another proof of Feynman's conjecture. Although closely related to Sherman's proof, our method is somewhat more geometric and, therefore, possibly more intuitive. The reason for rederiving Onsager's formula by yet another method is that, hopefully, it may shed some light on either the unsolved three-dimensional problem or the (perhaps physically more interesting and also unsolved) case of the two-dimensional lattice in a magnetic field.

To make the paper reasonably self-contained, an appendix gives the connection between the partition function, Eq. (1), and Onsager's formula.

II. THE THEOREM

Consider a two-dimensional square lattice, either finite or infinite. Upon this stage we distinguish two actors—paths, and admissible graphs.

An *admissible graph* is simply some finite set of lines (a line being the segment joining two neighboring sites) such that either 0, 2, or 4 lines meet at every site. Those sites where 4 lines meet are of special interest and are called *crossings*. A graph is not necessarily connected.

FIG. 3. A graph which is not admissible.

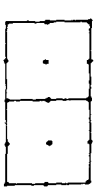
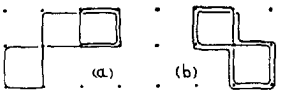


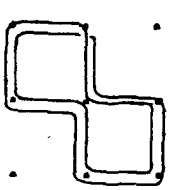
FIG. 4. Some nonperiodic paths.



By a *path* is meant a definite sequence of lines. Each succeeding line starts at the site where the previous line ended and may continue straight ahead or to either side but not backwards over the previous line. We only consider closed paths, and therefore the last line must end at the site from which the first line started. The direction in which the sequence of lines is traversed, and also the particular starting point are both immaterial. There is no restriction on the number of times a line may be traversed in a path. By a *periodic* path is meant one which can be constructed by exactly repeating some subpath of it two or more times. Note that a path may have all its lines repeated twice or more without necessarily being periodic [see Fig. 4(b)]. (In drawing paths it is necessary both to separate slightly the repeated lines and to distinguish the behavior at crossings. Otherwise it is not always clear from the drawing in which sequence the lines are traversed. The sign of the path, defined below, is independent of the manner in which this is done.[8])

Define the *sign* of a path to be $+1$ or -1 depending on whether the path crosses itself an even (+) or odd (−) number of times. For example the paths in Fig. 4(a) and (b) both have the sign $+1$ while Fig. (5) has -1.

FIG. 5. A periodic path of period 2.



To each line in the lattice associate a symbol d_i (i is an index distinguishing different lines). Then for any admissible graph G, define $I(G) = \prod d_i$, where the product is taken over those lines occurring in G. The order of the d_i's in the product is irrelevant. Similarly, for any path p, define $I(p) = \prod d_i^{\mu_i}$, where μ_i is the number of times the line d_i is traversed by p. Lastly let $W(p) = \pm I(p)$ where the $\pm$ is the sign of p.

Theorem:

$$1 + \sum_G I(G) = \prod_p [1 + W(p)]. \qquad (1)$$

[4] M. Kac and J. C. Ward, Phys. Rev. **88**, 1332 (1952).
[5] S. Sherman, J. Math. Phys. **1**, 202 (1960).
S. Sherman, Bull. Am. Math. Soc. **68**, 225 (1962).
[6] C. A. Hurst and H. S. Green, J. Chem. Phys. **33**, 1059 (1960) and Math. Rev. **22**, No. 7333 (1961).
[7] Sherman has recently given a new version of the second half of his proof. I am indebted to Dr. Sherman for showing me these results, as well as for several helpful discussions.
[8] H. Whitney, Comp. Math. **4**, 276 (1937).

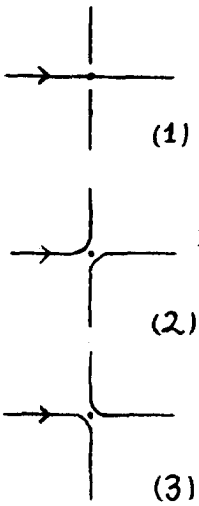


FIG. 6. The three possible behaviors of a path at a crossing.

The sum is taken over all admissible graphs on the lattice, the product over all *nonperiodic* paths.

Expand the right-hand side of (1) as a sum of terms: $W(p)W(p')\cdots = \pm\prod d_i^{\nu_i}$. The $\pm$ sign is the product[9] of the signs of the paths $p, p', \cdots$. The proof breaks naturally into two parts. In the first and easier part, we show that those terms having only $\nu_i = 0$ or 1 add up to $\sum I(G)$. The second part proves that all other terms add up to zero.

Proof, part 1: Consider those terms in the expansion of $\prod (1 + W(p))$ having no repeated lines ($\nu_i = 0, 1$). Each such term covers a certain set of lines which is clearly an admissible graph. Conversely let G be any admissible graph having m crossings. We first show that there are exactly 3^m terms covering G, i.e., satisfying $W(p)W(p')\cdots = \pm I(G)$.

Start at any point of G and trace out some path on G (without repeating lines). On first approaching each crossing of G, there are three possible choices of direction (see Fig. 6). Once the crossing has

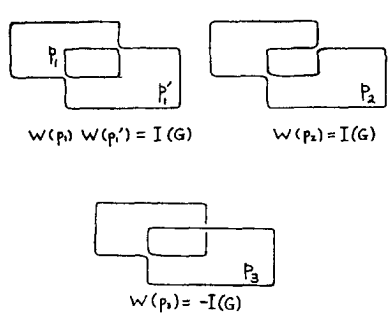


FIG. 7.

[9] Since two paths always intersect each other an even number of times, the sign of a term is also just $(-1)^c$ where c = total number of path crossings in the term.

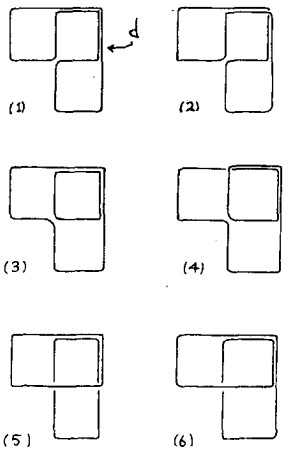


FIG. 8.

been traversed, no further choice remains there. Proceeding in this way until the whole of G is covered, one obtains a set of distinct, nonperiodic paths $p, p', \cdots$, and thus a term $W(p)W(p')\cdots$ covering G. Two such terms can differ only in their behavior at crossings. Since there are three choices at each crossing there are thus 3^m terms.

Next, observe that a given term with m_i "intersections" of type j (see Fig. 6, $j = 1, 2, 3$, $m_1 + m_2 + m_3 = m$) has sign $(-1)^{m_1}$ and there are $m!/m_1!\, m_2!\, m_3!$ such terms. Summing over all 3^m terms, we obtain

$$\sum (-1)^{m_1} m!/m_1!\, m_2!\, m_3! = (-1 + 1 + 1)^m = +1.$$

Thus G is counted correctly by the rhs of (1).

Example: For G as in Fig. 2 there are nine terms covering it, of which three are shown in Fig. 7.

Proof, part 2: All remaining terms in the expansion of $\prod (1 + W(p))$ are paths or products of paths in which at least one line is covered twice. Group together those terms which cover the same lines the same number of times. In such a group all the terms will have the identical product of d_i's, differing only in the sign. We now fix attention on a particular such group.

Choose some line, call it d, which is repeated more than once, say N times, by every term of this group. Then in any term $W(p)W(p')\cdots$ in the group, the paths $p, p', \cdots$ fall into one of two classes: (i) those containing d; (ii) those not containing d. For a path of type (i) imagine d

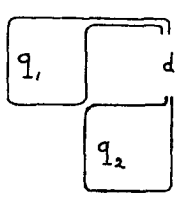


FIG. 9. The path segments corresponding to the first subgroup of Fig. 8.

to be temporarily removed from wherever it occurs. What remains is a set of *path segments* starting and finishing at the ends of d. Proceeding thus for each path of type (i) in the term we obtain a certain set of N path segments. Now there will be other terms in the group leading to the same set of path segments and also having the same paths of type (ii). Such terms will differ only in the manner in which the path segments are joined through d. Collect these terms into a *subgroup*. We shall show that the terms in a subgroup cancel out with each other.

Example: The six terms sketched in Fig. 8 belong to the same group (assuming each lies in the same place on the lattice). They are the only terms in the group. With d as shown ($N = 2$) the group breaks up into three subgroups (1)(2), (3)(4), and (5)(6). The subgroup (1)(2) has the path segments $\{q_1, q_2\}$ shown in Fig. 9, etc.

Consider then some definite subgroup with path segments $\{q_1, q_2, \cdots q_N\}$, [the paths of type (ii) play no role and can be ignored]. First we assume all the q_i to be distinct. Imagine the ends of the path segments to be slightly separated and fixed in some order at both the top[10] and bottom end of d. Then any way of joining the top ends of the $\{q_i\}$, through d, to the bottom ends leads to a term in the subgroup. This is not entirely obvious because the term might have a periodic or repeated path and would thus not occur in $\prod (1 + W(p))$, (such terms will be called *forbidden*). That this actually does not happen here is due to all the q_i being distinct. For the same reason, each of the $N!$ ways of joining the ends leads to a different term. To obtain the sign of a term in the subgroup, relative to the signs of other terms in the subgroup, one need only count the number of times the N d's cross each other in that term. To do this, suppose the ends to be numbered consecutively (at both top and bottom) from 1 to N. If the jth top end is joined to the π_jth bottom end, then the number of crossings at d in that term is equal to the minimum number of interchanges needed to put $123 \cdots N$ in the new order $\pi_1\pi_2\pi_3 \cdots \pi_N$. This number is even or odd depending on whether the permutation $\pi : j \to \pi_j$ is even or odd (by definition!). It is well known that of the $N!$ permutations, half are even the other half are odd. Thus half of the $N!$ ways lead to an even number of crossings and the other half to an odd number. This proves that cancellation occurs for this case.

[10] Arbitrarily call one end of d the "top" and the other end the "bottom."

Now in general not all the path segments will be distinct. Suppose that we have a subgroup Σ with path segments $\{q_1, \cdots, q_1, q_2, \cdots, q_2, \cdots, q_K, \cdots, q_K\}$ with q_1 occurring r_1 times, q_2 occurring r_2 times, $\cdots$, q_K occurring r_K times, and where $r_1 + r_2 \cdots + r_K = N$ and $r_1 > 1$. We use induction. Assume that cancellation has been proved for the subgroup Σ_0 with path segments $\{q_0, q_1, \cdots, q_1, q_2, \cdots, q_2, \cdots, q_K, \cdots, q_K\}$ with q_1 occurring $(r_1 - 1)$ times, q_2 occurring r_2 times, etc. Then we shall show that cancellation also occurs when q_0 is identified with q_1 to give Σ.

Note first that to any term in Σ must correspond r_1 terms in Σ_0. Each such term in Σ_0 is obtained by replacing one of the r_1 occurrences of q_1 by q_0. To see that this leads to r_1 different terms in Σ_0, recall that in general a term is forbidden, i.e., has repeated paths and/or periodic paths, if and only if one can start in two (or more) different places and still trace out the same path using the identical sequence of lines. Thus if some of the r_1 terms in Σ_0 were the same, it would mean that the original term was forbidden and therefore not in Σ.

The last step of the proof consists of showing that those terms in Σ_0, which do lead to forbidden terms, cancel in pairs. If, upon putting $q_0 = q_1$, some term $W(p)W(p') \cdots$ in Σ_0 gives a forbidden term, this means that the path containing q_0, say p, becomes either (1) periodic, of period $\omega \geq 2$, or (2) equal to some other path in the term. In case (1), p is of the form

$$p = (q_0 x q_1 x \cdots q_1 x),$$

where $q_1 x$ occurs $(\omega - 1)$ times and where $p_a = (q_1 x)$ is a nonperiodic subpath of p. These symbols are meant to imply the sequence and relative direction in which the various path segments, occurring in p and p_a, respectively, are traversed. x stands for some ordered sequence of q_i's and $\bar{q}_i$'s, including possibly further q_1's as well as $\bar{q}_1$'s. Here q_i represents not just the path segment but also a specific direction of traversing it. $\bar{q}_i$ represents the opposite direction. Although the overall direction of traversing a path is immaterial, it is of course essential to specify the relative direction in which the successive path segments are traversed. This explains the need for distinguishing q_i from $\bar{q}_i$ in such a symbol. Since neither cyclic permutation of the symbols nor interchanging all $q_i \leftrightarrow \bar{q}_i$ changes the path, we can, for convenience, always arrange to start with q_0.

Define

$$p_b = (q_0 x q_1 x \cdots q_1 x),$$

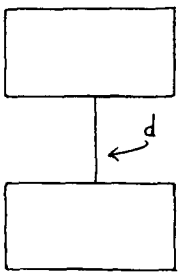


FIG. 10.

where q_1x occurs $(\omega - 2)$ times. Then to the above term $W(p)W(p') \cdots$ associate the term $W(p_a)W(p_b)W(p') \cdots$. This term is in Σ_0 and also leads to a forbidden term upon putting $q_0 = q_1$. Moreover, one easily sees that these two terms have opposite signs and therefore cancel. This method of pairing includes all such terms. The terms having paths of type (2) being cancelled by those of type (1) when $\omega = 2$.

Thus in going from Σ_0 to Σ, those terms in Σ_0 which do not give terms in Σ cancel among themselves. The remaining terms in Σ_0 collect in sets of r_1 elements, each element having the same sign and each set corresponding to one term in Σ. Therefore the terms in the subgroup Σ cancel.

This completes the proof of the theorem.

Example: Consider Fig. 10 with the line d covered four times and all other lines each covered twice. As is readily shown, this group consists of just four terms all in one subgroup (with d as shown). The path segments are $\{q_1, q_1, q_2, q_2\}$ as in Fig. 11. The four terms can be represented as follows:

(i) $W(p_1)W(p_1')$ where, $p_1 = (q_1q_2)$, $p_1' = (q_1\bar{q}_2)$,
(ii) $W(p_2)$,, $p_2 = (q_1q_2q_1\bar{q}_2)$,
(iii) $W(p_3)$,, $p_3 = (q_1q_2\bar{q}_1q_2)$,
(iv) $W(p_4)$,, $p_4 = (q_1q_2\bar{q}_1\bar{q}_2)$.

Term (ii) is sketched in Fig. 12. (i) and (iv) have sign -1, while (ii) and (iii) have sign $+1$.

If one considers a subgroup $\{q_0, q_1, q_2, q_2\}$, where q_0 differs from q_1, it is found to have 12 terms. On putting $q_0 = q_1$, four of these terms become

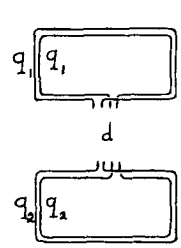


FIG. 11.

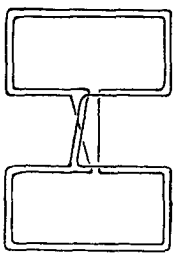


FIG. 12. The term (ii).

forbidden and cancel in pairs, namely

$(q_0q_2q_1q_2)$ cancels with $(q_0q_2)(q_1q_2)$,
$(q_0\bar{q}_2q_1\bar{q}_2)$,, ,, $(q_0\bar{q}_2)(q_1\bar{q}_2)$.

The remaining eight terms collect in sets of two elements and give the terms (i)–(iv), e.g., $(q_0q_2q_1\bar{q}_2)$ and $(q_1q_2q_0\bar{q}_2)$ both become (ii), etc.

III. CONCLUSIONS

The above result, and also its particular usefulness in solving the Ising problem exist only for a two-dimensional lattice. The theorem, however, can be extended to a three-dimensional lattice.

One approach would be to replace lines by *sides* (a side being the segment of area, bounded by four lines, in the three-dimensional lattice). Then an admissible "2-graph" would be a set of sides such that at every line, 0, 2, or 4 sides meet. The connection between such 2-graphs and the partition function of the three-dimensional Ising model is easily established [see Newell and Montroll (reference 1, p. 359)]. For the '2-paths" consider ordered "sequences" of sides such that to each of the four edges of every side is assigned a definite adjoining side. (At each edge there are just three possibilities since doubling back is forbidden as in the two-dimensional case.) Finally define the sign of a 2-path to be $(-1)^n$, where n is the number of times the path intersects itself [i.e. in cross section like Fig. 6, case (1)] at each edge of every side. With the obvious generalization of $I(G)$, $I(p)$, and $W(p)$, Eq. (1) again holds. The proof is virtually identical with that for the two-dimensional case.

Unfortunately, this result would not appear to be of much use in the Ising problem. One reason being that there seems to be no analogue to $\mathfrak{M}$ [see formula (A5) in the appendix] for counting 2-paths.

APPENDIX

For a square lattice with $\mathfrak{N}$ sites, the partition function for the Ising model is

$$Z(T) = \sum_{\{\sigma\}} \exp\left(K \sum_{\alpha\beta} \sigma_\alpha\sigma_\beta\right), \qquad \text{(A1)}$$

where $K = J/\kappa T$, T is the temperature, κ the Boltzmann constant, and J a parameter describing the strength of the interaction between neighboring sites. The first sum is over all $2^{\mathfrak{N}}$ possible configurations of the σ's.

There exists a well-known method, due to van der Waerden, [see Newell and Montroll, (reference 1, pp. 357–58)] for expressing (A1) as a sum over admissible graphs. The result is[11]

$$Z(T) = 2^{\mathfrak{N}}(\cosh K)^{2\mathfrak{N}}(1 + \sum_{G} I(G)), \qquad \text{(A2)}$$

where all $d_i = \tanh K$ in $I(G)$.

From (A2) and (1) we can write

$$\frac{1}{\mathfrak{N}} \ln Z(T) - \ln (2 \cosh^2 K) = \frac{1}{\mathfrak{N}} \sum_{p} \ln (1 + W(p)), \qquad \text{(A3)}$$

where all $d_i = \tanh K$ in $W(p)$. Expanding the logarithm, the rhs of (A3) becomes

$$\frac{1}{\mathfrak{N}} \sum_{p} (W(p) - \tfrac{1}{2}[W(p)]^2 + \tfrac{1}{3}[W(p)]^3 - \cdots). \qquad \text{(A4)}$$

In $\sum_p$, the first term is the sum of W over all nonperiodic paths, while the other terms just give the sum over all periodic paths. This follows since (a) any periodic path is made by repeating some nonperiodic path ω times, $\omega = 2, 3, \cdots$, and(b) from Whitney's result (see end of next paragraph), the sign of a periodic path is -1 if ω is even and equals the sign of its nonperiodic subpath if ω is odd. Those terms with period ω are weighted with the factor $1/\omega$. We now give a simple method of evaluating (A4)[12]

Define the 4×4 matrix

$$\mathfrak{M} = (\tanh K) \times \begin{vmatrix} h & 0 & \bar{\alpha}v & \alpha\bar{v} \\ 0 & \bar{h} & \alpha v & \bar{\alpha}\bar{v} \\ \alpha h & \bar{\alpha}\bar{h} & v & 0 \\ \bar{\alpha}h & \alpha\bar{h} & 0 & \bar{v} \end{vmatrix}, \qquad \text{(A5)}$$

where $\alpha = e^{-i\pi/4}$, $h = e^{i\theta_1}$, $v = e^{i\theta_2}$, and $\bar{\alpha}$, $\bar{h}$, $\bar{v}$ are their complex conjugates. Consider the following object:

$$-\tfrac{1}{2} \text{ Trace } \frac{1}{(2\pi)^2} \iint_0^{2\pi} \mathfrak{M}^l \, d\theta_1 \, d\theta_2. \qquad \text{(A6)}$$

[11] Assuming the lattice to be $\mathfrak{N}^{\frac{1}{2}} \times \mathfrak{N}^{\frac{1}{2}}$, the exponent of cosh K should more correctly be $2(\mathfrak{N} - \mathfrak{N}^{\frac{1}{2}})$. As $\mathfrak{N} \to \infty$ this distinction vanishes.

[12] See for example, H. N. V. Temperley, Phys. Rev. **103**, 1 (1956) and also reference 5. The above method is somewhat similar to that of R. P. Feynman in some lecture notes (unpublished).

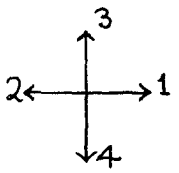


FIG. 13. Directions corresponding to values of the index i_j in Eq. (A7).

I maintain that this equals the sum of $W(p)$ over *all* (closed) paths of length l passing through some fixed site P_1. To see this, look at a term

$$\mathfrak{M}_{i_1 i_2}\mathfrak{M}_{i_2 i_3} \cdots \mathfrak{M}_{i_l i_{l+1}} \qquad (1 \leq i_j \leq 4) \qquad \text{(A7)}$$

in $\mathfrak{M}^l$. This defines the (not necessarily closed) path starting at P_1 and moving in direction i_2 (see Fig. 13) to the neighboring site P_2; and then moving in direction i_3 to P_3, etc. (see example in Fig. 14). The path corresponding to (A7) is closed if and only if there are an equal number of h's and $\bar{h}$'s, and an equal number of v's and $\bar{v}$'s. Thus the integral over θ_1 and θ_2 removes all nonclosed paths. The Trace ensures that the path returns to P_1 in the correct direction. The factor of $\frac{1}{2}$ occurs since each path is counted twice—once in each direction. Finally we have to see that each $W(p)$ is counted by (A6) with its correct sign [namely $(-1)^{n_p}$ where n_p is the number of times p crosses itself]. This is the purpose of the α's. Since $\mathfrak{M}$ contributes an α for each right-hand turn and an $\bar{\alpha}$ for each left-hand turn, the net effect in (A6) is a $(-1)^{t_p}$ where t_p = (number of clockwise revolutions) − (number of counterclockwise revolutions) made in traversing p. Now there is a simple but beautiful result, due to H. Whitney,[8] which says that $(-1)^{n_p} = -(-1)^{t_p}$. Thus, taking account of the sign of (A7), we see that the assertion following Eq. (A6) is justified.

On multiplying (A6) by $\mathfrak{N}/l$ and summing over l, we obtain all paths irrespective of starting point P_1.[13] Moreover, the periodic paths have a factor $1/\omega$, since a periodic path has only l/ω distinct starting points. Thus,

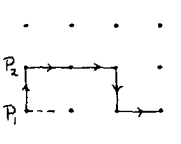


FIG. 14. The path corresponding to the element $\mathfrak{M}_{23}\mathfrak{M}_{31}\mathfrak{M}_{11}\mathfrak{M}_{14}\mathfrak{M}_{41}$ in $\mathfrak{M}^5$.

[13] This is only true if boundary effects are neglected, or if one assumes the lattice to be wound on a torus. In the latter case, Eq. (1) must be modified (see reference 5). The final result as $\mathfrak{N} \to \infty$ is the same.

$$\frac{1}{\mathfrak{N}} \sum_{p} \ln [1 + W(p)]$$

$$= \frac{1}{2} \frac{1}{(2\pi)^2} \text{Trace} \iint_0^{2\pi} \ln (I - \mathfrak{M})\, d\theta_1\, d\theta_2. \quad \text{(A8)}$$

Now Trace $[\ln (I - \mathfrak{M})] = \ln [\det (I - \mathfrak{M})]$, and the 4×4 determinant is easily evaluated, giving from (A3),

$$\frac{1}{\mathfrak{N}} \ln Z(T) = \ln 2 + \frac{1}{2(2\pi)^2} \iint_0^{2\pi} \ln [(\cosh 2K)^2 - \sinh 2K(\cos \theta_1 + \cos \theta_2)]\, d\theta_1\, d\theta_2. \quad \text{(A9)}$$

This is Onsager's formula.

(A9) is exact only if one ignores effects due to the lattice having a boundary (see references 11 and 13). This is of course a crucial point. It can be shown, however, that as $\mathfrak{N} \to \infty$ these effects vanish.[14]

The above method is readily extended to more complicated two-dimensional lattices, and also to the case where J differs for vertical and horizontal interactions.

The partition function is an analytic function of T (or K) with, among others, a branch point at $K = \ln (1 + \sqrt{2})$. This is the only singularity for real $T > 0$, and is the temperature at which the phase transition occurs.

[14] R. B. Potts and J. C. Ward, Progr. Theoret. Phys. (Kyoto) **13**, 38 (1955).

Approximate Evaluation of Feynman Functional Integrals

GERALD ROSEN
Southwest Research Institute, San Antonio, Texas
(Received 7 March 1963)

By considering a new set of functions for the integration domain, Feynman's functional-integral representation of the propagation kernel is evaluated approximately for values of the time that are not too large. The approximation theory is applied to systems of uncoupled and strongly coupled oscillators.

I. INTRODUCTION

AUGMENTING the canonical quantization procedure, the functional- or path-integral formalism of Feynman[1] offers an alternative way of solving dynamical problems in quantum mechanics. However, there are only a few significant applications of the Feynman formalism, mainly because it is difficult to extract numerical results from the functional integrals. Similar functional integrals provide formal (but not generally useful) solutions to problems in other branches of statistical physics.

This paper reports an approximation scheme that seems useful for certain dynamical problems, especially for dynamical problems in so-called strong coupling theories. Since the interaction energy is of the order of the "unperturbed" energy in strong coupling theories, the characteristic time for transitions between "unperturbed" states of the system is of the order of the "unperturbed" period of the system. Therefore, approximate transition rates ought to follow from an expression for the propagation kernel, valid approximately for values of the time that are not too large. Likewise, approximate energy eigenvalues of stationary states in strong coupling theories ought to follow from knowledge of the propagation kernel for values of the time that are not too large. Such an approximate expression for the propagation kernel is obtained from the associated Feynman functional integral in the work that follows. It should be noted that the point of view taken here contrasts with and complements the point of view taken in perturbation theory and S-matrix theory, where attention is focused on dynamical changes for values of the time that are very large in relation to the unperturbed period of the system. The approximation procedure in this paper is applicable to problems that are not handled appropriately by the other theories.

The arrangement of the paper is as follows. A brief review of the Feynman formalism for quantum dynamics is presented in Sec. II, pursuant to fix notation and facilitate reference to the equations of Feynman quantization theory. The approximation theory is developed in Sec. III with some supporting analysis in Appendix A. In Sec. IV, for purposes of illustration and corroboration, the approximation procedure is applied to a system of uncoupled simple harmonic oscillators. Then we turn to a useful application of the theory in Sec. V and treat a system of oscillators which are coupled strongly. An alternative but related approximation scheme is outlined in Appendix B, the latter procedure working better for some dynamical systems.

II. FEYNMAN FORMALISM

Consider a dynamical system with n degrees of freedom, described by a real n-tuple of coordinates, $q = (q_1, \cdots, q_m, \cdots, q_n)$. Each q_m ranges over the real numbers $[-\infty < q_m < +\infty]$ for $m = 1$ to n. In terms of q, the kinetic energy of the system reduces to the form

$$\frac{1}{2}\left(\frac{dq}{dt}\right)^2 \equiv \frac{1}{2}\sum_{m=1}^{n}\left(\frac{dq_m}{dt}\right)^2. \tag{2.1}$$

Also, the potential energy of the system $V(q)$ is a real scalar function of the coordinates q; $V(q)$ does not depend explicitly on the time t, nor depend on time derivatives of q. For such a dynamical system, the quantum mechanical state, represented by a wavefunction $\psi(q, t)$, evolves in time according to the Schrödinger equation

$$\left[i\frac{\partial}{\partial t} + \frac{1}{2}\left(\frac{\partial}{\partial q}\right)^2 - V(q)\right]\psi(q, t) = 0, \tag{2.2}$$

where

$$\left(\frac{\partial}{\partial q}\right)^2 \equiv \sum_{m=1}^{n}\frac{\partial^2}{\partial q_m^2}. \tag{2.3}$$

[1] R. P. Feynman, Rev. Mod. Phys. **20**, 367 (1948). Feynman quantization has been surveyed recently by: I. M. Gel'fand and A. M. Yaglom, J. Math. Phys. **1**, 48 (1960); S. G. Brush, Rev. Mod. Phys. **33**, 79 (1961). Also see the papers by: J. R. Klauder, Ann. Phys. (N. Y.) **11**, 123 (1960); S. S. Schweber, J. Math. Phys. **3**, 831 (1962).

Equation (2.2) is supplemented with the initial condition

$$\psi(q, 0) = \psi_0(q). \tag{2.4}$$

Then the solution to (2.2), (2.4) is obtained in terms of a propagation kernel

$$\psi(q, t) = \int K(q, q'; t)\psi_0(q')\, dq', \tag{2.5}$$

in which

$$dq' \equiv \prod_{m=1}^{n} dq'_m, \tag{2.6}$$

and the integration in (2.5) is over the entire physical domain of q, $[-\infty < q_m < +\infty]$ for $m = 1$ to n. In order for (2.5) to hold, the propagation kernel K satisfies the equations

$$\left[i\frac{\partial}{\partial t} + \frac{1}{2}\left(\frac{\partial}{\partial q}\right)^2 - V(q)\right]K(q, q'; t) = 0 \qquad (t > 0), \tag{2.7}$$

$$K(q, q'; 0) = \delta(q - q') \equiv \prod_{m=1}^{n} \delta(q_m - q'_m). \tag{2.8}$$

A useful property of K follows from (2.7) and (2.8), namely, the *semigroup composition law*

$$K(q, q'; t) = \int K(q, q''; t - t')K(q'', q'; t')\, dq'' \qquad (t \geq t' \geq 0). \tag{2.9}$$

By repeated application of (2.9), the propagation kernel for finite values of t is decomposed into an iteration of the propagation kernel for infinitesimal values of t, the propagation kernel for infinitesimal values of t being obtained easily from (2.7) and (2.8). In this way the solution to (2.7), (2.8) is expressed as a functional integral

$$K(q, q'; t) = (2\pi i t)^{-\frac{1}{2}n} \times \int_{\mathfrak{F}} \left\{\exp i \int_0^t \left[\frac{1}{2}\left(\frac{d\xi}{dS}\right)^2 - V(\xi)\right] dS\right\} D(\xi). \tag{2.10}$$

The domain for the ξ integration in (2.10), denoted by $\mathfrak{F}$, is the set of all real single-valued n-tuple functions $\xi(S)$, $0 \leq S \leq t$, such that $\xi(0) = q'$ and $\xi(t) = q$. A particular representation of the integration measure $D(\xi)$ is determined to within normalization (that may depend on t) by the property of translation invariance, $D(\xi) = D(\xi + \omega)$, in which $\omega = \omega(S)$, $\omega(0) = \omega(t) = 0$. Consistency with the composition law (2.9) fixes the normalization for a particular representation of $D(\xi)$. Usually the normalization is fixed automatically by the symmetry of the representation of $D(\xi)$ and the required existence of K as a finite integral.

By introducing some new variables $\tau \equiv S/t$, $\bar{q} \equiv q' + \tau(q - q')$, and $\alpha \equiv \xi - \bar{q}$, (2.10) becomes

$$K(q, q'; t) = K_0(q, q'; t)F(q, q'; t), \tag{2.11}$$

where

$$K_0(q, q'; t) = (2\pi i t)^{-\frac{1}{2}n} \exp\frac{i}{2t}\sum_{m=1}^{n}(q_m - q'_m)^2, \tag{2.12}$$

$$F(q, q'; t) \equiv \int_{\mathfrak{S}} \left\{\exp i \int_0^1 \left[\frac{1}{2t}\left(\frac{d\alpha}{d\tau}\right)^2 - tV(\bar{q} + \alpha)\right] d\tau\right\} D(\alpha). \tag{2.13}$$

The domain for the α integration in (2.13), denoted by $\mathfrak{S}$, is the set of all real single-valued n-tuple functions $\alpha(\tau)$, $0 \leq \tau \leq 1$, such that $\alpha(0) = \alpha(1) = 0$. Notice that K_0 given by (2.12) is the solution to (2.7), (2.8) for $V(q) \equiv 0$, that is, for the case of free motion.

III. APPROXIMATION THEORY

Equations (2.5) and (2.11) express the quantum dynamics of the system formally, but the application of the theory requires evaluation of the functional integral F defined by (2.13). Exact evaluation of F is possible with a few especially simple $V(q)$, but not in general. As a preliminary to the approximate evaluation of F for general $V(q)$, first consider $\mathfrak{K}(\lambda)$, a subset of $\mathfrak{S}$ depending on the real n-tuple parameter $\lambda = (\lambda_1, \cdots, \lambda_m, \cdots, \lambda_n)$, $[\lambda_m \geq 0]$. By definition, $\mathfrak{K}(\lambda)$ contains all real single-valued n-tuple functions $\alpha(\tau)$, $0 \leq \tau \leq 1$, such that

$$\alpha(0) = \alpha(1) = 0, \qquad \int_0^1 \left(\frac{d\alpha_m}{d\tau}\right)^2 d\tau \leq 2t\lambda_m, \quad \text{for } \alpha \in \mathfrak{K}(\lambda). \tag{3.1}$$

With respect to $\mathfrak{S}$, the *characteristic function* for the subset $\mathfrak{K}(\lambda)$ is

$$\chi(\lambda) = \prod_{m=1}^{n} \sigma\left(\lambda_m - \frac{1}{2t}\int_0^1 \left(\frac{d\alpha_m}{d\tau}\right)^2 d\tau\right), \tag{3.2}$$

in which σ is the stepfunction

$$\sigma(\nu) \equiv \begin{matrix} 0 & (\nu < 0), \\ 1 & (\nu \geq 0). \end{matrix} \tag{3.3}$$

A functional integrated over $\mathfrak{K}(\lambda)$ equals the functional integrated over $\mathfrak{S}$ after first being multiplied by (3.2). In particular, we have

$$W(q, q', \lambda; t)$$
$$\equiv \int_{\mathfrak{C}(\lambda)} \left\{\exp -it \int_0^1 V(\bar{q} + \alpha)\, d\tau\right\} D(\alpha)$$
$$= \int_{\mathcal{G}} x(\lambda) \left\{\exp -it \int_0^1 V(\bar{q} + \alpha)\, d\tau\right\} D(\alpha). \quad (3.4)$$

Also note that (3.2) gives

$$\Lambda x(\lambda) = \exp \frac{i}{2t} \sum_{m=1}^{n} \int_0^1 \left(\frac{d\alpha_m}{d\tau}\right)^2 d\tau$$
$$\equiv \exp \frac{i}{2t} \int_0^1 \left(\frac{d\alpha}{d\tau}\right)^2 d\tau, \quad (3.5)$$

where the integration operator Λ, essentially an n-fold Laplace transformation, is defined by

$$\Lambda f(\lambda) \equiv (-i)^n \left(\prod_{m=1}^{n} \operatorname*{Lim}_{\epsilon_m \to 0+}\right) \int_0^\infty \cdots \int_0^\infty f(\lambda)$$
$$\times \prod_{m=1}^{n} \{\exp (i\lambda_m - \epsilon_m \lambda_m)\}\, d\lambda_m. \quad (3.6)$$

By letting Λ act on (3.4) and making use of (3.5), the functional integral (2.13) is expressed in the form

$$F(q, q'; t) = \Lambda W(q, q', \lambda; t). \quad (3.7)$$

Now consider $\hat{\mathfrak{C}}(\lambda)$, another subset of $\mathcal{G}$ that, like $\mathfrak{C}(\lambda)$, depends on the real n-tuple parameter λ. By definition, $\hat{\mathfrak{C}}(\lambda)$ contains all real single-valued n-tuple functions $\alpha(\tau)$, $0 \leq \tau \leq 1$, such that

$$|\alpha_m(\tau)| \leq L_m(\tau) \equiv [3t\lambda_m(\tau - \tau^2)]^{\frac{1}{2}}$$
$$\text{for} \quad \alpha \in \hat{\mathfrak{C}}(\lambda). \quad (3.8)$$

$\mathfrak{C}(\lambda)$ is in fact a subset of $\hat{\mathfrak{C}}(\lambda)$. We prove a stronger relationship, that $\mathfrak{C}(\lambda)$ is a subset of $\hat{\mathfrak{C}}(\frac{1}{3}2\lambda)$. Letting μ denote a number greater than zero but less than unity, the defining Eqs. (3.1) for $\mathfrak{C}(\lambda)$ imply

$$2t\lambda_m \geq \int_0^1 \left(\frac{d\alpha_m}{d\tau}\right)^2 d\tau$$
$$= \int_0^\mu \left(\frac{d\alpha_m}{d\tau} - \frac{\alpha_m(\mu)}{\mu} + \frac{\alpha_m(\mu)}{\mu}\right)^2 d\tau$$
$$+ \int_\mu^1 \left(\frac{d\alpha_m}{d\tau} + \frac{\alpha_m(\mu)}{1-\mu} - \frac{\alpha_m(\mu)}{1-\mu}\right)^2 d\tau$$
$$= \int_0^\mu \left(\frac{d\alpha_m}{d\tau} - \frac{\alpha_m(\mu)}{\mu}\right)^2 d\tau$$
$$+ \frac{[\alpha_m(\mu)]^2}{\mu} + \int_\mu^1 \left(\frac{d\alpha_m}{d\tau} + \frac{\alpha_m(\mu)}{1-\mu}\right)^2 d\tau$$
$$+ \frac{[\alpha_m(\mu)]^2}{1-\mu} \geq \frac{[\alpha_m(\mu)]^2}{\mu - \mu^2} \qquad (0 < \mu < 1). \quad (3.9)$$

By combining (3.9) with the conditions $\alpha_m(0) = \alpha_m(1) = 0$, we have for $0 \leq \tau \leq 1$,

$$|\alpha_m(\tau)| \leq [2t\lambda_m(\tau - \tau^2)]^{\frac{1}{2}} \quad (3.10)$$

if α is contained in $\mathfrak{C}(\lambda)$. Hence, a comparison of (3.10) and (3.8) shows that $\mathfrak{C}(\lambda)$ is a subset of $\hat{\mathfrak{C}}(\frac{1}{3}2\lambda)$, which is obviously a subset of $\hat{\mathfrak{C}}(\lambda)$.

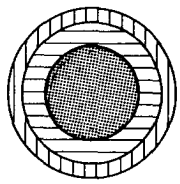


FIG. 1. Schematic representation of $\hat{\mathfrak{C}}(\lambda)$. In order to elucidate the mathematical result (3.13) in a physical way, the set of functions $\hat{\mathfrak{C}}(\lambda)$ is partitioned into the three disjoint subsets $[\hat{\mathfrak{C}}(\lambda) - \hat{\mathfrak{C}}(\frac{1}{3}2\lambda)]$ (the region with vertical shading), $[\hat{\mathfrak{C}}(\frac{1}{3}2\lambda) - \mathfrak{C}(\lambda)]$ (the region with horizontal shading), and $\mathfrak{C}(\lambda)$ (the region with dotted shading). Each subset of $\hat{\mathfrak{C}}(\lambda)$ makes a contribution to $\hat{F}$ defined by (3.12) and (3.11). However, these subsets are similar to half-period zones in diffraction theory, in the sense that contributions to $\hat{F}$ from the two *outer* subsets effectively cancel and leave only the contribution from the *inner* subset $\mathfrak{C}(\lambda)$.

In analogy to (3.4) and (3.7), let us introduce

$$\hat{W}(q, q', \lambda; t)$$
$$\equiv \int_{\hat{\mathfrak{C}}(\lambda)} \left\{\exp -it \int_0^1 V(\bar{q} + \alpha)\, d\tau\right\} D(\alpha), \quad (3.11)$$

and

$$\hat{F}(q, q'; t) \equiv \Lambda \hat{W}(q, q', \lambda; t). \quad (3.12)$$

Then some straightforward analysis reveals our main result

$$F(q, q'; t) = \hat{F}(q, q'; t) + O(t^3), \quad (3.13)$$

so $\hat{F}$ approximates F for values of t that are not too large (see Fig. 1). No matter what form is prescribed for the potential energy $V(q)$, the quantities F and $\hat{F}$ have the same initial "position," initial "velocity," and initial "acceleration."

To prove (3.13), first the functional integration in (3.11) is performed exactly. With customary symbolic notation, we have

$$\left\{\exp -it \int_0^1 V(\bar{q} + \alpha)\, d\tau\right\}$$
$$= \prod_{0<\tau<1} [1 - itV(\bar{q} + \alpha)\, d\tau], \quad (3.14)$$

$$D(\alpha) = \prod_{0<\tau<1} \left(\prod_{m=1}^{n} \frac{d\alpha_m(\tau)}{A_m(\tau)}\right). \quad (3.15)$$

Equation (3.11) becomes

$$\hat{W}(q, q', \lambda; t)$$
$$= \prod_{0<\tau<1} \int_{-L_1(\tau)}^{L_1(\tau)} \cdots \int_{-L_n(\tau)}^{L_n(\tau)} [1 - itV(\bar{q} + \alpha)\, d\tau]$$
$$\times \prod_{m=1}^{n} \frac{d\alpha_m(\tau)}{A_m(\tau)} = \exp -it \int_0^1 [\Omega V(\bar{q} + \alpha)]\, d\tau, \quad (3.16)$$

in which

$$\Omega V(\bar{q} + \alpha) \equiv \int_{-L_1(\tau)}^{L_1(\tau)} \cdots \int_{-L_n(\tau)}^{L_n(\tau)} V(\bar{q} + \alpha) \prod_{m=1}^{n} \frac{d\alpha_m}{2L_m(\tau)}, \quad (3.17)$$

and the normalization factors are fixed by

$$A_m(\tau) = 2L_m(\tau) \quad (3.18)$$

in order for (3.16) to exist as a finite integral.

The proof of (3.13) follows by expanding $V(\bar{q} + \alpha)$ in (3.16) as a Taylor series in α. One obtains

$$\hat{W}(q, q', \lambda; t) = \left\{ \exp\left[-it \int_0^1 V(\bar{q})\, d\tau - \frac{it^2}{2} \times \sum_{m=1}^{n} \lambda_m \int_0^1 (\tau - \tau^2) \frac{\partial^2 V(\bar{q})}{\partial \bar{q}_m^2}\, d\tau \right] \right\} + O(t^3), \quad (3.19)$$

so that (3.12) is

$$\hat{F}(q, q'; t) = 1 - it \int_0^1 V(\bar{q})\, d\tau - \frac{t^2}{2} \left[\int_0^1 V(\bar{q})\, d\tau \right]^2 + \frac{t^2}{2} \int_0^1 (\tau - \tau^2) \rho(\bar{q})\, d\tau + O(t^3), \quad (3.20)$$

where

$$\rho(q) \equiv \sum_{m=1}^{n} \frac{\partial^2 V(q)}{\partial q_m^2}. \quad (3.21)$$

Hence, by comparing (3.20) and the result for F, Eq. (A14) in Appendix A, we find the relation (3.13).

A first approximation to the propagation kernel is suggested by (2.11) and (3.13). Let us introduce

$$K^{(1)}(q, q'; t) \equiv K_0(q, q'; t) \hat{F}(q, q'; t), \quad (3.22)$$

where, by (3.12) and (3.16), we have

$$\hat{F}(q, q'; t) = \Lambda \exp -it \int_0^1 [\Omega V(\bar{q} + \alpha)]\, d\tau, \quad (3.23)$$

with the (n-fold ordinary) integration operators Λ and Ω given by (3.6), and (3.17), respectively. By virtue of (3.13), $K^{(1)}$ defined by (3.22) approximates K for values of t that are not too large. Refinements of this estimate follow from (3.22) by evoking the composition law (2.9) with $t' = \frac{1}{2}t$:

$$K^{(N+1)}(q, q'; t) \equiv \int K^{(N)}(q, q''; \tfrac{1}{2}t) \times K^{(N)}(q'', q'; \tfrac{1}{2}t)\, dq'' \qquad (N \geq 1). \quad (3.24)$$

The first approximation (3.22) is ordinarily sufficient for applications of the theory, as exemplified in the following sections.

IV. HARMONIC OSCILLATORS

The potential energy for a system of uncoupled simple harmonic oscillators is

$$V(q) = \frac{1}{2} \sum_{m=1}^{n} \omega_m^2 q_m^2, \quad (4.1)$$

in which the ω's are positive constants. Exact evaluation of the propagation kernel (2.10) with the potential energy (4.1) gives the well-known result[2]

$$K(q, q'; t) = \prod_{m=1}^{n} K_{\omega_m}(q_m, q_m'; t), \quad (4.2)$$

where

$$K_\omega(x, y; t) = (\omega/2\pi i \sin \omega t)^{\frac{1}{2}} \exp\{i\omega \times [(\tan \omega t)^{-1} \tfrac{1}{2}(x^2 + y^2) - (\sin \omega t)^{-1} xy]\}. \quad (4.3)$$

Let us check the approximation theory for the case of uncoupled simple harmonic oscillators. Equation (3.17) with (4.1) yields

$$\Omega V(\bar{q} + \alpha) = \frac{1}{2} \sum_{m=1}^{n} \omega_m^2 (\bar{q}_m^2 + t(\tau - \tau^2)\lambda_m). \quad (4.4)$$

By recalling that $\bar{q}_m \equiv q_m' + \tau(q_m - q_m')$, we obtain

$$\int_0^1 [\Omega V(\bar{q} + \alpha)]\, d\tau = \frac{1}{6} \sum_{m=1}^{n} \omega_m^2 \left(q_m^2 + q_m q_m' + q_m'^2 + \frac{t}{2} \lambda_m \right), \quad (4.5)$$

and (3.23) becomes

$$\hat{F}(q, q'; t) = \prod_{m=1}^{n} \left[\left(1 - \frac{\omega_m^2 t^2}{12} \right)^{-1} \times \exp -\frac{i\omega_m^2 t}{6} (q_m^2 + q_m q_m' + q_m'^2) \right]. \quad (4.6)$$

Thus, the first approximation to the propagation kernel, given by (3.22), is

$$K^{(1)}(q, q'; t) = \prod_{m=1}^{n} K_{\omega_m}^{(1)}(q_m, q_m'; t), \quad (4.7)$$

[2] In spite of formal similarity between associated Feynman and Wiener propagation kernels, the former is not determined completely by letting $t \to -it$ in an (evaluated) expression for the latter, as illustrated by considering the square root in (4.3). Expansion of the propagation kernel (4.3) in terms of a complete orthonormal set of energy eigenfunctions [see (A17) in the Appendix] with the energy eigenvalues $E_k = (k + \frac{1}{2})\omega$ shows that (4.3) must be antiperiodic in t:

$$K_\omega(x, y; t + (2\pi/\omega)) = -K_\omega(x, y; t).$$

Hence the physical branch of the square root in (4.3) is rather complicated, for the sign of the square root changes at every other singularity as t increases, namely, values of t for which $\omega t/\pi$ equals an even integer. Obtainable by careful evaluation of the Feynman functional integral, the physical branch of the square root in (4.3) cannot be determined from the associated Wiener propagation kernel with just the simple replacement $t \to -it$.

where

$$K_\omega^{(1)}(x, y; t) = (2\pi i t)^{-\frac{1}{2}}\left(1 - \frac{\omega^2 t^2}{12}\right)^{-1} \times \exp i\omega\left[\left(\frac{1}{\omega t} - \frac{\omega t}{3}\right)\frac{(x^2 + y^2)}{2} - \left(\frac{1}{\omega t} + \frac{\omega t}{6}\right)xy\right]. \tag{4.8}$$

Equation (4.7) expresses a feature of the approximation theory that is quite general and desirable: Uncoupled classical motion gives uncoupled quantum propagation, in accord with exact quantum dynamics.[3] A comparison of (4.3) and (4.8) reveals the accuracy of the approximation, indicating that transition probability coefficients (i.e., the transition rate per unit time) between nonstationary states of the system can be computed by using (4.7), (4.8). Approximate stationary states with their energy eigenvalues can be obtained by solving the associated integral equation which involves (4.7) and (4.8)

V. OSCILLATORS WITH STRONG COUPLING

Let us apply the approximation procedure to a dynamical system with the potential energy

$$V(q) = \frac{1}{2}\sum_{m=1}^{n}\omega_m^2 q_m^2 + \frac{1}{3}\sum_{p,r,s=1}^{n}\beta_{prs}q_p q_r q_s. \tag{5.1}$$

As in the preceding section, the ω's in (5.1) are positive constants. The real and constant β's in (5.1) are completely symmetrical with respect to permutations of their three subscript indices. For meaningful application of the approximation theory, it is assumed that $|\beta_{prs}|$ is at least of the order $(\omega_p\omega_r\omega_s)^{5/6}$ for each set of indices. Thus the potential energy (5.1) characterizes a system of essentially coupled oscillators with cubic interaction.

Equation (3.17) with (5.1) produces

$$\Omega V(\bar{q} + \alpha) = V(\bar{q}) + \frac{t}{2}(\tau - \tau^2)\sum_{m=1}^{n}(\omega_m^2 + 2\beta_m\bar{q}_m)\lambda_m, \tag{5.2}$$

in which

$$\beta_m \equiv \sum_{p=1}^{n}\beta_{mpp}. \tag{5.3}$$

Thus we get

$$\int_0^1 [\Omega V(\bar{q} + \alpha)]\,d\tau = \int_0^1 V(\bar{q})\,d\tau + \frac{t}{12}\sum_{m=1}^{n}[\omega_m^2 + \beta_m(q_m + q'_m)]\lambda_m, \tag{5.4}$$

with

$$\int_0^1 V(\bar{q})\,d\tau = \frac{1}{6}\sum_{m=1}^{n}\omega_m^2(q_m^2 + q_m q'_m + q'^2_m) + \frac{1}{12}\sum_{p,r,s=1}^{n}\beta_{prs}(q_p q_r q_s + q_p q_r q'_s + q_p q'_r q'_s + q'_p q'_r q'_s). \tag{5.5}$$

Then (3.23) gives

$$\hat{F}(q, q'; t) = \left[\prod_{m=1}^{n}\left(1 - \frac{t^2}{12}[\omega_m^2 + \beta_m(q_m + q'_m)]\right)^{-1}\right] \times \exp -it\int_0^1 V(\bar{q})\,d\tau, \tag{5.6}$$

and (3.22), the first approximation to the propagation kernel, is

$$K^{(1)}(q, q'; t) = (2\pi i t)^{-n/2} \times \left[\prod_{m=1}^{n}\left(1 - \frac{t^2}{12}[\omega_m^2 + \beta_m(q_m + q'_m)]\right)^{-1} \times \exp\left\{\frac{i}{2t}(q_m - q'_m)^2 - \frac{i\omega_m^2 t}{6}(q_m^2 + q_m q'_m + q'^2_m)\right\}\right] \times \exp\left\{-\frac{it}{12}\sum_{p,r,s=1}^{n}\beta_{prs}(q_p q_r q_s + q_p q_r q'_s + q_p q'_r q'_s + q'_p q'_r q'_s)\right\}. \tag{5.7}$$

With the aid of numerical methods, the transition probability between prescribed states of the system can be computed from (5.7), $2n$ ordinary integrations being involved in such a numerical computation. In order to get convergent results, the replacement $\omega_m^2 \to (\omega_m^2 - i\epsilon)$, $(\epsilon > 0)$, is indicated, and then well-defined contributions are obtained from the n poles in (5.7), poles that now appear at

$$\beta_m(q_m + q'_m) + \omega_m^2 = (12/t^2) + i\epsilon. \tag{5.8}$$

The transition probability coefficient can be estimated numerically from the slope of the transition probability vs time curve for values of t that are not too large.

APPENDIX A. TEMPORAL SERIES SOLUTION FOR F

Consider the solution to Eqs. (2.7), (2.8) in the form (2.11):

$$K(q, q'; t) = K_0(q, q'; t)F(q, q'; t), \tag{A1}$$

[3] By no means an automatic feature of a useful approximation procedure, the less complicated method in Appendix B (although more appropriate for some dynamical systems with n large) injects a certain amount of nonphysical coupling in the approximate propagation kernel.

where K_0 is the solution for $V(q) \equiv 0$,

$$K_0(q, q'; t) = (2\pi i t)^{-\frac{1}{2}n} \exp \frac{i}{2t} \sum_{m=1}^{n} (q_m - q'_m)^2, \quad \text{(A2)}$$

and F is analytic in t in the neighborhood of $t = 0$:

$$F(q, q'; t) = \sum_{k=0}^{\infty} (it)^k f_k(q, q'). \quad \text{(A3)}$$

Equation (2.8) is satisfied most simply by setting

$$f_0(q, q') \equiv 1. \quad \text{(A4)}$$

By substituting (A1) into (2.7), we obtain equations for $f_k(q, q')$ with $k \geq 1$:

$$\left[k + \sum_{m=1}^{n} (q_m - q'_m) \frac{\partial}{\partial q_m}\right] f_k(q, q') = g_{k-1}(q, q'), \quad \text{(A5)}$$

$$g_{k-1}(q, q') \equiv \left[\frac{1}{2}\left(\frac{\partial}{\partial q}\right)^2 - V(q)\right] f_{k-1}(q, q'). \quad \text{(A6)}$$

For application of (A1) in the formula (2.5), each $f_k(q, q')$ is required to be integrable in the neighborhood of $q = q'$. The solutions of (A5) that may have this property are[4]

$$f_k(q, q') = \int_0^1 \tau^{k-1} g_{k-1}(\bar{q}, q')\, d\tau \quad (k \geq 1), \quad \text{(A7)}$$

in which

$$\bar{q} \equiv q' + \tau(q - q'). \quad \text{(A8)}$$

To prove that (A7) satisfies (A5), the Euler homogeneity equation for a function of $\bar{q}$,

$$\left[\tau \frac{\partial}{\partial \tau} - \sum_{m=1}^{n} (q_m - q'_m) \frac{\partial}{\partial q_m}\right] g_{k-1}(\bar{q}, q') = 0, \quad \text{(A9)}$$

is multiplied by τ^{k-1} and then integrated with respect to the parameter τ over the range zero to unity. After an integration by parts, we have

$$\left[k + \sum_{m=1}^{n} (q_m - q'_m) \frac{\partial}{\partial q_m}\right] \int_0^1 \tau^{k-1} g_{k-1}(\bar{q}, q')\, d\tau = g_{k-1}(q, q') \quad (k \geq 1), \quad \text{(A10)}$$

which proves that (A7) satisfies (A5). For $k = 1$, (A7) and (A6) produce

$$f_1(q, q') = -\int_0^1 V(\bar{q})\, d\tau, \quad \text{(A11)}$$

and for $k = 2$, with a few obvious transformations of the definite integrals, we obtain

$$f_2(q, q') = \frac{1}{2}\left[\int_0^1 V(\bar{q})\, d\tau\right]^2 - \frac{1}{2}\int_0^1 (\tau - \tau^2)\rho(\bar{q})\, d\tau \quad \text{(A12)}$$

by putting

$$\rho(q) \equiv \sum_{m=1}^{n} \frac{\partial^2 V(q)}{\partial q_m^2}. \quad \text{(A13)}$$

Hence, (A3) becomes

$$F(q, q'; t) = 1 - it\int_0^1 V(\bar{q})\, d\tau - \frac{t^2}{2}\left[\int_0^1 V(\bar{q})\, d\tau\right]^2 + \frac{t^2}{2}\int_0^1 (\tau - \tau^2)\rho(\bar{q})\, d\tau + O(t^3). \quad \text{(A14)}$$

For most $V(q)$ the series (A3) converges rapidly, provided that t is not too large. In fact for large values of k, (A7) gives the asymptotic relation

$$f_k(q, q') \doteq (1/k) g_{k-1}(q, q') \quad (k \text{ large}), \quad \text{(A15)}$$

and the appropriate asymptotic solution of (A6) is

$$f_k(q, q') \doteq \frac{(2\pi)^{\frac{1}{2}n}(-E_0)^{k-\frac{1}{2}n}}{\Gamma(k + 1 - \frac{1}{2}n)} u_0(q) u_0^*(q') \quad (k \text{ large}), \quad \text{(A16)}$$

with the constant parameter E_0 interpreted as the energy eigenvalue of the ground state of the system, $u_0(q)$. Note that (A16) is consistent with the asymptotic behavior of the formal solution,

$$K(q, q'; t) = \sum_{k=0}^{\infty} e^{-iE_k t} u_k(q) u_k^*(q'), \quad \text{(A17)}$$

in which the energy eigenvalues are arranged in a nondecreasing sequence $E_k \leq E_{k+1}$, and the u's are a complete orthonormal set of energy eigenfunctions.

APPENDIX B. ALTERNATIVE APPROXIMATION THEORY

For a dynamical system with many degrees of freedom (i.e., n large) and physically coupled motion, the n-fold Laplace transformation Λ defined by (3.6) and appearing in (3.23) is not always practicable. A closely related approximation procedure features a onefold Laplace transformation (denoted below by Δ) in place of Λ for all values of n. This alternative approximation procedure, described here, seems more promising for some dynamical systems with n large.

First consider $\mathcal{G}(\delta)$, a subset of $\mathcal{G}$ depending on the real nonnegative parameter δ. By definition, $\mathcal{G}(\delta)$ contains all real single-valued n-tuple functions $\alpha(\tau)$, $0 \leq \tau \leq 1$, such that

[4] Conditions on $V(q)$ are required in order to guarantee the existence of an $f_k(q, q')$ that is integrable in the neighborhood of $q = q'$, or equivalently, the existence of a solution with the form (A7).

$$\int_0^1 \left(\frac{d\alpha}{d\tau}\right)^2 d\tau \equiv \sum_{m=1}^{n} \int_0^1 \left(\frac{d\alpha_m}{d\tau}\right)^2 d\tau \leq 2t\delta$$

$$\alpha(0) = \alpha(1) = 0 \quad \text{for} \quad \alpha \in \mathcal{G}(\delta). \tag{B1}$$

By comparing (B1) and (3.1), we see that $\mathcal{G}(\delta) = \mathcal{K}(\lambda)$ for $n = 1$ and $\delta \equiv \lambda_1$. The relation analogous to (3.4) is

$$W(q, q'; \delta, t) \equiv \int_{\mathcal{G}(\delta)} \left\{\exp -it \int_0^1 V(\bar{q} + \alpha)\, d\tau\right\} D(\alpha), \tag{B2}$$

from which it follows that

$$F(q, q'; t) = \Delta W(q, q'; \delta, t), \tag{B3}$$

where the Laplace transformation Δ is given by

$$\Delta f(\delta) \equiv -i(\lim_{\epsilon \to 0+}) \int_0^\infty f(\delta)\{\exp(i\delta - \epsilon\delta)\}\, d\delta. \tag{B4}$$

Now consider $\tilde{\mathcal{G}}(\delta)$, another subset of $\mathcal{G}$ that, like $\mathcal{G}(\delta)$, depends on the real nonnegative parameter δ. By definition, $\tilde{\mathcal{G}}(\delta)$ contains all real single-valued n-tuple functions $\alpha(\tau)$, $0 \leq \tau \leq 1$, such that

$$\left.\begin{aligned} &||\alpha(\tau)|| \\ &\quad \equiv \left[\sum_{m=1}^{n} (\alpha_m(\tau))^2\right]^{\frac{1}{2}} \leq r(\tau) \\ &r(\tau) \equiv [(n+2)t\delta(\tau - \tau^2)]^{\frac{1}{2}} \end{aligned}\right\} \text{for } \alpha \in \tilde{\mathcal{G}}(\delta). \tag{B5}$$

Introducing

$$\tilde{W}(q, q'; \delta, t) \equiv \int_{\tilde{\mathcal{G}}(\delta)} \left\{\exp -it \int_0^1 V(\bar{q} + \alpha)\, d\tau\right\} D(\alpha), \tag{B6}$$

relations similar to (3.14)–(3.16) produce

$$\tilde{W}(q, q'; \delta, t) = \exp -it \int_0^1 [\Phi V(\bar{q} + \alpha)]\, d\tau, \tag{B7}$$

in which

$$\Phi V(\bar{q} + \alpha) \equiv \left\{\frac{\Gamma(\frac{1}{2}n + 1)}{[\pi^{\frac{1}{2}} r(\tau)]^n}\right\} \times \int \cdots \int_{||\alpha(\tau)|| \leq r(\tau)} V(\bar{q} + \alpha) \prod_{m=1}^{n} d\alpha_m. \tag{B8}$$

Note that the prefactor in curly brackets which normalizes the integral in (B8) is just the reciprocal of the volume of an n sphere of radius $r(\tau)$. Finally, by introducing

$$\tilde{F}(q, q'; t) \equiv \Delta \tilde{W}(q, q'; \delta, t), \tag{B9}$$

we have

$$\tilde{F}(q, q'; t) = \Delta \exp -it \int_0^1 [\Phi V(\bar{q} + \alpha)]\, d\tau, \tag{B10}$$

and then Appendix A and relations similar to (3.19), (3.20) yield

$$F(q, q'; t) = \tilde{F}(q, q'; t) + O(t^3), \tag{B11}$$

so $\tilde{F}$ approximates F for values of t that are not too large. Consequently, $\tilde{F}$ may be used in place of $\hat{F}$ in the approximation to the progapation kernel (3.22).

Equation (B10), the formula for $\tilde{F}$, can be applied more easily than Eq. (3.23), the formula for $\hat{F}$, to some dynamical systems with n large and physically coupled motion. However, the approximate propagation kernel obtained with (B10) contains a certain amount of nonphysical coupling, an undesirable byproduct of the alternative approximation procedure. The approximation method in the body of the paper is generally preferable, because it gives uncoupled quantum dynamics for systems with uncoupled classical motion.

ACKNOWLEDGMENT

The author is grateful to Professor C. W. Misner for a critical examination of the manuscript and several helpful suggestions.

New Small-Angle X-Ray Scattering Calculation

W. S. ROTHWELL
Materials Sciences Laboratories, Research Laboratories, Lockheed Missiles and Space Company, Palo Alto, California
(Received 16 April 1963)

A method due to Hardy for evaluating finite integrals is used to obtain the structure factors for several rotationally symmetric shapes. The factor for particles formed by joining paraboloids of revolution is shown to be proportional to a second-order Lommel function of two variables.

INTRODUCTION

THE intensity of coherent scattering of radiation from a single fixed particle is given by[1]

$$I = I_e n^2 f^2, \tag{1}$$

where I_e is the scattered intensity from a single electron, n is the total number of electrons in the particle, $f^2 = ff^*$ and f is a structure factor characteristic of the particle which can be expressed as

$$f = \int_V e^{i\delta} P\, dV, \tag{2}$$

where PdV is the probability of finding an electron in volume dV with phase δ and the integration is taken over the volume of the particle. By obtaining f as a function of the orientation of the particle, the intensity of scattering from N particles of arbitrary orientation sufficiently far apart to scatter independently of one another is given by N times the average of the intensity over all orientations:

$$I = (I_e Nn^2)\frac{1}{4\pi}\int_\Omega f^2\, d\Omega.$$

Choosing $d\Omega = 2\pi R(\sin\gamma)R\, d\gamma/R^2$,

$$I = I_e Nn^2 \int_0^{\pi/2} f^2 \sin\gamma\, d\gamma. \tag{3}$$

Structure factors and scattering intensities have been obtained for spheres, cylinders, and ellipsoids of revolution.[2] It is the purpose of the following calculations to show how to obtain the structure factors for these shapes and a previously untreated shape by using f in a general form for particles of cylindrical and centro-symmetry[3] and Hardy's method of evaluating finite integrals.[4]

[1] A. H. Compton and S. K. Allison, *X-Rays in Theory and Experiment* (Van Nostrand: D. Van Nostrand, Inc., Princeton, New Jersey, 1948), p. 138.
[2] A. Guinier and G. Fournet, *Small-Angle Scattering of X-Rays* (John Wiley & Sons, Inc., 1955), p. 19.
[3] P. W. Schmidt (unpublished).
[4] G. N. Watson, *Theory of Bessel Functions* (Cambridge University Press, New York, 1952), p. 382.

I. GENERALIZED FORM OF STRUCTURE FACTOR

In the customary notation,

$$\delta = (4\pi R/\lambda)\sin\tfrac{1}{2}\phi = kR, \tag{4}$$

where λ is the wavelength of the radiation, ϕ is the scattering angle, and R is the distance from the origin to a point on the coplanar bisector of the angle between the paths of the incident and scattered radiation. This point is the intersection with the bisector of a perpendicular plane which passes through the position of the volume element, dV. In Fig. 1 the relationship is shown between R and a cylindrical coordinate system (r, α, z) for the particle with the z axis inclined at an angle, γ, to the bisector and α measured from the plane containing z and R. It can be seen that

$$R = z\cos\gamma + r\cos\alpha\sin\gamma. \tag{5}$$

Assuming a uniform electron density,

$$P\, dV = r\, dr\, d\alpha\, dz/V. \tag{6}$$

For shapes with cylindrical symmetry about z and centrosymmetry about the origin, let z vary between $\pm F(r)$ where $F(r)$ is a continuous single-valued function and r varies between 0 and a. Substituting (4), (5), and (6) into (2),

$$f = \frac{1}{V}\int_0^{2\pi} d\alpha \int_0^a r\, dr \times e^{ikr\sin\gamma\cos\alpha}\int_{-F(r)}^{F(r)} e^{ikz\cos\gamma}\, dz. \tag{7}$$

The last integral is just

$$2\sin[kF(r)\cos\gamma]/k\cos\gamma.$$

Noting that

$$J_0(x) = \frac{1}{\pi}\int_0^\pi e^{ix\cos\psi}\, d\psi, \tag{8}$$

(7) becomes

$$f = \frac{4\pi}{Vk\cos\gamma}\int_0^a J_0(kr\sin\gamma)\sin[kF(r)\cos\gamma]r\,dr. \quad (9)$$

II. HARDY'S METHOD

If $g(u)$ is a continuous function of u when $u > 0$ (and does not contain t), such that for all sufficiently large positive values of t

$$\int_0^\infty e^{-tu^2}g(u)\,du = 0, \quad (10)$$

then $g(u)$ is identically zero. Let $g(u) = g_1(u) - g_2(u)$. Then, if

$$\int_0^\infty e^{-tu^2}g_1(u)\,du = \int_0^\infty e^{-tu^2}g_2(u)\,du, \quad (11)$$

it follows that

$$g_1(u) \equiv g_2(u).$$

Given a definite integral to evaluate,

$$\int_0^{\frac{1}{2}\pi} G(\sin\theta, \cos\theta)\,d\theta, \quad (12)$$

define

$$g_1(u) = \int_0^{\frac{1}{2}\pi} G(u\sin\theta, u\cos\theta)u\,d\theta. \quad (13)$$

Change to Cartesian coordinates:

$$x = u\cos\theta, \quad y = u\sin\theta, \quad x^2 + y^2 = u^2,$$
$$u\,du\,d\theta = dx\,dy,$$

and the limits

$$0 \le \theta \le \tfrac{1}{2}\pi, \quad 0 \le u \le \infty$$

become

$$0 \le x \le \infty, \quad 0 \le y \le \infty.$$

Then, provided that $G(x, y)$ is separable,

$$\int_0^\infty e^{-tu^2}g_1(u)\,du = \int_0^\infty e^{-tx^2}G^x(x)\,dx\int_0^\infty e^{-ty^2}G^y(y)\,dy, \quad (14)$$

where

$$G(x, y) = G^x(x)G^y(y).$$

Upon integrating (14), if one can choose a $g_2(u)$ to satisfy (11), then (12) is determined by letting $u = 1$.

To apply this method to (9) it will be convenient for the cases we consider to take

$$F(r) = H(a^2 - r^2)^{\frac{1}{2}},$$

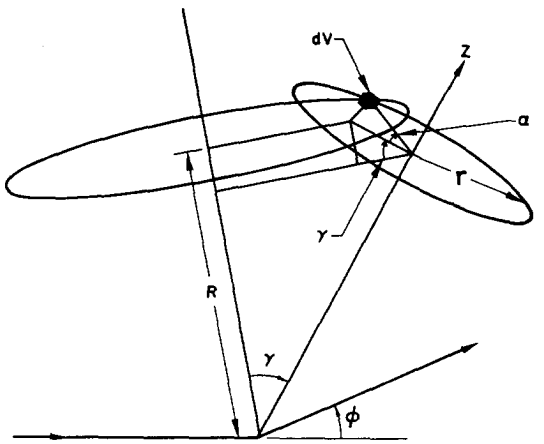


FIG. 1. Relationship between scattering coordinate system and cylindrical coordinate system of particle.

and letting $r = a\cos\theta$, we have

$$f(\gamma) = \frac{4\pi a^3}{c_1 V}\int_0^{\frac{1}{2}\pi} J_0(c_2\cos\theta)\sin\left[\frac{c_1}{a}H(a\sin\theta)\right] \times \sin\theta\cos\theta\,d\theta, \quad (15)$$

where

$$c_1 = ka\cos\gamma, \quad c_2 = ka\sin\gamma.$$

Then

$$g_1(u) = \frac{4\pi a^3}{c_1 v}\int_0^{\frac{1}{2}\pi} J_0(c_2u\cos\theta) \times \sin\left[\frac{c_1}{a}H(au\sin\theta)\right]u\sin\theta(u\cos\theta)u\,d\theta, \quad (16)$$

$$g_1(x, y) = \frac{4\pi a^3}{c_1 V}\int_0^\infty xJ_0(c_2x)\,dx \times \int_0^\infty \sin\left[\frac{c_1}{a}H(ay)\right]y\,dy. \quad (17)$$

So that form (14) becomes

$$\frac{4\pi a^3}{c_1 V}I_1I_2 \equiv \frac{4\pi a^3}{c_1 V}\int_0^\infty e^{-tx^2}J_0(c_2x)x\,dx \times \int_0^\infty e^{-ty^2}\sin\left[\frac{c_1}{a}H(ay)\right]y\,dy, \quad (18)$$

where I_1 and I_2 equal the first and second integrals, respectively. I_1 is in the form of Weber's first exponential integral[5] and is given as

$$I_1 = (1/2t)e^{-c_2^2/4t}. \quad (19)$$

In the next section, the solution will be carried to completion using specific forms of $F(r)$ and obtaining $f(\gamma) = g_2(1)$.

[5] Reference 4, p. 393.

III. STRUCTURE FACTOR DETERMINATIONS FOR VARIOUS SHAPES

In the table below, $F(r)$, $H(w)$, and V are listed for ellipsoids of revolution, cylinders, and a new shape—paraboloids of revolution (base to base).

	$F(r)$	$H(w)$	V
(1) Ellipsoids	$v(a^2 - r^2)^{\frac{1}{2}}$	vw	$4\pi va^3/3$
(2) Cylinders	h	h	$2\pi a^2 h$
(3) Paraboloids	$c_0(a^2 - r^2)$	$c_0 w^2$	$\pi c_0 a^4$.

Substituting in the integrals, I_2 and f now may be obtained for these three cases.

(1) Ellipsoids:

$$I_2 = \int_0^\infty e^{-tv^2} \sin(c_1 vy) y\, dy = \frac{\pi^{\frac{1}{2}} c_1 v}{4t^{\frac{3}{2}}} e^{-c_1^2 v^2/4t}. \quad (20)$$

Then from (11),

$$\int_0^\infty e^{-tu^2} g_2(u)\, du = \frac{3\pi^{\frac{1}{2}}}{8t^{5/2}} e^{-(c_1^2 v^2 + c_2^2)/4t}. \quad (21)$$

Noting that[6]

$$\int_0^\infty e^{-tu^2} J_n(au) u^{n+1}\, du = \frac{a^n}{(2t)^{n+1}} e^{-a^2/4t}, \quad (22)$$

and letting $n + 1 = \frac{5}{2}$ and $a = (c_1^2 v^2 + c_2^2)^{\frac{1}{2}}$, (22) will be equivalent to (21) if

$$g_2(u) = \left[\frac{3(2\pi)^{\frac{1}{2}} u^{5/2}}{2(c_1^2 v^2 + c_2^2)^{\frac{3}{4}}}\right] J_{\frac{3}{2}}[(c_1^2 v^2 + c_2^2)^{\frac{1}{2}} u]. \quad (23)$$

Thus, letting

$$\beta \equiv (c_1^2 v^2 + c_2^2)^{\frac{1}{2}} = ka(v^2 \cos^2\gamma + \sin^2\gamma)^{\frac{1}{2}},$$

$$f(\gamma) = g_2(1) = 3(\tfrac{1}{2}\pi)^{\frac{1}{2}} \frac{J_{\frac{3}{2}}(\beta)}{\beta^{\frac{3}{2}}} \equiv \Phi(\beta), \quad (24)$$

for spheres, $v = 1$ and $\beta = ka$.

(2) Cylinders:

$$I_2 = \sin\left(\frac{c_1 h}{a}\right) \int_0^\infty e^{-tv^2} y\, dy = \frac{1}{2t} \sin\left(\frac{c_1 h}{a}\right), \quad (25)$$

$$\int_0^\infty e^{-tu^2} g_2(u)\, du = \frac{2a \sin(c_1 h/a)}{c_1 h (2t)^2} e^{-c_2^2/4t}. \quad (26)$$

Using (22) again,

$$g_2(u) = \frac{2a \sin(c_1 h/a)}{c_1 c_2 h} J_1(c_2 u) u^2, \quad (27)$$

$$f(\gamma) = \frac{2 \sin(kh \cos\gamma)}{k^2 ah \sin\gamma \cos\gamma} J_1(ka \sin\gamma). \quad (28)$$

(3) Paraboloids:

$$I_2 = \int_0^\infty e^{-tv^2} \sin(c_0 c_1 a y^2) y\, dy = \frac{c_0 c_1 a}{2[t^2 + (c_0 c_1 a)^2]}. \quad (29)$$

Let $c_0 c_1 a = s$;

$$\int_0^\infty e^{-tu^2} g_2(u)\, du = \frac{1}{t(t^2 + s^2)} e^{-c_2^2/4t}. \quad (30)$$

Since the only requirement on t is that it be sufficiently large, choose $t^2 > s^2$. Then,

$$(t^2 + s^2)^{-1} = \frac{1}{t^2} \sum_{i=0}^\infty (-1)^i \left(\frac{s}{t}\right)^{2i}.$$

Thus, the right side of (30) can be written

$$\frac{1}{t^3} \sum_{i=0}^\infty (-1)^i \left(\frac{s}{t}\right)^{2i} e^{-c_2^2/4t}. \quad (31)$$

For the ith term,

$$\int_0^\infty e^{-tu^2} g_{2i}(u)\, du = (-1)^i \frac{s^{2i}}{t^{2i+3}} e^{-c_2^2/4t}; \quad (32)$$

and using (22),

$$g_2(u) = \sum_{i=0}^\infty g_{2i}(u)$$

$$= \frac{2}{s^2} \sum_{i=0}^\infty (-1)^i \left(\frac{2s}{c_2}\right)^{2i+2} J_{2i+2}(c_2 u) u^{2i+3}. \quad (33)$$

It can be shown, using the asymptotic expression for $J_n(x)$,[7] that this series converges absolutely.

Thus,

$$f(\gamma) = g_2(1) = \frac{2}{s^2} \sum_{i=0}^\infty (-1)^i \left(\frac{2s}{c_2}\right)^{2i+2} J_{2i+2}(c_2); \quad (34)$$

and using the definition of the second-order Lommel function,[8]

$$f(\gamma) = (2/s^2) U_2(2s, c_2), \quad (35)$$

where $s = c_0 ka^2 \cos\gamma \quad c_2 = ka \sin\gamma$.

This solution for the structure factor in terms of a tabulated function should make it possible, using numerical methods, to obtain the scattered intensity curve for base-to-base paraboloids.

The general approach used in the above solutions makes it relatively easy to inspect for integrability the expressions for various conceivable cylindrically symmetric shapes. Such an examination has revealed no other shapes for which the structure factor formula can be integrated.

[6] Reference 4, p. 394.

[7] Reference 4, p. 225.

[8] Reference 4, p. 537.

Note on the Global Validity of the Baker–Hausdorff and Magnus Theorems

JAMES WEI
Princeton University, Princeton, New Jersey and *Socony Mobil Oil Co., Paulsboro, New Jersey*
(Received 26 April 1963)

The Baker–Hausdorff theorem states that for two given elements x and y in an associative algebra, the equation $e^x e^y = e^z$ has a solution z which lies in the Lie algebra generated by x and y. The Magnus continuous analog gives an exponential solution to a linear operator differential equation. Both theorems are valid globally for free Lie algebras of formal power series. For algebras that are not free, however, both theorems are locally but not globally valid. Some examples are given. Necessary and sufficient conditions for global validity are discussed. A superior representation in terms of a finite product of exponentials is also given.

I. THE BAKER–HAUSDORFF AND MAGNUS THEOREMS

THE Baker–Hausdorff theorem and its continuous analog of Magnus have been applied to several problems of mathematical physics.[1,2,3] The question of the global validity of these theorems is important in the problem of the long-time behavior of quantum mechanical systems with a time-dependent Hamiltonian and of linear stochastic motions. The two theorems are usually stated as global theorems in free Lie algebras. Since all the applications are in Lie algebras that are not free, it is important to examine the domain of validity.

The Baker–Hausdorff theorem can be rephrased as four statements of increasing strength concerning an associative algebra R (which may be of linear operators) over an arbitrary field F of characteristic zero. Given two elements x and y in this algebra R,

(A) The equation $e^x e^y = e^z$ has a solution z in R.

(B) The solution z lies in the Lie algebra L generated by x and y over F. (The elements of L form a subset of R.)

(C) The solution z is a continuous analytic function of x and y.

(D) The solution z can be computed by the series

$$z = x + z_1 + z_2 + \cdots \tag{1}$$

where z_m is a Lie element of order m in y,

$$z_1 = y - \tfrac{1}{2}[y, x] - \sum_1^\infty (-1)^n B_n \{y, x^{2n}\}/(2n)!, \tag{2}$$

or

$$z_1 = \left\{ y \frac{x}{e^x - 1} \right\}.$$

In Eq. (2) we use the convention

$$[y, x] = y(adx) = yx - xy$$

$$\{y, x^n\} = y(adx)^n = [\cdots [[\overbrace{[y, x], x], x] \cdots x}^{n}].$$

Also in Eq. (2), B_n designates the nth Bernoulli number.

The other terms in Eq. (1) are given by the recursive formula

$$z_{m+1} = \frac{1}{m+1} \left(z_1 \frac{\partial}{\partial x} \right) z_m, \tag{3}$$

where $(z_1\ \partial/\partial x)$ is the "derivative" or "polar" operator of Hausdorff.

Statement (D) was obtained by Baker[4] and Hausdorff,[5] and an excellent review has been given by Magnus.[6] For a free Lie algebra (which is necessarily infinite-dimensional), each element $\{y, x^{2n}\}$ is linearly independent of all other elements and the problem of convergence of the series (D) does not enter. Even $\sum n! x^n$ is an element of good standing in the associative algebra R. Statement (D) is globally valid for a free Lie algebra over any field of characteristic zero, for instance, the field of rational numbers.

Statement (B) merely asserts that z is a Lie element; it is also a corollary of the elegant theorems

[1] E. H. Wichmann, J. Math. Phys. **2**, 876 (1961).
[2] F. J. Murray, J. Math. Phys. **3**, 451 (1962).
[3] G. H. Weiss and A. A. Maradudin, J. Math. Phys. **3**, 771 (1962).
[4] H. F. Baker, Proc. London Math. Soc. **35**, 333 (1903); **34**, 347 (1902).
[5] F. Hausdorff, Ber. Verhandl. Sächs. Akad. Wiss. Leipzig, Math.-naturo Kl. **58**, 19 (1906).
[6] W. Magnus, Commun. Pure Appl. Math. **7**, 649 (1954).

of Friedrichs, Cohn and Finkelstein[1,6,7,8,9] for free Lie algebras.

The continuous analog of Magnus concerns the differential equation[6]

$$dU/dt = AU, \tag{4}$$

where $A(t)$ is a linear operator depending on a real variable t, and $U(t)$ is another linear operator. The statements for the Magnus theorem are:

(A′) There exists an operator $\Omega(t)$ that satisfies the equation

$$U(t) = \exp \Omega(t).$$

(B′) The operator Ω is in the Lie algebra generated by $A(t)$.

(C′) The operator Ω is a continuous differentiable function of $A(t)$ and t.

(D′) The operator Ω satisfies the nonlinear differential equation

$$d\Omega/dt = A + \tfrac{1}{2}[A, \Omega] - \sum (-1)^n B_n \{A, \Omega^{2n}\}/(2n)!, \tag{5}$$

or

$$\frac{d\Omega}{dt} = \{A, \Omega/(1 - e^{-\Omega})\}.$$

Equations (2) and (5) are formally identical except for a change of sign in the second element.

II. EXAMPLES OF THE FAILURES OF THE THEOREMS

When an algebra is not free it is the homomorphic image of some free algebra. We now begin to encounter the problem of convergence of the series in Eqs. (2) and (5). For a nilpotent Lie algebra, the series become polynomials; the Baker–Hausdorff and Magnus theorems are globally valid. If the Lie algebra is not nilpotent, series (2) and (5) do not converge unless x and y are small in a suitable sense. The problem was investigated by Hausdorff[5] who gave a criterion for convergence. Take an example of a two-dimensional Lie algebra such that $[y, x] = ay$, where a is a scalar. The series of Eq. (1) reduces to $z = x + z_1$, and z_1 is given by

$$z_1 = y(1 - a/2 - \sum (-1)^n B_n a^{2n}/(2n)!) = ya(e^a - 1)^{-1}. \tag{6}$$

The series given by Eq. (6) is equal to $(a/2i)$ cot $(a/2i) - (a/2)$ and converges if a is less than 2π. In general, let n be the dimension of the Lie algebra and let M be the absolute value of the largest element in the adjoint representation of x; then the series converges if $M < 2\pi/n$.

Thus we see that statements (D) and (D′) are locally valid, but cannot be used to compute z and Ω in the large. However the other statements need not depend on the validity of (D) and (D′). It is well-known that for the full linear group and the unitary group over the complex numbers, any nonsingular matrix is the exponential of another matrix, so statements (A) and (A′) are never in doubt. If statements (B) and (C) are globally valid, we can still investigate many of the properties of z and Ω even though we cannot compute them with the aid of (D) and (D′).

Equation (6) shows that as long as $a \neq 2\pi ik$, where k is any integer number that is not zero, the element z_1 is well-defined and an analytic function of x and y. The numbers $2\pi ik$ play a central role on the global validity of the exponential mapping of the Lie algebra L into the Lie group G, which is a many-to-one mapping since $\exp(x) = \exp(2\pi ikI + x)$. The mapping is also singular about a point x of L if the representation of x has two eigenvalues that differ by $2\pi ik$. For instance, if

$$x = \begin{bmatrix} 0 & a \\ 0 & 2\pi ik \end{bmatrix}, \tag{7}$$

then

$$e^x = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix};$$

thus e^x is always the identity matrix regardless of the value of a.

Magnus has shown that statement (C′) is not globally valid. We shall show here that statement (C) is not globally valid for the same reason. Consider a Lie algebra with a subalgebra spanned by the following 2×2 matrices:

$$A = \begin{bmatrix} 0 & 0 \\ 0 & i \end{bmatrix}, \qquad B = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}. \tag{8}$$

Let $X = At$, $Y = At + B$; then

$$e^x e^y = \begin{bmatrix} 1 & 0 \\ \dfrac{\exp(2it) - \exp(it)}{it} & \exp(2it) \end{bmatrix}.$$

Let the solution z be given by

[7] R. Lyndon, Mich. Math. J. 3, 27 (1955).
[8] P. M. Cohn, Compt. Rend. 239, 743 (1954).
[9] H. Zassenhaus, Trans. Roy. Soc. Can. Sec. III 51, 55 (1957).

$$z = \begin{bmatrix} 0 & 0 \\ a & b \end{bmatrix}.$$

Then

$$e^z = \begin{bmatrix} 1 & 0 \\ (e^b - 1)a/b & e^b \end{bmatrix}.$$

Thus we have $a = 2(e^{2it} - e^{it})/(e^{2it} - 1)$, and $b = 2it$. Statement (C) is not valid since both $a(t)$ and $b(t)$ have discontinuities at $t = \pi$; in fact, as $t \to \pi$, $a \to \infty$ and $b \to 2\pi i$, but $a(\pi) = 2/\pi i$ and $b(\pi) = 0$. Now every nonnilpotent Lie algebra over the complex field has a subalgebra $[A, B] = iB$; thus the adjoint representation of $2\pi A$ will have two eigenvalues that differ by $2\pi i$. For such algebras, statement (C) is not globally valid.

It is more surprising that statements (B) and (B′) are also not valid in the large. We may use the example in equation (6) to show that (B) is not valid for an arbitrary field of characteristic zero. We may commence with the field of rational numbers, but z lies in the Lie algebra over the field of real numbers. The next example will show that, for some systems, one must extend to the complex numbers.

The set of all 2×2 *nonsingular* real matrices form a Lie group which has the property that not every element of the group can be written as $\exp(z)$ for some z in the Lie algebra.[10,11] Let

$$x = 5\pi/4\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \qquad y = 2\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \tag{9}$$

$$u = e^x e^y = \begin{bmatrix} -1/\sqrt{2} & 0 \\ -1/\sqrt{2} & -\sqrt{2} \end{bmatrix}.$$

This matrix is not the exponential of any real matrix, since it has no square root in real 2×2 matrices. If a matrix u is of the form e^z, then surely $e^{\frac{1}{2}z}$ is the square root of u. The square of any 2×2 real matrix v is

$$v^2 = \begin{bmatrix} a & b \\ c & d \end{bmatrix}^2 = \begin{bmatrix} a^2 + bc & b(a + d) \\ c(a + d) & bc + d^2 \end{bmatrix}.$$

If $v^2 = e^x e^y$, then $b(a + d) = 0$, which means that either $b = 0$ or $(a + d) = 0$. The former is not acceptable since it makes $a^2 = -1/\sqrt{2}$ and $(a + d) = 0$ is not acceptable because it makes $-1/\sqrt{2} = 0$.

The matrix u also lies on the solution space to the differential equation (4) if $A(t)$ is given by

$$A(t) = 2(1 - t)\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

when $0 \leq t \leq 1$, and by

$$A(t) = (t - 1)\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

when $t > 1$. Of course the element z exists and is in the Lie algebra over the field of complex numbers. But z can never be computed by the series (D) and (D′) since these series contain only real rational coefficients.

One may easily form the conjecture that statements (B) and (B′) are nevertheless valid for algebras over the complex numbers because they are valid for the full linear group and for the unitary group. This is not true, however, as the following example shows. Let

$$x = \begin{bmatrix} 0 & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{10}$$

and

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

This generates a two-dimensional Lie algebra with the commutation relation $[x, y] = iy$. In general,

$$e^{ax+by} = \begin{bmatrix} 1 & 0 & 0 \\ \dfrac{e^{ai} - 1}{ai} b & e^{ai} & 0 \\ 0 & 0 & e^a \end{bmatrix}.$$

Now the element $\exp(\pi x) \exp(\pi x + by)$ is given by

$$\exp(\pi x) \exp(\pi x + by) = \begin{bmatrix} 1 & 0 & 0 \\ -2i/\pi & 1 & 0 \\ 0 & 0 & e^{2\pi} \end{bmatrix},$$

and does not lie in the set of elements $\exp(ax + by)$. In this case, the Baker–Hausdorff and Magnus theorems completely break down, and no extension of field can force z to be a Lie element.

[10] S. Helgason, *Differential Geometry and Symmetric Spaces* (Academic Press Inc., New York, 1962) p. 227.

[11] D. Montgomery and L. Zippin, *Topological Transformation Groups* (Interscience Publishers, Inc., New York, 1955), p. 189.

III. NECESSARY AND SUFFICIENT CONDITIONS FOR THE GLOBAL VALIDITY OF THE THEOREMS

The global validity of the Baker–Hausdorff and Magnus theorems depend on the singularities of the exponential mapping of the Lie algebra L into the Lie group G. The validity of (C) and (C′) can be examined by rewriting Eqs. (2) and (5) as

$$z_1 = y(adx)(e^{adx} - 1)^{-1}, \tag{11}$$

where the analytic function of an operator (adx) is defined by a Cauchy integral type

$$F(adx) = \frac{1}{2\pi i}\int \frac{F(\gamma)}{\gamma I - adx}\, d\gamma.$$

It is easy to see that as long as (adx) does not have any eigenvalue equal to $2\pi ik$, $k \neq 0$, z_1 is an analytic function of x and y. This condition is necessary and sufficient for the validity of (C′) and is necessary for (C).

Dixmier[12] and Nono[13] investigated the conditions under which the exponential mapping will be locally nonsingular and onto. Saito[14] defined a Lie algebra to be of type (E) if for every element u of the Lie group G, there exists an element z of the Lie algebra L so that $u = e^z$ is satisfied. For every Lie algebra over the complex numbers, the exponential mapping is locally nonsingular if and only if the algebra is nilpotent. The condition is necessary, since for any complex nonnilpotent Lie algebra, there always exists elements x and y so that $[x, y] = 2\pi iy$. For a solvable algebra over the real numbers, the necessary and sufficient condition is that L contains no subalgebra isomorphic to the so-called D algebra[14] spanned by elements x, y, and z with the commutation rules $[x, y] = z$; $[x, z] = -y$.

IV. A SUPERIOR PARAMETRIZATION OF THE LIE GROUP

One sees that it is inadequate to parametrize a Lie group by a simple exponential mapping from the Lie algebra, even for the case of a solvable Lie algebra. A superior parametrization in terms of a product of exponentials has been given by Wei and Norman[15]; it is a generalization of a theorem of Chevalley. Consider a complex solvable algebra with the composition series

$$L = L_1 \supset L_2 \supset L_3 \cdots L_n \supset 0,$$

where each L_i is an ideal spanned by the basis elements x_i, x_{i+1}, $\cdots$ x_n. Let G_m be the set of all elements of the form

$$\exp(a_m x_m)\exp(a_{m+1}x_{m+1})\cdots\exp(a_n x_n),$$

and G_m' is the set of elements of the form

$$\exp(a_n x_n)\exp(a_{n-1}x_{n-1})\cdots\exp(a_m x_m).$$

We shall show that $G_m = G_m'$ and is a Lie group. Furthermore, we have the composition series

$$G = G_1 \supset G_2 \supset \cdots G_n \supset 1,$$

where each G_m is a normal subgroup of G_{m-1}, $\cdots$ G_2, G_1. It was shown by Wei and Norman that $\exp(L_m) \subset G_m$, or each element of the form

$$u = \exp(a_m x_m + a_{m+1}x_{m+1} + \cdots a_n x_n),$$

can be written in the form

$$u = \exp(b_m x_m)\exp(b_{m+1}x_{m+1})\cdots\exp(b_n x_n).$$

The parameters b are continuous functions of the parameters a.

First we prove the lemma that if $u \in L_m$, then $\exp(u)G_m = G_m\exp(u)$. This is due to the fact that $e^u e^v e^{-u} = \exp(e^u v e^{-u}) = \exp(e^{adu}v) = \exp(v')$. If v is in an ideal L_m, so is v'. Thus we have $\exp(L_m)\exp(u) = \exp(u)\exp(L_m)$. Since an element in G_m is a product of elements of the form $\exp(L_m)$, the lemma is proven.

Now we proceed with an inductive proof that $G_m = G_m'$ and is a Lie group. First $G_n = G_n'$ and is a Lie group. Next we show that if $G_m' = G_m$ and is a Lie group, then $G_{m-1}' = G_{m-1}$ and is a Lie group. G_{m-1} is closed under multiplication since

$$\exp(a_{m-1}x_{m-1})G_m\exp(b_{m-1}x_{m-1})G_m$$
$$= \exp(a_{m-1}x_{m-1})\exp(b_{m-1}x_{m-1})G_m G_m \subset G_{m-1}.$$

Because a typical element in G_{m-1} is

$$\exp(a_{m-1}x_{m-1})G_m = G_m\exp(a_{m-1}x_{m-1})$$
$$= G_m'\exp(a_{m-1}x_{m-1}) \subset G_{m-1}',$$

we have then that $G_{m-1} = G_{m-1}'$. Since G_m possesses the inverse of every element, and is closed under multiplication, G_m is a group.

For a Lie group that is not solvable, this representation by a product of one-parameter subgroups is not global. If we are willing to settle for the more modest goal of global approximation, then this representation is adequate for a large family of groups. For a suitable choice of basis and ordering,

---

[12] J. Dixmier, Bull. Soc. Math. France **85**, 113 (1957).
[13] T. Nono, J. Sci. Hiroshima Univ. Ser. A **20**, 115 (1956/57) **21**, 163 (1957/58).
[14] Masahiko Saito, Sci. Papers Coll. Gen. Educ., Univ. Tokyo **7**, 1 (1957).
[15] J. Wei and E. Norman, Proc. Am. Math. Soc. (to be published).

the product of one-parameter subgroups is open and everywhere dense in a complex connected semisimple Lie group; thus any element of the group can be approximated as close as we please by a product of one-parameter subgroups.

Consider first the full linear group $GL(n, C)$ which contains the subgroups: N of all lower triangular matrices with diagonal elements all equal to unity, D of all diagonal matrices, and P of all upper triangular matrices with diagonal elements all equal to unity. A theorem of Godement[16] shows that any element $g \in G$ with the property that the sequence of principal minors all have nonvanishing determinant,

$$g_{11} \neq 0, \quad \begin{vmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{vmatrix} \neq 0, \quad \begin{vmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{33} \\ g_{31} & g_{32} & g_{33} \end{vmatrix} \neq 0, \text{ etc.},$$

can be uniquely represented as $g = ndp$, where $n \in N$, $d \in D$, and $p \in P$. Since N, D, and P are all nilpotent connected Lie groups, they are all products of one-parameter subgroups; N is generated by the nilpotent Lie algebra of lower niltriangular matrices (negative roots), D by the Abelian Lie algebra of diagonal matrices (Cartan subalgebra), and P by the nilpotent Lie algebra of upper niltriangular matrices (positive roots). Since the complement of NDP is a closed nowhere dense set, the closure of NDP is G, we have established that $GL(n, C)$ can be globally approximated by a product of exponentials.

For the solution to Eq. (4), a theorem of Levi enables one to decompose the solution to the product of a solvable and a semisimple subgroup.[15] The "Gaussian decomposition" as given by Zelobenko[17] showed that any semisimple connected complex Lie group is equal to the closure of the product NDP. It appears that if we restrict ourselves to the field of complex numbers, the product of one-parameter subgroups with a suitable choice of basis and ordering is a global approximation to the solution space of Eq. (4).

ACKNOWLEDGMENT

The author is indebted to Professor Armand Borel for his stimulating discussions and suggestions.

[16] R. Godement, Trans. Am. Math. Soc. **73**, 496 (1952).

[17] Transl. Soviet Math. Dokl. Akad. Nauk. USSR **2**, 1076 (1961).

Errata: Nonrelativistic S-Matrix Poles for Complex Angular Momenta

[J. Math. Phys. 3, 867 (1962)]

ROGER G. NEWTON
Department of Physics, Indiana, University, Bloomington, Indiana

---

pp. 867 and 868: Equations (1.2), (1.3), and (1.5): Multiply right-hand sides by π; also change minus sign in front of sum in (1.3) to plus sign, and insert minus sign in right-hand side of equation below (1.5).

p. 869: Insert a minus sign in front of i in the exponential.

p. 870: Eq. (3.2): Eliminate the i; also replace $>$ by $<$ in the inequality above (3.3).

p. 871: Third line from bottom, right-hand column: Replace k_0 by k_0^2.

p. 872: lines 11 and 14, left-hand column: Replace $k^{\frac{1}{2}+\lambda}$ by $k^{-\frac{1}{2}+\lambda}$; line above Eq. (4.7): Replace k_0^2 by k_0.

p. 874: line 12, left-hand column, should read: "real part of λ fixed but allowing k to vary while . . ."

p. 877, line 10, left-hand column: For negative integral λ it is not true that v_λ can be expanded in a power series in r and that it is an entire function of r. It contains a logarithmic term. However, since the corresponding logarithmic term in h is multiplied by sufficient powers of r, the proof goes through nevertheless.

p. 879: The last sentence of Sec. 7 is incorrect and should be eliminated.