JOURNAL OF
MATHEMATICAL
PHYSICS

VOLUME 6, NUMBER 4 APRIL 1965

Microcausality in Asymptotic Quantum Field Theory

F. ROHRLICH AND J. C. STODDART
Department of Physics, Syracuse University, Syracuse, New York
(Received 26 May 1964)

Asymptotic quantum field theory can be formulated without the axiom of microcausality, (strong local commutativity). It is shown here that in certain cases microcausality is nevertheless a consequence of the remaining assumptions of the theory. These cases are the ones for which solutions of the basic equations can be obtained in perturbation expansion, i.e., the scalar, spin ½, and neutral vector field. For the charged vector boson, for example, the microscopic causality condition cannot be proved from the basic postulates. In the provable cases the functional derivative in the Bogoliubov causality condition can be defined explicitly, and this condition can then be derived from the postulates of asymptotic quantum field theory. In the other cases, Bogoliubov causality cannot be defined explicitly.

I. INTRODUCTION

IN the recent formulation of asymptotic quantum field theory[1] the microscopic causality condition,

$$[A(x), A(y)]_- = 0 \quad \text{for} \quad (x-y)^2 > 0,$$

is not explicity assumed, even though this condition is sufficient to assure the Lorentz invariance of the time-ordered product. [Since the theory involves time-ordered products in vacuum expectation values only, strong local commutativity is not required for their Lorentz invariances. For example, $\langle 0|\ [A(x), A(y)]\ |0\rangle = 0$ for $(x-y)^2 > 0$ on the basis of Lorentz invariance of $A(x)$ only, so that $\langle 0|\ T(A(x)A(y))\ |0\rangle$ is Lorentz-invariant without any assumption on commutativity.] Instead, a fundamental integral equation is proposed which determines all S-matrix elements. Only those solutions of this equation which are Lorentz-invariant are admitted as physically meaningful solutions.

The basic assumptions of asymptotic field theory (AQFT) are therefore only the validity of the usual quantum field theory of the free fields for the in and out field (i.e., the usual quantum field theory is valid asymptotically), the unitarity and Lorentz invariance of the S operator, and an interpolating field characterized by a weak asymptotic limit. This interpolating field can be completely eliminated, resulting in the integral equation for the S matrix and a particular way of going off the mass shell. The dynamics of AQFT is thereby characterized.

In order to connect this approach with the older formulations of quantum field theory, we investigate the question to what extent the commutation relations of the interpolating field can be recovered. We find that strong microcausality can indeed be recovered, but characteristically only in those cases which are amenable to a perturbation solution, e.g., the charged and neutral scalar field, the Dirac field and the Maxwell field. The charged vector field does not lead to microcausality.

The Bogoliubov causality condition thereby plays an important role. We derive it from the assumptions of AQFT for the microcausal cases (Secs. II–IV). However, this causality condition is no longer defined in the other cases (charged vector bosons, for example, Sec. V). This fact can be regarded as responsible for the breakdown of the microcausality proof.

[1] R. Pugh, Ann. Phys. (N. Y.) 23, 335 (1963).

II. SCALAR FIELDS

We consider first the real scalar field satisfying

$$[A_{in}(x), A_{in}(y)]_- = -i\Delta(x-y), \quad K(x)A_{in}(x) = 0,$$
$$A_{in}(x) = A_{in}^{\dagger}(x). \tag{1}$$

The interacting field operator connecting the in and out operator is

$$A(x) = S^{\dagger}(A_{in}(x)S)_+ = A_{in}(x) + S^{\dagger}[A_{in}(x), S]_R. \tag{2}$$

The subscripts $+$ and R denote the positive time ordered and the retarded product. Expressed in terms of functional derivatives[2] this field becomes

$$A(x) = A_{in}(x) - i\int \Delta_R(x-y)S^{\dagger}\frac{\delta S}{\delta A_{in}(y)}\,d^4y. \tag{3}$$

We consider the commutator $[A(x), A(y)]_-$,

$$\begin{aligned}[A(x), A(y)]_- &= -i\Delta(x-y)\\ &- \int \Delta_R(y-z_2)\Delta(x-z_1)\\ &\times \frac{\delta}{\delta A_{in}(z_1)}\left[S^{\dagger}\frac{\delta S}{\delta A_{in}(z_2)}\right]d^4z_1\,d^4z_2\\ &+ \int \Delta_R(x-z_1)\Delta(y-z_2)\\ &\times \frac{\delta}{\delta A_{in}(z_2)}\left[S^{\dagger}\frac{\delta S}{\delta A_{in}(z_1)}\right]d^4z_1\,d^4z_2\\ &+ \int \Delta_R(x-z_1)\Delta_R(y-z_2)\left[\frac{\delta S^{\dagger}}{\delta A_{in}(z_1)}\frac{\delta S}{\delta A_{in}(z_2)}\right.\\ &\left.- \frac{\delta S^{\dagger}}{\delta A_{in}(z_2)}\frac{\delta S}{\delta A_{in}(z_1)}\right]d^4z_1\,d^4z_2. \end{aligned}\tag{4}$$

Using $\theta(x-z_1)+\theta(z_1-x)=\theta(y-z_2)+\theta(z_2-y)=1$ in the first and second integrals of Eq. (6) we obtain, after cancellations,

$$\begin{aligned}[A(x), A(y)] &= -i\Delta(x-y)\\ &- \int \Delta(x-z_1)\Delta(y-z_2)\theta(y-z_2)\theta(z_1-x)\\ &\times \frac{\delta}{\delta A_{in}(z_1)}\left(S^{\dagger}\frac{\delta S}{\delta A_{in}(z_2)}\right)d^4z_1\,d^4z_2\\ &+ \int \Delta(x-z_1)\Delta(y-z_2)\theta(x-z_1)\theta(z_2-y)\\ &\times \frac{\delta}{\delta A_{in}(z_2)}\left(S^{\dagger}\frac{\delta S}{\delta A_{in}(z_1)}\right)d^4z_1\,d^4z_2. \end{aligned}\tag{5}$$

We have used[2]

$$[\delta/\delta A_{in}(z_1), \delta/\delta A_{in}(z_2)]_- = 0 \tag{6}$$

to obtain Eq. (5). The right side of (5) vanishes on the spacelike plane $x^0 = y^0$, $\mathbf{x} \neq \mathbf{y}$, provided the following equation holds:

$$\delta/\delta A_{in}(z_1)(S^{\dagger}\delta S/\delta A_{in}(z_2)) = 0$$
$$\text{for } z_1^0 > z_2^0 \text{ and for } (z_1-z_2)^2 = 0. \tag{7}$$

That this is indeed the case can be seen as follows. Equation (7) causes the first integrand to vanish for all values of z_1 which are not in the past light cone emanating from z_2; but the Δ and θ functions of that integrand will vanish unless $z_1^0 \geq x^0 = y^0 \geq z_2^0$. Since $\mathbf{x} \neq \mathbf{y}$, this implies that the first integral in (5) always vanishes. Similarly, the second integral vanishes because the first factors require $z_1^0 \leq x^0 = y^0 \leq z_2^0$ while (7) makes the last factor zero under these conditions.

Equation (7) is just the Bogoliubov "causality condition" which we now proceed to prove.

We can easily show that the condition (7) is a consequence of the Lehmann, Symanzik, Zimmerman (LSZ) formalism. Any interacting field obeying the asymptotic conditions can be expanded in the form[3]

$$\begin{aligned}A(x) &= A_{in}(x) + \sum_n \frac{1}{n!}\int (d^4x)\{K(x_1)\cdots K(x_n)\\ &\times \langle 0|\, R(xx_1\cdots x_n)\,|0\rangle\} : A_{in}(x_1)\cdots A_{in}(x_n):. \end{aligned}\tag{8}$$

The $R \neq$ product of $(n+1)$ scalar field operators is defined by

$$\begin{aligned}R(A(x)A(x_1)\cdots A(x_n)) &\equiv R(xx_1\cdots x_n)\\ &= (-i)^n \sum_P \theta(x-x_1')\theta(x_1'-x_2')\cdots\theta(x_{n-1}'-x_n')\\ &\times [\cdots [A(x), A(x_1')]_-, \cdots, A(x_n')]_- \end{aligned}\tag{9}$$

where P indicates a summation over all permutations $x_1'\cdots x_n'$ of $x_1\cdots x_n$. Using Eq. (3) we can write condition (7) in the form

$$\delta j_s(z_2)/\delta A_{in}(z_1) = 0$$
$$\text{for } z_1^0 > z_2^0 \text{ and for } (z_1-z_2)^2 > 0 \tag{10}$$

where $j_s(z_2) = (\Box - m^2)A(z_2) \equiv K(z_2)A(z_2)$. We obtain from Eq. (8)

$$\begin{aligned}\frac{\delta j_s(z_2)}{\delta A_{in}(z_1)} &= \sum_n \frac{1}{n!}\int (d^4x)\{K(z_1)K(z_2)K(x_1)\cdots K(x_n)\\ &\times \langle 0|\, R(A(z_2)A(z_1)A(x_1)\cdots A(x_n)\,|0\rangle\}\\ &\times : A_{in}(x_1)\cdots A_{in}(x_n):. \end{aligned}\tag{11}$$

[2] F. Rohrlich, J. Math. Phys. 5, 324 (1964).

[3] H. Lehmann, K. Symanzik, and W. Zimmermann, Nuovo Cimento 6, 319 (1957).

Only terms with both z_1^0 and $z_2^0 > x_i^0$ are kept in the R-product in Eq. (11). From the properties of the R-product we see that the right-hand side of Eq. (11) is indeed zero for $z_1^0 > z_2^0$. The Lorentz invariance of the S operator and of its derivatives with respect to the free fields implies that the right side vanishes for *all* $(z_1 - z_2)^2 > 0$ if it vanishes for some. Thus Eq. (10) follows and the Bogslinbov condition (7) is proved.

These results are generalized to the complex scalar field satisfying

$$\begin{aligned}[A_{\text{in}}(x),\, A_{\text{in}}(y)]_- &= 0,\\ [A_{\text{in}}(x),\, A_{\text{in}}^\dagger(y)]_- &= -i\Delta(x-y),\\ (\Box - m^2)A_{\text{in}}(x) &= 0.\end{aligned} \tag{12}$$

The interacting field operator is given by

$$A(x) = A_{\text{in}}(x) - i\int \Delta_R(x-z)S^\dagger \frac{\delta S}{\delta A_{\text{in}}^\dagger(z)}\, d^4z. \tag{13}$$

In analogy to the real scalar case, the interacting field commutator is

$$\begin{aligned}&[A(x),\, A^\dagger(y)]_-\\ &= -\int \Delta(y-z_2)\Delta(x-z_1)\theta(y-z_2)\theta(z_1-x)\\ &\quad\times \frac{\delta}{\delta A_{\text{in}}^\dagger(z_1)}\left[S^\dagger \frac{\delta S}{\delta A_{\text{in}}(z_2)}\right] d^4z_1\, d^4z_2\\ &\quad+\int \Delta(x-z_1)\Delta(y-z_2)\theta(x-z_1)\theta(z_2-y)\\ &\quad\times \frac{\delta}{\delta A_{\text{in}}(z_2)}\left[S^\dagger \frac{\delta S}{\delta A_{\text{in}}^\dagger(z_1)}\right] d^4z_1\, d^4z_2.\end{aligned} \tag{14}$$

This expression vanishes on the spacelike plane $(x-y)^2 > 0$ with $x^0 = y^0$ if

$$\frac{\delta}{\delta A_{\text{in}}^\dagger(z_1)}\left(S^\dagger \frac{\delta S}{\delta A_{\text{in}}(z_2)}\right) = \frac{\delta j_{cs}^\dagger(z_2)}{\delta A_{\text{in}}^\dagger(z_1)} = 0$$

$$\text{for}\quad z_1^0 > z_2^0 \quad\text{and for}\quad (z_1 - z_2)^2 > 0, \tag{15}$$

where $j_{cs}(z_2) = K(z_2)A(z_2)$. The previous argument concerning this applies as follows.

The LSZ expansion for the interacting field gives the following current operator:

$$\begin{aligned}j_{cs}(z_2) &= \sum_n \frac{1}{(n!)^2}\int (d^4x)\{K(z_2)K(x_1)\cdots K(x_{2n})\\ &\times \langle 0|R(A(z_2)A^\dagger(x_1)\cdots A^\dagger(x_n)A(x_{n+1})\cdots A(x_{2n}))\,|0\rangle\}\\ &\times :A_{\text{in}}(x_1)\cdots A_{\text{in}}(x_n)A_{\text{in}}^\dagger(x_{n+1})\cdots A_{\text{in}}^\dagger(x_{2n}):.\end{aligned} \tag{16}$$

The R-product is defined in a way similar to Eq. (9), but including both $A^\dagger$ and A fields. We can apply here the same considerations as in the real scalar case in order to establish (15).

This completes the proof that for the scalar case the usual form of microscopic causality follows from the basic assumptions of asymptotic quantum field theory as given in Ref. 1 and need not be given as a further axiom.

III. SPINOR FIELDS

We consider the case of the Dirac spin-$\frac{1}{2}$ field. The free-field relations are,

$$\begin{aligned}[\psi_{\text{in}}^\alpha(x),\, \psi_{\text{in}}^\beta(y)]_+ &= [\bar\psi_{\text{in}}^\alpha(x),\, \bar\psi_{\text{in}}^\beta(y)]_+ = 0\\ [\psi_{\text{in}}^\alpha(x),\, \bar\psi_{\text{in}}^\beta(y)]_+ &= iS_{\alpha\beta}(x-y)\\ D\psi_{\text{in}}^\alpha &= (\gamma^\mu\partial_\mu + m)\psi_{\text{in}}(x) = 0,\\ \psi_{\text{in}}(x)\mathcal{D} &= \bar\psi_{\text{in}}(x)(-\gamma^\mu\partial_\mu + m) = 0\\ S_{\alpha\beta}(x) &= (\gamma^\mu\partial_\mu - m)_{\alpha\beta}\Delta(x).\end{aligned} \tag{17}$$

In functional derivative form the interacting fields are

$$\begin{aligned}\psi^\alpha(x) &= \psi_{\text{in}}^\alpha(x) + i\int S_{\alpha\lambda}^R(x-z)S^\dagger \frac{\delta S}{\delta\bar\psi_{\text{in}}^\lambda(z)}\, d^4z,\\ \bar\psi^\alpha(x) &= \bar\psi_{\text{in}}^\alpha(x) + i\int S^\dagger \frac{\delta S}{\delta\psi_{\text{in}}^\lambda(z)}\, S_{\lambda\alpha}^A(z-x)\, d^4z.\end{aligned} \tag{18}$$

The fact that we are now dealing with free-field *anti*commutators leads to complications in the definitions of functional derivatives of functions of spinor operators.[2] In particular the rules depend on whether the functional transforms like the product of an odd or even number of spinors. These rules are given in Ref. 2 and we note that S transforms like an even number of spinor operators.

The interacting-field anticommutator is

$$\begin{aligned}&[\psi^\alpha(x),\, \bar\psi^\beta(y)]_+ = iS_{\alpha\beta}(x-y)\\ &-\int S_{\alpha\lambda}(x-z_1)S_{\pi\beta}^A(z_2-y)\Big[S^\dagger \frac{\delta^2 S}{\delta\bar\psi_{\text{in}}^\lambda(z_1)\delta\psi_{\text{in}}^\pi(z_2)}\\ &+\frac{\delta S^\dagger}{\delta\bar\psi_{\text{in}}^\lambda(z_1)}\cdot\frac{\delta S}{\delta\psi_{\text{in}}^\pi(z_2)}\Big]\, d^4z_1\, d^4z_2\\ &-\int S_{\alpha\lambda}^R(x-z_1)S_{\pi\beta}(z_2-y)\Big[S^\dagger \frac{\delta^2 S}{\delta\psi_{\text{in}}^\pi(z_2)\delta\bar\psi_{\text{in}}^\lambda(z_1)}\\ &-\frac{\delta S^\dagger}{\delta\psi_{\text{in}}^\pi(z_2)}\cdot\frac{\delta S}{\delta\bar\psi_{\text{in}}^\lambda(z_1)}\Big]\, d^4z_1\, d^4z_2\\ &+\int S_{\alpha\lambda}^R(x-z_1)S_{\pi\beta}^A(z_2-y)\Big[\frac{\delta S^\dagger}{\delta\bar\psi_{\text{in}}^\lambda(z_1)}\cdot\frac{\delta S}{\delta\psi_{\text{in}}^\pi(z_2)}\\ &+\frac{\delta S^\dagger}{\delta\psi_{\text{in}}^\pi(z_2)}\cdot\frac{\delta S}{\delta\bar\psi_{\text{in}}^\lambda(z_1)}\Big]\, d^4z_1\, d^4z_2.\end{aligned} \tag{19}$$

Using the θ-function indentities as in the scalar case, we obtain

$$[\psi^\alpha(x), \bar{\psi}^\beta(y)]_+ = iS_{\alpha\beta}(x-y)$$
$$+\int S_{\alpha\lambda}(x-z_1)S_{\pi\beta}(z_2-y)\theta(y-z_2)\theta(z_1-x)$$
$$\times\frac{\delta}{\delta\bar{\psi}^\lambda_{\text{in}}(z_1)}\left(S^\dagger\frac{\delta S}{\delta\psi^\pi_{\text{in}}(z_2)}\right)d^4z_1\,d^4z_2$$
$$-\int S_{\pi\lambda}(x-z_1)S_{\pi\beta}(z_2-y)\theta(x-z_1)\theta(z_2-y)$$
$$\times\frac{\delta}{\psi^\pi_{\text{in}}(z_2)}\left(S^\dagger\frac{\delta S}{\bar{\psi}^\lambda_{\text{in}}(z_1)}\right)d^4z_1\,d^4z_2. \tag{20}$$

We have used

$$[\delta/\delta\psi^\pi_{\text{in}}(z_2),\ \delta/\delta\bar{\psi}^\lambda_{\text{in}}(z_1)]_- S = 0 \tag{21}$$

and $S^\dagger S = 1$. Again, the conditions,

$$\frac{\delta\bar{j}_D(z_2)}{\delta\bar{\psi}^\lambda_{\text{in}}(z_1)} = \frac{\delta}{\delta\bar{\psi}^\lambda_{\text{in}}(z_1)}\left(iS^\dagger\frac{\delta S}{\delta\psi^\pi_{\text{in}}(z_1)}\right) = 0,$$
$$\text{for}\quad z_1^0 > z_2^0 \quad\text{and for}\quad (z_1-z_2)^2 > 0; \tag{22}$$
$$\frac{\delta j^\lambda_D(z_1)}{\delta\psi^\pi_{\text{in}}(z_2)} = \frac{\delta}{\delta\psi^\pi_{\text{in}}(z_2)}\left(S^\dagger\frac{\delta S}{\delta\bar{\psi}^\lambda_{\text{in}}(z_1)}\right) = 0,$$
$$\text{for } z_2^0 > z_1^0, \text{ and for } (z_1-z_2)^2 > 0$$

are sufficient to prove that the anticommutator (20) vanishes on the spacelike surface $(x-y)^2 > 0$ with $x^0 = y^0$. The anticommutator then vanishes for all points x and y with spacelike separation.

Expression (22) vanishes on the chosen spacelike surface as a consequence of the basic postulates, just as in the scalar case. By the LSZ method we obtain the following expression for the functional derivative of the current operator:

$$\frac{\delta j_D(z_1)}{\delta\psi_{\text{in}}(z_2)}$$
$$= \sum_n\frac{1}{(n!)^2}\int(d^4x)(d^4y)\{D(z_1)D(z_2)D(x_2)\cdots D(x_n)$$
$$\times\langle 0|\,R(\psi(z_1)\psi(z_2)\psi(x_2)\cdots\psi(x_n)\bar{\psi}(y_1)\cdots\bar{\psi}(y_n))\,|0\rangle$$
$$\times\mathcal{D}(y_1)\cdots\mathcal{D}(y_n)\}$$
$$\times\,:\bar{\psi}_{\text{in}}(x_2)\cdots\bar{\psi}_{\text{in}}(x_n)\psi_{\text{in}}(y_1)\cdots\psi_{\text{in}}(y_n):. \tag{23}$$

$D(x)$ and $\mathcal{D}(x)$ are defined in (17).

The R-product of $(n+1)$ spinor field operators $\psi(x)$ or $\bar{\psi}(x)$ is defined as

$$R(\psi(x)\psi(x_1)\cdots\psi(x_n))$$
$$= (i)^{n_1}(-i)^{n_2}\sum_p\theta(x-x_1')\theta(x_1'-x_2')\cdots\theta(x_{n-1}'-x_n')$$
$$\times[[\cdots[\psi(x),\psi(x_1')]_-,\psi(x_2')]_+,\psi(x_3')]_-\cdots,\psi(x_n')]_+. \tag{24}$$

The number of ψ and $\bar{\psi}$ fields in the product is denoted by n_1 and n_2, respectively. The R function properties imply that $\delta j_D(z_1)/\delta\psi_{\text{in}}(z_2)$ is zero for $z_1^0 > z_2^0$. As the LSZ expansions for the field and current operators depend only on the free-field relations, the Lorentz invariance of the scattering operator and the asymptotic properties, we have proven that together with the interacting fields (18), these imply the usual microscopic causality condition.

IV. THE ELECTROMAGNETIC FIELD

The special problems associated with the auxiliary condition on the electromagnetic field have been discussed in Ref. 1. We have

$$[A^{\text{in}}_\mu(x), A^{\text{in}}_\nu(y)]_- = -ig_{\mu\nu}D(x-y), \tag{25}$$

and the interacting field is written in the form[1]

$$A_\mu(x) = A^{\text{in}}_\mu(x)$$
$$-i\int g_{\mu\lambda}D_R(x-z)S^\dagger\frac{\delta S}{\delta A^{\text{in}}_\lambda(z)}d^4z + M\,\partial_\mu\partial_\lambda A^\lambda_{\text{in}}(x).$$

We adopt here the treatment in Ref. 1 of the supplementary condition.

To prove microcausality we consider all matrix elements,

$$\langle\phi^{(\alpha)}_{\text{in}}|\,[A_\mu(x), A_\nu(y)]_-\,|\phi^{(\beta)}_{\text{in}}\rangle. \tag{26}$$

With state vectors $\phi^{(\alpha)}_{\text{in}}$ constructed by the technique given in Ref. 1, it follows that all parts of (26) containing an M term vanish because

$$\langle\phi^{(\alpha)}_{\text{in}}|\,\partial_\mu A^\mu_{\text{in}}(x)\,|\phi^{(\beta)}_{\text{in}}\rangle = 0 \tag{27}$$

and because of the completeness of the in or out states. The remaining parts of the matrix element reduce, as in previous cases, to

$$\langle\phi^{(\alpha)}_{\text{in}}|\,[A_\mu(x), A_\nu(y)]_-\,|\phi^{(\beta)}_{\text{in}}\rangle$$
$$= \langle\phi^{(\alpha)}_{\text{in}}|\,[A^{\text{in}}_\mu(x), A^{\text{in}}_\nu(y)]_-\,|\phi^{(\beta)}_{\text{in}}\rangle$$
$$-\int g_{\mu\lambda}g_{\nu\pi}D(x-z_1)D(y-z_2)\theta(y-z_2)\theta(z_1-x)$$
$$\times\langle\phi^{(\alpha)}_{\text{in}}|\,\frac{\delta}{\delta A^{\text{in}}_\lambda(z_1)}\left(S^\dagger\frac{\delta S}{\delta A^{\text{in}}_\pi(z_2)}\right)|\phi^{(\beta)}_{\text{in}}\rangle\,d^4z_1\,d^4z_2$$
$$+\int g_{\mu\lambda}g_{\nu\pi}D(x-z_1)D(y-z_2)\theta(x-z_1)\theta(z_2-y)$$
$$\times\langle\phi^{(\alpha)}_{\text{in}}|\,\frac{\delta}{\delta A^{\text{in}}_\pi(z_2)}\left(S^\dagger\frac{\delta S}{\delta A^{\text{in}}_\lambda(z_1)}\right)|\phi^{(\beta)}_{\text{in}}\rangle\,d^4z_1\,d^4z_2. \tag{28}$$

The right-hand side of Eq. (28) vanishes on the spacelike domain $(x-y)^2 > 0$ if

$$\frac{\delta}{\delta A_\lambda^{\text{in}}(z_1)}\left(iS^\dagger \frac{\delta S}{\delta A_{\text{in}}^\nu(z_2)}\right) = \frac{\delta j_E^\nu(z_2)}{\delta A_\lambda^{\text{in}}(z_1)} = 0$$

$$\text{for } z_1^0 > z_2^0 \text{ and for } (z_1 - z_2)^2 > 0. \quad (29)$$

The LSZ expansion for the current operator is

$$\begin{aligned} j_E^\nu(z_2) &= \Box(z_2)A^\nu(z_2) \\ &= \sum_n \frac{1}{n!}\int (d^4x)\{\Box(z_2)\Box(x_1)\cdots\Box(x_n) \\ &\quad \times \langle 0|\, R(A^\nu(z_2)A_{\mu_1}(x_1)\cdots A_{\mu_n}(x_n))\,|0\rangle\} \\ &\quad \times :A_{\text{in}}^{\mu_1}(x_1)\cdots A_{\text{in}}^{\mu_n}(x_n):, \end{aligned} \quad (30)$$

and (20) follows upon functional differentiation in analogy to the previous cases.

Thus, strong local commutativity follows also for the electromagnetic field up to possibly gauge-variant terms which disappear for matrix elements between physical states.

V. THE CHARGED VECTOR BOSON

The free vector boson fields satisfy

$$\begin{aligned} &[\phi_\mu^{\text{in}}(x), \phi_\nu^{\text{in}}(y)]_- = 0, \quad [\phi_\mu^{\dagger\text{in}}(x), \phi_\nu^{\text{in}}(y)]_- = -i\Delta_{\mu\nu}(x-y), \\ &K(x)\phi_\mu^{\text{in}}(x) = 0, \quad \partial^\mu\phi_\mu^{\text{in}}(x) = 0, \\ &\Delta_{\mu\nu}(x) = (g_{\mu\nu} - \partial_\mu\partial_\nu/m^2)\Delta(x). \end{aligned} \quad (31)$$

We assume the interacting field in the form

$$\begin{aligned} \phi_\mu(x) &= \phi_\mu^{\text{in}}(x) \\ &\quad - i\int \theta(x-z)\Delta_{\mu\lambda}(x-z)S^\dagger \frac{\delta S}{\delta\phi_\lambda^{\dagger\text{in}}(z)}\, d^4z. \end{aligned} \quad (32)$$

The interacting-field commutator is, therefore,

$$\begin{aligned} [\phi_\mu^\dagger(x), \phi_\nu(y)] &= -i\Delta_{\mu\nu}(x-y) \\ &\quad - \int \theta(z_1 - x)\theta(y - z_2)\Delta_{\mu\tau}(x - z_1)\Delta_{\nu\lambda}(y - z_2) \\ &\quad \times \frac{\delta}{\delta\phi_\tau^{\text{in}}(z_1)}\left(S^\dagger \frac{\delta S}{\delta\phi_\lambda^{\dagger\text{in}}(z_2)}\right) d^4z_1\, d^4z_2 \\ &\quad + \int \theta(x - z_1)\theta(z_2 - y)\Delta_{\mu\tau}(x - z_1)\Delta_{\nu\lambda}(y - z_2) \\ &\quad \times \frac{\delta}{\delta\phi_\lambda^{\dagger\text{in}}(z_2)}\left(S^\dagger \frac{\delta S}{\delta\phi_\tau^{\text{in}}(z_1)}\right) d^4z_1\, d^4z_2. \end{aligned} \quad (33)$$

Again, the Bogoliubov condition,

$$\frac{\delta}{\delta\phi^{\text{in}}(z_1)}\left(S^\dagger \frac{\delta S}{\delta\phi_\lambda^{\dagger\text{in}}(z_2)}\right) = 0,$$

$$\text{for } z_1^0 > z_2^0 \text{ and for } (z_1 - z_2)^2 > 0, \quad (34)$$

is sufficient to prove the vanishing of the right-hand side of Eq. (33) for spacelike points $(x-y)^2 > 0$. However, the functional derivative with respect to the charged vector boson field cannot be explicitly defined.[2] The implicit definition is

$$[\phi_\mu^{\dagger\text{in}}(x), F(\phi\phi^\dagger)]_- = -i\int \Delta_{\mu\lambda}(x-z)\frac{\delta F}{\delta\phi_\lambda^{\text{in}}(z)}\, d^4z. \quad (35)$$

Therefore, we cannot give mathematical meaning to the Bogoliubov condition (34) for the charged vector boson. According to Eq. (35), we can only specify commutators and not functional derivatives explicitly. Although the LSZ expansion for the interacting vector field can be made, the techniques used for other fields break down in this case. Therefore, it does not seem to follow from the basic assumptions that Eq. (32) defines a microcausal theory.

VI. CONCLUSION

The Bogoliubov causality condition is not explicitly defined in general. It is therefore not meaningful to postulate it in all cases. For those cases where a definition exists this causality condition can be derived from the basic postulates of AQFT which do not involve microcausality. Strong local commutativity then follows by means of the Bogoliubov condition.

The fields that can thus not be proven to be microcausal are exactly those fields which spoil the possibility of a perturbation solution of the integral equation for the S matrix.

ACKNOWLEDGMENT

The partial support of this work by the National Science Foundation is gratefully acknowledged.

High-Order Limit of Para-Bose and Para-Fermi Fields

O. W. GREENBERG*
Department of Physics and Astronomy, University of Maryland, College Park, Maryland
AND
A. M. L. MESSIAH
Département de Physique Nucléaire et Physique du Solide, Centre d'Etudes Nucléaires de Saclay, Gif-sur-Yvette (Seine-et-Oise), France
(Received 26 August 1964)

Para-Bose and para-Fermi quantization are schemes of second quantization which generalize the usual Bose and Fermi schemes. The different cases of para-Bose (para-Fermi) quantization are labeled by a positive integer p, which is called the "order." For $p = 1$, the usual Bose (Fermi) quantization is recovered. The high-p limit of para-Bose (para-Fermi) quantization is studied, and it is shown that, to a large extent, the high-p limit of a para-Bose (para-Fermi) field is a Fermi (Bose) field. The nature of the limit is studied in detail, and a paradox relating to the connection of spin and statistics is resolved.

INTRODUCTION

PARA-BOSE and para-Fermi quantization, which generalize Bose and Fermi quantization, were invented by Green.[1] One of the most striking features of this generalization is that a doubly infinite number of new possibilities appears. For both the para-Bose and para-Fermi cases, the different possibilities are labeled by a number p, the "order," which can take any positive integral value. For $p = 1$, the para-Bose and para-Fermi fields reduce to Bose and Fermi fields, respectively. Knowing this fact brings to mind the question of the limiting behavior for $p \to \infty$, which is the subject of this article. Although with their original normalization the para operators do not approach a limit for large p, a simple change of normalization gives operators which do have such a limit. The amusing result is that, *to a large extent*, the high-p limits of the (renormalized) para-Bose and para-Fermi theories are the Fermi and Bose theories, respectively. What we mean by "to a large extent" should become clear in the following pages, and is summarized at the end of the paper.

* Supported in part by the U. S. Air Force Office of Scientific Research under contract AFOSR 500-64 and by the National Science Foundation under contract NSF GP 3221.

[1] H. S. Green, Phys. Rev. **90**, 270 (1953). A number of other papers study the mathematics of parafields. The requirement that the interaction Hamiltonian density be paralocal, together with certain experimental information, excludes all presently known particles from being para. A derivation of the selection rules on para particles necessary for this argument is given in O. W. Greenberg and A. M. L. Messiah, "Selection Rules for Parafields and the Absence of Para Particles in Nature," Phys. Rev. (to be published). This last paper also contains references to the literature.

RESUMÉ OF RELEVANT INFORMATION ABOUT PARAQUANTIZATION

Consider a set of annihilation and creation operators, a_k and $a_k^\dagger$, respectively, where k is a discrete quantum number, which satisfy the commutation relations

$$[[a_k^\dagger, a_l]_\pm, a_m]_- = -2\delta_{km}a_l, \tag{1}$$

$$[[a_k, a_l]_\pm, a_m]_- = 0, \tag{2}$$

and the no-particle condition

$$a_k\Phi_0 = 0, \tag{3}$$

where Φ_0 is the unique no-particle vector, and the upper (lower) signs are for para-Bose (para-Fermi) quantization. Then

Theorem[2]*:* All irreducible representations in a Hilbert space of Eqs. (1) and (2) which satisfy Eq. (3) also satisfy the additional condition

$$a_k a_l^\dagger\Phi_0 = p\delta_{kl}\Phi_0, \quad p \text{ a positive integer}, \tag{4}$$

and are characterized, up to unitary equivalence, by Eqs. (3) and (4). For $p = 1$, the para-Bose and para-Fermi cases reduce to Bose and Fermi, respectively.

For later reference, we define a commutative set of number operators n_k by

$$n_k = \tfrac{1}{2}([a_k^\dagger, a_k]_\pm \mp p). \tag{5}$$

These number operators have integer eigenvalues, and satisfy the usual commutation rules

$$[n_k, a_l^\dagger]_- = \delta_{kl}a_l^\dagger$$

with the creation operators.

[2] O. W. Greenberg and A. M. L. Messiah, cited in Ref. 1.

CHOICE OF OPERATORS WHICH HAVE A FINITE LIMIT

Because of Eq. (4), the operators $a_k^\dagger$ do not have a finite limit for $p \to \infty$. Consider, instead, the operators

$$c_k = p^{-\frac{1}{2}} a_k, \qquad c_k^\dagger = p^{-\frac{1}{2}} a_k^\dagger$$

for which Eq. (1), (2), (3) and (4) are replaced by

$$[[c_k^\dagger, c_l]_\pm, c_m]_- = (-2/p)\delta_{km} c_l, \tag{1'}$$

$$[[c_k, c_l]_\pm, c_m]_- = 0, \tag{2'}$$

$$c_k \Phi_0 = 0, \tag{3'}$$

$$c_k c_l^\dagger \Phi_0 = \delta_{kl} \Phi_0. \tag{4'}$$

For the c operators, p appears only in Eq. (1'), while for the a operators, p appears only in Eq. (4).

THE HILBERT SPACES $\mathfrak{C}(p)$ AND $\mathfrak{A}(p)$

At this point we emphasize the dependence of the theory of the para-Bose and para-Fermi operators on the order p. From now on we show the p dependence of all objects in the theory, and write the operators as $c_k(p)$ and $c_k(p)^\dagger$, and the no-particle state as $\Phi_0(p)$. The Hilbert space on which these operators act, $\mathfrak{C}(p)$, is defined to be the closure of vectors of the form

$$\mathcal{P}[c_k(p)^\dagger]\Phi_0(p),$$

where $\mathcal{P}$ is an arbitrary polynomial. Similar remarks apply for the original operators $a_k(p)$ and $a_k(p)^\dagger$, and, in particular for their Hilbert space, $\mathfrak{A}(p)$, which is defined to be the closure of vectors of the form $\mathcal{P}[a_k(p)^\dagger]\Phi_0(p)$. The Hilbert spaces $\mathfrak{C}(p)$ are the same as the $\mathfrak{A}(p)$ for each finite p, but not in the limit $p \to \infty$. Since the a and $a^\dagger$ operators are an irreducible set for all finite p, the same is true of the c and $c^\dagger$ operators. The operators $c(p)$ for a given p are defined only on $\mathfrak{C}(p)$; therefore operators for different values of p cannot be compared directly. However the expectation value of a polynomial in the c's for a given p in the no-particle state for the same p is a numerical function which can be compared to the corresponding function for a different value of p. Since the set of all vacuum expectation values uniquely characterizes the theory,[3] we base our discussion of the limit, for $p \to \infty$, of the para-Bose and para-Fermi theories of order p on the limits, for $p \to \infty$, of all the vacuum expectation values. The reconstruction of the Hilbert space and the operators from the vacuum expectation values, the "inverse problem," has been described by the authors cited in Ref. 3.

[3] A. S. Wightman, Phys. Rev. **101**, 860 (1956); W. Schmidt and K. Baumann, Nuovo Cimento **4**, 860 (1956); R. Haag and B. Schroer, J. Math. Phys. **3**, 248 (1962); R. F. Streater and A. S. Wightman, *PCT, Spin and Statistics and All That* (W. A. Benjamin, Inc., New York, 1964), Chap. 3; M. A. Naimark, *Normed Rings*, translated by L. F. Boron (P. Noordhoff, Ltd., Groningen, 1960), especially Chap. IV.

DEFINITION OF THE LIMIT, FOR $p \to \infty$

We define the limit, for $p \to \infty$, of $\mathfrak{C}(p)$ in terms of vectors and operators without explicit p dependence. For such vectors and operators, we define the limiting elements to be the corresponding vectors and operators constructed from the limit of the vacuum expectation values. In general, the limiting space need not be a Hilbert space; however the limiting spaces for the para-Bose and para-Fermi theories turn out to be well-known Hilbert spaces. By our definition, the limit of every vector or operator without explicit p dependence is a vector or operator in the limiting Hilbert space, and conversely every vector or operator in the limiting space is the limit of some vector or operator without explicit p dependence in the spaces $\mathfrak{C}(p)$.

DISCUSSION OF THE LIMIT FOR EXPRESSIONS WITH EXPLICIT p DEPENDENCE

In this case, the definition of the limit is more subtle. Let $O(p)$ be such an expression. For the limit $p \to \infty$ of $O(p)$ to exist we require: (1) that the limits of all vacuum expectation values of $\mathcal{P}_1[c_k(p)]O(p)\mathcal{P}_2[c_k(p)^\dagger]$ exist, and (2) that there exist an operator $O(\infty)$ in the limit space which has corresponding vacuum expectation values, i.e., such that

$$\lim_{p\to\infty} (\Phi_0(p), \mathcal{P}_1[c_k(p)]O(p)\mathcal{P}_2[c_k(p)^\dagger]\Phi_0(p))$$
$$= (\Phi_0(\infty), \mathcal{P}_1[c_k(\infty)]O(\infty)\mathcal{P}_2[c_k(\infty)^\dagger]\Phi_0(\infty)),$$

for all $\mathcal{P}_1$, $\mathcal{P}_2$.

We emphasize that $c_k(\infty)$, $c_k(\infty)^\dagger$, and $\Phi_0(\infty)$ have already been defined in the paragraph above. We find that there are expressions with explicit p dependence for which Condition (1) holds, but for which there is no operator $O(\infty)$ which can satisfy Condition (2). In such cases we say that $O(p)$ does not have a limit.

THE LIMIT FOR THE SPACES $\mathfrak{C}(p)$

The limits of our two families of Hilbert spaces are

$$\lim_{p\to\infty} \mathfrak{C}^{(B)}(p) = \mathfrak{C}^{(F)}(1), \tag{6}$$

and

$$\lim_{p\to\infty} \mathfrak{C}^{(F)}(p) = \mathfrak{C}^{(B)}(1), \tag{7}$$

where, for clarity, we have distinguished the para-Bose and para-Fermi families by the superscripts (B) and (F), respectively. In words, Eq. (6) and (7) mean that the limit, for $p \to \infty$, of the para-Bose

theory is the Fermi theory, and the limit of the para-Fermi theory is the Bose theory. We remark that in the limit the spaces $\mathcal{C}(p)$ become smaller. For example, $\mathcal{C}^{(F)}(1)$ has only antisymmetric states, while $\mathcal{C}^{(B)}(p)$, $p > 1$, has states with a large variety of different permutation symmetries. The limits of the spaces $\mathcal{A}(p)$ are degenerate (one dimensional) since the only vector in $\mathcal{A}(p)$ which has a limit is $\Phi_0(p)$.

The results for the $\mathcal{C}(p)$ can be seen by studying the algebraic manipulations used in the calculation of the vacuum expectation value $(\Phi_0(p), \mathcal{P}[c_k(p), c_k(p)^\dagger]\Phi_0(p))$ of an arbitrary polynomial in the c and $c^\dagger$. The calculation consists in moving the annihilation operators from left to right as many times as necessary by applying Eqs. (1′) and (2′) and then using the no-particle conditions, Eqs. (3′) and (4′). The reader should convince himself that the calculation can actually be completed in the way just described. Since the right-hand sides in Eqs. (1′) and (2′) introduce nothing but higher-order terms in p^{-1}, the leading term is obtained by performing the same algebraic manipulations with modified commutation relations obtained from Eqs. (1′) and (2′) by replacing all the right-hand sides by zero. Now, these modified commutation relations together with the (unmodified) no-particle conditions are satisfied by the Fermi operators in the para-Bose case, and by the Bose operators in the para-Fermi case. Consequently, the limit for $p \to \infty$ of the above expectation value for a para-Bose (para-Fermi) operator is equal to the expectation value of the same expression for a Fermi (Bose) operator.

For illustration, we give the exact expression for a typical vacuum expectation value in the para-Bose case:

$$(\Phi_0(p), c_k(p)c_l(p)c_m(p)^\dagger c_n(p)^\dagger \Phi_0(p)) = \delta_{kn}\delta_{lm} + (2/p - 1)\delta_{km}\delta_{ln}. \quad (8)$$

For $p = 1$, this is a Bose matrix element; the $p \to \infty$ limit is the corresponding Fermi matrix element.

To give a geometric insight to what is happening here, let us consider the subspace of $\mathcal{C}^{(B)}(p)$ spanned by the two vectors $c_k(p)^\dagger c_l(p)^\dagger \Phi_0(p)$ and $c_l(p)^\dagger c_k(p)^\dagger \Phi_0(p)$, $k \neq l$. According to Eq. (8), these are unit vectors and the angle between them, $\alpha = \arccos(2/p - 1)$, increases from 0 to π when p goes from 1 to ∞. Thus, for $p = 1$, the two vectors are equal and span a one-dimensional space. For greater values of p, the angle between the two vectors opens up, and the space becomes two dimensional. But, in the $p \to \infty$ limit, the two vectors become opposite and they again span a one-dimensional space.

It follows from our discussion that the limit of any polynomial of the para-Bose (para-Fermi) operators which does not have explicit p dependence (that is p occurs only implicitly in the c's) exists as an operator in the Fermi (Bose) space, and is the same polynomial of the Fermi (Bose) operators.

THE LIMIT FOR POLYNOMIALS IN THE NUMBER OPERATORS $n_k(p)$

However the limit of a polynomial which has explicit p dependence need not exist, and if it does exist it can have a different functional form. An important class of polynomials with explicit p dependence is polynomials in the number operators $n_k(p)$, as defined by Eq. (5). In terms of the c and $c^\dagger$ operators,

$$n_k(p) = \tfrac{1}{2}p\{[c_k(p)^\dagger, c_k(p)]_\pm \mp 1\}. \quad (5')$$

The Hamiltonian for free parabosons,

$$H_0^{(B)}(p) = \sum_k \omega_k n_k(p) = \tfrac{1}{2}p \sum_k \omega_k\{[c_k(p)^\dagger, c_k(p)]_+ - 1\},$$

is a good example of a polynomial in the number operators. The limit of such polynomials exists and is given by the following theorem: the limit for $p \to \infty$ of $\mathcal{P}[c_k(p), c_k(p)^\dagger, n_k(p)]$ in the operators c, $c^\dagger$, and n of the para-Bose (para-Fermi) theory of order p is the same polynomial in the corresponding operators of the Fermi (Bose) theory. To show that Condition (1) for existence of the limit holds, we note that the commutation relations

$$[n_k(p), c_l(p)]_- = -\delta_{kl}c_k(p), \quad (9a)$$

$$[n_k(p), c_l(p)^\dagger]_- = \delta_{kl}c_k(p)^\dagger, \quad (9b)$$

and the property of the vacuum

$$n_k(p)\Phi_0(p) = 0 \quad (10)$$

hold for all parafield theories as well as for the ordinary Bose and Fermi theories. In order to calculate the vacuum expectation value of the above polynomial $\mathcal{P}$ we first make the number operators travel toward the extreme right through repeated application of Eq. (9), and then use Eq. (10). The expression we are left with is the vacuum expectation value of a polynomial in the $c_k(p)$ and $c_k(p)^\dagger$. Since the algebraic manipulations for a para-Bose (para-Fermi) field of order p and for a Fermi (Bose) field are exactly the same, we obtain the same polynomial in both cases. We have proved above that the $p \to \infty$ limit of the expectation value of the former is equal to the expectation value of the latter, so Condition

(1) holds. Clearly Condition (2) also holds since the Fermi (Bose) number operators are the required limiting operators acting in the limit space $\mathfrak{C}^{(F)}(1)[\mathfrak{C}^{(B)}(1)]$.

Since the above result applies to any polynomial in the n's it applies in particular to the n's themselves and to the free Hamiltonian, leading to the following results for their limits:

$$\lim_{p\to\infty} n_k^{(B)}(p) = n_k^{(F)}(1), \tag{11}$$

$$\lim_{p\to\infty} H_0^{(B)}(p) = H_0^{(F)}(1). \tag{12}$$

We write out Eq. (11) in terms of the c-operators to show the qualitative change in form which occurs,

$$\lim_{p\to\infty} \tfrac{1}{2}p([c_k^{(B)}(p)^\dagger, c_k^{(B)}(p)]_+ - 1) = \tfrac{1}{2}\{[c_k^{(F)}(1)^\dagger, c_k^{(F)}(1)]_- + 1\}. \tag{11'}$$

Similar results are obtained in the case of parafermions.

Our statements about the $p \to \infty$ limits of the parafield theories are schematically illustrated in Fig. 1.

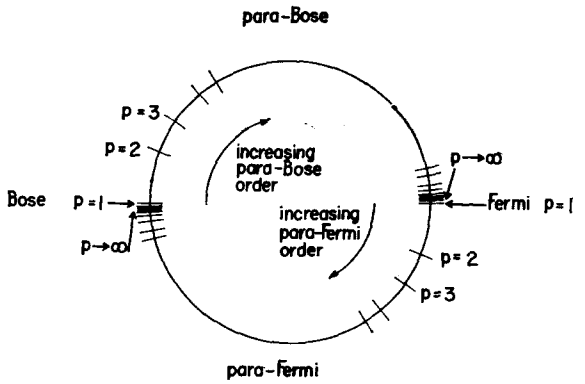

FIG. 1. The parafield clock.

PARADOX RELATING TO THE CONNECTION OF SPIN AND STATISTICS

We expect that the para-Bose and para-Fermi operators satisfy the connection of spin and statistics; that is, para-Bose operators must have integral spin and para-Fermi operators must have odd half-integral spin. On the other hand, in the limit $p \to \infty$, the para-Bose and para-Fermi operators become Fermi and Bose operators, respectively, and so violate the connection of spin and statistics. Since the field does not change its spin in this limit, there is a paradox. To formulate this paradox more precisely in a specific case, we consider the limit of the free spin-0 para-Bose theories. For all of these,

$$[\mathfrak{H}^{(B)}(x, p), \mathfrak{H}^{(B)}(y, p)]_- = 0, \qquad (x - y)^2 < 0,$$

where $\mathfrak{H}^{(B)}$ is the Hamiltonian density of the para-Bose theory of order p. However the limiting theory violates the connection of spin and statistics, and

$$[\mathfrak{H}^{(F)}(x, 1), \mathfrak{H}^{(F)}(y, 1)]_- \neq 0, \qquad (x - y)^2 < 0,$$

where $\mathfrak{H}^{(F)}$ is the free spin-0 Fermi Hamiltonian density. In fact no inconsistency occurs here, since as we now show, $\mathfrak{H}^{(B)}(x, p)$ has no limit, and *a fortiori*, its limit is not $\mathfrak{H}^{(F)}(x, 1)$. The Hamiltonian density $\mathfrak{H}^{(B)}$ is

$$\mathfrak{H}^{(B)}(x, p) = \tfrac{1}{2}p\{[\partial\phi(x)/\partial x^0]^2 + [\nabla\phi(x)]^2 + m^2[\phi(x)]^2\} - \langle\text{same}\rangle_0,$$

where

$$\phi(x) = (2\pi)^{-\frac{3}{2}} \int d^3\mathbf{k}[2\omega(\mathbf{k})]^{-\frac{1}{2}} \times [c^{(B)}(\mathbf{k}, p)e^{-ikx} + c^{(B)}(\mathbf{k}, p)^\dagger e^{ikx}],$$

$k^0 = \omega(\mathbf{k}) = (\mathbf{k}^2 + m^2)^{\frac{1}{2}}$, and the $c(\mathbf{k}, p)$ have Dirac δ functions rather than Kronecker δ functions in their commutation relations. When expressed in terms of the c's, $\mathfrak{H}^{(B)}$ contains expressions such as

$$\tfrac{1}{2}p([c^\dagger, c]_+ - \delta), \quad \tfrac{1}{2}p[c, c]_+, \quad \tfrac{1}{2}p[c^\dagger, c^\dagger]_+.$$

The reader can verify that

$$\lim_{p\to\infty} \tfrac{1}{2}p\{[c^{(B)}(\mathbf{k}, p)^\dagger, c^{(B)}(1, p)]_+ - \delta(\mathbf{k} - 1)\} = \tfrac{1}{2}\{[c^{(F)}(\mathbf{k}, 1)^\dagger, c^{(F)}(1, 1)]_- + \delta(\mathbf{k} - 1)\}$$

in our sense of the limit. However the limit of

$$O^{(B)}(\mathbf{k}, 1, p) \equiv \tfrac{1}{2}p[c^{(B)}(\mathbf{k}, p), c^{(B)}(1, p)]_+$$

and the limit of the corresponding adjoint do not exist. Condition (1) for the existence of the limit of $O^{(B)}$ is satisfied. A demonstration of this can be given in a way similar to the discussion of the limit of polynomials in the number operators given above. Here, for brevity, we exhibit only the simplest non-vanishing matrix element of $O^{(B)}$:

$$(\Phi_0^{(B)}(p), O^{(B)}(\mathbf{k}, 1, p)c^{(B)}(\mathbf{q}_1, p)^\dagger c^{(B)}(\mathbf{q}_2, p)^\dagger \Phi_0^{(B)}(p)) = \delta(1 - \mathbf{q}_1)\delta(\mathbf{k} - \mathbf{q}_2) + \delta(\mathbf{k} - \mathbf{q}_1)\delta(1 - \mathbf{q}_2). \tag{13}$$

Since the right-hand side is independent of p, the $p \to \infty$ limit of the left-hand side of Eq. (13) exists. Condition (2) for the existence of the limit requires that there exist an operator $O^{(F)}$ on the Fermi space $\mathfrak{C}^{(F)}(1)$ which has corresponding matrix elements; for our specific matrix element $O^{(F)}$ must satisfy

$$(\Phi_0^{(F)}(1), O^{(F)}(\mathbf{k}, 1, 1)c^{(F)}(\mathbf{q}_1, 1)^\dagger c^{(F)}(\mathbf{q}_2, 1)^\dagger \Phi_0^{(F)}(1)) = \delta(1 - \mathbf{q}_1)\delta(\mathbf{k} - \mathbf{q}_2) + \delta(\mathbf{k} - \mathbf{q}_1)\delta(1 - \mathbf{q}_2). \tag{14}$$

Clearly there is no such operator $O^{(F)}$ on $\mathfrak{C}^{(F)}(1)$,

since in $\mathcal{C}^{(F)}(1)$ the matrix element must be antisymmetric under interchange of $\mathbf{q}_1$, and $\mathbf{q}_2$. Therefore the limit of $O^{(B)}$, and also of $\mathcal{K}^{(B)}$, does not exist, and no inconsistency occurs.

SUMMARY

We can summarize this situation by saying that in the high-p limit a spin-0 para-Bose field has (a) annihilation and creation operators and an anticommutator which approach the annihilation and creation operators and c-number anticommutator of a spin-0 Fermi field and (b) a Hamiltonian and a set of number operators which approach those of the spin-0 Fermi field, but that, on the contrary, the Hamiltonian density of the high-order para-Bose theory does not converge to an operator in the Fermi theory. Analogous statements hold for the high-p limit of a spin-$\frac{1}{2}$ para-Fermi field.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 6, NUMBER 4 APRIL 1965

Statistical Mechanics of Assemblies of Coupled Oscillators*

G. W. FORD†
Department of Physics, University of Michigan, Ann Arbor, Michigan
M. KAC
The Rockefeller Institute, New York, New York
AND
P. MAZUR
Lorentz Institute for Theoretical Physics, Leiden, The Netherlands
(Received 25 September 1964)

It is shown that a system of coupled harmonic oscillators can be made a model of a heat bath. Thus a particle coupled harmonically to the bath and by an arbitrary force to a fixed center will (in an appropriate limit) exhibit Brownian motion. Both classical and quantum mechanical treatments are given.

1. INTRODUCTION

OUR aim here is to study a simple mechanical model, a chain of coupled harmonic oscillators, in order to come to a deeper understanding of some of the phenomena associated with Brownian motion.[1,2] With this model we are able to carry through the program one would like to achieve with more realistic interactions. This program, which really goes back to Gibbs, goes as follows[3]:

(i) Solve the equations of motion of the mechanical system consisting of a Brownian particle coupled to heat bath. The solution consists of expressions for the coordinates and momenta at time t in terms of the initial coordinates and momenta.

(ii) Assume the initial coordinates and momenta of the heat bath to be distributed according to some statistical distribution, e.g., that of the canonical ensemble.

(iii) Show that the coordinate and momentum of the Brownian particle, as functions of time, will then represent stochastic processes (whose statistical properties arise from the initial distribution of the heat bath) of the kind familiar from standard theories.

This is a very ambitious program, and it is no wonder that it can be carried out only for the simplest models.

* Part of the work reported in this paper was done while one of us (M. K.) was H. A. Lorentz Visiting Professor at the University of Leiden. When this work was reported at Yeshiva University in the Fall of 1963 Dr. N. L. Balazs informed us that he had independently obtained some of the results concerning the classical case.

† Supported in part by a grant from the National Science Foundation.

[1] There is an extensive literature on the motion of coupled oscillators, mostly concerned with motion in a lattice with nearest-neighbor interactions. Some of the more recent articles which have a bearing on our work are: P. Mazur and E. Montroll, J. Math. Phys. **1**, 70 (1960); P. C. Hemmer, "Dynamic and Stochastic Types of Motion in the Linear Chain," thesis, Norges Tekniske Høgskole, Trondheim, Norway (1959); R. J. Rubin, J. Math. Phys. **1**, 309 (1960); **2**, 373 (1961); M. Toda and Y. Koguri, Suppl. Progr. Theoret. Phys. (Kyoto) **23**, 157 (1962); R. E. Turner, Physica **26**, 274 (1960).

[2] The classic papers on the phenomological theory of Brownian motion are: G. E. Uhlenbeck and L. S. Ornstein, Phys. Rev. **36**, 823 (1930); M. C. Wang and G. E. Uhlenbeck, Rev. Mod. Phys. **17**, 323 (1945).

[3] See the article by H. Wergeland in *Fundamental Problems in Statistical Mechanics*, edited by E. G. D. Cohen (North-Holland Publishing Company, Amsterdam, 1962).

since in $\mathcal{C}^{(F)}(1)$ the matrix element must be antisymmetric under interchange of $\mathbf{q}_1$, and $\mathbf{q}_2$. Therefore the limit of $O^{(B)}$, and also of $\mathcal{K}^{(B)}$, does not exist, and no inconsistency occurs.

SUMMARY

We can summarize this situation by saying that in the high-p limit a spin-0 para-Bose field has (a) annihilation and creation operators and an anticommutator which approach the annihilation and creation operators and c-number anticommutator of a spin-0 Fermi field and (b) a Hamiltonian and a set of number operators which approach those of the spin-0 Fermi field, but that, on the contrary, the Hamiltonian density of the high-order para-Bose theory does not converge to an operator in the Fermi theory. Analogous statements hold for the high-p limit of a spin-$\frac{1}{2}$ para-Fermi field.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 6, NUMBER 4 APRIL 1965

Statistical Mechanics of Assemblies of Coupled Oscillators*

G. W. FORD†
Department of Physics, University of Michigan, Ann Arbor, Michigan
M. KAC
The Rockefeller Institute, New York, New York
AND
P. MAZUR
Lorentz Institute for Theoretical Physics, Leiden, The Netherlands
(Received 25 September 1964)

It is shown that a system of coupled harmonic oscillators can be made a model of a heat bath. Thus a particle coupled harmonically to the bath and by an arbitrary force to a fixed center will (in an appropriate limit) exhibit Brownian motion. Both classical and quantum mechanical treatments are given.

1. INTRODUCTION

OUR aim here is to study a simple mechanical model, a chain of coupled harmonic oscillators, in order to come to a deeper understanding of some of the phenomena associated with Brownian motion.[1,2] With this model we are able to carry through the program one would like to achieve with more realistic interactions. This program, which really goes back to Gibbs, goes as follows[3]:

(i) Solve the equations of motion of the mechanical system consisting of a Brownian particle coupled to heat bath. The solution consists of expressions for the coordinates and momenta at time t in terms of the initial coordinates and momenta.

(ii) Assume the initial coordinates and momenta of the heat bath to be distributed according to some statistical distribution, e.g., that of the canonical ensemble.

(iii) Show that the coordinate and momentum of the Brownian particle, as functions of time, will then represent stochastic processes (whose statistical properties arise from the initial distribution of the heat bath) of the kind familiar from standard theories.

This is a very ambitious program, and it is no wonder that it can be carried out only for the simplest models.

* Part of the work reported in this paper was done while one of us (M. K.) was H. A. Lorentz Visiting Professor at the University of Leiden. When this work was reported at Yeshiva University in the Fall of 1963 Dr. N. L. Balazs informed us that he had independently obtained some of the results concerning the classical case.

† Supported in part by a grant from the National Science Foundation.

[1] There is an extensive literature on the motion of coupled oscillators, mostly concerned with motion in a lattice with nearest-neighbor interactions. Some of the more recent articles which have a bearing on our work are: P. Mazur and E. Montroll, J. Math. Phys. **1**, 70 (1960); P. C. Hemmer, "Dynamic and Stochastic Types of Motion in the Linear Chain," thesis, Norges Tekniske Høgskole, Trondheim, Norway (1959); R. J. Rubin, J. Math. Phys. **1**, 309 (1960); **2**, 373 (1961); M. Toda and Y. Koguri, Suppl. Progr. Theoret. Phys. (Kyoto) **23**, 157 (1962); R. E. Turner, Physica **26**, 274 (1960).

[2] The classic papers on the phenomological theory of Brownian motion are: G. E. Uhlenbeck and L. S. Ornstein, Phys. Rev. **36**, 823 (1930); M. C. Wang and G. E. Uhlenbeck, Rev. Mod. Phys. **17**, 323 (1945).

[3] See the article by H. Wergeland in *Fundamental Problems in Statistical Mechanics*, edited by E. G. D. Cohen (North-Holland Publishing Company, Amsterdam, 1962).

We have a good idea of what the results of this program should be, since, after all, Brownian motion is a throughly studied experimental phenomenon with a satisfactory phenomenological theory. In general we expect to show:

(i) The approach to equilibrium. In particular, the distribution of momentum of the Brownian particle should approach the Maxwellian distribution.

(ii) The description of this approach to equilibrium should be *contracted*, i.e., should involve only a small number of the possible variables describing the system. Another way of saying this is that there should be a reduced description in terms of which the stochastic process is Markoffian.

(iii) What changes occur when one adopts a quantum description of the system. Here we have less of an idea of what we should expect, but somehow the basic features of the stochastic process should be preserved in the quantum description.

We can be still more explicit about what we mean by a contracted description of Brownian motion; we mean the Langevin equation of motion. For a Brownian particle of mass m acted upon by an outside force $F(x)$ this equation is

$$\dot{p} = -fp/m + E(t) + F(x), \tag{1}$$

where $p = m\dot{x}$ is the momentum of the Brownian particle, f the friction constant, and $E(t)$ is the random force due to the heat bath. This random force is a purely random Gaussian process characterized by

$$\langle E(t)\rangle = 0, \quad \langle E(t)E(t')\rangle = 2fkT\delta(t - t'),$$

where T is the temperature of the heat bath and k is Boltzmann's constant. Note that the Langevin equation is a contracted description in the sense that the heat bath is described by only two parameters, the friction constant and the temperature, and that only the first two time derivatives of the position x of the Brownian particle appear.[4]

In Sec. 2 we discuss the dynamics of a system of coupled oscillators. There we *formally* carry through the program for the case of an arbitrary coupling of the oscillators. In Sec. 3 the arbitrary coupling of a linear chain is considered, and we show that there is a coupling for which, in the limit of an infinite chain, the resulting stochastic process is Markoffian. Taking the chain of oscillators with this coupling as the heat bath, we derive in Sec. 4 the Langevin equation for a Brownian particle. In Sec. 5 we discuss the quantum description of the system, and in Sec. 6 we discuss the quantum Langevin equation. Finally, in Sec. 7 we consider the Brownian motion of a quantum oscillator.

[4] That is, there are no memory effects.

2. DYNAMICS OF A SYSTEM OF COUPLED OSCILLATORS

Consider a system of $(2N + 1)$ coupled oscillators with Hamiltonian:

$$H = \frac{1}{2}\sum_{i=-N}^{N} p_i^2 + \frac{1}{2}\sum_{i,k=-N}^{N} q_i A_{ik} q_k. \tag{2}$$

Here q_i and p_i are, respectively, the canonical coordinate and momentum of the jth oscillator. The mass of each oscillator has been taken to be unity. The interactions of the oscillators are characterized by the $(2N + 1) \times (2N + 1)$ symmetric matrix $\mathbf{A}$, whose elements are the A_{ik}. At present, we make no special assumptions about this matrix except that it has no negative eigenvalues. The canonical equations of motion may be conviently written in matrix notations as follows:

$$\dot{\mathbf{q}} = \mathbf{p}, \qquad \dot{\mathbf{p}} = -\mathbf{A}\mathbf{q}. \tag{3}$$

Here $\mathbf{p}$ and $\mathbf{q}$ are $(2N + 1)$-rowed column matrices whose elements are the p_i and q_i, respectively. The formal solution of the equations of motion is

$$\begin{aligned} \mathbf{q}(t) &= \cos(\mathbf{A}^{\frac{1}{2}}t)\mathbf{q}(0) + \mathbf{A}^{-\frac{1}{2}}\sin(\mathbf{A}^{\frac{1}{2}}t)\mathbf{p}(0), \\ \mathbf{p}(t) &= -\mathbf{A}^{\frac{1}{2}}\sin(\mathbf{A}^{\frac{1}{2}}t)\mathbf{q}(0) + \cos(\mathbf{A}^{\frac{1}{2}}t)\mathbf{p}(0), \end{aligned} \tag{4}$$

where, e.g.,

$$\cos \mathbf{A}^{\frac{1}{2}}t = \sum_{n=0}^{\infty} \frac{(-)^n}{(2n)!}\mathbf{A}^n t^{2n}. \tag{5}$$

We now assume that at $t = 0$ the system is in equilibrium at temperature T. That is, we assume that the $q_i(0)$ and $p_i(0)$ are distributed according to the canonical distribution

$$D(\mathbf{q}(0), \mathbf{p}(0)) = (2\pi/\beta)^{2N+1}(\det \mathbf{A})^{-\frac{1}{2}}e^{-\beta H(\mathbf{q}(0),\mathbf{p}(0))}, \tag{6}$$

where $\beta = (kT)^{-1}$ and $\det \mathbf{A}$ is the determinant of $\mathbf{A}$. Note that there is a difficulty here, since $\det \mathbf{A} = 0$ if $\mathbf{A}$ has zero eigenvalues. We therefore assume for the time being that $\mathbf{A}$ has no zero eigenvalues. The expectation of any function $F(\mathbf{q}(0), \mathbf{p}(0))$ is given by

$$\begin{aligned} \langle F\rangle = \int \cdots \int dq_{-N}(0) \cdots dq_N(0)\, dp_{-N}(0) \cdots dp_N(0) \\ \times F(\mathbf{q}(0), \mathbf{p}(0))D(\mathbf{q}(0), \mathbf{p}(0)). \end{aligned} \tag{7}$$

Now we ask, what are the properties of the stochastic variables $q_i(t)$ and $p_i(t)$ which result from (4) under the distribution (6)? First of all, it is clear that the process is Gaussian. This follows from the fact that the distribution (6) is Gaussian and that the relation (4) is linear. That the process is sta-

tionary follows from the Liouville theorem of mechanics, which states that

$$D(\mathbf{q}(t), \mathbf{p}(t)) = D(\mathbf{q}(0), \mathbf{p}(0)). \tag{8}$$

It is well known that the statistical properties of such a stationary Gaussian process are completely described by the pair correlation functions. In our case these are obtained in Appendix 1; the results are[5]

$$\langle p_j(t)p_k(t+\tau)\rangle = kT\,||\cos \mathbf{A}^{\frac{1}{2}}\tau||_{jk}, \tag{9a}$$

$$\langle q_j(t)p_k(t+\tau)\rangle = -kT\,||\mathbf{A}^{-\frac{1}{2}}\sin \mathbf{A}^{\frac{1}{2}}\tau||_{jk}, \tag{9b}$$

$$\langle q_j(t)q_k(t+\tau)\rangle = kT\,||\mathbf{A}^{-1}\cos \mathbf{A}^{\frac{1}{2}}\tau||_{jk}. \tag{9c}$$

Note that the position correlation (9c) involves the inverse of **A**, which does not exist if **A** has zero eigenvalues.

If we fix our attention on a single oscillator, say the one with index 0, the momentum autocorrelation is

$$\langle p_0(t)p_0(t+\tau)\rangle = kT\,||\cos \mathbf{A}^{\frac{1}{2}}\tau||_{00}. \tag{10}$$

This is the autocorrelation of a stationary Gaussian process in one variable. It is well known that such a process is Markoffian if and only if the autocorrelation is an exponential, i.e.,

$$\langle p_0(t)p_0(t+\tau)\rangle = kTe^{-f|\tau|}, \tag{11}$$

where f is a positive constant. The question we turn to in the next section is that of finding an interaction matrix **A** for which (10) assumes the form (11).

3. THE INTERACTION MATRIX

In our model we assume the $(2N+1)$ oscillators are identical and that they are arranged in a chain with cyclic boundary conditions. This means that the interaction matrix A is a symmetric cyclic matrix.[6] The elements of such a matrix can be written in the form

$$A_{mn} = \frac{1}{2N+1}\sum_{k=-N}^{N}\omega_k^2 \times \exp\left\{i\frac{2\pi}{2N+1}k(m-n)\right\}, \tag{12}$$

where the symmetry of A requires

$$\omega_k^2 = \omega_{-k}^2. \tag{13}$$

The eigenvalues of this matrix are the quantities ω_s^2, $s = -N, -N+1, \cdots, N$. That is

$$\mathbf{A}\xi^{(s)} = \omega_s^2\xi^{(s)}, \tag{14}$$

where the eigenvector $\xi^{(s)}$ is a $(2N+1)$-rowed column matrix whose elements are

$$\xi_n^{(s)} = (2N+1)^{-\frac{1}{2}}\exp\{i[2\pi/(2N+1)]sn\}. \tag{15}$$

These properties follow from the elementary formula

$$\sum_{k=-N}^{N}\exp\left(ik(m-n)\frac{2\pi}{2N+1}\right) = \delta_{m,n}, \quad -N \le m, n \le N. \tag{16}$$

With this formula we can also readily demonstrate that if $F(\mathbf{A})$ is a function of the matrix **A**, then

$$||F(\mathbf{A})||_{m,n} = \frac{1}{2N+1}\sum_{k=-N}^{N}F(\omega_k^2) \times \exp\left\{i\frac{2\pi}{2N+1}k(m-n)\right\}. \tag{17}$$

Note, incidentally, that the special case of nearest-neighbor interactions is that for which

$$\omega_s^2 = \tilde{\omega}^2\sin^2[\pi s/(2N+1)]. \tag{18}$$

Consider next the limit $N \to \infty$, the infinite chain. If we make the additional assumption that ω_s^2 is slowly varying function of s, then (12) becomes

$$A_{mn} = \frac{1}{2\pi}\int_{-\pi}^{\pi}d\theta\, f(\theta)e^{i(m-n)\theta} = \frac{1}{2\pi}\int_{-\pi}^{\pi}d\theta\, f(\theta)\cos(m-n)\theta, \tag{19}$$

where

$$f(\theta) = \{\omega_s^2\}_{s=(2N+1)\theta/2\pi}. \tag{20}$$

The relation (17) becomes in this limit

$$||F(\mathbf{A})||_{m,n} = \frac{1}{2\pi}\int_{-\pi}^{\pi}d\theta\, F(f(\theta))\cos(m-n)\theta. \tag{21}$$

We are now ready to turn to the problem posed at the end of the last section; that of finding an interaction matrix A for which

$$||\cos \mathbf{A}^{\frac{1}{2}}t||_{00} = e^{-f|t|}. \tag{22}$$

Using the result (17) we see that for a finite matrix

$$||\cos \mathbf{A}^{\frac{1}{2}}t||_{00} = \frac{1}{2N+1}\sum_{k=-N}^{N}\cos\omega_k t. \tag{23}$$

For any choice of the ω_k this is a quasiperiodic function and cannot be of the form (22). However, in the limit of large N, we can use (21) which gives

$$||\cos \mathbf{A}^{\frac{1}{2}}t||_{00} = \frac{1}{2\pi}\int_{-\pi}^{\pi}d\theta\cos\{[f(\theta)]^{\frac{1}{2}}t\}. \tag{24}$$

Taken with our requirement (22), this becomes an integral equation for $f(\theta)$. The answer is essentially

[5] We use the notation $||\mathbf{M}||_{jk}$ for the element in the jth row and kth column of a matrix **M**.

[6] See, e.g., G. Kowalewski, *Determinantentheorie* (Chelsea Publishing Company, New York, 1948), 3rd ed., p. 105.

unique and is[7]

$$f(\theta) = f^2 \tan^2 \tfrac{1}{2}\theta. \tag{25}$$

There is a difficulty here, however; when (25) is inserted in the expression (19) for the matrix elements, the expression diverges! What we must do is employ a second limiting process (after the limit $N \to \infty$), defining

$$f_{\omega_L}(\theta) = \begin{cases} f^2 \tan^2 \dfrac{\theta}{2}, & |\theta| < \theta_L, \\ 0, & \theta_L \leq |\theta| \leq \pi. \end{cases} \tag{26}$$

Here

$$\omega_L \equiv f \tan \tfrac{1}{2}\theta_L \tag{27}$$

is a high-frequency cutoff in the spectrum of eigenfrequencies which ensures that the matrix elements (19) are finite. This frequency cutoff corresponds to a "microscopic interaction time" ω_L^{-1} which we assume is very small compared with the "macroscopic relaxation time" f^{-1}. The result (22) holds strictly only in the limit $\omega_L \to \infty$. Alternatively, we can say that, for $\omega_L \gg f$ the result (22) holds for times long compared with ω_L^{-1}.

Our model, then, is that the interaction matrix elements are given by (19) with $f(\theta)$ given by (26) with $\omega_L \gg f$. If in (24) we make the change of variable $\omega = f \tan \frac{1}{2}\theta$, we find

$$||\cos \mathbf{A}^{\frac{1}{2}}t||_{00} = \frac{1}{\pi}\int_{-\omega_L}^{\omega_L} d\omega \frac{f}{\omega^2 + f^2} \cos \omega t. \tag{28}$$

In the limit $\omega_L \to \infty$ this becomes $e^{-f|t|}$ and, therefore, the Gaussian process $p_0(t)$ becomes also Markoffian.

4. THE LANGEVIN EQUATION

Having seen that our model leads to a Gaussian Markoffian stochastic process for the collection of coupled oscillators, we are led to ask whether it also leads to the Langevin equation for the motion of a single particle coupled to a heat bath consisting of such oscillators. In this section we see that this is indeed the case.

We select from the chain of $(2N + 1)$ oscillators, the particle with index 0 to be the Brownian particle; the remaining $2N$ oscillators represent the heat bath. The outside force on this particle we denote by

$$F(t) \equiv F(q_0(t)). \tag{29}$$

If we define $\mathbf{F}(t)$ to be a $(2N + 1)$-rowed column matrix whose elements are all zero except for the zeroth element which is $F(t)$, then in the notation of Sec. 2 the equations of motion for coupled "particle and heat path" are

$$\dot{\mathbf{q}} = \mathbf{p}, \qquad \dot{\mathbf{p}} = -\mathbf{A}\mathbf{q} + \mathbf{F}(t). \tag{30}$$

The formal solution of these equations is

$$\mathbf{q}(t) = \cos \mathbf{A}^{\frac{1}{2}}t\mathbf{q}(0) + \mathbf{A}^{-\frac{1}{2}} \sin \mathbf{A}^{\frac{1}{2}}t\mathbf{p}(0) + \int_0^t dt' \frac{\sin \mathbf{A}^{\frac{1}{2}}(t - t')}{\mathbf{A}^{\frac{1}{2}}} \mathbf{F}(t'), \tag{31a}$$

$$\mathbf{p}(t) = -\mathbf{A}^{\frac{1}{2}} \sin \mathbf{A}^{\frac{1}{2}}t\mathbf{q}(0) + \cos \mathbf{A}^{\frac{1}{2}}t\mathbf{p}(0) + \int_0^t dt' \cos \mathbf{A}^{\frac{1}{2}}(t - t')\mathbf{F}(t'). \tag{31b}$$

If, now, we take the zeroth element of Eq. (30) for $\dot{\mathbf{p}}$ and substitute (31a) we get

$$\dot{p}_0 = -\sum_i ||\mathbf{A} \cos \mathbf{A}^{\frac{1}{2}}t||_{0i}\, q_i(0) - \sum_i ||\mathbf{A}^{\frac{1}{2}} \sin \mathbf{A}^{\frac{1}{2}}t||_{0i}\, p_i(0) - \int_0^t dt' \, ||\mathbf{A}^{\frac{1}{2}} \sin \mathbf{A}^{\frac{1}{2}}(t - t')||_{00} F(t') + F(t).$$

Next, we eliminate $p_0(0)$ between this equation and the zeroth element of Eq. (31b). The result can be written in the form

$$\dot{p}_0 - F(t) = -\gamma(t)p_0 + E(t) + \int_0^t dt' \, [\gamma(t) - \gamma(t - t')] \times ||\cos \mathbf{A}^{\frac{1}{2}}(t - t')||_{00} F(t'), \tag{32}$$

where

$$\gamma(t) = \frac{||\mathbf{A}^{\frac{1}{2}} \sin \mathbf{A}^{\frac{1}{2}}t||_{00}}{||\cos \mathbf{A}^{\frac{1}{2}}t||_{00}} = -\frac{d}{dt} \log ||\cos \mathbf{A}^{\frac{1}{2}}t||_{00}, \tag{33}$$

and

$$E(t) = -\sum_i \{\gamma(t) \, ||\mathbf{A}^{\frac{1}{2}} \sin \mathbf{A}^{\frac{1}{2}}t||_{0i} + ||\mathbf{A} \cos \mathbf{A}^{\frac{1}{2}}t||_{0i}\} q_i(0) + \sum_i \{\gamma(t) \, ||\cos \mathbf{A}^{\frac{1}{2}}t||_{0i} - ||\mathbf{A}^{\frac{1}{2}} \sin \mathbf{A}^{\frac{1}{2}}t||_{0i}\} p_i(0). \tag{34}$$

Note that the coefficient of $p_0(0)$ vanishes in this expression. Equation (32) is the equation of motion for the Brownian particle. The right-hand side is the net force exerted on the Brownian particle by the other particles, i.e., by the heat bath. The first term represents a frictional force with time-dependent "friction coefficient" $\gamma(t)$, the second term represents a fluctuating force $E(t)$ depending upon the initial state of the heat bath, and the third

[7] What is unique is the *spectrum* of eigenfrequencies: $g(\omega) = 2\omega/(\pi f'(\theta))$, where θ is the function of ω obtained by inverting the equation $\omega^2 = f(\theta)$. For (25), $g(\omega) = (2f/\pi) \cdot (\omega^2 + f^2)^{-1}$.

term represents a memory effect depending upon the past history of the motion of the Brownian particle.

If we assume that the interaction between the Brownian particle and the heat bath is invariant under translations, we have

$$\sum_j A_{ij} = 0, \tag{35}$$

and this in turn implies that (34) may be written

$$E(t) = -\sum_j{}' \{\gamma(t)\, ||\mathbf{A}^{\frac{1}{2}} \sin \mathbf{A}^{\frac{1}{2}}t||_{0j} + ||\mathbf{A} \cos \mathbf{A}^{\frac{1}{2}}t||\}[q_j(0) - q_0(0)] + \sum_j{}' \{\gamma(t)\, ||\cos \mathbf{A}^{\frac{1}{2}}t||_{0j} - ||\mathbf{A}^{\frac{1}{2}} \sin \mathbf{A}^{\frac{1}{2}}t||_{0j}\}p_j(0), \tag{36}$$

where the prime on the sum denotes the omission of the term $j = 0$. Thus the fluctuating force depends only upon the initial coordinates of the particles of the heat bath *relative* to the initial coordinate of the Brownian particle, and is independent of the initial coordinate and momentum of the Brownian particle.

Consider now what happens when the matrix of interactions is that of the model discussed in Sec. 3, in which the matrix elements are given by (19) with $f(\theta)$ given by (26) in the limit $\omega_L \gg f$. For this model

$$||\cos \mathbf{A}^{\frac{1}{2}}t||_{00} = e^{-f|t|} \tag{37}$$

and therefore, from (33) we find that

$$\lim \gamma(t) = f, \tag{38}$$

which is a constant.

This in turn implies that the last term on the right-hand side of (32) (the memory-effect term) becomes, in the limit, identically 0!

With these results, (32) takes the form

$$\dot{p}_0 - F(t) = -fp_0 + E(t), \tag{39}$$

with

$$E(t) = -\sum_j ||f\mathbf{A}^{\frac{1}{2}} \sin \mathbf{A}^{\frac{1}{2}}t + \mathbf{A} \cos \mathbf{A}^{\frac{1}{2}}t||_{0j}q_j(0) + \sum_j ||f \cos \mathbf{A}^{\frac{1}{2}}t - \mathbf{A}^{\frac{1}{2}} \sin \mathbf{A}^{\frac{1}{2}}t||_{0j}p_j(0). \tag{40}$$

Equation (39) is the Langevin equation.

It remains to prove that the statistical properties of $E(t)$ become (again in the limit $N \to \infty$, $\omega_L \gg f$) those of a purely random Gaussian process. This depends, of course on the statistical assumptions concerning initial positions and momenta.

We would like to require that at $t = 0$ the heat bath is in equilibrium at temperature T and the simplest way of doing this is to assume that the initial distribution is the canonical distribution (6). This however is, strictly speaking, impossible since (35) implies that $\omega_0 = 0$ so that $\det \mathbf{A} = 0$ and the canonical distribution becomes improper. The difficulty is not serious and can be remedied, e.g., by slightly modifying the matrix $\mathbf{A}$; i.e., replacing it by ($\mathbf{I}$ is the unit matrix)

$$\mathbf{A} + \varepsilon_N \mathbf{I}, \tag{41}$$

where ε_N, though positive for every finite N, approaches 0 as $N \to \infty$ (as fast as one pleases). Now (35) is only approximately true and the canonical distribution (6) is proper. We hope that the reader does not become unduly confused by our use of the symbol $\mathbf{A}$ to denote three different matrices. We use it to denote the finite cyclic matrix (12), the modification (41), and last but not least for the infinite cyclic matrix (defined only with the cutoff ω_L). Clearly, since the distribution (6) of the $q_j(0)$ and $p_j(0)$ is Gaussian, $E(t)$ is a Gaussian process. We can form its covariance with the help of the results of Appendix 1, and, in the limit considered throughout this paper, we find that

$$\langle E(t)E(t')\rangle = kT\, ||(f^2 + \mathbf{A}) \cos \mathbf{A}^{\frac{1}{2}}(t - t')||_{00}. \tag{42}$$

But the matrix $\mathbf{A}$ is given by the model of Sec. 3; hence, using (21), we find

$$\langle E(t)E(t')\rangle = \frac{kT}{2\pi}\int_{-\pi}^{\pi} d\theta \left(f^2 + f^2 \tan^2 \frac{\theta}{2}\right) \cos f \tan \frac{\theta}{2}(t - t') = \frac{kTf}{\pi}\int_{-\infty}^{\infty} d\omega \cos \omega(t - t').$$

This last integral is the well known expression for the Dirac delta function, so we have

$$\langle E(t)E(t')\rangle = 2fkT\delta(t - t'). \tag{43}$$

Thus $E(t)$ is a purely random, Gaussian, stochastic process and Eq. (39) is the Langevin equation for Brownian motion.

In order that the equation of motion (32) become the Langevin equation it is necessary that (i) the friction constant $\gamma(t)$ be independent of time; (ii) the stochastic process $E(t)$ be a purely random Gaussian process; (iii) the memory effects disappear.

We feel that it is striking that, for our model, these three properties are intimately related. Undoubtedly this fact is of much more general significance.

5. THE MOTION OF COUPLED QUANTUM OSCILLATORS

We turn now to the question of the changes required for a quantum mechanical description of

the motion of coupled oscillators. The answer is that much of our previous discussion is, formally at any rate, entirely unchanged. Thus in our discussion of Sec. 2, the Hamiltonian (2) is unchanged, but the coordinates q_i and momenta p_i are now operators whose commutation rules are

$$[q_i, q_k] = [p_i, p_k] = 0, \qquad (44)$$
$$[q_i, p_k] = i\hbar\delta_{ik}.$$

The equations of motion (3) are the equations of motion in the Heisenberg picture, and their solution (4) relates the Heisenberg operators at time t to the operators at the initial time.

At $t = 0$, we assume the system is in equilibrium at temperature T. In the quantum description this means that the initial state of the system is described by the density matrix corresponding to the canonical ensemble:

$$\rho(\mathbf{q}(0), \mathbf{p}(0)) = \exp\{-\beta H(\mathbf{q}(0), \mathbf{p}(0))\}. \qquad (45)$$

The expectation of any function $F(\mathbf{q}(0), \mathbf{p}(0))$ of the operators $\mathbf{q}(0)$, $\mathbf{p}(0)$ given by

$$\langle F\rangle = \frac{\mathrm{Tr}\,\{F(\mathbf{q}(0), \mathbf{p}(0))\rho(\mathbf{q}(0), \mathbf{p}(0))\}}{\mathrm{Tr}\,\{\rho(\mathbf{q}(0), \mathbf{p}(0))\}}. \qquad (46)$$

Just as in the classical case, we now consider the properties of the *stochastic operators* $q_i(t)$ and $p_i(t)$ which result from the equation of motion and the initial density matrix. These properties we describe in terms of the correlation functions, the simplest being the pair correlation functions. These are obtained in Appendix 2; the results are

$$\langle p_i(t)p_k(t+\tau)\rangle = \left\|\frac{\hbar\mathbf{A}^{\frac{1}{2}}}{2}\left[\coth\frac{\hbar\mathbf{A}^{\frac{1}{2}}}{2kT}\cos\mathbf{A}^{\frac{1}{2}}\tau + i\sin\mathbf{A}^{\frac{1}{2}}\tau\right]\right\|_{ik} \qquad (47a)$$

$$\langle q_i(t)q_k(t+\tau)\rangle = \left\|\frac{\hbar}{2\mathbf{A}^{\frac{1}{2}}}\left[\coth\frac{\hbar\mathbf{A}^{\frac{1}{2}}}{2kT}\cos\mathbf{A}^{\frac{1}{2}}\tau + i\sin\mathbf{A}^{\frac{1}{2}}\tau\right]\right\|_{ik} \qquad (47b)$$

$$\langle q_i(t)p_k(t+\tau)\rangle = \left\|\frac{\hbar}{2}\left[-\coth\frac{\hbar\mathbf{A}^{\frac{1}{2}}}{2kT}\sin\mathbf{A}^{\frac{1}{2}}\tau + i\cos\mathbf{A}^{\frac{1}{2}}\tau\right]\right\|_{ik}. \qquad (47c)$$

Note that in the limit $\hbar = 0$, these expressions become identical with the classical expressions (9). As indicated in Appendix 2, the higher correlations are given by the rule:

Correlations of an odd number of q's and p's vanish. The correlation of an even number of q's and p's is equal to the sum of products of pair correlations, the sum being over all possible pairings of the operators, *with order preserved.*

Except for the italicized proviso that the order be preserved, this rule is identical with the classical rule for a Gaussian random process.[8] The stochastic operators also have the classical *stationarity property*; the correlations depend only upon the time differences. There are, however, obvious differences between the properties of the stochastic operators and the properties of the corresponding classical stochastic process. The principal differences arise from the fact that the quantum operators do not commute. Thus there are many (indeed, an infinity!) of correlations of the quantum operators which can be associated with a given classical correlation, corresponding to different orderings of the operators. As a simple example,

$$\langle \epsilon qp + (1-\epsilon)pq\rangle_{\text{quantum}} \to \langle qp\rangle_{\text{classical}},$$

where ϵ is an arbitrary complex number. Another difficulty comes from the fact that the product of two noncommuting Hermitian operators is not Hermitian, and, therefore, does not correspond to a physical observable. We see this difficulty explicitly in the correlations (47), which are complex functions, whereas the expectation value of a physical observable should be real.

The difficulties we mention would be largely resolved if we had a conventional definition of the product of operators with the following properties:

(a) The product is independent of the order of the operators.

(b) The product of a number of Hermitian operators is itself Hermitian.

(c) The classical pair decomposition rule for expressing higher correlations in terms of pair correlations holds.

For our system of coupled oscillators there is such a conventional product, namely the ordered product or *normal product* introduced in quantum field theory[9]; it is defined as follows. In Appendix 2 we show how the operators $\mathbf{q}(0)$ and $\mathbf{p}(0)$ can be expanded in terms of creation and annihilation operators for the normal modes;

$$\mathbf{q}(0) = i\sum_s \xi^{(s)}\left(\frac{\hbar}{2\omega_s}\right)^{\frac{1}{2}}(a_s - a_s^*),$$
$$\mathbf{p}(0) = \sum_s \xi^{(s)}\left(\frac{\hbar\omega_s}{2}\right)^{\frac{1}{2}}(a_s + a_s^*). \qquad (48)$$

[8] See, e.g., Wang and Uhlenbeck, Ref. 2.
[9] See, e.g., G.-C. Wick, Phys. Rev. 80, 268 (1950).

Here $\xi^{(s)}$ is the eigenvector of the interaction matrix $\mathbf{A}$, and ω_s^2 is the associated eigenvalue

$$\mathbf{A}\xi^{(s)} = \omega_s^2 \xi^{(s)}. \tag{49}$$

The operator a_s is the annihilation operator for the sth normal mode and a_s^* is the corresponding creation operator. Their commutation rules are

$$[a_s, a_r^*] = \delta_{sr}, \quad [a_s, a_r] = [a_s^*, a_r^*] = 0. \tag{50}$$

The time-dependent operators are expressed in terms of the a_s and a_s^* by inserting (48) in (4). We find

$$\begin{aligned} \mathbf{q}(t) &= i \sum_s \xi^{(s)} \left(\frac{\hbar}{2\omega_s}\right)^{\frac{1}{2}} (a_s e^{-i\omega_s t} - a_s^* e^{i\omega_s t}), \\ \mathbf{p}(t) &= \sum_s \xi^{(s)} \left(\frac{\hbar\omega_s}{2}\right)^{\frac{1}{2}} (a_s e^{-i\omega_s t} + a_s^* e^{i\omega_s t}). \end{aligned} \tag{51}$$

The normal product of a number of the operators a_s and a_s^* is defined to be that product in which all the a_s^* are written to the left of all the a_s. Because of the commutation rules (50), this defines a unique order. The normal product of a number of operators $q_i(t)$ and $p_i(t)$ is the product in which the expansions (51) are used and each product of the a_s and a_s^* is written in normal form. We denote the normal product by a colon placed before and after the product of factors. As an example,

$$\begin{aligned} :q_i(t_1)p_k(t_2): \equiv i\frac{\hbar}{2} \sum_s \sum_r \left(\frac{\omega_r}{\omega_s}\right)^{\frac{1}{2}} \xi_i^{(s)} \xi_k^{(r)} [a_s a_r e^{-i(\omega_s t_1 + \omega_r t_2)} \\ - a_s^* a_r e^{i(\omega_s t_1 - \omega_r t_2)} + a_r^* a_s e^{-i(\omega_s t_1 - \omega_r t_2)} \\ - a_s^* a_r^* e^{i(\omega_s t_1 + \omega_r t_2)}]. \end{aligned}$$

The normal product fulfills our requirements. It us clearly independent of the order of the factors, and, since a_s^* is the Hermitian conjugate of a_s, the normal product of a number of Hermitian operators is Hermitian. Using the results given in Appendix 2, we can show that the pair correlations of normal products of the $q_i(t)$ and $p_k(t)$ are

$$\begin{aligned} \langle :p_i(t)p_k(t+\tau): \rangle &= ||P(\hbar\mathbf{A}^{\frac{1}{2}}/kT) \cos \mathbf{A}^{\frac{1}{2}}\tau||_{ik}, \\ \langle :q_i(t)q_k(t+\tau): \rangle &= ||P(\hbar\mathbf{A}^{\frac{1}{2}}/kT)\mathbf{A}^{-1} \cos \mathbf{A}^{\frac{1}{2}}\tau||_{ik}, \\ \langle :q_i(t)p_k(t+\tau): \rangle &= -||P(\hbar\mathbf{A}^{\frac{1}{2}}/kT)\mathbf{A}^{-\frac{1}{2}} \sin \mathbf{A}^{\frac{1}{2}}\tau||_{ik}. \end{aligned} \tag{52}$$

Here we have introduced the Planck function

$$P(x) \equiv kTx/(e^x - 1). \tag{53}$$

When $x = \hbar\omega/kT$ the Planck function is the average energy, relative to the ground state, of a quantum oscillator of frequency ω. As $x \to 0$, $P(x) \to kT$, the classical equipartition energy.

The higher correlations of normal products are given by the well-known rule for a Gaussian random process:

The correlation of the normal product of an odd number of q's and p's vanishes. The correlation of the normal product of an even number of q's and p's is equal to the sum of products of pair correlation of normal products, the sum being over all possible pairings.

What we can say, then, is that the correlations of normal products of our stochastic operators are identical with those of a *stationary Gaussian process* whose pair correlations are given by (52).[10]

Finally, if we fix our attention on a single oscillator, the one with index 0, the momentum autocorrelation is

$$\langle :p_0(t)p_0(t+\tau): \rangle = ||P(\hbar\mathbf{A}^{\frac{1}{2}}/kT) \cos \mathbf{A}^{\frac{1}{2}}\tau||_{00}. \tag{54}$$

Just as in the corresponding classical process, we can ask whether there is an interaction matrix $\mathbf{A}$ for which (54) is an exponential and, therefore, the corresponding Gaussian process is Markoffian. The answer is that we can, but that it is temperature-dependent, and, therefore, not of physical interest. If we use the model discussed in Sec. 3, in which the matrix elements are given by (19) with $f(\theta)$ given by (26) in the limit $\omega_L \gg f$, we find (54) becomes

$$\begin{aligned} &\langle :p_0(t)p_0(t+\tau): \rangle \\ &\quad = \frac{2}{\pi} \int_0^\infty d\omega\, P\left(\frac{\hbar\omega}{kT}\right) \frac{f}{\omega^2 + f^2} \cos \omega\tau. \end{aligned} \tag{55}$$

In the limit $\hbar \to 0$, this becomes identical with the classical result obtained earlier, and the process becomes Markoffian.

6. THE QUANTUM LANGEVIN EQUATION

The formal manipulations we used in Sec. 4 to derive the Langevin equation are unchanged when we interpret the q's and p's as quantum operators. In particular, when the matrix of interactions is that given by the model of Sec. 3, in which the matrix elements are given by (19) with $f(\theta)$ given by (26) in the limit $\omega_L \gg f$, we obtain an operator equation of motion which is formally identical with Langevin equation (39). That is

$$\dot{p}_0 - F(t) = -fp_0 + E(t), \tag{39'}$$

where

$$\begin{aligned} E(t) = &- \sum_j ||f\mathbf{A}^{\frac{1}{2}} \sin \mathbf{A}^{\frac{1}{2}}t + \mathbf{A} \cos \mathbf{A}^{\frac{1}{2}}t||_{0j}\, q_j(0) \\ &+ \sum_j ||f \cos \mathbf{A}^{\frac{1}{2}}t - \mathbf{A}^{\frac{1}{2}} \sin \mathbf{A}^{\frac{1}{2}}t||_{0j}\, p_j(0). \end{aligned} \tag{40'}$$

[10] Essentially the same point is made in connection with the quantum description of statistical light beams by E. C. G. Sudarshan, Phys. Rev. Letters **10**, 277 (1963). See also R. J. Glauber, Phys. Rev. Letters **10**, 84 (1963).

The operator Langevin equation is an equation of motion for the time-dependent Heisenberg operators $p_0(t)$ and $q_0(t)$. The operator $F(t)$ is the external force operator,

$$F(t) \equiv (i\hbar)^{-1}[p_0(t),\ V(q_0(t),\ t)], \tag{56}$$

with $V(q_0,\ t)$ the (time-dependent) potential of the external force. The random-force operator $E(t)$ is in fact independent of the operators $p_0(0)$ and $q_0(0)$ since their coefficients in the expression (40′) vanish for our model of the interaction matrix. Because of the commutation rules (44), this means that

$$[q_0(0), E(t)] = [p_0(0), E(t)] = 0. \tag{57}$$

We assume that the statistical state of the initial coordinates and momenta of the heat bath is described by the density matrix (45). Just as in the classical case, this means that at $t = 0$ the heat bath (i.e., the oscillators other than the Brownian particle) is in equilibrium with a fictitious force-free Brownian particle. Since the random-force operator $E(t)$ is independent of the initial coordinate and momentum of the Brownian particle, its stochastic properties are unaffected by the dependence of the density matrix upon these operators. The covariance of the normal product is readily obtained using the results (52), we find

$$\langle :E(t)E(t+\tau):\rangle = ||(f^2 + \mathbf{A})P(\hbar\mathbf{A}^{\frac{1}{2}}/kT)\cos \mathbf{A}^{\frac{1}{2}}\tau||_{00}. \tag{58}$$

But for our model, we may use the general expression (21), with $f(\theta)$ given by (25). Hence

$$\begin{aligned}\langle :E(t)E(t+\tau):\rangle &= \frac{1}{2\pi}\int_{-\pi}^{\pi} d\theta\, f^2(1 + \tan^2 \tfrac{1}{2}\theta) \\ &\quad \times P\left(\frac{\hbar f}{kT}\,|\tan \tfrac{1}{2}\theta|\right)\cos (f\tau \tan \tfrac{1}{2}\theta) \\ &= \frac{2f}{\pi}\int_0^{\infty} d\omega\, P\left(\frac{\hbar\omega}{kT}\right)\cos \omega\tau.\end{aligned} \tag{59}$$

Correlations of higher normal products of $E(t)$ are again given by the rule for a Gaussian random process. Hence, the stochastic properties of the random operator $E(t)$, as expressed by the correlations of normal products, are identical with those of a stationary Gaussian process whose covariance is given by (59). In the limit $\hbar \to 0$, this covariance approaches the corresponding classical covariance (43), which is the covariance of a purely random Gaussian process. However, for finite $\hbar$, (59) is the covariance of a Gaussian process which is not even Markoffian. This is the chief difference between the quantum and classical Langevin equations for our model.

As an elementary application, consider the motion in a constant field of force,

$$F(t) = \mathcal{E}. \tag{60}$$

The solution of the operator Langevin equation is

$$p_0(t) = e^{-ft}p_0(0) + f^{-1}\mathcal{E}(1 - e^{-ft}) + \int_0^t dt'\, e^{-f(t-t')}E(t'). \tag{61}$$

If we average this expression over the initial state of the heat bath, we find[11]

$$\langle p_0(t)\rangle = e^{-ft}p_0(0) + f^{-1}\mathcal{E}(1 - e^{-ft}), \tag{62}$$

since $\langle E(t)\rangle = 0$. After a long time ($t \gg f^{-1}$) we find

$$\langle p_0(t)\rangle \sim f^{-1}\mathcal{E}. \tag{63}$$

This is the analog for our model of Ohm's law. The left-hand side is the "current," which in the steady state is proportional to the applied field, and inversely proportional to the "resistance," i.e., the friction constant.

The mean-square "fluctuation current" may also be obtained from (60)

$$\langle :[p_0(t) - \langle p_0(t)\rangle]^2:\rangle = \int_0^t dt'\, e^{-f(t-t')}\int_0^t dt''\, e^{-f(t-t'')}\langle :E(t')E(t''):\rangle. \tag{64}$$

Note that we have used the normal product in the definition of the squared fluctuation. Using the expression (58) for the covariance of $E(t)$, we find

$$\langle :[p_0(t) - \langle p_0(t)\rangle]^2:\rangle = \frac{2f}{\pi}\int_0^{\infty} d\omega\, P\left(\frac{\hbar\omega}{kT}\right)\left|\frac{1 - e^{-(f-i\omega)t}}{f - i\omega}\right|^2, \tag{65}$$

where we have used the elementary result

$$\int_0^t dt'\int_0^t dt''\, e^{-f(t-t')}e^{-f(t-t'')}\cos \omega(t' - t'') = |(1 - e^{-(f-i\omega)t})/(f - i\omega)|^2. \tag{66}$$

In the steady state, i.e., when $t \gg f^{-1}$, (65) becomes

$$\langle :[p_0(t) - \langle p_0(t)\rangle]^2:\rangle = \frac{2f}{\pi}\int_0^{\infty} d\omega\, P\left(\frac{\hbar\omega}{kT}\right)\frac{1}{\omega^2 + f^2}. \tag{67}$$

This is the analog for our model of the Nyquist formula, which relates the noise power spectrum

[11] The result of the average over the states of the heat bath is still in general an operator function of $q_0(0)$ and $p_0(0)$. We trust that our definition of this average, which involves an average over the initial coordinate and momentum of a fictitious Brownian particle, is not confusing.

to the resistance and the absolute temperature.[12] Note that if we had used the ordinary product instead of the normal product, we would find

$$\begin{aligned}&\langle[p_0(t) - \langle p_0(t)\rangle]^2\rangle \\ &\equiv \langle[p_0(t) - \langle p_0(t)\rangle]\cdot[p_0(t) - \langle p_0(t)\rangle]\rangle \\ &= \langle:[p_0(t) - \langle p_0(t)\rangle]^2:\rangle + \frac{2f}{\pi}\int_0^{\omega_L} d\omega\,\frac{\hbar\omega}{2}\,\frac{1}{\omega^2 + f^2}.\end{aligned} \tag{68}$$

The added integral is clearly related to the zero-point fluctuations of the heat bath. It is also divergent in the limit $\omega_L \to \infty$, so we have retained this cutoff explicitly. Which of the possible definitions of the product corresponds to an experimental measurement of the fluctuation spectrum? This, in the last analysis, must be determined by the experiment itself. We tend to the opinion that the normal product is the physically appropriate definition, since it leads to a noise spectrum which vanishes at absolute zero.

7. BROWNIAN MOTION OF A QUANTUM OSCILLATOR

As a second application of the quantum Langevin equation, we consider the case of the harmonic oscillator, for which the external force is[13]

$$F(t) = -\kappa^2 q_0(t), \tag{69}$$

where κ is the natural frequency of the oscillator. The quantum Langevin equation (39′) becomes

$$\dot{p}_0 + \kappa^2 q_0 = -f p_0 + E(t), \tag{70}$$

to which we must append the equation

$$\dot{q}_0 = p_0. \tag{71}$$

The solution of this pair of coupled equations is

$$\begin{aligned}q_0(t) &= e^{-\frac{1}{2}ft}\{[\cos\nu t + (f/2\nu)\sin\nu t]q_0(0) \\ &\quad + \nu^{-1}\sin\nu t p_0(0)\} \\ &\quad + \int_0^t dt'\, e^{-\frac{1}{2}f(t-t')}\,\frac{1}{\nu}\sin\nu(t-t')E(t'), \\ p_0(t) &= e^{-\frac{1}{2}ft}\{-(\kappa^2/\nu)\sin\nu t q_0(0) \\ &\quad + [\cos\nu t - (f/2\nu)\sin\nu t]p_0(0)\} \\ &\quad + \int_0^t dt'\, e^{-\frac{1}{2}(t-t')}[\cos\nu(t-t') \\ &\quad - (f/2\nu)\sin\nu(t-t')]E(t').\end{aligned} \tag{72}$$

Here

$$\nu = (\kappa^2 - \tfrac{1}{4}f^2)^{\frac{1}{2}}, \tag{73}$$

and we restrict our discussion to the underdamped case, where $\kappa > \frac{1}{2}f$, so ν is real and positive.

The mean motion of the oscillator is described by the operators obtained by averaging (72) and (73) over the initial states of the heat bath. Since $\langle E(t)\rangle = 0$, we have

$$\begin{aligned}\langle q_0(t)\rangle &= e^{-\frac{1}{2}ft}\{[\cos\nu t + (f/2\nu)\sin\nu t]q_0(0) \\ &\quad + \nu^{-1}\sin\nu t p_0(0)\}, \\ \langle p_0(t)\rangle &= e^{-\frac{1}{2}ft}\{-\kappa^2\nu^{-1}\sin\nu t q_0(0) \\ &\quad + [\cos\nu t - (f/2\nu)\sin\nu t]p_0(0)\}.\end{aligned} \tag{74}$$

These are just the operator solutions of the average of the equations of motion (70) and (71). We see from (74) that these operators vanish for $t \gg f^{-1}$; the mean motion of the oscillator vanishes for times long compared with the "macroscopic relaxation time" f^{-1}.

To see more precisely what we mean by the operators (74), consider the coordinate representation. The Heisenberg state of the oscillator is then specified by a time-independent wavefunction $\psi[q_0(0)]$, and the initial momentum operator is represented by $-i\hbar\,\partial/\partial q_0(0)$. The simplest kind of question we can ask about the Heisenberg operators is their *expectation value*, denoted by a subscript "ex" to the operator. For example,

$$\begin{aligned}\langle q_0(t)\rangle_{ex} &\equiv \int_{-\infty}^{\infty} dq_0(0)\;\psi^*[q_0(0)]\langle q_0(t)\rangle\psi[q_0(0)] \\ &= e^{-\frac{1}{2}ft}\{(\cos\nu t + (f/2\nu)\sin\nu t)[q_0(0)]_{ex} \\ &\quad + \nu^{-1}\sin\nu t[p_0(0)]_{ex}\}.\end{aligned} \tag{75}$$

As a simple illustration, if the wavefunction is

$$\psi[q_0(0)] = (\alpha/\pi)^{\frac{1}{4}}e^{-\frac{1}{2}\alpha[q_0(0)-x_0]^2}, \tag{76}$$

then

$$[q_0(0)]_{ex} = x_0, \qquad [p_0(0)]_{ex} = 0, \tag{77}$$

and

$$\begin{aligned}\langle q_0(t)\rangle_{ex} &= e^{-\frac{1}{2}ft}[\cos\nu t + (f/2\nu)\sin\nu t]x_0, \\ \langle p_0(t)\rangle_{ex} &= -e^{-\frac{1}{2}ft}\kappa^2\nu^{-1}\sin\nu t x_0.\end{aligned} \tag{78}$$

The operator describing the fluctuation of the displacement of the oscillator about the mean displacement is

$$\begin{aligned}&q_0(t) - \langle q_0(t)\rangle \\ &\quad = \int_0^t dt'\, e^{-\frac{1}{2}f(t-t')}\,\frac{\sin\nu(t-t')}{\nu}\,E(t').\end{aligned} \tag{79}$$

[12] For elementary discussion of the Nyquist formula and its quantum generalization see C. Kittel, *Elementary Statistical Physics* (John Wiley & Sons, Inc., New York, 1958), pp. 141–153. See also J. Lawson and G. E. Uhlenbeck, *Threshold Noise Signals* (McGraw-Hill Book Company, Inc., New York, 1950), especially pp. 64–79.

[13] The Brownian motion of a quantum oscillator is considered in a paper by J. Schwinger, J. Math. Phys. **2**, 407 (1961).

The mean-square fluctuation in displacement we express as the mean of the square of (79), written as a normal product. Using (59), we find

$$\langle :[q_0(t) - \langle q_0(t)\rangle]^2:\rangle = \frac{2f}{\pi}\int_0^\infty d\omega\, P\left(\frac{\hbar\omega}{kT}\right) \times \left|\frac{1 - e^{-(\frac{1}{2}f - i\omega)t}[\cos\nu t + (\frac{1}{2}f - i\omega)\nu^{-1}\sin\nu t]}{\omega^2 - \kappa^2 + if\omega}\right|^2. \quad (80)$$

Here we have used the following result of an elementary but tedious integration

$$\int_0^t dt' \int_0^t dt''\, e^{-\frac{1}{2}f(t-t')}\frac{\sin\nu(t-t')}{\nu} \times e^{-\frac{1}{2}f(t-t'')}\frac{\sin\nu(t-t'')}{\nu}\cos\omega(t'-t'') = \left|\frac{1 - e^{-(\frac{1}{2}f - i\omega)t}[\cos\nu t + (\frac{1}{2}f - i\omega)\nu^{-1}\sin\nu t]}{\omega^2 - \kappa^2 + if\omega}\right|^2. \quad (81)$$

The time dependence of the mean fluctuation in displacement, as expressed by (80), is rather complicated. It does, however, have the simple feature that, for times long compared with the relaxation time f^{-1}, it approaches an equilibrium value

$$\langle :[q_0(t) - \langle q_0(t)\rangle]^2:\rangle_{eq} = \frac{2f}{\pi}\int_0^\infty d\omega\, \frac{P(\hbar\omega/kT)}{(\omega^2 - \kappa^2)^2 + \omega^2 f^2}. \quad (82)$$

In the classical limit we use the property of the Planck function (53)

$$P(\hbar\omega/kT) \to kT, \quad \text{as} \quad \hbar \to 0. \quad (83)$$

The remaining integral in (82) is elementary:

$$\frac{2f}{\pi}\int_0^\infty d\omega\, \frac{1}{(\omega^2 - \kappa^2)^2 + \omega^2 f^2} = \kappa^{-2}, \quad (84)$$

and we find

$$\langle :[q_0(t) - \langle q_0(t)\rangle]^2:\rangle_{eq} \to kT/\kappa^2, \quad \text{as} \quad \hbar \to 0, \quad (85)$$

which is the classical equipartition result.

Another simple limit of (82) is the weak-coupling limit, in which the coupling of the Brownian particle to the heat bath is weak compared with the oscillator coupling. That is, $\kappa \gg f$. In this limit the resonance denominator in (82) becomes sharply peaked at $\omega = \kappa$, with a width $\approx f$. Hence, we can evaluate the Planck function at $\omega = \kappa$ and perform the remaining integral using (84). We find

$$\langle :[q_0(t) - \langle q_0(t)\rangle]^2:\rangle_{eq} \to \frac{1}{\kappa^2}P\left(\frac{\hbar\kappa}{kT}\right), \quad \text{for} \quad \kappa \gg f. \quad (86)$$

This is the well-known Planck result for the mean-square displacement of an independent oscillator of frequency κ at temperature T.

APPENDIX 1: CORRELATIONS IN A SYSTEM OF COUPLED CLASSICAL OSCILLATORS

Consider first the correlations of the initial values of the coordinates and momenta, whose distribution is [c.f. Eq. (6)]

$$D(\mathbf{q}(0), \mathbf{p}(0)) = (2\pi/\beta)^{2N+1}(\det \mathbf{A})^{-\frac{1}{2}} \times \exp\left\{-\frac{\beta}{2}\left[\sum_i p_i^2(0) + \sum_{j,k} q_j(0)A_{jk}q_k(0)\right]\right\}. \quad (1)$$

Since this is a Gaussian distribution, all higher correlations can be expressed in terms of the pair correlations. These are ($\beta = 1/kT$):

$$\begin{aligned}\langle p_j(0)p_k(0)\rangle &= kT\delta_{jk},\\ \langle p_j(0)q_k(0)\rangle &= 0,\\ \langle q_j(0)q_k(0)\rangle &= kT\,||\mathbf{A}^{-1}||_{jk}.\end{aligned} \quad (2)$$

These expressions are consequences of well-known integral formulas for Gaussian distributions.[14]

The pair correlations for the time-dependent coordinates and momenta are found from (4), using (2). We have for the momentum correlation:

$$\begin{aligned}\langle p_j(t)p_k(t+\tau)\rangle &= \sum_{m,n}\{||\mathbf{A}^{\frac{1}{2}}\sin\mathbf{A}^{\frac{1}{2}}t||_{jm} \times ||\mathbf{A}^{\frac{1}{2}}\sin\mathbf{A}^{\frac{1}{2}}(t+\tau)||_{kn}\langle q_m(0)q_n(0)\rangle\\ &\quad + ||\cos\mathbf{A}^{\frac{1}{2}}t||_{jm}\,||\cos\mathbf{A}^{\frac{1}{2}}(t+\tau)||_{kn}\langle p_m(0)p_n(0)\rangle\}\\ &= kT\{||\sin\mathbf{A}^{\frac{1}{2}}t\sin\mathbf{A}^{\frac{1}{2}}(t+\tau)\\ &\quad + \cos\mathbf{A}^{\frac{1}{2}}t\cos\mathbf{A}^{\frac{1}{2}}(t+\tau)||_{jk},\end{aligned}$$

where we have used the fact that the matrix $\mathbf{A}$ is symmetric. Using the formula for the cosine of the difference of two angles, we find

$$\langle p_j(t)p_k(t+\tau)\rangle = kT\,||\cos\mathbf{A}^{\frac{1}{2}}\tau||_{jk}. \quad (3)$$

In a similar way, we can also show

$$\langle q_j(t)p_k(t+\tau)\rangle = -kT\,||\mathbf{A}^{-\frac{1}{2}}\sin\mathbf{A}^{\frac{1}{2}}\tau||_{jk}, \quad (4)$$

$$\langle q_j(t)q_k(t+\tau)\rangle = kT\,||\mathbf{A}^{-1}\cos\mathbf{A}^{\frac{1}{2}}\tau||_{jk}. \quad (5)$$

Since the process is Gaussian, the higher correlations are given by the following rule[8]:

The correlation of an odd number of coordinates and momenta vanish. The correlation of an even number of coordinates and momenta is equal to the

[14] See, e.g., H. Cramer, *Mathematical Methods of Statistics* (Princeton University Press, Princeton, New Jersey, 1945), p. 118.

sum of products of pair correlations, the sum being over all pairings.

For example,

$$\langle q_i(t_1)q_k(t_2)q_l(t_3)p_m(t_4)\rangle = \langle q_i(t_1)q_k(t_2)\rangle\langle q_l(t_3)p_m(t_4)\rangle + \langle q_i(t_1)q_l(t_3)\rangle\langle q_k(t_2)p_m(t_4)\rangle + \langle q_i(t_1)p_m(t_4)\rangle\langle q_k(t_2)q_l(t_3)\rangle.$$

APPENDIX 2: CORRELATIONS IN A SYSTEM OF COUPLED QUANTUM OSCILLATORS

The system of oscillators is described by the Hamiltonian

$$H = \frac{1}{2}\sum_i p_i^2 + \frac{1}{2}\sum_{i,k} A_{ik}q_iq_k, \tag{1}$$

where the operators q_i and p_i satisfy the commutation relations

$$[q_i, p_k] = i\hbar\delta_{ik}, \qquad [q_i, q_k] = [p_i, p_k] = 0. \tag{2}$$

The expectation value of any operator F with respect to the canonical ensemble at temperature T is defined by

$$\langle F\rangle = \mathrm{Tr}\,\{Fe^{-\beta H}\}/\mathrm{Tr}\,\{e^{-\beta H}\}, \tag{3}$$

where the trace operation is in the space of the eigenfunctions of the Hamiltonian operator. We are interested mainly in the case where the operator F is a product of q's and p's.

As a first step in the evaluation of such traces, consider the eigenvalues and eigenvectors of the matrix **A**:

$$\sum_k A_{ik}\xi_k^{(s)} = \omega_s^2\xi_i^{(s)}. \tag{4}$$

The eigenvectors are assumed normalized so that

$$\sum_i \xi_i^{(s)}\xi_i^{(r)} = \delta_{sr}, \qquad \sum_r \xi_i^{(r)}\xi_k^{(r)} = \delta_{ik}. \tag{5}$$

We now introduce the operators

$$a_s = (2\hbar\omega_s)^{-\frac{1}{2}}\sum_i \xi_i^{(s)}(p_i - i\omega_s q_i),$$
$$a_s^* = (2\hbar\omega_s)^{-\frac{1}{2}}\sum_i \xi_i^{(s)}(p_i + i\omega_s q_i). \tag{6}$$

The commutation relations for these operators follow from (2) and (5). We find

$$[a_s, a_r^*] = \delta_{sr}, \qquad [a_s, a_r] = [a_s^*, a_r^*] = 0. \tag{7}$$

The inversion of (6) is readily accomplished using the relations (5). We find

$$q_i = i\sum_s \xi_i^{(s)}\left(\frac{\hbar}{2\omega_s}\right)^{\frac{1}{2}}(a_s - a_s^*),$$
$$p_i = \sum_s \xi_i^{(s)}\left(\frac{\hbar\omega_s}{2}\right)^{\frac{1}{2}}(a_s + a_s^*). \tag{8}$$

Inserting these expressions in the Hamiltonian (1) and using the relations (4) and (5) we find

$$H = \sum_s \hbar\omega_s(a_s^*a_s + \tfrac{1}{2}). \tag{9}$$

The operator $a_s^*a_s$ is the number operator for the sth normal mode; its eigenvalues are the nonnegative integers. The operator a_s is the step-down (annihilation) operator, and the operator a_s^* is the step-up (creation) operator; they have matrix elements only between eigenstates of the number operator which differ by unity.[15]

The evaluation of the expectation value (3) when F is a product of the a's and a^*'s is straightforward. Clearly, the only nonvanishing expectation values are for products containing an equal number of a^*'s and a's. The simplest of these are the pair expectation values. Thus

$$\langle a_s^*a_r\rangle = \delta_{sr}\,\mathrm{Tr}\,\{a_s^*a_se^{-\beta H}\}/\mathrm{Tr}\,\{e^{-\beta H}\} = \delta_{sr}\frac{\sum_{n=0}^{\infty} n\exp\left[-\frac{\hbar\omega_s}{kT}\left(n+\frac{1}{2}\right)\right]}{\sum_{n=0}^{\infty}\exp\left[-\frac{\hbar\omega_s}{kT}\left(n+\frac{1}{2}\right)\right]} = \delta_{sr}\left[\exp\left(\frac{\hbar\omega_s}{kT}\right) - 1\right]^{-1}$$

or

$$\langle a_s^*a_r\rangle = \tfrac{1}{2}\delta_{sr}[\coth\ \hbar\omega_s/2kT) - 1]. \tag{10}$$

Using the commutation relation (7), we have

$$\langle a_sa_r^*\rangle = \tfrac{1}{2}\delta_{sr}[\coth\,(\hbar\omega_s/2kT) + 1]. \tag{11}$$

The results for higher products are summarized by the following rule: The expectation value of a product of a's and a^*'s is equal to the sum of products of pair expectation values, the sum being over all possible pairings with the order of each pair preserved.

For example:

$$\langle a_s^*a_ra_ua_v^*\rangle = \langle a_s^*a_r\rangle\langle a_ua_v^*\rangle + \langle a_s^*a_u\rangle\langle a_ra_v^*\rangle.$$

We do not prove this rule here, since the most convincing demonstration is by example.

We turn now to the consideration of expectation

[15] These properties of the operators are discussed in many textbooks on quantum mechanics. See, e.g., A. Messiah, *Quantum Mechanics* (North-Holland Publishing Company, Amsterdam, 1961), Chap. 12.

values of products of q's and p's. Because of the linear relations (8) we have the same rule for these expectations:

The expectation value of an odd number of q's and p's vanishes. The expectation value of an even number of q's and p's is equal to the sum of products of pair expectation values, the sum being over all pairings which preserve the order of the pair.

For example:

$$\langle q_i q_j p_k q_l \rangle = \langle q_i q_j \rangle \langle p_k p_l \rangle + \langle q_i p_k \rangle \langle q_j q_l \rangle + \langle q_i q_l \rangle \langle q_j q_k \rangle.$$

The pair correlations are readily obtained from (10) and (11), using (8). Thus

$$\langle q_i q_k \rangle = \sum_s \frac{\hbar}{2\omega_s} \coth \frac{\hbar\omega_s}{2kT} \xi_i^{(s)} \xi_k^{(s)}.$$

Using (4) we see this can be written

$$\langle q_i q_k \rangle = \left|\left| \frac{\hbar}{2\mathbf{A}^{\frac{1}{2}}} \coth \frac{\hbar \mathbf{A}^{\frac{1}{2}}}{2kT} \right|\right|_{ik}. \tag{12}$$

Similarly, we find

$$\langle p_i p_k \rangle = ||\tfrac{1}{2}\hbar \mathbf{A}^{\frac{1}{2}} \coth (\hbar \mathbf{A}^{\frac{1}{2}}/2kT)||_{ik}, \tag{13}$$

$$\langle q_i p_k \rangle = -\langle p_i q_k \rangle = \tfrac{1}{2} i\hbar \delta_{ik}. \tag{14}$$

Consider now the time-dependent correlation functions in which the operators at time t are expressed in terms of the initial operators through the relations (4), considered here as formal solutions of the Heisenberg equations of motion. Again we have the rule that correlations of an odd number of operators vanish while a correlation of an even number of operators is equal to a sum of products of pair correlations, the sum being over all pairings which preserve the order of the pair. For example,

$$\begin{aligned}\langle q_i(t_1) q_k(t_2) p_l(t_3) q_m(t_4) \rangle &= \langle q_i(t_1) q_k(t_2) \rangle \langle p_l(t_3) q_m(t_4) \rangle \\ &+ \langle q_i(t_1) p_l(t_3) \rangle \langle q_k(t_2) q_m(t_4) \rangle \\ &+ \langle q_i(t_1) q_m(t_4) \rangle \langle q_k(t_2) p_l(t_3) \rangle.\end{aligned}$$

The pair correlations are

$$\langle p_i(t) p_k(t+\tau) \rangle$$

$$= ||\tfrac{1}{2}\hbar \mathbf{A}^{\frac{1}{2}} [\coth (\hbar \mathbf{A}^{\frac{1}{2}}/2kT) \cos \mathbf{A}^{\frac{1}{2}}\tau + i \sin \mathbf{A}^{\frac{1}{2}}\tau]||_{ik}, \tag{15}$$

$$\langle q_i(t) q_k(t+\tau) \rangle$$

$$= ||\tfrac{1}{2}\hbar \mathbf{A}^{-\frac{1}{2}} [\coth (\hbar \mathbf{A}^{\frac{1}{2}}/2kT) \cos \mathbf{A}^{\frac{1}{2}}\tau + i \sin \mathbf{A}^{\frac{1}{2}}\tau]||_{ik}, \tag{16}$$

$$\langle q_i(t) p_k(t+\tau) \rangle$$

$$= ||\tfrac{1}{2}\hbar [-\coth (\hbar \mathbf{A}^{\frac{1}{2}}/2kT) \sin \mathbf{A}^{\frac{1}{2}}\tau + i \cos \mathbf{A}^{\frac{1}{2}}\tau]||_{ik}. \tag{17}$$

The derivation of these expressions goes exactly as the derivation of the corresponding classical correlations, obtained in Appendix 1, but using the expressions (12), (13), and (14) for the initial expectation values.

Proof that Scattering Implies Production in Quantum Field Theory*

STANLEY Ø. AKS
Institute for Fluid Dynamics and Applied Mathematics† and Department of Physics and Astronomy, University of Maryland, College Park, Maryland
(Received 11 September 1964)

A proof of the necessity of production processes in quantum field theory is carried out in the axiomatic framework. It is shown that a field theory that is assumed to have a nontrivial scattering amplitude violates crossing symmetry, if production processes are absent. The proof is based on the rigorous analytic properties of the scattering amplitude, particularly the analyticity in the invariant-scattering variables s, t, and u. In the case of a scalar theory with pairing symmetry, the scattering amplitude is known to be analytic within the domain $|stu| < 7168m^6$, except for the usual cuts. Under the working assumption that production processes are null, it is shown that this domain can be enlarged by applying the elastic unitarity conditions beyond the (usual) elastic region. The domain is enlarged sufficiently to include the first Landau singularity of the absorptive part of the scattering amplitude. This singularity is not symmetric in s and t within the extended domain, and this is incompatible with the crossing symmetry of the scattering amplitude. In order to avoid a contradiction, the discontinuity across this Landau singularity must be null. It follows that the scattering amplitude must itself be null.

In the course of the proof it is shown that the conclusion is valid for a scattering amplitude satisfying the requirements of an S-matrix theory embodied in the Mandelstam representation.

1. INTRODUCTION

THE production of one or more particles in the collision of two elementary particles is a radical departure from the classical conception of matter. Originally anticipated in the early attempts to join quantum mechanics with electrodynamics, production processes have been encountered in virtually all of the interactions of elementary particles. Furthermore, production processes seem to follow quite naturally in the theories which have been used to describe the interactions of the elementary particles. In these theories it is assumed that the elementary particles can be described by quantized fields and, furthermore, that their interactions are communicated by particles, likewise described by quantized fields, in a straightforward generalization of quantum electrodynamics.

The precise relationship between production processes and the basic elements of these field theories are not entirely evident. For example, it is not easy to give a reassuring proof of the necessity of production processes in such theories. The perturbative methods which characterize calculations in these are ill suited to existence proofs. Not only is the convergence of the perturbation expansions in question (particularly in the case of strong interactions) but, if the perturbation series is examined term by term, unusual cancellations among the terms may occur and these are not easily ruled out.

To circumvent the many difficulties encountered in the usual (quantum electrodynamics) formulation of field theory, attention has been shifted to more general investigations of field theories in which no reference is made to specific interactions. This approach involves a system of axioms which characterizes those fundamental properties that all field theories are expected to share.[1] The axiomatic approach lends itself to existence theorems and, since the equations of motion are not specified, the theorems proved have considerable generality.

In this paper we give a proof of the necessity of production processes in a relativistic quantum field theory describing interacting particles. By interacting we mean that the elastic scattering amplitude is nonvanishing, and by relativistic quantum field theory we mean the formal scheme embodied in the axioms. The proof is given in detail for the case of self-interacting neutral scalar bosons of mass m, charge and spin zero. The extension of the proof to particles with nonzero spin and isospin is nontrivial

* This research was supported at the Department of Physics and Astronomy by the U. S. Air Force Office of Scientific Research and the National Science Foundation through grants and fellowships, and at the Institute of Fluid Dynamics and Applied Mathematics by the National Aeronautics and Space Administration.

† Present address.

[1] The axiomatic systems we have in mind have been proposed by H. Lehmann, K. Symanzik, and W. Zimmermann; N. N. Bogoliubov and A. S. Wightman. Our results apply to the Wightman formulation in the strict sense only after the axioms are suitably amended to admit the proof of the usual dispersion relations, etc., for the scattering amplitude. For a survey of the axiomatic systems see S. S. Schweber, *An Introduction to Relativistic Quantum Field Theory* (Row, Peterson, and Company, Evanston, Illinois, 1961), p. 721.

and is not to be considered. The generalization to the unequal mass case is less difficult, but is also not considered here.[2]

A number of simplifying assumptions are made in passing to avoid inessential complications which would tend to obscure the more significant aspects of the proof. Among these, we mention the simplifying assumptions of the pairing symmetry [i.e., the invariance of the field $A(x)$ under the transformation $A(x) \rightarrow -A(x)$]. Pairing symmetry is similar to G parity which occurs in the context of the interactions of pions and, like G parity, rules out processes involving an odd number of particles. In particular, production processes, when they occur, have their threshold at energy $4m$ in the center of momentum coordinate system. The assumption of pairing symmetry enables us to avoid a degeneracy which occurs in the equal-mass case only. The study of the equal-mass case without pairing symmetry is similar to the unequal-mass case and is also not considered here.[3]

The proof of the necessity of production processes in field theory is carried out "*reductio ad absurdum.*" The proof is split into two major parts for reasons of clarity of presentation and to permit the reader who is more interested in S-matrix theory than field theory to follow the main lines of reasoning. In the first part we prove the necessity of production processes in a relativistic scattering theory with a nontrivial elastic scattering amplitude which satisfies the analytic and crossing symmetry properties of the Mandelstam representation. The proof in this case includes most of the essential ideas and steps which are needed to prove the necessity of production in axiomatic field theory and it does avoid certain complications in field theory which tend to obscure the line of the proof. In the second part, the Mandelstam representation is replaced by the analytic properties of the elastic scattering amplitude which follow from the axioms, and the proof is modified to suit the restricted analytic properties which obtain.

Although words such as proof, theorem, etc. are used, they should not be regarded as used in the strict mathematical sense. Wherever possible, strict mathematical terminology and rigorous treatment are dropped in favor of simple language which, it is hoped, will give the arguments continuity and clarity.

2. THE MANDELSTAM REPRESENTATION AND THE ELASTIC UNITARITY CONDITION

In the following we are primarily concerned with the elastic scattering amplitude $T(p_3p_4;\ p_1p_2)$ describing the elastic scattering of particles with energy-momenta p_1 and p_2 before collision and p_3 and p_4 after collision. The scattering amplitude is related to the S-matrix through

$$(\Omega^{\text{out}}_{p_3p_4},\ \Omega^{\text{in}}_{p_1p_2}) = (\Omega^{\text{out}}_{p_3p_4},\ \Omega^{\text{out}}_{p_1p_2}) + 2iT(p_3p_4;\ p_1p_2). \quad (2.1)$$

The conservation of energy and momentum in elastic collisions is described by

$$T(p_3p_4;\ p_1p_2) = \delta(p_1 + p_2 - p_3 - p_4)\tilde{T}(p_3p_4;\ p_1p_2), \quad (2.2)$$

where $\delta(x)$, whenever it appears, is to be taken as the Dirac delta function appropriate to the argument; in this case it is the four-dimensional delta function. In the following we refer to both $T(p_3p_4;\ p_1p_2)$ and $\tilde{T}(p_3p_4;\ p_1p_2)$ as the elastic scattering amplitude; no confusion should result.

As we have noted above, the proof of the necessity of production processes in axiomatic field theory is complicated. For this reason we (temporarily) introduce the simplifying assumption that the elastic scattering amplitude is the boundary value of a function which has the analytic properties described by the Mandelstam representation,[4]

$$\tilde{T}(p_3p_4;\ p_1p_2) = \phi(s,\ t),$$

where

$$\begin{aligned}\phi(s,\ t) = \phi(s,\ t,\ u) &= \frac{1}{\pi^2}\int_{4m^2}^{\infty} ds' \int_{4m^2}^{\infty} dt' \frac{\rho_1(s',\ t')}{(s' - s)(t' - t)} \\ &+ \frac{1}{\pi^2}\int_{4m^2}^{\infty} dt' \int_{4m^2}^{\infty} du' \frac{\rho_2(t',\ u')}{(t' - t)(u' - u)} \\ &+ \frac{1}{\pi^2}\int_{4m^2}^{\infty} du' \int_{4m^2}^{\infty} ds' \frac{\rho_3(u',\ s')}{(u' - u)(s' - s)}, \end{aligned} \quad (2.3)$$

and $s = (p_1 + p_2)^2$, $t = (p_3 - p_1)^2$, and $u = (p_4 - p_1)^2$. We are using the Minowski quadratic form in taking the square of the energy-momentum 4-vector (i.e., $p^2 \equiv p_0^2 - \bar{p}^2$, $\bar{p}^2 \equiv p_1^2 + p_2^2 + p_3^2$). Since we are using $p_1^2 = p^2 = p_3^2 = p_4^2 = m^2$, we have $s + t + u = 4m^2$, and u is dependent on s and t.

[2] For a discussion of the unequal mass case (when anomolous thresholds do not occur) see, S. Ø. Aks, *A Proof that Scattering Implies Production in Quantum Field Theory* (Technical Report no. 345, Department of Physics and Astronomy, University of Maryland, College Park, Maryland, 1964), p. 39.

[3] The equal-mass case without pairing symmetry can be treated like the equal-mass case with pairing symmetry by subtracting the box diagram (with three-particle vertices) from the scattering amplitude. For details see Ref. 2, p. 34.

[4] S. Mandelstam, Phys. Rev. **112**, 1344 (1958).

However, in some instances, it is useful to display the u variable explicitly since $\phi(s, t, u)$ has properties readily described in terms of u [e.g., we shall see that $\phi(s, t, u)$ is symmetric under permutation of s, t, u]. When p_1 and p_2 are the energy-momenta of the particles before collision (i.e., particles one and two are incoming), we have $s \geq 4m^2$ and, $t \leq 0$ and $u \leq 0$. Henceforth, we speak of this configuration as the s channel.

In the s channel, s is the square of the energy in the center-of-momentum coordinate system. If θ is the scattering angle in this coordinate system, we have

$$t = -2\bar{q}^2(1 - \cos\theta), \tag{2.4a}$$

$$u = -2\bar{q}^2(1 + \cos\theta), \tag{2.4b}$$

where $\bar{q}$ is the relative momentum of the incoming particles, and,

$$s = 4\bar{q}^2 + 4m^2. \tag{2.4c}$$

The ambiquity in taking the boundary value in (2.3) is resolved by simply requiring that s tends to the cut $s \geq 4m^2$ (the s cut) from the upper half of the complex s plane.

The amplitude relating to the S-matrix element $(\Omega^{\text{out}}_{p_2p_4}, \Omega^{\text{in}}_{p_1p_3})$ is given by

$$(\Omega^{\text{out}}_{p_2p_4}, \Omega^{\text{in}}_{p_1p_3}) = (\Omega^{\text{out}}_{p_2p_4}, \Omega^{\text{out}}_{p_1p_3}) + 2iT(p_2p_4; p_1p_3),$$

where

$$T(p_2p_4; p_1p_3) = \delta(p_1 + p_3 - p_2 - p_4)\phi(s, t, u),$$

and where $s = (p_2 - p_1)^2$, $t = (p_1 + p_3)^2$ and $u = (p_4 - p_1)^2$. The scattering amplitude relating to the S-matrix element $(\Omega^{\text{out}}_{p_2p_3}, \Omega^{\text{in}}_{p_1p_3})$ is found similarly. When the particles with p_1 and p_3 are incoming, we speak of scattering in the t channel, and with p_1 and p_4 incoming, the u channel. In each channel the scattering amplitude is a boundary value (taken on the upper edge of the s cut in the s channel; taken on the upper edge of the t cut in the t channel; taken on the upper edge of the u cut in the u channel) of the analytic function $\phi(s, t, u)$. The existence of a single, although somewhat more complicated, analytic function whose boundary values along relevant cuts determines the scattering amplitudes in the various channels has been demonstrated in the context of field theory. However, for the present we use Mandelstam representation because of it's convenient domain of analyticity.

The proof of the necessity of production processes is carried out "*reductio ad absurdum*," all production processes are assumed to be null, and, under this working hypothesis, we show that $\phi(s, t, u)$ is the null function. However, before applying the working hypothesis we describe certain restrictions on $\phi(s, t, u)$ that follow simply from the physical properties of the scattering amplitude. We know that the s, t, and u channels differ only by the labeling of the particles, hence, the related boundary values must be equal. Furthermore, since the particles are bosons, they obey Bose–Einstein statistics and the scattering amplitude is an even function of $\cos\theta$ when expressed in center-of-momentum variables. It follows that $\phi(s, t, u)$ must be completely symmetric under permutations of s, t, and u. This we call the crossing symmetry; it occurs quite naturally for the scattering amplitude in the context of the axiomatic scheme. In order for $\phi(s, t, u)$ to be totally symmetric, the (density) functions $\rho_1(s, t)$, $\rho_2(t, u)$, and $\rho_3(u, s)$ must all be equal [$\rho(x, y) \equiv \rho_1(x, y) = \rho_2(x, y) = \rho_3(x, y)$], and symmetric [$\rho(x, y) = \rho(y, x)$]. It is a consequence of TCP invariance that $\phi(s, t, u)$ is a real analytic function, $\phi^*(s, t, u) = \phi(s^*, t, u)$ for $t \leq 4m^2$ and $u \leq 4m^2$, and similarly for the other variables. It follows that the density function $\rho(s, t)$ is real.

We note that $\phi(s, t)$, as defined in (2.3), implies boundedness properties at infinity which need not be satisfied physically. It is usually assumed that the scattering amplitude is bounded by a symmetric polynomial, that is, it has the form (2.3) up to an additive symmetric polynomial. We will not be concerned with this polynomial (called a subtraction term from the procedure used to compare it) and so it is omitted. In field theory the polynomial boundedness assumption is formally included in the technical axioms.

By assumption, no processes involving an uneven number of particles occur, therefore two-particle scattering is elastic in the s (t or u) channel for $4m^2 \leq s < 16m^2$. In this range the scattering amplitude satisfies the elastic unitarity condition. Under the working assumption that no production processes occur, the scattering amplitude then satisfies the elastic unitarity condition for all $s \geq 4m^2$. We find this incompatible with the above described analytic properties and crossing symmetry, so we must give up the possibility that scattering can occur without production. We have noted the relationship (2.1) between S-matrix elements and the scattering amplitude. This can be re-expressed as the operator relationship $S = I + 2iT$, where I is the identity operator, and the expectation values of S, for the outgoing particles states, are the S-matrix elements. When the incoming and outgoing particle states are complete, the S-matrix transforms one orthonormal

basis into another and is, therefore, a unitary operator[5] satisfying

$$SS^\dagger = S^\dagger S = I.$$

Coupling these operator equations, we obtain

$$(2i)^{-1}(T - T^\dagger) = TT^\dagger = T^\dagger T.$$

In terms of the outgoing states this becomes

$$(\Omega^{\text{out}}_{p_3p_4},\ (2i)^{-1}[T - T^\dagger]\Omega^{\text{out}}_{p_1p_2}) = (\Omega^{\text{out}}_{p_3p_4},\ T^\dagger T\Omega^{\text{out}}_{p_1p_2}).$$

Introducing the complete set of outgoing states between $T^\dagger$ and T in the right-hand side permits us to rewrite this as

$$(\Omega^{\text{out}}_{p_3p_4},\ (2i)^{-1}[T - T^\dagger]\Omega^{\text{out}}_{p_1p_2}) = \sum_{m=0}^{\infty} \sum_{\text{phase space}}^{k_1 \cdots k_m} (\Omega^{\text{out}}_{p_3p_4},\ T^\dagger\Omega^{\text{out}}_{k_1,\ldots,k_m})(\Omega^{\text{out}}_{k_1\cdots k_m},\ T\Omega^{\text{out}}_{p_1p_2}),$$

where the double sum formally denotes that we are splitting the right-hand side into contributions from the vacuum state, one-particle states, two-particle states, etc.[6] Since each of the terms on the right-hand side contains an energy–momentum delta function, the contributions from $n = 0$ and $n = 1$ are null. In the elastic region the contributions for $n \geq 3$ also vanish. The elastic region is either the interval $4m^2 \leq s < 16m^2$ or, in the case where it is assumed that no production occurs, $s \geq 4m^2$. In the elastic region we get

$$\delta(p_1 + p_2 - p_3 - p_4)\ \text{Im}\ \phi(s, t, u) = \delta(p_1 + p_2 - p_3 - p_4) \sum_{\text{phase space}}^{k_1k_2} \phi^*(s, t'', u'') \times \phi(s, t', u')\delta(k_1 + k_2 - p_1p_2),$$

where

$$t' = (k_1 - p_1)^2, \qquad u' = (k_2 - p_1)^2,$$
$$t'' = (p_3 - k_1)^2, \qquad u'' = (p_4 - k_1)^2.$$

On introducing the explicit summation over two-particle phase space, this becomes

$$\text{Im}\ \phi(s, t, u) = \int dk_1 \int dk_2 \theta(k_1^0)\delta(k_1^2 - m^2)\theta(k_2^0)\delta(k_2^2 - m^2) \times \phi^*(s, t'', u'')\phi(s, t', u')\delta(k_1 + k_2 - p_1 - p_2),$$

where the conservation delta functions which occur on both sides are omitted, but should, nevertheless, not be forgotten. The function $\theta(x)$ is the usual characteristic function of the positive axis [$\theta(x) = 1$, $x \geq 0$ and $\theta(x) = 0$, $x \leq 0$].

Some of the integrations can be carried out in the center-of-momentum system, $\bar{p}_1 + \bar{p}_2 = 0$. It is convenient to use the center-of-momentum variables s, $\bar{q}$, and $\cos\theta$, and introduce $\cos\theta_1$ and $\cos\theta_2$ in a similar way through $t' = -2\bar{q}^2(1 - \cos\theta_1)$ and $t'' = -2\bar{q}^2(1 - \cos\theta_2)$.

After performing integrations we can write the elastic unitarity condition in terms of these variable as

$$\text{Im}\ \phi(s, \cos\theta) = \frac{1}{8}\theta(s - 4m^2)\left(\frac{s - 4m^2}{s}\right)^{\frac{1}{2}} \int_{-1}^{+1} d\cos\theta_1 \int_{-1}^{+1} d\cos\theta_2 \times \frac{\theta[k(\cos\theta, \cos\theta_1, \cos\theta_2)]}{[k(\cos\theta, \cos\theta_1, \cos\theta_2)]^{\frac{1}{2}}} \phi(s; \cos\theta_1)\phi^*(s; \cos\theta_2),$$

where

$$k(\cos\theta, \cos\theta_1, \cos\theta_2) = 1 - \cos^2\theta - \cos^2\theta_1 - \cos^2\theta_2 + 2\cos\theta\cos\theta_1\cos\theta_2 \qquad (2.5)$$

and

$$\phi(s, t, u) = \phi(s; \cos\theta).$$

In much of our work we are concerned with the analytic properties of (2.5) and we make use of the ample literature on the subject, in particular, we find the paper of Zimmermann[7] of great value. Since the imaginary part of the scattering amplitude has analytic properties and, moreover, since the analytic continuation of Im $\phi(s; \cos\theta)$ does not remain the imaginary part of the scattering amplitude, we follow the custom of denoting the analytic continuation by $A(s; \cos\theta)$, and call it the absorptive part. Because of the extensive literature we shall but quickly survey the properties of the absorptive part limiting our attention to finding necessary conditions for the existence of a singularity. The question of the actual existence of a singularity is examined closely in a relevant case in Sec. 4.

The singularities of the absorptive part with which we are concerned are connected with the somewhat more general expression

$$U(z) = \int_{-1}^{+1} dz_1 \int_{-1}^{+1} dz_2 \frac{\theta[k(z, z_1, z_2)]}{[k(z, z_1, z_2)]^{\frac{1}{2}}} V(z_1)W(z_2), \qquad (2.6)$$

with $V(z_1)$ and $W(z_2)$ both analytic in domains which include the integration intervals. If we replace $V(z_1)$ and $W(z_2)$ by their Cauchy representations on contours C_1 and C_2 (within their domains of

[5] We are assuming, of course, that no bound states occur.
[6] All contributions for m odd are already null as a consequence of the pairing symmetry.
[7] W. Zimmermann, Nuovo Cimento **21**, 268 (1961).

analyticity), after interchanging the order of integrations, we obtain

$$U(z) = \frac{-1}{4\pi^2} \oint_{C_1} dz_1' \oint_{C_2} dz_2' \, H(z_1', z_2'; z) V(z_1') W(z_2'), \quad (2.7)$$

where

$$H(z_1', z_2'; z) = \left(\frac{1}{2\pi i}\right)^2 \int_{-1}^{+1} \frac{dz_1}{z_1 - z_1'} \times \int_{-1}^{+1} \frac{dz_2}{z_2 - z_2'} \frac{\theta[k(z, z_1, z_2)]}{[k(z, z_1, z_2)]^{\frac{1}{2}}}. \quad (2.8)$$

If we take the positive root of $[k(z, z_1, z_2)]^{\frac{1}{2}}$ when $-1 \leq z \leq +1$, $H(z_1', z_2'; z)$ is analytic in z_1' and z_2' in the complement of the cuts $-1 \leq z_1' \leq +1$ and $-1 \leq z_2' \leq +1$.

The z_1 and z_2 integrations can be carried out. First we note that the z_2 integration covers the interval on which $k(z, z_1, z_2)$ is positive. The endpoints of this interval are $z_2^{\pm} = zz_1 \pm [(z^2-1)(z_1^2-1)]^{\frac{1}{2}}$, the zeros of $k(z, z_1, z_2)$. Since the square root in the denominator of (2.8) takes on opposite signs on the upper and lower edges of the cut $z_2^- \leq z \leq z_2^+$, we can replace the z_2 integration by a contour integration around the cut (in the clockwise direction) and take half of the result. Within this contour the integrand is singular at $z_2 = z_2'$ with residue

$$-\pi[-k(z, z_1, z_2')]^{\frac{1}{2}}.$$

The z_1 integration then gives

$$H(z_1', z_2'; z) = \pi[-k(z, z_1', z_2')]^{-\frac{1}{2}} \times \log \frac{z_1'z_2' - z + [-k(z, z_1', z_2')]^{\frac{1}{2}}}{z'z_2' - z - [-k(z, z_1', z_2')]^{\frac{1}{2}}}. \quad (2.9)$$

The square roots are to be taken so that $H(z_1', z_2'; z)$ is singular in z when $z = z_1'z_2' + [(z_1'^2 - 1)(z_2'^2 - 1)]^{\frac{1}{2}}$. The other zero of the square-root factor occurs when the logarithm vanishes, and is, therefore, removable. A function with this analyticity satisfies the following representation (in the z variable):

$$H(z_1', z_2'; z) = -2\pi \int_{z_1'z_2' + (z_1'^2-1)^{\frac{1}{2}}(z_2'^2-1)^{\frac{1}{2}}}^{\infty} \frac{dz'}{z' - z} \times (-1 + z'^2 + z_1'^2 + z_2'^2 - 2z'z_1'z_2')^{-\frac{1}{2}}, \quad (2.10)$$

the roots $(z_1'^2 - 1)^{\frac{1}{2}}$ and $(z_2'^2 - 1)^{\frac{1}{2}}$ are taken as positive for z_1' and z_2' real, and greater than one.

We do not anticipate being able to carry out the integrations in Eq. (2.7), but, with what we already know, we are able to determine certain analytic properties of $U(z)$ which follow from the analytic properties of $V(z_1)$ and $W(z_2)$. A convenient way to describe the connection between these functions is to introduce the variables

$$\alpha_1 = \cosh^{-1} z_1', \quad \alpha_2 = \cosh^{-1} z_2', \quad \alpha_3 = \cosh^{-1} z.$$

In terms of these variables the initial point of the cut of $H(z_1', z_2'; z)$ is given by

$$\cosh \alpha = \cosh (\alpha_1 + \alpha_2). \quad (2.11)$$

The function $U(z)$ is analytic at all points inside the curve $z = \cosh \alpha$, where $\alpha_1(z_1')$ and $\alpha_2(z_2')$ run over their contours. Let us suppose the contours C_1 and C_2 are ellipses with foci ± 1 and semimajor axis $\cosh \lambda_1$ and $\cosh \lambda_2$. The equations for these ellipses are $\cosh (\lambda_1 + i_{\mu_1})$ and $\cosh (\lambda_2 + i_{\mu_2})$ with μ_1 and μ_2 real; the domain of analyticity in $U(z)$ includes the interior of the curve

$$z = \cosh (\lambda_1 + \lambda_2 + i\mu_1 + i\mu_2). \quad (2.12)$$

This is the equation of an ellipse with foci ± 1 and semimajor axis z_0, where $z_0 = \cosh (\lambda_1 + \lambda_2)$.

As we are frequently concerned with ellipses with foci ± 1, it is convenient to denote them by $E(a)$, where a is the semimajor axis. Thus the ellipse with foci $\pm$ 1 passing through the point $z = \cosh \alpha$ is E (cosh Re α).

Although (2.6) is modeled after the elastic unitarity condition (2.5), there is one essential difference: the function $\phi^*(s; \cos \theta)$ is not obviously analytic in $\cos \theta$ because of the complex conjugation; hence, the analyticity of the absorptive part $A(s; \cos \theta)$ cannot be inferred as for $U(z)$. This difficulty can be resolved since $\phi(s; \cos \theta)$ has been shown[8] to have a square-root singularity in the s variable at $4m^2$, and the analytic extension through the cut $s \geq 4m^2$ onto the second sheet, denoted by $\phi^{(2)}(s; \cos \theta)$, is, for $s \geq 4m^2$, just $\phi^*(s; \cos \theta)$. Then $\phi^*(s; \cos \theta)$ can be replaced by $\phi^{(2)}(s; \cos \theta)$ in (2.5), and the analyticity of $A(s; \cos \theta)$, following from the analyticity of $\phi(s; \cos \theta)$ and $\phi^{(2)}(s; \cos \theta)$, can be treated similar to $V(z)$ in (2.5). Nevertheless, the problem is formidable since the analytic extension of $\phi(s; \cos \theta)$ to $\phi^{(2)}(s; \cos \theta)$ itself depends on the elastic unitarity condition [specifically the analyticity of $A(s; \cos \theta)$]. A full determination of the analyticity of the absorptive part requires a step-by-step discussion. First, analyticity in $A(s; \cos \theta)$ must be found in some domain, then the analyticity of $\phi^{(2)}(s; \cos \theta)$ following from it can be determined. By using this analyticity of $\phi^{(2)}(s; \cos \theta)$ in the elastic unitarity condition, along with the analyticity of $\phi(s; \cos \theta)$, further analyticity is found for $A(s; \cos \theta)$. Using this it follows that further analyticity in $\phi^{(2)}(s; \cos \theta)$ results, and, the argument can be repeated to further extend the analyticity of analyticity

[8] W. Zimmermann, Nuovo Cimento **21**, 268 (1961).

of the absorptive part. This technique was introduced by Mandelstam[9] and refined by Zimmermann. In the following we make use of both versions.[10]

3. THE LANDAU SINGULARITIES AND THE MANDELSTAM REPRESENTATION

We now survey Zimmermann's work on the behavior of the scattering amplitude at $s = 4m^2$ with particular regard to the singularities on the second sheet of the square root found there. The scattering amplitude $\phi(s; \cos\theta)$ determines the partial wave amplitudes

$$\phi_l(s) = \frac{1}{2}\int_{-1}^{+1} d\cos\theta P_l(\cos\theta)\phi(s;\cos\theta), \qquad l = 0, 1, 2, \cdots;$$

the $P_l(\cos\theta)$ are Legendre polynomials. As we have previously noted, $\phi(s; \cos\theta)$ is a real analytic function, and, since the Legendre polynomials are real-valued functions in the interval $-1 \leq \cos\theta \leq +1$, it follows that $\phi_l(s)$ is a real analytic function, hence, $\phi_l^*(s) = \phi_l(s^*)$. From (2.3) it follows that each of the functions $\phi_l(s)$ is analytic in the s plane, except for the cuts $s \leq 0$ and $s \geq 4m^2$.

It is possible to reconstruct the scattering amplitude $\phi(s; \cos\theta)$ from the partial waves. To do this we must first examine the analytic properties in the $\cos\theta$ plane for fixed s not on the cut $s \leq 0$. The scattering amplitude is analytic in the $\cos\theta$ plane except for the images of the $t \geq 4m^2$ and $u \geq 4m^2$ cuts; the cuts start at $\cos\theta = \pm[1 + 8m^2/(s-4m^2)]$. These points lie on the ellipse $E[a_1(s)]$, where

$$a_1(s) = [1 + 8m^2(|s| + \operatorname{Re} s)/|s - 4m^2|^2]^{\frac{1}{2}}.$$

Within this ellipse, the partial wave expansion

$$\phi(s;\cos\theta) = \sum_{l=0}^{\infty}(2l+1)\phi_l(s)P_l(\cos\theta)$$

converges uniformly and absolutely.[11]

The two-particle unitarity condition can be reexpressed as a condition on the partial wave amplitude using[12]

$$\frac{\theta(1 - z_1^2 - z_2^2 - z_3^2 + 2z_1z_2z_3)}{(1 - z_1^2 - z_2^2 - z_3^2 + 2z_1z_2z_3)^{\frac{1}{2}}} = \frac{\pi}{2}\sum_{l=0}^{\infty}(2_l + 1)P_l(z_1)P_l(z_2)P_l(z_3).$$

In the elastic region,

$$\operatorname{Im}\phi_l(s) = \tfrac{1}{4}\pi[(s - 4m^2)/s]^{\frac{1}{2}}\,|\phi_l(s)|^2.$$

It follows from this, and the reality property, that each $\phi_l(s)$ can be extended through the cut $s \geq 4m^2$ onto a second sheet, the singularity at $s = 4m^2$ being of square-root type. The partial wavefunction can be written as

$$\phi_l(s) = F_l(s) + \tfrac{1}{4}i\pi[(s - 4m^2)/s]^{\frac{1}{2}}G_l(s),$$

where both $F_l(s)$ and $G_l(s)$ are (real) analytic functions in a neighborhood of $s = 4m^2$. The singular part $[(s - 4m^2)/s]^{\frac{1}{2}}$ is chosen positive on the upper edge of the cut $s \geq 4m^2$ and is analytic in the intervals between the left- and right-hand cuts. The analytic extension to the second sheet at $s = 4m^2$ is denoted by $\phi_l^{(2)}(s)$ and is given by

$$\phi_l^{(2)}(s) = F_l(s) - \tfrac{1}{4}i\pi[(s - 4m^2)/s]^{\frac{1}{2}}G_l(s).$$

The functions $F_l(s)$ and $G_l(s)$ can have isolated singularities but, since we know that $\phi_l(s)$ is analytic in the complement of the cuts $s \leq 0$ and $s \geq 4m^2$, their singularities must compensate for each other on the first sheet. On the second sheet they do not compensate (according to the change of sign between terms) and isolated sigularities may occur. However, we neglect this possibility as their presence only serves to complicate our work without materially affecting the results.[13]

We have already mentioned that the partial wave expansion converges uniformly in the ellipse $E[a_1(s)]$. The function $F_l(s)$ defines a similar expansion which is also convergent in the ellipse $E[a_1(s)]$ for s off the cut $s \leq 0$, however, the expansion

$$G(s;\cos\theta) = \sum_{l=0}^{\infty}(2l+1)G_l(s)P_l(\cos\theta)$$

is analytic in a larger ellipse. We note that $G_l(s)$ is, up to the factor $\frac{1}{4}\pi[(s - 4m^2)/s]^{\frac{1}{2}}$, equal to $\operatorname{Im}\phi_l(s)$. We noted above that the absorptive part of $\phi(s;\cos\theta)$ [the analytic extension of $\operatorname{Im}\phi(s;\cos\theta)$] is analytic in the ellipse $E(2a^2 - 1)$ if the scattering amplitude is analytic in the ellipse $E(a)$.[14] It follows that $G(s;\cos\theta)$ is analytic in the ellipse $E(a_2)$, where $a_2(s) = 2a_1(s)^2 - 1$. Zimmermann gives a direct proof of this in his paper. On summing the Legendre series with coefficients $\phi_l(s)$ and $\phi_l^{(2)}(s)$ we obtain

$$\phi(s;\cos\theta) = F(s;\cos\theta) + \tfrac{1}{4}i\pi[(s - 4m^2)/s]^{\frac{1}{2}}G(s;\cos\theta), \qquad (3.1a)$$

[9] S. Mandelstam, Nuovo Cimento **15**, 658 (1960).

[10] (To be published). The problem has recently been studied by S. Ø. Aks, R. Gilbert, and H. Howard using more general theorems which may simplify the arguments involved.

[11] E. T. Whittaker and G. N. Watson, *Modern Physics* (The MacMillan Company, New York, 1948), p. 332.

[12] The θ function on the left-hand side serves to define the range of validity of the expression.

[13] A discussion of these singularities is given in R. Blankenbeckler, M. L. Goldberger, S. W. McDowell, and S. B. Treiman, Phys. Rev. **122**, 983 (1961).

[14] This follows directly from (2.12) and the trigonometric identity $\cosh 2\lambda = 2\cosh^2\lambda - 1$.

$$\phi^{(2)}(s; \cos\theta) = F(s; \cos\theta) - \tfrac{1}{4}i\pi[(s-4m^2)/s]^{\frac{1}{2}}G(s; \cos\theta), \quad (3.1b)$$

for s not on the cut $s \leq 0$, and $\cos\theta$ in $E[a_1(s)]$. Each of the terms on the right-hand side of (3.1a) is analytic in the s variable. $\phi(s; \cos\theta)$ has a square-root singularity at $s = 4m^2$ and $\phi^{(2)}(s; \cos\theta)$ is the analytic continuation onto the second sheet. Similarly, $\phi^{(2)}(s; \cos\theta)$ is analytic in s and $\cos\theta$ within a domain described below. We can eliminate $F(s; \cos\theta)$ from the pair of equations (3.1) to get

$$\phi^{(2)}(s; \cos\theta) = \phi(s; \cos\theta) - \tfrac{1}{4}i\pi[(s-4m^2)/s]^{\frac{1}{2}}G(s; \cos\theta). \quad (3.2)$$

It follows that $\phi^{(2)}(s; \cos\theta)$ has the singularities of $G(s; \cos\theta)$ as well as the s, t, and u cuts of $\phi(s; \cos\theta)$. Hence, $\phi^{(2)}(s; \cos\theta)$ is analytic within the ellipse $E[a_2(s)]$ (s not on the cut $s \leq 0$), except for the t and u cuts, which, in the $\cos\theta$ plane, are given by

$$\cos\theta = \pm[1 + 2t/(s-4m^2)], \quad t \geq 4m^2.$$

In terms of the complex angle $\alpha = i\theta$, this becomes

$$\alpha(s, t) = \cosh^{-1}[1 + 2t/(s-4m^2)], \quad t \geq 4m^2. \quad (3.3)$$

The cut starts at

$$\alpha(s) = \alpha(s, 4m^2) = \cosh^{-1}[1 + 8m^2/(s-4m^2)]. \quad (3.4)$$

The semimajor axes $a_1(s)$ and $a_2(s)$ (defined above) are given in terms of $\alpha(s)$, by $a_1(s) = \cosh \operatorname{Re}\alpha(s)$, $a_2(s) = \cosh 2 \operatorname{Re}\alpha(s)$. We also have occasion to use the semimajor axes $a_n(s)$ defined by

$$a_n(s) = \cosh n \operatorname{Re}\alpha(s). \quad (3.5)$$

For $s \geq 4m^2$, $G(s; \cos\theta)$ is, up to a factor, the imaginary (or absorptive) part of the scattering amplitude. Therefore, in the elastic region we have

$$G(s; \cos\theta) = \frac{1}{2\pi}\int_{-1}^{+1} d\cos\theta_1 \int_{-1}^{+1} d\cos\theta_2 \frac{\theta(1 - \cos^2\theta_1 - \cos^2\theta_2 - \cos^2\theta + 2\cos\theta_1\cos\theta_2\cos\theta)}{(1 - \cos^2\theta_1 - \cos^2\theta_2 - \cos^2\theta + 2\cos\theta_1\cos\theta_2\cos\theta)^{\frac{1}{2}}} \times \phi(s; \cos\theta_1)\ \phi^{(2)}(s; \cos\theta_2). \quad (3.6)$$

Since $\phi(s; \cos\theta_1)$ and $\phi^{(2)}(s; \cos\theta_2)$ possess analytic continuations into the s plane (except for the cut $s \leq 0$), (3.6) is valid except on the cut $s \leq 0$. Equation (3.6) can be combined with (3.2) to provide a linear integral equation in $\phi^{(2)}(s; \cos\theta)$.

From (3.6) we see that for each s (not on the cut $s \leq 0$) $G(s; \cos\theta)$ has the form of $U(z)$ in (3.6). Furthermore, $\phi(s; \cos\theta)$ is an analytic within the $\cos\theta$ plane, except for the t and u cuts, and (for s not on the cut $s \leq 0$) $\phi^{(2)}(s; \cos\theta)$ is analytic within $E[a_2(s)]$, except for the t and u cuts. The arguments used in Sec. 2 on $U(z)$ can be applied directly to $G(s; \cos\theta)$ to show that $G(s; \cos\theta)$ is analytic within $E[a_3(s)]$, except for two regions denoted by $D_2^{\pm}$, which are related to the t and u cuts of $\phi(s; \cos\theta_1)$ and $\phi^{(2)}(s; \cos\theta_2)$ through the singularity combination formula (2.11). From (3.2) we find that $\phi^{(2)}(s; \cos\theta)$ is also analytic in $E[a_3(s)]$, but here the t and u cuts as well as $D_2^{\pm}$ must be excluded.

We show this as follows: Both $\phi(s; \cos\theta_1)$ and $\phi^{(2)}(s; \cos\theta_2)$ are analytic in the ellipses $E[a_2(s)]$ except for cuts in the variables $\cos\theta_1$ and $\cos\theta_2$ (with initial points located at $\cosh\alpha_1$, and $\cosh\alpha_2$). Using (2.11) the singularity points may be characterized by $\pm\cosh[\alpha(s, t_1) + \alpha(s, t_2)]$ and $\cosh[\alpha' + \alpha(s, t)]$, where $\alpha = \cosh^{-1}\cos\theta_1$, with $\cos\theta_1$ running over the ellipse $E[a_2(s)]$. The first set of singularity points determines the domains D_2^+ and D_2^-; the second set determines the ellipse $E[a_3(s)]$. The points $\pm\cosh[\alpha(s, 4m^2) + \alpha(s, 4m^2)]$ of D_2^+ and D_2^- lie on $E[a_2(s)]$.

Similarly, $\phi^{(2)}(s; \cos\theta)$ can be extended into the ellipse $E[a_3(s)]$ if the points of $D_2^{\pm}$ are excluded along with the images of the t and u cuts. Also, for $\cos\theta$ fixed within $E[a_3(s)] - D_2^+ - D_2^-$, $\phi^{(2)}(s; \cos\theta)$ is analytic in s, except for the cuts $s \geq 4m^2$ and $s \leq 0$.

Zimmermann repeats the arguments again and again to extend the functions $G(s; \cos\theta)$ and $\phi^{(2)}(s; \cos\theta)$ into the ellipses $E[a_4(s)]$, $E[a_5(s)]$, etc. He uses the analyticity found in the previous step to make the next step, and, in this manner, he finds that $\phi^{(2)}(s; \cos\theta)$ is analytic in s and $\cos\theta$ except for:

(a) the normal cuts:

$$\left.\begin{aligned} &s \geq 4m^2 \\ &t \geq 4m^2 \\ &u \geq 4m^2 \end{aligned}\right\} \cos\theta = \pm\left(1 + \frac{2t}{s-4m^2}\right), \quad t \geq 4m^2; \quad (3.7)$$

(b) the cut $s \leq 0$;

(c) the domains $D_n^{\pm}$ given by

$$D_n^{\pm} = \{\cos\theta \mid \cos\theta = \pm\cosh[\alpha(s, t_1) + \cdots + \alpha(s, t_n)], \quad t_i \geq 4m^2,\ i = 1, 2, \cdots, n\};$$

(d) the poles, which we are disregarding.

The function $G(s; \cos\theta)$ also has these singularities except for the normal cuts.

We can now discuss the Landau singularities.[15] For $s > 0$, the domains reduce to cuts starting at the points $\cos\theta = \pm\cosh[n\alpha(s)]$ on the ellipses $E[a_n(s)]$. These are the points which we refer to as Landau singularities, they lie on the Landau curves as s varies. For $n = 1$ the Landau singularities occur at the initial points of the t and u cuts. For $n = 2, 3, \cdots$ trigonometric identities are used to locate the Landau singularities in the t plane. For $n = 2$, we use $\cosh 2x = 2\cosh^2 x - 1$ and get

$$\cos\theta = \pm\{1 + 32m^2s/(s - 4m^2)^2\}$$

which give (for the positive sign)

$$t_2^+(s) = 16m^2 + 64m^4/(s - 4m^2). \qquad (3.8)$$

For the negative sign we obtain $t_2^-(s) = 4m^2 - s - t_2^+(s)$; this correspondence between the singularities of both signs is valid for all n.

The Landau singularity $t_3^+(s)$ is given by

$$t_3^+(s) = 36m^2 + 384m^4/(s - 4m^2) + 1024m^6/(s - 4m^2)^2. \qquad (3.9)$$

In general, $t_n^+(s)$ is given by a polynomial in $(s - 4m^2)^{-1}$ with positive coefficients; the constant term is $4n^2m^2$. It follows that the Landau curves are asymptotic to $s = 4m^2$ and $t = 4n^2m^2$. Furthermore, for $s \geq 4m^2$ the Landau curves lie on one behind the other, that is, starting from $s = 4m^2$ one reaches the curve $t_{n+1}^+(s)$ in the s–t plane only by traversing the curve $t_n^+(s)$.

It is important to remember that the Landau singularities are the initial points of the regions $D_n^\pm$ and that outside of these regions no other singularities occur. In the next section, we show that this result is inconsistent with the requirements of crossing symmetry and the analytic properties of the Mandelstam representation. In Zimmermann's work, elastic unitarity is used in the interval $4m^2 \leq s < 16m^2$, and, outside of this interval, other Landau singularities due to the inelastic contributions to the unitarity equation occur. These other singularities fulfill the requirements of crossing symmetry.

Later we show that, in the somewhat more restricted domain of analyticity which obtains for the elastic scattering amplitude in axiomatic field theory, the arguments used here (to determine the Landau singularities) can be suitably modified. There are, however, some new problems to be faced.

[15] The Landau singularities were first encountered in perturbation theory and studied systematically in this context by L. D. Landau, Nuclear Phys. **13**, 181 (1959).

4. THE NECESSITY OF PRODUCTION PROCESSES: THE CASE OF THE MANDELSTAM REPRESENTATION FOR EQUAL MASSES

The proof of the necessity of production processes (when nontrivial elastic scattering occurs) is carried out "*reductio ad absurdum*," all production processes are assumed null.[16] It follows that all scattering must be elastic, hence the elastic unitarity condition obtains for all $s \geq 4m^2$. It follows that for $s \geq 4m^2$, the function $G(s; \cos\theta)$ is analytic in $\cos\theta$, except for cuts whose initial points lie on the Landau curves. We now show that this singularity structure is not consistent with properties of the scattering amplitude, particularly, the crossing symmetry, unless the elastic scattering amplitude is also null.

The density function $\rho(s, t)$ in (2.3) is given by

$$\rho(s, t) = \lim_{\epsilon\to 0^+}\{\phi(s + i\epsilon, t + i\epsilon) - \phi(s - i\epsilon, t + i\epsilon) - \phi(s + i\epsilon, t - i\epsilon) + \phi(s - i\epsilon, t - i\epsilon)\} \qquad (4.1)$$

for $s \geq 4m^2$ and $t \geq 4m^2$. From (3.1) and the reality of the scattering amplitude we have

$$\phi(s + i\epsilon, t) - \phi(s - i\epsilon, t) = \phi(s + i\epsilon, t) - \phi^{(2)}(s + i\epsilon, t) = \tfrac{1}{2}i\pi[(s - 4m^2)/s]^{\frac{1}{2}}G(s + i\epsilon, t),$$

where $G(s, t)$ is just $G(s; \cos\theta)$ expressed in the s and t variables. No confusion should result from this notation. It follows that

$$\rho(s, t) = \tfrac{1}{2}i\pi[(s - 4m^2)/s]^{\frac{1}{2}} \times \lim_{\epsilon\to 0^+}\{G(s, t + i\epsilon) - G(s, t - i\epsilon)\}. \qquad (4.2)$$

The density function is (up to a factor) the discontinuity of the absorptive part. We now examine this discontinuity in detail. For $s \geq 4m^2$, and

$$4m^2 \leq t < t_2^+(s) = 16m^2 + 64m^4/(s - 4m^2),$$

$G(s, t)$ is analytic and the discontinuity vanishes. It follows that $\rho(s, t)$ vanishes for $4m^2 \leq t < t_2^+(s)$. See Fig. 1.

The density function $\rho(s, t)$ is symmetric in s and t, hence, the points at which $\rho(s, t)$ vanishes is likewise symmetric; see Fig. 1. However, the region $t > t_2^+(s)$ in which $\rho(s, t)$ can be nonvanishing is not symmetric. In particular, $\rho(s, t)$ vanishes in the strip $4m^2 \leq t \leq 16m^2$ and therefore, by symmetry, in the strip $4m^2 \leq s \leq 16m^2$ as well. Thus $\rho(s, t)$ also vanishes for values of s and t above the Landau curve $t_2^+(s)$. Using the elastic unitarity condition along with the analytic properties, we show

[16] This condition is stronger than actually required for the proof. All we use is the absence of production processes below the second inelastic threshold at $s = 36m$.

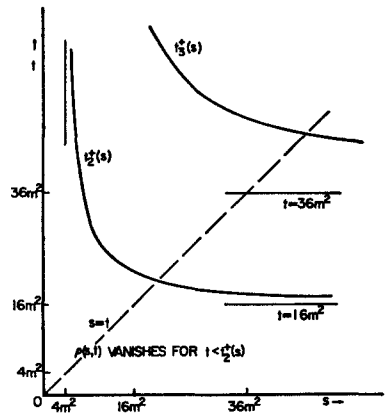

FIG. 1. The support of $\rho(s, t)$.

that the elastic scattering amplitude must be null in this circumstance.

The essence of the proof is the demonstration that the vanishing of $\rho(s, t)$ for $4m^2 \leq s \leq 16m^2$ implies the absorptive part of the scattering amplitude is null. Elastic unitarity is then used to show that the scattering amplitude is also null. Since $\rho(s, t)$ is (up to a factor) the discontinuity of the absorptive part, we will consider an expression for the discontinuity in order to determine the consequences of the vanishing of $\rho(s, t)$.

Previously, we found that for $s \geq 4m^2$ both $\phi(s; \cos\theta)$ and $\phi^{(2)}(s; \cos\theta)$ are analytic except for cuts starting at $\cos\theta = a_1(s) = \pm[1+8m^2/(s-4m^2)]$. In what follows we only make use of their analyticity within the ellipse $E[a_1(s) + \tau]$, where $a_1(s) + \tau < a_2(s)$; the ellipse must be just large enough for the cuts to penetrate. Starting from (3.6), and using (2.6) and (2.7), $G(s; \cos\theta)$ can be written as

$$G(s; \cos\theta) = -\frac{1}{8\pi^3}\oint_C dx_1 \times \oint_C dx_2\, H(x_1, x_2; \cos\theta)\phi(s; x_1)\phi^{(2)}(s; x_2), \quad (4.3)$$

C is the boundary of $E[a_1(s) + \tau]$ minus loops around the cuts starting at $\pm a_1(s)$. The singularities of $G(s; \cos\theta)$ occur according to the formula $\cos\theta = \cosh(\alpha_1 + \alpha_2)$, where $\alpha_1 = i\theta_1$, $\alpha_2 = i\theta_2$, and $\cos\theta_1$ and $\cos\theta_2$ vary over cut ellipses. The domain of analyticity of $G(s; \cos\theta)$ following from (4.3) consists of an ellipse and that part of the $D_2^{\pm}$ cuts which penetrate the ellipse. The ellipse is well within $E[a_3(s)]$, and, aside from the regions $D_2^{\pm}$, we know it is free of singularities. Hence, the only part of the contours of (4.3) which contribute to the discontinuity of $G(s; \cos\theta)$ are the loops around these cuts. The regular terms cancel on taking the discontinuity and we do not retain them. The discontinuity of $G(s; \cos\theta)$ [in the vicinity of $t_2^+(s)$] is, therefore, the same as the discontinuity of $C(s; \cos\theta)$, where

$$C(s; \cos\theta) = -\frac{1}{4\pi^3}\int_{a_1(s)}^{a_1(s)+\tau} dx_1 \times \int_{a_1(s)}^{a_1(s)+\tau} dx_2\, H(x_1, x_2; \cos\theta) \times [\phi(s; x_1 + i\epsilon) - \phi(s; x_1 - i\epsilon)] \times [\phi^{(2)}(s; x_2 + i\epsilon) - \phi^{(2)}(s; x_2 - i\epsilon)]. \quad (4.4)$$

The integrations around $-a_1(s)$ are accounted for by the extra factor of 2 in (4.4). The contributions to (4.3) given by integrating x_1 around the cut starting at $a_1(s)$, and x_2 around the cut starting at $-a_1(s)$ (and conversely), are singular near $t_2^-(s)$ and regular near $t_2^+(s)$ and, therefore, do not affect the discontinuity near $t_2^+(s)$.

The discontinuities of $\phi(s; x_1)$ and $\phi^{(2)}(s; x_2)$ appearing in (4.4) are simply the discontinuities across the $t \geq 4m^2$ cut transposed into angular variables. They are given by

$$\phi(s; x_1 + i\epsilon) - \phi(s; x_1 - i\epsilon) = \tfrac{1}{2}i\pi\{[x_1 - a_1(s)]/[x_1 - 1]\}^{\frac{1}{4}}G^t(x_1; s) \quad (4.5a)$$

and

$$\phi^{(2)}(s; x_2 + i\epsilon) - \phi^{(2)}(s; x_2 - i\epsilon) = \tfrac{1}{2}i\pi\{[x_2 - a_1(s)]/[x_2 - 1]\}^{\frac{1}{4}}G^t(x_2; s). \quad (4.5b)$$

The function $G^t(x; s)$ is just the absorptive part in the t channel, expressed in terms of the s channel center-of-momentum variables.[17] The factors $\{[x_1 - a_1(s)/[x_1 - 1]\}^{\frac{1}{4}}$ and $\{[x_2 - a_1(s)]/[x_2 - 1]\}^{\frac{1}{4}}$ are equivalent to $[(t - 4m^2)/t]^{\frac{1}{4}}$ in these variables.

Introducing the discontinuities (4.5) into (4.4) gives

$$C(s; \cos\theta) = \frac{1}{16\pi}\int_{a_1(s)}^{a_1(s)+\tau} dx_1 \times \int_{a_1(s)}^{a_1(s)+\tau} dx_2\, H(x_1, x_2; \cos\theta) \times \left(\frac{x_1 - a_1(s)}{x_1 - 1}\right)^{\frac{1}{4}}G^t(x_1; s)\left(\frac{x_2 - a_1(s)}{x_2 - 1}\right)^{\frac{1}{4}}G^t(x_2; s). \quad (4.6)$$

The discontinuity of $C(x; \cos\theta)$ is determined by replacing $H(x_1, x_2; \cos\theta)$ by its discontinuity, and this we get directly from the integral representation (2.10)

[17] Here $x = 1 + 2t/(s - 4m^2)$ and $G^t(x; s)$ is not to be confused with $G(s; x)$.

$$H(x_1, x_2; \cos\theta + i\epsilon) - H(z_1, z_2; \cos\theta - i\epsilon)$$

$$= \begin{cases} 0, & \cos\theta < x_1x_2 + [(x_1^2 - 1)(x_2^2 - 1)]^{\frac{1}{2}} \\ -4\pi^2(-1 + x_1^2 + x_2^2 + \cos\theta - 2x_1x_2\cos\theta)^{-\frac{1}{2}}, & \\ \cos\theta \geq x_1x_2 + [(x_1^2 - 1)(x_2^2 - 1)]^{\frac{1}{2}}. & \end{cases}$$

From this it follows that the integrations in (4.5) which contribute to the discontinuity do not cover the entire square $a_1(s) < x_1 < a_1(s) + \tau$, $a_1(s) < x_2 < a_1(s) + \tau$; for t near $t_2^+(s)$, $\rho(s, t)$ is given by

$$\rho(s, t) = i\frac{\pi}{2}\left(\frac{s - 4m^2}{s}\right)^{\frac{1}{2}}$$

$$\times \left\{\frac{1}{16\pi}\int_{a_1(s)}^{a_1(s)+\tau} dx_1 \int_{a_1(s)}^{a_1(s)+\tau} dx_2\,(-4\pi^2 i)\right.$$

$$\times \left[\{[x_1 - a_1(s)]/[x_1 - 1]\}^{\frac{1}{2}}G^{t_1}(x_1; s)\right.$$

$$\left.\left.\times \{[x_2 - a_1(s)]/[x_2 - 1]\}^{\frac{1}{2}}G^{t_2}(x_2; s)\right]\right.$$

$$\left.\times (-1 + x_1^2 + x_2^2 + \cos^2\theta - 2x_1x_2\cos\theta)^{-\frac{1}{2}}\right\}$$

$$\cos\theta \geq x_1x_2 + [(x_1^2 - 1)(x_2^2 - 1)]^{\frac{1}{2}},$$

$$\cos\theta = 1 + 2t/(s - 4m^2). \tag{4.7}$$

Each of the terms in the integrand of (4.7) are analytic within the domain of integration, hence $\rho(s, t)$ is analytic for $t_2^+(s) < t < t_3^+(s)$.

We already noted that $\rho(s, t)$ vanishes for $4m^2 \leq s \leq 16m^2$ and, since it is analytic, it must vanish for $t < t_3^+(s)$.

From (4.7) we see that the vanishing of $\rho(s, t)$ implies that $G^t(x; s) = 0$ within the region of integration. Now $G^t(x; s)$ is (up to a regular factor) the absorptive part in the t channel, x runs along the t cut ($t \geq 4m^2$), s is on the s cut ($s \geq 4m^2$), and these points are within the domain of analyticity of the absorptive part. We note that the absorptive part is not singular within $E[a_2(t)]$ and these points lie within this ellipse. Hence, the t-channel absorptive part vanishes identically, and, since the t-channel and s-channel absorptive parts are equal under the transformation $s \leftrightarrow t$, $A(s, \cos\theta)$ vanishes for $s \geq 4m^2$.

We now use the elastic unitarity condition again to show that the scattering amplitude is itself null when the absorptive part vanishes. This is easily seen using the partial wave form of the elastic unitarity condition. The vanishing of the absorptive part implies that each of the terms $\mathrm{Im}\,\phi_l(s)$ vanishes, hence $\phi_l(s)$ vanishes. It follows that $\phi(s; \cos\theta)$ vanishes. It vanishes throughout its domain of analyticity, hence it is the null function. This is what we set out to prove, the no-production assumption is incompatible with a nontrivial elastic scattering amplitude.

We have shown that elastic scattering and production processes are strongly connected. In particular, there can be no elastic scattering without production. Since (full) unitarity connects both types of amplitudes, this is not surprising. What is surprising is that (full) unitarity was never used. We used the *elastic* unitarity condition which involves only the elastic scattering amplitude. However, the union of the crossing symmetry and analyticity with the elastic unitarity condition does provide a connection between elastic unitarity and (full) unitarity. This connection may be viewed as a constraint on the (full) unitarity condition.

5. THE SCATTERING AMPLITUDE IN FIELD THEORY

The above proof of the necessity of production processes is now modified to account for the more restricted analytic properties which follow from the axioms of field theory. The basic organization of the proof is retained but there are essential differences in detail. The most significant modification is connected with the fact that the domain of analyticity of the scattering amplitude is so small that the Landau curves (for $s \geq 4m^2$) do not enter it. However, under the assumption that production processes are null, the domain can be enlarged using the elastic unitarity condition beyond the usual elastic region. Were it not for this enlargement, the proof would fail.

The domain of analyticity of the scattering amplitude is considered first. Although the actual domain is at present unknown, Mandelstam has shown that the elastic scattering amplitude $\phi(s, t, u)$ is analytic within the region $|stu| < 7168m^6$, except for the usual s, t, and u cuts starting at $4m^2$.[18] Since the detailed structure of the domain is so very important for what is to come we shall survey Mandelstam's proof. This is also valuable preparation for handling some of the details of the proof.

Within the framework of axiomatic field theory, Bogoluibov, Medvedev, and Polivanov[19] have derived dispersion relations for the scattering amplitude. Thereafter, a number of simpler derivation were given which extended the range of validity of dispersion relations. The derivation of Lehmann[20] is

[18] S. Mandelstam, Nuovo Cimento **15**, 658 (1960).

[19] N. N. Bogoluibov, B. V. Medvedev, and M. K. Polivanov, *Problems in the Theory of Dispersion Relations* (Institute for Advanced Studies, Princeton, New Jersey, 1959).

[20] H. Lehman, Nuovo Cimento **10**, 579 (1958); Suppl. Nuovo Cimento **14**, 153 (1959).

of particular interest for he provides not only the dispersion relations, but in the course of his derivation he shows that the scattering amplitude and its absorptive part have significant analyticity in the momentum transfer variable. We start with Lehmann's results and follow Mandelstam's derivation.

For the momentum transfer variable t (in the s channel) fixed in the interval $-28m^2 < t < 0$, the scattering amplitude is analytic in the complex s plane, except for the cuts $s \leq -t$ and $4m^2 \leq s$. If the scattering amplitude tends to zero sufficiently rapidly for all contours such that $|s| \to \infty$, the Cauchy representation for this domain involves only contributions from the cuts. The result is the dispersion relation

$$\phi(s, t, u) = \frac{1}{\pi}\int_{4m^2}^{\infty} ds' \frac{\operatorname{Im}\phi(s', t)}{s' - s} + \frac{1}{\pi}\int_{4m^2}^{\infty} du' \frac{\operatorname{Im}\phi(u', t)}{u' - u}\ ; -28m^2 \leq t.$$

For $s \geq 4m^2$, the imaginary part of the scattering amplitude is itself an analytic function in the t variable which we refer to as the absorptive part $A(s, t)$, or $A(s; \cos\theta)$, in terms of the center of momentum variables. Lehmann has shown that (for $s \geq 4m^2$) $A(s; \cos\theta)$ is analytic in the $\cos\theta$ plane within the ellipse $E[2a'(s)^2 - 1]$, where

$$a'(s) = (1 + 64m^4/sq^2)^{\frac{1}{2}}. \tag{5.1}$$

The ellipse $E[2a'(s)^2 - 1]$ is known as the "large Lehmann ellipse."

It would appear that the analyticity of the absorptive part could be combined with the above dispersion relation to provide a representation of the scattering amplitude analytic over a domain in the (complex) variables s and t. However, it is easily seen that, because the large Lehmann ellipse shrinks as s increases, there is no such domain. A simple way out of the difficulty is to restrict attention to a domain which shrinks in s as t grows and, conversely, so that arbitrarily large values of s do not occur in such a way that all analyticity in t is lost. Furthermore, since the crossing symmetry plays a significant part in what follows, the domain should be symmetric in s, t, and u. It is also required that, other than the usual s, t, and u cuts, no singularities occur within the domain. In this respect the domain is like the domain of the Mandelstam representation.

A domain satisfying these conditions has been determined by S. Mandelstam.[21] He considers the set of all points satisfying

$$|stu| < R \tag{5.2}$$

except for points on the cuts $s \geq 4m^2$, $t \geq 4m^2$, and $u \geq 4m^2$.[22] It has the property of decreasing in the s dimension as t increases, and, conversely. It is also a symmetric domain. Mandelstam has determined a value of R in (5.1) such that the scattering amplitude may be analytically continued into the domain. In terms of the s-channel center-of-momentum variables, (5.2) becomes

$$4\tilde{q}^4 s\,|1 - \cos^2\theta| < R. \tag{5.3}$$

The points satisfying (5.3) also satisfy

$$\cos\theta = \pm(1 + \zeta/4\tilde{q}^4 s)^{\frac{1}{2}}, \tag{5.4}$$

where $|\zeta| < R$. From this we see that, for $s \geq 4m^2$, a domain of the form of (5.1) is composed of two parts, one part about $\cos\theta = 1$, and the other part about $\cos\theta = -1$. However, for sufficiently small values of $s \geq 4m^2$ the parts overlap and the domain contains the entire interval $-1 \leq \cos\theta \leq +1$. We also note that the domain shrinks rapidly as s grows. Comparison with (5.1) shows that the domain shrinks by the factor $1/s$ faster than $a'(s)$. The shrinkage of the domain should get us around the difficulty mentioned above. Since (5.2) is symmetric, this is also the case in the t and u channels when examined in terms of their center-of-momentum variables.

The value of R is determined from the requirement that the projection of the domain $|stu| < R$ on the t plane (or $\cos\theta$ plane) is to be contained within the large Lehmann ellipse. A point satisfying (5.2), and hence (5.4), lies within this ellipse if the sum of the distances from the point to each of the foci is less than twice the semimajor axis. This sum is given by

$$\mathcal{S} = |\pm(1 + \zeta/4\tilde{q}^4 s)^{\frac{1}{2}} - 1| + |\pm(1 + \zeta/4\tilde{q}^4 s)^{\frac{1}{2}} + 1|. \tag{5.5}$$

It is clear that $\mathcal{S}$ is independent of the root chosen in (5.4). If we let $y = (1 + \zeta/4\tilde{q}^4 s)^{\frac{1}{2}}$, Eq. (5.6) becomes

$$\mathcal{S} = |1 + y| + |1 - y|,$$

and it follows that

$$\mathcal{S}^2 = 2 + 2\,|y|^2 + 2\,|1 - y^2|.$$

On returning to the original variables,

$$\mathcal{S}^2 = 2 + 2\,|1 + \zeta/4\tilde{q}^4 s|^2 + 2\,|\zeta/4\tilde{q}^4 s|. \tag{5.6}$$

If we denote ζ by $\zeta = re^{ik}$ with $r > 0$, we find that $\mathcal{S}^2$

[21] S. Mandelstam, Nuovo Cimento 15, 658 (1960).

[22] This choice of domain is primarily dictated by the fact that it works. Another possibility is $|st| < R$ but for our purposes this is less satisfactory.

takes on its maximum value for $k = 0$, that is, for $\zeta > 0$. Furthermore, the larger r is, the larger s^2 and, hence, s is. We conclude from this that all points of the form (5.4) lie within the large Lehmann ellipse, if the point with $\zeta = R$ lies within this ellipse. Another way of saying this is: All of the points of the form (5.4) lie within the ellipse if the largest real point (in the cos θ plane) is within the ellipse.

The maximum value of R is found by requiring that for each $s \geq 4m^2$, the point with $\zeta = R$ lies within the large Lehmann ellipse corresponding to that value of s. Hence, we require (for $s \geq 4m^2$)

$$4q^4 s\{1 - [2a'(s)^2 - 1]^2\} = R(s) \geq R,$$

where $a'(s)$ is given by (5.1). On introducing (5.1), this becomes

$$256m^4[s - 4m^4 + 256m^4/s] = R(s) \geq R. \quad (5.7)$$

The function $R(s)$ takes on its minimum value in the range $4m^2 \leq s$ when $s = 16m^2$. Placing $s = 16$ in (5.7), we find that the minimum of $R(s)$ is $7168m^6$. We take $7168m^6$ as the value of R, it is the largest value satisfying the requirements set above.

Since the domain $|stu| < 7168m^6$ is symmetric it follows immediately that the projection in the t and u channels are also contained within the corresponding large Lehmann ellipses.

We have still to show that it is possible to extend the scattering amplitude simultaneously in s and t into this domain. The domain determined by $|stu| < 7168m^6$ is bounded by a segment of an analytic hypersurface and the s, t, and u cuts. The cuts are also segments of analytic hypersurfaces. Using a formula due to Bergmann and Weil,[23] it is possible to give an integral representation for the most general function analytic within a domain whose boundary is composed of analytic hypersurfaces. The Bergmann–Weil representation involves integration over the intersections of these hypersurfaces, that is, over a domain of lower dimension than the boundary of the domain but contained in the boundary. In the case where the number of analytic hypersurfaces which compose the boundary equals the number of independent complex variables the Bergmann–Weil representation is little more than a repeated application of the Cauchy representation in each variable.

In this representation it is most convenient to use the variables s and ζ, instead of s and t, since the boundary of the domain is best described in terms of ζ. However, t (or cos θ) is not a single-valued function of ζ and this change of variable would not be analytic. Actually, there is no difficulty at this point as we know that the scattering amplitude is an even function of cos θ. Since it is analytic in cos θ, it must be an analytic function of $\cos^2\theta$. With $\cos^2\theta$ as the variable, there is no sign ambiquity, and the change of coordinates is analytic. In terms of $\zeta = -stu$, the scattering amplitude has the representation

$$\begin{aligned}\phi(s, t, u) &= \frac{1}{2\pi^2 i}\int_{4m^2}^{\infty} ds' \int_{|\zeta'|=7168m^6} d\zeta' \frac{\tilde{A}(s';\zeta')}{(s'-s)(\zeta'-\zeta)} \\ &+ \frac{1}{2\pi^2 i}\int_{4m^2}^{\infty} dt' \int_{|\zeta'|=7168m^6} d\zeta' \frac{\tilde{A}(t';\zeta')}{(t'-t)(\zeta'-\zeta)} \\ &+ \frac{1}{2\pi^2 i}\int_{4m^2}^{\infty} du' \int_{|\zeta'|=7168m^6} d\zeta' \frac{\tilde{A}(u';\zeta')}{(u'-u)(\zeta'-\zeta)}, \quad (5.8)\end{aligned}$$

where $\tilde{A}(s;\zeta)$ is the absorptive part in the s channel (similarly for the t and u channels) in terms of s and ζ. We note that the absorptive part is analytic in ζ, for all $s \geq 4m^2$, since the projection of $\{|\zeta| < 7168m^6\}$ on the cos θ plane was chosen to lie within the large Lehmann ellipse. Furthermore, the domain of analyticity of the absorptive part does not shrink with increasing s, as would be the case with the straightforward use of the large Lehmann ellipse. The difficulty connected with the shrinkage of the large Lehmann ellipse is thus circumvented.

The proof that $\phi(s, t, u)$ actually satisfies (5.8) is quite simple. The function $\phi(s, t, u)$ given by (5.8) and all its derivatives in t, at $t = 0$, agree with the forward scattering dispersion relations and the forward scattering dispersion relations for the various derivatives of the scattering amplitude found by Symanzik.[24] It follows that (5.8) is the desired analytic continuation of the elastic scattering amplitude.

The analytic representation (5.8) replaces the Mandelstam representation (2.3) in field theory. The proof of the necessity of production must be modified to suit this representation. However, as we shall see, the representation will itself require change in the process. In particular, the Landau curves which were so important to the previous case do not even enter the domain $|stu| < 7168m^6$ for $s \geq 4m^2$. As a result, (5.8) does not have contributions from the double discontinuity functions $\rho(s, t)$, as in the case of the Mandelstam representation. We return to this question after we have examined the Landau singularities as they occur in field theory.

[23] For a concise survey of the Bergmann and Weil formulas see F. Sommer, Math. Ann. **125**, 172 (1952).

[24] K. Symanzik, Phys. Rev. **105**, 743 (1957).

6. THE LANDAU SINGULARITIES IN FIELD THEORY

The field-theoretic scattering amplitude can be analytically continued onto the second sheet of the square-root branch point at $s = 4m^2$ using essentially the same method as previously used in connection with the Mandelstam representation. As before, we introduce the partial wave amplitudes

$$\phi_l(s) = \frac{1}{2}\int_{-1}^{+1} d\cos\theta P_l(\cos\theta)\phi(s;\cos\theta);$$

$$l = 0, 1, 2 \cdots, \qquad (6.1)$$

but, in the present case, $\phi_l(s)$ is analytic in s within the domain determined by

$$|s(s - 4m^2)| < 4R = 4(7168m^6), \qquad (6.2)$$

except for the cuts $4m^2 \leq s$ and $-\infty < s \leq 0$.

The elastic unitarity condition (in partial wave form) and the reality of the scattering amplitude are used to define the merimorphic functions $F_l(s)$, and $G_l(s)$, satisfying

$$\phi_l(s) = F_l(s) + i\pi/4[(s - 4m^2)/s]^{\frac{1}{2}}G_l(s),$$

for s fulfilling (6.2). The singularities of $F_l(s)$ are canceled by the $G_l(s)$ term and, as before, we disregard these singularities as inessential to the proof. We can, therefore, regard

$$\phi_l^{(2)}(s) = F_l(s) - i\pi/4[(s - 4m^2)/s]^{\frac{1}{2}}G_l(s),$$

the analytic extension of $\phi_l(s)$ to the second sheet, as analytic, for s satisfying (6.2), and off the cut $s \leq 0$.

Summation of the partial wave amplitudes $\phi_l(s)$, $\phi_l^{(2)}(s)$, and $G_l(s)$, presents more difficulty than was encountered with the Mandelstam representation. For s fixed, the partial wave series converges uniformly and absolutely to

$$\phi(s;\cos\theta) = \sum_{l=0}^{\infty}(2l + 1)\phi_l(s)P_l(\cos\theta), \qquad (6.3)$$

within the largest ellipse with foci ± 1 which can be fitted into the domain of analyticity of $\phi(s;\cos\theta)$. We seek that ellipse.

Previously, we noted that the projection (for s fixed) satisfying $|stu| < R$ could be fitted within an ellipse in the $\cos\theta$ variable (with foci ± 1) if the largest (real) value of the projection was contained within the ellipse. In this case we are interested in the largest ellipse which can be fitted into the projection. The function $\mathfrak{s}$ given by (5.5) provides a criterion for this case also. The projection contains an ellipse with foci ± 1 if the largest pure imaginary value of $\cos\theta$ satisfying $|stu| < R$ is greater than the magnitude of the semiminor axis of the ellipse. From (5.6) we see that $\mathfrak{s}$ takes on its minimum value for $k = \pi$, and the criterion follows from the usual arguments concerning the sum of the distances from a point on an ellipse to the foci. The semiminor axis of an ellipse $E[\hat{a}_1(s)]$ is located at

$$\cos\theta = i[\hat{a}_1(s)^2 - 1]^{\frac{1}{2}}.$$

Using this with the criterion, we obtain

$$\hat{a}(s) = R^{\frac{1}{2}}/2\,|\bar{q}^2 s^{\frac{1}{2}}|. \qquad (6.4)$$

For large values of $|s|$ we see that $\hat{a}_1(s) < 1$, and it follows that, for $|s(s - 4m^2)| > 4R$, it is not possible to fit an ellipse within the region given by $|stu| < R$. Hence, a nonvanishing ellipse can be found only for values of s satisfying (6.2), the condition for analyticity of partial wave amplitudes as functions of the s variable.

We have noted that the region $|stu| < R$ is composed of a component about $\cos\theta = 1$, and another about $\cos\theta = -1$, and that the parts do not overlap for large $|s|$. Our result is a symptom of this disjointness of the components of the projection of the $|stu| < R$.

On the other hand, when $|s|$ is sufficiently small, the ellipse may include parts of the t and u cuts; hence $\phi(s;\cos\theta)$ is not analytic throughout $E[\hat{a}_1(s)]$. If we denote the initial points of the t and u cuts (in the $\cos\theta$ plane) by $\pm\alpha(s)$, we have $\phi(s;\cos\theta)$ analytic within $E[a_1(s)]$, where $a_1(s) = \cosh \mathrm{Re}\ \alpha(s)$. The partial wave series (6.3) converges uniformly and absolutely within $E[\hat{a}_1(s)]$ when the cuts do not penetrate the ellipse. When they do, convergence is restricted to the ellipse $E[a_1(s)]$.

For s satisfying (6.2), and not on the cut $s \leq 0$, the Legendre coefficients $F_l(s)$ can be summed to the function $F(s;\cos\theta)$, where

$$F(s;\cos\theta) = \sum_{l=0}^{\infty}(2l + 1)F_l(s)P_l(\cos\theta),$$

within the relevant ellipse $E[\hat{a}_1(s)]$, or $E[a_1(s)]$. For s satisfying (6.2), and not on the cut $s \leq 0$, the function $G(s;\cos\theta)$, given by

$$G(s;\cos\theta) = \sum_{l=0}^{\infty}(2l + 1)G_l(s)P_l(\cos\theta),$$

converges uniformly and absolutely within the ellipse $E[\hat{a}_2(s)]$, where $\hat{a}_2(s) = 2\hat{a}_1(s)^2 - 1$, when $\phi(s;\cos\theta)$ converges within $E[\hat{a}_1(s)]$. When $\phi(s;\cos\theta)$ converges within $E[a_1(s)]$, the series converges within $E[a_2(s)]$, where $a_2(s) = 2a_1(s)^2 - 1$. The relationship between the ellipses $E[\hat{a}_1(s)]$, and

$E[\hat{a}_2(s)]$, and $E[a_1(s)]$, and $E[a_2(s)]$, is once again a consequence of the elastic unitarity condition.

The function $F_l(s)$ can be removed from the expressions for $\phi_l(s)$ and $\phi_l^{(2)}(s)$ to give

$$\phi^{(2)}(s; \cos\theta) = \phi(s; \cos\theta) - \tfrac{1}{2}i\pi[(s - 4m^2)/s]^{\frac{1}{2}}G(s; \cos\theta). \quad (6.5)$$

Since this result depends on the elastic unitarity condition it is at first valid only for s in the elastic region. However, as both terms on the right-hand side are analytic for s satisfying (6.2), except for the cuts $s \leq 0$ and $s \geq 4m^2$, and for $\cos\theta$ within the intersection of the domain $|stu| < R$ with the relevant ellipse $E[\hat{a}_2(s)]$ or $E[a_2(s)]$, Eq. (6.5) is valid throughout this region. The function $\phi^{(2)}(s; \cos\theta)$ is, therefore, the analytic continuation of the scattering amplitude onto the second sheet of the square-root singularity at $s = 4m^2$. It has the cut $-\infty < s \leq 0$, as well as the s, t, and u cuts.

The domain of analyticity we have found for $\phi^{(2)}(s; \cos\theta)$ is more complicated than was the case for the Mandelstam representation. For $|s|$ sufficiently small (off the cut $s \leq 0$), the t and u cuts penetrate so far into the ellipse $E[\hat{a}_1(s)]$ that the ellipse $E[a_2(s)]$ itself is contained within $E[\hat{a}_1(s)]$. We can also analytically continue $G(s; \cos\theta)$ into the ellipse $E[a_3(s)]$, if $E[a_3(s)]$ is contained within $E[\hat{a}_1(s)]$. The method is the same as was used in connection with the Mandelstam representation. Similarly, the regions $D^{\pm}$ given by (3.7) may contain singularities of $G(s; \cos\theta)$ within $E[a_3(s)]$. If not all of $E[a_3(s)]$ is contained within $E[\hat{a}_1(s)]$, the continuation stops at $E[\hat{a}(s)]$; the parts of the regions $D_2^{\pm}$ which lie within $E[\hat{a}_1(s)]$ may contain singularities of $G(s; \cos\theta)$.

If $E[a_3(s)]$ is contained within $E[\hat{a}_1(s)]$, this method of analytic extension can be used again to continue into $E[a_4(s)]$, if $E[a_4(s)]$ is contained within $E[\hat{a}_1(s)]$. The regions $D_2^{\pm}$ and $D_3^{\pm}$ must be removed from $E[a_4(s)]$ as $G(s; \cos\theta)$ may be singular within these regions. If $E[a_4(s)]$ is not entirely contained within $E[\hat{a}_1(s)]$, the continuation process ends at $E[\hat{a}_1(s)]$ from which $D_2^{\pm}$ and $D_3^{\pm}$ must be removed. For any value of s satisfying (6.2), the extension process stops after a finite number of steps. However, if we take $|s|$ sufficiently small we can reach any ellipse $E[a_n(s)]$ and have a part of $D_n^{\pm}$ as a (potential) singular region for the function $G(s; \cos\theta)$.

It follows that $\phi^{(2)}(s; \cos\theta)$ is analytic within $E[\hat{a}_1(s)]$ for s satisfying (6.2), except for the s, t, and u cuts, the cut $-\infty < s \leq 0$, and those parts of the regions $D_n^{\pm}$ which fall within $E[\hat{a}_1(s)]$. The domain of analyticity of $G(s; \cos\theta)$ can be extended beyond $E[\hat{a}_1(s)]$, but that is left for later.

The initial points of the regions $D_m^{\pm}$ determine the Landau curves. Hence, it is possible to reach any of the Landau singularities by taking $|s|$ sufficiently small. However, when $s \geq 4m^2$, the Landau curves do not enter the domain[25] of analyticity of $\phi(s; \cos\theta)$. This follows from the fact that $E[\hat{a}(s)]$ is contained within the domain determined by $|stu| < R$ which is, for $s \geq 4m^2$, itself contained within the large Lehmann ellipse for that value of s. The Landau singularities do not occur within the large Lehmann ellipse as the absorptive part, and, hence $G(s; \cos\theta)$, is analytic within this ellipse.

The Landau curves for $s \geq 4m^2$ played a central role in the proof using the Mandelstam representation; however, in the present case we find that the domain of analyticity is too small to reach these curves, and the proof will not work in this case, unless the domain $|stu| < R$ with $R = 7168m^6$ can be enlarged to include the relevant parts of the Landau curves. In the following we use the elastic unitarity condition to enlarge the domain. However, we first take a short look at the Landau curve for $n = 2$, given by the equation $(s - 4m^2)(t - 4m^2) = 64m^4$, to get an idea of just how much we have to enlarge the domain. The values of $\tilde{R} = |stu|$ on the curve $(s - 4m^2)(t - 4m^2) = 64m^4$ are given by

$$\tilde{R} = s\left(16m^2 + \frac{64m^4}{s - 4m^2}\right)\left(12m^2 + \frac{64m^4}{s - 4m^2} + s\right).$$

For $s \geq 4m^2$, we note some values of $\tilde{R}$: for $s = 4m^2$, $\tilde{R} = \infty$; for $s = 6m^2$, $\tilde{R}$ exceeds $12500m^6$; for $s = 8m^2$, $\tilde{R} = 9216m^6$; for $s = 10m^2$, $\tilde{R} = 8720m^6$; for $s = 12m^2$, $R = 9216m^6$; for $s > 12m^2$, R becomes still greater. As expected each of these values exceeds $7168m^6$. We conclude that R must be enlarged from $7168m^6$ to a value in excess of $8720m^6$, if a proof similar to the proof used in connection with the Mandelstam representation is to be used.

7. THE NECESSITY OF PRODUCTION PROCESSES: THE CASE OF FIELD THEORY FOR EQUAL MASSES

Mandelstam[26] has given a method for enlarging the domain of analyticity of the scattering amplitude using the elastic unitarity condition in somewhat the same way that Zimmermann used it to study the Landau singularities. The proof that scattering

[25] By domain we mean the domain before the cuts are removed.

[26] S. Mandelstam, Nuovo Cimento **15**, 658 (1960). Mandelstam's method does not work in this case when the elastic unitarity condition obtains only within the elastic interval $4m^2 \leq s \leq 16m^2$.

implies production is carried out *"reductio ad absurdum"* starting with the working assumption that production processes are null. We are then free to apply the elastic unitarity condition beyond the usual elastic region $4m^2 \leq s < 16m^2$. The domain of analyticity can then be enlarged sufficiently for the proof to be carried out.

We start the enlargement process by noting that for $s \geq 4m^2$ (but sufficiently small) the scattering amplitude is analytic within an ellipse which is larger than the small Lehmann ellipse for that value of s. It then follows from the elastic unitarity condition that the absorptive part is analytic in an ellipse which is larger than the large Lehmann ellipse. Since the value of R is determined by the domain of analyticity of the absorptive part, it may be that R can be enlarged using the analyticity in the larger ellipse. We found that $R = 7168m^6$ was determined at $s = 16m^2$; that is, the boundary of the projection of the domain determined by $|stu| < 7168m^6$ touches the large Lehman ellipse at $s = 16m^2$. If, by using the elastic unitarity condition for s about $16m^2$, the absorptive part is analytic in a larger ellipse, R can be increased. However, if the elastic unitarity condition is restricted to the interval $4m^2 \leq s < 16m^2$, the method does not work, since we could not claim that the absorptive part is analytic beyond the large Lehmann ellipse. As we are free to use the elastic unitarity condition for $s > 16m^2$, and since it follows from this that the absorptive part is analytic beyond the large Lehmann ellipse for s about $16m^2$, we can use Mandelstam's method even though it does not apply in the usual case.

We found above that the ellipse $E(\hat{a}_1(s))$, with s satisfying (6.2), can be fitted into the projection of the domain determined by $|stu| < 7168m^6$. Both $\phi(s; \cos\theta)$ and $\phi^{(2)}(s; \cos\theta)$ are analytic within this ellipse, for $s \geq 4m^2$ (but sufficiently small), except for the images of the t and u cuts. For $s \geq 4m^2$ and sufficiently large, the cuts no longer enter the ellipse $E[\hat{a}_1(s)]$ since, for increasing s, the ellipse shrinks faster than the initial points of the cuts move in toward the foci. In this case the absorptive part is analytic within $E[\hat{a}_2(s)]$, where $\hat{a}(s) = 2\hat{a}_1(s)^2 - 1$. When the initial points of the t and u cuts penetrate $E[\hat{a}_1(s)]$ the absorptive part is analytic within the ellipse $E[\tilde{a}_2(s)]$, where

$$\tilde{a}_2(s) = \cosh[\alpha(s) + \hat{\alpha}(s)], \tag{7.1}$$

and where $\alpha(s) = \cosh^{-1} a_1(s)$, and $\hat{\alpha}(s) = \cosh^{-1} \hat{a}_1(s)$, except for cuts with initial points $\pm a_2(s)$. Equation (7.1) follows from (2.12) using $\phi(s; \cos\theta)$ analytic within $E[a_1(s)]$, and $\phi^{(2)}(s; \cos\theta)$ analytic within $E[\hat{a}_1(s)]$, in consequence of the elastic unitarity condition.

We use the usual trigonometric identities to rewrite (7.1) as

$$\tilde{a}_2(s) = \hat{a}_1(s)a_1(s) + [\hat{a}_1(s)^2 - 1]^{\frac{1}{2}}[a_1(s)^2 - 1]^{\frac{1}{2}}. \tag{7.2}$$

The absorptive part is analytic within $E[\hat{a}_2(s)]$, or $E[\tilde{a}_2(s)]$, whichever is relevant, except for the cuts beginning at $\pm a_2(s)$ and running off to $\pm\infty$. For each value $s \geq 4m^2$, one of the ellipses $E[a_2(s)]$, $E[\tilde{a}_2(s)]$, or $E[\hat{a}_2(s)]$, is used to determine R. The largest of the ellipses is used, providing, of course, that it is relevant. The large Lehmann ellipse is always relevant, but for some values of s, particularly for s not too much greater than $4m^2$, $E[\tilde{a}_2(s)]$ and $E[\hat{a}_2(s)]$ may be larger.

We proceed as before by fitting the projections of the domain determined by $|stu| < R$ into these ellipses for $s \geq 4m^2$. We remember that the largest domain of this form (that can be projected into an ellipse with foci ± 1) is determined by requiring that the largest real value of the domain occurs at the endpoint of the semimajor axis of the ellipse. Hence R, in the case where $E[\tilde{a}_2(s)]$ is relevant, is dominated by $\tilde{R}_1(s)$, which is given by

$$\tilde{R}_1(s) = 4\bar{q}^4 s\,|1 - \tilde{a}_2(s)^2|. \tag{7.3}$$

Placing (7.2) into (7.3) gives

$$\tilde{R}_1(s) = 4\bar{q}^4 s\{\hat{a}_1(s)[a_1(s)^2 - 1]^{\frac{1}{2}} + a_1(s)[\hat{a}_1(s)^2 - 1]^{\frac{1}{2}}\}^2. \tag{7.4}$$

Since $a_1(s) = 1 + 2m^2/\bar{q}^2$, and $\hat{a}_1(s) = 16\cdot 7^{\frac{1}{2}} m^3/\bar{q}^2 s^{\frac{1}{2}}$, $\tilde{R}_1(s)$ is given by

$$\tilde{R}_1(s) = \left[64\cdot 7^{\frac{1}{2}}\left(\frac{m^2}{\bar{q}^2} + \frac{m^4}{\bar{q}^4}\right)^{\frac{1}{2}} + \left(1 + 2\frac{m^2}{\bar{q}^2}\right)\left(7168 - \frac{4\bar{q}^4 s}{m^6}\right)^{\frac{1}{2}}\right] m^6. \tag{7.5}$$

We note that $\tilde{R}_1(s)$ is monotonically decreasing for $4m^2 \leq s < 7168m^6/4\bar{q}^4$. The last restriction is just (6.2), the condition that must be satisfied if an ellipse is to be fitted into the projection of the domain determined by $|stu| < 7168m^6$.

The ellipse $E[\tilde{a}_2(s)]$ is relevant up to $s = 28m^2$. By simple computation at $s = 28m^2$, we find $\hat{a}_1(28m^2) = 1.335$ and $a_1(28m^2) = 1.333$; the initial points of the cuts just penetrate the ellipse $E[\hat{a}_1(28m^2)]$, and it follows that $E[\hat{a}_1(28m^2)]$ is still relevant for the absorptive part. For $s > 28m^2$ (actually a value slightly greater than $28m^2$), the ellipse $E[\hat{a}_2(s)]$ is relevant and $E[\tilde{a}_2(s)]$ is not. We note that $E[\tilde{a}_2(28m^2)]$ is larger than $E[a_2(28m^2)]$.

Furthermore, the ellipse $E[\hat{a}_2(30m^2)]$ is larger than the large Lehmann ellipse at $s = 30m^2$. In the interval $28m^2 < s < 30m^2$, $\hat{R}$ is dominated by $\hat{R}_1(s)$ which is given by

$$\hat{R}_1(s) = 4\bar{q}^4 s\ |1 - \hat{a}_2(s)^2|. \tag{7.6}$$

Using $\hat{a}_2(s) = 2\hat{a}_1(s)^2 - 1$, we find

$$\hat{R}_1(s) = 16\bar{q}^4 s\ |\hat{a}_1(s)^4 - \hat{a}_1(s)^2|. \tag{7.7}$$

Placing $\hat{a}_1(s) = 16\cdot 7^{\frac{1}{2}} m^3/\bar{q}^2 s^{\frac{1}{2}}$ into (7.7) gives

$$\hat{R}_1(s) = 4(7168)m^6\ |7168m^6/4\bar{q}^4 s - 1|. \tag{7.8}$$

The new value of R is dominated by $R_1(s)$ where

$$R_1(s) = \max\{R(s); \tilde{R}_1(s)\}, \quad 4m^2 \leq s < \bar{s}, \tag{7.9a}$$

$$R_1(s) = \max\{R(s); \hat{R}_1(s)\}, \quad \bar{s} \leq s < \hat{s}, \tag{7.9b}$$

$$R_1(s) = R(s), \quad \hat{s} < s, \tag{7.9c}$$

and $R(s)$ is given by (5.7). The value of $\bar{s}$ is determined by $\hat{a}_1(\bar{s}) = a(\bar{s})$, and, as we have seen, $\bar{s}$ is slightly greater than $28m^2$. The value of s at which the ellipse $E[\hat{a}_2(s)]$ becomes smaller than the large Lehmann ellipse is denoted by $\hat{s}$ and it is greater than $30m^2$.

The minimum value of $R_1(s)$ is not found at a stationary point, as was previously the case. However, we do know that $\tilde{R}_1(s)$ and $\hat{R}_1(s)$ are monotonically decreasing in the interval $4m^2 \leq s < 7168^6/4\bar{q}^4$, and, on the other hand, $R(s)$ is monotonically increasing for $s > 16m^2$. From this it follows that $R = R_1(\bar{s})$, where $\bar{s}$ is the value at which $R_1(s) = R(s)$ for all $s \geq \bar{s}$. Of course, we can use a smaller value for R; for example, the value of $R(s)$ at some $s < \bar{s}$. At $s = 30m^2$, $R(30m^2) < R < R_1(s)$ for all $s \geq 4m^2$. Calculation gives $R(30m^2) = 8840m^6$, and $R_1(30m^2) = 11900m^6$. Hence, we can use $8840m^6$ in place of $7168m^6$. We note that the new value exceeds $8720m^6$, hence, the Landau singularity $t_2^+(s)$ will enter the domain determined by $|stu| < 8840m^6$ for s in the range $s \geq 4m^2$.

We could use this domain as the starting point for repeating the extension process. An even larger ellipse can be fitted into this domain, however, the present domain is satisfactory for completing the proof that scattering implies production. We note that the value $8840m^6$ is of no significance other than that it exceeds $8720m^6$.

The scattering amplitude has a Bergmann–Weil representation in the new domain. It is not sufficient to replace $7168m^6$ by $8840m^6$ in (5.8) since the absorptive part $\tilde{A}(s'; \zeta)$ is singular on the Landau curve $t_2^+(s)$ within $|\zeta| = 8840m^6$. The contributions from contours around the cuts of the absorptive part must be included in (5.8). The cuts begin at $+a_2(s)$ $[-a_2(s)]$ and run off to the right (and left). We include contributions only from these parts of the cut which enter $|\zeta| = 8840m^6$. The contributions from the upper and lower sides of the cuts can be combined, in the usual way, in the form of the discontinuity across the cuts. Hence we write

$$\begin{aligned}
\phi(s, t, u) = {} & \frac{1}{\pi^2}\int_{4m^2} ds' \int_{4m^2} dt' \frac{\rho(s', t')}{(s' - s)(t' - t)} \\
& + \frac{1}{\pi^2}\int_{4m^2} dt' \int_{4m^2} du' \frac{\rho(t', u')}{(t' - t)(u' - u)} \\
& + \frac{1}{\pi^2}\int_{4m^2} du' \int_{4m^2} ds' \frac{\rho(u', s')}{(u' - u)(s' - s)} \\
& + \frac{1}{2\pi^2 i}\int_{4m^2}^{\infty} ds' \int_{|\zeta'|=8840m^6} d\zeta' \frac{\tilde{A}(s'; \zeta')}{(s' - s)(\zeta' - \zeta)} \\
& + \frac{1}{2\pi^2 i}\int_{4m^2}^{\infty} dt' \int_{|\zeta'|=8840m^6} d\zeta' \frac{\tilde{A}(t'; \zeta')}{(t' - t)(\zeta' - \zeta)} \\
& + \frac{1}{2\pi^2 i}\int_{4m^2}^{\infty} du' \int_{|\zeta'|=8840m^6} d\zeta' \frac{\tilde{A}(u'; \zeta')}{(u' - u)(\zeta' - \zeta)},
\end{aligned} \tag{7.10}$$

where $\rho(s', t')$ is the double discontinuity of the scattering amplitude within the domain, and $\tilde{A}(s'; \zeta')$ is the absorptive part as described previously. The functions $\rho(s', t')$, $\rho(t', u')$, and $\rho(u', s')$ are all equal and symmetric since $\phi(s, t, u)$ is totally symmetric. The integrations over these functions vanish over part of the indicated regions as we see below. The situation at this point is quite similar to the situation we started with in the case of the Mandelstam representation. From this point the proof goes the same way.

We are now ready to complete the proof following the arguments used in Sec. 3; we borrow results from there freely. The essential difference between our present situation, and that encountered with the Mandelstam representation, is the presence of the boundary $|stu| = 8840m^6$. Nevertheless, within this boundary we find the same incompatibility of crossing symmetry and the two-particle unitarity condition on which the previous proof was based.

Under the working assumption that production process are null, we can use the elastic unitarity condition for all $s \geq 4m^2$. Actually, we need it only for s below the second production threshold ($s = 36m^2$), but, since our results are not essentially strengthened by working with this weaker condition, we do not stress this point. With the elastic unitarity condition valid beyond $s = 16m^2$, the partial wave amplitudes (6.1) are defined for s satisfying (6.2) with $R = 8840m^6$. It follows that the analytic continua-

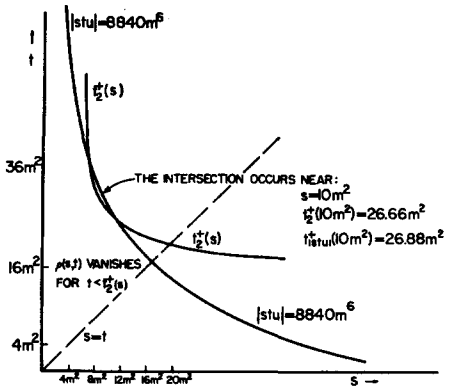

FIG. 2. The support of $\rho(s, t)$ in field theory.

tion $\phi^{(2)}(s; \cos\theta)$ is defined over a larger region, and the continuation is the cos θ variable can be pushed further. The function $G(s; \cos\theta)$ is analytic within an ellipse, which for each $s \geq 4m^2$ contains the entire domain $|stu| < 8840m^6$, except for the regions of singularities $D_2^{\pm}$ given by (3.8), some of which may occur within the ellipse. Only points of $D_2^{\pm}$ fall within $|stu| = 8840m$, the other singularities of the absorptive part occur outside the enlarged domain. The question of whether further enlargement of the domain will enable other singularities of the absorptive part to enter the domain of analyticity of the scattering amplitude does not concern us. The singularities already present are incompatible with the crossing symmetry.

The double density function $\rho(s, t)$ of (7.10) is given by (4.1) at points within the domain determined by $|stu| < 8840m^6$. For such points (4.2) is also valid, hence $\rho(s, t)$ is proportional to the discontinuity of the absorptive part. For $s \geq 4m^2$ and within the domain, the absorptive part can be singular only within the regions $D_2^{\pm}$ which, in this case, reduce to cuts in the cos θ plane.

Wherever the absorptive part is analytic, the double density function vanishes. It follows that $\rho(s, t)$ vanishes for $4m^2 \leq t < t_2^+(s)$, where $t_2^+(s) = 16m^2 + 64m^4/(s - 4m^2)$. We remember that $t_2^+(s)$ is the first Landau singularity occurring for $s \geq 4m^2$ and $t \geq 4m^2$. Its points correspond to the initial points of the cut D_2^+ as s varies.

The Landau curve $t_2^+(s)$ is not symmetric in s and t. It follows that the (potential) support [i.e., the region where $\rho(s, t)$ may be nonvanishing] is not symmetric although the crossing symmetry requires that $\rho(s, t)$ be symmetric; see Fig. 2. The situation differs from that encountered with the Mandelstam representation by the presence of the presence of the boundary $|stu| = 8840m^6$. However, this boundary is symmetric in s and t; hence, it cannot account for the asymmetry of $\rho(s, t)$.

It is essential to verify that the Landau curve $t_2^+(s)$ actually enters the domain $|stu| < 8840m^6$. We find that it does enter about $s = 10m^2$. If we look at the curve $|stu| = 8840m^6$ in the vicinity of $t = 10m^2$, we do not find a Landau singularity. The singularities of the absorptive part are not symmetric under the interchange of s and t within the symmetric boundary $|stu| = 8840m^6$.

Thus, once again the no-production assumption has placed us in the desired circumstance. Either we renounce the possibility that production processes can be null (in the presence of nontrivial scattering), or we admit the possibility that $\rho(s, t)$ vanishes within the relevant part of $t > t_2^+(s)$. In either case it follows that nontrivial scattering implies production, for, if the second alternative obtains, it follows that the absorptive part vanishes. The arguments are given for the case of the Mandelstam representation above and no complications arise from the domain restrictions. It then follows from the elastic unitarity condition that the scattering amplitude is also null. Hence, we find that scattering implies production.

ACKNOWLEDGMENTS

The author would like to note the many helpful discussions with Professor John Toll, Professor Young Kim, and Dr. Alex Dragt during the course of this work.

Parts of the investigation were begun at the Institute for Theoretical Physics, Lund, Sweden and the author wishes to thank Professor Torsten Gustafson and Professor Gunnar Källén for their hospitality.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 6, NUMBER 4 APRIL 1965

Classification of Three-Particle States According to SU_3

A. J. DRAGT
Institute for Advanced Study, Princeton, New Jersey
(Received 19 June 1964; final manuscript received 8 September 1964)

It is shown that the set of states for three noninteracting particles can be put into a one-to-one correspondence with the set of irreducible representations of SU_3. This classification leads to three-particle angular momentum states which treat all particles on an equal footing. The states exhibit the maximum localization compatible with a given total energy, momentum, and angular momentum. Only nonrelativistic particles are treated.

INTRODUCTION

ELEMENTARY processes involving three or more interacting particles exhibit an extremely rich and complicated structure. In view of this complexity, it is important to at least have a complete kinematic understanding of systems composed of several noninteracting particles. For example, noninteracting particle states may be used in the case of short-range interactions to form a complete set of basis vectors for an S matrix formulation of scattering theory. In the complementary bound-state problem, regular and irregular free-particle solutions are of use in the generation of trial wavefunctions.

In any classification of multiparticle states it is important to diagonalize those variables which are known to be constants of motion from general invariance principles. Thus plane-wave states $|q_1q_2\cdots q_n\rangle$ diagonal in the individual 4-momenta q_i and the total 4-momentum $Q = \sum q_i$ are useful in exploiting translational invariance. To benefit from rotational invariance as well, it is necessary to abandon plane-wave states in favor of states diagonal in the total angular momentum. A beautiful discussion of how such states may be constructed as linear combinations of plane-wave states has been given for the case of three particles by Wick.[1] He forms states $|QJJ_z,\ wjj_z\rangle$ which are diagonal in the total 4-momentum and the angular momentum about the center of mass. In addition, his states are also diagonal in the total energy w and angular momentum j of particles 1 and 2 in their center-of-mass frame.

Besides diagonalizing J, the Wick states enjoy a further advantage over plane-wave states in that the particles are partially localized. A knowledge of interparticle distance is very important when dealing with short-range interactions. For example, state vectors with moderate values of w and large j describe a situation in which particles 1 and 2 are almost always very far apart and therefore very unlikely to interact. An elaboration of this line of reasoning leads to cluster decomposition properties for the S matrix.[2]

Unfortunately, the type of localization exhibited by Wick states is not always convenient since the third particle is treated asymmetrically. Thus, while the spatial correlation between particles 1 and 2 is clearly defined, that between particles 1 and 3 or 2 and 3 is not quite so evident and must be obtained with the aid of recoupling coefficients.

There are instances in which it would be advantageous to have explicit information about all interparticle correlations simultaneously. We may cite as an example the study on a Dalitz plot of particle distributions arising from three-body decays. There one asks for those distributions in three-body momentum space which are compatible with a given total angular momentum and which also imply a high degree of spatial correlation among all three particles.

From what has been said, it is evident that it would be useful to have a method for constructing multiparticle angular momentum states which treat all particles on an equal footing. Major advances in this direction have been made for the case of three nonrelativistic particles in a series of important papers by Smith.[3,4] In his first paper, he discusses the three-body operators which commute with the total energy and momentum, and shows that they generate the six-dimensional orthogonal group $\mathcal{O}_6$.[5] From $\mathcal{O}_6$ he is able to select a subgroup $\mathfrak{D}$ which treats all three particles symmetrically. In a second paper he applies operators from $\mathfrak{D}$ to the problem

[1] G. C. Wick, Ann. Phys. **18**, 65 (1962).

[2] E. H. Wichmann and J. H. Crichton, Phys. Rev. **132**, 2788 (1963).

[3] F. T. Smith, Phys. Rev. **120**, 1058 (1960).

[4] F. T. Smith, J. Math. Phys. **3**, 735 (1962).

[5] In general, script letters will be used to denote groups or group elements; block letters will be used for the corresponding Lie algebras.

of three particles confined to a plane and obtains a complete solution for this special case.

It is the purpose of this paper to extend the results of Smith to the general case of three particles in space. Sections 1 and 2 comprise a review of the pertinent results on $\mathcal{O}_6$ and $\mathfrak{D}$ contained in Smith's work. In Sec. 3 it is shown that $\mathfrak{D}$ is isomorphic to the three-dimensional unitary group U_3. Section 4 provides a résumé of well-known facts concerning representations of U_3 and its unimodular subgroup SU_3. Section 5 is devoted to the choice of a symmetric three-particle coordinate system. The stage is then set for Sec. 6 where it is shown that three-particle states can be completely classified according to their transformation properties under the action of the SU_3 subgroup of $\mathfrak{D}$. That is, the set of three-particle states can be put into one-to-one correspondence with the set of irreducible representations of SU_3. Section 7 treats some aspects of the embedding of SU_3 in $\mathcal{O}_6$. A final section briefly treats the application of SU_3 states to three-body decays.

Our discussion in this paper is limited to non-relativistic particles. The relativistic case will be treated in a second paper.

1. THE INVARIANCE GROUP FOR THREE PARTICLES

A. The Introduction of O_6

The quantal specification of a three-particle state requires nine independent commuting operators. It is natural to take four of them to be the total 4-momentum Q. One then asks for those operators which commute with Q. In the case of equal-mass particles, this is equivalent to asking for the real linear transformations which leave invariant the forms

$$\mathbf{q}_1 + \mathbf{q}_2 + \mathbf{q}_3 = \mathbf{Q} \tag{1.1}$$

and

$$q_1^2 + q_2^2 + q_3^2 = 2mT. \tag{1.2}$$

The general mass case is treated in an Appendix. Let $\mathbf{P}$ be a nine-dimensional vector with the components $\mathbf{q}_1$, to $\mathbf{q}_3$,

$$\mathbf{P} = (\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3), \tag{1.3}$$

and introduce the vectors $\mathbf{u}_i (i = 1, 2, 3)$ defined by

$$\mathbf{u}_1 = (100; 100; 100), \tag{1.4}$$

$$\mathbf{u}_2 = (010; 010; 010), \text{ etc.} \tag{1.5}$$

Then Eqs. (1.1), (1.2) are equivalent to the forms

$$\mathbf{P}\cdot\mathbf{u}_i = \mathbf{Q}_i \tag{1.6}$$

$$P^2 = 2mT, \tag{1.7}$$

and it is evident that the required transformations consist of all nine-dimensional orthogonal transformations $\mathcal{O}_9$ subject to the constraint that the vectors $\mathbf{u}_i$ be left invariant. These transformations form a subgroup $\mathcal{G}$ of $\mathcal{O}_9$. More formally, $g \in \mathcal{G}$ if

$$(g\mathbf{P})\cdot(g\mathbf{P}) = \mathbf{P}\cdot\mathbf{P} \tag{1.8}$$

and

$$g\mathbf{u}_i = \mathbf{u}_i. \tag{1.9}$$

To examine $\mathcal{G}$ further, it is convenient to make a linear transformation $\mathcal{O}$ on the basis vectors,

$$\mathcal{O} = \begin{pmatrix} -(2^{-\frac{1}{2}}) & 2^{-\frac{1}{2}} & 0 \\ -(6^{-\frac{1}{2}}) & -(6^{-\frac{1}{2}}) & 2\cdot 6^{-\frac{1}{2}} \\ 3^{-\frac{1}{2}} & 3^{-\frac{1}{2}} & 3^{-\frac{1}{2}} \end{pmatrix}. \tag{1.10}$$

Each entry represents a 3×3 unit matrix multiplied by the indicated coefficient. $\mathcal{O}$ is orthogonal, and has determinant $+1$. Under the action of $\mathcal{O}$ the vectors $\mathbf{u}_i$ and $\mathbf{P}$ are transformed to

$$\mathbf{u}_1 \to \mathbf{u}_1' = (0; 0; 3^{\frac{1}{2}}\, 00), \tag{1.11a}$$

$$\mathbf{u}_2 \to \mathbf{u}_2' = (0; 0; 03^{\frac{1}{2}}\, 0), \text{ etc.}, \tag{1.11b}$$

$$\mathbf{P} \to \mathbf{P}' = (\mathbf{p}^{(1)}, \mathbf{p}^{(2)}, \mathbf{p}^{(3)}), \tag{1.12}$$

where

$$\mathbf{p}^{(1)} = (\mathbf{q}_2 - \mathbf{q}_1)/2^{\frac{1}{2}}, \tag{1.13a}$$

$$\mathbf{p}^{(2)} = (2\mathbf{q}_3 - \mathbf{q}_1 - \mathbf{q}_2)/6^{\frac{1}{2}}, \tag{1.13b}$$

$$\mathbf{p}^{(3)} = (\mathbf{q}_1 + \mathbf{q}_2 + \mathbf{q}_3)/3^{\frac{1}{2}} = \mathbf{Q}/3^{\frac{1}{2}}. \tag{1.13c}$$

The inverse transformation is given by

$$\mathbf{q}_1 = -\mathbf{p}^{(1)}/2^{\frac{1}{2}} - \mathbf{p}^{(2)}/6^{\frac{1}{2}} + \mathbf{p}^{(3)}/3^{\frac{1}{2}}, \tag{1.14a}$$

$$\mathbf{q}_2 = \mathbf{p}^{(1)}/2^{\frac{1}{2}} - \mathbf{p}^{(2)}/6^{\frac{1}{2}} + \mathbf{p}^{(3)}/3^{\frac{1}{2}}, \tag{1.14b}$$

$$\mathbf{q}_3 = 2\mathbf{p}^{(2)}/6^{\frac{1}{2}} + \mathbf{p}^{(3)}/3^{\frac{1}{2}}. \tag{1.14c}$$

At the same time, the group $\mathcal{G}$ undergoes the automorphism

$$\mathcal{G} \to \mathcal{G}' = \mathcal{O}\mathcal{G}\mathcal{O}^{-1}. \tag{1.15}$$

Since $\mathcal{O}$ is orthogonal, the group $\mathcal{G}'$ is also orthogonal and isomorphic to $\mathcal{G}$. Moreover, just as the elements of $\mathcal{G}$ are constrained to leave the $\mathbf{u}_i$ invariant, $\mathcal{G}'$ must leave the $\mathbf{u}_i'$ invariant,

$$g'\mathbf{u}_i' = \mathbf{u}_i' \quad \text{for all } g' \in \mathcal{G}'. \tag{1.16}$$

Referring to the form of $\mathbf{u}_i'$, Eq. (1.11), one sees that every element g' of $\mathcal{G}'$ must have the form

$$g' = \begin{pmatrix} \mathfrak{M} & 0 \\ 0 & I \end{pmatrix}, \tag{1.17}$$

where $\mathfrak{M}$ is an arbitrary 6×6 orthogonal matrix, and I denotes a 3×3 identity. Thus $\mathcal{G}'$ (or $\mathcal{G}$) is reducible and is in reality the six-dimensional orthogonal group $\mathcal{O}_6$.

B. The Permutation Group on Three Objects

At this point it would seem natural to initiate a study of the general properties of $\mathcal{G}'$. However, we choose instead to first study a discrete subgroup of $\mathcal{G}'$, the group S_3 of permutations on three objects. As becomes evident later, our choice is motivated by the desire to treat all three particles symmetrically. That S_3 is in fact a subgroup of $\mathcal{G}'$ is clear, for the forms (1.1) and (1.2) are clearly invariant under arbitrary permutations of the three particles.

The permutation group on three objects embraces six elements. They may conveniently be taken to be the transpositions $\mathcal{P}_{12}$, $\mathcal{P}_{23}$, and $\mathcal{P}_{31}$ which interchange particle pairs, the cyclic permutations $\mathcal{C}$ and $\mathcal{C}^2$ which effect the operations $1\,2\,3 \rightarrow 2\,3\,1$ etc., and the identity operation I. Any product of elements can be reduced by the group laws to a single element. For example,

$$\mathcal{P}_{23}\mathcal{P}_{12} = \mathcal{C} \tag{1.18}$$

and

$$\mathcal{C}^3 = I. \tag{1.19}$$

Each transposition forms with the identity a subgroup of order 2. The cyclic permutations and the identity form a subgroup of order 3.

Since $\mathcal{G}'$ is reducible, its elements can be taken to be the group of 6×6 orthogonal matrices acting on the vector space spanned by the six-component vectors $\mathbf{p}$,

$$\mathbf{p} = (\mathbf{p}^{(1)}, \mathbf{p}^{(2)}). \tag{1.20}$$

The elements of S_3 are then also represented by 6×6 matrices. These matrices may be divided into two classes depending on whether their determinant is plus or minus one. For example, under the transposition $\mathcal{P}_{12}$,

$$\mathbf{p}^{(1)} \rightarrow -\mathbf{p}^{(1)}, \tag{1.21a}$$

$$\mathbf{p}^{(2)} \rightarrow \mathbf{p}^{(2)}, \tag{1.21b}$$

so that $\mathcal{P}_{12}$ has the matrix representation

$$\mathcal{P}_{12} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \tag{1.22}$$

with determinant -1. Under the cyclic permutation $\mathcal{C}$,

$$\mathbf{p}^{(1)} \rightarrow 2^{-\frac{1}{2}}(\mathbf{q}_3 - \mathbf{q}_2) = -\mathbf{p}^{(1)}/2 + (3^{\frac{1}{2}}/2)\mathbf{p}^{(2)}, \tag{1.23a}$$

$$\mathbf{p}^{(2)} \rightarrow 6^{-\frac{1}{2}}(2\mathbf{q}_1 - \mathbf{q}_2 - \mathbf{q}_3) = (-3^{\frac{1}{2}}/2)\mathbf{p}^{(1)} - \mathbf{p}^{(2)}/2, \tag{1.23b}$$

giving $\mathcal{C}$ the representation

$$\mathcal{C} = -\frac{1}{2}\begin{pmatrix} 1 & -3^{\frac{1}{2}} \\ 3^{\frac{1}{2}} & 1 \end{pmatrix} \tag{1.24}$$

with determinant $+1$.

From the group relations of the form (1.18) and (1.19) we may deduce that all the transpositions have determinant -1; and the cyclic permutations, including the identity, all have determinant $+1$. Consequently, the cyclic permutations enjoy a continuous connection to the identity which lies entirely within the group $\mathcal{G}'$,

$$\mathcal{C} = \exp(\tfrac{2}{3}\pi S) = I\cos\tfrac{2}{3}\pi + S\sin\tfrac{2}{3}\pi \tag{1.25}$$

with

$$S = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{1.26}$$

C. The Subgroup $\mathfrak{D}$

Armed with these few facts about S_3, we return to a more general study of $\mathcal{G}'$. Let $\mathfrak{R}$ be that portion of $\mathcal{G}'$ which is infinitesimally generated. In the neighborhood of the identity, an arbitrary element of $\mathfrak{R}$ is of the form

$$\mathfrak{R} = I + \epsilon R. \tag{1.27}$$

For $\mathfrak{R}$ to be real, orthogonal, R must be real and antisymmetric,

$$R^* = R, \tag{1.28a}$$

$$\tilde{R} = -R. \tag{1.28b}$$

There are 15 linearly independent antisymmetric 6×6 matrices. Let $|i)$ denote a six-component column vector in a real vector space whose ith component is one, and whose remaining components are zero; let $(i|$ be the corresponding row vector. Then the space of 6×6 real antisymmetric matrices is spanned by the matrices R_{ij},[6]

$$R_{ij} = |i)(j| - |j)(i|; \quad i, j = 1, 2, \cdots, 6; \quad i \neq j. \tag{1.29}$$

The R_{ij} together form a 15-dimensional Lie algebra L_0, subject to the rules

$$[R_{ij}, R_{mn}] = 0, \qquad i \neq j \neq m \neq n, \tag{1.30a}$$

$$[R_{ij}, R_{jk}] = R_{ik}, \tag{1.30b}$$

$$R_{ij} = -R_{ji}. \tag{1.30c}$$

[6] The subscripts i, j are labels for the different matrices, and are not to be confused with matrix elements.

By definition, L_0 is the Lie algebra of $\mathcal{O}_6$. Note that S defined by Eq. (1.26) belongs to L_0.

Not all the elements of L_0 treat all three particles equivalently. For example, transformations generated by R_{ij} with $i, j \leq 3$ affect only the vector $\mathbf{p}^{(1)}$, and by Eq. (1.14c) the momentum $\mathbf{q}_3$ of particle 3 is left unchanged. There is, however, a subset L_1 of elements in L_0 which does. Moreover, L_1 is a Lie algebra in its own right and may be exponentiated to yield a subgroup $\mathfrak{D}$ of $\mathcal{G}'$. To see how this comes about, consider the automorphisms of L_0 generated by the permutation group. That is, look at the effect of the operation

$$L_0 \rightarrow s^{-1}L_0 s, \tag{1.31}$$

where s is an element of S_3. Since s is orthogonal and L_0 is the set of antisymmetric 6×6 matrices, it is clear that the operation maps L_0 onto itself so we do in fact get an automorphism.

We first deal with the cyclic permutations. Define L_1 to be the set of all elements F in L_0 with the property

$$\mathcal{C}^{-1}F\mathcal{C} = F. \tag{1.32}$$

Equation (1.32) can also be written in the form

$$[\mathcal{C}, F] = 0. \tag{1.32'}$$

It is now evident that L_1 is a Lie algebra. For if F_1 and F_2 are in L_1, it follows from the Jacobi identity that $[\mathcal{C}, [F_1, F_2]] = 0$. Therefore $[F_1, F_2]$ is also in L_1.

By Eq. (1.25), we may equivalently look for F's in L_0 with the property

$$[S, F] = 0. \tag{1.33}$$

Let M be a 6×6 matrix written in 3×3 block form,

$$M = \begin{pmatrix} M^{11} & M^{12} \\ M^{21} & M^{22} \end{pmatrix}. \tag{1.34}$$

For M to be in L_0, it must be antisymmetric,

$$\tilde{M}^{11} = -M^{11}, \quad \tilde{M}^{22} = -M^{22}, \quad \tilde{M}^{12} = -M^{21}. \tag{1.35}$$

To be in L_1 it must also commute with S,

$$[S, M] = \begin{bmatrix} M^{12} + M^{21} & M^{22} - M^{11} \\ M^{22} - M^{11} - M^{12} - M^{21} \end{bmatrix} = 0, \tag{1.36}$$

or

$$M^{12} = -M^{21}, \qquad M^{11} = M^{22}. \tag{1.37}$$

Therefore, a general element in L_1 is of the form

$$M = \begin{pmatrix} A & B \\ -B & A \end{pmatrix} \tag{1.38}$$

with

$$\tilde{A} = -A, \qquad \tilde{B} = B. \tag{1.39}$$

It is now easily verified from Eq. (1.29) that L_1 is spanned by the elements J_{ij} and K_{ij} defined by the relations

$$\begin{aligned} J_{ij} &= R_{ij} + R_{i+3,j+3}, \qquad i, j \leq 3, \quad i \neq j, \\ K_{ij} &= R_{i,j+3} - R_{i+3,j}, \qquad i, j \leq 3. \end{aligned} \tag{1.40}$$

From Eq. (1.30c) it follows that

$$J_{ij} = -J_{ji}, \qquad K_{ij} = K_{ji}. \tag{1.41}$$

Consequently L_1 is nine dimensional. Note that the element S used to define L_1 is itself a member of L_1. It is related to the K_{ij} by the equation

$$S = \tfrac{1}{2}\sum_i K_{ii}. \tag{1.42}$$

Since L_0 has 15 elements, there are still six elements to be accounted for. They may conveniently be taken to be matrices N of the form

$$N = \begin{pmatrix} A & B \\ B & -A \end{pmatrix} \tag{1.43}$$

with

$$\tilde{A} = -A, \qquad \tilde{B} = -B. \tag{1.44}$$

The space of matrices N is spanned by the elements V_{ij} and W_{ij} defined as

$$\begin{aligned} V_{ij} &= R_{ij} - R_{i+3,j+3}, \qquad i, j \leq 3, \quad i \neq j. \\ W_{ij} &= R_{i,j+3} + R_{i+3,j}, \end{aligned} \tag{1.45}$$

Inspection of Eqs. (1.40) and (1.45) shows that they may be inverted to give R in terms of J, K, V, and W; and our accounting is complete.

It is of interest to find the effect of $\mathcal{C}$ on V and W. To do this, calculate the commutators of S with V and W,

$$[S, V_{ij}] = -2W_{ij}, \qquad [S, W_{ij}] = 2V_{ij}. \tag{1.46}$$

It follows that

$$\begin{aligned} \mathcal{C}^{-1}V_{ij}\mathcal{C} &= \cos\tfrac{4}{3}\pi\, V_{ij} + \sin\tfrac{4}{3}\pi W_{ij}, \\ \mathcal{C}^{-1}W_{ij}\mathcal{C} &= \cos\tfrac{4}{3}\pi W_{ij} - \sin\tfrac{4}{3}\pi\, V_{ij}. \end{aligned} \tag{1.47}$$

Next, consider transpositions. To study them, it is sufficient to examine the effect of $\mathcal{P}_{12}$ since together $\mathcal{P}_{12}$ and $\mathcal{C}$ generate all of S_3 and the effect of $\mathcal{C}$ is already known. Referring to Eq. (1.22), we readily obtain

$$\mathcal{P}_{12}^{-1}J_{ij}\mathcal{P}_{12} = J_{ij}, \tag{1.48a}$$

$$\mathcal{P}_{12}^{-1}K_{ij}\mathcal{P}_{12} = -K_{ij}, \tag{1.48b}$$

$$\mathcal{P}_{12}^{-1}V_{ij}\mathcal{P}_{12} = V_{ij}, \tag{1.48c}$$

$$\mathcal{P}_{12}^{-1}W_{ij}\mathcal{P}_{12} = -W_{ij}. \tag{1.48d}$$

Thus, except for a change in sign of some of its elements, L_1 remains unaffected. Moreover, the effect on L_1 is the same for each transposition.

We conclude that L_1 has all the desired attributes. For if we were to relabel the particles, i.e., permute them, L_1 either remains unaffected or undergoes the sign change (1.48b). If there is a sign change, it is still impossible to tell which pair of particles has been interchanged and which particle has been left alone. Thus L_1 treats all three particles with complete symmetry.

2. REALIZATION OF $\mathcal{G}'$ BY UNITARY TRANSFORMATIONS

In the work so far, the Lie algebras L_0 and L_1 have been defined in terms of 6×6 matrices. To continue with a quantum mechanical discussion, it is necessary to have a realization of L_0 and L_1 as Lie algebras of Hermitian operators acting on three-particle state vectors. That is, the quantum analog to the R_{ik} is a set of operators Λ_{ik} with the property

$$[\Lambda_{ik}, \mathbf{p}] = iR_{ik}\mathbf{p}. \tag{2.1}$$

Let $\mathbf{r}_1$, $\mathbf{r}_2$, and $\mathbf{r}_3$ denote the position vectors of particles 1 to 3. Corresponding to Eqs. (1.13), we define relative coordinates $\mathbf{r}^{(1)}$ to $\mathbf{r}^{(3)}$,

$$\mathbf{r}^{(1)} = 2^{-\frac{1}{2}}(\mathbf{r}_2 - \mathbf{r}_1), \tag{2.2a}$$

$$\mathbf{r}^{(2)} = 6^{-\frac{1}{2}}(2\mathbf{r}_3 - \mathbf{r}_1 - \mathbf{r}_2), \tag{2.2b}$$

$$\mathbf{r}^{(3)} = 3^{-\frac{1}{2}}(\mathbf{r}_1 + \mathbf{r}_2 + \mathbf{r}_3). \tag{2.2c}$$

Let $\mathbf{r}$ be a six-component position vector defined analogously to $\mathbf{p}$,

$$\mathbf{r} = (\mathbf{r}^{(1)}, \mathbf{r}^{(2)}). \tag{2.3}$$

It is easily verified that $\mathbf{r}$ and $\mathbf{p}$ are Hermitian and canonically conjugate,[7]

$$[r_i, p_k] = i\delta_{ik}. \tag{2.4}$$

Consequently, we may take for Λ_{ik} the expression

$$\Lambda_{ik} = r_i p_k - r_k p_i. \tag{2.5}$$

One easily checks that Eq. (2.1) is satisfied. A similar relation holds for $\mathbf{r}$,

$$[\Lambda_{ik}, \mathbf{r}] = iR_{ik}\mathbf{r}. \tag{2.6}$$

Finally, the commutators of the Λ_{ik} with themselves obey the Hermitian analog of Eq. (1.30),

$$[\Lambda_{ik}, \Lambda_{lm}] = 0, \qquad j \neq k \neq l \neq m, \tag{2.7a}$$

$$[\Lambda_{jk}, \Lambda_{kl}] = -i\Lambda_{jl}, \tag{2.7b}$$

$$\Lambda_{ik} = -\Lambda_{ki}. \tag{2.7c}$$

By their definition in Eq. (2.5), the Λ_{ij} are subject to a bilinear identity,

$$\Lambda_{ij}\Lambda_{kl} + \Lambda_{il}\Lambda_{jk} + \Lambda_{ik}\Lambda_{lj} \equiv 0, \; i \neq j \neq k \neq l. \tag{2.8}$$

Therefore, the Λ_{ij} are not the most general objects satisfying the commutation rules (2.7). That is, the representations of $\mathcal{G}' = \mathcal{O}_6$ carried by unitary transformations on three-particle state vectors are not the most general ones. Consider the quadratic operator Λ^2 defined by

$$\Lambda^2 = \tfrac{1}{2}\sum_{i,j}\Lambda_{ij}^2. \tag{2.9}$$

From Eq. (2.7) it follows that Λ^2 commutes with all the Λ_{ij},

$$[\Lambda^2, \Lambda_{ij}] = 0. \tag{2.10}$$

Consequently, Λ^2 is a multiple of the quadratic Casimir operator[8] for L_0, and its eigenvalues can be employed as one of the parameters necessary to specify a representation of $\mathcal{G}'$. We shall see that as a consequence of the identity (2.8), the value of Λ^2 is completely determined in terms of an operator constructed out of elements in L_1. In Sec. 1 where we dealt with 6×6 matrices, L_1 was defined by Eq. (1.40) to be certain linear combinations of elements in L_0. In exactly the same way, we define Hermitian operators J_{ij} and K_{ij} in terms of the Λ_{ij},

$$J_{ij} = \Lambda_{ij} + \Lambda_{i+3,j+3}, \tag{2.11a}$$

$$K_{ij} = \Lambda_{i,j+3} - \Lambda_{i+3,j}, \tag{2.11b}$$

$$S = \tfrac{1}{2}\sum_i K_{ii}. \tag{2.11c}$$

It follows that

$$\Lambda^2 = \frac{1}{2}\sum_{i,j=1}^{6}\Lambda_{ij}^2 = \frac{1}{2}\sum_{i,j=1}^{3}(J_{ij}^2 + K_{ij}^2) + \sum_{i,j=1}^{3}(\Lambda_{i,j+3}\Lambda_{i+3,i} - \Lambda_{i,j}\Lambda_{i+3,j+3}). \tag{2.12}$$

But by Eq. (2.8),

$$-\Lambda_{ij}\Lambda_{i+3,j+3} = \Lambda_{i,j+3}\Lambda_{j,i+3} + \Lambda_{i,i+3}\Lambda_{j+3,j}. \tag{2.13}$$

Consequently,

$$\Lambda^2 = \tfrac{1}{2}\sum_{i,j}(J_{ij}^2 + K_{ij}^2) - S^2. \tag{2.14}$$

[7] The units are such that $\hbar = 1$.

[8] M. Hamermesh, *Group Theory* (Addison-Wesley Publishing Company, Inc., Reading, Massachusetts, 1962), p. 317.

The precise implications of this result for representations of $\mathcal{G}'$ will be discussed in Sec. 7.

It is of interest to have the operator realizations of L_1 displayed directly in terms of the three-particle coordinates $\mathbf{q}_i$ and $\mathbf{r}_i$. Since J_{ik} is antisymmetric in its indices, it is convenient to introduce J_i defined by

$$J_i = \tfrac{1}{2}\epsilon_{ijk}J_{jk}. \tag{2.15}$$

Using Eqs. (2.11), (2.5), (2.3), (2.2), etc., we find that $\mathbf{J}$ is indeed a vector as the notation suggests,

$$\mathbf{J} = \mathbf{r}_1 \times \mathbf{q}_1 + \mathbf{r}_2 \times \mathbf{q}_2 + \mathbf{r}_3 \times \mathbf{q}_3 - \mathbf{r}^{(3)} \times \mathbf{p}^{(3)}. \tag{2.16}$$

In the center-of-mass frame, $\mathbf{r}^{(3)} = 0$; correspondingly $\mathbf{p}^{(3)} = 0$ in the center-of-momentum frame.[9] Therefore $\mathbf{J}$ may be interpreted either as the total angular momentum about the center-of-mass, or the total angular momentum in the center-of-momentum frame. Similarly, the elements K_{ij} form a symmetric tensor. In dyad notation,

$$3^{\frac{1}{2}}\mathbf{K} = (\mathbf{r}_1 - \mathbf{r}_2)\mathbf{q}_3 + (\mathbf{r}_2 - \mathbf{r}_3)\mathbf{q}_1 + (\mathbf{r}_3 - \mathbf{r}_1)\mathbf{q}_2 + \text{transpose}. \tag{2.17}$$

Note that $\mathbf{J}$ and $\mathbf{K}$ involve the three particles equivalently as anticipated.

3. CLASSIFICATION OF L_1

We know that L_0 is the Lie algebra of $\mathcal{O}_6$. What is the corresponding group for L_1? From the definition (2.11a) and the relations (2.7), the commutators of $\mathbf{J}$ with itself are the expected ones for angular momentum,

$$[J_i, J_k] = i\epsilon_{ikl}J_l. \tag{3.1}$$

Similarly, the commutation rules of $\mathbf{K}$ with $\mathbf{J}$ are those expected for a tensor,

$$[J_i, K_{kl}] = i\epsilon_{ikm}K_{ml} + i\epsilon_{ilm}K_{km}. \tag{3.2}$$

By Eqs. (2.11c), (1.33), the trace of $\mathbf{K}$ commutes with all of L_1. This circumstance, and the fact that $\mathbf{K}$ transforms under $\mathbf{J}$ as a tensor, suggests the introduction of a spherical basis. Since $\mathbf{K}$ is symmetric, it may be decomposed into a scalar T^0_0 and a five-component object of spin 2, T^2_m:

$$T^0_0 = \tfrac{1}{2}\sum_i K_{ii} = S, \tag{3.3a}$$

$$T^2_2 = \tfrac{1}{2}(K_{11} - K_{22} + 2iK_{12}), \tag{3.3b}$$

$$T^2_1 = -K_{13} - iK_{23}, \tag{3.3c}$$

$$T^2_0 = 6^{-\frac{1}{2}}(2K_{33} - K_{11} - K_{22}), \tag{3.3d}$$

$$T^2_{-1} = K_{13} - iK_{23}, \tag{3.3e}$$

$$T^2_{-2} = \tfrac{1}{2}(K_{11} - K_{22} - 2iK_{12}). \tag{3.3f}$$

Equations (3.1) and (3.2) now take the form

$$[J_0, J_\pm] = \pm J_\pm, \tag{3.4a}$$

$$[J_0, T^i_m] = mT^i_m, \tag{3.4b}$$

$$[J_+, J_-] = 2J_0, \tag{3.4c}$$

$$[J_\pm, T^i_m] = [(j \mp m)(j \pm m + 1)]^{\frac{1}{2}}T^i_{m\pm 1}, \tag{3.4d}$$

where

$$J_\pm = J_1 \pm iJ_2, \qquad J_0 = J_3. \tag{3.5}$$

Finally, the nonvanishing commutators of T^i_m with itself are given by

$$\begin{aligned} &[T_2, T_{-2}] = 4J_0, \quad [T_2, T_{-1}] = -2J_+, \\ &[T_{-2}, T_1] = -2J_-, \quad [T_1, T_{-1}] = -2J_0, \\ &[T_1, T_0] = 6^{\frac{1}{2}}J_+, \quad [T_{-1}, T_0] = 6^{\frac{1}{2}}J_-, \end{aligned} \tag{3.6}$$

where the superscript $j = 2$ has been dropped for notational convenience.

Notice that T^0_0 nowhere appears on the right-hand side of a commutation relation. Consequently, if we separate out T^0_0, the remaining elements T^2_m, $J_\pm$, and J_0 form a Lie algebra L_2 of dimension 8.

We assert that L_2 is simple. That is, L_2 contains no invariant subalgebra L_3 such that all commutators of the form $[L_2, L_3]$ are also in L_3. The proof is by contradiction. Suppose that L_3 is an invariant subalgebra. Let F be an element in L_3. By definition, F can be expressed as a linear combination of elements in L_2,

$$F = \sum_m c_m T^2_m + d_m J_m. \tag{3.7}$$

Form successive commutators of F with $J_\pm$. Then by Eq. (3.4d) either all the c_m vanish or L_3 contains all the T^2_m. In the latter case, L_3 must also contain all the J's by Eq. (3.6) and $L_3 = L_2$ contrary to assumption. Therefore all the c_m vanish. Form once again successive commutators of F with $J_\pm$. By Eq. (3.4) either all the d_m vanish or L_3 contains all the J's. In the latter case, form the commutators $[T^2_2, L_3]$. From Eq. (3.4d), L_3 must contain all the T^2_m and $L_3 = L_2$, again contrary to assumption. Therefore, the d_m are also zero and L_3 vanishes identically.

It now follows that L_2 is isomorphic to SU_3, the Lie algebra of traceless Hermitian 3×3 matrices, for there is only one simple Lie algebra of dimension

[9] Note that $\mathbf{p}^{(3)}$ and $\mathbf{r}^{(3)}$ cannot vanish simultaneously since they do not commute.

8.[10,11] Adjoining T_0^0 to L_2, we get U_3, the algebra of 3×3 Hermitian matrices. Thus L_1 is isomorphic to U_3 and $\mathfrak{D}$ is isomorphic to the group of 3×3 unitary matrices.

This result can also be seen directly without recourse to the theory of simple Lie algebras by examining the space of matrices M defined in Eqs. (1.38), (1.39). Making the unitary transformation U with

$$U = 2^{-\frac{1}{2}}\begin{pmatrix} 1 & -i \\ -i & 1 \end{pmatrix}, \tag{3.8}$$

one finds

$$UMU^{-1} = \begin{pmatrix} A + iB & 0 \\ 0 & A - iB \end{pmatrix}. \tag{3.9}$$

The space of matrices M is thus reducible to two sets of 3×3 matrices. Matrices of the form $A + iB$ with A and B real and obeying (1.39) span the space of 3×3 anti-Hermitian matrices, and therefore upon exponentiation generate the group U_3. The matrices $A - iB$ then generate the conjugate representation U_3^*.

4. THEOREMS ABOUT U_3 AND SU_3

In the following sections we will discover that three-particle states can be completely classified by their transformation properties under L_2. We forearm ourselves by reviewing some well known facts about U_3, SU_3, and their representations.

There are two popular ways of treating SU_3. In the first, one observes that SU_3 is simple and thereby subject to the methods developed by Cartan for all simple Lie algebras. In the second, one embeds SU_3 within $\mathcal{GL}_3$ and then uses the Weyl classification for the general linear group.

A. Cartan's Method

The Cartan method[10,11] employs as a basis for the Lie algebra operators which are a generalization of the ladder and diagonal operators, $J_\pm$ and J_0, of angular momentum. In the case of SU_3, there are two commuting operators which play roles analogous to J_0 and six ladder operators. Call the commuting operators H_1 and H_2. The number of mutually commuting operators within an algebra is called the rank of an algebra. Thus SU_3 is of rank two. The 6 ladder operators are conveniently labeled by 3 two-component vectors called "root vectors", and their negatives. Let $\mathbf{e}_1$ and $\mathbf{e}_2$ be orthogonal unit vectors. Define the vectors

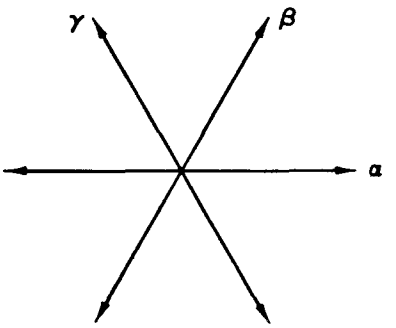

FIG. 1. Root vectors for SU_3.

$$\begin{aligned} \boldsymbol{\alpha} &= (3^{-\frac{1}{2}})\mathbf{e}_1, \\ \boldsymbol{\beta} &= (3^{-\frac{1}{2}}/2)\mathbf{e}_1 + \mathbf{e}_2/2, \\ \boldsymbol{\gamma} &= -(3^{-\frac{1}{2}}/2)\mathbf{e}_1 + \mathbf{e}_2/2. \end{aligned} \tag{4.1}$$

They are shown schematically in Fig. 1. We denote the ladder operators by $E(\boldsymbol{\mu})$ where $\boldsymbol{\mu}$ is one of the vectors (4.1) or their negatives.

The virtue of this labeling is that the commutation rules take a particularly illuminating form. The commutator of $E(\boldsymbol{\mu})$ with H_i is

$$[H_i, E(\boldsymbol{\mu})] = \boldsymbol{\mu}\cdot\mathbf{e}_i E(\boldsymbol{\mu}). \tag{4.2}$$

The H_i thus serve to establish the coordinate system for the root vectors. The commutators between pairs of E's, $E(\boldsymbol{\mu})$ and $E(\boldsymbol{\nu})$, are of two types. If the root vectors $\boldsymbol{\mu}$ and $\boldsymbol{\nu}$ are equal and opposite,

$$[E(\boldsymbol{\mu}), E(-\boldsymbol{\mu})] = \sum_i (\mathbf{e}_i\cdot\boldsymbol{\mu})H_i. \tag{4.3}$$

If the sum of $\boldsymbol{\mu}$ and $\boldsymbol{\nu}$ is again a root vector, the commutator takes the form

$$[E(\boldsymbol{\mu}), E(\boldsymbol{\nu})] = N(\boldsymbol{\mu}, \boldsymbol{\nu})E(\boldsymbol{\mu} + \boldsymbol{\nu}), \tag{4.4}$$

where $N(\boldsymbol{\mu}, \boldsymbol{\nu})$ is a numerical factor equal to $\pm 6^{-\frac{1}{2}}$. The positive N's are $N(\boldsymbol{\alpha}, \boldsymbol{\gamma})$, $N(-\boldsymbol{\gamma}, -\boldsymbol{\alpha})$, $N(\boldsymbol{\gamma}, -\boldsymbol{\beta})$, $N(-\boldsymbol{\beta}, \boldsymbol{\alpha})$, $N(\boldsymbol{\beta}, -\boldsymbol{\gamma})$, and $N(-\boldsymbol{\alpha}, \boldsymbol{\beta})$.[10,11] All other commutators vanish.

From an algebraic viewpoint, the choice of the two commuting operators is somewhat arbitrary. For example, one could have choosen the elements now labeled $E(\boldsymbol{\alpha})$ and $E(\boldsymbol{\beta})$ in place of H_1 and H_2. The particular choice one makes depends on which elements are to be diagonalized. In the case at hand, it is natural to diagonalize J_0. Therefore we take H_1 to be a multiple of J_0. The Cartan basis is then related to the basis already given for L_2 by the expressions

$$H_1 = (12)^{-\frac{1}{2}}J_0, \tag{4.5a}$$

[10] C. Fronsdal, "Elementary Particle Physics and Field Theory" in *1962 Brandeis University Summer Institute Lectures in Theoretical Physics* (W. A. Benjamin Company Inc., New York, 1963), Vol. 1, p. 461 ff.

[11] R. Behrends, J. Dreitlein, C. Fronsdal, and B. W. Lee, Rev. Mod. Phys. **34**, 1 (1962) and references cited therein.

$$H_2 = (24)^{-\frac{1}{2}}T_0, \tag{4.5b}$$

$$E(\pm\alpha) = (24)^{-\frac{1}{2}}T_{\pm 2}, \tag{4.5c}$$

$$E(\pm\beta) = \pm(48)^{-\frac{1}{2}}(T_{\pm 1} \mp J_{\pm}), \tag{4.5d}$$

$$E(\pm\gamma) = \pm(48)^{-\frac{1}{2}}(T_{\mp 1} \mp J_{\mp}). \tag{4.5e}$$

Suppose that H_1 and H_2 are diagonal. Let $|m_1m_2\rangle$ denote an eigenstate with the property

$$H_i\,|m_1m_2\rangle = m_i\,|m_1m_2\rangle. \tag{4.6}$$

Since the H_i are Hermitian by Eqs. (4.5a, b) the m_i are real. It is convenient to treat them together as the components of a single vector $\mathbf{m}$ called a weight. From the commutation rules (4.2) it follows that if $E(\mathfrak{y})\,|\mathbf{m}\rangle$ is different from zero, it is an eigenstate with weight $\mathbf{m} + \mathfrak{y}$. Consequently, a single weight generates a whole set of weights. The set of weights can be ordered by means of the following definitions:

1) A weight is *positive* if its first nonvanishing component is positive.

2) A weight $\mathbf{m}$ is *higher* than $\mathbf{m}'$ if $\mathbf{m} - \mathbf{m}'$ is positive.

We can now state the fundamental theorems of Cartan concerning representations:

1) In an irreducible representation, the eigenvector with highest weight is unique, i.e., nondegenerate.

2) Two irreducible representations are equivalent if they have the same highest weight.

3) Every highest weight $\mathbf{m}_h$ for SU_3 is of the form

$$\mathbf{m}_h = (\lambda_1/6)(3^{\frac{1}{2}}\mathbf{e}_1 + \mathbf{e}_2) + (\lambda_2/6)(3^{\frac{1}{2}}\mathbf{e}_1 - \mathbf{e}_2), \tag{4.7}$$

where λ_1 and λ_2 are arbitrary nonnegative integers.

Taken together, the theorems show that an irreducible representation of SU_3 is completely characterized by the two integers λ_1 and λ_2. We denote this representation by $\Gamma(\lambda_1, \lambda_2)$. It is easily shown that the conjugate representation is given by $\Gamma(\lambda_2, \lambda_1)$. That is, $\Gamma^*(\lambda_1, \lambda_2) = \Gamma(\lambda_2, \lambda_1)$.

B. Weyl's Method

In contrast to the infinitesimal methods of Cartan, the method of Weyl[12] deals with finite group elements. Consider the lowest-order representation of $\mathcal{GL}_3$, the general linear group of dimension 3, provided by arbitrary linear transformations on a three-dimensional vector space. By taking repeated Cartesian products to form tensors of higher rank, one obtains representations of higher order. The set of tensor representations can be reduced by forming linear combinations of tensors with permuted indices. Thus, the space of rth-rank tensors is reducible into subspaces consisting of tensors with different symmetry. For each symmetry there is a corresponding Young tableau. In the case of $\mathcal{GL}_3$, the tableaux for tensors of rank r can contain at most three rows of lengths f_1, f_2, and f_3 with $\sum f_i = r$ and $f_1 \geq f_2 \geq f_3 \geq 0$. Consequently, an irreducible representation of $\mathcal{GL}_3$ is characterized by the partition $(f_1f_2f_3)$.

Next consider the unitary subgroup U_3 of $\mathcal{GL}_3$. Its Lie algebra is spanned by the nine 3×3 Hermitian matrices. Since there are only 9 linearly independent 3×3 matrices, the Lie algebra of U_3 when taken over the complex field is isomorphic to that of $\mathcal{GL}_3$. Therefore, $(f_1f_2f_3)$ also serves as a label for irreducible representations of U_3.

If we pass from U_3 to SU_3, all representations of the form $(f_1 + e, f_2 + e, f_3 + e)$, where e is an arbitrary integer, become equivalent since tensors belonging to the different partitions differ in their transformation properties only by powers of the modulus which is now required to be one. Consequently, for SU_3, the partition $(f_1f_2f_3)$ can be replaced by the differences $k_1 = f_1 - f_3$ and $k_2 = f_2 - f_3$. It follows from the inequalities on the f_i that $k_1 \geq k_2 \geq 0$. We denote these representations by $\hat{\Gamma}(k_1k_2)$.

One may wonder if there is some way to get at k_1 and k_2 by infinitesimal means. The question has been answered in the affirmative in a series of papers by Bargmann and Moshinsky.[13,14] The Lie algebra of $\mathcal{GL}_3$ is spanned by the matrices X_{ij},

$$X_{ij} = |i)(j|, \qquad i, j \leq 3, \tag{4.8}$$

with the commutation rules

$$[X_{ij}, X_{kl}] = \delta_{jk}X_{il} - \delta_{li}X_{kj}. \tag{4.9}$$

The commutation rules for L_1 can be put into the same form by a linear transformation,

$$C_{jk} = \tfrac{1}{2}(K_{jk} + iJ_{jk}). \tag{4.10}$$

In analogy to the Cartan state of highest weight, Bargmann and Moshinsky define a state $|g_1g_2g_3\rangle$ of highest weight by the requirements

$$\begin{aligned} C_{ii}\,|g\rangle &= g_i\,|g\rangle, \\ C_{ij}\,|g\rangle &= 0, \qquad i < j. \end{aligned} \tag{4.11}$$

They then pass to SU_3 by discarding the trace $C_{11} + C_{22} + C_{33}$ and setting $k_1 = g_1 - g_3$, $k_2 = g_2 - g_3$.

[12] Reference 8, p. 377 ff.

[13] V. Bargmann and M. Moshinsky, Nucl. Phys. **18**, 697 (1960); *ibid.* **23**, 177 (1961).

[14] M. Moshinsky, Rev. Mod. Phys. **34**, 813 (1962).

The numbers k_1 and k_2 are evidently similar to λ_1 and λ_2. The precise relation is given by

$$\Gamma(\lambda_1\lambda_2) = \hat{\Gamma}(k_1, k_2) \tag{4.12a}$$

when

$$k_1 = \lambda_1 + \lambda_2, \tag{4.12b}$$

$$k_2 = \lambda_1. \tag{4.12c}$$

Since $\Gamma^*(\lambda_1\lambda_2) = \Gamma(\lambda_2, \lambda_1)$, we have the result $\hat{\Gamma}^*(k_1k_2) = \hat{\Gamma}(k_1, k_1 - k_2)$.

C. Embedding of SU_2 within SU_3

SU_3 contains SU_2 as a subalgebra in two algebraically distinct ways. Consider the two sets of operators $X_\pm$, X_0 and $Y_\pm$, Y_0 defined by the relations

$$\begin{aligned} X_\pm &= 6^{\frac{1}{2}}E(\pm\alpha), \\ X_0 &= 3^{\frac{1}{2}}H_1, \\ Y_\pm &= -2(3^{\frac{1}{2}})[E(\pm\beta) + E(\mp\gamma)], \\ Y_0 &= 2(3)^{\frac{1}{2}}H_1. \end{aligned} \tag{4.13}$$

It is easily verified that both sets satisfy the usual commutation rules for SU_2. Now look at the remaining elements of SU_3. For the X set we see that there is a fourth element, $X_4 = H_2$, of SU_3 which commutes with all the elements of the X SU_2. However, for the Y set there is no fourth commuting element. This is most easily verified by examining Eqs. (3.4) and (4.5). Thus the embedding of the X and Y SU_2's within SU_3 are algebraically distinct.

The X and Y sets also differ in their eigenvalue spectrum. If we look at the state of highest weight defined by Eq. (4.7), we see that X_0 has the eigenvalue $\frac{1}{2}(\lambda_1 + \lambda_2)$. That is, the eigenvalues are integral and half-integral as expected for SU_2. By contrast, Y_0 has only integral eigenvalues. Thus the Y set generates only the single-valued representations for the rotation group. For this reason, it is sometimes referred to as R_3.

D. Specification of States within a Representation

The procedures of Weyl or Cartan provide two indices for SU_3 which serve to distinguish different representations. Once one has a given representation, it is necessary to have further indices in order to specify the different vectors within a representation. The situation is analogous to the rotation group (SU_2) where one uses j to specify the representation and the additional index j_z to specify a state within the representation.

In the case of SU_3, we need three additional indices. They may be obtained in two different ways. The choice of which is to be used is dictated by physical considerations.

The first method, familiar from unitary symmetry schemes,[15] employs an SU_2 of the X variety along with its commuting operator X_4. The vectors within a representation are then labeled by the eigenvalues of $X^2 = X_0^2 + \frac{1}{2}(X_+X_- + X_-X_+)$, X_0, and X_4. In the "eightfold way" they have the physical interpretation of isotopic spin, its z component, and hypercharge, respectively.

The second method of indexing uses an SU_2 of the Y type. In this case one gets with relative ease only the two labels provided by the eigenvalues of Y^2 and Y_0. For the three-body problem these operators represent the total angular momentum and its z component.

The necessary third label is much harder to obtain. One therefore wonders what a third index is good for. That is, in a representation characterized by λ_1 and λ_2(or k_1 and k_2), how many states have the same J and J_z? This problem has been studied by Racah.[16] His results are summarized in Table I. We see that a third label is necessary only for the states with $J \geq 2$. Furthermore, the degeneracy depends essentially only upon J.

To produce a third index, it is necessary to find an operator which commutes with R_3. We already know that there is no element of SU_3 which does so. The next objects to consider are quadratic and higher-order operators formed by products of elements in SU_3. One is then led to the cubic and biquadratic operators $\Omega = \mathbf{J}\cdot\mathbf{T}\cdot\mathbf{J}$ and $\Omega' = \mathbf{J}\cdot\mathbf{T}\cdot\mathbf{T}\cdot\mathbf{J}$ where $\mathbf{T}$ is the Cartesian counterpart of T_m^2. It is readily verified that Ω and Ω' commute with all the elements of R_3. Moreover, one can show that they are essentially unique in that all other operators which commute with R_3 are either functions of Ω and Ω' or Casimir operators whose value depends only on λ_1 and λ_2. Finally, Bargmann and Moshinsky

TABLE I. Multiplicity of $R_3(J)$ within $SU_3(\lambda_1\lambda_2)$.

J	λ_1, λ_2	Multiplicity
even	both even	$\frac{1}{2}J + 1$
	one or both odd	$\frac{1}{2}J$
odd	both even	$\frac{1}{2}(J - 1)$
	one or both odd	$\frac{1}{2}(J + 1)$

[15] See, for example, J. J. deSwart, Rev. Mod. Phys. **35**, 916 (1963) and references cited therein.
[16] G. Racah, Rev. Mod. Phys. **21**, 494 (1949).

have shown that the use of Ω alone is sufficient to break all degeneracies.[13]

5. DALITZ-FABRI COORDINATES

In order to construct representations for L_2, it is necessary to have a convenient system of coordinates. Since L_2 treats the three particles equivalently, the coordinate system should also.

A. Momentum Coordinates

First consider momentum coordinates. It is natural to work in the center-of-momentum frame, and then pass to other frames by Lorentz transformations. We are then faced with the problem of parameterizing 3-momenta, $\mathbf{q}_1$ to $\mathbf{q}_3$, subject to the constraints

$$\mathbf{q}_1 + \mathbf{q}_2 + \mathbf{q}_3 = (3^{\frac{1}{2}})\mathbf{p}^{(3)} = 0, \tag{5.1}$$

$$\mathbf{q}_1^2 + \mathbf{q}_2^2 + \mathbf{q}_3^2 = 2mT. \tag{5.2a}$$

Note that by using Eq. (5.1), Eq. (5.2a) can be written in two other forms,

$$\mathbf{q}_1^2 + \mathbf{q}_2^2 + \mathbf{q}_3^2 = \mathbf{p}^{(1)2} + \mathbf{p}^{(2)2} = \mathbf{p}^2, \tag{5.2b}$$

$$(\mathbf{q}_1 - \mathbf{q}_2)^2 + (\mathbf{q}_2 - \mathbf{q}_3)^2 + (\mathbf{q}_3 - \mathbf{q}_1)^2 = 3\mathbf{p}^2. \tag{5.2c}$$

Consider the "momentum" triangle whose vertices are the endpoints of the three momentum vectors directed from a common origin, the center-of-momentum. By Eq. (5.1), the vectors lie entirely within the plane of the momentum triangle. A triangle is kinematically equivalent under rotations to a rigid body. Consequently, it requires three parameters to specify its orientation in space. These may conveniently be taken to be the three Euler angles α, β, γ of the rotation $\mathcal{R}$ required to transform the momentum triangle from some standard "reference" orientation to its actual orientation. The specific choice of body-fixed axes will be described later.

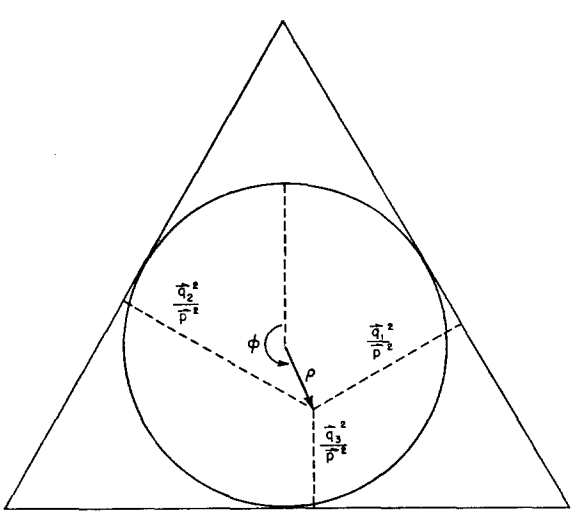

FIG. 2. Dalitz–Fabri coordinates.

Having dealt with the orientation of the momentum triangle, we are left with the problem of parameterizing the triangle itself. This is most easily done by the elegant method of Dalitz.[17] Consider an equilateral triangle (not to be confused with the momentum triangle) of unit altitude (see Fig. 2). It is easily verified that the sum of the three distances from an interior point to each of the three sides is the same for each interior point, and equal to one. We identify the three distances with the quantities $\mathbf{q}_i^2/\mathbf{p}^2$. By this construction, Eq. (5.2) is automatically satisfied. Now, following Fabri,[18] we introduce polar coordinates ρ and ϕ within the Dalitz triangle by the relations

$$\mathbf{q}_i^2 = \tfrac{1}{3}\mathbf{p}^2(1 + \rho\xi_i), \tag{5.3}$$

with

$$\xi_1 = \cos(\phi - \tfrac{2}{3}\pi), \tag{5.4a}$$

$$\xi_2 = \cos(\phi + \tfrac{2}{3}\pi), \tag{5.4b}$$

$$\xi_3 = \cos\phi. \tag{5.4c}$$

The ξ_i satisfy the relations

$$\sum \xi_i = 0, \tag{5.5a}$$

$$\sum \xi_i^2 = \tfrac{3}{2}, \tag{5.5b}$$

$$\sum \xi_i\xi_{i+1} = -\tfrac{3}{4}, \tag{5.5c}$$

$$\prod \xi_i = \tfrac{1}{4}\cos 3\phi. \tag{5.5d}$$

We know that Eq. (5.2) is automatically satisfied for all values of ρ and ϕ by construction. What about Eq. (5.1)? For the momenta to be real, they must satisfy the inequalities

$$(|\mathbf{q}_j| - |\mathbf{q}_k|)^2 \leq \mathbf{q}_i^2 \leq (|\mathbf{q}_j| + |\mathbf{q}_k|)^2 \tag{5.6}$$

or

$$4\mathbf{q}_j^2\mathbf{q}_k^2 \geq (\mathbf{q}_i^2 - \mathbf{q}_j^2 - \mathbf{q}_k^2)^2. \tag{5.7}$$

Insertion of the relations (5.3) into (5.7) transforms the inequalities to the form

$$4(1 + \rho\xi_j)(1 + \rho\xi_k) \geq (2\rho\xi_i - 1)^2. \tag{5.8}$$

A summation over cyclic permutations and use of Eqs. (5.5) gives the final result

$$\rho^2 \leq 1. \tag{5.9}$$

[17] R. H. Dalitz, Repts. Progr. Phys. **20**, 163 (1957).
[18] E. Fabri, Nuovo Cimento **11**, 479 (1954).

Consequently, the kinematically allowed region for our new coordinates is the unit disk inscribed within the Dalitz triangle.

We still have to specify the choice of body-fixed axes. Let ψ be an angle defined by the relation

$$\rho = \cos 2\psi. \tag{5.10}$$

We define the body-fixed X, Y plane to be the plane of the momentum triangle. The choice of unit vectors $\mathbf{e}_x$ and $\mathbf{e}_y$ within the plane is made according to the prescription[19]

$$\begin{aligned} \mathbf{p}^{(1)}\cdot\mathbf{e}_x &= p\cos\psi\sin\tfrac{1}{2}\phi, \\ \mathbf{p}^{(1)}\cdot\mathbf{e}_y &= -p\sin\psi\cos\tfrac{1}{2}\phi, \\ \mathbf{p}^{(2)}\cdot\mathbf{e}_x &= p\cos\psi\cos\tfrac{1}{2}\phi, \\ \mathbf{p}^{(2)}\cdot\mathbf{e}_y &= p\sin\psi\sin\tfrac{1}{2}\phi. \end{aligned} \tag{5.11}$$

The third unit vector, $\mathbf{e}_z$, is chosen by the right-hand rule,

$$\mathbf{e}_z = \mathbf{e}_x \times \mathbf{e}_y. \tag{5.12}$$

The origin of the unit vectors is, of course, the center of momentum. It is sufficient to use the difference vectors $\mathbf{p}^{(1)}$ and $\mathbf{p}^{(2)}$ since $\mathbf{p}^{(3)}$ vanishes in the center-of-momentum frame.

With this fact in mind, it is easy to verify that Eqs. (5.11) are compatible with our earlier definitions in Eqs. (5.3) and (5.4). Define quantities q_{ij}^2 by the relations

$$q_{ij}^2 = (\mathbf{q}_i - \mathbf{q}_j)^2. \tag{5.13}$$

Using Eqs. (5.1) and (5.3), we find

$$q_{ij}^2 = \mathbf{p}^2(1 - \rho\xi_k), \tag{5.14}$$

or

$$3\mathbf{q}_k^2 - q_{ij}^2 = 2\mathbf{p}^2\rho\xi_k. \tag{5.15}$$

The left-hand side of Eq. (5.15) can be evaluated with the aid of Eqs. (1.14, 5.11). One finds, after some algebra, that

$$\begin{aligned} \rho\xi_1 &= \cos 2\psi\cos(\phi - \tfrac{2}{3}\pi), \\ \rho\xi_2 &= \cos 2\psi\cos(\phi + \tfrac{2}{3}\pi), \\ \rho\xi_3 &= \cos 2\psi\cos\phi, \end{aligned} \tag{5.16}$$

as expected.

Finally, the unit vectors have been so chosen that

$$\sum_i (\mathbf{e}_x\cdot\mathbf{q}_i)(\mathbf{e}_y\cdot\mathbf{q}_i) = 0. \tag{5.17}$$

Thus, if we imagine equal masses to be placed at the vertices of the momentum triangle, the body-fixed axes become the principle axes of intertia.

In summary, the result of the discussion so far has been to replace the nine variables $\mathbf{q}_1$ to $\mathbf{q}_3$ subject to the constraint (5.1) by the variables $\mathbf{p}^2$, ρ, ϕ, and three Euler angles. The constraint can now be removed by performing a Lorentz transformation and adding $\mathbf{p}^{(3)}$ to the list of new variables. A routine calculation shows that the volume elements of the two sets of variables are related by the equation

$$d^3\mathbf{q}_1\, d^3\mathbf{q}_2\, d^3\mathbf{q}_3 = \tfrac{1}{8}\, d^3\mathbf{p}^{(3)} p^5\, dp\, d\mathcal{R}_3 \rho\, d\rho\, d\phi, \tag{5.18}$$

where $d\mathcal{R}_3$ is the volume element for the three-dimensional rotation group. In terms of the Euler angles, $d\mathcal{R}_3$ is given by the expression

$$d\mathcal{R}_3 = d\alpha\, d\gamma \sin\beta\, d\beta. \tag{5.19}$$

B. Position Coordinates

The choice of Dalitz–Fabri position coordinates can be made in complete analogy to the momentum case. We first pass to the center-of-mass frame defined by

$$\mathbf{r}_1 + \mathbf{r}_2 + \mathbf{r}_3 = (3^{\frac{1}{2}})\mathbf{r}^{(3)} = 0, \tag{5.20}$$

and define $\mathbf{r}^2$ by relations analogous to Eqs. (5.2b, c),

$$\mathbf{r}_1^2 + \mathbf{r}_2^2 + \mathbf{r}_3^2 = \mathbf{r}^{(1)^2} + \mathbf{r}^{(2)^2} = \mathbf{r}^2, \tag{5.21a}$$

or

$$(\mathbf{r}_1 - \mathbf{r}_2)^2 + (\mathbf{r}_2 - \mathbf{r}_3)^2 + (\mathbf{r}_3 - \mathbf{r}_1)^2 = 3\mathbf{r}^2. \tag{5.21b}$$

We then introduce quantities ρ' and ϕ' defined by the counterparts of Eqs. (5.3), (5.4), and (5.11), and three-position Euler angles. In this case the body-fixed axes are in fact the principle axes of inertia for the three-body system. Finally, we remove the constraint (5.20) by performing a translation and adding $\mathbf{r}^{(3)}$ to the list of new variables. The relation between the volume elements is given by an expression analogous to Eq. (5.18),

$$d^3\mathbf{r}_1\, d^3\mathbf{r}_2\, d^3\mathbf{r}_3 = \tfrac{1}{8}\, d^3\mathbf{r}^{(3)} r^5\, dr\, d\mathcal{R}_3' \rho'\, d\rho'\, d\phi'. \tag{5.22}$$

6. THE CONSTRUCTION OF EIGENSTATES FOR SU_3

We now have at our disposal all the necessary apparatus for a classification of three-body states according to SU_3. Our aim is to demonstrate the existence of a complete set of three-particle states $|\mathbf{p}^{(3)}, p; n\rangle$ which are diagonal in the total momentum $\mathbf{p}^{(3)}$ and center-of-mass energy p, and which simultaneously form a carrier space for irreducible representations of L_2. The elements of L_2 commute, of

[19] Similar coordinates have been employed by other authors. See, for example, Ref. 4 or B. F. Bayman, *Proceedings of the Eastern Theoretical Physics Conference, 1962*, edited by M. Rose (Gordon and Breach, New York, 1963), p. 81 ff.

course, with $\mathbf{p}^{(3)}$ and p by construction. For the moment, we use an index n to distinguish between states having the same energy and momentum but different properties under L_2.

A. The Radial Wavefunction

Let $|\mathbf{r}^{(1)}, \mathbf{r}^{(2)}, \mathbf{r}^{(3)}\rangle$ denote the state diagonal in the position operators $\mathbf{r}^{(1)}$ to $\mathbf{r}^{(3)}$, or equivalently $\mathbf{r}_1$ to $\mathbf{r}_3$. The state $|\mathbf{p}^{(3)}, p; n\rangle$ is completely specified by the function $\langle \mathbf{r}^{(1)}, \mathbf{r}^{(2)}, \mathbf{r}^{(3)} \mid \mathbf{p}^{(3)}, p; n\rangle$. Since $\mathbf{p}^{(3)}$ and $\mathbf{r}^{(3)}$ are canonically conjugate, we have the immediate simplification

$$\langle \mathbf{r}^{(1)}, \mathbf{r}^{(2)}, \mathbf{r}^{(3)} \mid \mathbf{p}^{(3)}, p; n\rangle = \exp(i\mathbf{p}^{(3)}\cdot\mathbf{r}^{(3)})\langle \mathbf{r}^{(1)}, \mathbf{r}^{(2)}, 0 \mid 0, p; n\rangle. \tag{6.1}$$

For notational simplicity, let

$$\psi = \langle \mathbf{r}^{(1)}, \mathbf{r}^{(2)}, 0 \mid 0, p; n\rangle. \tag{6.2}$$

The wavefunction ψ must satisfy the differential equation

$$-\nabla_6^2\psi = p^2\psi, \tag{6.3}$$

where $(1/i)\nabla_6$ is the differential-operator analog of $\mathbf{p}$.

Consider the Casimir operator Λ^2 of Sec. 2. It is easily verified that $\mathbf{p}$ and Λ^2 are related by the identity

$$\mathbf{p}^2 = \frac{\Lambda^2}{r^2} + p_r^2 - 5ir^{-1}p_r \tag{6.4}$$

where p_r is defined by the equation

$$p_r = r^{-1}\mathbf{r}\cdot\mathbf{p}. \tag{6.5}$$

Therefore, ψ may be written as a product of radial and angular wavefunctions in the form

$$\psi = f(r)g(\rho\phi\alpha\beta\gamma) \tag{6.6}$$

with f and g satisfying the equations

$$\left(\frac{\partial^2}{\partial r^2} + \frac{5}{r}\frac{\partial}{\partial r} - \frac{\lambda(\lambda+4)}{r^2} + p^2\right)f = 0, \tag{6.7a}$$

$$\hat{\Lambda}^2 g = \lambda(\lambda+4)g. \tag{6.7b}$$

The arguments of g are position Dalitz–Fabri coordinates with primes omitted for convenience. The quantity $\hat{\Lambda}^2$ is the differential operator analog of Λ^2. We have written its eigenvalues in the form $\lambda(\lambda+4)$ in anticipation of future results.

The fundamental solutions of Eq. (6.7a) are $r^{-2}J_{\lambda+2}(pr)$ and $r^{-2}N_{\lambda+2}(pr)$ where J and N are Bessel functions of the first and second kind. If we stipulate that f be regular at the origin, the solution containing N is excluded. We conclude that

$$\langle \mathbf{r}^{(1)}, \mathbf{r}^{(2)}, \mathbf{r}^{(3)} \mid \mathbf{p}^{(3)}, p; n\rangle = g\exp(i\mathbf{p}^{(3)}\cdot\mathbf{r}^{(3)})r^{-2}J_{\lambda+2}(pr), \tag{6.8}$$

where g is still undetermined.

B. Angular Wavefunctions and the Harmonic Oscillator.

The determination of g is facilitated by a small trick. Consider the "harmonic oscillator" generalization of Eq. (6.3) given by

$$(-\nabla_6^2 + r^2)\psi_1 = 2E\psi_1. \tag{6.9}$$

Using Eq. (6.4) we may again write the solution in the form

$$\psi_1 = f_1 g_1. \tag{6.10}$$

This time f_1 must satisfy the equation

$$\left(\frac{\partial^2}{\partial r^2} + \frac{5}{r}\frac{\partial}{\partial r} - \frac{\lambda(\lambda+4)}{r^2} - r^2 + 2E\right)f_1 = 0. \tag{6.11}$$

However, g_1 satisfies the same equation as g. Therefore, g can be determined by solving the harmonic oscillator problem and then factoring out the radial part of the wavefunction.

For our purposes, the harmonic oscillator is most easily handled in the Fock representation. We define vector creation operators $\mathbf{A}^\dagger$ and $\mathbf{B}^\dagger$ by the relations

$$\mathbf{A}^\dagger = 2^{-\frac{1}{2}}(\mathbf{r}^{(1)} - i\mathbf{p}^{(1)}), \tag{6.12a}$$

$$\mathbf{B}^\dagger = 2^{-\frac{1}{2}}(\mathbf{r}^{(2)} - i\mathbf{p}^{(2)}). \tag{6.12b}$$

They obey the usual commutation rules with their conjugate destruction operators,

$$[A_i, A_j^\dagger] = \delta_{ij}, \text{ etc.} \tag{6.13}$$

Inversion of Eqs. (6.12) and their Hermitian counterparts gives the expressions

$$\mathbf{p}^{(1)} = (i/2^{\frac{1}{2}})(\mathbf{A}^\dagger - \mathbf{A}), \qquad \mathbf{r}^{(1)} = (1/2^{\frac{1}{2}})(\mathbf{A}^\dagger + \mathbf{A}), \text{ etc.} \tag{6.14}$$

Consequently, the Hamiltonian leading to Eq. (6.9) is

$$\mathcal{H} = (\mathbf{A}^\dagger\cdot\mathbf{A} + \mathbf{B}^\dagger\cdot\mathbf{B}). \tag{6.15}$$

Finally, we recall that the states generated by arbitrary polynomials in $\mathbf{A}^\dagger$ and $\mathbf{B}^\dagger$ acting on the vacuum state $|0\rangle$ form a complete basis. The vacuum state is uniquely characterized by the property

$$\mathcal{H}\,|0\rangle = 0, \tag{6.16}$$

and has the position representation

$$\langle \mathbf{r}^{(1)}, \mathbf{r}^{(2)} \mid 0\rangle \sim \exp(-\tfrac{1}{2}r^2). \tag{6.17}$$

The first step is to express the elements of L_1 in terms of $\mathbf{A}^\dagger$ and $\mathbf{B}^\dagger$. This is easily done from the definitions given in Eqs. (2.5), (2.11), and (2.15), with the aid of Eq. (6.14). One finds

$$\mathbf{J} = i\mathbf{A}\times\mathbf{A}^\dagger + i\mathbf{B}\times\mathbf{B}^\dagger, \tag{6.18}$$

$$K_{ik} = i[A_iB_k^\dagger + A_kB_i^\dagger - (A_i^\dagger B_k + A_k^\dagger B_i)]. \tag{6.19}$$

We observe that the expression for $\mathbf{K}$ is quite complicated in that it contains products of $\mathbf{A}^\dagger$ with $\mathbf{B}$ etc. It is therefore natural to pass to new variables $\mathbf{a}^\dagger$ and $\mathbf{b}^\dagger$ by the canonical transformation

$$\begin{aligned}\mathbf{a}^\dagger &= 2^{-\frac{1}{2}}(\mathbf{B}^\dagger + i\mathbf{A}^\dagger),\\ \mathbf{b}^\dagger &= 2^{-\frac{1}{2}}(\mathbf{B}^\dagger - i\mathbf{A}^\dagger),\end{aligned} \tag{6.20}$$

or

$$\begin{aligned}\mathbf{A}^\dagger &= (i/2^{\frac{1}{2}})(\mathbf{b}^\dagger - \mathbf{a}^\dagger),\\ \mathbf{B}^\dagger &= (1/2^{\frac{1}{2}})(\mathbf{b}^\dagger + \mathbf{a}^\dagger).\end{aligned} \tag{6.21}$$

The operators $\mathfrak{IC}$, $\mathbf{J}$, and $\mathbf{K}$ now take the more pleasing form

$$\mathfrak{IC} = (\mathbf{a}^\dagger\cdot\mathbf{a} + \mathbf{b}^\dagger\cdot\mathbf{b}), \tag{6.22}$$

$$\mathbf{J} = i\mathbf{a}\times\mathbf{a}^\dagger + i\mathbf{b}\times\mathbf{b}^\dagger, \tag{6.23}$$

$$K_{jk} = b_j^\dagger b_k + b_k^\dagger b_j - (a_j^\dagger a_k + a_k^\dagger a_j). \tag{6.24}$$

Our task is to construct eigenstates of $\mathfrak{IC}$ which simultaneously transform irreducibly under L_2. We first look for the bilinear operators which commute with $\mathbf{J}$ and $\mathbf{K}$. They are easily found to be

$$N_a = \mathbf{a}^\dagger\cdot\mathbf{a}, \tag{6.25a}$$

$$N_b = \mathbf{b}^\dagger\cdot\mathbf{b}, \tag{6.25b}$$

$$\Delta_+ = \mathbf{a}^\dagger\cdot\mathbf{b}^\dagger, \tag{6.26a}$$

$$\Delta_- = \mathbf{a}\cdot\mathbf{b}. \tag{6.26b}$$

The quantities N_a and N_b are number operators measuring the number of excitations of types $\mathbf{a}$ and $\mathbf{b}$. They of course commute with each other. Since they commute with $\mathfrak{IC}$ and the elements of L_1 as well, we may require that our states be simultaneous eigenstates of N_a and N_b with eigenvalues n_a and n_b.

What are $\Delta_\pm$ good for? From the commutators of $\Delta_\pm$ with the number operators, one sees that $\Delta_\pm$ creates or destroys pairs of $\mathbf{a}$ and $\mathbf{b}$ excitations,

$$[N_{a,b}, \Delta_\pm] = \pm\Delta_\pm. \tag{6.27}$$

Therefore, if we have a representation of L_2 labeled by n_a and n_b, we get back another one labeled by $n_a \pm 1$ and $n_b \pm 1$ by applying $\Delta_\pm$. Moreover, since $\Delta_\pm$ commutes with L_2, we actually get back the same representation. The only exception is the case in which Δ_- annihilates all the vectors in a representation. Since we want the numbers n_a and n_b to specify something unique about a representation, we will therefore impose the condition that all our states be annihilated by Δ_-.

This condition does something more for us. From Eq. (6.22) we see that application of Δ_- to a state lowers the energy by two units. On the other hand, Δ_- commutes with Λ^2 by Eq. (2.14) so that the value λ remains unchanged. Therefore, the radial part f_1 of ψ_1 must be that solution to Eq. (6.11) which has the lowest energy compatible with a given value of λ. Thus, f_1 is given by the expression

$$f_1 = r^\lambda e^{-\frac{1}{2}r^2}. \tag{6.28}$$

We can now compute the eigenvalues of Λ^2. Using Eqs. (2.14), (6.23), and (6.24), we find that

$$\Lambda^2 = (N_a + N_b)(N_a + N_b + 4) - 4\Delta_+\Delta_-. \tag{6.29}$$

Since Δ_- has eigenvalue zero for the states of interest, Λ^2 has eigenvalues of the anticipated form $\lambda(\lambda + 4)$ with

$$\lambda = n_a + n_b. \tag{6.30}$$

We note also that

$$S = \tfrac{1}{2}\sum_i K_{ii} = N_b - N_a. \tag{6.31}$$

Therefore S has the eigenvalues $n_b - n_a$. Consequently, the numbers n_a and n_b for a representation can be obtained completely in terms of operators constructed out of elements in L_1.

So far, our states are characterized by the numbers n_a and n_b. That is, they are eigenstates of Λ^2 and S. Since both these operators commute with all the elements of L_2, the numbers n_a and n_b must be related to the Cartan classification indices λ_1 and λ_2 described in Sec. 4A. They are, in fact, identical to the Cartan indices as will be shown later. Assuming this to be the case, our next step is to label states within a representation. For this purpose we will use the operators J^2 and J_z as outlined in Sec. 4D. Our states will then be labeled by n_a, n_b, j, and m; and will satisfy the relations

$$N_{a,b}\,|n_an_bjm\rangle = n_{a,b}\,|n_an_bjm\rangle, \tag{6.32a}$$

$$J^2\,|n_an_bjm\rangle = j(j+1)\,|n_an_bjm\rangle, \tag{6.32b}$$

$$J_z\,|n_an_bjm\rangle = m\,|n_an_bjm\rangle. \tag{6.32c}$$

In addition, there is of course the annihilation requirement,

$$\Delta_-\,|n_an_bjm\rangle = 0. \tag{6.32d}$$

As is well known in the theory of angular momentum, it is really only necessary to have the state with $j_z = j$. The remaining states can then be obtained by successive applications of J_-. We therefore replace Eqs. (6.32b, c) by the requirements

$$\begin{aligned} J_+ \,|n_a n_b jj\rangle &= 0, \\ J_z \,|n_a n_b jj\rangle &= j\,|n_a n_b jj\rangle. \end{aligned} \tag{6.33}$$

In constructing explicit solutions to Eqs. (6.32, 6.33), it will prove useful to have creation and destruction operators in a spherical basis. They are defined by the relations

$$a_1^\dagger = -2^{-\frac{1}{2}}(a_x^\dagger + ia_y^\dagger), \qquad a_0^\dagger = a_z^\dagger, \tag{6.34}$$

$$a_{-1}^\dagger = 2^{-\frac{1}{2}}(a_x^\dagger - ia_y^\dagger), \qquad a_m = (a_m^\dagger)^\dagger, \tag{6.35}$$

and similarly for the b's. The spherical commutation rules are

$$[a_m, a_n^\dagger] = \delta_{m,n}. \tag{6.36}$$

Finally, we note that Δ_- has the spherical form

$$\Delta_- = \sum_m (-1)^m a_m b_{-m}. \tag{6.37}$$

The states for small values of n_a and n_b can now easily be written down. The state with $n_a = n_b = 0$ is the vacuum state itself. Thus

$$|00\ 00\rangle = |0\rangle. \tag{6.38}$$

Naturally, it has zero angular momentum. Next come the states $n_a = 1$, $n_b = 0$ and $n_a = 0$, $n_b = 1$ given by

$$\begin{aligned} |10\ 11\rangle &= a_1^\dagger\,|0\rangle, \\ |01\ 11\rangle &= b_1^\dagger\,|0\rangle. \end{aligned} \tag{6.39}$$

They both have angular momentum 1.

Proceeding to the case $n_a = n_b = 1$, we expect to have three sets of states with angular momenta 0, 1, and 2:

$$|11\ 00\rangle = \mathbf{a}^\dagger\cdot\mathbf{b}^\dagger\,|0\rangle, \tag{6.40a}$$

$$|11\ 11\rangle = (\mathbf{a}^\dagger\times\mathbf{b}^\dagger)_1\,|0\rangle, \tag{6.40b}$$

$$|11\ 22\rangle = a_1^\dagger b_1^\dagger\,|0\rangle. \tag{6.40c}$$

However, Δ_- fails to annihilate the state (6.40a) with zero angular momentum. One finds instead that

$$\Delta_-\,|11\ 00\rangle = 3\,|00\ 00\rangle. \tag{6.41}$$

Therefore, the zero angular momentum state must be excluded. This is in accord with Table I. Put another way, the wavefunctions for the $|11\ 00\rangle$ and $|00\ 00\rangle$ states differ only in the radial part f_1. Since we are only interested in the angular part g, we get all the information we need by considering states with the simplest radial dependence.

The number of states grows very rapidly for larger values of n_a and n_b. For the present, therefore, we shall only write down states for which $n_a \geq n_b$. The remaining states can easily be obtained merely by interchanging the roles of $\mathbf{a}$ and $\mathbf{b}$. In addition, we shall present the states in order of increasing λ. With this caveat, the remaining states with $\lambda = 2$ are given by the $n_a = 2$, $n_b = 0$ representation,

$$|20\ 00\rangle = \mathbf{a}^\dagger\cdot\mathbf{a}^\dagger\,|0\rangle, \tag{6.42a}$$

$$|20\ 22\rangle = a_1^\dagger a_1^\dagger\,|0\rangle. \tag{6.42b}$$

They have angular momenta 0 and 2.

The set of states for $\lambda = 3$ has either $n_a = 3$, $n_b = 0$ or $n_a = 2$, $n_b = 1$. They are given by the expressions

$$\begin{aligned} |30\ 33\rangle &= (a_1^\dagger)^3\,|0\rangle, \\ |30\ 11\rangle &= (\mathbf{a}^\dagger\cdot\mathbf{a}^\dagger)a_1^\dagger\,|0\rangle, \\ |21\ 33\rangle &= (a_1^\dagger)^2 b_1^\dagger\,|0\rangle, \\ |21\ 22\rangle &= a_1^\dagger(\mathbf{a}^\dagger\times\mathbf{b}^\dagger)_1\,|0\rangle, \\ |21\ 11\rangle &= (2\mathbf{a}^\dagger\cdot\mathbf{a}^\dagger b_1^\dagger - \mathbf{a}^\dagger\cdot\mathbf{b}^\dagger a_1^\dagger)\,|0\rangle, \end{aligned} \tag{6.43}$$

and have angular momenta 3, 2, and 1.

It is now obvious from the examples just given that the maximum value of j occurring in any given representation is restricted by the inequality

$$j \leq n_a + n_b = \lambda. \tag{6.44}$$

In fact, for any specified representation, the state with highest j is unique and given by the expression

$$|n_a, n_b, n_a + n_b, n_a + n_b\rangle = (a_1^\dagger)^{n_a}(b_1^\dagger)^{n_b}\,|0\rangle. \tag{6.45}$$

Note that by Eqs. (6.36) and (6.37), Δ_- annihilates the state as required.

We assert that the state with $j = n_a + n_b$ is the state of highest weight in the sense of Cartan. This follows immediately from the observation that successive application of elements in L_2 cannot change the values of n_a and n_b. Therefore, since the state is unique, all other states must have lower values of $J_z = (12)^{\frac{1}{2}}H_1$. Employing Eqs. (3.3d), (4.5b), (6.24), we find H_2 is given by the expression

$$H_2 = \tfrac{1}{6}[3(b_0^\dagger b_0 - a_0^\dagger a_0) + N_a - N_b]. \tag{6.46}$$

Consequently, the highest state has the weights

$$m_1 = (12)^{\frac{1}{2}}(n_a + n_b), \tag{6.47a}$$

$$m_2 = \tfrac{1}{6}(n_a - n_b). \tag{6.47b}$$

Comparing with Eq. (4.7), we find $n_a = \lambda_1$ and $n_b = \lambda_2$ as anticipated.

One may also look for the highest state in the Bargmann–Moshinsky scheme.[20] We need the op-

[20] I am indebted to Professor Moshinsky for discussions and correspondence on his treatment of SU_3.

erators C_{ik} defined by Eq. (4.10). They are given in terms of the a's and b's by the expressions

$$C_{ij} = b_i^\dagger b_j - a_i^\dagger a_j. \tag{6.48}$$

Equations (4.11) and (6.32d) have as their solution the state given by

$$|g_1 g_2 g_3\rangle = (a_x^\dagger)^{n_a}(b_x^\dagger)^{n_b}\,|0\rangle \tag{6.49}$$

with

$$g_1 = n_b, \qquad g_2 = 0, \qquad g_3 = -n_a. \tag{6.50}$$

Therefore k_1 and k_2 have the values $n_a + n_b$ and n_a, respectively.

It is now possible to proceed from the state $|g_1 g_2 g_3\rangle$ to all other states within the representation by successive applications of the C_{ij} with $i > j$. The process is similar in spirit to the familiar "raising" and "lowering" operations applied in the theory of angular momentum. It is, however, considerably more complicated since one has to deal with several different shift operations. An extensive discussion of the entire process has been given by Moshinsky in a series of papers concerned with the construction of polynomial representations for SU_3.[14] Although his particular polynomials differ from the ones required here due to a different realization of the C_{ij}, the underlying algebraic structure is the same. Thus, there already exists a complete prescription for constructing all the states in SU_3, and our problem is, in principle, completely solved.

We shall limit our further discussion to the construction of a restricted subset of SU_3 states. In that circumstance it is still convenient to proceed in a rather more pedestrian way. We shall stop for the moment with the states for $\lambda = 4$. They are given by the expressions

$$\begin{aligned}
|40\ 44\rangle &= (a_1^\dagger)^4\,|0\rangle,\\
|40\ 22\rangle &= \mathbf{a}^\dagger\cdot\mathbf{a}^\dagger(a_1^\dagger)^2\,|0\rangle,\\
|40\ 00\rangle &= (\mathbf{a}^\dagger\cdot\mathbf{a}^\dagger)^2\,|0\rangle,\\
|31\ 44\rangle &= (a_1^\dagger)^3 b_1^\dagger\,|0\rangle,\\
|31\ 33\rangle &= (a_1^\dagger)^2(\mathbf{a}^\dagger\times\mathbf{b}^\dagger)_1\,|0\rangle,\\
|31\ 22\rangle &= (5\mathbf{a}^\dagger\cdot\mathbf{a}^\dagger a_1^\dagger b_1^\dagger - 2\mathbf{a}^\dagger\cdot\mathbf{b}^\dagger a_1^\dagger a_1^\dagger)\,|0\rangle,\\
|31\ 11\rangle &= \mathbf{a}^\dagger\cdot\mathbf{a}^\dagger(\mathbf{a}^\dagger\times\mathbf{b}^\dagger)_1\,|0\rangle,\\
|22\ 44\rangle &= (a_1^\dagger)^2(b_1^\dagger)^2\,|0\rangle,\\
|22\ 33\rangle &= a_1^\dagger b_1^\dagger(\mathbf{a}^\dagger\times\mathbf{b}^\dagger)_1\,|0\rangle,\\
|22\ 22+\rangle &= [5(\mathbf{a}^\dagger\cdot\mathbf{a}^\dagger b_1^\dagger b_1^\dagger + \mathbf{b}^\dagger\cdot\mathbf{b}^\dagger a_1^\dagger a_1^\dagger)\\
&\quad - 8\mathbf{a}^\dagger\cdot\mathbf{b}^\dagger a_1^\dagger b_1^\dagger]\,|0\rangle,\\
|22\ 22-\rangle &= (\mathbf{a}^\dagger\cdot\mathbf{a}^\dagger b_1^\dagger b_1^\dagger - \mathbf{b}^\dagger\cdot\mathbf{b}^\dagger a_1^\dagger a_1^\dagger)\,|0\rangle,\\
|22\ 00\rangle &= [2\mathbf{a}^\dagger\cdot\mathbf{a}^\dagger\mathbf{b}^\dagger\cdot\mathbf{b}^\dagger - (\mathbf{a}^\dagger\cdot\mathbf{b}^\dagger)^2]\,|0\rangle.
\end{aligned} \tag{6.51}$$

We see that the states with $n_a = 2$, $n_b = 2$, and $j = 2$ are doubly degenerate as predicted by Table I. In accordance with our previous discussion, we may break the degeneracy by application of the operator Ω. However, in this instance it is more convenient to employ a different operation, the interchange of a and b. The vectors chosen above are eigenvectors for the interchange operation with eigenvalues ± 1. It should be remarked at this point that it may be best to avoid the use of Ω or Ω' wherever possible, for their eigenvalues are in general irrational. That is, the eigenvalue equations for Ω are algebraic equations whose order is equal to the degeneracy and whose solutions, in general, cannot be expressed in closed form.[21]

C. Explicit Angular Wavefunctions

We shall now find the explicit angular wavefunctions for the harmonic oscillator states obtained so far. As discussed earlier, these angular functions are those needed for the SU_3 classification of three-particle states.

The harmonic oscillator states are of the general form

$$|n_a n_b j j_z\rangle = P(\mathbf{a}^\dagger, \mathbf{b}^\dagger)\,|0\rangle, \tag{6.52}$$

where P is a polynomial in $\mathbf{a}^\dagger$ and $\mathbf{b}^\dagger$. It is first necessary to undo the canonical transformation (6.21),

$$P(\mathbf{a}^\dagger, \mathbf{b}^\dagger) = P[2^{-\frac{1}{2}}(\mathbf{B}^\dagger + i\mathbf{A}^\dagger),\ 2^{-\frac{1}{2}}(\mathbf{B}^\dagger - i\mathbf{A}^\dagger)]. \tag{6.53}$$

Following our previous discussion, the desired angular functions g are now given by the expressions

$$\begin{aligned}
g = r^{-\lambda} e^{+\frac{1}{2}r^2}\langle \mathbf{r}^{(1)},\\
\mathbf{r}^{(2)}|\, P[2^{-\frac{1}{2}}(\mathbf{B}^\dagger + i\mathbf{A}^\dagger),\ 2^{-\frac{1}{2}}(\mathbf{B}^\dagger - i\mathbf{A}^\dagger)]\,|0\rangle.
\end{aligned} \tag{6.54}$$

The matrix element obeys the simple relation

$$\begin{aligned}
&\langle \mathbf{r}^{(1)}, \mathbf{r}^{(2)}|\, P[2^{-\frac{1}{2}}(\mathbf{B}^\dagger + i\mathbf{A}^\dagger),\ 2^{-\frac{1}{2}}(\mathbf{B}^\dagger - i\mathbf{A}^\dagger)]\,|0\rangle\\
&\qquad = c\, e^{-\frac{1}{2}r^2} P(\mathbf{r}^{(2)} + i\mathbf{r}^{(1)},\ \mathbf{r}^{(2)} - i\mathbf{r}^{(1)}),
\end{aligned} \tag{6.55}$$

where c is an immaterial constant. To see this, we note that it is composed of a sum of terms of the general form

$$\langle \mathbf{r}^{(1)}, \mathbf{r}^{(2)}|\, (A_i^\dagger)^m (B_j^\dagger)^n\,|0\rangle = c\, e^{-\frac{1}{2}r^2} He_m(r_i^{(1)}) He_n(r_j^{(2)}), \tag{6.56}$$

where He_m denotes the Hermite polynomial of order m.[22] The leading term in $He_m(x)$ is x^m. By construc-

[21] Various aspects of the situation have recently been discussed by G. Racah, *Proceedings of the Istanbul International Summer School, 1962* (Gordon and Breach, New York, 1965).

[22] W. Magnus and F. Oberhettinger, *Functions of Mathematical Physics* (Chelsea Publishing Company, New York, 1954), p. 80.

tion, P is a homogeneous polynomial of order λ,

$$P(\alpha x, \alpha y) = \alpha^{\lambda} P(x, y). \tag{6.57}$$

Therefore,

$$\langle \mathbf{r}^{(1)}, \mathbf{r}^{(2)} | \, P[2^{-\frac{1}{2}}(\mathbf{B}^{\dagger} + i\mathbf{A}^{\dagger}), 2^{-\frac{1}{2}}(\mathbf{B}^{\dagger} - i\mathbf{A}^{\dagger})] \, |0\rangle$$
$$= c\, e^{-\frac{1}{2}r^2}[P(\mathbf{r}^{(2)} + i\mathbf{r}^{(1)}, \mathbf{r}^{(2)} - i\mathbf{r}^{(1)}) + P'], \tag{6.58}$$

where P' is a polynomial of lower order in r. But by Eq. (6.28), the radial part of the wavefunction must contain only an exponential multiplied by r^{λ}. This result followed upon imposition of Eq. (6.32d). Therefore P' vanishes identically, and we obtain the simple relation

$$g = cr^{-\lambda} P(\mathbf{r}^{(2)} + i\mathbf{r}^{(1)}, \mathbf{r}^{(2)} - i\mathbf{r}^{(1)}). \tag{6.59}$$

There remains the algebraic task of converting our results to Dalitz–Fabri variables. We shall illustrate the process for a few of the lowest states.

The lowest state is $|00\ 00\rangle$. In this case, g is simply a constant. The next state is $|10\ 1\ m\rangle$. Using Eqs. (6.39) and (6.59), we find

$$\begin{aligned} g(10\ 1m) &= r^{-1}(\mathbf{r}^{(2)} + i\mathbf{r}^{(1)})_m \\ &= r^{-1}\, \mathbf{n}_m \cdot (\mathbf{r}^{(2)} + i\mathbf{r}^{(1)}), \end{aligned} \tag{6.60}$$

where $\mathbf{n}_m$ denotes a member of a triad of space-fixed spherical unit vectors,

$$\mathbf{n}_1 = -2^{-\frac{1}{2}}(\mathbf{n}_x + i\mathbf{n}_y), \qquad \mathbf{n}_0 = \mathbf{n}_z, \tag{6.61}$$
$$n_{-1} = 2^{-\frac{1}{2}}(\mathbf{n}_x - i\mathbf{n}_y).$$

The normalization is

$$\mathbf{n}_m \cdot \mathbf{n}_{m'} = (-1)^m \delta_{m,-m'}. \tag{6.62}$$

We shall also need a spherical body-fixed triad defined by

$$\mathbf{e}_1 = -2^{-\frac{1}{2}}(\mathbf{e}_x + i\mathbf{e}_y), \qquad \mathbf{e}_0 = \mathbf{e}_z, \tag{6.63}$$
$$\mathbf{e}_{-1} = 2^{-\frac{1}{2}}(\mathbf{e}_x - i\mathbf{e}_y).$$

Let I be the unit dyad composed of the unit vectors $\mathbf{e}_m$,

$$I = \sum_m (-1)^m \mathbf{e}_m \mathbf{e}_{-m}. \tag{6.64}$$

Equation (6.60) can now be written in the form

$$g(10\ 1m') = \frac{1}{r} \sum_m (-1)^m \mathbf{n}_{m'} \cdot \mathbf{e}_{-m} \mathbf{e}_m \cdot (\mathbf{r}^{(2)} + i\mathbf{r}^{(1)}). \tag{6.65}$$

Let $\mathcal{R}$ be that rotation which takes the space-fixed triad into the body-fixed triad,

$$\mathbf{e}_i = \mathcal{R}\mathbf{n}_i, \qquad i = 1, 2, 3. \tag{6.66}$$

The scalar products $\mathbf{n}_{m'} \cdot \mathbf{e}_{-m}$ are related to the spherical matrix elements of $\mathcal{R}$. One finds

$$(-1)^m \mathbf{n}_{m'} \cdot \mathbf{e}_{-m} = D^{1*}_{m',m}(\alpha\beta\gamma), \tag{6.67}$$

where α, β, and γ denote the three Euler angles parameterizing $\mathcal{R}$.[23] Combining Eqs. (6.65) and (6.67), we may write

$$g(10\ 1m') = \frac{1}{r} \sum_m D^{1*}_{m',m}(\alpha\beta\gamma) \mathbf{e}_m \cdot (\mathbf{r}^{(2)} + i\mathbf{r}^{(1)}). \tag{6.68}$$

The scalar products $\mathbf{e}_m \cdot (\mathbf{r}^{(2)} + i\mathbf{r}^{(1)})$ can be evaluated by using the position-space analog of Eqs. (5.11). Thus,

$$\begin{aligned} \mathbf{e}_1 \cdot (\mathbf{r}^{(2)} + i\mathbf{r}^{(1)}) &= -2^{-\frac{1}{2}} r(\cos\psi + \sin\psi)\, e^{i\frac{1}{2}\phi}, \\ \mathbf{e}_0 \cdot (\mathbf{r}^{(2)} + i\mathbf{r}^{(1)}) &= 0, \\ \mathbf{e}_{-1} \cdot (\mathbf{r}^{(2)} + i\mathbf{r}^{(1)}) &= 2^{-\frac{1}{2}} r(\cos\psi - \sin\psi)\, e^{i\frac{1}{2}\phi}. \end{aligned} \tag{6.69}$$

Inserting Eqs. (6.69) into Eq. (6.68), we find

$$\begin{aligned} g(10\ 1m) = -2^{-\frac{1}{2}}[&(\sin\psi + \cos\psi) D^{1*}_{m,1}(\alpha\beta\gamma) \\ &+ (\sin\psi - \cos\psi) D^{1*}_{m,-1}(\alpha\beta\gamma)]\, e^{i\frac{1}{2}\phi}. \end{aligned} \tag{6.70}$$

A similar calculation may be carried out for the $|01\ 1m\rangle$ state.

We continue with the state $|11\ 1m\rangle$. By Eqs. (6.40b), (6.59),

$$g(11\ 1m) = r^{-2} \mathbf{n}_m \cdot (\mathbf{r}^{(2)} + i\mathbf{r}^{(1)}) \times (\mathbf{r}^{(2)} - i\mathbf{r}^{(1)}). \tag{6.71}$$

Inserting the unit dyad as before, we find

$$g(11\ 1m') = \frac{1}{r^2} \sum_m D^{1*}_{m',m} \mathbf{e}_m \cdot (\mathbf{r}^{(2)} + i\mathbf{r}^{(1)}) \times (\mathbf{r}^{(2)} - i\mathbf{r}^{(1)}). \tag{6.72}$$

In this case there is only one nonvanishing scalar product,

$$\mathbf{e}_0 \cdot (\mathbf{r}^{(2)} + i\mathbf{r}^{(1)}) \times (\mathbf{r}^{(2)} - i\mathbf{r}^{(1)}) = ir^2 \sin 2\psi. \tag{6.73}$$

Therefore,

$$g(11\ 1m) = i \sin 2\psi D^{1*}_{m0}(\alpha\beta\gamma). \tag{6.74}$$

The angular momentum 2-state for the same representation, $|11\ 2m\rangle$, has for its g the expression

$$g(11\ 22) = r^{-2}\, [\mathbf{n}_{+1} \cdot (\mathbf{r}^{(2)} + i\mathbf{r}^{(1)})][\mathbf{n}_{+1} \cdot (\mathbf{r}^{(2)} - i\mathbf{r}^{(1)})]. \tag{6.75}$$

---

[23] We use the notation and conventions of M. E. Rose, *Elementary Theory of Angular Momentum* (John Wiley & Sons, Inc., New York, 1957).

TABLE II. SU_3 states for $\lambda \leq 3$.

λ	n_a	n_b	j	$g(n_a n_b jm; \rho\phi\alpha\beta\gamma)$
0	0	0	0	Constant
1	1	0	1	$-1/2^{\frac{1}{2}}[(\sin\psi + \cos\psi)D^{1*}_{m,1} + (\sin\psi - \cos\psi)D^{1*}_{m,-1}]e^{i\frac{1}{2}\phi}$
1	0	1	1	$1/2^{\frac{1}{2}}[(\sin\psi - \cos\psi)D^{1*}_{m,1} + (\sin\psi + \cos\psi)D^{1*}_{m,-1}]e^{-i\frac{1}{2}\phi}$
2	2	0	0	$\rho e^{i\phi}$
2	0	2	0	$\rho e^{-i\phi}$
2	2	0	2	$\frac{1}{2}[(1 + \sin 2\psi)D^{2*}_{m,2} - 2\rho/6^{\frac{1}{2}}D^{2*}_{m,0} + (1 - \sin 2\psi)D^{2*}_{m,-2}]e^{i\phi}$
2	0	2	2	$\frac{1}{2}[(1 - \sin 2\psi)D^{2*}_{m,2} - 2\rho/6^{\frac{1}{2}}D^{2*}_{m,0} + (1 + \sin 2\psi)D^{2*}_{m,-2}]e^{-i\phi}$
2	1	1	1	$i \sin 2\psi D^{1*}_{m0}$
2	1	1	2	$\frac{1}{2}[\rho D^{2*}_{m,2} - 2/6^{\frac{1}{2}}D^{2*}_{m,0} + \rho D^{2*}_{m,-2}]$
3	3	0	1	$-1/2^{\frac{1}{2}}\rho[(\sin\psi + \cos\psi)D^{1*}_{m,1} + (\sin\psi - \cos\psi)D^{1*}_{m,-1}]e^{i\frac{1}{2}\phi}$
3	0	3	1	$1/2^{\frac{1}{2}}\rho[(\sin\psi - \cos\psi)D^{1*}_{m,1} + (\sin\psi + \cos\psi)D^{1*}_{m,-1}]e^{-i\frac{1}{2}\phi}$
3	3	0	3	$-1/8^{\frac{1}{2}}[(\sin\psi + \cos\psi)^3 D^{3*}_{m,3} - (3/5)^{\frac{1}{2}}\rho(\sin\psi + \cos\psi)D^{3*}_{m,1}$ $-(3/5)^{\frac{1}{2}}\rho(\sin\psi - \cos\psi)D^{3*}_{m,-1} + (\sin\psi - \cos\psi)^3 D^{3}_{m,-3}]e^{i\frac{3}{2}\phi}$
3	0	3	3	$1/8^{\frac{1}{2}}[(\sin\psi - \cos\psi)^3 D^{3*}_{m,3} - (3/5)^{\frac{1}{2}}\rho(\sin\psi - \cos\psi)D^{3*}_{m,1}$ $-(3/5)^{\frac{1}{2}}\rho(\sin\psi + \cos\psi)D^{3*}_{m,-1} + (\sin\psi + \cos\psi)^3 D^{3*}_{m,-3}]e^{-i\frac{3}{2}\phi}$
3	2	1	1	$1/2^{\frac{1}{2}}\{[(2\rho + 1)\sin\psi - (2\rho - 1)\cos\psi]D^{1*}_{m,1} + [(2\rho + 1)\sin\psi + (2\rho - 1)\cos\psi]D^{1*}_{m,-1}\}e^{i\frac{1}{2}\phi}$
3	1	2	1	$-1/2^{\frac{1}{2}}\{[(2\rho + 1)\sin\psi + (2\rho - 1)\cos\psi]D^{1*}_{m,1} + [(2\rho + 1)\sin\psi - (2\rho - 1)\cos\psi]D^{1*}_{m,-1}\}e^{-i\frac{1}{2}\phi}$
3	2	1	2	$-i/2^{\frac{1}{2}}\{[(\sin\psi + \cos\psi)\sin 2\psi]D^{2*}_{m,1} + [(\sin\psi - \cos\psi)\sin 2\psi]D^{2*}_{m,m-1}\}e^{i\frac{1}{2}\phi}$
3	1	2	2	$i/2^{\frac{1}{2}}\{[(\sin\psi - \cos\psi)\sin 2\psi]D^{2*}_{m,1} + [(\sin\psi + \cos\psi)\sin 2\psi]D^{2*}_{m,-1}\}e^{-i\frac{1}{2}\phi}$
3	2	1	3	$g(20\ 22) \times g(01\ 11)$
3	1	2	3	$g(10\ 11) \times g(02\ 22)$

This time it is necessary to insert two unit dyads. Thus,

$$g(11\ 22) = \frac{1}{r^2}\sum_{m,m'} D^{1*}_{1m}\mathbf{e}_m\cdot(\mathbf{r}^{(2)} + i\mathbf{r}^{(1)}) \times D^{1*}_{1m'}\mathbf{e}_{m'}\cdot(\mathbf{r}^{(2)} - i\mathbf{r}^{(1)}), \tag{6.76}$$

or

$$g(11\ 22) = g(10\ 11)g(01\ 11). \tag{6.77}$$

The product of the two D's may be reduced with the aid of the Clebsch–Gordan series,[23]

$$D^1_{1m}D^1_{1m'} = C(112; 112) \times C(112; m, m', m + m')D^2_{2,m+m'}. \tag{6.78}$$

Therefore,

$$g(11\ 2m) = \tfrac{1}{2}[\rho D^{2*}_{m2}(\alpha\beta\gamma) - (2/6^{\frac{1}{2}})D^{2*}_{m0}(\alpha\beta\gamma) + \rho D^{2*}_{m,-2}(\alpha\beta\gamma)]. \tag{6.79}$$

The remaining wavefunctions can be found by similar calculations. Explicit results are listed in Table II for $\lambda \leq 3$. The states for $\lambda = 4$ are too complicated for easy tabulation, but can be easily found when necessary using Eqs. (6.51). None of the functions are normalized. The proper normalization is easily obtained with the aid of Eq. (5.22) and the relation

$$2\delta(x^2 - y^2) = \int_0^\infty z\,dz J_\nu(xz)J_\nu(yz). \tag{6.80}$$

It is possible to show that all the $j = 0$ wavefunctions can be expressed in terms of Jacobi polynomials. The first step is to observe that certain key states can easily be found explicitly. For example, the state $|n_a 0\ 00\rangle$ is given by

$$|n_a 0\ 00\rangle = (\mathbf{a}^\dagger\cdot\mathbf{a}^\dagger)^{\frac{1}{2}n_a}|0\rangle, \tag{6.81}$$

and

$$g(n_a 0\ 00) = \rho^{\frac{1}{2}n_a}e^{i\frac{1}{2}n_a\phi}. \tag{6.82}$$

Equation (6.32d) is evidently satisfied. Suppose, now, that we know the polynomial in $\mathbf{a}^\dagger$ and $\mathbf{b}^\dagger$ corresponding to the state $|n_a n_b 00\rangle$. The polynomial for the adjacent state, $|n_a + 2, n_b + 2, 00\rangle$, must be composed of a linear combination of the terms from that of the previous state multiplied by pairs of $\mathbf{a}^\dagger$'s and $\mathbf{b}^\dagger$'s. This process of passing from a state to the adjacent state is completely specified by the orthogonality requirement

$$\int g^*(n_a n_b 00)g(n_a' n_b' 00)\rho\,d\rho\,d\phi = c\,\delta_{n_a n_a'}\,\delta_{n_b n_b'}. \tag{6.83}$$

It is thus possible to obtain all the states starting from the known key states by essentially employing

the Schmidt orthogonalization procedure.[24,25] The recursion relations so obtained are identical to those for the Jacobi polynomials. Using one of the standard notations,[26] one finds

$$g(n_a n_b 00) = e^{i\frac{1}{2}(n_a - n_b)\phi} \rho^\sigma \mathcal{J}_\tau(\sigma + 1, \sigma + 1, \rho^2), \quad (6.84)$$

where

$$\sigma = \tfrac{1}{2} |n_a - n_b|, \quad (6.85)$$

$$\tau = \tfrac{1}{4}(n_a + n_b - 2\sigma). \quad (6.86)$$

The method just outlined can also be used to find all the $j = 1$ wavefunctions in terms of Jacobi polynomials. The procedure is more complicated, however, and requires the use of the Christoffel formula.[27] We shall therefore not pursue the matter any further here.

Our discussion so far has been concerned with the construction of spatial wavefunctions. We may summarize our results by saying that we have found states $|\mathbf{p}^{(3)}p; n_a n_b jj_z\rangle$ with the spatial wavefunctions

$$\langle \mathbf{r}^{(1)}\mathbf{r}^{(2)}\mathbf{r}^{(3)} \mid \mathbf{p}^{(3)}p; n_a n_b jm\rangle$$

$$= c\, e^{i\mathbf{p}^{(3)}\cdot\mathbf{r}^{(3)}} r^{-2} J_{n_a+n_b+2}(pr) g(n_a n_b jm; \rho\phi\alpha\beta\gamma), \quad (6.87)$$

where c is a normalization constant. The functions g have been determined for $n_a + n_b \leq 4$ and for all values of n_a, n_b when $j = 0$.

For many applications, it is equally important to have momentum-space wavefunctions,

$$\langle \mathbf{p}^{(1)\prime}\mathbf{p}^{(2)\prime}\mathbf{p}^{(3)\prime} \mid \mathbf{p}^{(3)}p; n_a n_b jm\rangle.$$

They are easily calculated for SU_3 states. The exponential and radial wavefunctions are replaced by momentum and energy delta functions. By Eqs. (2.1), (2.5), and (2.6), the operators Λ_{ij} out of which L_2 has been constructed treat momentum and position operators on an equal footing. Consequently, the "angular" functions g remain unchanged. It is only necessary to replace the position Euler angles and position Dalitz–Fabri coordinates by their momentum counterparts. Therefore,

$$\langle \mathbf{p}^{(1)\prime}\mathbf{p}^{(2)\prime}\mathbf{p}^{(3)\prime} \mid \mathbf{p}^{(3)}p; n_a n_b jm\rangle$$

$$= c_1\, \delta_3(\mathbf{p}^{(3)\prime} - \mathbf{p}^{(3)}) \frac{\delta(p'^2 - p^2)}{p'p} g(n_a n_b jm; \rho\phi\alpha\beta\gamma), \quad (6.88)$$

where $\rho, \phi, \alpha, \beta$, and γ now denote momentum Dalitz–Fabri coordinates and momentum Euler angles. We see that SU_3 states are particularly well adapted to situations requiring a simultaneous knowledge of correlations both in momentum and position space.

D. Symmetry Properties of SU_3 States

We close this section with a discussion of the properties of SU_3 states under the operations of parity and the permutation group S_3. By Eqs. (6.12) and (6.20), the operators $\mathbf{a}^\dagger$ and $\mathbf{b}^\dagger$ change sign under spatial inversion. Therefore, the spatial parity of a general SU_3 state is given by $(-1)^{n_a+n_b}$. According to Table I, states with $j = 0$ can only occur for even values of n_a and n_b. Thus, all $j = 0$ states have even spatial parity. Similarly, the $j = 1$ odd–odd states have even parity while the odd–even states have odd parity.

To study the effect of the permutation group, it is only necessary to consider the transposition $\mathcal{P}_{12}$ and the cyclic permutation $\mathcal{C}$. By the unitary analog of Eq. (1.25) and Eq. (6.31),

$$\mathcal{C}\, |n_a n_b jm\rangle = e^{\frac{2}{3}\pi i(n_b - n_a)} |n_a n_b jm\rangle. \quad (6.89)$$

Therefore, the only effect of a cyclic permutation is multiplication by a phase. Using Eqs. (1.21) and (6.12), one sees that $\mathcal{P}_{12}$ changes the sign of $\mathbf{A}^\dagger$ while leaving $\mathbf{B}^\dagger$ unaffected. Inserting this information into Eq. (6.20), we find that $\mathcal{P}_{12}$ interchanges the roles of $\mathbf{a}^\dagger$ and $\mathbf{b}^\dagger$. Thus, since n_a and n_b are the Cartan indices, $\mathcal{P}_{12}$ maps a given representation of SU_3 into its conjugate representation. By Eqs. (1.48a,b), the value of $\mathbf{J}$ is unchanged while $\mathbf{K}$ changes sign. Therefore, if we denote a general state by $|n_a n_b jm\omega\rangle$ where ω is the eigenvalue of the operator $\Omega = \mathbf{J}\cdot\mathbf{T}\cdot\mathbf{J}$ necessary to break the degeneracy described in Sec. 4D, we find

$$\mathcal{P}_{12}\, |n_a n_b jm\omega\rangle = |n_b n_a jm - \omega\rangle. \quad (6.90)$$

Alternatively, in the case of a self-conjugate representation, we may use $\mathcal{P}_{12}$ to lower the degeneracy as was done in Eq. (6.51).

5. THE EMBEDDING OF SU_3 IN O_6

By Eq. (2.14), the Casimiro perator for $\mathcal{G}'$ can be completely expressed in terms of elements contained within L_1. We shall now explore the consequences of this result for representations of $\mathcal{G}'$ carried by three-particle states.

The Lie algebra L_0 of $\mathcal{G}'$ is simple. It is, in fact, isomorphic to the Lie algebra of the group of unitary unimodular 4×4 matrices, SU_4. We shall therefore

[24] F. Riez and B. Sz.-Nagy, *Functional Analysis* (Frederick Ungar Publishing Company, New York, 1955), p. 67.

[25] The $j = 0$ solutions can also be obtained by a direct separation of variables for the $j = 0$ Schrödinger equation. See V. Gallina, P. Nata, L. Bianchi, and G. Viano, Nuovo Cimento **24**, 835 (1962).

[26] Reference 22, p. 83.

[27] G. Szegö, *Orthogonal Polynomials* (American Mathematical Society, New York, 1959), p. 28.

use the Cartan procedure to classify its representations. The first step is to bring L_0 into Cartan normal form. Let $\mathbf{e}_3$ be a third unit vector orthogonal to the vectors $\mathbf{e}_1$ and $\mathbf{e}_2$ defined in Sec. 4A. We define six more root vectors by the equations

$$\begin{aligned}\boldsymbol{\delta} &= -\tfrac{1}{3}\mathbf{e}_2 + (2^{\frac{1}{2}}/3)\mathbf{e}_3,\\ \boldsymbol{\varepsilon} &= +[2(3^{\frac{1}{2}})]^{-1}\mathbf{e}_1 + \tfrac{1}{6}\mathbf{e}_2 + (2^{\frac{1}{2}}/3)\mathbf{e}_3,\\ \boldsymbol{\zeta} &= -[2(3^{\frac{1}{2}})]^{-1}\mathbf{e}_1 + \tfrac{1}{6}\mathbf{e}_2 + (2^{\frac{1}{2}}/3)\mathbf{e}_3,\end{aligned} \tag{7.1}$$

and their negatives. These vectors, along with those defined by Eqs. (4.1) and their negatives, provide the twelve root vectors for the Cartan normal form of O_6 or SU_4.[28] We also define six more operators by relations similar to Eqs. (1.45) and (2.15),

$$\begin{aligned}V_i &= \tfrac{1}{2}\epsilon_{ijk}(\Lambda_{jk} - \Lambda_{j+3,k+3}),\\ W_i &= \tfrac{1}{2}\epsilon_{ijk}(\Lambda_{j,k+3} + \Lambda_{j+3,k}),\end{aligned} \qquad i \leq 3. \tag{7.2}$$

The six ladder operators corresponding to the root vectors (7.1) are now given by the expressions

$$\begin{aligned}E(\pm\boldsymbol{\delta}) &= (24)^{-\frac{1}{2}}(V_3 \pm iW_3),\\ E(\pm\boldsymbol{\varepsilon}) &= (48)^{-\frac{1}{2}}[V_1 - W_2 \pm i(W_1 + V_2)],\\ E(\pm\boldsymbol{\zeta}) &= (48)^{-\frac{1}{2}}[V_1 + W_2 \pm i(W_1 - V_2)].\end{aligned} \tag{7.3}$$

Finally, the third commuting operator is

$$H_3 = (18)^{-\frac{1}{2}}S. \tag{7.4}$$

The 15 operators of Eqs. (4.5), (7.3), and (7.4) together form the Cartan basis for L_0.

Let $|m_1m_2m_3\rangle$ denote the state of highest weight for a given representation of L_0. We have the relations

$$H_i\,|\mathbf{m}\rangle = m_i\,|\mathbf{m}\rangle, \qquad i = 1, 2, 3; \tag{7.5}$$

and

$$E(\boldsymbol{\mu})\,|\mathbf{m}\rangle = 0, \tag{7.6}$$

where $\boldsymbol{\mu}$ is one of the 6 root vectors $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, $-\boldsymbol{\gamma}$, $-\boldsymbol{\delta}$, $\boldsymbol{\varepsilon}$, and $-\boldsymbol{\zeta}$. Equation (7.6) follows from the stipulation that $|\mathbf{m}\rangle$ be the state of highest weight, for if $E(\boldsymbol{\mu})\,|\mathbf{m}\rangle$ did not vanish, it would be a state of still higher weight. In the Cartan basis, Eq. (2.14) takes the form

$$\begin{aligned}&6\Big[\sum_1^3 H_i^2 + \sum_{\boldsymbol{\mu}\in SU_4} E(\boldsymbol{\mu})E(-\boldsymbol{\mu})\Big]\\ &\quad = 12\Big[\sum_1^2 H_i^2 + \sum_{\boldsymbol{\mu}\in SU_3} E(\boldsymbol{\mu})E(-\boldsymbol{\mu})\Big] - 6H_3^2.\end{aligned} \tag{7.7}$$

Now apply both sides of Eq. (7.7) to the state of highest weight. Using Eqs. (7.5, 6) and the commutation rules (4.3), one finds

$$2m_3^2 - (m_1^2 + m_2^2) - \left(\frac{m_1}{3^{\frac{1}{2}}} - \frac{m_2}{3} + \frac{2^{\frac{1}{2}}m_3}{3}\right) = 0. \tag{7.8}$$

The highest weight for SU_4 or O_6 is given by a formula analogous to Eq. (4.7),

$$\begin{aligned}\mathbf{m} = {}&\tfrac{1}{6}\lambda_1(3^{\frac{1}{2}}\mathbf{e}_1 + \mathbf{e}_2 + 2^{-\frac{1}{2}}\mathbf{e}_3)\\ &+ \tfrac{1}{6}\lambda_2(3^{\frac{1}{2}}\mathbf{e}_1 - \mathbf{e}_2 - 2^{-\frac{1}{2}}\mathbf{e}_3)\\ &+ \tfrac{1}{6}\lambda_3(3^{\frac{1}{2}}\mathbf{e}_1 + \mathbf{e}_2 - 2^{\frac{1}{2}}\mathbf{e}_3),\end{aligned} \tag{7.9}$$

where λ_1, λ_2, and λ_3 are arbitrary nonnegative integers. Thus, Eq. (7.8) implies the relation

$$(\lambda_1 + \lambda_2)(\lambda_1 + \lambda_2 + 2) + 4\lambda_1\lambda_3 = 0 \tag{7.10}$$

among the Cartan indices. Since the indices cannot be negative, we must have $\lambda_1 = \lambda_2 = 0$. Therefore, the representation of O_6 or SU_4 carried by three-particle states can have only one nonzero index, λ_3. Applying the expression for Λ^2 in the Cartan basis [the left-hand side of Eq. (7.7)] to the state of highest weight, one finds

$$\lambda_3 = \lambda = n_a + n_b. \tag{7.11}$$

We may summarize our results by saying that the spherical polynomials for the sphere embedded in six-dimensional Euclidean space can be put into one-to-one correspondence with the irreducible representations of SU_3.[29]

8. APPLICATION TO THREE-BODY DECAYS

As an example of the utility of SU_3 states, we shall treat the decay process $K^+ \to \pi^+ + \pi^+ + \pi^-$. For simplicity, we shall assume that the K meson has spin zero and that a $\Delta I = \frac{1}{2}$ rule is in force so that the three pions have $I = 1$.

The treatment of isospin is facilitated by the introduction of isospin creation operators $\boldsymbol{\alpha}^\dagger$, $\boldsymbol{\beta}^\dagger$, $\boldsymbol{\gamma}^\dagger$ which create the isospin states for particles 1, 2, and 3, respectively, when acting on an isospin vacuum state, $|\text{vac}\rangle$. A linearly independent set of $I = 1$ states is easily seen to be $\boldsymbol{\alpha}^\dagger\boldsymbol{\beta}^\dagger\cdot\boldsymbol{\gamma}^\dagger\,|\text{vac}\rangle$, $\boldsymbol{\beta}^\dagger\boldsymbol{\gamma}^\dagger\cdot\boldsymbol{\alpha}^\dagger\,|\text{vac}\rangle$, and $\boldsymbol{\gamma}^\dagger\boldsymbol{\alpha}^\dagger\cdot\boldsymbol{\beta}^\dagger\,|\text{vac}\rangle$.

To enforce Bose statistics, we shall need isospin states of definite symmetry. Under the cyclic permu-

[28] The end points of the root vectors for SU_4 (or O_6) lie on the vertices of a cuboctahedron. See H. S. M. Coxeter, *Regular Polytopes* (The Macmillan Company, New York, 1963), plate I. Fig. 9.

[29] The construction of spherical harmonics which transform irreducibly under O_6 but carry pieces of several representations of SU_3 can be carried out by a direct separation of variables in the three-particle Schrödinger equation. For this, one takes as diagonal operators $J_z^{(1)} = \frac{1}{2}(J_z + V_z)$, $J_z^{(2)} = \frac{1}{2}(J_z - V_z)$, $J^{(1)2}$, $J^{(2)2}$, and Λ^2. See ref. 3, and Z. Koba, Acta Phys. Polonica, **22** (Suppl.), 103 (1962) and references cited therein.

tation $\mathcal{C}$, we have $\alpha^\dagger$, $\beta^\dagger$, and $\gamma^\dagger \to \beta^\dagger$, $\gamma^\dagger$, and $\alpha^\dagger$, respectively. We therefore introduce states $|s\rangle$ with $s = 0, \pm 1$ by the prescription

$$\begin{aligned}
|s = 0\rangle &= (15)^{-\frac{1}{2}}[\alpha^\dagger\beta^\dagger\cdot\gamma^\dagger + \beta^\dagger\gamma^\dagger\cdot\alpha^\dagger + \gamma^\dagger\alpha^\dagger\cdot\beta^\dagger]\,|\text{vac}\rangle,\\
|s = 1\rangle &= 6^{-\frac{1}{2}}[e^{-\frac{2}{3}\pi i}\alpha^\dagger\beta^\dagger\cdot\gamma^\dagger\\
&\quad + e^{+\frac{2}{3}\pi i}\beta^\dagger\gamma^\dagger\cdot\alpha^\dagger + \gamma^\dagger\alpha^\dagger\cdot\beta^\dagger]\,|\text{vac}\rangle,\\
|s = -1\rangle &= 6^{-\frac{1}{2}}[e^{+\frac{2}{3}\pi i}\alpha^\dagger\beta^\dagger\cdot\gamma^\dagger\\
&\quad + e^{-\frac{2}{3}\pi i}\beta^\dagger\gamma^\dagger\cdot\alpha^\dagger + \gamma^\dagger\alpha^\dagger\cdot\beta^\dagger]\,|\text{vac}\rangle. \qquad (8.1)
\end{aligned}$$

It is easily verified that the new states are orthonormal,

$$\langle s' \mid s\rangle = \delta_{s's}, \qquad (8.2)$$

and satisfy the relation

$$\mathcal{C}\,|s\rangle = e^{\frac{2}{3}\pi i s}\,|s\rangle. \qquad (8.3)$$

Under the transposition $\mathcal{P}_{12}$, $\alpha^\dagger \leftrightarrow \beta^\dagger$ and $\gamma^\dagger \leftrightarrow \gamma^\dagger$, so that

$$\mathcal{P}_{12}\,|s\rangle = |-s\rangle. \qquad (8.4)$$

By combining isospin space and momentum space, we may write an arbitrary three-pion ($I = 1$ and $j = 0$) state in the form

$$|3\pi\rangle = \sum_{n_a, n_b, s} d^s_{n_a n_b}\,|n_a n_b\rangle\,|s\rangle, \qquad (8.5)$$

where we have used the shorthand notation $|n_a n_b\rangle$ for the SU_3 states $|\mathbf{p}^{(3)}p; n_a n_b j = 0\, m = 0\rangle$. If we require Bose statistics, the 3π state must be invariant under both $\mathcal{P}_{12}$ and $\mathcal{C}$. Combining Eqs. (6.89), (6.90), (8.3), and (8.4), gives the relations

$$d^s_{n_a n_b} = d^{-s}_{n_b n_a} \qquad (8.6)$$

and

$$d^s_{n_a n_b} \sin \tfrac{1}{3}\pi(n_b - n_a + s) = 0, \qquad (8.7)$$

or

$$d^s_{n_a n_b} = 0 \qquad (8.8)$$

unless

$$n_b - n_a + s = 0 \text{ modulo } 3. \qquad (8.9)$$

Turning to the specific problem of K^+ decay, we employ an S matrix formalism and write

$$\frac{1}{2i}(S - 1)\,|K^+\rangle = |3\pi\rangle + \text{other decay modes}, \qquad (8.10)$$

where the 3π state has the expansion given in Eq. (8.5) subject to the conditions (8.6) to (8.9). The probability amplitude A for the three emitted pions to have coordinates ρ and ϕ on the momentum Dalitz plot and charges $(++-)$ is given by

$$\begin{aligned}
A &= \langle ++-\rho\phi|\,\frac{1}{2i}(S - 1)\,|K^+\rangle\\
&= \sum_{n_a n_b s} d^s_{n_a n_b}\langle\rho\phi \mid n_a n_b\rangle\langle ++- \mid s\rangle, \qquad (8.11)
\end{aligned}$$

where the $\langle\rho\phi \mid n_a n_b\rangle$ are the functions $g(n_a n_b j m; \rho\phi\alpha\beta\gamma)$ of Sec. 6C with suitable normalization and $|++-\rangle$ denotes the isospin state in which particles 1 to 3 have the charges $(++-)$ respectively. One easily finds from Eq. (8.1) that

$$\begin{aligned}
\langle ++- \mid s = 0\rangle &= 2(15)^{-\frac{1}{2}},\\
\langle ++- \mid s = \pm 1\rangle &= -6^{-\frac{1}{2}}.
\end{aligned} \qquad (8.12)$$

In K decay, the kinetic energy of the pions is about 75 MeV, so that by Eq. (5.2),

$$p \simeq m_\pi. \qquad (8.13)$$

If we assume that the pions are essentially produced within a volume of radius R, we may conclude that SU_3 states with large values of λ are very unlikely to be produced since the spatial wavefunctions are of the form (6.87) and $J_{\lambda+2}$ vanishes strongly at the origin for large values of λ. Our argument is simply a generalization of the notion of a centrifugal barrier. We shall therefore retain only those terms in Eq. (8.11) for which $\lambda \leq 2$ to obtain

$$A \simeq 2[(15\pi)^{-\frac{1}{2}}]e - 2[(3\pi)^{-\frac{1}{2}}]f\rho\cos\phi. \qquad (8.14)$$

Here we have inserted normalized expressions for the g's and set $e = d^0_{00}$ and $f = d^{-1}_{20} = d^1_{02}$. Neglecting $|f|^2$ compared to $|e|^2$, the decay distribution is given by

$$|A|^2 = 4(15\pi)^{-1}\,|e|^2\,[1 + 2(5^{\frac{1}{2}})h\rho\cos\phi], \qquad (8.15)$$

where

$$h = -\text{Re}\,(f/e). \qquad (8.16)$$

Comparing with experimental data, we find $h = 0.06$.[30] We conclude that the pions are produced almost entirely in the SU_3 state with $\lambda = 0$ as anticipated by the centrifugal-barrier argument.

It is of interest to ask what value of R for the radius of the production volume leads to a value of h with the observed magnitude. We may implement the notion of a production volume by using the function $e^{-r/R}$ and setting

$$\left|\frac{f}{e}\right|^2 \simeq \int_0^\infty dr\, r\, e^{-r/R}J_4^2(pr)\left[\int_0^\infty dr\, r\, e^{-r/R}J_2^2(pr)\right]^{-1}. \qquad (8.17)$$

[30] M. Ferro-Luzzi, D. H. Miller, J. J. Murray, A. H. Rosenfeld, and R. D. Tripp, Nuovo Cimento **22**, 1087 (1961). The available data for K decay is not yet sufficient to detect terms in the matrix element going as ρ^2.

One finds

$$\left|\frac{f}{e}\right| \simeq 4.6(\tfrac{1}{2}pR)^2, \tag{8.18}$$

or, using Eq. (8.13),

$$R \sim (4m_\pi)^{-1}. \tag{8.19}$$

This value for R is quite in keeping with one's physical intuition.

ACKNOWLEDGMENTS

I am indebted to Professor V. Bargmann, Professor M. Moshinsky, and Professor F. Dyson for helpful discussions, and to Professor J. R. Oppenheimer for his kind hospitality at the Institute for Advanced Study.

APPENDIX

For the general mass case, it is convenient to introduce the quantities

$$M = (m_1 + m_2 + m_3), \tag{A1}$$

$$\mu = \left(\frac{m_1 m_2 m_3}{M}\right)^{\frac{1}{2}}, \tag{A2}$$

$$d_3 = \left(\mu \frac{m_1 + m_2}{m_1 m_2}\right)^{\frac{1}{2}}, \tag{A3}$$

$$c = 3^{\frac{1}{2}}. \tag{A4}$$

The definitions for $\mathbf{p}^{(1)}$ to $\mathbf{p}^{(3)}$ and $\mathbf{r}^{(1)}$ to $\mathbf{r}^{(3)}$ previously given by Eqs. (1.13) and (2.2) now take the general form

$$\begin{aligned}
\mathbf{p}^{(1)} &= \frac{c\mu}{d_3}\left(\frac{\mathbf{q}_2}{m_2} - \frac{\mathbf{q}_1}{m_1}\right), \\
\mathbf{p}^{(2)} &= c\mu\, d_3\left(\frac{\mathbf{q}_3}{m_3} - \frac{\mathbf{q}_1 + \mathbf{q}_2}{m_1 + m_2}\right), \\
\mathbf{p}^{(3)} &= (1/c^2)(\mathbf{q}_1 + \mathbf{q}_2 + \mathbf{q}_3),
\end{aligned} \tag{A5}$$

and

$$\begin{aligned}
\mathbf{r}^{(1)} &= (1/cd_3)(\mathbf{r}_2 - \mathbf{r}_1), \\
\mathbf{r}^{(2)} &= \frac{d_3}{c}\left(\mathbf{r}_3 - \frac{m_1\mathbf{r}_1 + m_2\mathbf{r}_2}{m_1 + m_2}\right), \\
\mathbf{r}^{(3)} &= (c^2/M)(m_1\mathbf{r}_1 + m_2\mathbf{r}_2 + m_3\mathbf{r}_3).
\end{aligned} \tag{A6}$$

The total energy takes the form

$$\frac{\mathbf{q}_1^2}{m_1} + \frac{\mathbf{q}_2^2}{m_2} + \frac{\mathbf{q}_3^2}{m_3} = \frac{1}{c^2\mu}(\mathbf{p}^{(1)2} + \mathbf{p}^{(2)2}) + \frac{c^4}{M}\mathbf{p}^{(3)2}, \tag{A7}$$

so that $\mathcal{O}_6$ invariance and all its consequences still hold. The generalizations of the remaining equations are straightforward. In particular, Eqs. (5.11) remain unchanged. For further details, the reader is referred to the original work of Smith.[4] In this connection, it should be noted that the definitions employed here sometimes differ from those of Smith by factors of c.

Diagram Renormalization, Variational Principles, and the Infinite-Dimensional Ising Model*

CLAUDE BLOCH†
Service de Physique Theorique, C. E. N. de Saclay, France
and
Palmer Physical Laboratory, Princeton University, Princeton, New Jersey
AND
JAMES S. LANGER
Physics Department, Carnegie Institute of Technology, Pittsburgh, Pennsylvania
(Received 27 August 1964)

The Ising model has been employed in a study of the variational principles which are associated with renormalized diagram expansions in statistical mechanics. For this model the variational functional is expressed in terms of the renormalized semi-invariants, which play the role of the one-particle density or the one-particle Green's function in quantum statistics. We review the derivation of this functional and discuss some of its properties. In order to examine the mathematical content of the variational principle in more detail, we specialize to the exactly soluble infinite-dimensional (infinite-range) model. We find that, by virtue of the Dyson relation, the variational theorem is not valid over all possible values of the renormalized semi-invariants but applies only within a restricted domain. Within this domain the variational functional is multiple valued. The renormalized expansion is asymptotically convergent to the branch of this functional which describes a single phase with uniform magnetization.

1. INTRODUCTION

THE most successful approach to a description of phase transitions, so far, has been based on exactly soluble models. This method yields rigorous results, but usually cannot be extended beyond the limitations of each model. In particular, it is very difficult to draw any conclusions about real systems.

On the other hand, the use of approximation methods in the case of phase transitions is made very difficult by the nonanalytic behavior characteristic of these phenomena. Of course, one can hope to find practical methods of approximation, such as rapidly convergent expansions, only for the evaluation of functions or functionals having a simple analytic behavior. The fundamental problem is then to find mathematical mechanisms through which the nonanalyticity associated with the phase transition may be generated from essentially analytic quantities.

The standard low-density expansions of statistical mechanics obviously diverge in the case of phase transitions. From physical arguments, on the other hand, one has a feeling that the methods based on the equations relating the distribution functions, or the Green's functions in quantum statistical mechanics, may be applicable to dense systems and may serve as a basis for a discussion of phase transitions.

A convenient way of generating such systems of equations is provided by the variational principles which follow from the renormalized diagram expansions.[1] In classical statistical mechanics, this method is based on the expression of $\ln Z$ as a functional of the density $n(r)$, or of the density and the correlation function $\mathfrak{C}(r, r')$, which goes through an absolute maximum at equilibrium.[2] In quantum statistical mechanics, the corresponding functional may be only stationary at equilibrium. In all cases, the stationarity conditions provide a system of equations for the determination of $n(r)$, or of $n(r)$ and $\mathfrak{C}(r, r')$, and therefore also of $\ln Z$.

The functional representing $\ln Z$ in these methods is given by an infinite expansion which necessarily must be truncated in any practical calculation. The purpose of the present work is to examine the mathematical nature of the approximation methods based on these expansions by applying them to a simple, exactly soluble model. We have used therefore the infinite-dimensional Ising model discussed by Tem-

* Supported in part by the National Science Foundation.
† Almost all of this work was performed at the Physics Department of Carnegie Institute of Technology during the fall semester, 1963.

[1] For a review of these methods and detailed references, see C. Bloch in *Studies in Statistical Mechanics*, edited by J. de Boer and G. E. Uhlenbeck (North-Holland Publishing Company, Amsterdam, 1964), Vol. III.
[2] T. Morita and K. Hiroike, Progr. Theoret. Phys. (Kyoto) 25, 537 (1961).

perly and others.[3] This model, in spite of the extreme simplicity of its exact solution, exhibits a kind of phase transition characterized by the appearance of a spontaneous magnetization below a certain critical temperature (Sec. 5).

Perturbation expansions for Ising models have been developed by Brout,[4] Englert,[5] and by Horwitz and Callen.[6] They involve the so-called semi-invariants $M_{n,i}$ ($n = 1, 2, \cdots$; i denotes the various spins of the system), which replace the density $n(r)$ of classical statistical mechanics, or the one-particle propagator of quantum statistical mechanics. The variational expression for ln Z appears then as a functional of the renormalized semi-invariants $M_{n,i}$ (Sec. 2). For studying the general properties of these functionals, it is convenient to introduce fictitious external fields whose variations induce variations of the $M_{n,i}$ (Sec. 3). The field corresponding to the first semi-invariant $M_{1,i}$ is similar to a magnetic field. The consequence of the physical nature of this field is that ln Z, considered as a functional of the $M_{1,i}$ only, is a concave functional going through an absolute maximum at equilibrium. Also the magnetization curve giving $\sum_i M_{1i}$ in terms of the external field is necessarily a monotonically increasing function. These calculations do not extend to the general functional depending on all semi-invariants $M_{n,i}$ which correspond to the introduction of unphysical fields resulting in nonpositive-definite probability distributions. It will therefore be particularly interesting to discuss the various approximations to the *reduced functional* of the $M_{1,i}$ obtained by substituting into the general functional the stationary values of the $M_{2,i}$, $M_{3,i}$, $\cdots$. The reduced functional should be concave.

In the infinite-dimensional model, the contributions of all individual renormalized diagrams vanish in the limit $N \to \infty$, except one of them (Sec. 6). This one gives a functional of M_1 which is identical with the molecular field approximation. It is not a concave functional, however, when the phase transition takes place, although it is exact outside the two-phase region.

As a first investigation into the nature of this difficulty, we have summed a particular infinite family of diagrams—the ring diagrams—for finite N, and carried out the limit $N \to \infty$ only after solving the variational problem with respect to the semi-invariant M_2. (Higher semi-invariants do not appear in this approximation.) This procedure indeed gives a different reduced functional of M_1 in part of the phase transition region, in spite of the fact that each ring diagram individually gives a vanishing contribution as $N \to \infty$. The resulting functional is still not exact and nonconcave in the two-phase region.

Summation of still larger families of diagrams (Secs. 7 and 8) does not produce further improvements of the reduced functional of M_1 in the same way as the ring diagrams. An unexpected aspect of the problem, however, becomes important. As a consequence of the Dyson relation, the renormalized semi-invariants M_1, M_2, $\cdots$ turn out to satisfy inequalities defining a restricted domain of variation. This restriction excludes precisely the region where the molecular field approximation is wrong. An even more surprising development is that the functional to which the variational theorem applies turns out to be multiple-valued in the allowed domain. The perturbation expansion determines directly only one branch of this functional.

The same analysis yields some information concerning the convergence of the expansions. This convergence, in the infinite-dimensional Ising model, is at best an asymptotic convergence for $N \to \infty$. This type of convergence disappears below the critical temperature for the unrenormalized expansion, whereas the renormalized expansion remains always asymptotically convergent, and converges to one branch of the correct functional provided the values of the M_1, M_2, $\cdots$ stay within the permitted domain of variation mentioned above.

The conclusion of this paper is to suggest a prescription for the calculation of the complete reduced functional: use the renormalized expansion in the region defined by the constraint (corresponding to a unique uniform phase), and deduce the functional in the remainder of the domain of M_1 (the two-phase region) by taking the concave envelope of the functional. How far this conclusion can be extended beyond the particular model treated here remains, of course, to be seen.

2. DIAGRAM EXPANSION FOR THE ISING MODEL

2.1 Elementary Rules

The partition function for the Ising model in the presence of an external field X_i/β is given by

$$Z = \sum_{\{\mu\}} \exp\left(\frac{\beta}{2} \sum_{ij} v_{ij}\mu_i\mu_j + \sum_i X_i\mu_i\right), \qquad (2.1)$$

[3] H. N. V. Temperley, Proc. Phys. Soc. (London) **A67**, 233 (1954). Very complete discussions of this model have been given by S. Katsura, Progr. Theoret. Phys. (Kyoto) **13**, 571 (1955) and N. Saito, J. Chem. Phys. **35**, 232 (1961).

[4] R. Brout, Phys. Rev. **115**, 824 (1959).

[5] F. Englert, Phys. Rev. **129**, 567 (1963).

[6] G. Horwitz and H. B. Callen, Phys. Rev. **124**, 1757 (1961).

where $\{\mu\}$ ($\mu_i = \pm 1$, $i = 1, 2, \cdots, N$) denotes the set of N spin variables of the system.

We consider first the expansion of Z in powers of the interaction v_{ij}. The result may conveniently be described by means of diagrams consisting of an arbitrary number of points labeled i, j, k, $\cdots$ and bonds joining pairs of points. The factors associated with the elements of a diagram are (see Appendix A):

βv_{ij} with each bond between points i and j;

$M^0_{n,i}$ with each point i at which n bonds arrive, where the quantities

$$M^0_{n,i} = (d/dX_i)^n \ln 2 \cosh X_i \tag{2.2}$$

are the so-called semi-invariants.

The contribution of a diagram is obtained by carrying out independent summations over the indices i, j, k, $\cdots$ from 1 to N and multiplying the result by the weight $1/S$ where S is the symmetry number of the diagram.

Then $\ln Z$ is the sum of $\ln Z_0$ and the contributions of all connected diagrams, where

$$\ln Z_0 = \ln \Big[\sum_{\{\mu\}} \exp \Big(\sum_i X_i \mu_i\Big)\Big] = \sum_i \ln 2 \cosh X_i \tag{2.3}$$

is the logarithm of the partition function in the absence of interaction.

A point i in a connected diagram is called an *articulation point* if the diagram splits into two or more disconnected parts when this point is removed. The points i, j, k on the diagram of Fig. 1, for instance, are articulation points. If we call a *star* any subdiagram which cannot be split into two disconnected subdiagrams by the removal of a single point, we see that the various articulation points of a diagram define the structure of a *tree* of *stars*. On the diagram of Fig. 1, the triangle i, j, k is a star; two other stars are attached to each of the points i and j; one additional star is attached to k.

2.2 Renormalization

We now consider the partial resummation obtained by attaching all possible subdiagrams to all the points of a given star considered as a skeleton. Let us call $G_{p,i}$ the sum of the contributions of all proper subdiagrams which are connected to i by p bonds. The term "proper" implies that i is not an articulation point of $G_{p,i}$. Consider a point i of the skeleton star at which n bonds arrive. We can attach to this point n_1 proper diagrams of $G_{1,i}$, n_2 of $G_{2,i}$, etc. $\cdots$. The total number of bonds arriving at i is then

$$n + n_1 + 2n_2 + 3n_3 + \cdots .$$

The total contribution of the skeleton star after this resummation has been performed is therefore obtained by associating with every point i at which n bonds of the skeleton arrive the *renormalized semi-invariant*

$$M_{n,i} = \sum_{n_1, n_2, \ldots} \frac{1}{n_1!\, n_2! \cdots} (G_{1,i})^{n_1} \times (G_{2,i})^{n_2} \cdots M^0_{n+n_1+2n_2+\cdots,i}, \tag{2.4}$$

where $n_1!\, n_2! \cdots$ is the supplementary symmetry factor introduced by the presence of identical factors $G_{1,i}$, $G_{2,i}$, $\cdots$.

The relation (2.4) is the equivalent of the Dyson relation. It expresses the renormalized semi-invariants (which correspond to the exact propagators of quantum statistical mechanics) in terms of the unperturbed ones and the $G_{n,i}$ (which correspond to the self-energy operators[7]).

The relation (2.4) takes a simpler form in terms of the *generating functions* for the semi-invariants:

$$\begin{aligned} M_i(z) &= \sum_{n=0}^{\infty} \frac{z^n}{n!} M_{n,i}, \\ M^0_i(z) &= \sum_{n=0}^{\infty} \frac{z^n}{n!} M^0_{n,i} = \ln 2 \cosh (X_i + z), \end{aligned} \tag{2.5}$$

where we have used (2.2). Then (2.4) may be written

$$M_i(z) = \Big[\exp \sum_{p=1}^{\infty} G_{p,i} \Big(\frac{\partial}{\partial z}\Big)^p\Big] M^0_i(z). \tag{2.6}$$

The bracket here is a symbolical representation of an integral operator; and this relation is actually equivalent to

$$M_i(z) = \int_{-\infty}^{+\infty} K_i(z - z') M^0_i(z')\, dz', \tag{2.7}$$

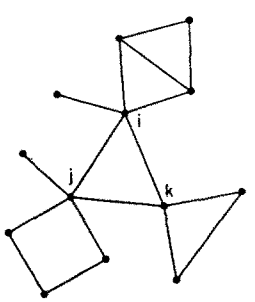

FIG. 1. A diagram containing three articulation points.

[7] This analogy suggests an interchange of the notations M_n and G_n. We have, however, kept the notation commonly used in this subject.

where

$$K_i(z - z') = \frac{1}{2\pi i}\int_{-i\infty}^{+i\infty} dx \times \exp\left[x(z - z') + \sum_{p=1}^{\infty} x^p G_{p,i}\right]. \quad (2.8)$$

The relation (2.6) or (2.7) takes its simplest form in terms of the Fourier transforms of $M_i(z)$ and $M_i^0(z)$:

$$\hat{M}_i(x) = \hat{M}_i^0(x) \exp\left[\sum_{p=1}^{\infty} x^p G_{p,i}\right], \quad (2.9)$$

where

$$\hat{M}_i(x) = \int_{-\infty}^{+\infty} e^{-xz} M_i^0(z)\, dz, \qquad \hat{M}_i^0(x) = \int_{-\infty}^{+\infty} e^{-xz} M_i^0(z)\, dz. \quad (2.10)$$

In these relations, x should be taken purely imaginary.

2.3 Variational Principle

Let us return now to the evaluation of the contribution of all closed connected diagrams. It is clear that by starting from all possible skeleton stars, and by attaching at all their points all possible subdiagrams in all possible ways, we shall reconstruct all closed connected diagrams. Each diagram, however, in this process appears a number of times equal to its number of stars. It is easy to correct this overcounting of diagrams by making use of the topological relation valid for any tree of stars,

$$N_s - N_i + N_a = 1, \quad (2.11)$$

where N_s is the number of stars of the diagram, N_a the number of articulation points, and N_i the sum of the indices of all articulation points (the index of an articulation point being defined here as the number of stars attached to it).

Let us then call $\mathfrak{D}\{M_{n,i}\}$ the total weighted contribution of all diagrams consisting of a single star, computed with the renormalized semi-invariants $M_{n,i}$. We have seen that each tree of stars in the complete perturbation expansion of $\mathfrak{D}\{M_{n,i}\}$ appears N_s times. On the other hand, if we consider the corresponding expansion for the expression

$$\sum_{n,i} G_{n,i} M_{n,i}, \quad (2.12)$$

it is easy to see that each tree of stars diagram appears N_i times. Finally, in the expansion of

$$\sum_{n_1, n_2, \ldots} \frac{1}{n_1!\, n_2! \cdots} (G_{1,i})^{n_1} (G_{2,i})^{n_2} \cdots M^0_{n_1+2n_2+\cdots,i} \equiv M_{0,i}, \quad (2.13)$$

each tree of stars appears N_a times. The notation $M_{0,i}$ for the expression (2.13) has been introduced in view of its identity with (2.4) when $n = 0$.

It follows now from (2.11) that the functional:

$$\Phi\{M_{n,i}\} = \mathfrak{D}\{M_{n,i}\} - \sum_{n,i} G_{n,i} M_{n,i} + \sum_i M_{0,i} \quad (2.14)$$

becomes equal to $\ln Z$ when the exact equilibrium values are substituted for the $M_{n,i}$. Moreover, $\Phi\{M_{n,i}\}$ is stationary around the equilibrium values of the $M_{n,i}$:

$$\partial\Phi/\partial M_{n,i} = 0. \quad (2.15)$$

This follows from (2.4), (2.13), (2.14), and the relation

$$G_{n,i} = \partial\mathfrak{D}/\partial M_{n,i}, \quad (2.16)$$

which follows immediately from a comparison of the diagram expansions of both sides.

The stationarity condition (2.15) provides a set of equations for the determination of the equilibrium values of the $M_{n,i}$. Substitution of these values into (2.14) yields the value of $\ln Z$ corresponding to each approximate expression used for the functional $\mathfrak{D}\{M_{n,i}\}$.

3. CONVEXITY CONDITIONS

In some cases, the functional Φ is not only stationary, but goes through an absolute maximum at equilibrium. This holds quite generally in classical statistical mechanics, and also in some formulations of quantum statistical mechanics. In the Ising model, although all dynamical variables commute, we see that the situation differs from the ordinary case of classical statistical mechanics as a consequence of the discrete nature of these variables.

3.1 Fictitious Fields

A convenient way to define the functionals Φ is to introduce external fields of a nature appropriate to each particular choice of the functions [$n(r)$ or $n(r)$ and $\mathcal{C}(r, r')$, etc. . . .] occurring as arguments in Φ. In the Ising model, the simplest field of this kind is the field X_i introduced in (2.1). This field, which is of the nature of a magnetic field, may be considered as a physical field, or rather, since we are essentially interested in the thermodynamics of

the system in the absence of an external field, as a fictitious field introduced only in order to force the semi-invariants to vary, and thus to allow a natural definition of Φ as a functional of these semi-invariants. The equilibrium situation corresponds therefore to

$$X_i = 0. \tag{3.1}$$

We adopt this point of view throughout the rest of the paper.

In accord with this point of view it is convenient to treat the external field X_i as part of the perturbation. Thus, instead of (2.2), we understand the unrenormalized semi-invariants to be defined by

$$M^0_{n,i} = (\partial/\partial X)^n \ln 2 \cosh X \,|_{X=0}. \tag{3.2}$$

As a perturbation, X_i appears only implicitly in $G_{1,i}$. That is,

$$G_{1,i} = X_i + G'_{1,i}, \tag{3.3}$$

where $G'_{1,i}$ is defined by the previous sum of proper diagrams. In other words, we think of X_i as the simplest possible proper part contributing to $G_{1,i}$. The appropriate variational functional (2.14) now should be written

$$\Phi_X = \mathfrak{D}_X\{M_{n,i}\} - \sum_{n,i} G_{n,i}M_{n,i} + \sum_i M_{0,i}. \tag{3.4}$$

The only place where Φ_X contains X_i explicitly is in $\mathfrak{D}_X$ which contains a term

$$\sum_i X_i M_{1,i}, \tag{3.5}$$

clearly the simplest possible skeleton diagram. Note that, if we subtract (3.5) from (3.4) we obtain a functional Φ which contains no explicit X_i's and is identical to our original functional (2.14) for $X_i = 0$. Thus we are led to consider

$$\begin{aligned}\Phi\{M_{n,i}\} &= \Phi_X\{M_{n,i}\} - \sum_i X_i M_i \\ &= \mathfrak{D}\{M_{n,i}\} - \sum_{n,i} G_{n,i}M_{n,i} + \sum_i M_{0,i}.\end{aligned} \tag{3.6}$$

3.2 The Reduced Functional

The reduced functional $\tilde{\Phi}(M_{1,i})$ is obtained from (3.6) by substituting for the $M_{n,i}$, $n = 2, 3, \cdots$, their values which make Φ stationary. That is,

$$\tilde{\Phi}(M_{1,i}) = \Phi\{M_{1,i}, M_{2,i}(M_{1,i}), \cdots\}, \tag{3.7}$$

where the $M_{n,i}(M_{1,i})$ are obtained from (2.15) for $n = 2, 3, \cdots$. This functional is simply the Legendre transform of $\ln Z\{X\}$. To see this, note that the renormalized semi-invariants $M_{1,i}$ are given by

$$M_{1,i} = M_i = \partial \ln Z/\partial X_i, \tag{3.8}$$

and

$$\tilde{\Phi} = \ln Z - \sum_i X_i M_{1,i}, \tag{3.9}$$

according to (3.6) and (3.7). It is assumed for the moment that we can invert the relations (3.8) and express the X_i in terms of the $M_{1,i}$. Then

$$\begin{aligned}d\tilde{\Phi} &= \sum_i \frac{\partial \ln Z}{\partial X_i} dX_i - \sum_i M_{1,i}\, dX_i - \sum_i X_i\, dM_{1,i} \\ &= -\sum_i X_i\, dM_{1,i}.\end{aligned} \tag{3.10}$$

This implies that

$$\partial\tilde{\Phi}(M_1)/\partial M_{1,i} = -X_i, \tag{3.11}$$

which shows again that $\tilde{\Phi}(M_1)$ is stationary at equilibrium when $X_i = 0$.

In order to discuss the possibility of inverting the relations (3.8), we must now examine the corresponding Jacobian, which is the determinant of the matrix,

$$\frac{\partial M_{1,i}}{\partial X_j} = \frac{\partial^2 \ln Z}{\partial X_i\, \partial X_j} = \langle(\mu_i - \langle\mu_i\rangle)(\mu_j - \langle\mu_j\rangle)\rangle. \tag{3.12}$$

The right-hand side of this relation, which follows readily from (2.1), is a fluctuation matrix. It is therefore positive-definite; and this shows that our Jacobian can never vanish. The fact that (3.12) is a positive-definite matrix implies that $\ln Z$ is a concave functional of the X_i. It follows that the relations (3.8) define one and only one set of values of the X_i for each set of values of the $M_{1,i}$ belonging to the domain of variation of these quantities. Clearly, from (3.8), the $M_{1,i}$ satisfy the conditions

$$-1 \leq M_{1,i} \leq 1. \tag{3.13}$$

A point on the boundary of the domain of variation of the $M_{1,i}$ is obtained by taking some of the X_i equal to $+\infty$ or $-\infty$. The corresponding $M_{1,i}$ are then clearly equal $+1$ or -1. This shows that the inequalities (3.13) define the actual domain of variation of the $M_{1,i}$.

Returning now to the functional $\tilde{\Phi}(M_1)$ defined by (3.7), we see from (3.11) that the matrix formed by its second derivatives may be written:

$$\partial^2\tilde{\Phi}(M_1)/\partial M_{1,i}\, \partial M_{1,j} = -\partial X_j/\partial M_{1,i}. \tag{3.14}$$

In this form it appears as minus the reciprocal of the positive-definite fluctuation matrix (3.12). It is therefore negative-definite, which shows that $\tilde{\Phi}$ is a concave functional of the $M_{1,i}$ in the domain

defined by (3.13). It follows that $\tilde{\Phi}(M_1)$ goes through an absolute maximum at equilibrium.

As long as N is finite, all functionals entering the problem are analytic. For $N \to \infty$, however, the limiting functional $\tilde{\Phi}(M_1)$ may become linear or constant in some range, in exactly the same way as a concave curve may have flat sections. Such a behavior is expected when a phase transition takes place. The functional $\tilde{\Phi}(M_1)$ then must have two equal maxima at values of the $M_{1,i}$ corresponding to two different phases. This is consistent with the concavity condition only if $\tilde{\Phi}(M_1)$ is constant along the straight segment joining the two points $M'_{1,i}$ and $M''_{n,i}$ corresponding to the two pure phases in equilibrium.

3.3 The Complete Functional

So far, we have not yet discussed the functional of Sec. 2, which depends not only on the $M_{1,i}$ but on all the $M_{n,i}$. Clearly, a field such as X_i does not possess sufficiently many degrees of freedom to force independent variations of all the $M_{n,i}$. We must therefore introduce modifications of the system of a more general nature than those induced by the field X_i. Let us assume that the discrete values ± 1 taken by the μ_i are replaced by a continuum with distribution functions $p_i(\mu_i)$. The corresponding partition function is then a functional of the $p_i(\mu_i)$:

$$Z\{p\} = \int p_1(\mu_1)p_2(\mu_2) \cdots d\mu_1\, d\mu_2 \cdots \times \exp\left(\frac{\beta}{2}\sum_{ij} v_{ij}\mu_i\mu_j\right). \tag{3.15}$$

Eq. (3.15) defines a *model system* which we shall find very useful in Sec. 7 of this paper.

The partition function in the absence of interaction is now given by

$$\ln Z_0 = \sum_i \ln\left[\int p_i(\mu)\, d\mu\right]. \tag{3.16}$$

The diagram expansion of $\ln (Z/Z_0)$, on the other hand, is given by the rules of Sec. 2, where the unrenormalized semi-invariants (2.2) are replaced by the $M^0_{n,i}$ defined by the relations

$$\sum_{n=0}^{\infty} \frac{z^n}{n!} M^0_{n,i} = M^0_i(z), \qquad \exp\,[M^0_i(z)] = \int e^{\mu z} p_i(\mu)\, d\mu. \tag{3.17}$$

Inversion of these relations yields, in principle, the distribution functions $p_i(\mu)$ corresponding to given $M^0_{n,i}$.

Let us now define a generalized external field $X_{n,i}$ by writing in analogy with (2.6)

$$M^0_i(z) = \exp\left[\sum_{p=1}^{\infty} X_{p,i}\left(\frac{\partial}{\partial z}\right)^p\right] \ln 2\cosh z. \tag{3.18}$$

Note that this relation reduces exactly to the second relation (2.5) when all $X_{p,i}$ for $p = 2, 3, \cdots$ vanish, and $X_{1,i} = X_i$. The relation (3.18) may be transformed into a form analogous to (2.7) or to (2.9). Thus we have

$$\hat{M}^0_i(x) = \hat{M}^{00}_i(x) \exp\left[\sum_{p=1}^{\infty} x^p X_{p,i}\right], \tag{3.19}$$

where

$$\hat{M}^0_i(x) = \int_{-\infty}^{+\infty} e^{-xz} M_i(z)\, dz, \qquad \hat{M}^{00}_i(x) = \int_{-\infty}^{+\infty} e^{-xz} \ln (2\cosh z)\, dz. \tag{3.20}$$

The relation (3.19) may readily be used to express the $X_{p,i}$ in terms of the $M^0_{n,i}$. Thus, at least formally, the $X_{p,i}$ appear as a system of variables equivalent to the $M_{n,i}$ or the $p_i(\mu)$ for the characterization of the external disturbance of the system under consideration. To be rigorous, however, one should still determine the permitted domain of variation of each system of variables.

The stationarity properties discussed in the first part of this section can be extended immediately to the more general functionals. By combining (2.6) and (3.18), we may write for the renormalized semi-invariants

$$M_i(z) = \exp\left[\sum_{p=1}^{\infty} (X_{p,i} + G'_{p,i})\left(\frac{\partial}{\partial z}\right)^p\right] \ln 2\cosh z, \tag{3.21}$$

where the $G'_{p,i}$ are the proper parts defined by the usual diagrams. This shows the analogy of $X_{p,i}$ with the $G'_{p,i}$ and generalizes Eq. (3.3). It follows that the expansion of $\ln Z\{X\}$ in powers of the interaction v_{ij} and of the $X_{p,i}$ is given by the rules of Sec. 2, with additional elements corresponding to the $X_{p,i}$. Each such element $X_{p,i}$ should be considered, in the evaluation of the semi-invariants associated with each point i, as equivalent to p bonds. It follows immediately that

$$\partial \ln Z\{X\}/\partial X_{p,i} = M_{p,i}, \tag{3.22}$$

which generalizes (3.8) and enables us to write the complete Φ as a Legendre transform:

$$\Phi = \ln Z - \sum_{n,i} X_{n,i} M_{n,i}. \tag{3.23}$$

Then from (3.22) and (3.23) we have

$$d\Phi = \sum_{n,i} \left(\frac{\partial \ln Z}{X_{n,i}} dX_{n,i} - X_{n,i}\, dM_{n,i} - M_{n,i}\, dX_{n,i} \right)$$
$$= -\sum_{n,i} X_{n,i}\, dM_{n,i}. \tag{3.24}$$

Thus, if we invert the relations (3.22) and express the $X_{p,i}$ as functions of the $M_{n,i}$, substitution of the resulting expressions into Φ will yield a functional $\Phi\{M\}$ such that

$$\partial\Phi\{M\}/\partial M_{n,i} = -X_{n,i}. \tag{3.25}$$

Therefore $\Phi\{M\}$ is stationary at equilibrium when the external field $X_{n,i}$ is set equal to zero.

The discussion of concavity, however, does not readily generalize because, for arbitrary $X_{n,i}$, the probability distributions $p_i(\mu_i)$ are not necessarily everywhere positive. The concavity condition could therefore hold only in a restricted domain of $X_{n,i}$. This makes the corresponding domain for the $M_{n,i}$ much more complicated than the domain (3.13) for the $M_{1,i}$.

4. SOME SIMPLE APPROXIMATIONS

In order to clarify the formalism developed in the last two sections, we examine now two simple approximations to the functional $\mathfrak{D}$. The first approximation yields the molecular field theory; the second recovers the results of the ring diagram expansions of Horwitz and Callen.[6]

4.1 Molecular-Field Approximation

We take for $\mathfrak{D}$ just the diagram consisting of two points joined by one bond (Fig. 2). This is actually the only diagram with no articulation point containing M_1. This gives

$$\mathfrak{D} = \tfrac{1}{2}KNM_1^2, \tag{4.1}$$

where

$$K = \beta \sum_j v_{ij} \tag{4.2}$$

is independent of i as a consequence of translational invariance, which is assumed for simplicity. Similarly, we have assumed uniformity in the system, i.e., that $M_{1,i}$ is independent of i and denoted it simply by M_1.

FIG. 2. Skeleton diagram which determines $\mathfrak{D}$ for the molecular field approximation.

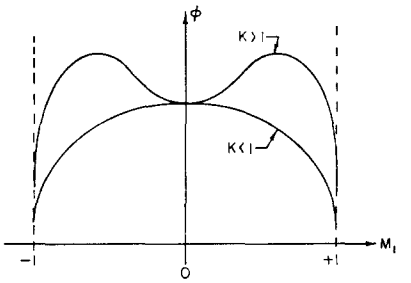

FIG. 3. Variational function $\phi(M_1)$ for molecular field approximation.

The stationarity relations (2.16) show immediately that

$$G_2 = G_3 = \cdots = 0. \tag{4.3}$$

The function Φ is given by

$$(1/N)\Phi(M_1) \equiv \phi(M_1)$$
$$= \tfrac{1}{2}KM_1^2 - G_1M_1 + \ln 2 \cosh G_1, \tag{4.4}$$

(the notation $\phi = \Phi/N$ will be convenient in what follows); and the Dyson relation reduces to

$$M_1 = \tanh G_1. \tag{4.5}$$

The stationarity relation (2.15) for $n = 1$, which follows from the stationarity condition of (4.4) taking into account (4.5), reads:

$$G_1 = KM_1. \tag{4.6}$$

Elimination of G_1 between (4.5) and (4.6) yields the familiar molecular field equation

$$M_1 = \tanh KM_1. \tag{4.7}$$

It is also easy to write $\phi(M_1)$ explicitly by eliminating G_1 between (4.4) and (4.5):

$$G_1 = \tfrac{1}{2} \ln [(1 + M_1)/(1 - M_1)], \tag{4.8}$$

and therefore

$$\phi(M_1) = \frac{1}{2}KM_1^2 - \frac{M_1}{2} \ln \left(\frac{1+M_1}{1-M_1} \right) - \frac{1}{2} \ln \left(\frac{1-M_1^2}{4} \right). \tag{4.9}$$

This function is plotted in Fig. 3. It is clearly not concave for $K > 1$, i.e., for low-enough temperatures.

4.2 Ring Diagrams

The next simple class of diagrams includes the ring diagrams shown in Fig. 4. Their contribution to Φ is given by

$$\sum_q \sum_{n=2}^{\infty} \frac{1}{2n} (\beta v_q M_2)^n = -\frac{1}{2} \sum_q \ln (1 - \beta v_q M_2), \tag{4.10}$$

FIG. 4. Some ring diagrams.

where we have used the Fourier representation of v_{ij}:

$$v_{ij} = \frac{1}{N}\sum_{\mathbf{q}} v_{\mathbf{q}}\, e^{i\mathbf{q}\cdot(\mathbf{r}_i-\mathbf{r}_j)}, \qquad v_{ii} = 0. \tag{4.11}$$

Thus we obtain

$$\mathfrak{D} = \tfrac{1}{2}KNM_1^2 - \tfrac{1}{2}\sum_{\mathbf{q}} \ln(1 - \beta v_{\mathbf{q}} M_2). \tag{4.12}$$

From the relations (2.16) we see that

$$G_3 = G_4 = \cdots = 0, \tag{4.13}$$

and we have

$$\phi(M_1, M_2) = \frac{1}{2}KM_1^2 - \left(\frac{1}{2N}\right)\sum_{\mathbf{q}} \ln(1 - \beta v_{\mathbf{q}} M_2) - M_1G_1 - M_2G_2 + M_0. \tag{4.14}$$

The kernel (2.8) occurring in the Dyson relation (2.7) is given by

$$K(z - z') = \frac{1}{2\pi i}\int_{-i\infty}^{i\infty} dx\, \exp[x(z-z') + xG_1 + x^2G_2] = (4\pi G_2)^{-\frac{1}{2}} \exp[-(z - z' + G_1)^2/4G_2], \quad (G_2 > 0). \tag{4.15}$$

It follows immediately that the Dyson relations reduce in this case to

$$M_1 = \int_{-\infty}^{+\infty} \frac{dz'}{(4\pi G_2)^{\frac{1}{2}}} \exp\left[-\frac{(G_1 - z')^2}{4G_2}\right] \tanh z', \tag{4.16}$$

$$M_2 = \int_{-\infty}^{+\infty} \frac{dz'}{(4\pi G_2)^{\frac{1}{2}}} \exp\left[-\frac{(G_1 - z')^2}{4G_2}\right] \operatorname{sech}^2 z',$$

(the expressions for M_3, M_4, $\cdots$ are not useful). Also,

$$M_0 = \int_{-\infty}^{+\infty} \frac{dz'}{(4\pi G_2)^{\frac{1}{2}}} \exp\left[-\frac{(G_1 - z')^2}{4G_2}\right] \ln 2\cosh z'. \tag{4.17}$$

The stationarity of $\phi(M_1, M_2)$ is now expressed by the relations

$$G_1 = (1/N)\partial\mathfrak{D}/\partial M_1 = KM_1, \qquad G_2 = \frac{1}{N}\frac{\partial\mathfrak{D}}{\partial M_2} = \frac{1}{2N}\sum_{\mathbf{q}} \frac{\beta v_{\mathbf{q}}}{1 - \beta v_{\mathbf{q}} M_2}, \tag{4.18}$$

which are the equations discussed by Horwitz and Callen.[6]

5. THE INFINITE-DIMENSIONAL MODEL

For a more detailed study of the various expansions we now restrict ourselves to a particularly simple case of the Ising model: the infinite-dimensional case studied for similar reasons by Temperley.[3] In this model, every spin interacts equally with every other spin, the interaction strength being proportional to $1/N$. The exact solution to this model in the presence of an external field is extremely simple.

The partition function for the infinite-dimensional model is given by

$$Z = \sum_{\{\mu\}} \exp\left[\frac{K}{2N}\sum_{ij}\mu_i\mu_j + X\sum_i \mu_i\right], \tag{5.1}$$

where $K = \beta J$ and $v = J/N$ is the interaction energy. Here we have introduced a uniform external field X (independent of i) for simplicity.

If we introduce the magnetization

$$M = \frac{1}{N}\sum_i \mu_i, \tag{5.2}$$

we can write (5.1) in the form[8]

$$Z = \sum_M g(M) \exp\left[N\left(\frac{KM^2}{2} + XM\right)\right], \tag{5.3}$$

where the summation is extended to the following values of M:

$$M = -1, -1 + \frac{2}{N}, -1 + \frac{4}{N}, \cdots, 1 - \frac{2}{N}, 1;$$

and $g(M)$ is the number of configurations with magnetization M:

$$g(M) = N!/[\tfrac{1}{2}N(1 + M)]!\,[\tfrac{1}{2}N(1 - M)]!. \tag{5.4}$$

For large N, we may replace (5.4) by its Stirling's approximation and replace the discrete summation over M by an integration. This gives

$$Z = \frac{1}{\pi^{\frac{1}{2}}}\left(\frac{2}{N}\right)^{\frac{1}{2}}\int_{-1}^{+1} \frac{dM}{(1 - M^2)^{\frac{1}{2}}} e^{N[f(M) + XM]}, \tag{5.5}$$

where

$$f(M) = \frac{KM^2}{2} - \left(\frac{1 + M}{2}\right)\ln\left(\frac{1 + M}{2}\right) - \left(\frac{1 - M}{2}\right)\ln\left(\frac{1 - M}{2}\right) = \frac{KM^2}{2} - \frac{1}{2}\ln\left(\frac{1 - M^2}{4}\right) - \frac{M}{2}\ln\left(\frac{1 + M}{1 - M}\right). \tag{5.6}$$

This function is identical with the function $\phi(M_1)$ of the molecular-field approximation defined by (4.9) and is plotted on Fig. 3. For $K < 1$, $f(M)$ has a maximum at $M = 0$. For $K > 1$, $M = 0$ becomes a local minimum and two symmetric maxima appear.

When N is very large, the dominant contributions

[8] In this expression we have not subtracted the terms $i = j$. This makes no difference in the limit $N \to \infty$.

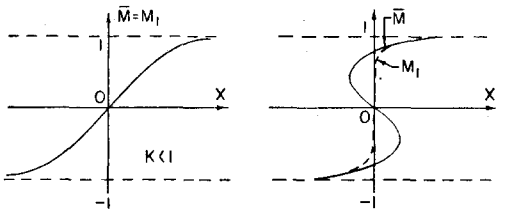

FIG. 5. Magnetization curves for the infinite-dimensional model. For $K > 1$, the narrow line represents the spurious solutions of Eq. (5.7). The dashed line is the exact $M_1(X)$ for finite N.

to (5.5) come from the neighborhood of $\bar{M}$ for which $f(M) + XM$ is maximum. Thus $\bar{M}$ is given by

$$f'(\bar{M}) + X = 0. \tag{5.7}$$

For $K > 1$, this equation may have several solutions in $\bar{M}$. One of them only, however, corresponds to the maximum of $f(M) + XM$. The variation of $\bar{M}$ as a function of X is shown on Fig. 5. In the case $K > 1$, the spurious solutions of (5.7) are represented by the narrow part of the curve. When $X = 0$ and $K > 1$, an exceptional situation arises as the function $f(M) + XM$ has two equal symmetric maxima.

When $f(M) + XM$ has a unique maximum at $M = \bar{M}$, we have

$$\lim_{N\to\infty} \frac{1}{N} \ln Z = f(\bar{M}) + X\bar{M}. \tag{5.8}$$

The corresponding curves are shown on Fig. 6 in the two cases $K < 1$ and $K > 1$. When $K > 1$, the derivative of $f(\bar{M}) + X\bar{M}$ with respect to X is discontinuous. This is clearly related to the discontinuity of $\bar{M}(X)$ at $X = 0$. From the original expression (5.5) of Z, however, it appears that for any *finite* N, Z is an analytic function of X. The nonanalytic curve of Fig. 6 is therefore the limit of an analytic curve whose curvature at $X = 0$ increases indefinitely as $N \to \infty$ (dotted curve).

Let us consider now the variable M_1 defined by (3.8) and the reduced functional $\tilde{\phi}(M_1)$. When $f(M) + XM$ has a unique maximum, we see immediately from (5.8) that

$$M_1 = (1/N)\, \partial \ln Z/\partial X = \bar{M} \tag{5.9}$$

and

$$\tilde{\phi}(M_1) = (1/N) \ln Z - XM_1 = f(M_1). \tag{5.10}$$

Thus for $K < 1$, the curve representing $\tilde{\phi}(M_1)$ is identical to Fig. 3. For $K > 1$, the same conclusion applies except when M_1 lies between the two maxima. The part of the curve corresponding to $X \sim 0$ requires a little more attention. The simplest way to see what happens then is to consider a large but finite value of N. The curve of Fig. 6 is then rounded off at $X = 0$. It follows that M_1 as a function of X is continuous, but increases very rapidly in a small interval around $X = 0$ (dashed curve on Fig. 5). On the other hand, (5.9) and (5.10) show immediately that the variation of $\tilde{\phi}(M_1)$ throughout this small interval is very small.

Another way to reach the same conclusion is to examine directly the integral (5.5) for N large but finite for $K > 1$ and $X \sim 0$. For $X < 0$ the integrand has a unique maximum, but as X approaches 0, the local maximum occurring for a nearly opposite value of M tends to give contributions of the same order of magnitude. As X becomes positive, the local maximum becomes the absolute maximum, and it becomes more and more dominant as X increases. This, however, has little effect on $\ln Z$. Indeed, when the two maxima give equal contributions ($X = 0$) Z is multiplied by 2, which does not modify

$$\lim_{N\to\infty} \frac{1}{N} \ln Z.$$

On the other hand, M_1 is proportional to

$$\frac{\partial Z}{\partial X} = \int_{-1}^{+1} \frac{dM}{(1 - M^2)^{\frac{1}{2}}}\, M\, e^{N[f(M)+XM]}.$$

In this integral the two symmetric maxima give opposite contributions. In particular, they cancel exactly for $X = 0$. Thus M_1 varies very rapidly from one value to its opposite as the weight shifts from one maximum to the other.

The result of this analysis is that in the limit of $N \to \infty$, the curve representing $\tilde{\phi}(M_1)$ is the concave envelope of the curve representing $f(M)$ (Fig. 7), i.e., the curve obtained by joining the two maxima by a straight line. For finite N, the corresponding curve would be a concave analytic curve (dashed curve).

6. DIAGRAM EXPANSION FOR THE INFINITE-DIMENSIONAL MODEL

6.1 Molecular-Field Approximation

The factors associated with the various elements of a diagram for $\mathfrak{D}\{M_n\}$ are, in the case of the

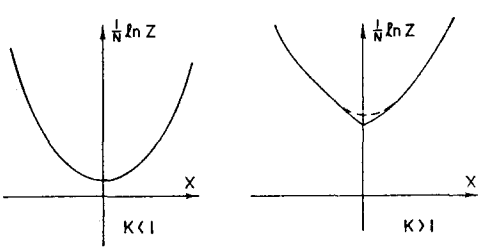

FIG. 6. The function $(1/N) \ln Z(X)$. The dashed line for $K > 1$ is the analytic function for finite N. The solid line is the nonanalytic limit function for $N \to \infty$.

infinite-dimensional model, K/N with each bond; NM_n with each point.

We can also decompose the factor K/N associated with each bond into two factors $(K/N)^{\frac{1}{2}}$ associated with each of the two points joined by the bond. The rule becomes then that we must associate a factor

$$N(K/N)^{n/2}M_n \tag{6.1}$$

with each point at which n bonds arrive.

This shows that all individual diagrams with no articulation points give contributions proportional to 1, $1/N$, $(1/N)^2$, $\cdots$ except the unique diagram consisting of two points joined by one bond, which gives a contribution proportional to N. Retaining only this diagram we obtain

$$\mathfrak{D} = \tfrac{1}{2}N^2(K/N)M_1^2 = \tfrac{1}{2}KNM_1^2. \tag{6.2}$$

This is identical with the molecular-field approximation (4.1). The corresponding function $\phi(M_1)$ is therefore given by (4.9):

$$\phi(M_1) = \frac{1}{2}KM_1^2 - \frac{M_1}{2}\ln\left(\frac{1+M_1}{1-M_1}\right) - \frac{1}{2}\ln\left(\frac{1-M_1^2}{4}\right). \tag{6.3}$$

It is identical with the function $f(M)$ defined by (5.6). It follows then from the discussion of the previous section that (6.3) is rigorously exact in the limit $N \to \infty$, except when $K > 1$ and M_1 is between the values of M which make $f(M)$ maximum. In this interval, the exact function is constant whereas (6.3) goes through a minimum (curve f on Fig. 7).

6.2 Ring Diagrams

The choice (6.2) for $\mathfrak{D}$ is based upon the following argument: we consider each term of the expansion of $\mathfrak{D}$ separately, consider its limit for $N \to \infty$, and sum the resulting contributions. In this procedure, all diagrams give a vanishing contribution in the limit $N \to \infty$, except the diagram containing a single bond and two points. A different result, however, might be obtained by summing first a family of diagrams, finding the maximum of the stationary functional and letting $N \to \infty$ at the end only.

This may be illustrated very easily by considering the ring diagrams. We return therefore to (4.14) which simplifies considerably since, in the infinite-dimensional model,

$$v_{\mathbf{q}} = 0 \quad \text{for} \quad \mathbf{q} \neq 0, \qquad v_0 = K/\beta.$$

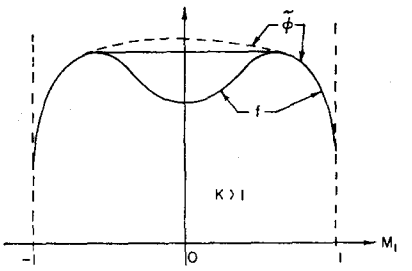

FIG. 7. The exact reduced functional $\tilde{\phi}(M_1)$ for the infinite-dimensional model. The dashed line represents $\tilde{\phi}(M_1)$ for finite N.

Thus we have

$$\phi(M_1, M_2) = \tfrac{1}{2}KM_1^2 - (1/2N)\ln(1 - KM_2) - M_1G_1 - M_2G_2 + M_0. \tag{6.4}$$

In addition, M_0, M_1, and M_2 are given in terms of G_1 and G_2 by the relations (4.16) and (4.17).

Let us discuss the reduced functional $\tilde{\phi}(M_1)$ obtained by substituting in (6.4) the value of M_2 which makes $\phi(M_1, M_2)$ stationary with respect to M_2:

$$\tilde{\phi}(M_1) = \phi(M_1, M_2), \quad \partial\phi(M_1, M_2)/\partial M_2 = 0. \tag{6.5}$$

The stationarity condition with respect to M_2 is equivalent to the second relation (4.18) which reduces here to

$$G_2 = (1/2N)K/(1 - KM_2),$$

or

$$1/K = 1/2NG_2 + M_2. \tag{6.6}$$

Elimination of M_2 between (6.6) and the Dyson relation (4.16) for M_2 yields for the determination of G_2 an equation which may be written

$$1/K = 1/2NG_2 + \mathcal{S}(G_1, G_2), \tag{6.7}$$

where

$$\mathcal{S}(G_1, G_2) \equiv \int_{-\infty}^{+\infty} \frac{dy}{(4\pi G_2)^{\frac{1}{2}}} e^{-y^2/4G_2} \operatorname{sech}^2(y + G_1). \tag{6.8}$$

Here we have replaced the integration variable z' of (4.16) by $y = z' - G_1$.

The behavior of the solution G_2 of Eq. (6.7) as $N \to \infty$ is most easily discussed by plotting $\mathcal{S}$ and $\mathcal{S} + 1/2NG_2$ as functions of G_2 (Fig. 8). For $G_2 = 0$, $\mathcal{S} = \operatorname{sech}^2 G_1$; and for G_2 large, $\mathcal{S} \sim G_2^{-\frac{1}{2}}$. When N is large, $\mathcal{S} + 1/2NG_2$ differs very little from $\mathcal{S}$, except for the small values of G_2. Let us denote by $\mu(G_1)$ the maximum of the function $\mathcal{S}(G_1, G_2)$ for G_2 varying between 0 and $+\infty$. We see from Fig. 8 that if

(a) $\dfrac{1}{K} > \mu(G_1)$, then $G_2 \to 0$, as $N \to +\infty$;

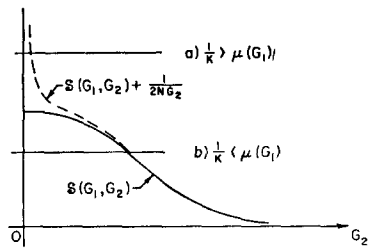

FIG. 8. Graphical solution of Eq. (6.7).

but that if

$$\text{(b)} \quad \frac{1}{K} < \mu(G_1),$$

then G_2 has a finite limit, the solution of the equation

$$\frac{1}{K} = \int_{-\infty}^{+\infty} \frac{dy}{(4\pi G_2)^{\frac{1}{2}}} e^{-y^2/4G_2} \operatorname{sech}^2(y + G_1) = S(G_1, G_2). \tag{6.9}$$

In the case (a), we obtain naturally the same answer for $\tilde{\phi}(M_1)$ as in the molecular-field approximation when the ring diagrams are omitted. But this is not true in the case (b).

It may be shown that if G_1 is sufficiently small

$$\mu(G_1) = S(G_1, 0) = \operatorname{sech}^2 G_1 \quad (G_1 \text{ small}). \tag{6.10}$$

According to (b), a nonvanishing G_2 then appears when

$$0 < K - \cosh^2 G_1 = K - (1 - M_1^2)^{-1} = \partial^2\phi(M_1)/\partial M_1^2, \tag{6.11}$$

where we have used (4.8), and where $\phi(M_1)$ is the molecular field approximation (6.3). The condition (6.11) is satisfied in the region limited by the two inflexion points of the molecular-field curve.

Thus $\tilde{\Phi}(M_1)$ is modified by the ring diagrams between the two inflection points of the molecular-field curve (Fig. 9). This applies when G_1 is sufficiently small in the condensation region, i.e., for small-enough K.

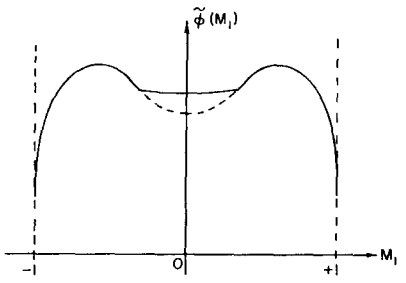

FIG. 9. The reduced functional in the ring-diagram approximation. (This curve does not represent a numerical calculation.)

7. GENERAL FAMILIES OF DIAGRAMS

7.1 The Model System

We now consider higher-order contributions to the functional $\mathfrak{D}\{M\}$. It should be emphasized that we do not mean "order" in the conventional perturbation theoretic sense, i.e., powers of the interaction strength K. In the last section we computed $\mathfrak{D}$ as a function of M_1 and M_2 and to all orders in K. Similarly, we now contemplate $\mathfrak{D}$ as a function of a larger but still finite number of semi-invariants, say, M_1, M_2, M_3, or M_1, M_2, M_4. Most generally we want to consider a set of M_n's with arbitrary indices n_i up to a maximum index ν. The set of all diagrams which can be formed with the given semi-invariants is what we call a "family of diagrams". By summing a complete family, we obtain in principle a non-perturbation-theoretic variational function which is still a function of only a finite number of variables.

The difficulty arises immediately that, for families with $\nu > 2$, we no longer are able to compute $\mathfrak{D}$ directly because we cannot sum these larger classes of skeleton diagrams. For practical purposes, then, we must investigate the convergence of the partial perturbation expansion required for approximation of $\mathfrak{D}$. To do this we take advantage of the exact solubility of the infinite dimensional model to obtain, at least implicitly, a generating function for this expansion.

Our strategy is based on the observation that, if we take a *model system* in which only a finite set of renormalized semi-invariants are nonzero, then the class of all skeleton diagrams for this system must be formally identical to the family of diagrams which define the desired function $\mathfrak{D}$ for the original system. This model system clearly must be the same as the one we introduced in Sec. 3, that is, one in which the μ_i are no longer restricted to ± 1, but may take any value μ with a probability distribution $p(\mu)$ such that all its semi-invariants vanish except those we have selected. Thus

$$\int_{-\infty}^{\infty} p(\mu)\, e^{\mu z}\, d\mu = \exp\left(\sum_n^{\nu} \frac{z^n}{n!} Q_n^0\right), \tag{7.1}$$

where Q_n^0 denotes the nonvanishing semi-invariants of $p(\mu)$. The partition function Ξ for this system is

$$\Xi = \int_{-\infty}^{\infty} d\mu_1 \cdots \int_{-\infty}^{\infty} d\mu_N p(\mu_1) \cdots p(\mu_N) \times \exp\left[\frac{K}{2N}\left(\sum_i \mu_i\right)^2\right]. \tag{7.2}$$

The diagram expansion for $(1/N) \ln \Xi \equiv \psi\{Q^0\}$ may

be generated in the usual way; and we may write, following Eq. (2.14),

$$\psi\{Q^0\} \equiv \frac{1}{N} \ln \Xi = \frac{1}{N} \mathfrak{D}\{Q\} - \sum_n Q_n J_n + Q_0, \quad (7.3)$$

where the Q_n, $n \leq \nu$, are the renormalized semi-invariants; and the J_n, related to the Q_n by

$$J_n = (1/N)\, \partial \mathfrak{D}/\partial Q_n, \quad (7.4)$$

are the contributions of the proper renormalization diagrams. Our point is that $\mathfrak{D}\{Q\}$ is in principle a known functional because $\psi\{Q^0\}$ is known. Expansion of $\mathfrak{D}\{Q\}$ in powers of the Q's generates the skeleton diagrams for the model system. But this set of diagrams is identical to the family of skeleton diagrams which determine $\mathfrak{D}\{M\}$. Furthermore, the variables Q_n, J_n are the same as the sets of variables M_n, G_n; thus we shall drop the notation Q_n, J_n from now on.

7.2 The Distribution $p(\mu)$

In the following analysis we need some properties of the distribution function $p(\mu)$.

The inverse of the Laplace transformation (7.1) is

$$p(\mu) = \frac{1}{2\pi i} \int_{a-i\infty}^{a+i\infty} dz \exp\left(\sum_n^{\nu} \frac{z^n}{n!} Q_n^0 - z \right). \quad (7.5)$$

The convergence of this integral depends on the parity of ν, the largest value of n such that $Q_n^0 \neq 0$, and on the sign of Q_ν^0. For ν even, convergence requires that

$$(-1)^{\nu/2} Q_\nu^0 < 0 \quad (\nu \text{ even}); \quad (7.6)$$

whereas for ν odd there are always values of a such that (7.5) converges.

It is interesting to examine the asymptotic behavior of $p(\mu)$ for large μ. The saddle-point method yields (Appendix B)

$$p(\mu) \approx \left(\frac{2\nu\, |\xi|}{\pi} \right)^{\frac{1}{2}} \frac{e^{\xi \sin\varphi}}{(\nu - 1)\, |\mu|} \cos\left(\frac{\varphi}{2} \pm \frac{\pi}{4} - \xi \cos\varphi \right) \quad (7.7)$$

for $\mu \to \pm\infty$, where

$$\xi = \frac{\nu - 1}{\nu} \mu \left[\frac{(\nu - 1)!\, |\mu|}{Q_\nu^0} \right]^{1/(\nu-1)} \sim |\mu|^{\nu/(\nu-1)}; \quad (7.8)$$

and the phase φ is defined by

$$\begin{aligned} \varphi &= [\pi/(\nu - 1)][n - \tfrac{1}{2}(\nu + 1)] && \text{for} \quad \mu > 0; \\ \varphi &= [\pi/(\nu - 1)][n - \tfrac{1}{2}(\nu + 1) + 2] && \text{for} \quad \mu < 0. \end{aligned} \quad (7.9)$$

Here n is the integer, even when $\mu/Q_\nu^0 < 0$ or odd when $\mu/Q_\nu^0 > 0$, such that

$$\begin{aligned} n \leq (\nu + 1)/2 < n + 2 \quad &\text{for} \quad \mu > 0; \\ n < (\nu + 1)/2 \leq n + 2 \quad &\text{for} \quad \mu < 0. \end{aligned} \quad (7.10)$$

Thus, roughly speaking, $p(\mu)$ exhibits for $\mu \to \pm\infty$ oscillations of decreasing amplitude proportional to an exponential function of $|\mu|^{\nu/(\nu-1)}$ except when $\varphi = 0$, i.e., if $\nu = 2n - 1$.

7.3 The Unrenormalized Expansion

Let us consider now the partition function Ξ given by Eq. (7.2). Since $p(\mu)$ always decreases more slowly than an exponential of μ^2, the convergence of (7.2) requires even for finite N that $K < 0$, corresponding to a repulsive (antiferromagnetic) interaction. We therefore leave out for the moment the case of an attractive interaction.

We use an integral representation of the delta function to write:

$$\begin{aligned} \Xi &= \int_{-\infty}^{+\infty} d\mu_1 \cdots \int_{-\infty}^{+\infty} d\mu_N p(\mu_1) \cdots p(\mu_N) N \\ &\quad \times \int_{-\infty}^{+\infty} dM\, e^{NKM^2/2}\, \delta\Big(\sum_i \mu_i - MN\Big) \\ &= \int_{-\infty}^{+\infty} d\mu_1 \cdots \int_{-\infty}^{\infty} d\mu_N p(\mu_1) \cdots p(\mu_N) N \\ &\quad \times \int_{-\infty}^{\infty} dM \frac{1}{2\pi i} \int_{-i\infty}^{+i\infty} dz \\ &\quad \times \exp\left[z\Big(\sum_i \mu_i - MN\Big) + \tfrac{1}{2} NKM^2 \right] \\ &= \left(\frac{N}{2\pi K} \right)^{\frac{1}{2}} \int_{-i\infty}^{i\infty} dz\, e^{NF(z)}, \end{aligned} \quad (7.11)$$

where

$$F(z) = \sum_{n=1}^{\nu} \frac{z^n}{n!} Q_n^0 - \frac{z^2}{2K}. \quad (7.12)$$

In deriving (7.11) we have carried out explicitly the integration in M and used (7.1) for the integrations over μ_i. It should be understood that for $K < 0$

$$K^{\frac{1}{2}} = +i(-K)^{\frac{1}{2}}. \quad (7.13)$$

The diagram expansion of Ξ for the model system is of the form of an expansion in powers of the Q_n^0, or, more precisely according to (6.1), in powers of

$$N(K/N)^{n/2} Q_n^0, \qquad n = 1, 2, \cdots.$$

Such an expansion is obtained immediately from (7.11) by introducing a new variable of integration y:

$$z = iy(-K/N)^{\frac{1}{2}},$$

which yields

$$\Xi = \frac{1}{(2\pi)^{\frac{1}{2}}} \int_{-\infty}^{\infty} dy\, e^{\theta - y^2/2}, \quad (7.14)$$

where

$$g(y) = \sum_{n=1}^{\nu} \frac{y^n}{n!} N\left(\frac{K}{N}\right)^{n/2} Q_n^0, \tag{7.15}$$

and by expanding e^g in powers of y. In (7.15) the convention (7.13) should be used for the half-integral powers of $K < 0$. This gives the expansion

$$\Xi = \sum_{n_1, n_2, \cdots} \frac{1}{(2\pi)^{\frac{1}{2}}} \int_{-\infty}^{\infty} dy \frac{e^{-y^2/2}}{n_1!\, n_2! \cdots} \times [(y/1!)(K/N)^{\frac{1}{2}} N Q_1^0]^{n_1} [(y^2/2!)\,(K/N) N Q_2^0]^{n_2} \cdots . \tag{7.16}$$

This expansion contains only integral powers of K because

$$\int_{-\infty}^{\infty} dy\, e^{-y^2/2} y^n = 0 \quad \text{for} \quad n \text{ odd}.$$

Thus Ξ may be considered as a formal generating function for the chosen family of diagrams. The form (7.14), (7.16) of Ξ shows that this applies for $K > 0$ as well as $K < 0$. When $K > 0$, we can bring Ξ back to a form similar to (7.11) by the change of variable

$$z = y(K/N)^{\frac{1}{2}},$$

which gives

$$\Xi = \left(\frac{N}{2\pi K}\right)^{\frac{1}{2}} \int_{-\infty}^{\infty} dz\, e^{NF(z)}, \tag{7.17}$$

The only difference between (7.11) and (7.17) is that the integration in (7.17) is taken now along the real axis instead of the imaginary axis.

The function Ξ cannot be considered as a generating function in the ordinary sense because the expansion (7.16) has a vanishing radius of convergence in $N(N/K)^{\nu/2} Q_\nu^0$. This follows from the fact that $Q_\nu^0 = 0$ is a branch point of Ξ considered as a function of the complex variable Q_ν^0 (Appendix C).

7.4 Renormalization Equations

The next step in the procedure is the transformation from the unrenormalized Q_n^0 to the renormalized variables M_n.

The Dyson equation is

$$M_n = \sum_{n_1, n_2, \cdots} \frac{1}{n_1!\, n_2! \cdots} G_1^{n_1} G_2^{n_2} \cdots Q_{n+n_1+2n_2+\cdots}^0 . \tag{7.18}$$

For example, if $\nu = 3$ we have

$$\begin{aligned} M_1 &= Q_1^0 + G_1 Q_2^0 + (\tfrac{1}{2} G_1^2 + G_2) Q_3^0; \\ M_2 &= Q_2^0 + G_1 Q_3^0; \\ M_3 &= Q_3^0. \end{aligned} \tag{7.19}$$

Rather than use Eq. (2.16) or (7.4) to evaluate the proper parts G_n, we must base the calculation on $\psi\{Q^0\}$ as computed in Sec. 7.3. This can be done by noting that

$$\frac{\partial \psi}{\partial Q_n^0} = \sum_{n_1, n_2, \cdots} \frac{1}{n_1!\, n_2! \cdots} G_1^{n_1} G_2^{n_2} \cdots , \qquad n_1 + 2n_2 + \cdots = n, \tag{7.20}$$

which is quite obvious in terms of diagrams. Indeed, differentiation of ψ with respect to Q_n^0 has the effect of selecting any of the points of index n. Such a point may be an articulation point for n_1 proper parts contributing one bond, n_2 contributing two bonds, etc., . . . , in such a way that

$$n_1 + 2n_2 + \cdots = n.$$

Each of these proper parts corresponds to a contribution to $G_1, G_2, \cdots$, respectively. The relation (7.20) follows immediately. Finally, we may combine Eqs. (7.18) and (7.20) to obtain the renormalization relations in closed algebraic form:

$$M_n = Q_n^0 + \sum_{m=1}^{\nu-1} \frac{\partial \psi}{\partial Q_m^0} Q_{m+n}^0. \tag{7.21}$$

The set of equations (7.21) has a remarkable integral representation. From (7.17) and (7.12) we obtain

$$\frac{\partial \psi}{\partial Q_n^0} = \int_{-\infty}^{\infty} dz \frac{z^n}{n!} e^{NF(z)} \Big/ \int_{-\infty}^{\infty} dz\, e^{NF(z)}. \tag{7.22}$$

Then using

$$\sum_{m=0}^{\nu-n} \frac{z^m}{m!} Q_{m+n}^0 = \left(\frac{d}{dz}\right)^n \left[F(z) + \frac{z^2}{2K}\right],$$

we obtain for the renormalization relations

$$\begin{aligned} M_1 &= \int_{-\infty}^{\infty} dz \frac{z}{K} e^{NF(z)} \Big/ \int_{-\infty}^{\infty} dz\, e^{NF(z)}; \\ M_2 &= \frac{1}{K} + \left[\int_{-\infty}^{\infty} dz F''(z)\, e^{NF(z)} \Big/ \int_{-\infty}^{\infty} dz\, e^{NF(z)}\right]; \\ M_n &= \int_{-\infty}^{\infty} dz F^{(n)}(z)\, e^{NF(z)} \Big/ \int_{-\infty}^{\infty} dz\, e^{NF(z)}, \\ & \qquad n = 3, 4, \cdots, \nu. \end{aligned} \tag{7.23}$$

In the first of these equations we have taken into account the identity

$$\int_{-\infty}^{\infty} dz F'(z)\, e^{NF(z)} = 0.$$

We need also the relation

$$M_0 = \int_{-\infty}^{\infty} dz\left[F(z) + \frac{z^2}{2K}\right] e^{NF(z)} \Big/ \int_{-\infty}^{\infty} dz\, e^{NF(z)}$$
$$= \int_{-\infty}^{\infty} dz M_0(z)\, e^{NF(z)} \Big/ \int_{-\infty}^{\infty} dz\, e^{NF(z)}. \qquad (7.24)$$

The equations for the G_n may conveniently be written

$$\exp\left(\sum_{n=1}^{\nu} x^n G_n\right) = \left[\int_{-\infty}^{\infty} dz\, e^{NF(z)+xz} \Big/ \int_{-\infty}^{\infty} dz\, e^{NF(z)}\right] + O(x^{\nu+1}), \qquad (7.25)$$

where it should be understood that the expansions in powers of x of both sides are identical up to order x^ν. Equation (7.25) follows immediately from the relations (7.20) and (7.22).

8. ASYMPTOTIC EVALUATIONS

8.1 Single Saddle-Point Evaluation

The function $F(z)$ as defined by Eq. (7.12) is just a finite polynomial in z. Therefore, except for very special choices of the Q_n^0 (which we shall see are very important), it is reasonable to assume that the integrals (7.23), (7.24), (7.25) are dominated by a single saddle-point in the limit $N \to \infty$. Let us call this dominant saddle-point s. We have then

$$\begin{aligned} M_1 &= s/K + O(N^{-\frac{1}{2}}); \\ M_2 &= 1/K + F''(s) + O(N^{-\frac{1}{2}}); \\ M_n &= F^{(n)}(s) + O(N^{-\frac{1}{2}}), \quad n = 3, 4, \cdots; \\ M_0 &= F(s) + s^2/2K + O(N^{-\frac{1}{2}}); \end{aligned} \qquad (8.1)$$

and

$$\exp\left(\sum_{n=1}^{\nu} x^n G_n\right) = e^{xs} + O(x^{\nu+1}) + O(N^{-\frac{1}{2}}). \qquad (8.2)$$

This last relation (8.2) gives

$$\begin{aligned} G_1 &= s + O(N^{-\frac{1}{2}}); \\ G_n &= O(N^{-\frac{1}{2}}), \quad n = 2, 3, \cdots, \nu. \end{aligned} \qquad (8.3)$$

Equations (8.1) show that the M_n in this approximation are the derivatives (or the derivative plus $1/K$ for $n = 2$) of $F(z)$ at the dominant saddle-point. This should be compared with the relations (7.12) according to which the unrenormalized Q_n^0 have the same meaning at $z = 0$.

Conversely, G_1 and the M_n determine $F(z)$ to order $N^{-\frac{1}{2}}$ by the Taylor series expansion

$$F(z) = F(s) - \frac{(z-s)^2}{2K} + \sum_{n=2}^{\nu} \frac{(z-s)^n}{n!} M_n + O(N^{-\frac{1}{2}}), \qquad (8.4)$$

where

$$F(0) = F(s) - \frac{s^2}{2K} + \sum_{n=2}^{\nu} \frac{(-s)^n}{n!} M_n = 0, \qquad (8.5)$$

valid to order $N^{-\frac{1}{2}}$, is required to complete the definition of $F(z)$ given by (8.4). We have in addition

$$s = G_1 \qquad (8.6)$$

to the same order of approximation.

Under the same assumption of uniqueness of the dominant saddle point, we obtain from (7.3) and (7.17)

$$\psi = F(s) + O(N^{-\frac{1}{2}}). \qquad (8.7)$$

Solving (7.3) for $\mathfrak{D}$ we find

$$\begin{aligned} \mathfrak{D}/N &= \psi + M_1 G_1 - M_0 \\ &= F(s) + \frac{s^2}{K} - F(s) - \frac{s^2}{2K} = \frac{s^2}{2K} = \frac{KM_1^2}{2}, \end{aligned} \qquad (8.8)$$

where we have used the relations (8.1), (8.3), and (8.7). This is precisely the molecular field approximation given by Eq. (6.2).

8.2 Domain of Variation of the M_n

It should be noted now that the single saddle-point assumption, although apparently very weak in terms of the Q_n^0, imposes a nontrivial constraint on the variation of the M_n. In particular, the M_n must be such that $z = s$ corresponds to the absolute maximum of $\mathrm{Re}F(z)$ along the contour of integration. For the case of an attractive (ferromagnetic) force, when the contour may be taken along the real axis, this condition requires that

$$-\frac{u^2}{2K} + \sum_{n=2}^{\nu} \frac{u^n}{n!} M_n < 0 \qquad (8.9)$$

for all real values of u. This means that when the unrenormalized semi-invariants Q_n^0 [i.e., the function $F(z)$] are allowed to take on all possible values, the renormalized M_n defined by the sum of the corresponding subdiagrams remain within a restricted domain, say, $\mathfrak{M}_1$.

The domain $\mathfrak{M}_1$ is clearly convex. If $\{M'\}$ and $\{M''\}$ (that is, $M_1', M_2', \cdots$ and $M_1'', M_2'', \cdots$) denote two points of $\mathfrak{M}_1$ such that (8.9) is satisfied for all u, then the same condition is satisfied also by

$$\{M\} = \lambda'\{M'\} + \lambda''\{M''\} \qquad (8.10)$$

with

$$\lambda', \lambda'' \geq 0, \qquad \lambda' + \lambda'' = 1.$$

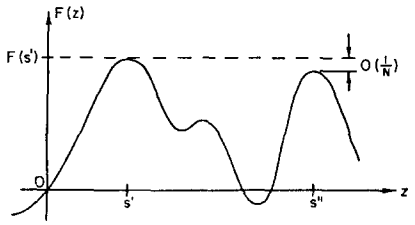

FIG. 10. A typical function $F(z)$ for $\{M'\}$ on the boundary of $\mathfrak{M}_1$.

In order to examine the domain $\mathfrak{M}_1$ for the infinite-dimensional model we return to the original function $\phi(M_1, M_2, \cdots)$ with $\mathfrak{D}$ given by (8.8). Because $G_2 = G_3 = \cdots = 0$, the Dyson relation is simply

$$M_n = (d/dG_1)^n \ln 2 \cosh G_1. \tag{8.11}$$

Condition (8.9) then reads

$$-u^2/2K + \ln 2 \cosh (u + G_1) - \ln 2 \cosh G_1 - u \tanh G_1 < 0, \tag{8.12}$$

where we have let $\nu \to \infty$. We may write (8.12) in the form

$$F_0(z) - F_0(G_1) - (z - G_1)F_0'(G_1) < 0, \tag{8.13}$$

where

$$z = G_1 + u$$

and

$$F_0(z) = -z^2/2K + \ln 2 \cosh z. \tag{8.14}$$

The inequality (8.13) for all z means that the curve representing $F_0(z)$ must be entirely under its tangent at $z = G_1$. This excludes precisely the values of G_1 corresponding to the two-phase region, in which the function ϕ following from (8.8) is manifestly wrong.

8.3 Two or More Equivalent Saddle Points

Now consider what happens when the point $\{M'\}$ approaches the boundary of the domain $\mathfrak{M}_1$. In the simplest case the function $F(z)$ has, in addition to its maximum at s', a single secondary maximum at, say, $z = s''$ such that $F(s'') \to F(s')$ as $\{M'\}$ approaches the boundary. A typical function of this sort is shown in Fig. 10. When $\{M'\}$ is so close to the boundary that $F(s') = F(s'')$ to order N^{-1}, then Eqs. (8.1) and (8.2) must be replaced by

$$\begin{aligned} M_1 &= (1/K)(\lambda' s' + \lambda'' s''); \\ M_2 &= 1/K + \lambda' F''(s') + \lambda'' F''(s''); \\ M_n &= \lambda' F^{(n)}(s') + \lambda'' F^{(n)}(s''), \qquad n = 3, 4, \cdots; \end{aligned} \tag{8.15}$$

$$\exp\left(\sum_{n=1}^{\nu} x^n G_n\right) = \lambda' e^{xs'} + \lambda'' e^{xs''} + O(x^{\nu+1}); \tag{8.16}$$

where λ' and λ'' clearly are positive numbers such that $\lambda' + \lambda'' = 1$.

In interpreting these equations it should be noted first that the point $\{M'\}$ on the boundary of $\mathfrak{M}_1$ is associated with a second point $\{M''\}$, also on the boundary. The set of semi-invariants $M_1'', M_2'', \cdots$ are the coefficients of the Taylor series expansion of $F(z)$ about the point s''. Conversely, had we started with $\{M''\}$ and let it approach the boundary, we should have obtained $\{M'\}$. Furthermore, from Eqs. (8.15) we find a whole new set of semi-invariants $\{M\}$ which lie on a straight line between $\{M'\}$ and $\{M''\}$. Because $\mathfrak{M}_1$ is convex, the new points $\{M\}$ all lie within the original domain, as shown in Fig. 11. They define a new domain $\mathfrak{M}_2$ which is contained in $\mathfrak{M}_1$ and may or may not cover it completely. On the other hand, Eq. (8.16) implies that the G_n are not just weighted averages of the G_n' and G_n'' but are much more complicated. In fact, we obtain a series of nonvanishing G_2, G_3, etc., and consequently, a functional $\mathfrak{D}$ quite different and very much more complicated than that of Eq. (8.8). Thus, at the points $\{M\}$ in $\mathfrak{M}_2$ we have at least two different functionals, say $\mathfrak{D}_1\{M\}$ and $\mathfrak{D}_2\{M\}$.

In general, we must consider the set of points on the boundary of $\mathfrak{M}_1$ such that $F(z)$ has, say, p equivalent saddle-points. Associated with any one such point there will be $p - 1$ other points on the boundary. By an obvious generalization of (8.15), these sets of p points generate a domain $\mathfrak{M}_p$ on which is defined a pth new functional $\mathfrak{D}_p\{M\}$. The convexity of $\mathfrak{M}_1$ implies that any $\mathfrak{M}_p$ lies entirely within $\mathfrak{M}_1$. It should be emphasized that the variational principle applies over the entire set of functionals $\mathfrak{D}_p\{M\}$.

8.4 Convergence of the Expansions

The convergence of the various expansions which have been used in the summation of general families of diagrams may be discussed by means of the generating functions (7.11) or (7.17).

Let us consider first the unrenormalized expansion of Sec. 7.3. It has already been mentioned that this expansion (7.16) is not a convergent expansion. The discussion given in Appendix C shows however that this is an asymptotic expansion for $N \to \infty$ provided the maximum of the real part of the expo-

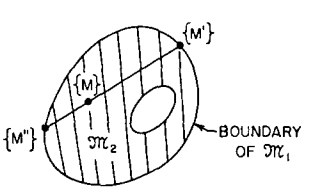

FIG. 11. The domain $\mathfrak{M}_1$ showing a point $\{M\}$ in $\mathfrak{M}_2$ generated according to Eqs. (8.15).

nent in the integrand takes place at the origin $z = 0$. This occurs only when $K < 1$, i.e., above the critical temperature.

The renormalized expansion is obtained less directly. The discussion of the integrals occurring in (7.23) is essentially the same as in Appendix C. One important difference, however, is the fact that we do not have to expand the exponent around $z = 0$ as in the unrenormalized case. Assuming the existence of one dominant saddle-point s, we can expand $F(z)$ around s:

$$F(z) = F(s) + \sum_{n=2} \frac{(z-s)^n}{n!} F^{(n)}(s).$$

The argument of Appendix C then yields asymptotic expansions of M_1, M_2, $\cdots$, M_n, and also of Ξ (from 7.11 or 7.17) in terms of s and the $F^{(n)}(s)$. It is essential to note here that the condition that the maximum of Re $F(z)$ along the contour should take place at $z - s = 0$ is now automatically satisfied.

It is still necessary to invert the relations giving the M_n in terms of the $F^{(n)}(s)$. Substitution of the resulting expressions of the $F^{(n)}(s)$ in terms of the M_n into Ξ, will, in principle, yield an asymptotic expansion for $\mathfrak{D}_1$ in terms of the M_n. This expansion is valid inside the domain $\mathfrak{M}_1$, where it generates the family of skeleton diagrams. The functionals $\mathfrak{D}_2$, $\mathfrak{D}_3$, $\cdots$ do not generate the perturbation expansion.

These considerations apply only inside the domain $\mathfrak{M}_1$. Outside $\mathfrak{M}_1$, the function $F(z)$ reconstructed from M_2, M_3, $\cdots$ has a saddle point giving a larger contribution than the initial saddle points. It is however possible to find a contour going through s, but not through the other saddle point, simply by staying in the valleys nearest to s. This defines a new function towards which the expansion is now asymptotically convergent. Thus, outside the domain $\mathfrak{M}_1$, the renormalized expansion remains asymptotically convergent, but it does not any longer represent the equilibrium partition function.

It is tempting to interpret the functional $\mathfrak{D}_1$ outside $\mathfrak{M}_1$ as giving the partition function of a metastable phase. In fact, for the infinite-dimensional model this interpretation is almost certainly correct. The analytic function $\phi(M_1)$ given by Eq. (6.3) is identical to $f(M)$ given by Eq. (5.6); and $f(M)$ clearly gives rise to a two-peaked probability distribution. More precisely, the function $f(M) + XM$ exhibits a stable and a metastable peak in the presence of the external field X. Note that the anomalous behavior associated with the divergence of the ring diagrams always occurs *outside* $\mathfrak{M}_1$. This singularity apparently corresponds to the limit of metastability.

In accord with the discussion of metastability, it is apparent that the renormalized expansion converges asymptotically to a functional $\mathfrak{D}_1$ describing a single phase with uniform magnetization. Presumably $\mathfrak{D}_2$ has something to do with two phases in equilibrium; but the connection is not really clear. The physical interpretation of the higher functionals $\mathfrak{D}_p$ is also doubtful. In any case, we know from the theorem of Sec. 3 that the reduced functional $\tilde{\Phi}(M_1)$ should be concave; thus, in order to compute the complete $\tilde{\Phi}$ from $\mathfrak{D}_1$, we need only take the concave envelope of the perturbation-theoretic result. A serious study of the two phase region for more realistic models must await further developments.

ACKNOWLEDGMENTS

One of the authors (C.B.) would like to thank Professor J. Ashkin and the Physics Department of Carnegie Institute of Technology for the hospitality and stimulation he has enjoyed during the fall semester of 1963.

APPENDIX A: THE SEMI-INVARIANTS

Straightforward expansion of (2.1) in powers of v_{ij} yields:

$$Z = Z_0 \sum_{m,i,j,\cdots} \frac{1}{m!} \left(\frac{\beta}{2}\right)^m v_{ij} v_{kl} \cdots \langle \mu_i \mu_j \mu_k \mu_l \cdots \rangle, \quad \text{(A1)}$$

where m denotes the number of factors v_{ij}, v_{kl}, $\cdots$, and the bracket the expectation value defined for any function of the μ_i by

$$\langle \varphi(\mu_1, \mu_2, \cdots) \rangle = \frac{1}{Z_0} \sum_{\{\mu\}} \varphi(\mu_1, \mu_2, \cdots) \exp\left(\sum_i X_i \mu_i\right). \quad \text{(A2)}$$

Here Z_0 denotes the partition function (2.3) in the absence of interaction. For instance,

$$\begin{aligned} \langle \mu_i^2 \rangle &= \langle \mu_i^4 \rangle = \cdots = \langle \mu_i^{2n} \rangle = 1, \\ \langle \mu_i \rangle &= \langle \mu_i^3 \rangle = \cdots = \langle \mu_i^{2n+1} \rangle = \tanh X_i. \end{aligned} \quad \text{(A3)}$$

Clearly

$$\begin{aligned} \langle \mu_i \mu_j \rangle &= \langle \mu_i \rangle \langle \mu_j \rangle, \qquad \text{if } i \neq j, \\ \langle \mu_i \mu_j \mu_k \rangle &= \langle \mu_i \rangle \langle \mu_j \rangle \langle \mu_k \rangle, \quad \text{if } i \neq j \neq k \neq i, \text{ etc.} \end{aligned} \quad \text{(A4)}$$

The relations (A3) and (A4) give immediately the values of all the brackets occurring in the expansion (A1). These values, however, depend on which groups of indices i, j, k, $\cdots$ are identical in each term of the summation. In order to overcome this

complication, let us define new quantities, denoted by curly brackets, through the following relations:

$$\begin{aligned}\langle\mu_i\rangle &= \{\mu_i\},\\ \langle\mu_i\mu_j\rangle &= \{\mu_i\mu_j\} + \{\mu_i\}\{\mu_j\},\\ \langle\mu_i\mu_j\mu_k\rangle &= \{\mu_i\mu_j\mu_k\} + \{\mu_i\}\{\mu_j\mu_k\} + \{\mu_j\}\{\mu_i\mu_k\}\\ &\quad + \{\mu_k\}\{\mu_i\mu_j\} + \{\mu_i\}\{\mu_j\}\{\mu_k\}, \cdots .\end{aligned} \tag{A5}$$

The right-hand side of these relations involves the sum of the products of curly brackets obtained in splitting the factors $\mu_1, \mu_j, \mu_k, \cdots$ in all possible ways. It is easy to see that, as a consequence of the relations (A4), the curly brackets satisfy the relations

$$\{\mu_i\mu_j\mu_k \cdots\} = 0, \quad \text{except if } i = j = k = \cdots . \tag{A6}$$

We can therefore write

$$\{\mu_i\mu_j\mu_k \cdots\} = \delta_{ij}\,\delta_{ik} \cdots M^0_{n,i}, \tag{A7}$$

where n is the number of factors $\mu_i, \mu_j, \cdots$ in the curly bracket, and

$$M^0_{n,i} = \{\mu_i^n\}. \tag{A8}$$

Substitution of the expressions (A5) for the brackets into (A1), taking into account the relations (A7), leads immediately to the diagram expansions of Z and $\ln Z$ described in Sec. 2.

The simplest way to derive the actual values of the semi-invariants $M^0_{n,i}$ is to consider a fictitious system without two-body interaction v_{ij}, where the field X_i is replaced by $X_i + \delta X_i$. The partition function is then

$$\begin{aligned}Z &= \sum_{\{\mu\}} \exp \sum_i (X_i + \delta X_i)\mu_i\\ &= \prod_i 2\cosh (X_i + \delta X_i),\end{aligned}$$

and therefore

$$\begin{aligned}\ln Z &= \ln Z_0 + \sum_i \ln 2\cosh (X_i + \delta X_i)\\ &= \sum_i \sum_{n=0}^{\infty} \frac{\delta X_i^n}{n!}\left(\frac{d}{dX_i}\right)^n \ln 2\cosh X_i.\end{aligned} \tag{A9}$$

On the other hand, the only connected diagrams in this case consist of a single point with $n = 1, 2, \cdots$ incoming δX_i lines. Thus

$$\ln Z = \ln Z_0 + \sum_i \sum_{n=1}^{\infty} \frac{\delta X_i^n}{n!} M^0_{n,i}, \tag{A10}$$

where $n!$ is the symmetry factor of the diagram. Comparison of (A9) and (A10) yields the relation (2.2).

Except for the particular values of the semi-invariants, the above argument clearly applies to any probability distribution $p_i(\mu_i)$ of the values of the spins.

APPENDIX B: ASYMPTOTIC EVALUATION OF EQUATION (7.5)

The saddle points of the exponent in (7.5) are given by

$$\sum_n \frac{z^{n-1}}{(n-1)!} Q^0_n - \mu = 0. \tag{B1}$$

When μ becomes very large, $|z|$ for the saddle points becomes very large also and the term of highest order in z dominates. The equation (B1) may therefore be reduced to

$$z^{\nu-1}Q^0_\nu/(\nu-1)! - \mu = 0. \tag{B2}$$

This gives $\nu - 1$ saddle points on a circle of radius

$$r = |(\nu-1)!\,\mu/Q^0_\nu|^{1/(\nu-1)}. \tag{B3}$$

The value of the exponent at the saddle points may be written, by using (B2),

$$zQ^0_\nu/\nu! - \mu z = -(\nu-1)\mu z/\nu. \tag{B4}$$

Similarly, the second derivative at the saddle points reads

$$z^{\nu-2}Q^0_\nu/(\nu-2)! = (\nu-1)\mu/z. \tag{B5}$$

We must now examine through which saddle points it is possible to draw a contour going from $-i\infty$ to $+i\infty$, in such a way that the real part of the exponent,

$$(\rho^\nu/\nu!)Q^0_\nu \cos\nu\theta - \mu\rho\cos\theta \qquad (z = \rho e^{i\theta}), \tag{B6}$$

has its maximum at the selected saddle point or points. Let us consider the variation of (B6) for fixed θ, as ρ goes from 0 to $+\infty$. This function goes through a minimum equal to

$$m(\theta) = -\frac{\nu-1}{\nu}\,\mu\cos\theta\left(\frac{(\nu-1)!\,\mu\cos\theta}{Q^0_\nu\cos\nu\theta}\right)^{1/(\nu-1)}, \tag{B7}$$

when θ is such that

$$\mu\cos\theta > 0, \qquad Q^0_\nu\cos\theta > 0. \tag{B8}$$

These conditions define sectors of the complex plane, each of which contains one of the saddle points (Fig. 12). Within each sector, the function (B7) goes through a maximum at the saddle point. It follows that for each contour going through one of the sectors defined by (B8), the real part of the exponent takes values at least equal to the saddle-point value.

On the other hand, the conditions

$$\mu \cos \theta < 0, \qquad Q_\nu^0 \cos \nu\theta > 0 \tag{B9}$$

define another set of sectors within which the real part of the exponent is positive, and therefore larger than at any of the saddle points contained in the sectors (B8). It is now clear that a contour C with a maximum value of the integrand in (7.5) as small as possible is obtained by crossing all the sectors (B8) at their respective saddle points (Fig. 12). From (B4), it follows that the two saddle points nearest to the imaginary axis give the dominant contributions. This yields immediately the expression (7.7).

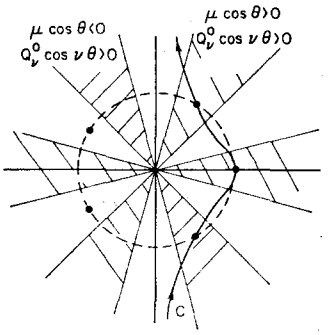

FIG. 12. Contour of integration for Eq. (7.5) in the case $\nu = 6$, $Q_6^0 > 0$, and $\mu > 0$.

APPENDIX C: EXPANSIONS OF CERTAIN CONTOUR INTEGRALS

Let us consider an integral of the form

$$I_N(a_2, a_3, \cdots, a_\nu) = \frac{1}{(2\pi)^{\frac{1}{2}}} \int_{-\infty}^{+\infty} dz\, e^{NF(z)}, \tag{C1}$$

where

$$F(z) = a_2 \frac{z^2}{2!} + a_3 \frac{z^3}{3!} \cdots + a_\nu \frac{z^\nu}{\nu!}. \tag{C2}$$

The integral (C1), when taken along the real axis is convergent only if

$$\nu \text{ even}, \qquad \operatorname{Re} a_\nu < 0. \tag{C3}$$

More generally, however, the integral may be taken along a contour C coming from infinity in the direction θ_-, and going to infinity in the direction θ_+. The convergence conditions (C3) are then replaced by

$$\operatorname{Re} \{a_\nu e^{i\nu\theta_\pm}\} < 0, \tag{C4}$$

which reduces to (C3) when $\theta_+ = 0$, $\theta_- = \pi$.

It is easy to see that when a_ν and the contour C are fixed according to (C3) or (C4), the integral (C1) is absolutely and uniformly convergent when $a_2, a_3, \cdots, a_{\nu-1}$ vary inside any bounded region. The integral $I_N(a_2, a_3, \cdots, a_\nu)$ is therefore an entire function of $a_2, a_3, \cdots, a_{\nu-1}$ for fixed a_ν. The expansion in powers of $a_2, a_3, \cdots, a_{\nu-1}$ obtained by carrying out the z integration in

$$I_N = \frac{1}{(2\pi)^{\frac{1}{2}}} \int dz\, e^{Na_\nu z^\nu/\nu!} \times \left[1 + NF_1(z) + \frac{N^2}{2!} F_1(z)^2 + \cdots\right], \tag{C5}$$

where

$$F_1(z) = a_2 \frac{z^2}{2!} + a_3 \frac{z^3}{3!} + \cdots + a_{\nu-1} \frac{z^{\nu-1}}{(\nu-1)!}$$

is convergent for all values of these variables. The successive terms of this expansion, however, involve increasing positive powers of N, and the expansion is therefore not appropriate to a study of the limit $N \to \infty$.

The analytic properties of (C1) with respect to a_ν are quite different. For each value of a_ν, the conditions (C4) define permitted sectors of angle π/ν for θ_+ and θ_-. This allows $\nu(\nu-1)/2$ choices of contour giving as many different values for I_N. The contours such that $\theta_+ - \theta_-$ is the same actually correspond to the different branches of the same analytic function around the point $a_\nu = 0$. Indeed the conditions (C4) require a rotation of the contour when a_ν turns around 0. To one turn of a_ν corresponds a decrease of θ_+ and θ_- equal to $2\pi/\nu$, which yields another contour C such that $\theta_+ - \theta_-$ has not changed. This gives altogether $(\nu-1)/2$ functions having ν determinations at the branch point $a_\nu = 0$.

It follows immediately that the expansion in powers of $a_3, a_4, \cdots, a_\nu$ obtained by carrying out the z integration in

$$I_N = \frac{1}{(2\pi)^{\frac{1}{2}}} \int dz\, e^{Na_2 z^2/2} \times [1 + NF_2(z) + N^2F_2(z)^2/2! + \cdots], \tag{C6}$$

where

$$F_2(z) = a_3 \frac{z^3}{3!} + a_4 \frac{z^4}{4!} + \cdots + a_\nu \frac{z^\nu}{\nu!}, \tag{C7}$$

has a vanishing radius of convergence. We are, however, led to investigate this expansion further in the case $N \to \infty$, by the fact that the successive terms are proportional to increasing powers of $N^{-\frac{1}{2}}$.

Let us show that (C6) actually leads to an asymptotic expansion for $N \to \infty$, provided the maximum of $\operatorname{Re} F(z)$ along the contour of integration takes place at $z=0$. Let α be a number such that $\frac{1}{3} < \alpha < \frac{1}{2}$, and let us divide the contour of integration in (C1) into two parts C_1 and C_2 by

$$C_1 : |z| > a/N^\alpha, \qquad C_2 : |z| < a/N^\alpha;$$

where a is an arbitrary positive number. Since the maximum of Re $F(z)$ is equal to 0 at $z = 0$, we have along C_1

$$\text{Re } F(z) < -b/N^{2\alpha},$$

where b is some positive number. It follows immediately that the contribution I_1 of the portion C_1 of the contour to I_N satisfies the inequality

$$|I_1| < c \exp(-bN^{1-2\alpha}), \tag{C8}$$

where C is some positive number. This assumes, of course, that the integral is convergent.

For z on C_2, there is clearly a positive number d such that

$$\left|\exp N\left(a_3\frac{z^3}{3!} + \cdots + a_\nu\frac{z^\nu}{\nu!}\right) - S_n(z)\right| < dN^{n/3}z^n, \tag{C9}$$

where $S_n(z)$ denotes the expansion in powers of z of the exponential occurring in (C9) up to the order $n - 1$. It follows now from (C8) and (C9) that

$$\begin{aligned}|I_N - I_N^{(n)}| &< c \exp(-bN^{1-2\alpha}) + \frac{dN^{n/3}}{(2\pi)^{\frac{1}{2}}}\left|\int_{C_2} dz\, z^n\, e^{Na_3z^3/2}\right| \\ &< c\exp(-bN^{1-2\alpha}) + dI_n\,|a_2|^{-(n+1)/2}N^{-(n+3)/6},\end{aligned} \tag{C10}$$

where

$$I_n = \frac{1}{(2\pi)^{\frac{1}{2}}}\int_{-\infty}^{+\infty} du\, u^n\, e^{-u^2/2}.$$

In (C10), $I_N^{(n)}$ denotes the expression obtained from (C6) by retaining in the bracket the terms of order lower than n in z. This relation shows that this procedure yields an asymptotic expansion for $N \to \infty$.

Note on Positronium*

GEORGE TIKTOPOULOS†
Palmer Physical Laboratory, Princeton University, Princeton, New Jersey
(Received 19 October 1964)

By means of an inequality providing an upper bound for the norm of integral operators, it is shown that the Bethe–Salpeter equation for bound states of the electron–positron system (in the ladder approximation) admits solutions associated with a discrete spectrum of binding energies. It is found that in the weak coupling limit the spectrum $B(\alpha)$ approaches asymptotically the Coulomb spectrum in the sense

$$\lim_{\alpha\to 0}\,(B(\alpha)/m\alpha^2) = \{(1/4n^2)\mid n = 1, 2, 3, \cdots \infty\},$$

where α is the fine-structure constant and m the electron rest mass.

ONE of the major successes of quantum electrodynamics has been the calculation of the energy levels of the electron–positron bound system[1] (poistronium) in agreement with experiments of high accuracy.[2] These calculations were based on the field-theoretic Bethe–Salpeter (BS) equation and in fact remain the only direct numerical confrontation of the BS formalism with experimental data. It is, accordingly, an annoying fact that no useful hints were obtained from these calculations as to the existence of solutions and the discreteness of the spectrum of the BS equation, because the approach used was essentially a perturbational scheme of corrections to the "instantaneous" Coulomb problem. It seems, therefore, desirable from the theoretical point of view to demonstrate that the BS equation for states of two spin-$\frac{1}{2}$ particles bound by the electromagnetic field admits solutions corresponding to a *discrete* energy spectrum which is approximated in lowest order (i.e., in the weak coupling limit) by the nonrelativistic Coulomb spectrum. It is understood, of course, that the interaction kernel is to be approximated by a finite number of irreducible graphs. The corresponding problem for spinless particles bound by a scalar field was essentially solved by Wick[3] who was able to reduce the problem to the eigenvalue equation for a completely continuous integral operator. However in the case of spin-$\frac{1}{2}$ particles interacting through a vector field the kernel of the integral equation is *not* completely continuous (because of the behavior of the spinor propagators for large momenta) and the problem has since remained open although it corresponds to the physically interesting situation. An embarrassing set of solutions indicating a continuous spectrum was found by Goldstein[4] in the "ladder" approximation for the special case of zero total energy but these solutions were shown to be inadmissible by Mandelstam's normalization condition.[5]

In this note we consider the BS equation for e^+e^- bound states in the "ladder" approximation (i.e., we approximate the interaction kernel by the one-photon-exchange diagram), although we believe that our approach can be applied to more general situations. It is shown that *the equation admits solutions compatible with Mandelstam's condition and associated with a discrete spectrum, which coincides asymptotically with the nonrelativistic Coulomb spectrum in the weak coupling limit in the sense* $\lim_{\alpha\to 0} B/m\alpha^2 = \frac{1}{4}n^2$; $n = 1, 2, \cdots \infty$, where B is the binding energy, α the fine-structure constant and m the electron rest mass. Any other solutions must violate a simple integrability condition, which although compatible with Mandelstam's criterion does not seem to follow directly from physical considerations.

We consider the Fourier transform $\Phi_{\alpha\beta}(p)$ of the BS wavefunction

$$\chi_{\alpha\beta}(x) = \langle 0\mid T[\psi_\alpha(\tfrac{1}{2}x)\bar{\psi}_\beta(-\tfrac{1}{2}x)]\mid B\rangle, \qquad (1)$$

for a e^+e^- bound state represented by $|B\rangle$. Here $\psi(x)$ is the electron field operator ($\bar{\psi} = \psi^\dagger\gamma_4$) and T is Wick's chronological operator. The BS equation in the ladder approximation reads

$$\Phi(p) = -\frac{e^2}{(2\pi)^4 i}\frac{1}{i\gamma(p+\frac{1}{2}P)+m} \times\int\frac{d^4k}{(p-k)^2}\gamma_\mu\Phi(k)\gamma_\mu\frac{1}{i\gamma(p-\frac{1}{2}P)+m}, \qquad (2)$$

* Work supported by the U. S. Air Force Office of Research, Air Research and Development Command.

† Part of this work was done during the author's stay at the Physics Department, Brookhaven National Laboratory, Upton, Long Island, New York.

[1] R. Karplus and A. Klein, Phys. Rev. **87**, 848 (1952).
[2] M. Deutsch, Phys. Rev. **87**, 212 (1952).
[3] G. C. Wick, Phys. Rev. **96**, 1124 (1954).
[4] J. Goldstein, Phys. Rev. **91**, 1516 (1953).
[5] S. Mandelstam, Proc. Roy. Soc. (London) **A233**, 248 (1955).

where Φ is the 4×4 matrix with components $\Phi_{\alpha\beta}(p)$ and P is the total 4-momentum of state $|B)$. We use the metric $pk = \mathbf{p}\cdot\mathbf{k} - p_0k_0 = \mathbf{p}\cdot\mathbf{k} + p_4k_4$ and γ matrices that are Hermitian and unitary. For concerteness we work in the c.m. frame in which $P_1 = P_2 = P_3 = 0$, $P_4 = iP_0 = iE$.

We first perform the Wick rotation[3] of the integration path of the relative energy to the imaginary axis in order to "cure" the singularities of the kernel due to the indefiniteness of the relativistic metric. Next we require that the function

$$f(p) = \Phi(p)[i\gamma(p - \tfrac{1}{2}P) + m] \tag{3}$$

be square-integrable in the sense

$$\int \mathrm{Tr}\,[f^\dagger(p)f(p)]\,d^4p < +\infty. \tag{4}$$

Although it would be very interesting to see if this requirement can be deduced from a physical argument, we have introduced it only in order to apply the methods of ordinary functional analysis. We note that (4) excludes the Goldstein solutions[4] and is compatible with Mandelstam's criterion[5] for the singularity of $\chi(x)$ at $x^2 = 0$.

It is clear that the 4×4 function matrices satisfying (4) form a Hilbert space $\mathfrak{H}$ with inner product

$$(\varphi_1, \varphi_2) = \int \mathrm{Tr}\,[\varphi_1^\dagger(p)\varphi_2(p)]\,d^4p$$

and the function $f(p)$ defined by (3) satisfies

$$f(p) = -\frac{e^2}{(2\pi)^4}\frac{1}{i\gamma(p + \frac{1}{2}P) + m} \times \int \frac{d^4k}{(p-k)^2}\gamma_\mu f(k)\frac{1}{i\gamma(k - \frac{1}{2}P) + m}\gamma_\mu, \tag{5}$$

or symbolically $f = Kf$ where K represents an operator on $\mathfrak{H}$. We now show that K is a *bounded* operator on $\mathfrak{H}$. By means of the inequality

$$|\mathrm{Tr}\,(A_1A_2\cdots A_n)| \le [\mathrm{Tr}\,(A_1^\dagger A_1)\,\mathrm{Tr}\,(A_2^\dagger A_2)\cdots\mathrm{Tr}\,(A_n^\dagger A_n)]^{\frac{1}{2}}, \tag{6}$$

we deduce that

$$|(\varphi_1, K\varphi_2)| \le \frac{e^2}{\pi^4}\int d^4p\,d^4k\,\mathrm{Tr}^{\frac{1}{2}}\,[\varphi_1^\dagger(p)\varphi_1(p)] \times \left[\frac{p^2 + m^2 + E^2/4}{(p^2 + m^2 - E^2/4)^2 + p_4^2E^2}\right]^{\frac{1}{2}}\frac{1}{(p-k)^2} \times \left[\frac{k^2 + m^2 + E^2/4}{(k^2 + m^2 - \frac{1}{4}E^2)^2 + k_4^2E^2}\right]^{\frac{1}{2}}\mathrm{Tr}^{\frac{1}{2}}\,[\varphi_2^\dagger(k)\varphi_2(k)].$$

The boundedness of K may thus be inferred from that of the operator with kernel

$$M(p,k) = \frac{e^2}{\pi^4}\left[\frac{p^2 + m^2 + E^2/4}{(p^2 + m^2 - \frac{1}{4}E^2)^2 + p_4^2E^2}\right]^{\frac{1}{2}} \times \frac{1}{(p-k)^2}\left[\frac{k^2 + m^2 + E^2/4}{(k^2 + m^2 - \frac{1}{4}E^2)^2 + k_4^2E^2}\right]^{\frac{1}{2}}$$

on the Hilbert space of (one-component) functions $g(p)$ with

$$\int |g(p)|^2\,d^4p < +\infty.$$

We now use an inequality for the norm of an operator which is our main mathematical tool in this paper and is derived in the Appendix. Denoting by $|M|$ the norm of the operator with kernel $M(p, k)$ we have

$$|M| \le \sup_{(p)}\int |M(p,k)|\,\sigma(k)\,d^4k/\sigma(p) \qquad \text{for any} \quad \sigma(k) > 0.$$

If we choose

$$\sigma(k) = \left[\frac{k^2 + m^2 + E^2/4}{(k^2 + m^2 - \frac{1}{4}E^2)^2 + k_4^2E^2}\right]^{\frac{1}{2}} \times \frac{(\mathbf{k}\cdot\mathbf{b})}{(k^2 + a^2)^2}\frac{\mathbf{k}^2 + m^2 - E^2/4}{(\mathbf{k}^2 + m^2)^{\frac{1}{2}}}$$

and majorize $M(p, k)$ by replacing $(p - k)^{-2}$ by $(\mathbf{p} - \mathbf{k})^{-2}$ we can carry out the integrations to obtain

$$|K| \le |M| \le \sup_{(y)}\frac{e^2}{\pi^2}\frac{(y^2 + m^2/a^2)^{\frac{1}{2}}}{y^2 + (m^2 - \frac{1}{4}E^2)/a^2} \times \left(1 + \frac{1}{y^2}\right)^2\left(y\tan^{-1}y - \frac{y^2}{y^2+1}\right) < \frac{e^2}{2\pi}\frac{m}{(m - \frac{1}{4}E^2)^{\frac{1}{2}}}. \tag{7}$$

This result already imposes a restriction on the values of E for which (5) is soluble,

$$(m^2 - \tfrac{1}{4}E^2)/2m^2 < 2(e^2/4\pi)^2,$$

to be compared to the nonrelativistic Coulomb spectrum

$$\frac{m^2 - \frac{1}{4}E_n^2}{2m^2} = \frac{1}{4n^2}\left(\frac{e^2}{4\pi}\right)^2; \qquad n = 1, 2, 3, \cdots.$$

We are using rationalized units in which $\alpha = e^2/4\pi = \frac{1}{137}$. Writing $K(E)$ to indicate the dependence of K on E we observe that, according to (7), $[1 - \lambda K(0)]^{-1}$ is bounded and analytic in λ for

$|\lambda| < 1/2\alpha$. Now it is straightforward to verify that $[K(E) - K(0)]^\dagger[K(E) - K(0)]$ is a Schmidt operator (i.e. has a square-integrable kernel) on $\mathfrak{K}$, although $K(E)$ is not. {This is because the large momentum contribution of $K(E)$ is cancelled by that of $K(0)$. The reason for considering $[K(E) - K(0)]^\dagger[K(E) - K(0)]$ rather than $K(E) - K(0)$ is that the latter is not of the Schmidt type because of the singularity of $(p - k)^{-2}$ at $p = k$}. It follows[6] that $K(E) - K(0)$ is completely continuous and consequently so is $[1 - \lambda K(0)]^{-1}[K(E) - K(0)]$ for $|\lambda| < 1/2\alpha$. Furthermore, from a general theorem[7] on completely continuous operators depending analytically on a parameter, it follows that

$$[1 - \lambda K(E)]^{-1} = \{1 - \lambda[1 - \lambda K(0)]^{-1} \times [K(E) - K(0)]\}^{-1}[1 - \lambda K(0)]^{-1}$$

is meromorphic in λ for $|\lambda| < 1/2\alpha$. Its poles at $\lambda_N = \gamma_N(E)$ correspond to the eigenvalues of $K(E)$. Now $K(E)$ is analytic in E in some region containing the open interval $(0, 2m)$; therefore the $\gamma_N(E)$'s are analytic[8] (with possibly several branch points) in that region. We conclude that only a discrete set of points of the region can satisfy $\gamma_N(E) = 1$; thus *the energy spectrum can only be discrete.*

We turn now to the weak coupling limit. It is convenient to make the following change of variables

$$\mathbf{p} = \epsilon\mathbf{q}, \quad p_4 = \epsilon^2 v, \quad \text{where} \quad \epsilon = (m^2 - E^2/4)^{\frac{1}{2}}/m,$$
$$\bar{f}(\mathbf{q}, v) = f(p).$$

For $\bar{f}$ we have the equation

$$\bar{f}_N(\mathbf{q}, v) = -\frac{\Lambda_N(\epsilon)}{4\pi^3} \frac{-i\epsilon\boldsymbol{\gamma}\cdot\mathbf{q} - i\epsilon^2\gamma_4 v + \gamma_4 m(1 - \epsilon^2)^{\frac{1}{2}} + m}{q^2 + \epsilon^2 v^2 + 2im(1 - \epsilon^2)^{\frac{1}{2}}v + m^2} \times \int \frac{d^3\mathbf{q}'\, dv'}{(\mathbf{q} - \mathbf{q}')^2 + \epsilon^2(v - v')^2} \gamma_\mu \bar{f}_N(\mathbf{q}', v') \times \frac{-i\epsilon\boldsymbol{\gamma}\cdot\mathbf{q}' - i\epsilon^2\gamma_4 v' - \gamma_4 m(1 - \epsilon^2)^{\frac{1}{2}} + m}{q'^2 + \epsilon^2 v'^2 - 2im(1 - \epsilon^2)^{\frac{1}{2}}v' + m^2} \gamma_\mu, \quad (8)$$

where $\Lambda_N(\epsilon) = \alpha_N(\epsilon)/\epsilon$ and $\alpha_N(\epsilon)$ is the ϵ-dependent value of α associated with the solution $\bar{f}_N(\mathbf{q}, v)$. The possibility of assigning to the solutions an index N running over discrete values follows from the discreteness of the spectrum.

We write Eq. (8) symbolically:

$$\bar{f}_N = \Lambda_N(\epsilon)\bar{K}(\epsilon)\bar{f}_N.$$

It will be shown below that $\lim_{\epsilon\to 0} \bar{K}(\epsilon) = \bar{K}(0)$ in the uniform operator topology (i.e., in the norm), from which it follows[9] that $\lim_{\epsilon\to 0} \Lambda_N(\epsilon) = \Lambda_N^0$, where Λ_N^0 corresponds to the equation

$$\bar{f}_N^0 = \Lambda_N^0 \bar{K}(0)\bar{f}_N^0. \quad (9)$$

It is easy to verify however that this last equation is equivalent to the Schrödinger equation. Indeed (9) is obtained by setting $\epsilon = 0$ in (8),

$$\bar{f}_N^0(\mathbf{q}, v) = -\frac{m^2}{4\pi^3}\Lambda_N^0 \frac{1 + \gamma_4}{q^2 + m^2 + 2imv} \times \int \frac{d^3\mathbf{q}'\, dv'}{(\mathbf{q} - \mathbf{q}')^2} \gamma_\mu \bar{f}_0(\mathbf{q}', v') \frac{1 - \gamma_4}{q'^2 + m^2 - 2imv'} \gamma_\mu. \quad (10)$$

The general form of $\bar{f}^0$ is

$$\bar{f}^0 = \left[a' + \sum_{i=1}^{3} a_i\gamma_i + \sum_{i,j=1}^{3} a_{ij}\gamma_i\gamma_j + a\gamma_5\right]\frac{1 - \gamma_4}{2} + \left[b' + \sum_{i=1}^{3} b_i\gamma_i + \sum_{i,j=1}^{3} b_{ij}\gamma_i\gamma_j + b\gamma_5\right]\frac{1 + \gamma_4}{2}$$

where a', a_i, $\cdots$, b', b_i, $\cdots$ are multiples of the unit matrix. From

$$\frac{1 + \gamma_4}{2}\gamma_\mu\bar{f}^0\frac{1 + \gamma_4}{2} = -\left[\sum_{i=1}^{3} a_i\gamma_i + a\gamma_5\right]\frac{1 - \gamma_4}{2} + \left[2a' + 4\sum_{i=1}^{3} a_{ii} - \sum_{i,j=1}^{3} a_{ij}\gamma_i\gamma_j\right]\frac{1 + \gamma_4}{2}$$

and (10) we have that

$$\bar{f}^0 = \left[\sum_{i=1}^{3} a_i\gamma_i + a\gamma_5\right]\frac{1 - \gamma_4}{2}.$$

Substitution into (10) shows that a_1, a_2, a_3 and a all satisfy the equation (this corresponds to the absence of spin–orbit coupling in the nonrelativistic limit),

$$a(\mathbf{q}, v) = \frac{m^2}{\pi^3}\Lambda_N^0 \frac{1}{q^2 + m^2 + 2imv} \times \int \frac{d^3\mathbf{q}'\, dv'}{(\mathbf{q} - \mathbf{q}')^2} \frac{1}{q'^2 + m^2 - 2imv'} a(\mathbf{q}', v'). \quad (11)$$

Finally, setting

$$\tilde{\varphi}(\epsilon\mathbf{q}) = (\mathbf{q}^2 + m^2 + 2imv)(\mathbf{q}^2 + m^2)^{-1}a(\mathbf{q}, v)$$

[6] N. I. Akhiezer and I. M. Glazman, *Theory of Linear Operators in Hilbert Space* (Frederick Ungar Publishing Company, New York, 1961), p. 57.

[7] See G. Tiktopoulos, Phys. Rev. **133**, B 1231 (1964), Appendix. Definitions of analyticity of operators, etc. are also given in this reference.

[8] See N. Dunford and J. Schwarz, *Linear Operators* (Interscience Publishers, Inc., New York, 1958), Vol. I, Sec. VII 6.

[9] This follows, e.g., from the discussion in Ref. 8 by changing the word "analytic" to "continuous", appropriately.

we obtain

$$(\mathbf{p}^2 + m^2 - \tfrac{1}{4}E^2)\bar{\varphi}(\mathbf{p}) = \frac{m\alpha_N^0}{2\pi^2}\int \frac{d^3\mathbf{p}'}{(\mathbf{p}-\mathbf{p}')^2}\,\bar{\varphi}(\mathbf{p}'),$$

which is the momentum-space form of the Schrödinger equation for the Coulomb potential

$$\left[-\frac{1}{m}\nabla^2 - \frac{\alpha_N^0}{r} - \frac{E^2/4 - m^2}{m}\right]\varphi(\mathbf{r}) = 0.$$

Accordingly, the values of Λ_N^0 for which (10) is soluble are given by $\Lambda_n^0 = 2n$ where $n = 1, 2, \cdots$; we have thus shown that *as* $\epsilon \to 0$ *the set of values of* $\alpha(\epsilon)/\epsilon$ *for which Eq. (5) is soluble approaches those given by the nonrelativistic Coulomb spectrum*: $\lim_{\epsilon\to 0} \alpha(\epsilon)/\epsilon = 2n$; $n = 1, 2, \cdots, \infty$. We emphasize that this result does not exclude the existence of solutions for which $\lim_{\epsilon\to 0}\alpha(\epsilon)/\epsilon = \infty$ since ∞ is an accumulation point of the set $\{2n\}$. Wick's "abnormal" solutions[3] belong to this category. We have nothing to say in this paper about these possibilities.

It remains now to complete the proof by showing that

$$\lim_{\epsilon\to 0} |\bar{K}(\epsilon) - \bar{K}(0)| = 0.$$

We write $\bar{K}(\epsilon) - \bar{K}(0) = A(\epsilon) + R(\epsilon)$, where

$$\begin{aligned}\langle \mathbf{q}, v|\, A(\epsilon)\, |\mathbf{q}', v'\rangle &= -\frac{m^2}{4\pi^3}\frac{1+\gamma_4}{q^2 + m^2 + 2imv}\left\{\frac{1}{(\mathbf{q}-\mathbf{q}')^2}\right.\\ &\left. - \frac{1}{(\mathbf{q}-\mathbf{q}')^2 + \epsilon^2(v-v')^2}\right\}\gamma_\mu \cdots\\ &\cdots \frac{1}{q'^2 + m^2 - 2imv'}\,\gamma_\mu, \qquad (13)\end{aligned}$$

where the dots indicate the order of matrix operations. The application of inequalities (6) and (A2) gives, apart from constant factors,

$$\begin{aligned}|A(\epsilon)| &\le \sup_{(\mathbf{q},v)} \sigma^{-1}(\mathbf{q}, v)[(q^2 + m^2)^2 + 4m^2v^2]^{-\frac{1}{2}}\\ &\times \int\left[\frac{1}{(\mathbf{q}-\mathbf{q}')^2} - \frac{1}{(q-q')^2 + \epsilon^2(v-v')^2}\right]\\ &\times [(q'^2 + m^2)^2 + 4m^2v'^2]^{\frac{1}{2}}\sigma(\mathbf{q}', v')\, d^3\mathbf{q}'\, dv'.\end{aligned}$$

If we take

$$\sigma(\mathbf{q}, v) = \tau(v)[(q^2 + m^2)^2 + 4m^2v^2]^{-\frac{1}{2}},$$

we have

$$\begin{aligned}|A(\epsilon)| &\le \sup_{(v)}\frac{4\pi\epsilon^2}{\tau(v)}\int \frac{(v-v')^2}{q'^2 + \epsilon^2(v-v')^2}\\ &\times \frac{\tau(v')}{(q'^2 + m^2)^2 + 4m^2v'^2}\, dq'\, dv'\\ &\le \frac{4\pi\epsilon^{1-\gamma}}{c(\gamma)c(\delta)}\left(\int_0^\infty \frac{q^{\gamma-1}\, dq}{[(q^2 + m^2)^2 - \frac{1}{2}m^4]^{(1-\delta)/2}}\right)\\ &\times \left(\sup_{(v)}\frac{1}{\tau(v)}\int_{-\infty}^{\infty}\frac{|v-v'|^{1-\gamma}\,\tau(v')\, dv'}{[4m^2v'^2 + \frac{1}{2}m^4]^{\frac{1}{2}(1+\delta)}}\right).\end{aligned}$$

We have used here twice the inequality

$$X + Y \ge c(\gamma)X^\gamma Y^{1-\gamma}; \qquad 0 < \gamma < 1,$$

where $c(\gamma)$ is a positive number depending on γ. It now suffices to take $\tau(v) = (v^2 + m^2)^{\frac{1}{2}(1-\gamma)}$ and choose γ and δ to satisfy $\delta + 2\gamma > 2 > \gamma + 2\delta$ (e.g., $\gamma = 0.7$, $\delta = 0.61$) in order to obtain

$$|A(\epsilon)| < \epsilon^{1-\gamma} \times (\text{const independent of } \epsilon).$$

The application of inequalities (6) and (A1) leads by similar methods to $|R(\epsilon)| < \epsilon \times$ (const independent of ϵ). The derivation is straightforward and is not given here.

Concluding, it should be emphasized that the smallness ($\alpha < \frac{1}{2}$) of the fine-structure constant is essential for our proof of the discreteness of the spectrum. For larger values of the coupling constant one probably encounters the continuum part of the spectrum of K (possibly related to Goldstein's solutions) and the situation no longer makes sense at least in the conventional way of looking at the bound state problem in field theory. The natural way out of this difficulty would be to say that as the coupling becomes stronger one must include more and more graphs in the interaction kernel to maintain a reasonable approximation and eventually the perturbation expansion for the kernel might fail to converge. This, however, is the notorious weakness of the Bethe–Salpeter scheme and stays well beyond our present scope.

ACKNOWLEDGMENT

I wish to thank the members of the staff of the Physics Department at the Brookhaven National Laboratory for their warm hospitality. It is also a pleasure to thank Dr. R. Rockmore for an interesting conversation.

APPENDIX

In this paper extensive use was made of an inequality for the norm of integral operators. We give its derivation here.[10]

Consider the Hilbert space of square-integrable functions $f(x)$ over some measure μ.

[10] It is in fact related to the Hilbert inequality. A detailed discussion of the Hilbert inequality is given by G. H. Hardy, J. E. Littlewood, and G. Polya, *Inequalities* (Cambridge University Press, New York, 1959), Chap. IX.

$$|f| = \int |f(x)|^2 \, d\mu(x) < \infty .$$

Let $K(x, x')$ be the kernel of the integral operator K and φ and ψ any two functions of the Hilbert space with $|\varphi| = |\psi| = 1$. Then for *any* two positive (measurable) functions $\sigma_{1,2}(x)$ we have, via the Schwarz inequality,

$$\begin{aligned} |(\varphi, K\psi)| &= \left| \int \varphi^*(x) K(x, x') \psi(x') \, d\mu(x) \, d\mu(x') \right| \\ &\leq \left\{ \int |\varphi(x)|^2 \, |K(x, x')| \, \frac{\sigma_1(x')}{\sigma_2(x)} \, d\mu(x) \, d\mu(x') \right\}^{\frac{1}{2}} \\ &\quad \times \left\{ \int |\psi(y)|^2 \cdot |K(y, y')| \, \frac{\sigma_2(y)}{\sigma_1(y')} \, d\mu(y) \, d\mu(y') \right\}^{\frac{1}{2}}, \end{aligned}$$

so that the norm of K defined by

$$|K| = \sup_{(\varphi, \psi)} |(\varphi, K\psi)|; \qquad |\varphi| = |\psi| = 1$$

satisfies the inequality

$$\begin{aligned} |K| &\leq \left\{ \sup_{(x)} \frac{1}{\sigma_2(x)} \int |K(x, x')| \, \sigma_1(x') \, d\mu(x') \right\}^{\frac{1}{2}} \\ &\quad \times \left\{ \sup_{(y)} \frac{1}{\sigma_1(y)} \int |K(y', y)| \, \sigma_2(y') \, d\mu(y') \right\}^{\frac{1}{2}}. \end{aligned} \tag{A1}$$

In particular, if

$$|K(x, x')| = |K(x', x)|,$$

we have ($\sigma_1 = \sigma_2 = \sigma$)

$$|K| \leq \sup_{(x)} \frac{1}{\sigma(x)} \int |K(x, x')| \, \sigma(x') \, d\mu(x'). \tag{A2}$$

It should perhaps be emphasized that $\sigma(x)$ does *not* have to belong to the Hilbert space. Intuitively speaking, one obtains better upper bounds for $|K|$ the more $\sigma(x)$ approximates the eigenfunction corresponding to the biggest (in modulus) eigenvalue. In practice one chooses $\sigma_{1,2}(x)$ to depend on a number of parameters and such that the integrals in (A1) or (A2) are elementary. One then optimizes the upper bound by varying the parameters. It appears that quite crude choices for $\sigma_{1,2}(x)$ may produce useful upper bounds for $|K|$.

An easy generalization of (A1) is obtained by using a general (positive) function of two variables $\tau(x, x')$ instead of just $\sigma_1(x)/\sigma_2(x')$. The result is

$$\begin{aligned} |K| &\leq \left\{ \sup_{(x)} \int |K(x, x')| \, \tau(x, x') \, d\mu(x') \right\}^{\frac{1}{2}} \\ &\quad \times \left\{ \sup_{(y)} \int |K(y', y) \tau^{-1}(y', y) \, d\mu(y') \right\}^{\frac{1}{2}}. \end{aligned}$$

A further generalization to operators acting on L_p spaces is also possible.

Algebraic Construction of the Basis for the Irreducible Representations of Rotation Groups and for the Homogeneous Lorentz Group

J. A. CASTILHO ALCARÁS
AND
P. LEAL FERREIRA
Instituto de Física Teórica, São Paulo, Brazil
(Received 24 August 1964)

The basic functions for a class of the irreducible representations of the rotation groups in n dimensions (R_n) are explicitly constructed by an algebraic method in which the basic functions are taken to be homogeneous polynomials in the variables of E_n. The solutions correspond to the hyperspherical harmonics of the mathematical literature and are of interest for problems exhibiting invariance under a certain R_n. The method is also applied to derive a basis for the infinite-dimensional irreducible representations of the homogeneous Lorentz group if we then look for homogeneous functions in the variables of the corresponding pseudo-Euclidean space.

I. INTRODUCTION

THE aim of the present work is to construct explicitly the basic functions for a class of irreducible representations of the rotation groups in n dimensions, R_n. These functions correspond to the hyperspherical harmonics of the mathematical literature.[1] The treatment is algebraic and is an extension of the method for deriving the basic functions for the three-dimensional rotation group—the well-known spherical harmonics—in terms of homogeneous polynomials in the variables of the corresponding Euclidean space. The algebraic treatment is an alternative to the analytical treatment of the hyperspherical harmonics and follows the paradigm familiar to physicists as the quantum theory of angular momentum.

As it is also shown in this paper, the hyperspherical harmonics may be represented as polynomials in boson creation operators applied to a "vacuum state," in the spirit of the Schwinger's treatment of the angular momentum.[2]

The hyperspherical harmonics are relevant for problems exhibiting invariance under a certain R_n. A well-known example is the Fock–Bargmann treatment of the nonrelativistic hydrogen atom.[3] Another field of application is the theory of the Bethe–Salpeter equation supplemented by Wick's stability conditions, through which an Euclidean metric is introduced.[4] Finally, we mention that rotation groups appear as intermediate symmetry groups for the classification of atomic or nuclear configurations.[5]

The present approach may also be applied to study the representations of the pseudorotation groups, if we then look for homogeneous functions in the variables of the corresponding pseudo-Euclidean space. As an example, the infinite irreducible representations of the homogeneous Lorentz group are derived in a close parallelism to the treatment of the R_4 group which is a particular case of the above mentioned general treatment.

II. THE CONSTRUCTION OF THE BASIS OF THE R_n GROUP

A.

The rotation group in n dimensions (R_n) is the group of all linear real transformations with determinant $+1$ that leave invariant the quadratic form

$$\mathbf{x}^2 = \sum_{\alpha=1}^{n} x_\alpha^2 = \sum_{\alpha,\beta=1}^{n} g_{\alpha\beta} x_\alpha x_\beta, \quad (g_{\alpha\beta} = \delta_{\alpha\beta}). \tag{2.1}$$

The group is generated by the $n(n-1)/2$ infinitesimal rotation operators $D_{\alpha\beta} = -D_{\beta\alpha}$ which obey the commutation relation

$$[D_{\alpha\beta}, D_{\gamma\delta}] = i(\delta_{\gamma\alpha} D_{\beta\delta} + \delta_{\alpha\delta} D_{\gamma\beta} - \delta_{\gamma\beta} D_{\alpha\delta} - \delta_{\delta\beta} D_{\gamma\alpha}). \tag{2.2}$$

A set of such operators $D_{\alpha\beta}$ is obtained by putting

$$D_{\alpha\beta} = x_\alpha p_\beta - x_\beta p_\alpha, \quad (\alpha, \beta = 1, 2, \cdots, n) \tag{2.3}$$

with

$$[x_\alpha, p_\beta] = i\delta_{\alpha\beta}, \qquad p_\alpha = -i\partial/\partial x_\alpha. \tag{2.4}$$

[1] A. Erdelyi, W. Magnus, F. Oberhettinger, and F. G. Tricomi, *Bateman Manuscript Project* (Addison-Wesley Publishing Company, Reading, Massachusetts, 1962), Chap. 9.
[2] J. Schwinger, "On Angular Momentum," NYO-3071 (Office of Technical Services, Department of Commerce, Washington, D. C., 1952).
[3] V. Fock, Z. Physic 98, 145 (1935); V. Bargmann, Z. Physic 99, 576 (1936).
[4] G. C. Wick, Phys. Rev. 94, 1124 (1954).
[5] M. Hamermesh, *Group theory and Its Application to Physical Problems* (Addison-Wesley Publishing Company, Inc., Reading, Massachusetts, 1962).

The group R_n has $[n/2]$ Casimir invariants, the rank of the group R_n.[6] In this work we restrict ourselves to a single-row representation of R_n, that is, one characterized by the Casimir invariant

$$I_1^{(n)} = \frac{1}{2}\sum_{\alpha,\beta=1}^{n} D_{\alpha\beta}^2. \tag{2.5}$$

The Casimir invariant $I_1^{(n)}$ may be written as

$$I_1^{(n)} = \mathbf{x}^2\mathbf{p}^2 - (\mathbf{x}\cdot\mathbf{p})^2 + i(n-2)\mathbf{x}\cdot\mathbf{p}, \tag{2.6}$$

n being the dimension of the Euclidean space E_n where the inner products appearing in (2.6) are defined.

Since the operator $I_1^{(n)}$ conserves the degree of the polynomials, we will look for homogeneous-polynomials solutions of the eigenvalue equation

$$I_1^{(n)}P_\lambda(x_1, x_2, \cdots, x_n) = \lambda P_\lambda(x_1, x_2, \cdots, x_n). \tag{2.7}$$

The homogeneity of $P_\lambda(x_1, x_2, \cdots, x_n)$ together with Eq. (2.6) implies that they must be harmonic and $\lambda = k(k+n-2)$ where k is a nonnegative integer which designates the degree of homogeneity of $P_\lambda(x_1, x_2, \cdots, x_n)$. Then Eq. (2.7) is equivalent to

$$\nabla_n^2 P_\lambda(x_1, x_2, \cdots, x_n) = 0$$

and

$$\lambda = k(k+n-2), \qquad (k = 0, 1, 2, \cdots). \tag{2.8}$$

Since the operator ∇_n^2 is invariant under n-dimensional rotations, and such rotations conserve the homogeneity and degree of polynomials, we see that if $P_\lambda(x_1, x_2, \cdots, x_n)$ is an eigenfunction of $I_1^{(n)}$ with eigenvalue λ, after "rotated" it is still an eigenfunction of $I_1^{(n)}$ with the same eigenvalue λ. So, the $D_n(\lambda)$ solutions of Eq. (2.7) form a $D_n(\lambda)$-dimensional representation of the group R_n, which is irreducible since $I_1^{(n)}$ is a Casimir invariant of R_n.

For classifying the $D_n(\lambda)$ solutions of Eq. (2.7) we consider the chain of subgroups of R_n

$$R_n \supset R_{n-1} \supset \cdots \supset R_3 \tag{2.9}$$

and to each of its members we attach the corresponding Casimir invariant

$$I_1^{(n)}, I_1^{(n-1)}, \cdots, I_1^{(3)}, \tag{2.10}$$

where

$$I_1^{(p)} = \frac{1}{2}\sum_{\alpha,\beta=1}^{p} D_{\alpha\beta}^2, \qquad (p = 3, 4, \cdots, n). \tag{2.11}$$

The basic functions are the homogeneous polynomials in the variables $x_1, x_2, \cdots, x_n$, labeled by the numbers $k_n, k_{n-1}, \cdots, k_3, m$, solutions of the system

$$\begin{aligned} I_1^{(p)}P_{k_n,k_{n-1},\cdots,k_3,m} &= k_p(k_p+p-2)P_{k_n,k_{n-1},\cdots,k_3,m}, \\ &\qquad (p = 3, 4, \cdots, n), \\ L_0P_{k_n,k_{n-1},\cdots,k_3,m} &= mP_{k_n,k_{n-1},\cdots,k_3,m}, \end{aligned} \tag{2.12}$$

where

$$L_0 = D_{12} = x_1p_2 - x_2p_1. \tag{2.13}$$

Furthermore, it is easily seen that the set of nonnegative integers $k_n, k_{n-1}, \cdots, k_3, |m|$ satisfies the branching rule

$$k_n \geq k_{n-1} \geq \cdots \geq k_3 \geq |m| \geq 0.$$

In fact, $I_1^{(p)} = I_1^{(p-1)} + \sum_{\alpha=1}^{p-1} D_{\alpha p}^2$. Since the $D_{\alpha\beta}$ are Hermitian, the eigenvalues of $\sum_{\alpha=1}^{p-1} D_{\alpha p}^2$ are nonnegative. Therefore

$$k_p(k_p+p-2) \geq k_{p-1}(k_{p-1}+p-3)$$

or $k_p \geq k_{p-1}$ for $p = 4, 5, \cdots, n$. The last inequality $l \geq |m| \geq 0$ is well known from the study of the R_3 group.

For constructing the solutions $P_{k_n,k_{n-1},\cdots,k_3,m}$ of Eq. (2.12), we try to relate them to the corresponding normalized polynomials of the R_{n-1} group, i.e., the solutions of the system (2.12) with $n-1$ instead of n. We then put

$$\begin{aligned} &P_{k_n,k_{n-1},\cdots,k_3,m}(x_1, x_2, \cdots, x_n) \\ &= C_0(k_n, k_{n-1})G_{k_n,k_{n-1}}(x_1, x_2, \cdots, x_n) \\ &\times P_{k_{n-1},k_{n-2},\cdots,k_3,m}(x_1, x_2, \cdots, x_{n-1}) \end{aligned} \tag{2.15}$$

where $C_0(k_n, k_{n-1})$ is a normalization constant.

Since the $P_{k_{n-1},k_{n-2},\cdots,k_3,m}(x_1, x_2, \cdots, x_{n-1})$ satisfy the system (2.12) for p up to $n-1$, we have to do only with

$$\begin{aligned} &I_1^{(n)}P_{k_n,k_{n-1},\cdots,k_3,m}(x_1, x_2, \cdots, x_n) \\ &= k_n(k_n+n-2)P_{k_n,k_{n-1},\cdots,k_3,m}(x_1, x_2, \cdots, x_n), \end{aligned} \tag{2.16}$$

if we take for $G_{k_n,k_{n-1}}(x_1, x_2, \cdots, x_n)$ an invariant of all the subgroups of R_n appearing in Eq. (2.9). On the other hand, since $P_{k_n,k_{n-1},\cdots,k_3,m}$ and $P_{k_{n-1},k_{n-2},\cdots,k_3,m}$ are homogeneous of degree k_n and k_{n-1}, $G_{k_n,k_{n-1}}(x_1, x_2, \cdots x_n)$ is homogeneous of degree $k_n - k_{n-1}$. Hence we put

$$\begin{aligned} G_{k_n,k_{n-1}}(x_1, x_2, \cdots, x_n) &= G_{k_n,k_{n-1}}(R_n, x_n) \\ &= \sum_{\mu=0}^{k_n-k_{n-1}} a_\mu R_n^\mu x_n^{k_n-k_{n-1}-\mu}, \end{aligned} \tag{2.17}$$

[6] L. C. Biedenharn, J. Math. Phys. 4, 436 (1963).

where

$$R_n = \left(\sum_{\alpha=1}^{n} x_\alpha^2\right)^{\frac{1}{2}}.$$

As pointed out before, Eq. (2.16) is equivalent to

$$\nabla_n^2 P_{k_n,k_{n-1},\cdots,k_3,m}(x_1, x_2, \cdots, x_n)$$
$$= C_0(k_n, k_{n-1})\nabla_n^2[G_{k_n,k_{n-1}}(R_n, x_n)$$
$$\times P_{k_{n-1},k_{n-2},\cdots,k_3,m}(x_1, x_2, \cdots, x_{n-1})] = 0, \quad (2.18)$$

which in view of Eq. (2.15) and the harmonicity of $P_{k_{n-1},k_{n-2},\cdots,k_3,m}(x_1, x_2, \cdots, x_{n-1})$ becomes

$$2k_{n-1}\sum_{\mu=0}^{k_n-k_{n-1}} \mu a_\mu R_n^{\mu-2} x_n^{k_n-k_{n-1}-\mu} + \nabla_n^2 G_{k_n,k_{n-1}} = 0. \quad (2.19)$$

This equation provides the recursion relations

$$a_1 = 0,$$
$$\mu(2k_n + n - 2 - \mu)a_\mu + (k_n - k_{n-1} - \mu + 2)(k_n - k_{n-1} - \mu + 1)a_{\mu-2} = 0, \quad (\mu = 2, 3, \cdots) \quad (2.20)$$

which allow us to write

$$a_{2\mu+1} = 0, \quad (2.21)$$
$$a_{2\mu} = \frac{(-)^\mu \Gamma(n/2 + k_n - \mu - 1)(k_n - k_{n-1})!\, a_0}{2^{2\mu}\mu!\, \Gamma(n/2 - 1 + k_n)(k_n - k_{n-1} - 2\mu)!} \quad (\mu = 0, 1, \cdots, [(k_n - k_{n-1})/2]).$$

Putting $a_0 = \Gamma(k_n + n/2 - 1)/(k_n - k_{n-1})!$, we have for $G_{k_n,k_{n-1}}(R_n, x_n)$ the expression

$$G_{k_n,k_{n-1}}(R_n, x_n) = \sum_{\mu=0}^{[(k_n-k_{n-1})/2]} \frac{(-)^\mu \Gamma(k_n+n/2-\mu-1)R_n^{2\mu}x_n^{k_n-k_{n-1}-2\mu}}{2^{2\mu}\mu!\,(k_n-k_{n-1}-2\mu)!}. \quad (2.22)$$

The function $G_{k_n,k_{n-1}}(R_n, x_n)$ has the following noticeable properties [the $C_{k_n-k_{n-1}}^{k_{n-1}+n/2-1}(x_n/R_n)$ appearing in property (2.23a) are Gegenbauer polynomials[7] which have also analogs to properties (2.23b) and (2.23e)]:

$$G_{k_n,k_{n-1}}(R_n, x_n) = [\Gamma(n/2 + k_{n-1} - 1)/2^{k_n-k_{n-1}}] \times R_n^{k_n-k_{n-1}} C_{k_n-k_{n-1}}^{k_{n-1}+n/2-1}(x_n/R_n), \quad (2.23a)$$

$$x_n G_{k_n,k_{n-1}}(R_n, x_n) = 2G_{k_n+1,k_{n-1}}(R_n, x_n) - (k_n + k_{n-1} + n - 3)G_{k_n,k_{n-1}-1}(R_n, x_n), \quad (2.23b)$$

$$R_n^2 G_{k_n,k_{n-1}}(R_n, x_n) = 4[G_{k_n+2,k_{n-1}}(R_n, x_n) - (k_n + n/2)G_{k_n+1,k_{n-1}-1}(R_n, x_n)], \quad (2.23c)$$

$$\partial G_{k_n,k_{n-1}}(R_n, x_n)/\partial x_i = -\tfrac{1}{2}x_i G_{k_n-1,k_{n-1}+1}(R_n, x_n) \quad (i < n), \quad (2.23d)$$

$$\partial G_{k_n,k_{n-1}}(R_n, x_n)/\partial x_n = G_{k_n,k_{n-1}}(R_n, x_n) - \tfrac{1}{2}x_n G_{k_n-1,k_{n-1}+1}(R_n, x_n) = \tfrac{1}{2}(k_n + k_{n-1} + n - 3)G_{k_n-1,k_{n-1}}(R_n, x_n). \quad (2.23e)$$

If we now express all the polynomials

$$P_{k_p,k_{p-1},\cdots,k_3,m} \qquad (p = 4, 5, \cdots, n)$$

in the form (2.15) and observe that [from Eq. (2.24)] we see that the present treatment holds also for $n=3$ if we put

$$P_{k_3,m} \equiv \mathcal{Y}_{k_3,m} = C_0(k_3, |m|)P_m(x_1, x_2)G_{k_3,|m|}(R_3, x_3),$$

with

$$P_m(x_1, x_2) = (2\pi)^{-\frac{1}{2}}(x_1 \pm ix_2)^{|m|},$$
$$P_{k_3,m} = \mathcal{Y}_{k_3,m} = 2^{k_3}(-)^{(m-|m|)/2}\left[\frac{(2k_3+1)(k_3-|m|)!}{2\pi(k_3+|m|)!}\right]^{\frac{1}{2}} \times \frac{(x_1 \pm ix_2)^{|m|}}{(2\pi)^{\frac{1}{2}}} G_{k_3,|m|}(R_3, x_3) = C_0(k_3, |m|) \times \frac{(x_1 \pm ix_2)^{|m|}}{(2\pi)^{\frac{1}{2}}} G_{k_3,|m|}(R_3, x_3), \quad (2.24)$$

it follows that

$$P_{k_n,k_{n-1},\cdots,k_3,m}(x_1, x_2, \cdots, x_n) = (2\pi)^{-\frac{1}{2}}(x_1 \pm ix_2)^{|m|} \times \prod_{\mu=3}^{n} C_0(k_\mu, k_{\mu-1})G_{k_\mu,k_{\mu-1}}(R_\mu, x_\mu), \quad (k_2 = |m|) \quad (2.25)$$

where the $G_{k_\mu,k_{\mu-1}}(R_\mu, x_\mu)$ are given by Eq. (2.22). Normalization to one over the unit hypersphere yields

$$C_0(k_\mu, k_{\mu-1}) = 2^{k_\mu+(\mu-4)/2}\left[\frac{2(k_\mu + \mu/2 - 1)(k_\mu - k_{\mu-1})!}{\pi(k_\mu + k_{\mu-1} + \mu - 3)!}\right]^{\frac{1}{2}} \quad (\mu = 3, 4, \cdots, n). \quad (2.26)$$

B.

An important particular case of Eq. (2.25), which is relevant in the Fock–Bargmann treatment of the bound states of a charged particle in a Coulomb field, is obtained for $n = 4$. The scattering states, however, are related to the homogeneous Lorentz group. The basic functions of the R_4 group are briefly discussed here in a way that makes easier

[7] E. D. Rainville, *Special Functions* (The Macmillan Company, New York, 1960).

the construction of the corresponding functions for the homogeneous Lorentz group (Sec. III).

For $n = 4$, we have to do with the simple chain $R_4 \supset R_3$, and the equations (2.12) reduce to

$$\begin{aligned} I_1^{(4)} P_{klm} &= I_1 P_{klm} = k(k+2) P_{klm}, \\ I_1^{(3)} P_{klm} &= L^2 P_{klm} = l(l+1) P_{klm}, \\ L_0 P_{klm} &= m P_{klm}, \end{aligned} \tag{2.27}$$

where $k = 0, 1, \cdots; l = 0, 1, \cdots, k; m = -l, -l+1, \cdots, +l$.

As pointed out before, the first of Eqs. (2.27) is equivalent to

$$\nabla_4^2 P_{klm} = 0.$$

The polynomials (2.25) become

$$\begin{aligned} &P_{klm}(x_1, x_2, x_3, x_4) \\ &= C_0(k, l) \mathcal{Y}_{lm}(x_1, x_2, x_3) G_{k,l}(R_4, x_4) \\ &= 2^k \left[\frac{2(k+1)(k-l)!}{\pi (k+l+1)!} \right]^{\frac{1}{2}} \mathcal{Y}_{l,m}(x_1, x_2, x_3) \\ &\quad \times \sum_{\mu=0}^{[(k-l)/2]} \frac{(-)^\mu (k-\mu)!\, R_4^{2\mu} x_4^{k-l-2\mu}}{2^{2\mu} \mu!\, (k-l-2\mu)!}. \end{aligned} \tag{2.28}$$

The functions (2.28) can be written in a more familiar form in terms of the polar coordinates of E_4, which can be defined as

$$\begin{aligned} x_1 &= R_4 \sin\lambda \sin\theta \cos\phi, \\ x_2 &= R_4 \sin\lambda \sin\theta \sin\phi, \\ x_3 &= R_4 \sin\lambda \cos\theta, \qquad x_4 = R_4 \cos\lambda. \end{aligned} \tag{2.29}$$

From Eqs. (2.29) and making use of

$$\begin{aligned} &\cos(k+1)\lambda \\ &= 2^k (k+1) \sum_{\mu=0}^{[(k+1)/2]} \frac{(-)^\mu (k-\mu)! \cos^{k+1-2\mu}\lambda}{2^{2\mu} \mu!\, (k+1-2\mu)!} \\ &\qquad (k = 0, 1, 2, \cdots), \end{aligned} \tag{2.30}$$

we get

$$\begin{aligned} &P_{klm}(R_\mu, \theta, \phi, \lambda) \\ &= \frac{R_\mu^k \mathcal{Y}_{lm}(\theta, \phi)}{T_{k+1,l}} \sin^l\lambda \frac{d^{l+1} \cos(k+1)\lambda}{d(\cos\lambda)^{l+1}}, \end{aligned} \tag{2.31}$$

where

$$T_{k,l}^2 = \frac{\pi}{2} \frac{k(k+l)!}{(k-l-1)!} = \frac{\pi}{2} k^2 (k^2 - 1^2)(k^2 - 2^2) \cdots (k^2 - l^2). \tag{2.32}$$

[The function (2.31), for $R_4 = 1$, was firstly obtained by Fock[3] and analytically derived by Kursunoglu.[8]]

C.

We will now show that the basic functions for the R_n group have in the boson representation the same expression (2.25).

We define the boson creation and destruction operators and a vacuum $|0\rangle$ by

$$\begin{aligned} \mathbf{a}^\dagger &= 2^{-\frac{1}{2}}(\mathbf{x} - i\mathbf{p}), \\ \mathbf{a} &= 2^{-\frac{1}{2}}(\mathbf{x} + i\mathbf{p}), \\ a_i |0\rangle &= 0, \qquad (i = 1, 2, \cdots, n). \end{aligned} \tag{2.33}$$

These operators obey the commutation relations

$$\begin{aligned} &[a_i, a_j^\dagger] = \delta_{ij}; \qquad [a_i, a_j] = [a_i^\dagger, a_j^\dagger] = 0 \\ &\qquad (i, j = 1, 2, \cdots, n). \end{aligned} \tag{2.34}$$

Expressing $I_1^{(q)}$ and L_0 in terms of $\mathbf{a}^\dagger$ and $\mathbf{a}$ we get

$$\begin{aligned} I_1^{(q)} &= -\mathbf{a}^{\dagger 2} \mathbf{a}^2 + (\mathbf{a}^\dagger \cdot \mathbf{a})^2 + (q-2)\mathbf{a}^\dagger \cdot \mathbf{a} \\ &\qquad (q = 3, 4, \cdots, n), \\ L_0 &= -i(a_1^\dagger a_2 - a_2^\dagger a_1). \end{aligned} \tag{2.35}$$

Since the commutation relations (2.34) allow us to interpret the a_i, when applied to a polynomial $P(\mathbf{a}^\dagger)|0\rangle$ as $a_i = \partial/\partial a_i^\dagger$, we see that the operators $I_1^{(q)}$ and L_0 have the same expression in terms of $x_1, x_2, \cdots, x_n$ or $a_1^\dagger, a_2^\dagger, \cdots, a_n^\dagger$. Therefore, the same occurs with the basic polynomials $P_{k_n, k_{n-1}, \cdots, k_3, m}$. The solution corresponding to (2.25), normalized in the sense that

$$\langle 0 | P^\dagger_{k_n, k_{n-1}, \cdots, k_3, m} P_{k_n, k_{n-1}, \cdots, k_3, m} |0\rangle = 1 \tag{2.36}$$

is

$$\begin{aligned} P_{k_n, k_{n-1}, \cdots, k_3, m} |0\rangle &= (2\pi)^{-\frac{1}{2}} (a_1^\dagger \pm i a_2^\dagger)^{|m|} \\ &\quad \times \prod_{\mu=3}^{n} C_0'(k_\mu, k_{\mu-1}) G_{k_\mu, k_{\mu-1}}(A_\mu^\dagger, a_\mu^\dagger) |0\rangle \end{aligned} \tag{2.37}$$

with

$$\begin{aligned} A_\mu^\dagger &= \left[\sum_{i=1}^{\mu} (a_i^\dagger)^2 \right]^{\frac{1}{2}}, \\ C_0'(k_\mu, k_{\mu-1}) &= \left[\frac{2^{k_\mu - k_{\mu-1}} (2k_{\mu-1} + \mu - 3)!\, (k_\mu - k_{\mu-1})!}{\Gamma(k_\mu - 1 + \mu/2)\Gamma(k_{\mu-1} - 1 + \mu/2)(k_\mu + k_{\mu-1} + \mu - 3)!} \right]^{\frac{1}{2}}. \end{aligned} \tag{2.38}$$

[8] B. Kursunoglu, *Modern Quantum Theory* (W. H. Freeman and Company, San Francisco, 1962).

The polynomials (2.37) have been used by two of the authors in the classification of the states of an n-dimensional isotropic harmonic oscillator,[9] in connection with nuclear vibrations.

III. THE INFINITE-DIMENSIONAL BASIS FOR REPRESENTATIONS OF THE HOMOGENEOUS LORENTZ GROUP

The present method may be also applied to pseudo-rotation groups in n dimensions. In the present section, as an example, we briefly discuss the case of the homogeneous Lorentz group, in view of its intrinsic interest.

The homogeneous Lorentz group, as it is well known, is generated by the six infinitesimal operators $J_{\mu\nu} = -J_{\nu\mu}$, obeying

$$[J_{\alpha\beta}, J_{\gamma\delta}] = i(g_{\gamma\beta}J_{\alpha\delta} - g_{\alpha\delta}J_{\gamma\beta} + g_{\beta\delta}J_{\gamma\alpha} - g_{\gamma\alpha}J_{\beta\delta}), \quad (3.1)$$

where the metric tensor $g_{\alpha\beta}$ is taken with $g_{44} = -g_{11} = -g_{22} = -g_{33} = 1$, all the other components being zero.

A particular set of those operators is obtained by putting

$$J_{\mu\nu} = x_\mu p_\nu - x_\nu p_\mu \quad (3.2)$$

with

$$[x_\mu, p_\nu] = ig_{\mu\nu}. \quad (3.3)$$

We shall consider the representation characterized by the Casimir invariant

$$I_1 = \tfrac{1}{2}J_{\alpha\beta}J^{\alpha\beta}. \quad (3.4)$$

In analogy with the previous treatment of the R_4 group, the basic functions $\psi_{\lambda lm}$ for the homogeneous Lorentz group satisfy equations formally identical to Eq. (2.27), where now $L_i = \frac{1}{2}\epsilon_{ijk}J^{jk}$. As the L_i are Hermitian and obey the usual commutation relations for angular momentum operators, it follows that $l = 0, 1, 2, \cdots$, and $m = -l, -l+1, \cdots, +l$.

Now the $\psi_{\lambda lm}$ are taken to be homogeneous functions of the (real) variables x_1, x_2, x_3, x_4.

The Casimir invariant I_1 may be written in the form of Eq. (2.6) where now, of course, the inner products refer to the pseudo-Euclidean space with metric tensor $g_{\alpha\beta}$. It follows then easily that $\psi_{\lambda lm}$ satisfy the d'Alembertian equation and the eigenvalues of I_1 are of the form $\lambda = k(k+2)$, where k is the degree of homogeneity of the $\psi_{\lambda lm}$, which is complex in general. On the other hand, I_1 is a Hermitian operator, so λ must be real, giving to k the two possibilities

$$k = \nu - 1; \quad k = i\nu - 1 \qquad (\nu \text{ real}), \quad (3.5)$$

By the same reasoning made in Sec. IIA, we see that the $\psi_{\lambda lm}$ form a basis for an irreducible representation of the homogeneous Lorentz group. This representation is usually infinite, because in general for a given value of k, all the values of l are allowed.

We see that our problem is very similar to that of the R_4 group. In fact, L^2 and L_3 have here the same formal expression as in that case. Besides, the d'Alembertian operator acting on $\psi_{\lambda lm}(x_1, x_2, x_3, x_4)$ gives the same result as ∇_4^2 gave acting on $P_{klm}(x_1, x_2, x_3, x_4)$, if we replace in $\psi_{klm}(x_1, x_2, x_3, x_4)$

$$x_j \to ix_j \quad (j = 1, 2, 3), \qquad x_4 \to x_4. \quad (3.6)$$

In complete analogy to the treatment of Sec. IB, we get

$$\psi_{klm}(x_1, x_2, x_3, x_4) = i^l \mathcal{Y}_{lm}(x_1, x_2, x_3) C_0(k, l) \times \sum_{\mu=0}^{\infty} \frac{(-)^\mu \Gamma(k-\mu+1) R^{2\mu} x_4^{k-l-2\mu}}{2^{2\mu}\mu!\,\Gamma(k-l-2\mu+1)}, \quad (3.7)$$

where

$$R = (x_4^2 - x_1^2 - x_2^2 - x_3^2)^{\frac{1}{2}}.$$

We wish now to find an equivalent expression for (3.7) in terms of the polar coordinates $(R, \theta, \phi, \lambda)$ of the pseudo-Euclidean space with the metric tensor $g_{\alpha\beta}$ here considered. The following parametrization is adopted[8]:

$$\begin{aligned} x_1 &= r\sin\theta\cos\phi, & 0 \le R < \infty, \\ x_2 &= r\sin\theta\sin\phi, & 0 \le \lambda < \infty, \\ x_3 &= r\cos\theta, & 0 \le \theta \le \pi, \\ x_4 &= \begin{cases} R\cosh\lambda \\ R\sinh\lambda \end{cases}, \quad r = \begin{cases} R\sinh\lambda \\ R\cosh\lambda \end{cases} & 0 \le \phi \le 2\pi, \end{aligned} \quad (3.8)$$

where the alternatives $(x_4 = R\cosh\lambda, r = R\sinh\lambda)$ and $(x_4 = R\sinh\lambda, r = R\cosh\lambda)$ correspond to timelike and spacelike intervals, respectively.

For the timelike case $(x_4^2 - r^2 > 0)$, we get

$$\psi_{klm} = i^l C_0(k, l) R^k \mathcal{Y}_{lm}(\theta, \phi) \sinh^l\lambda \times \sum_{\mu=0}^{\infty} \frac{(-)^\mu \Gamma(k-\mu+1)\cosh^{k-l-2\mu}\lambda}{2^{2\mu}\mu!\,\Gamma(k-l-2\mu+1)}. \quad (3.9)$$

Making use of the relation

$$\cos(k+1)i\lambda = (k+1)2^{k+1}\sum_{\mu=0}^{\infty}\frac{(-)^\mu\Gamma(k+1-\mu)\cosh^{k+1-2\mu}\lambda}{2^{2\mu}\mu!\,\Gamma(k+2-2\mu)} \qquad (k+1 \neq -1, -2, \cdots), \quad (3.10)$$

[9] P. Leal Ferreira, J. A. Castilho Alcarás, and V. C. Aguilera Navarro, Phys. Rev. **136**, B1243 (1964).

which is easily gotten by power series method, we finally have

$$\psi_{klm} = \frac{i^l C_0(k,l)}{(k+1)2^{k+1}} R^k Y_{lm}(\theta,\phi) \times \sinh^l\lambda \frac{d^{l+1}\cos(k+1)i\lambda}{d(\cosh\lambda)^{l+1}}. \quad (3.11)$$

From Eq. (3.5), we have two kinds of solutions

(1) $k = i\nu - 1$ (ν real),

$$\psi^{(1)}_{\nu lm} = \frac{i^l C_0^{(1)}(\nu,l)}{i\nu 2^{i\nu}} R^{i\nu-1} Y_{lm}(\theta,\phi)\sinh^l\lambda \frac{d^{l+1}\cos\nu\lambda}{d(\cosh\lambda)^{l+1}}, \quad (3.12)$$

(2) $k = \nu - 1$ (ν real),

$$\psi^{(2)}_{\nu lm} = \frac{i^l C_0^{(2)}(\nu,l)}{\nu 2^{\nu}} R^{\nu-1} Y_{lm}(\theta,\phi)\sinh^l\lambda \frac{d^{l+1}\cosh\nu\lambda}{d(\cosh\lambda)^{l+1}}. \quad (3.13)$$

For the spacelike interval, the substitution (3.6) would be

$$x_j \to x_j \quad (j = 1, 2, 3), \qquad x_4 \to ix_4, \quad (3.14)$$

giving for ψ_{klm}, apart a phase factor, the results of replacing in Eq. (3.12) and Eq. (3.13)

$$\lambda \to \lambda + i\pi/2; \qquad R \to iR. \quad (3.15)$$

The present approach has the interesting feature that it gives rise, in a natural way, to two kinds of basic functions $\psi^{(i)}_{\nu lm}$ ($i = 1, 2$) for the Lorentz group representation, according to the alternatives contained in Eq. (3.5). These representations are infinite dimensional, since for a given value of k, no restriction on l exists, i.e., $l = 0, 1, 2, \cdots, \infty$.

The basic functions $\psi^{(1)}_{\nu lm}$ correspond to a particular case of what is known as the principal series representation of the homogeneous Lorentz group.[10] Using for the constant $C_0^{(1)}(\nu, l)$ the value

$$C_0^{(1)}(\nu,l) = i^{1-l}2^{i\nu} \times [(\pi/2)(\nu^2+1^2)(\nu^2+2^2)\cdots(\nu^2+l^2)]^{-\frac{1}{2}}, \quad (3.16)$$

we have the orthonormality relation[11]

$$\iiint \psi^{(1)*}_{\nu lm}(1,\theta,\phi,\lambda)\psi^{(1)}_{\nu' l' m'}(1,\theta,\phi,\lambda)\sinh^2\lambda \times \sin\theta\, d\theta\, d\phi\, d\lambda = \delta_{ll'}\,\delta_{mm'}\,\delta(\nu-\nu'), \quad (3.17)$$

if we restrict the real variable ν to nonnegative values. The $\psi^{(1)}_{\nu lm}$ may be useful for describing the continuum of positive energy states of a charged particle in a Coulomb field.

In contrast, the functions $\psi^{(2)}_{\nu lm}$ are not orthogonal and for this reason they lose part of their interest. On the other hand, if we restrict the real variable ν to nonnegative integers, as considered by Dolginov[11] they form a basis for a finite-dimensional representation of the homogeneous Lorentz group. In fact, in this case the series (3.10) becomes a polynomial, and for $l \geq \nu$

$$d^{l+1}\cosh\nu\lambda/d(\cosh\lambda)^{l+1} = 0. \quad (3.18)$$

Therefore for a given ν nonnegative integer, l can assume only the values $l = 0, 1, 2, \cdots, \nu$.

Finally, we notice that the $\psi^{(i)}_{\nu lm}$ may be put in terms of Gegenbauer functions,[1] which are extensions of the Gegenbauer polynomials $C_p^q(\cos\lambda)$ for noninteger values of p.

ACKNOWLEDGMENTS

We are much indebted to Professor M. Moshinsky, of the Universidad de México, whose lectures and discussions in the V Latin School of Physics, held in Rio de Janeiro, July 1963, provided the necessary basis for the present work.

We are also pleased to thank V. C. Aguilera Navarro and H. V. Fagundes for helpful discussions.

Finally, one of the authors (J. A. Castilho Alcarás) is grateful to the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) for support.

[10] W. Pauli, *Continuous Groups in Quantum Mechanics* (CERN, Geneva, 1956).

[11] A. Z. Dolginov, Zh. Eksperim. i Teor. Fiz **30**, 746 (1956) [English transl.: Soviet Phys. **3**, 589 (1956)]; A. Z. Dolginov and I. N. Toptygin, *ibid.* **37**, 1441 (1959) [English transl.: *ibid.*, **10**, 1022 (1960)]; A. Z. Dolginov and A. N. Moscalev *ibid.*, p. 1697 [English transl.: *ibid.*, p. 1202].

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 6, NUMBER 4 APRIL 1965

On a Special Class of Type-I Gravitational Fields*

MANFRED TRÜMPER
Yeshiva University, New York, New York†
(Received 9 July 1964)

This paper contains an investigation of the algebraic structure and the analytic properties of a class of normal hyperbolic Riemannian 4-spaces restricted by the following condition: There exists a timelike unit vector u^a such that the Riemann tensor satisfies ${}^*R_{abcd}u^bu^d = 0$. This condition is shown to be equivalent to the statement that the conform tensor is Petrov type I with real eigenvalues, u^a being a principal vector and an eigenvector of the Ricci tensor. This means that there is no flux of nongravitational energy relative to an observer travelling with 4-velocity u^a.

The eigen null directions (Debever vectors) of the conform tensor lie in a timelike hyperplane spanned by u^a and the two eigenvectors of $\epsilon_{ac} \equiv -C_{abcd}u^bu^d$ belonging to the eigenvalues with largest absolute value. The conform tensor is degenerate (type D) if and only if a Debever direction projected into the rest space of an observer u^a is an eigendirection of ϵ_{ab}.

The complete set of Bianchi identities is examined. It yields an expression for the covariant eigentime derivative of ϵ_{ab} and an algebraic relation linking the rotation and shear of u^a to the curvature tensor of the Riemannian space.

The general results are applied to special Einstein spaces ($R_{ab} = 0$) admitting a congruence of timelike curves without shear and rotation. We get a new simplified proof of the theorem that in the case of nondegeneracy such spaces are static, the curves of the congruence being paths of an isometric motion.

INTRODUCTION

WE study the class of Riemannian 4-spaces which admit a congruence of timelike curves with unit tangent u^a satisfying

$$ {}^*R_{abcd}u^bu^d = 0. \tag{1}$$

The condition (1) means that there exists an observer with 4-velocity u^a to whom the structure of the curvature tensor appears to be particularly simple. The fields defined by (1) correspond formally to Maxwell fields of the electric type[1] as far as the algebraic structure is concerned.

Our aim is to exhibit general properties of these Riemannian spaces without specializing to any sort of field equation. Therefore we leave the Ricci tensor as an unspecified object in the equations to be derived. We are particularly interested in relations between the curvature of space–time and the kinematics of the congruence u^a. This interest stems from the fact that in a Riemannian space, there exist relations between the curvature and the kinematical behavior of a congruence, no matter whether this congruence represents the motion of sources of the gravitational field, or only of test material.

The main tools in our investigation are the timelike congruences and the Bianchi identities. In the same way as these identities have been proven to be useful for fields with a distinguished congruence of null curves, we can expect them to help us in our problem where a timelike congruence is singled out. Such congruences are of importance for the description of a gravitational field in terms of the history of an observer's rest space. Though given in covariant form, this description yields, of course, no information about intrinsic properties of the field unless the observer himself is uniquely singled out by the field. This prerequisite of any useful application of the concept of congruences is satisfied in our class of Riemannian spaces and, therefore, our results will reflect properties of the gravitational field.

In Sec 1 we deal with algebraic properties of the curvature tensor implied by the defining equation (1). In Sec. 2 we give the complete set of Bianchi identities which are, in the case of exterior fields, analyzed with respect to the triad defined by the eigenvectors of ϵ_{ab} (see summary). In the same section we also derive the main results of this paper, which are (i) an equation relating the proper time derivative of $E_{ac} \equiv -R_{abcd}u^bu^d$ to the Ricci tensor and to the kinematical quantities of the vector field u^a and (ii) an algebraic relation between the rotation and shear of u^a and the curvature tensor. In Sec. 3 we apply our results to give a simple proof of a theorem on exterior gravitational fields admitting a congruence of timelike curves without shear and rotation.

* Supported by the U. S. Air Force under Grant No. 509.64.
† Present address: Universität Hamburg, I. Institut für Theoretische Physik, Jungiustrasse 9, Hamburg, Germany.
[1] By this we mean that at every event there exists an observer with 4-velocity v^a to whom the field at that event has no magnetic component, i.e., the electromagnetic field tensor F_{ab} obeys $F^*_{ab}v^b = 0$.

NOTATION

Latin indices $a, b, c, \cdots$ run from 1 to 4
Greek indices $\mu, \nu, \lambda, \cdots$ run from 1 to 3 (throughout this paper we use Greek or numerical indices to denote nonholonomic components of a tensor with respect to an orthonormal triad.)
signature of the metric $+ + + -$
equality by definition $\equiv$
symmetric part $F_{(ab)} \equiv \frac{1}{2}(F_{ab} + F_{ba})$
skew part $F_{[ab]} \equiv \frac{1}{2}(F_{ab} - F_{ba})$
metric tensor g_{ab}, its determinant g,
$g_{abcd} \equiv g_{ac}g_{bd} - g_{ad}g_{bc}$
curvature tensor (Riemann tensor) R_{abcd}
conform tensor (Weyl tensor) C_{abcd}
contracted curvature tensor (Ricci tensor)
$R_{ab} \equiv R^c_{acb}$, Ricci scalar $R \equiv R^c_c$
alternating pseudotensor density η_{abcd}, $\eta^{1234} = |g|^{-\frac{1}{2}}$
duality operation, e.g.,
$^*R_{abcd} \equiv \frac{1}{2}\eta_{ab}{}^{rs}R_{rscd}$, $F^*_{ab} = \frac{1}{2}\eta_{abcd}F^{cd}$
covariant differentiation is denoted by a semicolon
kinematical quantities[2] for a congruence of timelike curves with unit tangent vector u^a

$$(h^c_a \equiv \delta^c_a + u_a u^c):$$

tensor of rotation $\omega_{ab} \equiv h^c_a u_{[c;d]}h^d_b$

vector of rotation $\omega^a \equiv \frac{1}{2}\eta^{abcd}u_b\omega_{cd}$

scalar of rotation $\omega \equiv (\frac{1}{2}\omega_{ab}\omega^{ab})^{\frac{1}{2}}$

scalar of expansion $\theta \equiv u^c_{;c}$

tensor of shear $\sigma_{ab} \equiv h^c_a(u_{(c;d)} - \frac{1}{3}\theta h_{cd})h^d_b$

scalar of shear $\sigma \equiv (\frac{1}{2}\sigma_{ab}\sigma^{ab})^{\frac{1}{2}}$

decomposition of $u_{a;b}$

$$u_{a;b} = \omega_{ab} + \sigma_{ab} + \tfrac{1}{3}\theta h_{ab} - \dot{u}_a u_b$$

covariant eigentime derivative $\dot{F}_{\cdots ab} \equiv F_{\cdots ab;c}u^c$ for any tensor $F_{\cdots ab}$.

1. ALGEBRAIC PROPERTIES

In this section we describe some algebraic properties of the curvature tensor which follow from

$$^*R_{abcd}u^b u^d = 0. \tag{1.1}$$

First of all it is clear that this equation is equivalent to

$$R_{abcd}u^d = 2E_{c[a}u_{b]}. \tag{1.2}$$

From (1.2) we get by contraction

$$q_a \equiv h^c_a R^d_c u_d = 0. \tag{1.3}$$

In order to express (1.2) in terms of the conform tensor we rewrite this equation using the decomposition

$$R_{abcd} = C_{abcd} - g_{abr[c}R^r_{d]} - \tfrac{1}{6}Rg_{abcd}$$

of the curvature tensor. Making use of the identity

$$R^a_c u^c = q^a + E^c_c u^a$$

and of (1.3) we obtain by a simple calculation

$$C_{abcd}u^d = 2\epsilon_{c[a}u_{b]}. \tag{1.4}$$

Also this calculation shows that from (1.4) and $q^a = 0$ we get (1.2). Thus we have

Theorem 1. The equation $^*R_{abcd}u^b u^d = 0$ is equivalent to the conditions

$$^*C_{abcd}u^b u^d = 0, \quad \text{and} \quad q_a = 0. \tag{1.5}$$

Because of (1.5) the conform tensor in our class of Riemannian spaces is Petrov type I (or type D) with real eigenvalues. Equation (1.3) means that there is no flux of nongravitational energy relative to the observer u^a. Since every Petrov type I conform tensor with real eigenvalues satisfies (1.4) where u^a is the timelike principal vector, we may characterize our class of fields by the following invariant conditions: The conform tensor is type I with real eigenvalues and the timelike principal vector is an eigenvector of the Ricci tensor.

As a consequence of (1.4) the conform tensor can be expressed by ϵ_{ab} according to the formula

$$C_{abcd} = -(g_{abpq}g_{cdrs} - \eta_{abpq}\eta_{cdrs})u^p u^r \epsilon^{qs}. \tag{1.6}$$

For the space projection $P_{abcd} \equiv h^k_a h^l_b h^m_c h^n_d R_{klmn}$ of the curvature tensor we get

$$P_{abcd} = \eta_{abpq}\eta_{cdrs}u^p u^r(E^{qs} - \tfrac{1}{2}P^{qs}) + h_{a[c}P_{d]b} - h_{b[c}P_{d]a} - \tfrac{1}{2}Eh_{abcd}, \tag{1.7}$$

where $P_{ab} \equiv h^c_a h^d_b R_{cd}$ is the space projection of the Ricci tensor and $h_{abcd} \equiv h_{ac}h_{bd} - h_{ad}h_{bc}$. The Riemann tensor itself becomes

$$R_{abcd} = P_{abcd} + 4u_{[a}E_{b][c}u_{d]}. \tag{1.8}$$

Since the conform tensor is algebraically equivalent to the rank-3 tensor ϵ_{ab} we can reduce the algebraic properties of C_{abcd} to those of ϵ_{ab}. Therefore we want to find out the way in which the Debever vectors of the conform tensor are determined by the tensor ϵ_{ab}.

It is well known that C_{abcd} determines in general

[2] For an introduction to the notion of the kinematical quantities see, e.g., J. Ehlers, Akad. Wiss. Lit. Mainz, Abhandl. Math. Nat. Kl. Nr. 11, **1961**, 798.

four null eigendirections (we call them also Debever vectors) k^a which satisfy the condition[3]

$$k_{[a}C_{b]\ cd[e}k_{f]}k^c k^d = 0. \tag{1.9}$$

We are free to choose k^a in such a way that $k_e u^e = 1$. If we multiply (1.9) with $u^a u^f$ and make use of (1.4) we get

$$C_{bcde}k^c k^d + 2k_b\epsilon_{c[d}u_{e]}k^c k^d - 2\epsilon_{d[b}u_{c]}k^c k^d k_e + \epsilon_{dc}k^c k^d k_b k_e = 0.$$

Now we insert (1.6) into this equation. We obtain

$$\epsilon_{be} - 2u_{(b}\epsilon^a_{e)}k_a + \epsilon_{cd}k^c k^d u_b u_e + 2\epsilon_{cd}k^c k^d k_{(b}u_{e)} - 2k_{(b}\epsilon^c_{e)}k_c + \epsilon_{cd}k^c k^d k_b k_e + \eta_{bcpq}\eta_{ders}k^c k^d u^p u^r \epsilon^{qs} = 0.$$

For the last term in this equation we introduce temporarily the abbreviation Y_{be}. Furthermore we call s^a the projection of k^a into the rest space of u^a, i.e., we define

$$s^a \equiv h^a_c k^c = k^a + u^a. \tag{1.10}$$

Then our equation reduces to (we rename indices)

$$\epsilon_{ab} - 2s_{(a}\epsilon^c_{b)}s_c + \epsilon_{cd}s^c s^d s_a s_b + Y_{ab} = 0. \tag{1.11}$$

To analyze this equation we introduce the eigenvectors e^a_ν and the eigenvalues λ_ν of ϵ^{ab}:

$$\epsilon^{ab} = \lambda_1 e^a_1 e^b_1 + \lambda_2 e^a_2 e^b_2 + \lambda_3 e^a_3 e^b_3, \quad \lambda_1 + \lambda_2 + \lambda_3 = 0,$$
$$e^c_\mu e^\lambda_c = \delta^\lambda_\mu.$$

If we multiply (1.11) by $e^a_1 e^b_2$ we have

$$-(\lambda_1 + \lambda_2)s_a e^a_1 s_c e^c_2 + \epsilon_{cd}s^c s^d s_a e^a_1 s_b e^b_2 + Y_{ab}e^a_1 e^b_2 = 0,$$

where the last term turns out to be $\lambda_3 s_a e^a_1 s_c e^c_2$. Therefore we have

$$(\epsilon_{ab}s^a s^b + 2\lambda_3)e^a_1 s_a e^b_2 s_b = 0,$$

and, similarly,

$$(\epsilon_{ab}s^a s^b + 2\lambda_2)e^a_3 s_a e^b_1 s_b = 0, \qquad (\epsilon_{ab}s^a s^b + 2\lambda_1)e^a_2 s_a e^b_3 s_b = 0. \tag{1.13}$$

From these equations we conclude that s^a is orthogonal to at least one of the eigenvectors e^a because otherwise the eigenvalues would be equal and, therefore, be zero. Let then $e^a_1 s_a = 0$. We consider first the case, where $e^a_2 s_a \neq 0$ and $e^a_3 s_a \neq 0$.

Then it follows from (1.13) that

$$2\lambda_1 + \epsilon_{ab}s^a s^b = 0. \tag{1.14}$$

Since we can write

$$s^a = \cos\Omega e^a_2 + \sin\Omega e^a_3, \tag{1.15}$$

(1.14) implies

$$\lambda_1 - \lambda_2 \sin^2\Omega - \lambda_3 \cos^2\Omega = 0. \tag{1.16}$$

These relations lead to a simplification of the last term in (1.11). Using (1.12), (1.14), and (1.15) we get after some manipulations

$$Y_{ab} = \lambda_1(s_a s_b - h_{ab}).$$

With this result we obtain from (1.11) and (1.14) finally

$$\epsilon_{ab} - 2s_{(a}\epsilon^c_{b)}s_c - \lambda_1(h_{ab} + s_a s_b) = 0. \tag{1.17}$$

Equation (1.16) permits us to derive an expression for cosΩ which determines the relative position of s^a with respect to the eigendirections of ϵ_{ab}. We have

$$\cos^2\Omega = (\lambda_1 - \lambda_2)/(\lambda_3 - \lambda_2), \qquad \sin^2\Omega = (\lambda_3 - \lambda_1)/(\lambda_3 - \lambda_2). \tag{1.18}$$

We observe that cosΩ and sinΩ are real only if $\lambda_2 < \lambda_1 < \lambda_3$ or if $\lambda_3 < \lambda_1 < \lambda_2$ holds; in other words, if λ_1 lies numerically between the two other eigenvalues. Because of the condition $\lambda_1 + \lambda_2 + \lambda_3 = 0$ this means that the absolute value of λ_1 is smaller than the absolute values of λ_2 and λ_3. The coefficients cosΩ and sinΩ themselves are determined up to a sign.

Now we consider the case where $e^c_1 s_c = 0$ and $e^c_2 s_c = 0$, say. Then s^a is collinear with e^a_3 and, therefore, is an eigenvector of ϵ_{ab} belonging to the eigenvalue λ_3. (1.11) now reduces to

$$\epsilon^{ab} - \lambda_3 s^a s^b - (\lambda_1 e^a_2 e^b_2 - \lambda_2 e^a_1 e^b_1) = 0$$

which implies $\lambda_1 = \lambda_2$ as seen by multiplication with $e^a_1 e^b_1$. Thus we have seen that the conform tensor is degenerate if s^a is an eigenvector of ϵ_{ab}. In such a case s^a is necessarily orthogonal to the degenerate eigenspace. On the other hand it is clear that degeneracy can only occur when s^a is orthogonal to the degenerate eigenspace because otherwise s^a would not be uniquely determined.

Note that our formulas are correct also in the case of degeneracy which is described by $\cos\Omega = 0$. We collect our results in the following

Theorem 2. If in a normal hyperbolic Riemannian 4-space there exists a timelike vector field u^a such that ${}^*C_{abcd}u^b u^d = 0$ then

(A) the four Debever vectors of the conform tensor lie in a hyperplane S,

(B) S is spanned by u^a and two eigenvectors of

[3] See e.g., R. Debever, Bull. Soc. Math. Belgique **10**, 112 (1959); R. Penrose, Ann. Phys. **10**, 171 (1960); P. Jordan, J. Ehlers, and R. K. Sachs, Akad. Wiss. Lit. Mainz, Abhandl. Math. Nat. Kl. No. 1, **1961**.

ϵ_{ab} which belong to the two eigenvalues with largest absolute value,

(C) the four null directions are given by

$$k_1^a = \cos\Omega e_2^a + \sin\Omega e_3^a + u^a,$$
$$k_2^a = \cos\Omega e_2^a - \sin\Omega e_3^a + u^a,$$
$$k_3^a = -\cos\Omega e_2^a + \sin\Omega e_3^a + u^a,$$
$$k_4^a = -\cos\Omega e_2^a - \sin\Omega e_3^a + u^a,$$

where

$$\cos\Omega = \left(\frac{\lambda_1 - \lambda_2}{\lambda_3 - \lambda_2}\right)^{\frac{1}{2}}, \qquad \sin\Omega = \left(\frac{\lambda_3 - \lambda_1}{\lambda_3 - \lambda_2}\right)^{\frac{1}{2}},$$

(D) the conform tensor is degenerate if and only if s^a is an eigenvector of ϵ_{ab}. In the case of degeneracy, formally described by $\cos\Omega = 0$, s^a is orthogonal to the degenerate eigenspace of ϵ_{ab}.

2. ANALYTIC PROPERTIES

In this section we derive some properties of our fields from the Bianchi identities. We do not make any assumption on the sources of our field but rather leave the Ricci tensor as an unspecified object in our equations.

The Bianchi identities, written in the form[4]

$$R^{abcd}{}_{;d} = 2R^{c[a;b]}, \tag{2.1}$$

can be decomposed into three sets of equations (the sign $\perp$ indicates that every free index in a tensor following this sign is projected with the help of h_a^c):

$$\text{set I} \quad u^a u^c R_{abcd}{}^{;d} = 2u^a u^c R_{c[a;b]}, \tag{2.2a}$$
$$\text{set II} \quad \perp u^a R_{abcd}{}^{;d} = \perp 2u^a R_{c[a;b]}, \tag{2.2b}$$
$$\text{set III} \quad \perp R_{abcd}{}^{;d} = \perp 2R_{c[a;b]}. \tag{2.2c}$$

The number of independent equations is three in set I, nine in set II, and eight in set III.

We start with an analysis of sets I and II. For this purpose we rewrite (1.2) as $R_{dabc}u^d = -2E_{a[b}u_{c]}$ and form the covariant derivative with respect to x^e. Then we obtain

$$R_{dabc}{}^{;e}u^d + R_{dabc}u^{d;e} = -(E_{ab}u_c - E_{ac}u_b)^{;e},$$

and, using (2.1), we have

$$2R_{b[d;a]}u^d + R_{dabc}u^{d;e} = -\dot{E}_{ab} - \theta E_{ab} + E^c_{a;c}u_b + E^c_a u_{b;c}.$$

If we insert the decomposition $R_{ab} = P_{ab} - E^c_c u_a u_b$ of the Ricci tensor into the last equation we get after partial integration

$$E_{,a}u_b + Eu_{b;a} - P^d_b u_{d;a} + \dot{E}_{ab} - \dot{P}_{ab} + \dot{E}u_a u_b + 2E\dot{u}_{(a}u_{b)} + R_{dabc}u^{d;e} + \theta E_{ab} - E^c_a u_{b;c} - E^c_{a;c}u_b = 0. \tag{2.3}$$

From this equation we obtain set I of the Bianchi identities by multiplying with u^b and projecting the result into the space orthogonal to u^a. Here and in the following formulas it will be convenient to use the tensor $D_{ac} \equiv E_{ac} - P_{ac} = -P_{abcd}h^{bd}$, which is obtained from the space projection of the Riemann tensor by contraction. We have

$$h^a_c D^b_{a;b} - D^b_c \dot{u}_b - \tfrac{1}{2}h^a_c D_{,a} = 0. \tag{2.4}$$

In order to obtain set II, we project (2.3) into the space orthogonal to u^a and form its symmetric and its skew part. This procedure yields

$$\perp\dot{D}_{ab} - D^c_{(a}\omega_{b)c} - E^c_{(a}\sigma_{b)c} - P^c_{(a}\sigma_{b)c} + \theta E_{ab} - \tfrac{2}{3}\theta P_{ab} + P_{dabc}\sigma^{dc} + E(\sigma_{ab} + \tfrac{1}{3}\theta h_{ab}) = 0, \tag{2.5}$$

and

$$E\omega_{ab} + E^c_{[a}\sigma_{b]c} + E^c_{[a}\omega_{b]c} - P^c_{[a}\sigma_{b]c} + P^c_{[a}\omega_{b]c} - P_{dabc}\omega^{dc} = 0. \tag{2.6}$$

Equation (2.5) relates the space projection of the covariant eigentime derivative of D_{ab} to the Riemann tensor and to the kinematical quantities. (2.6) has the remarkable feature to be an algebraic condition relating the shear and rotation of u^a to the curvature of space–time. It is clear that this equation will be useful in the investigation of gravitational fields which belong to the class treated here and which have a congruence with shear and rotation.

Taking the dual of (2.6) we obtain the equivalent relation

$$3D^a_c\omega^c + 2P^a_c\omega^c + \eta^{abcd}D^m_b\sigma_{cm}u_d = 0. \tag{2.7}$$

To analyze set III of the Bianchi identities we insert (1.8) into (2.1) and form the space projection of the resulting equation. Then we have after some simplifications

$$\eta_{abpq}\eta_{cdrs}u^p u^r(E^{qs;d} - \tfrac{1}{2}P^{qs;d}) - \eta_{abpq}\eta_{cdrs}u^p u^d \dot{u}^r(E^{qs} - \tfrac{1}{2}P^{qs}) + \perp(h_{c[a}P_{b]d})^{;d} - (h_{d[a}P_{b]c})^{;d} - \tfrac{1}{2}E^{;d}h_{abcd} - \tfrac{1}{2}\perp E h_{abcd}{}^{;d} - E_{ac}\dot{u}_b + E_{bc}\dot{u}_a = 2\perp R_{c[a;b]}. \tag{2.8}$$

The equations (2.4), (2.5), (2.7), and (2.8) are the complete set of Bianchi identities.

So far we did not make any assumption on the source of the gravitational field apart from (1.3)

[4] This form of the Bianchi identities is valid only in a 4-dimensional space; see W. Kundt and M. Trümper; Akad. Wiss. (Mainz) Abhandl. Math. Nat. Kl. Nr. 12, 1962, 970.

which is an immediate consequence of our basic assumption (1.1). Since we are mainly interested in the investigation of exterior solutions of Einstein's field equations we now specialize our formulas to the case of vacuum fields. For the remainder of this section we assume $R_{ab} = 0$. It turns out to be convenient to use the eigenvalues λ_μ and the eigenvectors e^a_μ of ϵ_{ab} as introduced by (1.12). Furthermore we define the Ricci coefficients

$$\gamma_{\mu\nu\rho} = e_{\mu a;c} e^a_\nu e^c_\rho,$$

and the nonholonomic components of the gradient of any scalar S with respect to the eigentriad of ϵ_{ab}:

$$S_{.\mu} = S_{,a} e^a_\mu.$$

The complete system of Bianchi identities reduces then to (in the following formulas no summation over repeated Greek indices!)

$$\text{set I} \quad \lambda_{\mu.\mu} + (\lambda_\mu - \lambda_\nu)\gamma_{\mu\nu\nu} + (\lambda_\mu - \lambda_\rho)\gamma_{\mu\rho\rho} = 0, \quad (2.9)$$

where μ, ν, ρ are unequal,

$$\text{set II} \quad \dot\lambda_\mu + \theta\lambda_\mu - 2\sigma_{\mu\mu}\lambda_\mu + \sigma_{\nu\nu}\lambda_\nu + \sigma_{\rho\rho}\lambda_\rho = 0, \quad (2.10)$$

$$(\lambda_\mu - \lambda_\nu) e^\mu_c e^c_\nu + \tfrac{1}{2}(\lambda_\mu - \lambda_\nu)\omega_\rho + \tfrac{3}{2}\lambda_\rho \sigma_{\mu\nu} = 0, \quad (2.11)$$

$$3\lambda_\mu\omega_\mu + (\lambda_\nu - \lambda_\rho)\sigma_{\nu\rho} = 0, \quad (2.12)$$

μ, ν, $\rho = 1, 2, 3$ and cyclic,

$$\text{set III} \quad (\lambda_3 - \lambda_1)\gamma_{312} = (\lambda_2 - \lambda_3)\gamma_{231} = (\lambda_1 - \lambda_2)\gamma_{123}, \quad (2.13)$$

$$\lambda_{\mu.\nu} + (\lambda_\mu - \lambda_\nu)\gamma_{\nu\mu\mu} + (\lambda_\mu - \lambda_\rho)\dot u_\nu = 0, \quad (2.14)$$

where μ, ν, ρ are unequal.

In these formulas we use tensor components in the nonholonomic system e^a_ν, i.e., we define

$$\sigma_{\mu\nu} \equiv \sigma_{ab} e^a_\mu e^b_\nu, \qquad \omega_\mu \equiv \omega_a e^a_\mu, \qquad \dot u_\mu \equiv \dot u_a e^a_\mu.$$

Since we are dealing with the Bianchi identities for exterior fields set I is a consequence of set III; in fact, (2.9) follows from (2.14).

It should be noted that (2.13) and (2.14) do not contain the shear, expansion, and rotation of u^a. These equations are in perfect analogy to those in the case of static exterior fields,[5] which are a very special case contained in the class of fields treated here.

3. APPLICATION: NORMAL HYPERBOLIC RIEMANNIAN 4-SPACES ADMITTING A CONGRUENCE OF TIMELIKE CURVES WITHOUT SHEAR AND ROTATION

We start this section with some general remarks about the type of space under consideration and then we show how the preceding results can be applied to these fields.

Since we assume $\sigma = 0$, $\omega = 0$ we have

$$u_{a;b} = \tfrac{1}{3}\theta h_{ab} - \dot u_a u_b. \quad (3.1)$$

The Ricci identities, applied to this formula, then yield

$$-\tfrac{1}{2}R_{dabc}u^d = \tfrac{1}{3}\theta_{,[c}h_{b]a} + \tfrac{1}{3}\dot\theta g_{a[b}u_{c]} + E_{a[b}u_{c]}$$

which simplifies to

$$-\tfrac{1}{2}R_{dabc}u^d = \tfrac{1}{3}\theta_{,d}h^d_{[c}h_{b]a} + E_{a[b}u_{c]}. \quad (3.2)$$

To replace the term $\theta_{,d}h^d_c$ we go back to the definition $\theta \equiv u^c{}_{;c}$ and form

$$\theta_{,d} = u^a{}_{;ad} = u^a{}_{;da} + R^{ca}{}_{da}u_c = \tfrac{1}{3}(\theta h^a_d)_{;a} - (\dot u^a u_d)_{;a} + R^c_d u_c.$$

From this it follows

$$h^d_c\theta_{,d} = \tfrac{3}{2}q_c. \quad (3.3)$$

This means that under our assumption $\sigma = \omega = 0$ the expansion scalar is constant on each hypersurface orthogonal to the trajectories of u^a if and only if u^a is an eigenvector of the Ricci tensor. From (3.3) and (3.2) we get

$$-R_{dabc}u^d = h_{a[b}q_{c]} + 2E_{a[b}u_{c]}. \quad (3.4)$$

With the last two equations in mind we can make a general statement on the conform tensor of these fields.

Theorem 3. The conform tensor of a normal hyperbolic 4-space admitting a congruence of timelike curves K without shear and rotation is Petrov type I (or type D) with real eigenvalues. The tangent vectors of K are principal vectors of the conform tensor.[6]

To prove this theorem we remark first that it holds true for any timelike congruence with $\sigma = \theta = \omega = 0$. This follows immediately from (3.3), (3.4), and Theorem 1.

Assume now we are given a congruence with unit tangent vector u^a such that $\sigma = \omega = 0$. Then we consider this congruence on the same manifold to which we assign a conformally transformed metric. It is clear that the conditions $\sigma = \omega = 0$ still hold true with respect to the new metric. However, by a suitable choice of the conform factor we can arrange that the expansion scalar with respect to the new metric also will be zero. Therefore the conform tensor has the algebraic structure (1.4) and since this state-

[5] Compare P. Jordan, J. Ehlers, and W. Kundt, Akad. Wiss. Lit. Mainz, Abhandl. Math. Nat. Kl. Nr. 2, 1960, 64.

[6] This theorem together with Theorem 1 generalizes and completes a theorem by Ehlers in Ref. 2, p. 811.

ment is invariant under conformal transformation C_{abcd} has this structure also in the ease $\theta \neq 0$. This proves our theorem.

Now we turn to the consideration of special Einstein spaces, i.e., we assume $R_{ab} = 0$. Equation (3.4) shows that then our field satisfies (1.1), the basic assumption of this paper. Our main objective is to give a proof of

Theorem 4[7]. Any four-dimensional nondegenerate special Einstein space which admits a congruence K of timelike curves with vanishing shear and rotation is a static field, i.e., K has necessarily zero expansion and its acceleration vector $\dot{u}^a$ is Fermi propagated along $u^a (h^c_a \ddot{u}_c = 0)$.

To prove this theorem we show first that there exists a timelike conform vector which is hypersurface orthogonal, i.e., a vector ξ^a satisfying

$$\xi_{(a;b)} - \tfrac{1}{4}\xi^c{}_{;c} g_{ab} = 0, \qquad \xi_{[a;b}\xi_{c]} = 0.$$

Then we prove that a special Einstein space admitting such a vector field is flat unless $\xi^c{}_{;c} = 0$.

Assuming that the three eigenvalues λ_μ are unequal we get immediately from (2.10) and (2.11)

$$\dot{\lambda}_\mu + \theta\lambda_\mu = 0 \tag{3.5}$$

and

$$e^c_\mu = \epsilon_\mu u_c, \tag{3.6}$$

where the ϵ_μ turn out to be the nonholonomic components of $\dot{u}_c$ with respect to the eigentriad of ϵ_{ab}. Now we form the derivative of the equations (2.14) with respect to the eigentime. A straightforward calculation and the use of our previous results yields

$$(\dot{u}_\mu)^{\cdot} = 0,$$

which is equivalent to

$$\ddot{u}_{[a} u_{b]} = 0. \tag{3.7}$$

This means that $\dot{u}^a$ is Fermi propagated along u^a. From this fact and from $\sigma = \omega = 0$ it follows that u^a is collinear to a conform vector as shown in this way:

We compute

$$\dot{u}_{a;c} = (u_{a;b}u^b)_{;c} = (u_{a;c})^{\cdot} - R_{dabc}u^d u^b + \tfrac{1}{3}\theta h_{ab}(\tfrac{1}{3}\theta h^b_c - \dot{u}^b u_c)$$

and take the skew part of this equation. Then we get because of (3.1) and (3.7)

$$\dot{u}_{[a;c]} = -\tfrac{1}{3}\theta \dot{u}_{[a} u_{c]}$$

which, by virtue of (3.3), can be written as

$$\dot{u}_{[a;b]} - \tfrac{1}{3}(\theta u_{[a})_{;b]} = 0.$$

Therefore $\dot{u}_a - \tfrac{1}{3}\theta u_a$ is the gradient of a scalar W and one checks easily that $\xi_a \equiv e^W u_a$ obeys $\xi_{(a;b)} - \tfrac{1}{4}\xi^c{}_{;c} g_{ab} = 0$. If $\theta = 0$, then we have $\xi^c{}_{;c} = 0$ and ξ_a is a Killing vector. In this case the field is static. If $\theta \neq 0$ we know that ξ_a is a conform Killing vector. Then we choose a time coordinate t in such a way that it is constant on every hypersurface orthogonal to u^a. The metric takes the form

$$G = e^{2\phi}(g_{\mu\nu}\, dx^\mu\, dx^\nu - dt^2) = e^{2\phi}(\tilde{G} - dt^2),$$

where $\phi = \phi(x^a)$ and $g_{\mu\nu} = g_{\mu\nu}(x^\lambda)$. This means that the metric is conform to the direct product of a positive-definite Riemannian 3-space $\tilde{G}$ and the t axis. In such a case the conditions $R_{ab} = 0$ require $\tilde{G}$ to be a space of constant curvature unless $\partial\phi/\partial t = 0$. Therefore, since the direct product of a three-dimensional space of constant curvature and the real axis is conformally flat, our space–time is flat. The only way to exclude such a trivial case is to assume $\partial\phi/\partial t = 0$ which is equivalent to $\theta = 0$.

Note added in proof: It has been shown by Dr. W. Kundt (private communication) that the exterior Schwarzschild solution (which is Petrov type D) admits a timelike congruence with vanishing shear and rotation, but with nonvanishing expansion. Therefore Theorem 4 may be stated in a simpler form as follows: If a special Einstein space admits a timelike congruence with vanishing shear and rotation, then it is either static or it is type D.

ACKNOWLEDGMENTS

The author is indebted to Professor Peter G. Bergmann for his kind hospitality and for many interesting discussions. Several remarks of Dr. Michel Cahen and Dr. Wolfgang Rindler have been of considerable value and are gratefully acknowledged.

[7] M. Trümper, Z. Physik **168**, 55 (1962).

New Homogeneous Solutions of Einstein's Field Equations with Incoherent Matter Obtained by a Spinor Technique

István Ozsváth
Southwest Center for Advanced Studies, Dallas, Texas
(Received 14 April 1964)

Einstein's field equations with incoherent matter are solved for the case of homogeneous space-time, i.e., for metrics allowing a four parametric simply transitive group of motions. Two families of new solutions are obtained by use of a spinor technique. As a special result a proof emerges for Gödel's theorem, which states that there exist only two homogeneous solutions of Einstein's field equations with incoherent matter and rigid rotation, namely the Gödel cosmos and the Einstein static universe.

I. INTRODUCTION

We begin with a brief summary of a method previously used[1,2] for solving Einstein's field equations, which will then be modified by the use of spinor techniques.

Suppose in a four-dimensional differentiable manifold M with the coordinates x^0, x^1, x^2, x^3 we have four contravariant vector fields

$$\xi^i{}_a = \xi^i{}_a(x^k), \tag{1.1}$$

where $a, b, \cdots = 0, 1, 2, 3$ label the vectors and $j, k, \cdots = 0, 1, 2, 3$ the coordinates. Suppose further that the condition

$$\det(\xi^i{}_a) \neq 0 \tag{1.2}$$

holds everywhere, and that the functions $\xi^i{}_a(x^k)$ satisfy the equations

$$\xi^i{}_a\xi^k{}_{b,i} - \xi^i{}_b\xi^k{}_{a,i} = C_{ab}{}^c\xi^k{}_c, \tag{1.3}$$

where a comma denotes the partial derivative, and the constants $C_{ab}{}^c$ have the properties

$$C_{ab}{}^c + C_{ba}{}^c = 0, \quad C_{fb}{}^a C_{cd}{}^f + C_{fc}{}^a C_{db}{}^f + C_{fd}{}^a C_{bc}{}^f = 0. \tag{1.4}$$

Then we say that we have a four-parametric simply transitive group operating in the manifold. This means that there exists for any two arbitrary points P and Q of M one and only one transformation of the group which carries P into Q. $C_{ab}{}^c$ are the structure constants of the group, (1.4) are the Jacobi identities, and the vector fields $\xi^i{}_a$ are called the infinitesimal generators of the group.

The invariant vectors of the group are defined by the equations

$$\xi^i{}_a e^k{}_{b,i} - e^i{}_b\xi^k{}_{a,i} = 0 \tag{1.5}$$

which must hold for each a and b, with the proviso that

$$\det(e^i{}_a) \neq 0 \tag{1.6}$$

in the whole manifold. If we impose the conditions $e^i{}_a(x_0^k) = \xi^i{}_a(x_0^k)$, where x_0^k are the coordinates of some arbitrarily chosen point in M, we have for the functions $e^i{}_a(x^k)$ the equations

$$e^i{}_a e^k{}_{b,i} - e^i{}_b e^k{}_{a,i} = -C_{ab}{}^c e^k{}_c, \tag{1.7}$$

as a consequence of (1.3) and (1.5). By means of the equivalent equations

$$e^i{}_a e^b{}_i = \delta^b{}_a \quad \text{or} \quad e^a{}_i e^k{}_a = \delta^k{}_i \tag{1.8}$$

we can define four covariant vector fields $e^a{}_i$ which then satisfy the equations

$$e^a{}_{i,k} - e^a{}_{k,i} = -C_{bc}{}^a e^b{}_i e^c{}_k \tag{1.9}$$

and the condition

$$\det(e^a{}_i) \neq 0 \tag{1.10}$$

in the whole manifold.[3,4] $e^a{}_i$ are the reciprocal vectors of the group.

The transformations of the group carry the vector fields $e^i{}_a$ and $e^a{}_i$ into themselves. We have therefore an invariant tetrad of contravariant vectors $e^i{}_a$ and an invariant tetrad of covariant vectors $e^a{}_i$ in each point of our manifold M, and it is possible to assign unique tetrad components to each tensor, e.g., $T^i{}_{kl}$, by the equations

$$T^i{}_{kl} = e^i{}_a T^a{}_{bc} e^b{}_k e^c{}_l \quad \text{or} \quad T^a{}_{bc} = e^a{}_i T^i{}_{kl} e^k{}_b e^l{}_c. \tag{1.11}$$

$T^a{}_{bc}$ are in general functions of the coordinates; but if $T^a{}_{bc}$ are constant we call $T^i{}_{kl}$ *invariant* with respect to the group.

We now extend our manifold M to a Riemannian

[1] O. Heckmann and E. L. Schücking, "Relativistic Cosmology," in *Gravitation* edited by L. Witten (John Wiley & Sons, Inc., New York (1962).
[2] I. Ozsváth, Akad. Wiss. Mainz, Abhandl. Math. Nat. Kl. Nr. 13, 1962.
[3] L. P. Eisenhardt, *Continuous Groups of Transformations* (Princeton University Press, Princeton, New Jersey, 1933).
[4] A. Taub, Ann. Math. 53, 472 (1951); A. Z. Petrov, *Prostranstva Einsteina* (GIPhML, Moscow, 1961).

space R by introducing a nonsingular symmetric tensor g_{ik} as a metric. Defining g^{ik} by

$$g^{ik}g_{kl} = \delta^i_l \tag{1.12}$$

and assigning tetrad components in accordance with (1.11) we have

$$g_{ab} = e^i{}_a g_{ik} e^k{}_b = e_{ak}e^k{}_b, \quad g^{ab} = e^a{}_i g^{ik} e^b{}_k = e^{ak}e^b{}_k \tag{1.13}$$

and

$$g^{ab}g_{bc} = \delta^a_c, \tag{1.14}$$

and

$$g_{ik} = e^a{}_i g_{ab} e^b{}_k. \tag{1.15}$$

To the raising and lowering the indices in R with g^{ik} and g_{ik} there exists a corresponding raising and lowering of the tetrad indices by means of the tetrad metric g^{ab} and g_{ab}. *If g_{ab} are constant, and we will make this assumption henceforth, we call R homogeneous.*

We now show that the covariant derivative with respect to g_{ik} can also be expressed in the tetrad. We define the Ricci rotation coefficients by

$$A_{abc} = e_{ai;k}e^i{}_b e^k{}_c, \tag{1.16}$$

where the semicolon denotes covariant differentiation with respect to the metric, and

$$e_{ai} = g_{ik}e^k{}_a = g_{ab}e^b{}_i. \tag{1.17}$$

From

$$g_{ab;k} = (e_{ai}e^i{}_b)_{;k} = e_{ai;k}e^i{}_b + e^i{}_a e_{bi;k} = 0$$

it follows that A_{abc} has the property

$$A_{abc} + A_{bac} = 0 \tag{1.18}$$

and therefore the following identity holds:

$$A_{abc} = A_{a[bc]} + A_{b[ca]} - A_{c[ab]}. \tag{1.19}$$

From (1.16) and (1.9) one sees immediately that

$$A_{a[bc]} = e_{a[i;k]}e^i{}_b e^k{}_c = e_{a[i,k]}e^i{}_b e^k{}_c = -\tfrac{1}{2}C_{bca}, \tag{1.20}$$

where

$$C_{bca} = C_{bc}{}^d g_{da} \tag{1.21}$$

and so we have

$$A_{abc} = -\tfrac{1}{2}(C_{bca} + C_{cab} - C_{abc}). \tag{1.22}$$

The covariant derivatives of the tetrad vectors are given by

$$e^i{}_{a;k} = A_a{}^b{}_c e^i{}_b e^c{}_k, \qquad e^a{}_{i;k} = A^a{}_{bc}e^b{}_i e^c{}_k \tag{1.23}$$

which immediately follows from (1.16). The covariant derivative of a vector $u_i = u_a e^a{}_i$ which is invariant (u_a constant) is given by

$$u_{i;k} = u_{a;b}e^a{}_i e^b{}_k, \qquad u_{a;b} = u_{i;k}e^i{}_a e^k{}_b, \tag{1.24}$$

where

$$u_{a;b} = u^f A_{fab}. \tag{1.25}$$

Using the Ricci identity

$$e_{ak;l;m} - e_{ak;m;l} = -e^i{}_a R_{iklm} \tag{1.26}$$

one can calculate the tetrad components of the Riemann tensor, which turn out to be

$$R_{abcd} = A_{fac}A^f{}_{bd} - A_{fad}A^f{}_{bc} + A_{abf}C_{cd}{}^f. \tag{1.27}$$

One defines η_{abcd} by the equation

$$\eta_{abcd} = (-g)^{\frac{1}{2}}\epsilon_{abcd}, \tag{1.28}$$

where $g = \det(g_{ab})$ and ϵ_{abcd} is totally skew symmetric and $\epsilon_{0123} = +1$. One then defines the tensor R^*_{abcd} by the equation

$$R^*_{abcd} = \tfrac{1}{2}R_{ab}{}^{fg}\eta_{cdfg}. \tag{1.29}$$

The cyclic symmetry of the Riemann tensor

$$R_{abcd} + R_{acdb} + R_{adbc} = 0 \tag{1.30}$$

can be expressed by the equation

$$R^*_{abc}{}^b = 0. \tag{1.31}$$

Substituting (1.27) into (1.31) one finds by a short straightforward calculation that

$$R^{*a}{}_b{}^{cb} = \tfrac{1}{2}\eta^{cfgh}C_{df}{}^a C_{gh}{}^d = 0 \tag{1.32}$$

and this is simply the Jacobi identity, as can be seen from (1.4). *Thus the cyclic symmetry of R_{abcd} is equivalent to the Jacobi identities.* This fact is of importance below.

The tetrad components of the Ricci tensor are given by

$$R_{bd} = g^{ac}R_{abcd}. \tag{1.33}$$

Using (1.27) we get the expression

$$R_{ab} = A^f{}_{ag}A^g{}_{bf} - C_{fg}{}^f A^g{}_{ab}, \tag{1.34}$$

where

$$A^a{}_{bc} = g^{af}A_{fbc}. \tag{1.35}$$

Einstein's field equations with incoherent matter are:

$$R_{ik} - \tfrac{1}{2}Rg_{ik} + \Lambda g_{ik} = -Au_i u_k, \quad u_i u^i = 1, \tag{1.36}$$

where $R = g^{ik}R_{ik}$, Λ is the cosmological constant, and $A > 0$ is the relativistic constant of gravitation multiplied by the density of the incoherent matter. These can be written in tetrad components as follows

$$R_{ab} - \tfrac{1}{2}Rg_{ab} + \Lambda g_{ab} = -Au_a u_b, \quad u_a u^a = 1, \tag{1.37}$$

where $R = g^{ik}R_{ik} = g^{ab}R_{ab}$ and u_a is defined by

$$u_a = u_i e^i{}_a. \tag{1.38}$$

Since the terms on the left-hand side of (1.37) are constant, it is easy to see that A and u_a are also necessarily constant. (To prove the constancy of A we raise b and contract with a). This means that *in a homogeneous space-time, the density of the incoherent matter is constant and the velocity vector of the matter is an invariant vector of the group.*

The problem of finding homogeneous solutions of the field equations (1.36) turns out to be equivalent to the problem of solving the algebraic equations (1.37). We deal with this algebraic problem in the next paragraphs, but at the moment we give some additional formulas which we need later. From the twice-contracted Bianchi identities $(R^a{}_b - \frac{1}{2}R\delta^a{}_b)_{:a} = 0$ and from (1.37) it follows that $u^a{}_{:a} = 0$ and $u_{a:b}u^b = 0$. These conditions imply (as can be seen from their tensor equivalents) that *the expansion vanishes and the world lines of the matter are geodesic.* Writing out the first of these equations more explicitly we get

$$u^a{}_{:a} = u^f A_f{}^a{}_a = -u^f A^a{}_{fa} = -u^f C^a{}_{fa} = -u^f k_f = 0, \tag{1.39}$$

where we introduced the vector k_f defined by

$$k_f = A^a{}_{fa} = C^a{}_{fa}. \tag{1.40}$$

Equation (1.39) means that k_a is either perpendicular to u_a or zero. The second equation yields

$$u_{a:b}u^b = -A_{a(bc)}u^b u^c = -C_{a(bc)}u^b u^c = 0. \tag{1.41}$$

Since the expansion vanishes and the world lines of the matter are geodesic, the tetrad components of the shear tensor[1] are given by

$$\sigma_{ab} = u_{(a:b)} \quad \text{(hence } \sigma^a{}_a = 0\text{)}. \tag{1.42}$$

If the shear vanishes, i.e.,

$$\sigma_{ab} = u_{(a:b)} = u^f A_{f(ab)} = u^f C_{f(ab)} = 0, \tag{1.43}$$

then the world lines of the matter are group trajectories, since $u_{(a:b)} = 0$ are the Killing equations The tetrad components of the rotation vector are given by

$$\omega^a = \tfrac{1}{2}\eta^{abcd}u_b\omega_{cd}, \tag{1.44}$$

where

$$\omega_{cd} = u_{[c:d]} = u^f A_{f[cd]} = -\tfrac{1}{2}C_{cdf}u^f. \tag{1.45}$$

After these preliminaries we come to the algebraic problem of solving the equations (1.37).

2. SPINOR FORMS OF THE BASIC EQUATIONS

In this paper we use spinors merely as an algebraic device. Little of the standard theory is used beyond the notation. See for example Refs. 5 and 6.

We have two different ways to solve the algebraic equations

$$R_{ab} = -Au_a u_b + (\Lambda + \tfrac{1}{2}A)g_{ab}, \quad u_a u^a = 1, \tag{2.1}$$

which are trivially equivalent to (1.37). *In method* (*i*) we regard the structure constants $C_{ab}{}^c$ as given and the constants g_{ab} as unknown, and we use Eqs. (2.1) to determine these unknowns. *In method* (*ii*) we regard the g_{ab}'s as given and the $C_{ab}{}^c$'s as unknown and we use the equations (2.1) and (1.31) to determine these unknowns. The first approach has been used in[2] where some solutions of the field equations in vacuum and also with incoherent matter are given. This method is only adapted to finding the solution, if such exists, belonging to a given group, and to study the properties of this solution. However, it does not allow us to make general statements about homogeneous solutions. Consequently in this paper the second approach is chosen.

Suppose we have a homogeneous space given by $g_{jk} = e^a{}_j g_{ab} e^b{}_k$ with the signature -2. Then we can always find a linear transformation with constant coefficients $A^a{}_{a'}$ such that $g_{a'b'} = A^a{}_{a'}g_{ab}A^b{}_{b'} = \text{diag}\,(1, -1, -1, -1)$. Then we can write g_{jk} in the form $g_{jk} = e^{a'}{}_j g_{a'b'} e^{b'}{}_k$, where $e^{a'}{}_j = A^{a'}{}_a e^a{}_j$ and $A^{a'}{}_a A^a{}_{b'} = \delta^{a'}{}_{b'}$ or $A^a{}_{a'}A^{a'}{}_b = \delta^a{}_b$. The vectors $e^{a'}{}_j$ satisfy the equations $e^{a'}{}_{j,k} - e^{a'}{}_{k,j} = -C_{b'c'}{}^{a'} e^{b'}{}_j e^{c'}{}_k$, where $C_{a'b'}{}^{c'} = A^a{}_{a'}A^b{}_{b'}C_{ab}{}^c A^{c'}{}_c$. That means that the $e^{a'}{}_j$ are the reciprocal vectors of the same group as $e^a{}_j$, since *the structure constants of a given group are only determined up to linear transformations.* Therefore we can suppose without loss of generality that g_{ab} already has the form

$$g_{ab} = g^{ab} = \eta_{ab} = \eta^{ab} = \text{diag}\,(+1, -1, -1, -1). \tag{2.2}$$

The tetrads $e^i{}_a$ and $e^a{}_j$ are fixed by (2.2) only up to the transformations of the homogeneous Lorentz group. We could still choose instead of $e^a{}_j$ the tetrad $e^{a'}{}_j$ given by $e^{a'}{}_j = L^{a'}{}_a e^a{}_j$ where the constants $L^{a'}{}_a$ satisfy the equations $L^{a'}{}_a L^{b'}{}_b \eta_{a'b'} = \eta_{ab}$. We could use these transformations to simplify the structure constants or the vectors u_a and k_a and try to solve the algebraic problem in this tetrad. But it appears to be easier to use a null tetrad and its related spinors.

[5] R. Penrose, Ann. Phys. (N. Y.) **10**, 171 (1960).

[6] W. L. Bade and H. Jehle, Rev. Mod. Phys. **25**, 714 (1953); E. M. Corson, *Introduction to Tensors, Spinor, and Relativistic Wave-Equations* (Blackie & Son, Ltd., London, 1953).

We introduce a null tetrad $\sigma^i{}_{AA'}$ as a special linear combination of the original orthonormal tetrad, namely

$$\sigma^i{}_{AA'} = e^i{}_a \sigma^a{}_{AA'} \tag{2.3}$$

$$\sigma^{AA'}{}_i = \sigma^{AA'}{}_a e^a{}_i, \tag{2.4}$$

where

$$\sigma^a{}_{AA'} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & i & -i & 0 \\ 1 & 0 & 0 & -1 \end{bmatrix} \tag{2.5}$$

$$\sigma^{AA'}{}_a = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & -i & 0 \\ 0 & 1 & i & 0 \\ 1 & 0 & 0 & -1 \end{bmatrix}. \tag{2.6}$$

In Formula (2.5) [(2.6)], a is the row (column) index and the pairs AA' (A, A', $\cdots = 1, 2$) are the column (row) indices, related to the indices 0, 1, 2, 3 by the following correspondence

$$\begin{array}{cccccc} a & 0 & 1 & 2 & 3 \\ AA' & 11' & 12' & 21' & 22'. \end{array} \tag{2.7}$$

(The numerical values of A and A' are independent of each other.) These matrices are taken from the forthcoming book by Penrose and Rindler[7] (see also Ref. 5). They are used to translate the tensors of the Minkowski space into spinors.

Since $\det(\sigma^a{}_{AA'}) = i$ and $\det(\sigma^{AA'}{}_a) = -i$, it follows that

$$\det(\sigma^i{}_{AA'}) \neq 0 \text{ and } \det(\sigma^{AA'}{}_i) \neq 0. \tag{2.8}$$

The equations corresponding to the equations (1.8) are

$$\sigma^i{}_{AA'} \sigma^{AA'}{}_k = \delta^i_k \text{ and } \sigma^i{}_{AA'} \sigma^{BB'}{}_i = \delta^B_A \delta^{B'}_{A'}. \tag{2.9}$$

Using $\sigma^i{}_{AA'}$ and $\sigma^{AA'}{}_i$ as basic tetrads, the tetrad components of the tensors of our Riemannian manifold become spinors—strictly speaking, dyad components of spinors (see Ref. 8). The translation is given by such typical equations as

$$\begin{aligned} T^{AA'}{}_{BB'CC'} &= \sigma^{AA'}{}_i T^i{}_{kl} \sigma^k{}_{BB'} \sigma^l{}_{CC'}, \\ T^i{}_{kl} &= \sigma^i{}_{AA'} T^{AA'}{}_{BB'CC'} \sigma^{BB'}{}_k \sigma^{CC'}{}_l, \end{aligned} \tag{2.10}$$

which are analogous to the equations (1.11). We note that this is the usual null tetrad formalism except that in place of each single tetrad index we write a pair of "spinor" indices AA'. The advantages becomes apparent below. It follows further that

$$\begin{aligned} T^{AA'}{}_{BB'CC'} &= \sigma^{AA'}{}_a T^a{}_{bc} \sigma^b{}_{BB'} \sigma^c{}_{CC'}, \\ T^a{}_{bc} &= \sigma^a{}_{AA'} T^{AA'}{}_{BB'CC'} \sigma^{BB'}{}_b \sigma^{CC'}{}_c. \end{aligned} \tag{2.11}$$

Equations (2.11) give the correspondence between the two sort of "tetrad components." The invariant tensors of our manifold evidently have constant spinor components. For example

$$\eta_{ab} \leftrightarrow \epsilon_{AB} \epsilon_{A'B'}, \qquad \eta^{ab} \leftrightarrow \epsilon^{AB} \epsilon^{A'B'}, \tag{2.12}$$

where ϵ_{AB}, $\epsilon_{A'B'}$ and ϵ^{AB}, $\epsilon^{A'B'}$ are skew symmetric and

$$\epsilon_{12} = \epsilon_{1'2'} = \epsilon^{12} = \epsilon^{1'2'} = 1. \tag{2.13}$$

The raising and lowering of spinor indices is carried out according to the typical rules

$$\begin{aligned} \varphi^A &= \epsilon^{AB} \varphi_B, & \varphi_B &= \varphi^A \epsilon_{AB}, \\ \psi^{A'} &= \epsilon^{A'B'} \psi_{B'}, & \psi_{B'} &= \psi^{A'} \epsilon_{A'B'}. \end{aligned} \tag{2.14}$$

If

$$F_{ab} = -F_{ba} \quad (F_{ab} \text{ real}), \tag{2.15}$$

then the spinor $F_{AA'BB'} \leftrightarrow F_{ab}$ can be split in the following way

$$F_{AA'BB'} = \tfrac{1}{2}(F_{AB} \epsilon_{A'B'} + \bar{F}_{A'B'} \epsilon_{AB}), \tag{2.16}$$

where

$$F_{AB} = F_{BA}. \tag{2.17}$$

We have used here the notation typified by

$$\overline{\phi_{ABC'}} = \bar{\phi}_{A'B'C} \tag{2.18}$$

for the complex conjugate spinor. (In *this* equation $A = A'$, etc., numerically.) Since the quantities A_{abc} are real tetrad tensors and skew symmetric in the first two indices, the spinor $A_{AA'BB'CC'} \leftrightarrow A_{abc}$ can be written in the form

$$A_{AA'BB'CC'} = \tfrac{1}{2}(A_{ABCC'} \epsilon_{A'B'} + \bar{A}_{A'B'C'C} \epsilon_{AB}), \tag{2.19}$$

where $A_{ABCC'}$, which we call the "Ricci rotation spinor," has the symmetry property

$$A_{ABCC'} = A_{BACC'}. \tag{2.20}$$

It is convenient later to split this further, thus:

$$A_{ABCC'} = \alpha_{ABCC'} + \tfrac{1}{3}(\epsilon_{AC} \alpha_{BC'} + \epsilon_{BC} \alpha_{AC'}), \tag{2.21}$$

where the spinor $\alpha_{ABCC'}$ is totally symmetric in the unprimed indices:

$$\alpha_{ABCC'} = \alpha_{(ABC)C'}$$

and $\alpha_{AC'}$ is defined by

$$\alpha_{AC'} = A_{AB}{}^B{}_{C'}. \tag{2.22}$$

[7] R. Penrose and W. Rindler, *Application of Spinors to Relativity* (Cambridge University Press, London and New York, to be published).

[8] E. Newman and R. Penrose, J. Math. Phys. **3**, 566 (1962).

The spinor $\alpha_{AC'}$ can be split into Hermitian and anti-Hermitian parts:

$$\alpha_{AC'} = k_{AC'} - i\mu_{AC'} \tag{2.23}$$

where

$$k_{AC} = \tfrac{1}{2}(\alpha_{AC} + \bar{\alpha}_{C'A}), \quad \mu_{AC'} = \tfrac{1}{2}i(\alpha_{AC'} - \bar{\alpha}_{C'A}) \tag{2.24}$$

and

$$\overline{k_{AC'}} = \bar{k}_{A'C} = k_{CA'}, \quad \overline{\mu_{AC'}} = \bar{\mu}_{A'C} = \mu_{CA'}. \tag{2.25}$$

The following correspondences hold:

$$\begin{aligned} k_a &= A_{fa}{}^{f} \leftrightarrow k_{AA'} \\ \mu_a &= A^{*}_{fa}{}^{f} = \tfrac{1}{2}A_{pq}{}^{f}\eta_{fa}{}^{pq} \leftrightarrow \mu_{AA'}. \end{aligned} \tag{2.26}$$

This can be seen easily using the definitions and the spinor equivalent of η_{abcd} [see (1.28)] given by

$$\begin{aligned} &\eta_{abcd} \leftrightarrow \eta_{AA'BB'CC'DD'} \\ &= -i\{\epsilon_{AC}\epsilon_{BD}\epsilon_{A'D'}\epsilon_{B'C'} - \epsilon_{AD}\epsilon_{BC}\epsilon_{A'C'}\epsilon_{B'D'}\} \end{aligned} \tag{2.27}$$

(See Ref. 7). The tetrad tensor C_{abc} [see (1.21)] has the same skew symmetry in its first two indices as A_{abc}; therefore the spinor $C_{AA'BB'CC'} \leftrightarrow C_{abc}$ can be split similarly:

$$C_{AA'BB'CC'} = \tfrac{1}{2}(C_{ABCC'}\epsilon_{A'B'} + \bar{C}_{A'B'C'C}\epsilon_{AB}) \tag{2.28}$$

and the "structure-constant spinor" $C_{ABCC'}$ is symmetric in its first two indices. Using (1.20) we get the expression

$$C_{ABCC'} = A_{C(AB)C'} + \bar{\alpha}_{C'(A}\epsilon_{B)C}. \tag{2.29}$$

We see later that in our algebraic problem the components of the spinor $A_{ABCC'}$ are the natural unknowns. We proceed to write down the equations which have to be fulfilled by these quantities. Since R_{abcd} are real and have the symmetry properties

$$R_{abcd} = -R_{bacd} = -R_{abdc},$$

the spinor $R_{AA'BB'CC'DD'}$ can be split[5] as follows:

$$\begin{aligned} R_{AA'BB'CC'DD'} &= \tfrac{1}{2}\{\chi_{ABCD}\epsilon_{A'B'}\epsilon_{C'D'} + \epsilon_{CD}\phi_{ABC'D'}\epsilon_{A'B'} \\ &+ \epsilon_{AB}\bar{\phi}_{A'B'CD}\epsilon_{C'D'} + \epsilon_{AB}\epsilon_{CD}\bar{\chi}_{A'B'C'D'}\} \end{aligned} \tag{2.30}$$

and the spinors χ_{ABCD} and $\phi_{ABC'D'}$ (curvature spinors) have the symmetry properties

$$\chi_{ABCD} = \chi_{BACD} = \chi_{ABDC}, \quad \phi_{ABC'D'} = \phi_{BAC'D'} = \phi_{ABD'C'}. \tag{2.31}$$

Applying this splitting process to the expression

$$\begin{aligned} R_{AA'BB'CC'DD'} &= A_{PP'AA'CC'}A^{PP'}{}_{BB'DD'} \\ &- A_{PP'AA'DD'}A^{PP'}{}_{BB'CC'} + A_{AA'BB'PP'}C_{CC'DD'}{}^{PP'}, \end{aligned}$$

which corresponds to (1.27), we get for the curvature spinors

$$\chi_{ABCD} = \tfrac{1}{4}(A_{PACP'}A^{P}{}_{BD}{}^{P'} + A_{PADP'}A^{P}{}_{BC}{}^{P'}) + \tfrac{1}{2}A_{ABPP'}C_{CD}{}^{PP'} \tag{2.32}$$

$$\phi_{ABC'D'} = \tfrac{1}{4}(A_{PAQC'}A^{P}{}_{B}{}^{Q}{}_{D'} + A_{PAQD'}A^{P}{}_{B}{}^{Q}{}_{C'}) + \tfrac{1}{2}A_{ABPP'}\bar{C}_{C'D'}{}^{P'P}. \tag{2.33}$$

[Naturally the right-hand sides satisfy (2.31).] [See Ref. 5]

Penrose proves[5] that the cyclic symmetry of the Riemann tensor is equivalent to the following symmetry properties of the curvature spinors

$$\chi_{ABCD} = \chi_{CDAB} \tag{2.34}$$

$$\lambda = \tfrac{1}{2}\chi_{AB}{}^{AB} = \tfrac{1}{2}\bar{\chi}_{A'B'}{}^{A'B'} = \bar{\lambda} \tag{2.35}$$

$$\phi_{ABC'D'} = \bar{\phi}_{C'D'AB}. \tag{2.36}$$

According to our remark in 1, the cyclic symmetry of (1.27) is equivalent to the Jacobi identities. Consequently the Jacobi identities are equivalent to the requirement that the Ricci rotation spinor fulfills the equations which we get if we insert the expressions (2.32) and (2.33) into (2.34)–(2.36). In other words (2.34)–(2.36) are the Jacobi identities in spinor form. [One can prove this statement by a straightforward calculation translating the equation

$$R^{*}{}_{abc}{}^{b} = \tfrac{1}{2}R_{abfg}\eta_{c}{}^{bfg} = 0$$

into spinors, i.e.,

$$R^{*}{}_{AA'BB'CC'}{}^{BB'} = \tfrac{1}{2}R_{AA'BB'FF'GG'}\eta_{CC'}{}^{BB'FF'GG'}$$

and using (2.27) and (2.30).]

Before proceeding with the field equations we rewrite (2.34) in a more convenient form. We want first to show that (2.34) is equivalent to

$$\chi_{ABC}{}^{B} + \chi_{CBA}{}^{B} = 0. \tag{2.37}$$

We see that

$$\chi_{ABC}{}^{B} + \chi_{CBA}{}^{B} = (\chi_{ABCD} + \chi_{CBAD})\epsilon^{BD} = 0$$

and this is equivalent to

$$\chi_{ABCD} + \chi_{CBAD} - \chi_{ADCB} - \chi_{CDAB} = 0.$$

Writing down this equation with the indices $BADC$ instead of $ABCD$ we have

$$\chi_{BADC} + \chi_{DABC} - \chi_{BCDA} - \chi_{DCBA} = 0.$$

Adding up these two equations and using (2.31) we get (2.34). This argument also works backwards and our assertion is proved.

Now inserting (2.32) into (2.37) and using (2.29) we get

$$\chi_{ABC}{}^{B} + \chi_{CBA}{}^{B} = C_{ACPP'}k^{PP'} = 0. \tag{2.38}$$

We refer to the equations (2.38), (2.35), and (2.36) as Jacobi identities.

We now write down the field equations in spinor form. The spinor

$$R_{BB'DD'} = \epsilon^{AC}\epsilon^{A'C'}R_{AA'BB'CC'DD'}$$

can be found using (2.30):

$$R_{BB'DD'} = \tfrac{1}{2}(\chi_{AB}{}^{A}{}_{D}\epsilon_{B'D'} - \phi_{BDB'D'} - \bar{\phi}_{B'D'BD} + \epsilon_{BD}\bar{\chi}_{A'B'}{}^{A'}{}_{D'}) \quad (2.39)$$

(see Ref. 5). It is easy to show from (2.34) and (2.35) that

$$\chi_{AB}{}^{A}{}_{D} = \lambda\epsilon_{BD}. \quad (2.40)$$

Using (2.35), (2.36), and (2.40) we get

$$R_{BB'DD'} = -\phi_{BDB'D'} + \lambda\epsilon_{BD}\epsilon_{B'D'}. \quad (2.41)$$

The right-hand side of (2.1), written in spinor form, is

$$-Au_{BB'}u_{DD'} + (\Lambda + \tfrac{1}{2}A)\epsilon_{BD}\epsilon_{B'D'}, \quad (2.42)$$

where $u_a \leftrightarrow u_{AA'}$ and the equation $u_a u^a = 1$ has the spinor form

$$u_{AA'}u^{AA'} = 1. \quad (2.43)$$

The expression (2.42) can be written in the form

$$-\tfrac{1}{2}A(u_{BB'}u_{DD'} + u_{BD'}u_{DB'}) - \tfrac{1}{2}A(u_{BB'}u_{DD'} - u_{BD'}u_{DB'}) + (\Lambda + \tfrac{1}{2}A)\epsilon_{BD}\epsilon_{B'D'}. \quad (2.44)$$

From (2.43) we see easily that

$$\tfrac{1}{2}(u_{BB'}u_{DD'} - u_{BD'}u_{DB'}) = \tfrac{1}{4}\epsilon_{BD}\epsilon_{B'D'}. \quad (2.45)$$

Combining these facts we get

$$-\phi_{BDB'D'} + \lambda\epsilon_{BD}\epsilon_{B'D'} = -\tfrac{1}{2}A(u_{BB'}u_{DD'} + u_{BD'}u_{DB'}) + (\Lambda + \tfrac{1}{4}A)\epsilon_{BD}\epsilon_{B'D'}.$$

Thus the field equations in spinor form are

$$\lambda = \Lambda + \tfrac{1}{4}A \quad (2.46)$$

$$\phi_{ABA'B'} = \tfrac{1}{2}A(u_{AA'}u_{BB'} + u_{AB'}u_{BA'}). \quad (2.47)$$

We now observe the following facts. If (2.46) is satisfied then (2.35) is also satisfied, since Λ and A are real. If (2.47) is satisfied then (2.36) is also satisfied, since the right-hand side of (2.47) is Hermitian ($u_{AA'}$ being Hermitian). *The spinor form of the basic equations* is therefore:

$$C_{ABCC'}k^{CC'} = 0 \quad (2.48a)$$

$$\Lambda + \tfrac{1}{4}A = \tfrac{1}{2}\chi_{AB}{}^{AB} \quad (2.48b)$$

$$\phi_{ABA'B'} = \tfrac{1}{2}A(u_{AA'}u_{BB'} + u_{AB'}u_{BA'}) \quad (2.48c)$$

$$u_{AA'}u^{AA'} = 1. \quad (2.48d)$$

Our algebraic problem is to solve these equations for the "unknowns" $A_{ABCC'}$, A, Λ, and $u_{AA'}$.

As a final preliminary we translate our previous equations (1.25), (1.39), (1.41), (1.42), (1.44), and (1.45) into their spinor equivalents. These are found to be, respectively,

$$u_{AA'}{}_{:BB'} = \tfrac{1}{2}(u^{P}{}_{A'}A_{PABB'} + u_{A}{}^{P'}\bar{A}_{P'A'B'B}), \quad (2.49)$$

$$u^{AA'}{}_{:AA'} = -u_{AA'}k^{AA'} = 0, \quad (2.50)$$

$$u_{AA'}{}_{:BB'}u^{BB'} = \tfrac{1}{2}(u^{P}{}_{A'}A_{PABB'} + u_{A}{}^{P'}\bar{A}_{P'A'B'B})u^{BB'} = 0, \quad (2.51)$$

$$\sigma_{AA'BB'} = \tfrac{1}{4}\{u^{P}{}_{A'}A_{PABB'} + u_{A}{}^{P'}\bar{A}_{P'A'B'B} + u^{P}{}_{B'}A_{PBAA'} + u_{B}{}^{P'}\bar{A}_{P'B'A'A}\}. \quad (2.52)$$

[$\sigma_{AA'BB'}$ is not to be confused with the σ's of (2.5) and (2.6).] The spinor $\sigma_{AA'BB'}$ has the following symmetry property

$$\sigma_{AA'BB'} = \sigma_{ABA'B'} = \sigma_{(AB)(A'B')} \quad (2.53)$$

since the trace vanishes. [See (1.42).] The spinor $\omega_{AA'BB'} \leftrightarrow \omega_{ab}$ splits according to the pattern (2.16) ($\omega_{ab} = -\omega_{ba}$). The corresponding symmetric spinor ω_{AB} is given by

$$\omega_{AB} = -\tfrac{1}{2}C_{ABCC'}u^{CC'}. \quad (2.54)$$

The spinor $\omega^{AA'} \leftrightarrow \omega^a$ [defined by (1.44)] has the form

$$\omega^{AA'} = \tfrac{1}{2}i\{\omega^{AB}\epsilon^{A'B'} - \bar{\omega}^{A'B'}\epsilon^{AB}\}u_{BB'}. \quad (2.55)$$

3. THE EXPLICIT ALGEBRAIC SYSTEM

In writing down our equations explicitly we can still use the freedom that the tetrad is given only up to the transformations of the homogeneous Lorentz group. Thus we are at liberty to assume that the tetrad components of u_i with respect to the orthonormal tetrad $e^a{}_i$ have the form

$$u_a = (1, 0, 0, 0). \quad (3.1)$$

It follows that the corresponding spinor is then given by

$$u_{AA'} = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}, \quad (3.2)$$

and (2.48d) is satisfied. We denote our unknowns, the components of $A_{ABCC'}$, by the following symbols:

$A_{ABCC'}$:

AB \ CC'	11′	12′	21′	22′
11	κ	σ	ρ	τ
(12)	ϵ	β	α	γ
22	π	μ	λ	ν

(3.3)

(See Ref. 8). Using (3.2) and (3.3) we get from (2.51) the equations

$$\gamma + \bar{\gamma} + \epsilon + \bar{\epsilon} = 0, \tag{3.4a}$$

$$\kappa - \bar{\pi} + \tau - \bar{\nu} = 0. \tag{3.4b}$$

The components of the spinors $k_{AA'}$ and $\mu_{AA'}$ defined by (2.24) are

$$k_{AA'} = \begin{bmatrix} \frac{1}{2}(\rho+\bar{\rho})-\frac{1}{2}(\epsilon+\bar{\epsilon}) & \frac{1}{2}(\tau-\beta)+\frac{1}{2}(\bar{\alpha}-\bar{\pi}) \\ \frac{1}{2}(\bar{\tau}-\bar{\beta})+\frac{1}{2}(\alpha-\pi) & \frac{1}{2}(\gamma+\bar{\gamma})-\frac{1}{2}(\mu+\bar{\mu}) \end{bmatrix}, \tag{3.5}$$

$$\mu_{AA'} = \begin{bmatrix} \frac{1}{2}i(\rho-\bar{\rho})-\frac{1}{2}i(\epsilon-\bar{\epsilon}) & \frac{1}{2}i(\tau-\beta)-\frac{1}{2}i(\bar{\alpha}-\bar{\pi}) \\ -\frac{1}{2}i(\bar{\tau}-\bar{\beta})+\frac{1}{2}i(\alpha-\pi) & \frac{1}{2}i(\gamma-\bar{\gamma})-\frac{1}{2}i(\mu-\bar{\mu}) \end{bmatrix}. \tag{3.6}$$

Note: Whenever in this paper we exhibit a matrix it will be understood that the first or upper index denotes the row and the other the column. Looking at the equation (2.48a) and (2.50) we see that we must distinguish between two different cases, namely

$$\text{(i)} \quad k_{AA'} \neq 0, \qquad \text{(ii)} \quad k_{AA'} = 0 \tag{3.7}$$

which we shall discuss separately.

Case (i): $k_{AA'} \neq 0$. Equation (2.50) tells us that k_a is perpendicular to u_a. Using Lorentz transformations we can arrange that k_a has with respect to the orthonormal tetrad $e^a{}_i$ the components

$$k_a = (0, 0, 0, \sqrt{2}\, r), \qquad r \neq 0 \tag{3.8}$$

whence the corresponding spinor is given by

$$k_{AA'} = \begin{pmatrix} r & 0 \\ 0 & -r \end{pmatrix}. \tag{3.9}$$

Comparing (3.5) and (3.9) we see that this choice of the tetrad allows us to impose the conditions

$$(\rho + \bar{\rho}) - (\epsilon + \bar{\epsilon}) = 2r, \tag{3.10a}$$

$$(\gamma + \bar{\gamma}) - (\mu + \bar{\mu}) = -2r, \tag{3.10b}$$

$$\tau - \beta + \bar{\alpha} - \bar{\pi} = 0, \tag{3.10c}$$

without loss of generality. The equation (2.50) is fulfilled. Using (2.29) and (3.8) we get from (2.48a) the equations

$$\rho - \bar{\rho} + \epsilon + \bar{\epsilon} = \mu - \bar{\mu} + \gamma + \bar{\gamma} \tag{3.11a}$$

$$\kappa = \bar{\alpha} + \beta - \bar{\pi} \tag{3.11b}$$

$$\nu = \alpha + \bar{\beta} - \bar{\tau}. \tag{3.11c}$$

By now we have fixed our orthonormal tetrad $e^a{}_i$ up to rotations in the 1–2 plane. Using this freedom we can annihilate either the μ_1, or the μ_2 component of μ_a. In other words, looking at (3.6), we can require that $i\{\tau - \beta - (\bar{\alpha} - \bar{\pi})\}$ is either imaginary or real, i.e.,

$$(\tau - \beta) - (\bar{\alpha} - \bar{\pi}) = (\bar{\tau} - \bar{\beta}) - (\alpha - \pi) \tag{3.12a}$$

or

$$(\tau - \beta) - (\bar{\alpha} - \bar{\pi}) = -(\bar{\tau} - \bar{\beta}) + (\alpha - \pi). \tag{3.12b}$$

We now proceed to solve these equations. From (3.4a), (3.10a), (3.10b), and (3.11a) we get

$$\gamma + \bar{\gamma} = 0, \qquad \epsilon + \bar{\epsilon} = 0, \tag{3.13a}$$

$$\mu = \rho, \tag{3.13b}$$

$$\mathrm{Re}\,(\rho) = r \neq 0. \tag{3.13c}$$

From (3.4b), (3.10c), (3.11b), and (3.11c) we get

$$\pi = \bar{\tau}, \tag{3.14a}$$

$$\beta = \bar{\alpha}, \tag{3.14b}$$

$$\kappa = 2\bar{\alpha} - \tau, \tag{3.14c}$$

$$\nu = 2\alpha - \bar{\tau}. \tag{3.14d}$$

From (3.12) we get

$$\tau - \bar{\tau} = -(\alpha - \bar{\alpha}) \tag{3.15a}$$

or

$$\tau + \bar{\tau} = \alpha + \bar{\alpha}. \tag{3.15b}$$

Summarizing our results, we have

$A_{ABCC'}$:

AB \ CC'	11′	12′	21′	22′
11	$2\bar{\alpha} - \tau$	σ	ρ	τ
(12)	ϵ	$\bar{\alpha}$	α	γ
22	$\bar{\tau}$	ρ	λ	$2\alpha - \bar{\tau}$

(3.16)

where

$$\gamma + \bar{\gamma} = 0, \quad \epsilon + \bar{\epsilon} = 0, \quad \mathrm{Re}\,(\rho) = r \neq 0$$

and

$$\tau - \bar{\tau} + \alpha - \bar{\alpha} = 0, \quad \text{or} \quad \tau + \bar{\tau} = \alpha + \bar{\alpha}. \tag{3.17}$$

Using (2.29), (2.32), (2.33), (3.16), and (3.17) we get from (2.48b) and (2.48c) the following equations for our problem in case (i):

$$8\alpha\bar{\alpha} - 2(\alpha\tau + \overline{\alpha\tau}) - \sigma\bar{\sigma} - \rho^2 = A, \tag{3.18a}$$

$$8\alpha\bar{\alpha} - 2(\alpha\tau + \overline{\alpha\tau}) - \lambda\bar{\lambda} - \rho^2 = A, \tag{3.18b}$$

$$4\alpha\bar{\alpha} - 4(\alpha\tau + \overline{\alpha\tau}) + \sigma\lambda - \rho^2 = A, \tag{3.18c}$$

$$\sigma(4\gamma - \rho) + 4\bar{\alpha}(\bar{\alpha} - \tau) - \rho\bar{\lambda} = 0, \tag{3.18d}$$

$$\lambda(4\epsilon - \rho) + 4\alpha(\alpha - \bar{\tau}) - \rho\bar{\sigma} = 0, \tag{3.18e}$$

$$(2\bar{\alpha} - \tau)\gamma - \rho(\bar{\alpha} + \tau) + \sigma(\alpha - \bar{\tau}) - \epsilon\tau = 0, \tag{3.18f}$$

$$(2\alpha - \bar{\tau})\epsilon - \rho(\alpha + \bar{\tau}) + \lambda(\bar{\alpha} - \tau) - \gamma\bar{\tau} = 0, \tag{3.18g}$$

$$(2\bar{\alpha} - \tau)\lambda + (\gamma + \epsilon)(2\alpha - \bar{\tau}) + \alpha(2\epsilon - \rho) - \rho\bar{\tau} - \bar{\alpha}\bar{\sigma} = 0, \tag{3.18h}$$

$$(2\alpha - \bar{\tau})\sigma + (\gamma + \epsilon)(2\bar{\alpha} - \tau) + \bar{\alpha}(2\gamma - \rho) - \rho\tau - \alpha\lambda = 0, \tag{3.18i}$$

$$\Lambda + \tfrac{1}{4}A = \tfrac{1}{4}\{4\alpha\bar{\alpha} - \sigma\lambda - 3\rho^2\}. \tag{3.18j}$$

Note that from (3.18a) and the condition that A has to be real, it follows that $\rho^2 - \bar{\rho}^2 = (\rho + \bar{\rho})(\rho - \bar{\rho}) = 0$; and since $\rho + \bar{\rho} = 2r \neq 0$, we have that $\rho = \bar{\rho} = r$. This fact has been used in the derivation of the later equations in (3.18).

Case (ii): ($k_{AA'} = 0$). From (3.5) and (3.7–ii) we get the equations

$$\rho + \bar{\rho} = \epsilon + \bar{\epsilon}, \tag{3.19a}$$

$$\gamma + \bar{\beta} = \mu + \bar{\mu}, \tag{3.19b}$$

$$\tau - \beta + \bar{\alpha} - \bar{\pi} = 0. \tag{3.19c}$$

Using Lorentz transformations we can arrange that

$$\tau - \beta - (\bar{\alpha} - \bar{\pi}) = 0, \tag{3.20}$$

which means that μ_a has the form $\mu_a = (\mu_0, 0, 0, \mu_3)$. This fixes our orthonormal tetrad $e^a{}_i$ up to rotations in the 1–2 plane. We still have the freedom to use these rotations, i.e., spinor transformations of the form of

$$L^A{}_B = \begin{pmatrix} e^{i\varphi} & 0 \\ 0 & e^{-i\varphi} \end{pmatrix} \tag{3.21}$$

to simplify $A_{ABCC'}$ at a convenient stage of our calculations. From (3.4a), (3.19a), and (3.19b) we have

$$\rho + \bar{\rho} = \epsilon + \bar{\epsilon} = -(\gamma + \bar{\gamma}) = -(\mu + \bar{\mu}) \tag{3.22}$$

and from (3.19c), (3.20), and (3.4b)

$$\tau = \beta, \tag{3.23a}$$

$$\pi = \alpha, \tag{3.23b}$$

$$\kappa - \bar{\nu} = \bar{\alpha} - \beta. \tag{3.23c}$$

Collecting our results we have

$A_{ABCC'}$:

AB	CC': 11′	12′	21′	22′
11	κ	σ	ρ	β
(12)	ϵ	β	α	γ
22	α	μ	λ	ν

(3.24)

with (3.22) and (3.23c). From (2.48b) and (2.48c) we get the following equations for our problem in case (ii):

$$(2\alpha + \bar{\beta})\kappa + \beta\bar{\kappa} - \sigma\bar{\sigma} - \rho(\rho + \epsilon + \bar{\epsilon}) = A, \tag{3.25a}$$

$$(2\beta + \bar{\alpha})\nu + \alpha\bar{\nu} - \lambda\bar{\lambda} - \mu(\mu + \gamma + \bar{\gamma}) = A, \tag{3.25b}$$

$$\kappa\nu - \rho\mu + \sigma\lambda - \alpha\beta + 2\epsilon\bar{\mu} + 2\gamma\bar{\rho} - 2\alpha\bar{\alpha} - 2\beta\bar{\beta} = A, \tag{3.25c}$$

$$\kappa\bar{\nu} - \rho\bar{\lambda} - \sigma(\mu + \bar{\gamma} - 3\gamma) + \beta(\bar{\alpha} - 2\beta) = 0, \tag{3.25d}$$

$$\nu\bar{\kappa} - \bar{\sigma}\mu - \lambda(\rho + \bar{\epsilon} - 3\epsilon) + \alpha(\bar{\beta} - 2\alpha) = 0, \tag{3.25e}$$

$$\kappa(\gamma + \bar{\mu}) - \rho(\bar{\alpha} + \beta) + \sigma(\alpha - \bar{\beta}) + \beta(\bar{\rho} - \epsilon) = 0, \tag{3.25f}$$

$$\nu(\epsilon + \bar{\rho}) - \mu(\alpha + \bar{\beta}) + \lambda(\beta - \bar{\alpha}) + \alpha(\bar{\mu} - \gamma) = 0, \tag{3.25g}$$

$$\kappa\lambda + \epsilon\bar{\beta} - \beta\bar{\sigma} - \alpha(2\rho - \epsilon + \bar{\epsilon}) + \gamma\bar{\kappa} = 0, \tag{3.25h}$$

$$\nu\sigma + \gamma\bar{\alpha} - \alpha\bar{\lambda} - \beta(2\mu - \gamma + \bar{\gamma}) + \epsilon\bar{\nu} = 0, \tag{3.25i}$$

$$\Lambda + \tfrac{1}{4}A = \tfrac{1}{4}\{\kappa\nu + 3\alpha\beta - \sigma\lambda + \rho\mu - 2\epsilon\mu - 2\gamma\rho\}. \tag{3.25j}$$

4. SOLUTIONS OF THE ALGEBRAIC SYSTEMS

In this section we obtain *all* the solutions of the equations (3.18) and *some* solutions of (3.25). In the next sections we give explicitly the corresponding line elements, which represent two new and two already known families of solutions of Einstein's field equations with incoherent matter.

Case (i): *Equations (3.18)*. We have seen that

$$\rho = \bar{\rho} = r. \tag{4.1}$$

From $A > 0$, (3.18a) and (4.1) it follows that

$$\alpha \neq 0, \tag{4.2}$$

and from (3.18a) (3.18b), and (3.18c) that

$$\sigma\bar{\sigma} = \lambda\bar{\lambda} \quad \text{and} \quad \sigma\lambda = \bar{\sigma}\bar{\lambda}. \tag{4.3}$$

We now observe that

$$\sigma \neq 0 \quad \text{and} \quad \lambda \neq 0. \tag{4.4}$$

For $\sigma = 0$ is equivalent to $\lambda = 0$ [see (4.3)]; using (3.18e) and (4.2) we get $\tau = \bar{\alpha}$, and from (3.18c) it follows that $A = -4\alpha\bar{\alpha} - r^2$, which violates the condition $A > 0$. The equations (4.3) have two solutions:

$$\lambda = -\bar{\sigma} \tag{4.5}$$

and

$$\lambda = \bar{\sigma}. \tag{4.6}$$

Equation (3.18) has no solutions in the case of (4.5). To show this, we substitute (4.5) into (3.18), using (4.2) and (4.4), and get the following system:

$$A = 12\alpha\bar{\alpha} - \sigma\bar{\sigma} - r^2, \qquad \Lambda + \tfrac{1}{4}A = \tfrac{1}{4}(4\alpha\bar{\alpha} + \sigma\bar{\sigma} - 3r^2),$$

$$\epsilon = \gamma, \qquad \tau = -\bar{\alpha},$$

$$\sigma\gamma + 2\bar{\alpha}^2 = 0, \quad \text{and} \quad 2\bar{\alpha}\gamma + \alpha\sigma = 0,$$

which is not solvable under the condition $A > 0$: For

$$\alpha(\sigma\gamma + 2\bar{\alpha}^2) = (\alpha\sigma)\gamma + 2\bar{\alpha}(\alpha\bar{\alpha}) = 2\bar{\alpha}(-\gamma^2 + \alpha\bar{\alpha})$$
$$= 2\bar{\alpha}(\gamma\bar{\gamma} + \alpha\bar{\alpha}) = 0 \rightarrow \alpha = 0,$$

which contradicts (4.2).

Thus we are left with only the possibility (4.6). Inserting (4.6) into (3.18) and using (4.2) and (4.4), we get the following system:

$$\epsilon = -\gamma, \tag{4.7a}$$

$$\sigma(2\gamma - r) + 2\bar{\alpha}(\bar{\alpha} - \tau) = 0, \tag{4.7b}$$

$$\sigma(\alpha - \bar{\tau}) + \bar{\alpha}(2\gamma - r) - r\tau = 0, \tag{4.7c}$$

$$\sigma\bar{\sigma} = 2\alpha\bar{\alpha} + \alpha\tau + \bar{\alpha}\bar{\tau}, \tag{4.7d}$$

$$A = 6\alpha\bar{\alpha} - 3(\alpha\tau + \bar{\alpha}\bar{\tau}) - r^2, \tag{4.7e}$$

$$\Lambda = \tfrac{1}{2}\{-2\alpha\bar{\alpha} + \alpha\tau + \bar{\alpha}\bar{\tau} - r^2\}. \tag{4.7f}$$

Using (3.15a) in the form $\tau - \bar{\alpha} = \bar{\tau} - \alpha = t$, i.e.,

$$\tau = \bar{\alpha} + t, \tag{4.8}$$

we get the equations

$$(2\gamma - r)\sigma - 2t\bar{\alpha} = 0, \tag{4.9a}$$

$$-t\sigma + 2(\gamma - r)\bar{\alpha} = rt, \tag{4.9b}$$

$$\sigma\bar{\sigma} = 4\alpha\bar{\alpha} + t(\alpha + \bar{\alpha}), \tag{4.9c}$$

$$A = -3(\alpha + \bar{\alpha})t - r^2, \tag{4.9d}$$

$$\Lambda = \tfrac{1}{2}\{(\alpha + \bar{\alpha})t - r^2\}. \tag{4.9e}$$

We observe that (4.9a) and (4.9b) is a linear system in the quantities σ and $\bar{\alpha}$. Since $t \neq 0$ because of (4.9d) the determinant $d = 2[(r - \gamma)(r - 2\gamma) - t^2]$ must be different from zero. The solution of this system is given by

$$\sigma = \frac{rt^2}{(r - \gamma)(r - 2\gamma) - t^2},$$
$$\alpha = -\frac{rt}{2}\frac{r + 2\gamma}{(r + \gamma)(r + 2\gamma) - t^2}. \tag{4.10}$$

One sees by substitution that (4.9c) is satisfied. From (4.9d) and (4.9e) we get

$$A = r^2\left\{\frac{3t^2[(r+2\gamma)(r-2\gamma)-t^2]}{[(r+\gamma)(r+2\gamma)-t^2][(r-\gamma)(r-2\gamma)-t^2]} - 1\right\} \tag{4.11}$$

and

$$\Lambda = -\frac{r^2}{2}\left\{\frac{t^2[(r+2\gamma)(r-2\gamma)-t^2]}{[(r+\gamma)(r+2\gamma)-t^2][(r-\gamma)(r-2\gamma)-t^2]}+1\right\}. \tag{4.12}$$

Collecting our results we have

$A_{ABCC'}$:

AB \ CC'	11′	12′	21′	22′
11	$\bar{\alpha} - t$	σ	r	$t + \bar{\alpha}$
(12)	$-\gamma$	$\bar{\alpha}$	α	γ
22	$t + \alpha$	r	$\bar{\sigma}$	$\alpha - t$

(4.13)

where α and σ are given by (4.10) as functions of the parameters r, t and γ (γ is imaginary). The range of the parameters is determined by the condition $A > 0$. (4.13) gives all the solutions of (3.18). Using (2.29) we can calculate the components of $C_{ABCC'}$; then, with the help of (2.28) those of $C_{AA'BB'CC'}$; and finally, using (2.11), we get C_{abc}. The nonvanishing components of $C_{ab}{}^{c}$ are shown in the following table:

$$C^{\beta}{}_{\delta} = C_{\delta 3}{}^{\beta} = \begin{bmatrix} 0 & -(\alpha + \bar{\alpha} - 2t)/\sqrt{2} & -i(\alpha - \bar{\alpha})/\sqrt{2} \\ (\alpha + \bar{\alpha} + 2t)/\sqrt{2} & (2r + \sigma + \bar{\sigma})/2\sqrt{2} & i[4\gamma - (\sigma - \bar{\sigma})]/2\sqrt{2} \\ i(\alpha - \bar{\alpha})/\sqrt{2} & -i(4\gamma + \sigma - \bar{\sigma})/2\sqrt{2} & [2r - (\sigma + \bar{\sigma})]/2\sqrt{2} \end{bmatrix} \tag{4.14}$$

(β and δ have the values 0, 1, 2). We see that the corresponding groups always have an Abelian invariant subgroup. The shear spinor defined by (2.52) has the components

$\sigma_{ABA'B'}$:

AB \ $A'B'$	1′1′	(1′2′)	2′2′
11	0	$-\tau/2\sqrt{2}$	0
(12)	$-\bar{\tau}/2\sqrt{2}$	0	$\tau/2\sqrt{2}$.
22	0	$\bar{\tau}/2\sqrt{2}$	0

(4.15)

Thus τ is responsible for the shear. The spinor ω_{AB} defined by (2.54) has the components

ω_{AB}:

A \ B	1	2
1	$-(\bar{\alpha} - t)/\sqrt{2}$	0
2	0	$-(\alpha - t)/\sqrt{2}$

(4.16)

and thus $\alpha - t$ is responsible for the rotation. We note that our solutions *always have rotation* since $t = \alpha$ would imply, because of (4.9d), that $A = -6t^2 - r^2 < 0$! The components of the spinors $k_{AA'}$, $\mu_{AA'}$, and $\omega^{AA'}$ [see (2.55)] are

$$k_{AA'} = \begin{bmatrix} r & 0 \\ 0 & -r \end{bmatrix}, \qquad \mu_{AA'} = \begin{bmatrix} i\gamma & it \\ -it & i\gamma \end{bmatrix},$$
$$\omega^{AA'} = \begin{bmatrix} 0 & \frac{1}{2}i(\alpha - t) \\ -\frac{1}{2}i(\bar{\alpha} - t) & 0 \end{bmatrix}. \tag{4.17}$$

Defining the eigenvalues ${}_{(a)}\sigma$ and the eigenspinors ${}_{(a)}\sigma^{AA'}$ of $\sigma_{AA'BB'}$ by the equations

$$\sigma^{AA'}{}_{BB'}\,{}_{(a)}\sigma^{BB'} = {}_{(a)}\sigma\,{}_{(a)}\sigma^{AA'} \tag{4.18}$$

($a = 0, 1, 2, 3$), we get after a short calculation

$${}_{(0)}\sigma = 0, \quad {}_{(1)}\sigma = 0, \quad {}_{(2)}\sigma = 2\,|\tau|, \quad {}_{(3)}\sigma = -2\,|\tau| \tag{4.19}$$

and, correspondingly,

$$\begin{aligned} {}_{(0)}\sigma^{AA'} &= \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = u^{AA'}, \\ {}_{(1)}\sigma^{AA'} &= \frac{i}{\sqrt{2}}\begin{pmatrix} 0 & e^{-i\varphi} \\ -e^{i\varphi} & 0 \end{pmatrix}, \\ {}_{(2)}\sigma^{AA'} &= \frac{1}{2}\begin{pmatrix} 1 & e^{-i\varphi} \\ e^{i\varphi} & -1 \end{pmatrix}, \\ {}_{(3)}\sigma^{AA'} &= \frac{1}{2}\begin{pmatrix} 1 & -e^{-i\varphi} \\ -e^{i\varphi} & -1 \end{pmatrix}, \end{aligned} \tag{4.20}$$

where $e^{i\varphi} = \tau/|\tau|$. These vectors are mutually perpendicular. [${}_{(0)}\sigma^{AA'}$ and ${}_{(1)}\sigma^{AA'}$ are not uniquely determined because of the multiplicity of the eigenvalue 0.] We see that

$$k_{AA'}\mu^{AA'} = 0, \tag{4.21}$$

which means that k_a and μ_a are always perpendicular to each other.

We now consider the special case

$$\gamma = 0. \tag{4.22}$$

Here α, σ, and τ are real [see (4.10)]. Using instead of the parameters r and t the new parameters a and s defined by the equations

$$r = a(2 - s^2)/\sqrt{2} \text{ and } t = as(2 - s^2)/2 \tag{4.23}$$

we get, from (4.14), (4.11), and (4.12),

$$C^{\beta}{}_{\delta} = C_{\delta 3}{}^{\beta} = \begin{bmatrix} 0 & as(3-s^2)/\sqrt{2} & 0 \\ as(1-s^2)/\sqrt{2} & a & 0 \\ 0 & 0 & a(1-s^2) \end{bmatrix} \tag{4.24}$$

and

$$A = a^2(2 - s^2)(2s^2 - 1), \quad \Lambda = -\tfrac{1}{2}a^2(2 - s^2),$$
$$\text{(therefore } \tfrac{1}{2} \leq s^2 \leq 2). \tag{4.25}$$

We see that in this case μ_a is perpendicular to u_a and parallel to ω^a and ${}_{(1)}\sigma^a$. This means that the rotation vector is an eigenvector of the shear tensor corresponding to the eigenvalue 0, which is not true in general. The line element for this case is obtained in Sec. 6.

Case (ii): Equations (3.25). We shall deal only with the special case

$$\kappa = 0, \quad \nu = 0, \quad \alpha = 0 \tag{4.26a}$$

$$\beta = 0, \quad \rho + \bar{\rho} = 0, \quad \epsilon + \bar{\epsilon} = 0, \quad \gamma + \bar{\gamma} = 0, \quad \mu + \bar{\mu} = 0. \tag{4.26b}$$

[Note that (4.26b) follows from (4.26a) and from (3.25).] Using the freedom (3.21), we can arrange that

$$\sigma + \bar{\sigma} = 0 \tag{4.27}$$

and

$$\lambda + \bar{\lambda} = 0 \tag{4.28}$$

as can be seen by the following reasoning: From (4.26), (4.27), (3.25c), and the fact that A has to be real, it follows that $\sigma(\lambda + \bar{\lambda}) = 0$. If σ is different from zero, then (4.28) follows; otherwise we can use (3.21) to make λ imaginary. Substituting (4.26)–(4.28) into (3.25) we get the following system:

$$(4\gamma - \mu)\sigma + \rho\lambda = 0, \tag{4.29a}$$
$$\mu\sigma + (4\epsilon - \rho)\lambda = 0, \tag{4.29b}$$
$$\sigma^2 - \rho^2 = A, \tag{4.29c}$$
$$\lambda^2 - \mu^2 = A, \tag{4.29d}$$
$$\sigma\lambda - \rho\mu - 2(\epsilon\mu + \gamma\rho) = A, \tag{4.29e}$$
$$\Lambda = \tfrac{1}{2}(\rho\mu - \sigma\lambda). \tag{4.29f}$$

The spinors $\sigma_{AB\,A'B'}$ [see (2.52)] and ω_{AB} [see (2.54)] have the nonvanishing components

$$\sigma_{112'2'} = \sigma_{221'1'} = (\sigma + \lambda)/2\sqrt{2},$$
$$\omega_{12} = -(\rho + \mu)/2\sqrt{2}. \tag{4.30}$$

Using $A > 0$ one sees from (4.29c) and (4.29d) that

$$\rho \neq 0 \text{ and } \mu \neq 0. \tag{4.31}$$

Considering (4.29a) and (4.29b) as a linear system for the quantities λ and σ with the determinant $d = 4(-4\epsilon\gamma + \gamma\rho + \epsilon\mu)$ we have two cases according to $d \neq 0$ and $d = 0$. If $d \neq 0$ we get $\lambda = \sigma = 0$ and the shear vanishes [see (4.30)]. We do not consider this case here: shear-free cases are discussed systematically in the last section. $d = 0$ gives the equation

$$4\epsilon\gamma = \epsilon\mu + \gamma\rho. \tag{4.32}$$

Because of (4.31) and (4.32), $\epsilon = 0$ is equivalent to $\gamma = 0$; therefore we have two different cases:

$$\gamma = 0 \text{ and } \epsilon = 0, \tag{4.33a}$$

$$\gamma \neq 0 \text{ and } \epsilon \neq 0, \tag{4.33b}$$

which we discuss separately. In case (4.33a), (4.29) reduces to the system

$$\rho\lambda - \sigma\mu = 0, \tag{4.34a}$$

$$\sigma\lambda - \rho\mu = A, \tag{4.34b}$$

$$A = \sigma^2 - \rho^2, \tag{4.34c}$$

$$A = \lambda^2 - \mu^2, \quad \Lambda = -\tfrac{1}{2}A. \tag{4.34d}$$

Considering (4.34a), (4.34b) as a system of linear equations for λ and μ with determinant given by (4.34c), we immediately find the solution

$$\lambda = \sigma, \ \mu = \rho \text{ and } A = \sigma^2 - \rho^2, \ \Lambda = -\tfrac{1}{2}A. \tag{4.35}$$

Introducing new parameters r and s by the equations

$$\rho = -ir \text{ and } \sigma = irs, \tag{4.36}$$

we get our results in the form

$A_{ABCC'}$:		CC'			
	AB	$11'$	$12'$	$21'$	$22'$
	11	0	irs	$-ir$	0
	(12)	0	0	0	0
	22	0	$-ir$	irs	0

(4.37)

$$A = r^2(1 - s^2) > 0 (-1 < s < 1) \text{ and } \Lambda = -\tfrac{1}{2}A. \tag{4.38}$$

The corresponding structure constant tensor has the nonvanishing components

$$C_{01}{}^{2} = \left(\frac{A}{2}\right)^{\frac{1}{2}} \frac{1 - s}{(1 - s^2)^{\frac{1}{2}}}, \quad C_{12}{}^{0} = -\left(\frac{A}{2}\right)^{\frac{1}{2}} \frac{2}{(1 - s^2)^{\frac{1}{2}}},$$

$$C_{20}{}^{1} = \left(\frac{A}{2}\right)^{\frac{1}{2}} \frac{1 + s}{(1 - s^2)^{\frac{1}{2}}}, \tag{4.39}$$

where we replaced the parameter r by A according to Eq. (4.38). The corresponding line element is obtained in Sec. 7.

We now discuss the case (4.33b). First we introduce a new parameter p by the equation

$$\Lambda = \tfrac{1}{2}Ap. \tag{4.40}$$

Using (4.40), (4.29), and (4.32) we obtain the system in the following form:

$$(4\gamma - \mu)\sigma + \rho\lambda = 0, \tag{4.41a}$$

$$\mu\sigma + (4\epsilon - \rho)\lambda = 0, \tag{4.41b}$$

$$\epsilon\mu + \gamma\rho = 4\epsilon\gamma (\neq 0), \tag{4.41c}$$

$$\epsilon\mu + \gamma\rho = -\tfrac{1}{2}A(p + 1), \tag{4.41d}$$

$$\rho\mu - \sigma\lambda = Ap, \tag{4.41e}$$

$$\lambda^2 - \mu^2 = A, \tag{4.41f}$$

$$\sigma^2 - \rho^2 = A. \tag{4.41g}$$

The solution of this system is

$$\sigma = -i\left(\frac{A}{2}\right)^{\frac{1}{2}} \{[(p + 1)(p - 1)]^{\frac{1}{2}} + [p(p - 1)]^{\frac{1}{2}}\},$$

$$\lambda = i\left(\frac{A}{2}\right)^{\frac{1}{2}} \{[(p + 1)(p - 1)]^{\frac{1}{2}} - [p(p - 1)]^{\frac{1}{2}}\},$$

$$\rho = -i\left(\frac{A}{2}\right)^{\frac{1}{2}} \{(p - 1) + [p(p + 1)]^{\frac{1}{2}}\},$$

$$\mu = -i\left(\frac{A}{2}\right)^{\frac{1}{2}} \{(p - 1) - [p(p + 1)]^{\frac{1}{2}}\},$$

$$\epsilon = \frac{i}{2}\left(\frac{A}{2}\right)^{\frac{1}{2}} \{(p + 1) + [p(p + 1)]^{\frac{1}{2}}\},$$

$$\gamma = \frac{i}{2}\left(\frac{A}{2}\right)^{\frac{1}{2}} \{(p + 1) - [p(p + 1)]^{\frac{1}{2}}\}, \tag{4.42}$$

where the parameters A and p have the range

$$A > 0 \text{ and } |p| \geq 1, \tag{4.43}$$

respectively. (4.42) gives the nonvanishing components of the spinor $A_{ABCC'}$. The corresponding structure constant tensor has the following nonvanishing components:

$$C_{01}{}^{2} = \tfrac{1}{2}A^{\frac{1}{2}}\{2p + [p(p - 1)]^{\frac{1}{2}}\},$$

$$C_{02}{}^{1} = -\tfrac{1}{2}A^{\frac{1}{2}}\{2p - [p(p - 1)]^{\frac{1}{2}}\},$$

$$C_{23}{}^{1} = \tfrac{1}{2}A^{\frac{1}{2}}\{2[p(p + 1)]^{\frac{1}{2}} - [(p + 1)(p - 1)]^{\frac{1}{2}}\},$$

$$C_{31}{}^{2} = \tfrac{1}{2}A^{\frac{1}{2}}\{2[p(p + 1)]^{\frac{1}{2}} + [(p + 1)(p - 1)]^{\frac{1}{2}}\},$$

$$C_{12}{}^{0} = -A^{\frac{1}{2}}(p - 1), \ C_{12}{}^{3} = A^{\frac{1}{2}}[p(p + 1)]^{\frac{1}{2}}. \tag{4.44}$$

The corresponding line elements are obtained in Sec. 8.

5. EXPLICIT CONSTRUCTION OF THE RECIPROCAL VECTORS

In the following three sections, we obtain the reciprocal vectors and the line elements corresponding to (4.24), (4.39), and (4.44). To this end we must examine in detail the group structures given by these expressions. The conventional method of exhibiting group structure is by the commutator relations

$$(X_a, X_b) = C_{ab}{}^{c}X_c, \tag{5.1}$$

where $X_a = \xi_a^i \partial/\partial x^i$ and $(X_a, X_b) = X_aX_b - X_bX_a$ (see Ref. 3). These relations correspond to Eq. (1.3). Using new operators Y_a given by

$$Y_a = X_bA^b{}_a \tag{5.2}$$

[see note following Eq. (3.6)], where $A^a{}_b$ are constant and $\det(A^a{}_b) \neq 0$ we get new commutator relations

$$(Y_a, Y_b) = A^d{}_a A^e{}_b C_{de}{}^f A^{-1}{}^c{}_f Y_c = \bar{C}_{ab}{}^c Y_c \qquad (5.3)$$

which give the same group [cf. remark preceding Eq. (2.2)]. We express this fact by saying that (5.1) and (5.3) are isomorphic. By suitable choice of $A^a{}_b$ one can arrange that $\bar{C}_{ab}{}^c$ has one of the normal forms first given by Lie.[9] One obtains the infinitesimal generators $\eta^i{}_a$ of (5.3) integrating (1.3) under the condition (1.2).[4] We mention that (1.3) does not define $\eta^i{}_a$ uniquely, but *using different solutions of* (1.3) *only means different choice of coordinates in our manifold.* Having found a set of infinitesimal generators one obtains the invariant vectors $\bar{e}^i{}_a$ of (5.3) integrating the equations (1.5) [putting $\bar{e}^i{}_a(x_0^k) = \eta^i{}_a(x_0^k)$ where x_0^k are the coordinates of some arbitrarily chosen point], and finally from (1.8) we obtain the reciprocal vectors $\bar{e}^a{}_i$ of (5.3). The connection between the reciprocal vectors $e^a{}_i$ of (5.1) and those of (5.3) is given by

$$e^a{}_i = A^a{}_b \bar{e}^b{}_i \qquad (5.4)$$

which can be seen easily. One can naturally obtain $e^a{}_i$ by integrating (1.9) directly (if the $C_{ab}{}^c$'s are simple enough). Here the remark that *different solutions only mean different coordinates applies equally.*[2]

In the case where our group has a three-parametric Abelian invariant subgroup we can always arrange to have

$$(X_0, X_1) = 0, \quad (X_1, X_2) = 0, \quad (X_2, X_0) = 0,$$
$$(X_\alpha, X_3) = C_{\alpha 3}{}^\gamma X_\gamma = X_\gamma C^\gamma{}_\alpha \quad (\alpha, \beta, \gamma = 0, 1, 2). \qquad (5.5)$$

Transformations (5.2) of the form

$$A^a{}_b = \begin{pmatrix} A^\alpha{}_\beta & 0 \\ 0 & A \end{pmatrix} \qquad (5.6)$$

preserve the pattern (5.5), leading to

$$(Y_0, Y_1) = 0, \quad (Y_1, Y_2) = 0, \quad (Y_2, Y_0) = 0,$$
$$(Y_\alpha, Y_3) = A Y_\beta (A^{-1}{}^\beta{}_\gamma C^\gamma{}_\delta A^\delta{}_\alpha). \qquad (5.7)$$

Thus the transformations (5.6) are in this case essentially equivalent to similarity transformations applied to the matrix $C^\gamma{}_\delta$. The normal forms (giving the type of the group) which one can thereby obtain are essentially given by the roots of the characteristic polynomial

$$\det(C^\alpha{}_\beta - \lambda \delta^\alpha{}_\beta) = 0. \qquad (5.8)$$

[9] S. Lie und G. Scheffers, *Vorlesungen ueber Continuerliche Gruppen* (B. G. Teubner, Leipzig, 1893).

6. LINE ELEMENTS CORRESPONDING TO (4.24)

For the structure constants (4.24), (5.1) has the form

$$(X_0, X_1) = 0, \quad (X_1, X_2) = 0, \quad (X_2, X_0) = 0,$$
$$(X_0, X_3) = a(1 - s^2) s X_1 / \sqrt{2},$$
$$(X_1, X_3) = a(3 - s^2) s X_0 + a X_1 / \sqrt{2},$$
$$(X_2, X_3) = a(1 - s^2) X_2. \qquad (6.1)$$

(6.1) has a three-dimensional Abelian invariant subgroup. Therefore our remark—made at the end of the previous section—applies. The characteristic equation is essentially the polynomial

$$\lambda^2 - \lambda - \tfrac{1}{2} s^2 (1 - s^2)(3 - s^2) = 0. \qquad (6.2)$$

Corresponding to its roots

$$\lambda_0 = \tfrac{1}{2}(1 - \beta), \quad \lambda_1 = \tfrac{1}{2}(1 + \beta),$$
$$\beta = [1 + 2s^2(1 - s^2)(3 - s^2)]^{\frac{1}{2}} \qquad (6.3)$$

we have to distinguish [in the range of $\frac{1}{2} \leq s^2 \leq 2$ cf. (4.25)] three different cases:

$$\lambda_0 \neq \lambda_1 \quad \text{real}; \qquad (6.4a)$$
$$\lambda_0 = \lambda_1; \qquad (6.4b)$$
$$\lambda_0 = \bar{\lambda}_1. \qquad (6.4c)$$

In case (6.4a), $A^a{}_b$ [see (5.2)] is given by

$$A^a{}_b = \begin{bmatrix} 1 & 1 & 0 & 0 \\ \nu_0 & \nu_1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \qquad (6.5)$$

where

$$\nu_0 = \sqrt{2}\lambda_0 / s(3 - s^2), \quad \nu_1 = \sqrt{2}\lambda_1 / s(3 - s^2). \qquad (6.6)$$

We see that (6.1) is isomorphic to

$$(Y_0, Y_1) = 0, \quad (Y_1, Y_2) = 0, \quad (Y_2, Y_0) = 0,$$
$$(Y_\alpha, Y_3) = a \lambda_\alpha Y_\alpha \qquad (\alpha = 0, 1, 2), \qquad (6.7)$$

where we used the notation

$$\lambda_2 = 1 - s^2. \qquad (6.8)$$

It is very easy to obtain the infinitesimal generators $\eta^i{}_a$, the invariant vectors $\bar{e}^i{}_a$ and the reciprocal vectors $\bar{e}^a{}_i$ of (6.7) by integrating the equations (1.3), (1.5), and using (1.8):

$$\eta^i{}_a = a \begin{bmatrix} 1 & 0 & 0 & \lambda_0 x^0 \\ 0 & 1 & 0 & \lambda_1 x^1 \\ 0 & 0 & 1 & \lambda_2 x^2 \\ 0 & 0 & 0 & 1 \end{bmatrix}; \qquad \bar{e}^i{}_a = a \begin{bmatrix} e^{\lambda_0 x^0} & 0 & 0 & 0 \\ 0 & e^{\lambda_1 x^1} & 0 & 0 \\ 0 & 0 & e^{\lambda_2 x^2} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

($\eta^i{}_a = \bar{e}^i{}_a$ at the point $x^k = 0$) and

$$\bar{e}^a{}_i = \frac{1}{a}\begin{bmatrix} e^{-\lambda_0 x^3} & 0 & 0 & 0 \\ 0 & e^{-\lambda_1 x^3} & 0 & 0 \\ 0 & 0 & e^{-\lambda_2 x^3} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{6.9}$$

Using (5.4), (6.5), and (6.9) we obtain the reciprocal vectors of (6.1) which are given by

$$e^a{}_i = \frac{1}{a}\begin{bmatrix} e^{-\lambda_0 x^3} & e^{-\lambda_1 x^3} & 0 & 0 \\ \nu_0 e^{-\lambda_0 x^3} & \nu_1 e^{-\lambda_1 x^3} & 0 & 0 \\ 0 & 0 & e^{-\lambda_2 x^3} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{6.10}$$

For the line element we find, using (1.15) and (2.2),

$$ds^2 = a^{-2}\{(1-\nu_0^2)e^{-2\lambda_0 x^3}(dx^0)^2 + 2(1-\nu_0\nu_1)e^{-x^3}\,dx^0\,dx^1 + (1-\nu_1^2)e^{-2\lambda_1 x^3}(dx^1)^2 - e^{-2\lambda_2 x^3}(dx^2)^2 - (dx^3)^2\}, \tag{6.11}$$

$$A = a^2(2-s^2)(2s^2-1), \quad \Lambda = -\tfrac{1}{2}a^2(2-s^2), \quad u_i = e^a{}_i.$$

In case (6.4b), we find by a similar procedure that (6.1) is isomorphic to

$$\begin{aligned} &(Y_0, Y_1) = 0, \ (Y_1, Y_2) = 0, \ (Y_2, Y_0) = 0, \\ &(Y_0, Y_3) = \tfrac{1}{2}Y_0, \ (Y_1, Y_3) = Y_0 + \tfrac{1}{2}Y_1, \\ &(Y_2, Y_3) = \lambda_2 Y_3. \end{aligned} \tag{6.12}$$

Here λ_2 is given by (6.8) and the reciprocal vectors of (6.1) are

$$e^a{}_i = \frac{1}{a}\begin{bmatrix} (\frac{1}{2}(b^2-1)^{\frac{1}{2}} - x^3/(b^2-1)^{\frac{1}{2}})\,e^{-x^3/2} & (b^2-1)^{-\frac{1}{2}}\,e^{-x^3/2} & 0 & 0 \\ ((b^2-1)^{\frac{1}{2}}/2b - bx^3/(b^2-1)^{\frac{1}{2}})\,e^{-x^3/2} & b\,e^{-x^3/2}/(b^2-1)^{\frac{1}{2}} & 0 & 0 \\ 0 & 0 & e^{-\lambda_2 x^3} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{6.13}$$

where $b = 2^{\frac{1}{2}}\,s(3-s^2)$. [One can verify this by substituting into (1.9)]. The line element has the form

$$ds^2 = a^{-2}\{[(b^2-1)/2b]^2 e^{-x^3}(dx^0)^2 - (x^3\,dx^0 - dx^1)^2 e^{-x^3} - e^{-2\lambda_0 x^3}(dx^2)^2 - (dx^3)^2\}. \tag{6.14}$$

In case (6.4c), (6.1) is isomorphic to

$$\begin{aligned} &(Y_0, Y_1) = 0, \quad (Y_1, Y_2) = 0, \quad (Y_2, Y_0) = 0, \\ &(Y_0, Y_3) = \tfrac{1}{2}Y_0 - \tfrac{1}{2}\beta' Y_1, \\ &(Y_1, Y_3) = \tfrac{1}{2}\beta' Y_0 + \tfrac{1}{2}Y_1, \quad (Y_2, Y_3) = \lambda_2 Y_2, \end{aligned} \tag{6.15}$$

and the reciprocal vectors of (6.1) are given by

$$e^a{}_i = \frac{1}{a}\begin{bmatrix} e^{-x^3/2}\cos\tfrac{1}{2}\beta' x^3 & -e^{-x^3/2}\sin\tfrac{1}{2}\beta' x^3 & 0 & 0 \\ b^{-1}e^{-x^3/2}(\cos\tfrac{1}{2}\beta' x^3 + \beta'\sin\tfrac{1}{2}\beta' x^3) & b^{-1}e^{-x^3/2}(\beta'\cos\tfrac{1}{2}\beta' x^3 + \sin\tfrac{1}{2}\beta' x^3) & 0 & 0 \\ 0 & 0 & e^{-\lambda_2 x^3} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{6.16}$$

where $\beta' = -[1 + 2s^2(1-s^2)(3-s^2)]^{\frac{1}{2}}$. The line element has the form

$$ds^2 = a^{-2}\{e^{-x^3}[\cos\tfrac{1}{2}\beta' x^3\,dx^0 - \sin\tfrac{1}{2}\beta' x^3\,dx']^2 - e^{-x^3}b^{-2}\,[(\cos\tfrac{1}{2}\beta' x^3 + \beta'\sin\tfrac{1}{2}\beta' x^3)\,dx^0 + (\beta'\cos\tfrac{1}{2}\beta' x^3 - \sin\tfrac{1}{2}\beta' x^3)\,dx^1]^2 - e^{-2\lambda_2 x^3}(dx^2)^2 - (dx^3)^2\}. \tag{6.17}$$

Keeping our notation (6.3) and (6.6) and introducing complex coordinates by $z^0 = \frac{1}{2}(x^0 + ix^1)$, $z^1 = \frac{1}{2}(x^0 - ix^1)$, we can write (6.17) in the form

$$\begin{aligned} ds^2 &= a^{-2}\,\{(e^{-\lambda_0 x^3}\,dz^0 + e^{-\lambda_1 x^3}\,dz^1)^2 - (\nu_0 e^{-\lambda_0 x^3}\,dz^0 + \nu_1 e^{-\lambda_1 x^3}\,dz^1)^2 - e^{2\lambda_2 x^3}(dx^2)^2 - (dx^3)^2\} \\ &\equiv a^{-2}\,\{(1-\nu_0^2)e^{-2\lambda_0 x^3}(dz^0)^2 + 2(1-\nu_0\nu_1)e^{-x^3}\,dz^0\,dz^1 + (1-\nu_1^2)e^{-2\lambda_1 x^3}(dz^1)^2 - e^{-2\lambda_2 x^3}(dx^2)^2 - (dx^3)^2\}, \end{aligned} \tag{6.18}$$

which shows a formal similarity to (6.11); this is because the types of (6.1) in cases (6.4a) and (6.4c) coincide in the complex field. [To emphasize these similarities we could use the following solution of

(1.9) which corresponds to (6.1)

$$e^{\alpha}{}_{i} = \frac{1}{a}\begin{bmatrix} (\lambda_0 - \lambda_1)^{-1}(\lambda_0 e^{-\lambda_1 x^3} - \lambda_1 e^{-\lambda_0 x^3}) & \alpha(3-\alpha^2)(e^{-\lambda_0 x^3} - e^{-\lambda_1 x^3})/\sqrt{2}\,(\lambda_0 - \lambda_1) & 0 & 0 \\ -\alpha(1-\alpha^2)(e^{-\lambda_0 x^3} - e^{-\lambda_1 x^3})/\sqrt{2}\,(\lambda_0 - \lambda_1) & (\lambda_0 - \lambda_1)^{-1}(\lambda_0 e^{-\lambda_0 x^3} - \lambda_1 e^{-\lambda_1 x^3}) & 0 & 0 \\ 0 & 0 & e^{-\lambda_2 x^3} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Here it is only supposed that $\lambda_0 \neq \lambda_1$, regardless of reality. We do not use this because the form of the line element in the case of real λ's would be more complicated than (6.11). We repeat that *the choice of different "corresponding" solutions of* (1.9) *means only the choice of different coordinate systems in our manifold (see Ref. 2).*]

Putting $s^2 = 1/2$ in (6.3) and (6.6) we get

$$\beta = \tfrac{3}{2}, \quad \lambda_0 = -\tfrac{1}{4}, \quad \lambda_1 = \tfrac{5}{4}, \quad \lambda_2 = \tfrac{1}{2},$$
$$\nu_0 = -\tfrac{1}{3}, \quad \nu_1 = 1$$

and after trivial scale changes, (6.11) yields the line element

$$ds^2 = a^{-2}\{e^{x^3/2}(dx^0)^2 + 2e^{-x^3}\,dx^0\,dx^1 - e^{-x^3}(dx^2)^2 - (dx^3)^2\} \qquad (6.19)$$
$$A = 0, \quad \Lambda = -\tfrac{3}{4}a^2.$$

This is a type N vacuum solution with negative Λ term,[5] having a five-parametric group as maximal group.

Putting $s^2 = 1$ we get

$$\beta = 1, \quad \lambda_0 = 0, \quad \lambda_1 = 1, \quad \nu_0 = 0, \quad \nu_1 = 1/\sqrt{2},$$
$$A = a^2, \quad \Lambda = -\alpha^2/2,$$

whereupon (6.11) yields

$$ds^2 = a^{-2}\{(dx^0 + e^{-x^3}\,dx^1)^2 - \tfrac{1}{2}e^{-2x^3}(dx^1)^2 - (dx^2)^2 - (dx^3)^2\}, \qquad (6.20)$$

which is Gödel's line element.[10]

Putting $s^2 = 2$ we get

$$\beta' = \sqrt{3}, \quad \lambda_0 = \nu_0 = \tfrac{1}{2}(1 - i\sqrt{3}),$$
$$\lambda_1 = \nu_1 = \tfrac{1}{2}(1 + i\sqrt{3}), \quad \lambda_2 = -1,$$
$$A = 0, \quad \Lambda = 0,$$

and after changing the scale of z^0 and z^1 (6.11) yields the line element

$$ds^2 = a^{-2}\{e^{-2\lambda_0 x^3}(dz^0)^2 + e^{-2\lambda_1 x^3}(dz^1)^2 - e^{-2x^3}(dx^2)^2 - (dx^3)^2\}, \qquad (6.21)$$

which is Petrov's vacuum solution of type I[11]; he writes it in the form

$$ds^2 = a^{-2}\{e^{-x^3}[((dx^0)^2 - (dx^1)^2)\cos\sqrt{3}x^3 - 2\sin\sqrt{3}x^3\,dx^0\,dx^1] - e^{2x^3}(dx^2)^2 - (dx^3)^2\}.$$

Thus, when s^2 is 1 or 2 our solutions reduce to known forms. When $s^2 = 1/2$, we found a new type N vacuum solution. For all other values of s^2 we have new solutions of the field equations with shear and rotation.

Now consider a line element of the following form

$$ds^2 = dt^2 + 2p_\alpha e^\alpha{}_\mu\,dx^\mu\,dt + e^\alpha{}_\mu g_{\alpha\beta} e^\beta{}_\nu\,dx^\mu\,dx^\nu$$
$$\equiv (dt + p_\alpha e^\alpha{}_\mu\,dx^\mu)^2 + e^\alpha{}_\mu(g_{\alpha\beta} - p_\alpha p_\beta)e^\beta{}_\nu\,dx^\mu\,dx^\nu,$$
$$\alpha, \beta, \cdots = 1, 2, 3; \quad \mu, \nu, \cdots = 1, 2, 3. \qquad (6.22)$$

We require that

$$p_\alpha = \text{const}, \quad g_{\alpha\beta} = g_{\alpha\beta}(t), \quad e^\alpha{}_\mu = e^\alpha{}_\mu(x^\lambda). \qquad (6.23)$$

(To give the correct signature, $g_{\alpha\beta} - p_\alpha p_\beta$ must be negative definite.) We require further that the functions $e^\alpha{}_\mu$ satisfy the equations

$$e^\alpha{}_{\mu,\nu} - e^\alpha{}_{\nu,\mu} = -c_{\beta\gamma}{}^\alpha e^\beta{}_\mu e^\gamma{}_\nu \qquad (6.24)$$

and the condition

$$\det(e^\alpha{}_\mu) \neq 0, \qquad (6.25)$$

which makes them reciprocal vectors of a three-parametric simply transitive group. Then the time-like geodesic unit vector $u^i = \delta^i_0$, having the covariant components $u_i = (1, p_\alpha e^\alpha{}_\mu)$, is an invariant vector of the group (6.24). Suppose that (6.22) is a solution of the field equations with incoherent matter such that u^i is the velocity vector of the matter; then we see that the matter is at rest with respect to this coordinate system, and the hypersurfaces $t = \text{const}$ are homogeneous, allowing the same group as the local rest frames [being characterized, respectively, by the last terms in the alternative expressions for (6.22)].

We now make a remark on solutions with λ_0, λ_1

[10] K. Gödel, Rev. Mod. Phys. **21**, 447 (1949).

[11] A. Z. Petrov, "Gravitational Field Geometry" in *Recent Development in General Relativity* (Pergamon Press, Inc., New York, 1962).

real. In this case (1.9) [corresponding to (6.1)] can also be satisfied by

$$e^{\alpha}{}_{i} = \frac{1}{a}\begin{vmatrix} 1 & -\dfrac{\sqrt{2}\lambda_0}{s(1-s^2)}\,e^{-\lambda_1 y^3} & 0 & 0 \\ 0 & e^{-\lambda_1 y^3} & 0 & \dfrac{s(1-s^2)}{\sqrt{2}}\,y^0 \\ 0 & 0 & e^{-\lambda_2 y^3} & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix}.$$

Note that if $s^2 \to 1$, $\lambda_0/s(1-s^2) \to 1$. This leads to a line element (after scale changes) of the form

$$\begin{aligned} ds^2 &= (dt - \sqrt{2}\lambda_0 e^{-a\lambda_1 y^3}/s(1-s^2)\,dy^1)^2 \\ &\quad - (e^{-a\lambda_1 y^3}\,dy^1 + as(1-s^2)t/\sqrt{2}\,dy^3)^2 \\ &\quad - e^{-2a\lambda_2 y^3}(dy^2)^2 - (dy^3)^2, \end{aligned} \tag{6.26}$$

which is isometric to (6.11) or (6.14), respectively. In this coordinate system the world lines of the matter are the t lines. We note that (6.26) has the form of (6.22). To see this we substitute into (6.22) the expressions

$$p_\alpha = (-\sqrt{2}\lambda_0/s(1-s^2), 0, 0),$$

$$g_{\alpha\beta} = \begin{vmatrix} \left(\dfrac{\sqrt{2}\lambda_0}{s(1-s^2)}\right)^2 - 1 & 0 & -a\,\dfrac{s(1-s^2)}{\sqrt{2}}\,t \\ 0 & -1 & 0 \\ -a\,\dfrac{s(1-s^2)}{\sqrt{2}}\,t & 0 & -\left(a\,\dfrac{s(1-s^2)}{\sqrt{2}}\,t\right)^2 - 1 \end{vmatrix},$$

$$e^{\alpha}{}_{\mu} = \begin{vmatrix} e^{-a\lambda_1 y^3} & 0 & 0 \\ 0 & e^{-a\lambda_2 y^3} & 0 \\ 0 & 0 & 1 \end{vmatrix},$$

which are in fact the reciprocal vectors of a group of Bianchi type 6^4 given by the commutator relations

$$(Y_1, Y_3) = a\lambda_1 Y_1, \quad (Y_2, Y_3) = a\lambda_2 Y_2, \quad (Y_1, Y_2) = 0 \tag{6.27}$$

and we get (6.26). Thus the hypersurfaces $t = \text{const}$ are homogeneous, allowing the group (6.27). We will give the omitted details and further investigations of the solutions in a later paper.

7. LINE ELEMENT CORRESPONDING TO (4.39)

Consider now (4.39); the commutator relations are

$$\begin{aligned} &(X_0, X_1) = \left(\frac{A}{2}\right)^{\frac{1}{2}}\left(\frac{1-s}{1+s}\right)^{\frac{1}{2}} X_2, \\ &(X_1, X_2) = -\left(\frac{A}{2}\right)^{\frac{1}{2}}\frac{2}{(1-s^2)^{\frac{1}{2}}}\,X_0, \\ &(X_2, X_0) = \left(\frac{A}{2}\right)^{\frac{1}{2}}\left(\frac{1+s}{1-s}\right)^{\frac{1}{2}} X_1, \\ &(X_\alpha, X_3) = 0 \quad (\alpha = 0, 1, 2), \qquad |s| < 1. \end{aligned} \tag{7.1}$$

Introducing the operators

$$\begin{aligned} Y_0 &= \left(\frac{2}{A}\right)^{\frac{1}{2}} X_0, \quad Y_1 = \left(\frac{1+s}{A}\right)^{\frac{1}{2}} X_1, \\ Y_2 &= \left(\frac{1-s}{A}\right)^{\frac{1}{2}} X_2, \quad Y_3 = X_3, \end{aligned}$$

we see that (7.1) is isomorphic to

$$\begin{aligned} &(Y_0, Y_1) = Y_2, \ (Y_1, Y_2) = -Y_0, \ (Y_2, Y_0) = Y_1, \\ &(Y_\alpha, Y_3) = 0 \qquad (\alpha = 0, 1, 2). \end{aligned} \tag{7.2}$$

The following solution of (1.9) [corresponding to (7.1)]

$$e^{\alpha}{}_{i} = \left(\frac{2}{A}\right)^{\frac{1}{2}} \times \begin{vmatrix} -e^{x^3} & -1 & 0 & 0 \\ [\frac{1}{2}(1+s)]^{\frac{1}{2}} e^{x^3}\sin x^1 & 0 & [\frac{1}{2}(1+s)]^{\frac{1}{2}}\cos x^1 & 0 \\ [\frac{1}{2}(1-s)]^{\frac{1}{2}} e^{x^3}\cos x^1 & 0 & -[\frac{1}{2}(1-s)]^{\frac{1}{2}}\sin x^1 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix}$$

leads, according to (1.15) and (2.2), to the line element

$$\begin{aligned} ds^2 &= 2A^{-1}\{(dx^1)^2 + 2e^{x^3}\,dx^0\,dx^1 \\ &\quad + \tfrac{1}{2}(1 + s\cos 2x^1)e^{2x^3}(dx^0)^2 \\ &\quad - \tfrac{1}{2}(1 + s\cos 2x^1)(dx^2)^2 \\ &\quad - se^{x^3}\sin 2x^1\,dx^0\,dx^2 - (dx^3)^2\}. \end{aligned} \tag{7.3}$$

This line element has been given previously by the author.[2] The equivalence of (7.3) with the line element (3.33) given on the p. 1018 of Ref. 2 can be seen by replacing the parameter k by $s + 1$ and constructing the metric according to (2.9) of that paper. Making the following simple transformations

$$\begin{aligned} &\left(\frac{2}{A}\right)^{\frac{1}{2}} x' = t, \quad \left(\frac{2}{A}\right)^{\frac{1}{2}} x^0 = y', \\ &\left(\frac{2}{A}\right)^{\frac{1}{2}} x^2 = y^2, \ \left(\frac{2}{A}\right)^{\frac{1}{2}} x^3 = y^3 \end{aligned}$$

we get, from (7.3)

$$\begin{aligned} ds^2 &= (dt)^2 + 2e^{(\frac{1}{2}A)^{\frac{1}{2}}y^3}\,dt\,dy' \\ &\quad + \tfrac{1}{2}[1 + s\cos(2A)^{\frac{1}{2}}t]e^{(2A)^{\frac{1}{2}}y^3}(dy^1)^2 \\ &\quad - \tfrac{1}{2}[1 + s\cos(2A)^{\frac{1}{2}}t](dy^2)^2 \\ &\quad - se^{(\frac{1}{2}A)^{\frac{1}{2}}y^3}\sin(2A)^{\frac{1}{2}}t\,dy^1\,dy^2 - (dy^3)^2. \end{aligned} \tag{7.4}$$

Note that $s = 0$ gives the Gödel cosmos. In this coordinate system the t lines are the world lines of the matter and (7.4) is of the form (6.22). One gets (7.4) by substituting into (6.22) the expressions

$$p_\alpha = (1, 0, 0)$$

$$g_{\alpha\beta} = \begin{bmatrix} \frac{1}{2}[1+s\cos(2A)^{\frac{1}{2}}t] & -\frac{1}{2}s\sin(2A)^{\frac{1}{2}}t & 0 \\ -\frac{1}{2}s\sin(2A)^{\frac{1}{2}}t & -\frac{1}{2}(1+s\cos(2A)^{\frac{1}{2}}t) & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

$$e^{\alpha}{}_{\mu} = \begin{bmatrix} e^{(\frac{1}{2}A)^{\frac{1}{2}}x^2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

which are in fact the reciprocal vectors of a group of Bianchi type 3^4 given by the commutator relations

$$(Y_1, Y_2) = -(A/2)^{\frac{1}{2}}Y_1, \quad (Y_2, Y_3) = 0, \quad (Y_3, Y_1) = 0. \tag{7.5}$$

Thus the hypersurfaces $t = \text{const}$ are homogeneous, allowing the group (7.5). This property of (7.4) has been pointed out by C. Behr (private communication), who has used it to find (7.4) independently.

8. LINE ELEMENTS CORRESPONDING TO (4.44)

The commutator relations corresponding to (4.44) are

$$\begin{aligned}
(X_0, X_1) &= \tfrac{1}{2}A^{\frac{1}{2}}\{2p + [p(p-1)]^{\frac{1}{2}}\}X_2, \\
(X_0, X_2) &= -\tfrac{1}{2}A^{\frac{1}{2}}\{2p - [p(p-1)]^{\frac{1}{2}}\}X_1, \\
(X_0, X_3) &= 0, \\
(X_2, X_3) &= \tfrac{1}{2}A^{\frac{1}{2}}\{2[p(p+1)]^{\frac{1}{2}} - [(p-1)(p+1)]^{\frac{1}{2}}\}X_1, \\
(X_3, X_1) &= \tfrac{1}{2}A^{\frac{1}{2}}\{2[p(p+1)]^{\frac{1}{2}} + [(p-1)(p+1)]^{\frac{1}{2}}\}X_2, \\
(X_1, X_2) &= A^{\frac{1}{2}}\{-(p-1)X_0 + [p(p+1)]^{\frac{1}{2}}X_3\}.
\end{aligned} \tag{8.1}$$

Already from (4.40) it is clear that the sign of p is critical (recall $|p| \geq 1$). In fact, the group (8.1) is isomorphic to

$$(Y_1, Y_2) = Y_3, \quad (Y_2, Y_3) = Y_1, \quad (Y_3, Y_1) = Y_2, \quad (Y_0, Y_\alpha) = 0 \quad (\alpha = 1, 2, 3) \tag{8.2a}$$

if $p \geq 1$, and isomorphic to

$$(Y_1, Y_2) = -Y_3, \quad (Y_2, Y_3) = Y_1, \quad (Y_3, Y_1) = Y_2, \quad (Y_0, Y_\alpha) = 0 \quad (\alpha = 1, 2, 3) \tag{8.2b}$$

if $p \leq -1$. This can be demonstrated by making the linear transformation (5.2), where in case $p \geq 1$, $A^b{}_c$ is given by

$$A^b{}_c = a\begin{bmatrix} [p(p+1)]^{\frac{1}{2}} & 0 & 0 & -(p-1) \\ 0 & \{2p - [p(p-1)]^{\frac{1}{2}}\}^{\frac{1}{2}} & 0 & 0 \\ 0 & 0 & \{2p + [p(p-1)]^{\frac{1}{2}}\}^{\frac{1}{2}} & 0 \\ -p & 0 & 0 & [p(p+1)]^{\frac{1}{2}} \end{bmatrix} \tag{8.3a}$$

and in case $p \leq -1$

$$A^b{}_c = a\begin{bmatrix} [p(p+1)]^{\frac{1}{2}} & 0 & 0 & (p-1) \\ 0 & \{-2p + [p(p-1)]^{\frac{1}{2}}\}^{\frac{1}{2}} & 0 & 0 \\ 0 & 0 & \{-2p - [p(p-1)]^{\frac{1}{2}}\}^{\frac{1}{2}} & 0 \\ -p & 0 & 0 & -[p(p+1)]^{\frac{1}{2}} \end{bmatrix} \tag{8.3b}$$

with a in both cases given by

$$a = [Ap(3p+1)]^{-\frac{1}{2}}. \tag{8.4}$$

Denoting the reciprocal vectors of (8.1) and (8.2) by $e^a{}_i$ and $\bar{e}^a{}_i$ respectively, we have [see (5.4)] $e^a{}_i = A^a{}_b\bar{e}^b_i$, and since the metric is given by $g_{ik} = e^c{}_i\eta_{cd}e^d{}_k$, we obtain

$$g_{ik} = \bar{e}^a{}_i g_{ab}\bar{e}^b{}_k, \tag{8.5}$$

where

$$g_{ab} = A^c{}_a\eta_{cd}A^d{}_b. \tag{8.6}$$

Case $p \geq 1$

Carrying out the indicated multiplications we have, in case $p \geq 1$, using for $A^b{}_c$ (8.3a), that

$$g_{ab} = a^2\begin{bmatrix} p & 0 & 0 & [p(p+1)]^{\frac{1}{2}} \\ 0 & -\{2p - [p(p-1)]^{\frac{1}{2}}\} & 0 & 0 \\ 0 & 0 & -\{2p + [p(p-1)]^{\frac{1}{2}}\} & 0 \\ [p(p+1)]^{\frac{1}{2}} & 0 & 0 & -(3p-1) \end{bmatrix}. \tag{8.7}$$

It is easy to see that

$$\bar{e}^a{}_i = \begin{bmatrix} (2a)^{-1}[(p+1)/p]^{\frac{1}{2}} & 0 & 0 & 0 \\ 0 & -\sin x & \sin x^1 \cos x & 0 \\ 0 & \cos x & \sin x^1 \sin x & 0 \\ (2a)^{-1} & 0 & \cos x^1 & 1 \end{bmatrix}, \tag{8.8}$$

where

$$x = x^3 + (2a)^{-1}t \qquad (t = x^0) \tag{8.9}$$

are reciprocal vectors of (8.2) (a). Now $\bar{e}^a{}_i$ can be written in the form

$$\bar{e}^a{}_i = B^a{}_b k^b{}_i, \tag{8.10}$$

where

$$B^a{}_b(t) = \begin{bmatrix} (2a)^{-1}[(p+1)/p]^{\frac{1}{2}} & 0 & 0 & 0 \\ 0 & \cos t/2a & -\sin t/2a & 0 \\ 0 & \sin t/2a & \cos t/2a & 0 \\ (2a)^{-1} & 0 & 0 & 1 \end{bmatrix} \tag{8.11}$$

and

$$k^a{}_i = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -\sin x^3 & \sin x^1 \cos x^3 & 0 \\ 0 & \cos x^3 & \sin x^1 \sin x^3 & 0 \\ 0 & 0 & \cos x^1 & 1 \end{bmatrix}. \tag{8.12}$$

Carrying out the multiplications indicated by $g_{ik} = B^a{}_c k^c{}_i g_{ab} B^b{}_d k^d{}_k$ we get the line element in the form of (6.22) with

$$p_\alpha = a(0, 0, -(p-1)), \tag{8.13a}$$

$$g_{\alpha\beta}(t) = -a^2 \begin{bmatrix} 2p - [p(p-1)]^{\frac{1}{2}} \cos t/a & [p(p-1)]^{\frac{1}{2}} \sin t/a & 0 \\ [p(p-1)]^{\frac{1}{2}} \sin t/a & 2p + [p(p-1)]^{\frac{1}{2}} \cos t/a & 0 \\ 0 & 0 & (3p-1) \end{bmatrix} \tag{8.13b}$$

and

$$e^\alpha{}_\mu = \begin{bmatrix} -\sin x^3 & \sin x^1 \cos x^3 & 0 \\ \cos x^3 & \sin x^1 \sin x^3 & 0 \\ 0 & \cos x^1 & 1 \end{bmatrix}, \tag{8.13c}$$

which are the reciprocal vectors of the group of Bianchi type 9^4 with the commutator relations

$$(Y_1, Y_2) = Y_3, \qquad (Y_2, Y_3) = Y_1, \qquad (Y_3, Y_1) = Y_2. \tag{8.14}$$

The world lines of the matter are the t lines, and the hypersurfaces $t = \text{const}$ are homogeneous, allowing the group (8.14). This is a finite rotating universe previously given in Ref. 2. The correspondence between the parameters here and in Ref. 2 is given by $R = 2a(2p)^{\frac{1}{2}}$ and $k = \frac{1}{2}[(p-1)/p]^{\frac{1}{2}}$. At $p = 1$ we have Einstein's static universe [see (9.31)].

Case p ≤ −1

In the case $p \leq -1$ [see (8.3b)] g_{ab}, defined by (8.6), has the form

$$g_{ab} = a^2 \begin{bmatrix} p & 0 & 0 & -[p(p+1)]^{\frac{1}{2}} \\ 0 & 2p - [p(p-1)]^{\frac{1}{2}} & 0 & 0 \\ 0 & 0 & 2p + [p(p-1)]^{\frac{1}{2}} & 0 \\ -[p(p+1)]^{\frac{1}{2}} & 0 & 0 & -(3p-1) \end{bmatrix}. \tag{8.15}$$

A solution of (1.9), corresponding to the group (8.2b), is given by

$$\bar{e}^{\alpha}{}_{i} = \begin{pmatrix} 1 & a^{-1}[(p+1)/p]^{\frac{1}{2}} & 0 & 0 \\ 0 & 0 & \cos x^1/2a\ e^{x^3} & \sin x^1/2a \\ 0 & 0 & -\sin x^1/2a\ e^{x^3} & \cos x^1/2a \\ 0 & -(2a)^{-1} & 0 & -e^{x^3} \end{pmatrix}. \tag{8.16}$$

Using (8.5), (8.15), and (8.16) and the notation

$$t = x^1, \quad y^1 = x^0, \quad y^2 = x^2, \quad y^3 = x^3,$$

we get the line element in the form

$$ds^2 = dt^2 + 2a\{[p(p+1)]^{\frac{1}{2}}\, dy^1 - (p-1)e^{y^3}\, dy^3\}\, dt$$
$$- a^2\{-p(dy^1)^2 - 2[p(p+1)]^{\frac{1}{2}}e^{y^3}\, dy^1\, dy^3$$
$$+ (-2p + [p(p-1)]^{\frac{1}{2}} \cos t/2a)(dy^2)^2$$
$$\times 2[p(p-1)]^{\frac{1}{2}}e^{y^3}\, dy^2\, dy^3$$
$$+ e^{2y^3}((p-1) - [p(p-1)]^{\frac{1}{2}} \cos t/2a)(dy^3)^2\}, \tag{8.17}$$

where a is given by (8.4). This is a new metric. [If $p = -1$, (8.17) goes over into (7.4) with the special value $s = 1/2^{\frac{1}{2}}$.] The world lines of the matter are the t-lines, and the motion of the matter in this space has shear and rotation. One sees immediately that (8.17) is of the form (6.22) with the following values of p_α, $g_{\alpha\beta}$, and $e^{\alpha}{}_{\mu}$

$$p_\alpha = a([p(p+1)]^{\frac{1}{2}}, 0, -(p-1)), \tag{8.18a}$$

$$g_{\alpha\beta} = -a^2 \begin{pmatrix} -p & 0 & -[p(p+1)]^{\frac{1}{2}} \\ 0 & -2p + p(p-1)\cos t/a & [p(p-1)]^{\frac{1}{2}} \sin t/a \\ -[p(p+1)]^{\frac{1}{2}} & [p(p-1)]^{\frac{1}{2}} \sin t/a & (p-1) - [p(p-1)]^{\frac{1}{2}} \cos t/a \end{pmatrix} \tag{8.18b}$$

and

$$e^{\alpha}{}_{\mu} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & e^{y^3} \end{pmatrix} \tag{8.18c}$$

$e^{\alpha}{}_{\mu}$ are the reciprocal vectors of a group of Bianchi type 3^4 given by the commutator relations

$$(Y_2, Y_3) = Y_3, \ (Y_3, Y_1) = 0, \ (Y_1, Y_2) = 0. \tag{8.19}$$

In (8.17) therefore the t lines are the world lines of the matter and the hypersurfaces t = const are homogeneous, allowing the group (8.19).

9. PROOF OF GÖDEL'S THEOREM

Gödel writes in Ref. 10: "I am mentioning without proof that, disregarding the connectivity in the large (which can be changed by identifying the points of certain point sets with each other), the solution given (Gödel's cosmos) and Einstein's static universe are the only spatially homogeneous cosmological solutions with nonvanishing density of matter and equidistant world lines of matter."

The conditions of the above theorem imply homogeneous space–time. We can see this by the following consideration: "cosmological solution" means a solution of (1.36). One sees from the twice contracted Bianchi identities that the world lines of the matter are geodesic. "Equidistant world lines" mean vanishing shear and vanishing expansion (see Ref. 12), but geodesic motion with vanishing shear and vanishing expansion leads to the equation

$$u_{(i;k)} = 0, \tag{9.1}$$

where u^i is the tangent vector to the world line. Equation (9.1) is the Killing equation, whence u^i is the infinitesimal generator of a one-parametric group, which leaves the metric invariant. By virtue of the assumption of spatial homogeneity we already have a three-parametric group G_3. Since u^i is necessarily an invariant vector of this G_3 (this follows from the field equations), u^i has vanishing commutators with the infinitesimal generators of G_3. That means that we have a four parametric group leaving the metric of the space–time invariant, i.e., we have homogeneous space–time. Nonvanishing density implies

$$A > 0. \tag{9.2}$$

After these preliminaries it is evident that in order to prove Gödel's theorem we have only to show that under the assumption [see (2.52)]

$$\sigma_{ABA'B'} = 0, \tag{9.3}$$

which corresponds to (9.1), our equations (3.18) and (3.25) have precisely the two solutions mentioned by Gödel.

We make a preliminary remark: If a normal hyperbolic Riemannian space R_4 with the signature -2

[12] J. Ehlers, Akad. Wiss. Mainz. Abh. Math. Natl. Kl.

contains a line congruence with the properties $u_i u^i = 1$ and $u_{(i;k)} = 0$, and satisfies the vacuum field equations $R_{ik} = 0$ [i.e., (1.36) with $A = 0$ and $\Lambda = 0$], then the space–time is flat. This can easily be seen by using, for example, the equations (2.3.2), (2.3.3) in Ref. 13 p. 56.

We now proceed to set up the equations corresponding to the condition (9.3).

Case (i): Eqs. (3.18). We have seen in Sec. 4 that the equations (3.18) reduce to the equations (4.7). According to (4.15) the condition (9.3) gives the single equation

$$\tau = 0. \tag{9.4}$$

Using (6.4) and $\alpha = \bar{\alpha} \neq 0$ [see (3.15a) and (4.2)], we get from (4.7) the following equations

$$\epsilon = -\gamma, \tag{9.5a}$$

$$\sigma(r - 2\gamma) = 2\alpha^2, \tag{9.5b}$$

$$\sigma = r - 2\gamma, \tag{9.5c}$$

$$\sigma\bar{\sigma} = 2\alpha^2, \tag{9.5d}$$

$$A = 6\alpha^2 - r^2, \tag{9.5e}$$

$$\Lambda = -\tfrac{1}{2}(2\alpha^2 + r^2). \tag{9.5f}$$

Inserting (9.5c) into (9.5b) we have $(r - 2\gamma)^2 = 2\alpha^2$, which can only be satisfied if $\gamma = 0$, since r and α are real. It follows further that

$$\sigma = r, \quad 2\alpha^2 = r^2, \quad A = 2r^2, \quad \Lambda = -\tfrac{1}{2}A. \tag{9.6}$$

Collecting our results we have

$A_{ABCC'}$:

AB \ CC'	11′	12′	21′	22′
11	2α	r	r	0
(12)	0	α	α	0
22	0	r	r	2α

. (9.7)

The nonvanishing components of the corresponding structure constant tensor are given by

$$C_{31}{}^0 = 2\sqrt{2}\alpha, \quad C_{31}{}^1 = -\sqrt{2}r. \tag{9.8}$$

We see later that this leads to the Gödel cosmos.

Case (ii): Eqs. (3.25). The components of $\sigma_{ABA'B'}$ [see (2.52)] are given by

$\sigma_{ABA'B'}$:

AB \ $A'B'$	1′1′	(1′2′)	2′2′
11	$(\gamma + \bar{\gamma})/2\sqrt{2}$	$[\kappa - (2\bar{\alpha} + \beta)]/4\sqrt{2}$	$(\sigma - \bar{\lambda})/2\sqrt{2}$
(12)	$[\bar{\kappa} - (2\alpha + \bar{\beta})]/4\sqrt{2}$	$-(\gamma + \bar{\gamma})/2\sqrt{2}$	$[\bar{\nu} - (\bar{\alpha} + 2\beta)]/4\sqrt{2}$.
22	$(\bar{\sigma} - \lambda)/2\sqrt{2}$	$-[\nu - (\alpha + 2\beta)]/4\sqrt{2}$	$(\gamma + \bar{\gamma})/2\sqrt{2}$

(9.9)

Condition (9.3) combined with (3.22) gives

$$\rho + \bar{\rho} = \epsilon + \bar{\epsilon} = \gamma + \bar{\gamma} = \mu + \bar{\mu} = 0,$$

$$\lambda = \bar{\sigma}, \quad \nu = \alpha + 2\bar{\beta}, \quad \kappa = 2\bar{\alpha} + \beta. \tag{9.10}$$

With the help of (3.21) we can arrange that

$$\alpha + \bar{\alpha} = 0, \tag{9.11}$$

and also that

$$\beta + \bar{\beta} = 0. \tag{9.12}$$

This can be seen by the following reasoning: Eqs. (9.10), (9.11), (3.25a) and the condition that A has to be real, imply $\alpha(\beta + \bar{\beta}) = 0$. If $\alpha \neq 0$, (9.12) follows; if $\alpha = 0$, we can use (3.21) to make β imaginary. Using these equations we get from (3.25) the following system:

$$-2(\alpha - \beta)(2\alpha - \beta) - \sigma\bar{\sigma} - \rho^2 = A, \tag{9.13a}$$

$$-2(\alpha - \beta)(\alpha - 2\beta) - \sigma\bar{\sigma} - \mu^2 = A, \tag{9.13b}$$

$$\sigma\bar{\sigma} - \rho\mu - 2(\epsilon\mu + \gamma\rho) + 4\alpha\beta = A, \tag{9.13c}$$

$$-\tfrac{1}{2}((\alpha - \beta)^2 + \sigma\bar{\sigma} - \rho\mu) = \Lambda, \tag{9.13d}$$

$$2\alpha(\alpha - 3\beta) - \sigma(\rho + \mu - 4\gamma) = 0, \tag{9.13e}$$

$$2\beta(3\alpha - \beta) - \sigma(\rho + \mu - 4\epsilon) = 0, \tag{9.13f}$$

$$\gamma(3\alpha - \beta) - \epsilon(\alpha - 3\beta) = 0, \tag{9.13g}$$

$$\sigma(\alpha + \beta) + \rho(\alpha - 2\beta) + \mu(2\alpha - \beta) - \tfrac{1}{2}(\alpha - \beta)(\gamma + \epsilon) = 0, \tag{9.13h}$$

$$\sigma\alpha - \rho\alpha + \tfrac{1}{2}(2\alpha - \beta)(\gamma + \epsilon) = 0, \tag{9.13i}$$

$$\sigma\beta + \mu\beta + \tfrac{1}{2}(\alpha - 2\beta)(\gamma + \epsilon) = 0, \tag{9.13j}$$

where (9.13g) and (9.13h) are the sum and the difference of (3.25f) and (3.25g). We now must find all its solutions. In order to do that we have to distinguish four different cases,

(1) $\alpha\beta \neq 0$; (2) $\alpha \neq 0, \beta = 0$;

(3) $\alpha = 0, \beta \neq 0$; (4) $\alpha = \beta = 0$

which we discuss separately.

Case (1): From (9.13i) and (9.13j) we get

$$\rho = \sigma + (2\alpha)^{-1}(2\alpha - \beta)(\gamma + \epsilon),$$

$$\mu = -\sigma - (2\beta)^{-1}(\alpha - 2\beta)(\gamma + \epsilon).$$

Inserting into (9.13h) we obtain

$$(\alpha - \beta)^3(\gamma + \epsilon) = 0$$

[13] P. Jordan, J. Ehlers, and W. Kundt, Akad. Wiss. Mainz, Abh. Math. Nat. Kl.

which can be considered together with (9.13g) as a system of linear equations for the quantities γ and ϵ, having the determinant

$$d = 4(\alpha - \beta)^4.$$

If $d \neq 0$ it follows that $\gamma = \epsilon = 0$, $\rho = \sigma$, $\mu = -\sigma$ and from (9.13e) and (9.13f) we obtain

$$\alpha - 3\beta = 0, \quad 3\alpha - \beta = 0,$$

whence $\alpha = \beta = 0$, which contradicts $d \neq 0$. If $d = 0$, $\alpha = \beta$, and, from (9.13g), $\gamma + \epsilon = 0$ whence $\rho = -\mu = \sigma$. Substituting into (9.13) we get

$$A = 0, \quad \Lambda = 0, \quad \sigma^2 - \sigma\gamma = 0,$$

which means that the space–time is flat, according to our earlier remark.

Case (2): From (9.13g) and (9.13j) we get

$$\gamma = 0, \quad \epsilon = 0,$$

and from (9.13i) and (9.13h)

$$\rho = \sigma, \quad \mu = -\sigma.$$

Then (9.13c) would require $\alpha = 0$, which is a contradiction.

Case (3): From (9.13g) and (9.13i) we get

$$\gamma = 0, \quad \epsilon = 0$$

and from (9.13j) and (9.13h)

$$\rho = \sigma, \quad \mu = -\sigma.$$

Then (9.13f) would require $\beta = 0$, which is a contradiction. We are therefore left with only one possibility:

Case (4):

$$\alpha = 0, \quad \beta = 0. \tag{9.14}$$

Equations (9.13) now read

$$-\sigma\bar{\sigma} - \rho^2 = A, \tag{9.15a}$$

$$-\sigma\bar{\sigma} - \mu^2 = A, \tag{9.15b}$$

$$\sigma\bar{\sigma} - \rho\mu - 2(\epsilon\mu + \gamma\rho) = A, \tag{9.15c}$$

$$-\tfrac{1}{2}(\sigma\bar{\sigma} - \rho\mu) = \Lambda, \tag{9.15d}$$

$$\sigma(\rho + \mu - 4\gamma) = 0, \tag{9.15e}$$

$$\sigma(\rho + \mu - 4\epsilon) = 0. \tag{9.15f}$$

They imply

$$\sigma = 0. \tag{9.16}$$

This can be seen by the following reasoning: If $\sigma \neq 0$ we have, from (9.15),

$$\gamma = \epsilon = \tfrac{1}{4}(\rho + \mu), \quad \Lambda = -\tfrac{1}{2}(\sigma\bar{\sigma} - \rho\mu),$$

$$A = \sigma\bar{\sigma} - \rho\mu - \tfrac{1}{2}(\rho + \mu)^2 \quad A = -\sigma\bar{\sigma} - \rho^2,$$

$$(\rho + \mu)(\rho - \mu) = 0.$$

Therefore

$$\mu = \rho \text{ or } \mu = -\rho.$$

If $\mu = \rho$ we have

$$\gamma = \epsilon = \tfrac{1}{2}\rho,$$

$$A = -\sigma\bar{\sigma} - \rho^2, \quad A = \sigma\bar{\sigma} - 3\rho^2, \quad \Lambda = -\tfrac{1}{2}(\sigma\bar{\sigma} - \rho^2).$$

This would require $\sigma\bar{\sigma} = \rho$ which is only possible if $\sigma = 0$ and $\rho = 0$ since $\sigma\bar{\sigma} \geq 0$ and $\rho^2 \leq 0$. But then the space–time is flat. If $\mu = -\rho$ we have

$$\gamma = \epsilon = 0,$$

$$A = -(\sigma\bar{\sigma} + \rho^2), \quad A = \sigma\bar{\sigma} + \rho^2, \quad \Lambda = -\tfrac{1}{2}(\sigma\bar{\sigma} + \rho^2)$$

but then $\sigma\bar{\sigma} + \rho^2 = 0$ and the space–time is flat.

Using (9.16) we get from (9.15) the equations

$$A = -\rho^2, \quad A = -\mu^2, \quad A = -\rho\mu - 2(\epsilon\mu + \gamma\rho), \quad \Lambda = \tfrac{1}{2}\rho\mu, \tag{9.17}$$

and we have the possibilities

$$\mu = \rho, \ \gamma + \epsilon = 0, \ A = -\rho^2, \ \Lambda = -\tfrac{1}{2}A, \tag{9.18}$$

$$\mu = -\rho, \ \gamma - \epsilon = 0, \ A = -\rho^2, \ \Lambda = \tfrac{1}{2}A. \tag{9.19}$$

Collecting our results we have, in the case (9.18),

$A_{ABCC'}$:

AB \ CC'	11′	12′	21′	22′
11	0	0	ir	0
(12)	$-ig$	0	0	ig
22	0	ir	0	0

($\gamma = ig$, $\rho = ir$) (9.20)

and the corresponding nonvanishing structure constants are

$$C_{12}{}^0 = \sqrt{2}r, \quad C_{20}{}^1 = -r/\sqrt{2}, \quad C_{01}{}^2 = -r/\sqrt{2},$$

$$C_{23}{}^1 = -\sqrt{2}g, \quad C_{31}{}^2 = -\sqrt{2}g. \tag{9.21}$$

In the case (9.19) we have

$A_{ABCC'}$:

AB \ CC'	11′	12′	21′	22′
11	0	0	ir	0
(12)	$i(g - r)$	0	0	ig
22	0	$-ir$	0	0

(9.22)

with the corresponding nonvanishing structure constants

$$C_{23}{}^1 = -\sqrt{2}r, \ C_{31}{}^2 = -\sqrt{2}r, \ C_{12}{}^3 = -\sqrt{2}r,$$

$$C_{20}{}^1 = \sqrt{2}(g - \tfrac{1}{2}r), \ C_{10}{}^2 = -\sqrt{2}(g - \tfrac{1}{2}r). \tag{9.23}$$

To conclude the proof of Gödel's theorem we give the line elements corresponding to (9.8), (9.21), and (9.23). One sees by substituting into (1.9) that the following are possible sets of reciprocal vectors:

$$e^a{}_i = \begin{bmatrix} -(r\sqrt{2})^{-1} & -e^{-x^3}/r\sqrt{2} & 0 & 0 \\ 0 & -e^{-x^3}/2\sqrt{2}\,\alpha & 0 & 0 \\ 0 & 0 & -(r\sqrt{2})^{-1} & 0 \\ 0 & 0 & 0 & -(r\sqrt{2})^{-1} \end{bmatrix}, \tag{9.24}$$

where

$$2\alpha^2 = r^2;$$

and

$$e^a{}_i = \begin{bmatrix} r^{-1} & 0 & e^{x^1}/r & 0 \\ 0 & -\cos\tau/r & -e^{x^1}\sin\tau/r\sqrt{2} & 0 \\ 0 & \sin\tau/r & -e^{x^1}\cos\tau/r\sqrt{2} & 0 \\ 0 & 0 & 0 & -(r)^{-1} \end{bmatrix}, \tag{9.25}$$

where

$$\tau = (x^0 - (2g/r)x^3)/\sqrt{2}; \tag{9.26}$$

and

$$e^a{}_i = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \sin\omega/r\sqrt{2} & -\sin x^1\cos\omega/r\sqrt{2} & 0 \\ 0 & -\cos\omega/r\sqrt{2} & -\sin x^1\sin\omega/r\sqrt{2} & 0 \\ 0 & 0 & -\cos x^1/r\sqrt{2} & -(r\sqrt{2})^{-1} \end{bmatrix}, \tag{9.27}$$

where

$$\omega = \sqrt{2}(g - \tfrac{1}{2}r)x^0 + x^3. \tag{9.28}$$

Using $g_{ik} = e^a{}_i\eta_{ab}e^b{}_i$, the corresponding line elements have the form

$$ds^2 = (2r^2)^{-1}\{(dx^0 + e^{-x^3}\,dx^1)^2 - \tfrac{1}{2}e^{-2x^3}(dx^1)^2 - (dx^2)^2 - (dx^3)^2\}, \tag{9.29}$$

$$ds^2 = r^{-2}\{(dx^0 + e^{x^1}\,dx^2)^2 - \tfrac{1}{2}e^{2x^1}(dx^2)^2 - (dx^1)^2 - (dx^3)^2\}, \tag{9.30}$$

$$ds^2 = (dx^0)^2 - (2r^2)^{-1}\{(dx^1)^2 + (dx^2)^2 + (dx^3)^2 + 2\cos x^1\,dx^2\,dx^3\}, \tag{9.31}$$

(9.29) and (9.30) are obviously the Gödel cosmos, (9.31) is the Einstein static universe; to see this we calculate the expression

$$(dy^1)^2 + (dy^2)^2 + (dy^3)^2 + (dy^4)^2$$

with both of the following alternative expressions for the y's:

$$y^1 = \cos\tfrac{1}{2}x^1\cos\tfrac{1}{2}(x^2 + x^3) = \sin\chi\cos\varphi\sin\theta,$$

$$y^2 = \cos\tfrac{1}{2}x^1\sin\tfrac{1}{2}(x^2 + x^3) = \sin\chi\sin\varphi\sin\theta,$$

$$y^3 = \sin\tfrac{1}{2}x^1\cos\tfrac{1}{2}(x^2 - x^3) = \sin\chi\cos\theta,$$

$$y^4 = \sin\tfrac{1}{2}x^1\sin\tfrac{1}{2}(x^2 - x^3) = \cos\chi. \tag{9.32}$$

[The formulas (9.32) have been computed by E. L. Schücking.] On one hand we get (9.31), and on the other hand

$$ds^2 = (dx^0)^2 - 2r^{-2}\{d\chi^2 + \sin^2\chi(d\theta^2 + \sin^2\theta\,d\varphi^2)\} \tag{9.33}$$

which is the usual form of Einstein's static universe. This completes the proof of Gödel's theorem.

ACKNOWLEDGMENTS

I am deeply indebted to the members of the Mathematics and Mathematical Physics Division of the Southwest Center for Advanced Studies, M. Cahen, I. Robinson, M. Trümper, for their interest in this work. Special thanks go to W. Rindler for valuable discussions on spinors and for his help in preparing this paper for publication. This research was partially supported by the Center's NASA grant, number NsG 269-62.

Uniform Asymptotic Estimates for Wave Packets in the Quantum Theory of Scattering*

CALVIN H. WILCOX
Argonne National Laboratory, Argonne, Illinois and The University of Wisconsin, Madison, Wisconsin
(Received 25 September 1964)

The theory of the scattering of particles by a potential field $V(x)$ according to the classical wave mechanics of Schrödinger is considered. Various conditions on $V(x)$ are known which imply that the wave packets $\psi(x, t)$ describing scattered particles are asymptotically equal in $L_2(R^n)$ to wave packets $\psi_0(x, t)$ for free particles when $t \to \infty$ or $-\infty$. In this paper, additional conditions on $V(x)$ and the initial values of the wave packets are given which imply that $\psi(x, t)$ and $\psi_0(x, t)$ have square-integrable (or continuous) partial derivatives of a prescribed order whose difference tends to zero in $L_2(R^n)$ (or uniformly in R^n) when $t \to \infty$ or $-\infty$.

1. INTRODUCTION

THIS paper deals with the theory of the scattering of particles (elementary particles, atoms, or molecules) by a potential field $V(x)$ according to the classical wave mechanics of Schrödinger. In the theory, the state of a particle at time t is described by a wave packet $\psi(x, t)$ which is a complex-valued function of the point $x \in R^n$ (n-dimensional Euclidean space). The wave packets are solutions of an initial-value problem for Schrödinger's wave equation which may be written (after an appropriate choice of units)

$$i\,\partial\psi/\partial t = -\Delta\psi + V(x)\psi, \quad x \in R^n, \quad -\infty < t < \infty,$$

and

$$\psi(x, 0) = \psi(x), \qquad x \in R^n.$$

Here Δ is the Laplace operator on R^n and $V(x)$ is the potential function. In particular, the wave packets $\psi_0(x, t)$ for free particles are solutions of

$$i\,\partial\psi_0/\partial t = -\Delta\psi_0, \qquad x \in R^n, \qquad -\infty < t < \infty,$$

and

$$\psi_0(x, 0) = \psi_0(x), \qquad x \in R^n.$$

This problem can be solved explicitly in various ways, so that the wave packets for free particles may be regarded as known.

Physical intuition suggests that when $t \to \infty$ or $-\infty$ the wave packet $\psi(x, t)$ for a scattered particle converges asymptotically to a wave packet $\psi_0(x, t)$ for a free particle. This behavior has been verified for various classes of potentials $V(x)$ by Cook,[1] Hack,[2] Kuroda,[3] Stankevich,[4] Brownell,[5] and others. In all of this work, the convergence is in the mean-square sense, i.e.,

$$\int_{R^n} |\psi(x, t) - \psi_0(x, t)|^2\, dx \to 0 \text{ when } t \to \infty \text{ or } -\infty.$$

The purpose of this paper is to present additional conditions on $V(x)$ and the initial values which guarantee that $\psi(x, t)$ and $\psi_0(x, t)$ have square-integrable (or continuous) partial derivatives of a prescribed order whose difference tends to zero in the mean-square sense (or uniformly in R^n) when $t \to \infty$ or $-\infty$. The results are described below, following a more precise formulation of the problem.

The quantum-mechanical scattering problem can be given the following abstract formulation.[6] The wave packets describing both the free and the scattered particles are represented by vectors ψ in a separable Hilbert space $\mathfrak{H}$ over the complex number field (the state space). The free and scattered particles are characterized by their Hamiltonian energy operators, denoted by H_0 and H, respectively, which are self-adjoint linear operators on $\mathfrak{H}$. H_0 and H have spectral resolutions[7]

$$H_0 = \int_{-\infty}^{\infty} \lambda\, dE_0(\lambda) \quad \text{and} \quad H = \int_{-\infty}^{\infty} \lambda\, dE(\lambda),$$

and the time dependences of the wave packets representing free and scattered particles are given by the corresponding unitary operators

$$e^{-itH_0} = \int_{-\infty}^{\infty} e^{-it\lambda}\, dE_0(\lambda) \text{ and } e^{-itH} = \int_{-\infty}^{\infty} e^{-it\lambda}\, dE(\lambda).$$

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

[1] J. M. Cook, J. Math. and Phys. **36**, 82 (1957).

[2] M. N. Hack, Nuovo Cimento **9**, 731 (1958).

[3] S. T. Kuroda, Nuovo Cimento **12**, 431 (1959).

[4] I. V. Stankevich, Dokl. Akad. Nauk SSSR **144**, 279 (1962) [English transl.: Soviet Phys.—Doklady **3**, 719 (1962)].

[5] F. H. Brownell, Pacific J. Math. **12**, 47 (1962).

[6] J. M. Jauch, Helv. Phys. Acta **31**, 127 (1958).

[7] F. Riesz and B. Sz.-Nagy, *Functional Analysis* (F. Ungar Publishing Company, New York, 1955), Chap. 8.

Thus

$$\psi_0(t) = e^{-itH_0}\psi_0 \quad \text{and} \quad \psi(t) = e^{-itH}\psi \tag{1.1}$$

describe the evolution in time of the wave packets for a free particle with initial state ψ_0 and a scattered particle with initial state ψ, respectively. The wave packets for the free and scattered particles are said to be asymptotically equal in $\mathfrak{H}$ for $t \to \infty$, written

$$\psi_0(t) \sim \psi(t), \qquad t \to \infty, \tag{1.2}$$

if

$$\lim_{t\to\infty} ||\psi_0(t) - \psi(t)||_{\mathfrak{H}} = 0.$$

If ψ is an eigenvector for H, so $H\psi = \lambda\psi$, then

$$\psi(t) = e^{-itH}\psi = e^{-it\lambda}\psi. \tag{1.3}$$

Thus $\psi(t)$ describes a stationary, or bound, state and cannot be expected to tend asymptotically to a free-particle wave packet. More generally, if ψ is not orthogonal to every eigenvector for H, then $\psi(t)$ will have a component which behaves like (1.3) and the asymptotic behavior (1.2) cannot be expected. Hence, it is natural to restrict the initial state ψ to a class $\mathfrak{M}$ of "scattered" or "unbound" states. $\mathfrak{M}$ is defined here, following Kuroda,[3] as the set

$$\mathfrak{M} = \{\psi : (E(\lambda)\psi, \psi) \text{ is absolutely continuous on } -\infty < \lambda < \infty\}.$$

Obviously, $\mathfrak{M}$ contains no eigenvectors for H, since $(E(\lambda)\psi, \psi)$ has a discontinuity when ψ is an eigenvector of H.

Kuroda[3] has proved that (i) $\mathfrak{M}$ is a closed linear subspace of $\mathfrak{H}$, and (ii) the direct sum decomposition $\mathfrak{H} = \mathfrak{M} \oplus \mathfrak{M}^{\perp}$ reduces H; i.e., $\mathfrak{M}$ and $\mathfrak{M}^{\perp}$ are mapped into themselves by H (and therefore by e^{-itH}). It follows that $\psi(t) = e^{-itH}\psi$ is always orthogonal to the bound states if $\psi \in \mathfrak{M}$.

In what follows the projection of $\mathfrak{H}$ onto $\mathfrak{M}$ is denoted by P. The corresponding subspace and projection for the operator H_0 are denoted by $\mathfrak{M}_0$ and P_0. The wave packets $\psi(t)$ and $\psi_0(t)$ are restricted to the subspaces $\mathfrak{M}$ and $\mathfrak{M}_0$, respectively, by requiring their initial values to have the form $\psi = P\phi$ and $\psi_0 = P_0\phi_0$.

Operator e^{-itH} is a unitary operator on $\mathfrak{H}$, with inverse e^{itH}. Hence

$$||e^{-itH}\psi - e^{-itH_0}\psi_0||_{\mathfrak{H}} = ||\psi - e^{itH}e^{-itH_0}\psi_0||_{\mathfrak{H}} \quad \text{for all} \quad \psi, \psi_0 \in \mathfrak{H}.$$

It follows that $\psi(t) \sim \psi_0(t)$ when $t \to \infty$, with $\psi_0 = P_0\psi_0$, if and only if

$$\lim_{t\to\infty} e^{itH}e^{-itH_0}P_0\phi_0 \text{ exists in } \mathfrak{H}.$$

If this limit exists for each $\phi_0 \in \mathfrak{H}$ it defines an operator

$$W_+ = W_+(H, H_0) = \text{s}\text{-}\lim_{t\to\infty} e^{itH}e^{-itH_0}P_0 \tag{1.4}$$

on $\mathfrak{H}$ (s − lim stands for "strong limit"). Similarly, $\psi(t) \sim \psi_0(t)$ when $t \to -\infty$, with $\psi_0 = P_0\phi_0$, if and only if

$$\lim_{t\to-\infty} e^{itH}e^{-itH_0}P_0\phi_0 \text{ exists in } \mathfrak{H}.$$

The corresponding operator is

$$W_- = W_-(H, H_0) = \text{s}\text{-}\lim_{t\to-\infty} e^{itH}e^{-itH_0}P_0. \tag{1.5}$$

W_+ and W_- are called the Møller wave operators for the scattering problem.[8]

Equations (1.4) and (1.5) may be written more concisely as

$$W_\pm = W_\pm(H, H_0) = \text{s-}\lim_{t\to\pm\infty} e^{itH}e^{-itH_0}P_0. \tag{1.6}$$

Henceforth this notation is used to discuss the cases $t \to +\infty$ and $t \to -\infty$ simultaneously.

Kuroda[3] has shown that the wave operators $W_\pm$ have a number of properties that follow directly from the definition (1.6). Two of these properties which are needed below are quoted here as

Theorem 1.1 (*Kuroda*). Let the wave operators $W_\pm = W_\pm(H, H_0)$ exist. Then $W_\pm$ is a partial isometry[3] with initial set $\mathfrak{M}_0$ and final set contained in $\mathfrak{M}$; i.e.,

$$W_\pm^* W_\pm = P_0 \quad \text{and} \quad W_\pm\mathfrak{H} \subset \mathfrak{M}. \tag{1.7}$$

The wave operators satisfy the operator equations

$$HW_\pm = W_\pm H_0 P_0. \tag{1.8}$$

It was shown above that $\psi_0(t) = e^{-itH_0}P_0\phi_0 \sim \psi(t) = e^{-itH}\psi$ when $t \to \pm\infty$ for each $\phi_0 \in \mathfrak{H}$ if and only if the wave operators $W_\pm$ exist and $\psi = W_\pm\phi_0$. The existence of $W_\pm$ is dependent on the structure of the operators H_0 and $V = H - H_0$. In the classical case, where $\mathfrak{H} = L_2(R^n)$, H_0 is (a self-adjoint extension of) the Laplacian on R^n and V is the operator corresponding to multiplication by a potential function $V(x)$, sufficient conditions for the existence of $W_\pm(H, H_0)$ have been given by several authors. In 1957, Cook[1] obtained the sufficient conditions $n = 3$, $V(x) \in L_2(R^3)$. Subse-

[8] T.-Y. Wu and T. Ohmura, *Quantum Theory of Scattering*, (Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1962), Chap. 4.

quently, Hack[2] obtained the sufficient conditions $n = 3$, V square-integrable on bounded sets and $V(x) = 0(|x|^{-\eta})$ with $\eta > 1$. Kuroda[3] obtained $V(x)(1 + |x|)^{-n/2+1+\epsilon} \in L_2(R^n)$, and further generalizations were given by Stankevich,[4] Brownell,[5] and others. These conditions are sufficient to guarantee that $\psi(x, t) - \psi_0(x, t) \to 0$ in $L_2(R^n)$. In this paper, the wave operators $W_\pm$ are assumed to exist and additional conditions on $V(x)$ and $\psi_0(x)$ are presented which guarantee that $\psi_0(x, t)$ and $\psi(x, t)$ have square-integrable (or continuous) derivatives up to a prescribed order whose differences tend to zero in $L_2(R^n)$ (or uniformly in R^n) when $t \to \pm\infty$. The remainder of this introduction contains a description of the principal results obtained in the paper, together with an outline of the methods that are used to obtain them.

The abstract scattering problem formulated above is considered in Sec. 2. It is shown that if $\psi \in D(H^m)$, the domain of definition of the mth power of H, then $\psi(t) = e^{-itH}\psi \in D(H^m)$ for every t and, moreover, $\psi(t)$ has time derivatives $D_t^k\psi(t)$ (with values in $\mathfrak{H}$) of orders $k = 0, 1, \cdots, m$. This result is used to show that if $\phi_0 \in D(H_0^m)$ and $W_\pm$ exist then $W_\pm\phi_0 \in D(H^m)$, so that $\psi_0(t) = e^{-itH_0}P_0\phi_0$ and $\psi_+(t) = e^{-itH}W_\pm\phi_0$ have time derivatives $D_t^k\psi_0(t)$ and $D_t^k\psi_\pm(t)$ for $k = 0, 1, \cdots, m$, and these are asymptotically equal in $\mathfrak{H}$ for $t \to \pm\infty$; i.e.,

$$D_t^k\psi_\pm(t) \sim D_t^k\psi_0(t) \quad \text{for} \quad t \to \pm\infty \quad \text{and} \quad k = 0, 1, \cdots, m.$$

In Secs. 3, 4, and 5, $\mathfrak{H} = L_2(R^n)$, $H_0 = -\Delta$ where Δ is a self-adjoint extension of the Laplace operator for R^n, and $V = H - H_0$ is the operator corresponding to multiplication by a scalar potential $V(x)$. Section 3 contains a discussion of the operators H_0^m and H^m. The principal results are the "coerciveness inequalities" which imply that functions $\psi(x)$ in $D(H_0^m)$ or $D(H^m)$ necessarily have derivatives of order $2m$ in $L_2(R^n)$.

The following notation is used in discussing partial derivatives of higher order:

$D_j = \partial/\partial x_j, \quad j = 1, 2, \cdots, n;$

$\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_n)$, where each α_j is a non-negative integer;

$D_x^\alpha = D_1^{\alpha_1}D_2^{\alpha_2} \cdots D_n^{\alpha_n}$;

$|\alpha| = \alpha_1 + \alpha_2 + \cdots + \alpha_n =$ order of D_x^α.

If $\psi(x) \in L_2(R^n)$ then $D_x^\alpha\psi(x)$ is said to exist in $L_2(R^n)$ and equal $\theta_\alpha(x) \in L_2(R^n)$ if

$$\int_{R^n} \psi(x)\, D_x^\alpha\phi(x)\, dx = (-1)^{|\alpha|}\int_{R^n} \theta_\alpha(x)\phi(x)\, dx$$

for all "test functions" $\phi(x)$ which have continuous derivatives of all orders and vanish outside a bounded set. The set

$$L_2^m(R^n) = \{\psi : D_x^\alpha\psi \in L_2(R^n) \text{ for } |\alpha| \le m\}$$

is a Hilbert space with respect to the norm

$$||\psi||_m = \left(\int_{R^n} \sum_{|\alpha|\le m} |D_x^\alpha\psi(x)|^2\, dx\right)^{\frac{1}{2}}.$$

Functions $\psi \in D(H_0^m)$ necessarily have square-integrable derivatives $D_x^\alpha\psi$ of order $|\alpha| \le 2m$; i.e., $D(H_0^m) \subset L_2^{2m}(R^n)$. This is a consequence of an inequality of the following form.

$$||\psi||_{2m}^2 \le c(||H_0^m\psi||_0^2 + ||\psi||_0^2) \text{ for } \psi \in D(H_0^m). \quad (1.9)$$

A systematic study of such inequalities was initiated by Aronszajn.[9] In his terminology the integro-differential form on the right-hand side of (1.9) is "coercive" if (1.9) holds. (If the form has a finite value then all the derivatives appearing in it are "forced" to have a finite norm.) In this paper (1.9) is called the "coerciveness inequality" for the operator H_0^m. Coerciveness inequalities have been proved for a variety of differential operators, boundary conditions and domains; see Agmon,[10] Schechter,[11] and de Figueiredo.[12]

In Sec. 3 a simple proof of (1.9), based on the Fourier transform, is given. The principal result of Sec. 3 is the corresponding inequality for $H^m = (H_0 + V)^m$.

In Sec. 4 the results of Secs. 2 and 3 are combined to find conditions on $V(x)$ and $\psi_0(x)$ which imply that $\psi_0(x, t)$ and $\psi(x, t)$ have space–time derivatives of a prescribed order which are asymptotically equal in $L_2(R^n)$. In particular, it is shown that if $W_\pm$ exist, $\psi_0 \in D(H_0^m)$ and the derivatives $D_x^\alpha V(x)$ are continuous and bounded for $|\alpha| \le 2m - 2$ then the derivatives $D_t^kD_x^\alpha\psi_0(x, t)$ and $D_t^kD_x^\alpha\psi_\pm(x, t)$ exist in $L_2(R^n)$ for $k + |\alpha| \le m$ and

$$D_t^kD_x^\alpha\psi_\pm(x, t) \sim D_t^kD_x^\alpha\psi_0(x, t) \quad \text{for} \quad t \to \pm\infty \quad \text{and} \quad k + |\alpha| \le m.$$

A well-known version of Sobolev's imbedding

[9] N. Aronszajn, *On Coercive Integro-Differential Quadratic Forms*, Conf. on Partial Differential Equations, Univ. of Kansas Rept. No. 14 (1954), 94–106.

[10] S. Agmon, J. Anal. Math. **6**, 183 (1958).

[11] M. Schechter, Comm. Pure Appl. Math. **11**, 153 (1958).

[12] D. G. de Figueiredo, Comm. Pure Appl. Math. **16**, 63 (1963).

theorem[13] states that if $\psi \in L_2^m(R^n)$ and $m \geq [n/2] + k + 1$, then ψ has bounded continuous derivatives of orders $\leq k$ in R^n. In Sec. 5 this theorem and the results of Sec. 4 are combined to obtain conditions which imply that $\psi_0(x, t)$ and $\psi_\pm(x, t)$ have continuous partial derivatives of a prescribed order whose difference tends to zero uniformly in R^n when $t \to \pm\infty$.

2. ASYMPTOTIC ESTIMATES FOR TIME DERIVATIVES

This section is concerned with the abstract scattering problem formulated in Sec. 1. It was shown there that if the wave operators $W_\pm = W_\pm(H, H_0)$ exist then the wave packets $\psi_\pm(t) = e^{-itH}W_\pm\phi_0$ and $\psi_0(t) = e^{-itH_0}P_0\phi_0$, corresponding to scattered particles and free particles, respectively, are asymptotically equal in $\mathfrak{H}$ for $t \to \pm\infty$. The purpose of this section is to derive additional conditions which imply that the time derivatives $D_t^m\psi_\pm(t)$ and $D_t^m\psi_0(t)$ (exist and) are asymptotically equal in $\mathfrak{H}$ for $t \to \pm\infty$.

The calculus of functions with values in a Hilbert space $\mathfrak{H}$ is used in what follows. A function $\psi(t)$ with values in $\mathfrak{H}$ has a derivative $D_t\psi(\tau) \in \mathfrak{H}$ at $t = \tau$ if and only if

$$h^{-1}[\psi(\tau + h) - \psi(\tau)] \to D_t\psi(\tau) \text{ in } \mathfrak{H} \text{ when } h \to 0.$$

Similarly, $\psi(t)$ is continuous at $t = \tau$ if and only if $\psi(t) \to \psi(\tau)$ in $\mathfrak{H}$ when $t \to \tau$. The class of functions on $-\infty < t < \infty$ with values in $\mathfrak{H}$ which are continuous together with their derivatives of orders $k = 1, 2, \cdots, m$ is denoted by $C^m(-\infty, \infty; \mathfrak{H})$. The usual rules of calculus hold for functions in this class.[14]

After these preliminaries, the main result of this section may be stated as follows.

Theorem 2.1. Let m be a positive integer, let $\phi_0 \in D(H_0^m)$, and assume that the wave operators $W_\pm = W_\pm(H, H_0)$ exist. Then

$$\psi_0(t) \text{ and } \psi_\pm(t) \text{ are in } C^m(-\infty, \infty; \mathfrak{H}), \tag{2.1}$$

$$\psi_0(t) \in D(H_0^m) \text{ and } \psi_\pm(t) \in D(H^m) \text{ for } -\infty < t < \infty, \tag{2.2}$$

$$D_t^k\psi_0(t) = (-iH_0)^k\psi_0(t) = e^{-itH_0}P_0(-iH_0)^k\phi_0, \tag{2.3}$$

$$D_t^k\psi_\pm(t) = (-iH)^k\psi_\pm(t) = e^{-itH}(-iH)^kW_\pm\phi_0$$

$$\text{for } k = 0, 1, \cdots, m,$$

and

$$\lim_{t\to\pm\infty} ||D_t^k\psi_\pm(t) - D_t^k\psi_0(t)||_{\mathfrak{H}} = 0$$

$$\text{for } k = 0, 1, \cdots, m. \tag{2.4}$$

The proof of Theorem 2.1 is based on the following two lemmas.

Lemma 2.1. If $W_\pm = W_\pm(H, H_0)$ exist then

$$H^mW_\pm = W_\pm H_0^mP_0 \quad \text{for} \quad m = 1, 2, 3, \cdots. \tag{2.5}$$

Proof. For $m = 1$ this result is contained in Kuroda's theorem (Theorem 1.1). The general case follows easily by induction on m.

Lemma 2.2. Let H be an arbitrary self-adjoint operator on $\mathfrak{H}$, and let $\psi \in D(H^m)$. Then $\psi(t) = e^{itH}\psi \in D(H^m)$ for $-\infty < t < \infty$. Moreover,

$$\psi(t) \in C^m(-\infty, \infty; \mathfrak{H}) \text{ and } D_t^k\psi(t) = (-iH)^k\psi(t)$$

$$\text{for } k = 1, 2, \cdots, m.$$

Proof. The following facts are used.[7] If $f(\lambda)$ is a complex-valued function, defined for $-\infty < \lambda < \infty$ and $E(\lambda)$ is the resolution of the identity corresponding to H, then $f(H)$ is the operator with domain

$$D(f(H)) = \left\{\psi : \int_{-\infty}^{\infty} |f(\lambda)|^2\, d\, ||E(\lambda)\psi||^2 < \infty\right\}$$

defined by

$$f(H) = \int_{-\infty}^{\infty} f(\lambda)\, dE(\lambda).$$

Moreover,

$$||f(H)\psi||^2 = \int_{-\infty}^{\infty} |f(\lambda)|^2\, d\, ||E(\lambda)\psi||^2$$

$$\text{for } \psi \in D(f(H)). \tag{2.3}$$

In particular,

$$\psi \in D(H^m) \text{ if and only if } \int_{-\infty}^{\infty} \lambda^{2m}\, d\, ||E(\lambda)\psi||^2 < \infty. \tag{2.4}$$

Now

$$E(\lambda)\psi(t) = E(\lambda)\, e^{-itH}\psi = e^{-itH}E(\lambda)\psi$$

because $E(\lambda)$ commutes with functions of H. Hence

$$||E(\lambda)\psi(t)|| = ||E(\lambda)\psi|| \tag{2.5}$$

since e^{-itH} is unitary. Equations (2.4) and (2.5) imply that $\psi(t) \in D(H^m)$ if and only if $\psi \in D(H^m)$.

Next, notice that if $\psi \in D(H^m)$ then $(-itH)^k\psi(t)$ satisfies

[13] S. L. Sobolev, *Applications of Functional Analysis in Mathematical Physics*, Translations of Mathematical Monographs (American Mathematical Society, Providence, Rhode Island, 1963), Vol. 7.

[14] E. Hille, and R. S. Phillips, *Functional Analysis and Semi-Groups*, Colloquium Publications (American Mathematical Society, Providence, Rhode Island, 1957), Vol. 31.

$$||(-iH)^k\psi(t+\tau) - (-iH)^k\tau(t)||^2 = \int_{-\infty}^{\infty} |e^{-i\tau\lambda} - 1|^2 \lambda^{2k}\, d\,||E(\lambda)\psi||^2 \qquad (2.6)$$

for $k = 0, 1, \cdots, m$ by (2.3). The integrand in (2.6) tends to zero when $\tau \to 0$ and is bounded by a constant multiple of λ^{2k} which is integrable with respect to $d\,||E(\lambda)\psi||^2$ for $k = 0, 1, \cdots, m$. Hence, (2.6) tends to zero with τ, by Lebesgue's dominated convergence theorem, and it follows that

$$(-iH)^k\psi(t) \in C(-\infty, \infty; \mathfrak{H}) \text{ for } k = 0, 1, \cdots, m.$$

The proof of Lemma 2.2 is completed by showing that $D_t^k\psi(t)$ exists and equals the continuous function $(-iH)^k\psi(t)$ for $k = 1, 2, \cdots, m$. For the case $k = 1$, (2.3) implies

$$||[\psi(t+\tau) - \psi(t)]/\tau - (-iH)\tau(t)||^2 = \int_{-\infty}^{\infty} \left|\frac{e^{-i\tau\lambda} - 1}{\tau\lambda} + i\right|^2 \lambda^2\, d\,||E(\lambda)\psi||^2.$$

Moreover, the integrand of this integral is bounded by a constant multiple of λ^2, and tends to zero when $\tau \to 0$. Hence if $\psi \in D(H)$ then $D_t\psi(t)$ exists and equals $(-iH)\psi(t)$, by Lebesgue's theorem. This completes the proof for the case $m = 1$. The general case follows easily by induction on m.

Proof of Theorem 2.1. If $\phi_0 \in D(H_0^m)$ then $\psi_\pm = W_\pm\phi_0 \in D(H^m)$, by Lemma 2.1. Hence, Lemma 2.2 implies that both $\psi_0(t)$ and $\psi_\pm(t)$ are in $C^m(-\infty, \infty; \mathfrak{H})$ which proves (2.1). Moreover, Lemma 2.2 implies $\psi_0(t) \in D(H_0^m)$, $\psi_\pm(t) \in D(H^m)$ and

$$D_t^k\psi_0(t) = (-iH_0)^k e^{-itH_0}P_0\phi_0 = e^{-itH_0}P_0(-iH_0)^k\phi_0,$$

and

$$D_t^k\psi_\pm(t) = (-iH)^k e^{-itH}W_\pm\phi_0 = e^{-itH}(-iH)^kW_\pm\phi_0,$$

for $k = 0, 1, \cdots, m$ which proves (2.2) and (2.3). But $H^kW_\pm = W_\pm H_0^kP_0$ for $k = 1, \cdots, m$ by Lemma 2.1. Thus for $k = 0, 1, \cdots, m$

$$D_t^k\psi_0(t) = e^{-itH_0}P_0\phi_k \text{ and } D_t^k\psi_\pm(t) = e^{-itH}W_\pm\phi_k, \qquad (2.7)$$

where $\phi_k = (-iH_0)^k\phi_0$. Hence

$$||D_t^k\psi_\pm(t) - D_t^k\psi_0(t)||_{\mathfrak{H}} = ||W_\pm\phi_k - e^{itH}e^{-itH_0}P_0\phi_k||_{\mathfrak{H}} \to 0 \text{ when } t \to \pm\infty$$

for $k = 0, 1, \cdots, m$ because of the assumption that $W_\pm$ exists. This completes the proof of Theorem 2.1.

3. COERCIVENESS ESTIMATES FOR THE OPERATORS H_0^m AND H^m

The remainder of this paper deals with the classical problem of the scattering of nonrelativistic particles by a potential. Thus $\mathfrak{H} = L_2(R^n)$, H_0 is (a self-adjoint extension of) the operator $-\Delta = -D_1^2 - \cdots - D_n^2$ and $H - H_0 = V$ is the operator corresponding to multiplication by a real-valued potential function $V(x)$.

In this section, the operators H_0 and H are defined and a number of their properties are derived. The main result is the coerciveness inequality for H^m, Theorem 3.2.

The free particle Hamiltonian H_0 is discussed by means of the Plancherel theory of the Fourier transform. The relevant facts concerning the transform are summarized here without proofs.[15]

The Fourier transform establishes a unitary mapping $\psi \to \psi^\star$ of $L_2(R^n)$ onto itself defined by

$$\psi^\star(p) = \lim_{R\to\infty} \frac{1}{(2\pi)^{n/2}} \int_{|x|\le R} e^{-i\langle p,x\rangle}\psi(x)\, dx,$$

$$\psi(x) = \lim_{R\to\infty} \frac{1}{(2\pi)^{n/2}} \int_{|p|\le R} e^{i\langle x,p\rangle}\psi^\star(p)\, dp,$$

where $p \in R^n$, $\langle p, x\rangle = p_1x_1 + \cdots + p_nx_n$ and the limits exist in the $L_2(R^n)$ sense (mean-square convergence). The unitary property of the transform is expressed by

Parseval's Theorem.

$$\int_{R^n} |\psi(x)|^2\, dx = \int_{R^n} |\psi^\star(p)|^2\, dp \text{ for every } \psi \in L_2(R^n).$$

If ψ and $D_j\psi$ are in $L_2(R^n)$ it can be shown that $\psi^\star$ and $p_j\psi^\star(p)$ are in $L_2(R^n)$ and $[D_j\psi(p)]^\star = ip_j\psi(p)$. This is a special case of

Lemma 3.1. Let $P(D) = \sum_{|\alpha|\le m} c_\alpha D^\alpha$ be a differential operator with constant coefficients, and let $\psi \in L_2(R^n)$. Then $P(D)\psi \in L_2(R^n)$ if and only if $P(ip)\psi^\star(p) = \sum_{|\alpha|\le m} c_\alpha i^{|\alpha|}p^\alpha\psi^\star(p) \in L_2(R^n)$, where $p^\alpha = p_1^{\alpha_1}\cdots p_n^{\alpha_n}$. Moreover, $[P(D)\psi(p)]^\star = P(ip)\psi^\star(p)$ if either $P(D)\psi$ or $P(ip)\psi^\star(p)$ is in $L_2(R^n)$.

This result is easy to verify by means of Parseval's theorem. The proof is omitted.

Lemma 3.1 implies that

$$L_2^m(R^n) = \{\psi : p^\alpha\psi^\star(p) \in L_2(R^n) \text{ for } |\alpha| \le m\}.$$

Moreover, by Parseval's theorem,

$$||\psi||_m^2 = \int_{R^n} \sum_{|\alpha|\le m} |D^\alpha\psi(x)|^2\, dx = \int_{R^n} \sum_{|\alpha|\le m} |p^\alpha|^2\, |\psi^\star(p)|^2\, dp. \qquad (3.1)$$

[15] For an exposition of the Plancherel theory see, for example, S. Bochner and K. Chandrasekharan, *Fourier Transforms*, Annals of Mathematical Studies (Princeton University Press, Princeton, New Jersey, 1949), No. 19.

Now, the free-particle Hamiltonian is formally $-\Delta = -D_1^2 - \cdots - D_n^2$. Thus if $\psi \in L_2(R^n)$ then $\Delta\psi \in L_2(R^n)$ if and only if

$$\psi^\star \text{ and } (p_1^2 + \cdots + p_n^2)\psi^\star(p) \in L_2(R^n),$$

which is obviously equivalent to $p^\alpha\psi^\star(p) \in L_2(R^n)$ for $|\alpha| \leq 2$. This motivates the

Definition. H_0 is the operator with domain $D(H_0) = L_2^2(R^n)$ defined by $H_0\psi = -D_1^2\psi - \cdots - D_n^2\psi$.

Lemma 3.2. The operator H_0 is a self-adjoint operator on $L_2(R^n)$, i.e., $H_0^* = H_0$.

Proof. The conclusion $H_0 \subset H_0^*$ is immediate because H_0 is symmetric.[16] To prove $H_0 \supset H_0^*$ let $\phi \in D(H_0^*)$; i.e.,

$$(H_0\psi, \phi)_{L_2(R^n)} = (\psi, \theta)_{L_2(R^n)} \quad \text{for some} \quad \theta \in L_2(R^n) \quad \text{and all} \quad \psi \in D(H_0).$$

Then Parseval's theorem gives

$$\int_{R^n} (p_1^2 + \cdots + p_n^2)\psi^\star(p)\bar{\phi}^\star(p)\,dp = \int_{R^n} \psi^\star(p)\bar{\theta}^\star(p), \quad \text{for all} \quad \psi \in L_2^2(R^n),$$

whence

$$\theta^\star(p) = (p_1^2 + \cdots + p_n^2)\phi^\star(p) \in L_2(R^n).$$

It follows that $\phi \in L_2^2(R^n) = D(H_0)$ and $\theta = H_0\phi$, which completes the proof.

The spectral family of the operator H_0 may be described, by means of the Fourier transform, by the formula

$$E_0(\lambda)\psi(x) = \begin{cases} \dfrac{1}{(2\pi)^{n/2}} \displaystyle\int_{|p|^2 \leq \lambda} e^{i(x,p)}\psi^\star(p)\,dp, & \lambda > 0, \\ 0, & \lambda \leq 0. \end{cases} \tag{3.2}$$

The correctness of this formula can be verified by direct computation. Notice that, by Parseval's theorem,

$$(E_0(\lambda)\psi, \psi) = \int_{|p|^2 \leq \lambda} |\psi^\star(p)|^2\,dp$$

is an absolutely continuous function of λ for each $\psi \in L_2(R^n)$. Hence $\mathfrak{M}_0 = L_2(R^n)$ and $P_0 = I$, the identity operator.

The spectral family may be used to define functions of the operator H_0. Thus if $f(\lambda)$ is defined for real $\lambda \geq 0$ then direct computation using (3.2) gives

$$f(H_0)\psi(x) = \int_{-\infty}^{\infty} f(\lambda)\,dE_0(\lambda)\psi(x) = \frac{1}{(2\pi)^{n/2}} \int_{R^n} e^{i(x,p)} f(|p|^2)\psi^\star(p)\,dp$$

and

$$||f(H_0)\psi||^2 = \int_{-\infty}^{\infty} |f(\lambda)|^2\,d\,||E_0(\lambda)\psi||^2 = \int_{R^n} |f(|p|^2)|^2\,|\psi^\star(p)|^2\,dp.$$

The domain of $f(H_0)$ is precisely the set of ψ for which the last integral is finite. In particular, the operator H_0^m has domain $D(H_0^m) = \{\psi : \psi^\star(p)$ and $|p|^{2m}\psi^\star(p)$ are in $L_2(R^n)\}$. Now $|p_i| \leq |p|$, whence $|p^\alpha| = |p_1^{\alpha_1} \cdots p_n^{\alpha_n}| \leq |p|^{|\alpha|}$. It follows that $\psi \in D(H_0^m)$ if and only if $p^\alpha\psi^\star(p) \in L_2(R^n)$ for $|\alpha| \leq 2m$; i.e., $D(H_0^m) = L_2^{2m}(R^n)$. Moreover, the elementary inequality $|p|^{2|\alpha|} \leq |p|^{4m} + 1$ which is valid for all p and $|\alpha| \leq 2m$ implies

$$||\psi||_{2m}^2 = \int_{R^n} \sum_{|\alpha| \leq 2m} |p^\alpha|^2\,|\psi^\star(p)|^2\,dp \leq c\int_{R^n} (|p|^{4m} + 1)\,|\psi^\star(p)|^2\,dp = c(||H_0^m\psi||^2 + ||\psi||^2),$$

where c is a constant which depends on m and n only. This completes the proof of

Lemma 3.3. For each integer $m \geq 1$, $D(H_0^m) = L_2^{2m}(R^n)$ and there exists a constant $c = c(m, n)$ such that

$$||\psi||_{2m}^2 \leq c(||H_0^m\psi||^2 + ||\psi||^2), \quad \text{for every} \quad \psi \in D(H_0^m). \tag{3.3}$$

A related result which is needed below is

Lemma 3.4. For each integer $m \geq 1$ and each $\epsilon > 0$, there exists a constant $c = c(\epsilon, m, n)$ such that

$$||\psi||_{2m-1}^2 \leq \epsilon^2\,||H_0^m\psi||^2 + c^2\,||\psi||^2 \quad \text{for every} \quad \psi \in D(H_0^m). \tag{3.4}$$

Proof. If $|\alpha| < 2m$ and $\delta > 0$ it can be shown by an elementary argument that there exists a constant $\gamma(\delta)$ such that $|p^\alpha|^2 \leq \delta|p|^{4m} + \gamma(\delta)$ for every p. If $\psi \in D(H_0^m)$ it follows that

$$||\psi||_{2m-1}^2 = \int_{R^n} \sum_{|\alpha| \leq 2m-1} |p^\alpha|^2\,|\psi^\star(p)|^2\,dp \leq c(m, n)(\delta\,||H_0^m\psi||^2 + \gamma(\delta)\,||\psi||^2),$$

which implies (3.4).

[16] For an explanation of this notation see Riesz and Nagy, Footnote 7, Chap. 8.

A final result concerning the operator H_0 is

Lemma 3.5. Let I denote the identity operator on $L_2(R^n)$ and let m be a positive integer. Then there exists a constant $c = c(m, n)$ such that

$$||\psi||_m^2 \leq c((I + H_0)^m\psi, \psi) \quad \text{for every} \quad \psi \in D(H_0^m). \tag{3.5}$$

Proof. By Lemma 3.1,

$$[(I + H_0)^m\psi(p)]^\star = (1 + |p|^2)^m \psi^\star(p).$$

Hence, by Parseval's theorem,

$$((I + H_0)^m\psi, \psi) = \int_{R^n} (1 + |p|^2)^m |\psi^\star(p)|^2 dp. \tag{3.6}$$

Now, $|p^\alpha| \leq |p|^{|\alpha|}$, whence

$$\sum_{|\alpha|\leq m} |p^\alpha|^2 \leq \sum_{|\alpha|\leq m} |p|^{2|\alpha|} \leq \sum_{|\alpha|\leq m} (1 + |p|^2)^{|\alpha|} \leq c(1 + |p|^2)^m$$

for a suitable value of c. Multiplying this by $|\psi(p)|^2$, integrating over R^n and using (3.1) and (3.6) gives (3.5).

The Hamiltonian operator $H = H_0 + V$ is discussed next. The precise definition of this operator is based on the following theorem due to Kato.[17]

Theorem of Kato. Suppose that there exist constants a and b, with $0 \leq a < 1$ and $b \geq 0$, such that

$$||V\psi|| \leq a\,||H_0\psi|| + b\,||\psi|| \text{ for every } \psi \in D(H_0). \tag{3.7}$$

Then the operator H defined by $D(H) = D(H_0)$, $H = H_0 + V$ is self-adjoint.

It is known[5] that (3.7) holds if $V \in L_p(R^n)$ with $p \geq 2$ and $p > n/2$. In particular, (3.7) holds (with $a = 0$) if V is bounded. Criteria that ensure $D(H^m) = D(H_0^m)$ also are needed below. They can be derived from

Lemma 3.6. Let m by a positive integer and let $V \in C^{2m-2}(R^n)$. Then

$$H^m\psi = H_0^m\psi + mVH_0^{m-1}\psi + \sum_{|\alpha|\leq 2m-3} P_\alpha^{(2m-2)} D_x^\alpha\psi \text{ for every } \psi \in D(H_0^m), \tag{3.8}$$

where $P_\alpha^{(2m-2)}$ is a polynomial in V and its derivatives $D^\beta V$ of order $|\beta| \leq 2m - 2$ with constant coefficients.

Proof. Direct calculation gives

$$H\psi = H_0\psi + V\psi,$$
$$H^2\psi = H_0^2\psi + 2VH_0\psi + \left(V^2\psi + (H_0V)\psi - 2\sum_{i=1}^n D_iVD_i\psi\right), \tag{3.9}$$

which verifies (3.8) for $m = 1$ and 2. The proof is completed by induction on m. The identity

$$H_0(\phi\psi) = (H_0\phi) - 2\sum_{i=1}^n D_i\phi\, D_i\psi + \phi(H_0\psi) \tag{3.10}$$

is used. If $Q^{(m)}$ is a generic symbol for a polynomial in V and its derivatives $D^\beta V$ of order $|\beta| \leq m$, with constant coefficients, then (3.10) implies

$$H_0(VH_0^{m-1}\psi) = VH_0^m\psi + \sum_{|\alpha|\leq 2m-1} Q_\alpha^{(2)} D_x^\alpha\psi$$

and

$$H_0 \sum_{|\alpha|\leq 2m-3} P_\alpha^{(2m-2)} D_x^\alpha\psi = \sum_{|\alpha|\leq 2m-3} Q_\alpha^{(2m)} D_x^\alpha\psi + \sum_{|\alpha|\leq 2m-2} Q_\alpha^{(2m-1)} D_x^\alpha\psi + \sum_{|\alpha|\leq 2m-1} Q_\alpha^{(2m-2)} D_x^\alpha\psi = \sum_{|\alpha|\leq 2m-1} Q_\alpha^{(2m)} D_x^\alpha\psi.$$

Thus, if (3.8) holds for a certain value of m then

$$H_0H^m\psi = H_0^{m+1}\psi + mVH_0^m\psi + \sum_{|\alpha|\leq 2m-1} Q_\alpha^{(2m)} D_x^\alpha\psi$$

and

$$VH^m\psi = VH_0^m\psi + mV^2H_0^{m-1}\psi + \sum_{|\alpha|\leq 2m-3} Q_\alpha^{(2m-2)} D_x^\alpha\psi = VH_0^m\psi + \sum_{|\alpha|\leq 2m-2} Q_\alpha^{(2m-2)} D_x^\alpha\psi.$$

Adding these expressions gives

$$H^{m+1}\psi = H_0^{m+1}\psi + (m + 1)VH_0^m\psi + \sum_{|\alpha|\leq 2m-1} Q_\alpha^{(2m)} D_x^\alpha\psi$$

which is equivalent to (3.8) with $m + 1$ in place of m.

Combining Lemmas 3.4 and 3.6 gives

Theorem 3.1. Let m be a positive integer and let V and its derivatives $D_x^\beta V$ of order $|\beta| \leq 2m - 2$ be continuous and bounded in R^n. Then for each $a > 0$ there exists a constant c, depending on a, m, n, and the bounds for the $D_x^\beta V$, $|\beta| \leq 2m - 2$, such that

$$||H^m\psi - H_0^m\psi|| \leq a\,||H_0^m\psi|| + c\,||\psi|| \quad \text{for every} \quad \psi \in D(H_0^m). \tag{3.11}$$

Proof. The hypotheses imply that Eq. (3.8) of Lemma 3.6 is valid and the functions V and $P_\alpha^{(2m-2)}$ are bounded. Moreover, H_0^{m-1} is a differential operator of order $2m - 2$. Thus, applying the triangle

[17] T. Kato, Trans. Amer. Math. Soc. 70, 195 (1951).

inequality to (3.8) and estimating V and $P_\alpha^{(2m-2)}$ by an upper bound and derivatives D_x^α with $|\alpha| \leq 2m-2$ by $||\psi||_{2m-2}$ gives

$$||H^m\psi - H_0^m\psi|| \leq M\ ||\psi||_{2m-2}, \tag{3.12}$$

where M is a constant. Now, Lemma 3.4 and the elementary inequality $(a^2 + b^2)^{\frac{1}{2}} \leq |a| + |b|$ imply that given any $\epsilon > 0$ there exists a constant $c' = c'(\epsilon, m, n)$ such that

$$||\psi||_{2m-2} \leq ||\psi||_{2m-1} \leq \epsilon\ ||H_0^m\psi|| + c'\ ||\psi||.$$

Combining this estimate with (3.12) gives (3.11) with $a = M\epsilon$ and $c = Mc'$.

Theorem 3.2. Under the hypotheses of Theorem 3.1, $D(H^m) = D(H_0^m) = L_2^{2m}(R^n)$ and there exists a constant c, depending on m, n and the bounds for the $D^\beta V$, $|\beta| \leq 2m - 2$, such that

$$||\psi||_{2m}^2 \leq c(||H^m\psi||^2 + ||\psi||^2) \text{ for every } \psi \in D(H^m). \tag{3.13}$$

Proof. The statement $D(H^m) = D(H_0^m)$ follows from Kato's theorem, applied to H_0^m and H^m instead of H_0 and H. The estimate (3.11) of Theorem 3.1 is equivalent to the hypothesis (3.7) for this case.

To prove (3.13) note that (3.11) with $a = \frac{1}{2}$ implies

$$\begin{aligned} ||H_0^m\psi|| &\leq ||H^m\psi - H_0^m\psi|| + ||H^m\psi|| \\ &\leq \tfrac{1}{2}\ ||H_0^m\psi|| + c_0\ ||\psi|| + ||H^m\psi||. \end{aligned}$$

Hence

$$||H_0^m\psi|| \leq 2(||H^m\psi|| + c_0\ ||\psi||),$$

which, with the elementary inequality $(a + b)^2 \leq 2(a^2 + b^2)$ implies

$$||H_0^m\psi||^2 \leq 8(||H^m\psi||^2 + c_0^2\ ||\psi||^2).$$

Combining this with estimate (3.3) of Lemma 3.3 gives (3.13).

4. ASYMPTOTIC ESTIMATES FOR SPACE DERIVATIVES

In this section, the results of Secs. 2 and 3 are combined to obtain conditions on $\psi_0(x)$ and $V(x)$ which guarantee that $\psi_\pm(x, t) - \psi_0(x, t) \to 0$ in $L_2^m(R^n)$ when $t \to \pm\infty$. The basic estimate is described by

Theorem 4.1. Let m be a positive integer and assume that

$$W_\pm = W_\pm(H, H_0) \text{ exist;} \tag{4.1}$$

$$\psi_0 \in D(H_0^m), \text{ and} \tag{4.2}$$

$$V(x) \text{ and its derivatives } D_x^\alpha V(x) \text{ are continuous and bounded in } R^n \text{ for } |\alpha| \leq 2m - 2. \tag{4.3}$$

Then, if

$$\psi_0(t) = e^{-itH_0}\psi_0 \quad \text{and} \quad \psi_\pm(t) = e^{-itH}W_\pm\psi_0,$$

$$\psi_0(t) \quad \text{and} \quad \psi_\pm(t) \text{ are in } \bigcap_{l=0}^{m} C(-\infty, \infty, L_2^{2m-2l}(R^n)), \tag{4.4}$$

and there exists a constant c, independent of t, such that

$$||D_t^l\psi_\pm(t) - D_t^l\psi_0(t)||_{m-l}^2 \leq c\ ||D_t^l\psi_\pm(t) - D_t^l\psi_0(t)||_0$$

$$\text{for} \quad -\infty < t < \infty \quad \text{and} \quad l = 0, 1, \cdots, m. \tag{4.5}$$

The right-hand term of the last inequality tends to zero when $t \to \pm\infty$, by Theorem 2.1, which implies

Corollary 4.1. Under the hypothesis of Theorem 4.1,

$$\lim_{t\to\pm\infty} \{D_t^l\psi_\pm(x, t) - D_t^l\psi_0(x, t)\} = 0 \quad \text{in} \quad L_2^{m-l}(R^n)$$

$$\text{for} \quad l = 0, 1, \cdots, m.$$

Proof of Theorem 4.1. Hypotheses (4.1) and (4.2) are the hypotheses of Theorem 2.1. Thus $\psi_0(t)$ and $\psi_\pm(t)$ are in $C^m(-\infty, \infty, L_2(R^n))$ and for $0 \leq l \leq m$

$$D_t^k\psi_0(t) \in D(H_0^{m-l}) \quad \text{and} \quad D_t^k\psi_\pm(t) \in D(H^{m-l})$$

$$\text{for} \quad k = 0, 1, \cdots, l.$$

Hence, by Lemma 3.3, $D_t^k\psi_0(t) \in L_2^{2m-2l}(R^n)$ for $k = 0, 1, \cdots, l$ and

$$\begin{aligned} &||D_t^k\psi_0(t) - D_t^k\psi_0(\tau)||_{2m-2l}^2 \\ &\quad \leq c[||H_0^{m-l}(D_t^k\psi_0(t) - D_t^k\psi_0(\tau))||_0^2 \\ &\qquad + ||D_t^k\psi_0(t) - D_t^k\psi_0(\tau)||_0^2] \\ &\quad = c[||D_t^{m-l+k}\psi_0(t) - D_t^{m-l+k}\psi_0(\tau)||_0^2 \\ &\qquad + ||D_t^k\psi_0(t) - D_t^k\psi_0(\tau)||_0^2]. \end{aligned}$$

For $k \leq l$ the last two terms tend to zero when $t \to \tau$, by Theorem 2.1. Thus

$$\psi_0(t) \in C^l(-\infty, \infty, L_2^{2m-2}(R^n)) \text{ for } l = 0, 1, \cdots, m$$

which proves (4.4) for $\psi_0(t)$. The same argument, with Theorem 3.2 in place of Lemma 3.3, applies to $\psi_\pm(t)$.

The proof of (4.5) makes use of Lemma 3.5.

Inequality (3.5) gives

$$||D_t^l\psi_\pm(t) - D_t^l\psi_0(t)||_{m-l}^2 \le c((I + H_0)^{m-l} \times [D_t^l\psi_\pm(t) - D_t^l\psi_0(t)], D_t^l\psi_\pm(t) - D_t^l\psi_0(t))$$
$$= c\sum_{k=0}^{m-l}\binom{m-l}{k}(H_0^kD_t^l\psi_\pm(t) - H_0^kD_t^l\psi_0(t), \times D_t^l\psi_\pm(t) - D_t^l\psi_0(t)).$$

Applying Schwarz's inequality and the triangle inequality to the inner products in the last sum gives

$$||D_t^l\psi_\pm(t) - D_t^l\psi_0(t)||_{m-l}^2 \le c\left\{\sum_{k=0}^{m-l}\binom{m-l}{k}\left(||H_0^kD_t^l\psi_\pm(t)||_0 + ||H_0^kD_t^l\psi_0(t)||_0\right)\right\} \times ||D_t^l\psi_\pm(t) - D_t^l\psi_0(t)||_0. \quad (4.6)$$

Hence to complete the proof of (4.5) it is sufficient to show that the sum in braces is bounded for $-\infty < t < \infty$. This is proved below with the help of

Lemma 4.1. Let $\Lambda = \sum_{|\alpha|\le k}\Lambda_\alpha(x)D_x^\alpha$ be a partial differential operator of order k, with coefficients $\Lambda_\alpha(x)$ which are defined and bounded in R^n. Then there exists a constant c such that

$$||\Lambda\psi||_0 \le c\,||\psi||_k, \quad \text{for all} \quad \psi \in L_2^k(R^n). \quad (4.7)$$

Proof of Lemma 4.1. If $\psi \in L_2^k(R^n)$, then

$$\Lambda\psi(x) = \sum_{|\alpha|\le k}\Lambda_\alpha(x)\,D_x^\alpha\psi(x).$$

Applying the triangle inequality gives

$$||\Lambda\psi||_0 \le M\sum_{|\alpha|\le k}||D_x^\alpha\psi||_0, \quad (4.8)$$

where M is an upper bound for the coefficients Λ_α. Moreover,

$$||D_x^\alpha\psi||_0 \le \Big(\sum_{|\beta|\le k}||D_x^\beta\psi||_0^2\Big)^{\frac{1}{2}} = ||\psi||_k$$

for $|\alpha| \le k$. Combining this with (4.8) gives (4.7).

Proof of Theorem 4.1 (concluded). H_0^k is a differential operator of order $2k$ with constant coefficients. Hence, by Lemma 4.1 and Theorem 2.1,

$$||H_0^kD_t^l\psi_\pm(t)||_0 \le c\,||D_t^l\psi_\pm(t)||_{2k} = c\,||H^l\psi_\pm(t)||_{2k}. \quad (4.9)$$

Moreover, H^l is a differential operator with bounded coefficients, by Lemma 3.6. Hence, (4.9) and Lemma 4.1 imply

$$||H_0^kD_t^l\psi_\pm(t)||_0 \le c\,||\psi_\pm(t)||_{2k+2l} \le c\,||\psi_\pm(t)||_{2m}. \quad (4.10)$$

The last inequality holds because $k \le m$ and $l \le m - k$ in (4.6). Next, (4.10) and Theorem 3.2 imply

$$||H_0^kD_t^l\psi_\pm(t)||_0^2 \le c(||H^m\psi_\pm(t)||_0^2 + ||\psi_\pm(t)||_0^2). \quad (4.11)$$

But, Theorem 2.1 implies

$$||\psi_\pm(t)||_0 = ||W_\pm\psi_0||_0 = \text{constant}$$

and

$$||H^m\psi_\pm(t)||_0 = ||H^mW_\pm\psi_0||_0 = \text{constant}.$$

Hence, (4.11) implies that

$$||H_0^kD_t^l\psi_\pm(t)||_0 \le \text{constant} \quad \text{for} \quad -\infty < t < \infty.$$

Similarly,

$$||H_0^kD_t^l\psi_0(t)||_0 = ||H_0^{l+k}\psi_0(t)||_0 = ||H_0^{l+k}\psi_0|| = \text{constant}.$$

Hence, in (4.6) the term in braces is bounded for $-\infty < t < \infty$ and the proof of (4.5) is complete.

5. UNIFORM ASYMPTOTIC ESTIMATES

In this section, the results of Sec. 4 and Sobolev's imbedding theorem are combined to obtain conditions which guarantee that $\psi_\pm(x, t)$ and $\psi_0(x, t)$ have continuous derivatives of prescribed orders, and that $D_x^\alpha D_t^k\psi_\pm(x, t) - D_x^\alpha D_t^k\psi_0(x, t) \to 0$ when $t \to \pm\infty$, uniformly for $x \in R^n$. The following version of Sobolev's theorem is used.

Sobolev's Imbedding Theorem. Let $\psi \in L_2^m(R^n)$. Then

(a) If $m \ge [n/2] + 1$, then $\psi(x)$ is bounded and continuous in R^n (after correction on a null set) and there exists a constant c, independent of ψ, such that

$$\max_{x\in R^n}|\psi(x)| \le c\,||\psi||_m. \quad (5.1)$$

(b) If $m \ge [n/2] + k + 1$ then $\psi(x)$ has continuous derivatives $D_x^\alpha\psi(x)$ of order $|\alpha| \le k$ (after correction on a null set) and there exists a constant c, independent of ψ, such that

$$\max_{\substack{x\in R^n\\ |\alpha|\le k}}|D_x^\alpha\psi(x)| \le c\,||\psi||_m. \quad (5.2)$$

Part (b) of the theorem is a simple corollary of Part (a). A proof of the analogous result with R^n replaced by a bounded domain $\Omega \subset R^n$ may be found in Sobolev's book.[13] A simple proof of Part (a) may be found in a paper by Nirenberg.[18]

Theorem 4.1 and Sobolev's Imbedding Theorem imply that the wave packets $\psi_\pm(x, t)$ and $\psi_0(x, t)$ have continuous space derivatives of order k when $m \ge [n/2] + k + 1$. The existence of continuous

[18] L. Nirenberg, Ann. Scuola Norm. Sup. Pisa, Ser. III **13**, 30 (1959).

mixed (space–time) derivatives is implied by the following generalization of Sobolev's theorem.

Theorem 5.1. If $m \geq [n/2] + k + 1$ then

$$\bigcap_{l=0}^{k} C^l(-\infty, \infty, L_2^{m-l}(R^n)) \subset C^k(R^{n+1}). \tag{5.3}$$

Proof. If $\psi \in C^l(-\infty, \infty, L_2^{m-l}(R^n))$ where $0 \leq l \leq k$ then derivatives $D_t^j\psi(x, t) \in L_2^{m-l}(R^n)$ for $-\infty < t < \infty$ and $j = 0, 1, \cdots, l$. Hence, by Sobolev's imbedding theorem, $D_t^j\psi(x, t) \in C^{k-l}(R^n)$ and

$$\max_{\substack{x\in R^n \\ |\alpha|\leq k-l}} |D_x^\alpha D_t^j\psi(x, t)| \leq c\, ||D_t^j\psi(t)||_{m-l}$$

$$\text{for} \quad -\infty < t < \infty, \qquad j = 0, 1, \cdots, l. \tag{5.4}$$

This implies that $D_x^\alpha D_t^j\psi(x, t) \in C(R^{n+1})$ for $|\alpha| \leq k - l$, $0 \leq j \leq l$. Indeed, applying (5.4) to the difference $\psi(x, t) - \psi(x, \tau)$ gives

$$\max_{\substack{x\in R^n \\ |\alpha|\leq k-l}} |D_x^\alpha D_t^j\psi(x, t) - D_x^\alpha D_t^j(x, \tau)| \leq c\, ||D_t^j\psi(t) - D_t^j\psi(\tau)||_{m-l}$$

$j \leq l$ and $\psi \in C^l(-\infty, \infty, L_2^{m-l}(R^n))$. Moreover, a second application of (5.4) gives

$$\max_{\substack{x\in R^n \\ |\alpha|\leq k-l}} |\{D_x^\alpha D_t^{j-1}\psi(x, t + \tau) - D_x^\alpha D_t^{j-1}\psi(x, t)\}/\tau - D_x^\alpha D_t^j\psi(x, t)|$$

$$\leq c\, ||\{D_t^{j-1}\psi(t + \tau) - D_t^{j-1}\psi(t)\}/\tau - D_t^j\psi(t)||_{m-l}$$

and the last term tends to zero with τ if $0 \leq j \leq l$ and $\psi \in C^l(-\infty, \infty, L_2^{m-l}(R^n))$. Thus, in particular, $D_x^\alpha D_t^j\psi(x, t)$ is a partial derivative of $\psi(x, t)$ in the classical sense, and $\psi(x, t)$ has continuous derivatives $D_x^\alpha D_t^j\psi(x, t)$ for $|\alpha| \leq k - l$, $0 \leq j \leq l$. If

$$\psi \in \bigcap_{l=0}^{k} C^l(-\infty, \infty, L_2^{m-l}(R^n))$$

it follows that ψ has continuous derivatives $D_x^\alpha D_t^l\psi(x, t)$ for $|\alpha| + l \leq k$ and $l = 0, 1, \cdots, k$; i.e., $\psi \in C^k(R^{n+1})$.

The principal results of this section are described by

Theorem 5.2. If the hypotheses of Theorem 4.1 hold with $m \geq [n/2] + k + 1$ then $\psi_0(x, t)$ and $\psi_\pm(x, t)$ are in $C^k(R^{n+1})$ and there exists a constant c, independent of t, such that

$$\max_{x\in R^n} |D_x^\alpha D_t^l\psi_\pm(x, t) - D_x^\alpha D_t^l\psi_0(x, t)| \leq c[||D_t^l\psi_\pm(t) - D_t^l\psi_0(t)||_0]^{\frac{1}{2}} \tag{5.5}$$

for $-\infty < t < \infty$ and all α and l with $|\alpha| + l \leq k$.

Combining (5.5) and Theorem 2.1 gives

Corollary 5.1. If the hypotheses of Theorem 4.1 hold with $m \geq [n/2] + k + 1$ then $\psi_\pm(x, t) - \psi_0(x, t)$ and its space–time derivatives of order $\leq k$ tend to zero when $t \to \pm\infty$, uniformly in R^n.

Proof of Theorem 5.2. Since $L_2^j(R^n) \subset L_2^k(R^n)$ if $j \geq k$, conclusion (4.4) of Theorem 4.1 implies that both ψ_0 and $\psi_\pm$ are in $\bigcap_{l=0}^{k} C^l(-\infty, \infty, L_2^{m-l}(R^n))$. Thus $\psi_0(x, t)$ and $\psi_\pm(x, t)$ are in $C^k(R^{n+1})$, by Theorem 5.1. Moreover, inequality (5.4) holds for the difference $\psi_\pm(x, t) - \psi_0(x, t)$, whence

$$\max_{x\in R^n} |D_x^\alpha D_t^l\psi_\pm(x, t) - D_x^\alpha D_t^l\psi_0(x, t)| \leq c\, ||D_t^l\psi_\pm(t) - D_t^l\psi_0(t)||_{m-l}$$

for $-\infty < t < \infty$ and $|\alpha| + l \leq k$. Combining this result with conclusion (4.5) of Theorem 4.1 gives (5.5). This completes the proof of Theorem 5.2.

Excitation Operators and Intrinsic Hamiltonians

R. K. NESBET
IBM San Jose Research Laboratory, San Jose, California
(Received 18 September 1964)

An operator $A^\dagger$ that satisfies $[H, A^\dagger] = \hbar\omega A^\dagger$ converts a stationary-state eigenfunction of the Hamiltonian H into another eigenfunction with energy eigenvalue increased by $\hbar\omega$. Such operators describe collective excitations of many-particle systems, and their properties can be used to construct an intrinsic Hamiltonian that is dynamically independent of the collective degrees of freedom, without introducing subsidiary conditions. The procedure developed by Lipkin, valid when $\hbar\omega$ is real and positive, is extended to make possible the construction of an intrinsic Hamiltonian when $\hbar\omega$ vanishes and $A^\dagger$ is Hermitian, and also when $\hbar\omega$ is complex. The nuclear cranking model is shown to be a special case of the proposed general method for vanishing $\hbar\omega$ in which effective moments of inertia occur as eigenvalues of linear equations. Several examples are worked out in detail, all dealing with an interacting phonon–electron system in the random-phase approximation. Results derived are the explicit screened Coulomb interaction resulting from electronic plasma excitations, a verification of the renormalized phonon frequency spectrum and phonon–electron interaction derived in the adiabatic approximation, and the resulting screened Coulomb and phonon-induced electronic interactions obtained when plasma and phonon excitations are treated simultaneously.

I. INTRODUCTION

THIS paper is concerned with "excitation operators," defined as operators that convert one stationary state of a quantum mechanical system into another state. The basic equation satisfied by an excitation operator $A^\dagger$ is

$$[H, A^\dagger] = \hbar\omega A^\dagger, \tag{1}$$

where $\hbar\omega$ is an elementary excitation energy. The technique of approximating such operators by linear combinations of simpler operators, then truncating the resulting linear equations to obtain a closed system, has been applied in recent years to a number of physical problems involving many-particle systems. From the structure of Eq. (1), this technique is most commonly referred to as the "equation of motion" method. The same method, when expressed in terms of matrix elements such as

$$(A^\dagger\Psi, a_i^\dagger a_j\Psi), \tag{2}$$

where $a_i^\dagger$, a_j are elementary fermion creation and annihilation operators, is a generalization of the Tamm–Dancoff method, as proposed by Dyson.[1] The matrix elements in Eq. (2) are referred to as Dyson–Tamm–Dancoff amplitudes.

Some of the more important applications of this method will be found in papers on the free-electron gas,[2,3] on the pairing interaction in superconductors,[4] on collective excitations of nuclear matter and of finite nuclei,[5] and on excited states of atoms.[6] A general discussion of the approximations that can be made in truncating these equations has been given by Suhl and Werthamer.[7]

The particular aspect of this theory that will be considered in the present paper is the problem of constructing an intrinsic Hamiltonian, defined so that it describes those degrees of freedom of a system that are dynamically independent of excitations defined by Eq. (1). A discussion of this subject has been given by Lipkin.[8] The need for such a Hamiltonian, expressed in terms of the coordinates of the original set of independent particles, is clear in the empirical context of nuclear physics, where collective motions of the nucleus as a whole are combined with the shell structure appropriate to an independent particle model.[9]

In the simplest case considered here, when $\hbar\omega$ is real and positive, the technique proposed by Lipkin[8] gives an intrinsic Hamiltonian that commutes with both $A^\dagger$ and A, establishing the required dynamical independence. A method due to Bohr and Mottelson[5]

[1] F. J. Dyson, Phys. Rev. **91**, 1543 (1953).
[2] K. Sawada, Phys. Rev. **106**, 372 (1957); K. Sawada, K. A. Brueckner, N. Fukuda, and R. Brout, *ibid.* **108**, 507 (1957); R. Brout, *ibid.* **108**, 515 (1957).
[3] A general review of this subject has been given recently by D. Pines, *Elementary Excitations in Solids* (W. A. Benjamin, Inc., New York, 1963).
[4] P. W. Anderson, Phys. Rev. **112**, 1900 (1958); G. Rickayzen, Phys. Rev. **115**, 795 (1959).
[5] A. Bohr and B. R. Mottelson, K. Norske Vidensk. Selsk. Forhandl. **31**, 71 (1958); A. E. Glassgold, W. Heckrotte, and K. M. Watson, Ann. Phys. **6**, 1 (1959); S. Takagi, Progr. Theoret. Phys. (Kyoto) **21**, 174 (1959); G. E. Brown, J. A. Evans, and D. J. Thouless, Nucl. Phys. **24**, 1 (1961); S. Fallieros, Ph.D. thesis, University of Maryland, 1959 (unpublished).
[6] P. L. Altick and A. E. Glassgold, Phys. Rev. **133**, A632 (1964).
[7] H. Suhl and N. R. Werthamer, Phys. Rev. **122**, 359 (1961).
[8] H. J. Lipkin, Phys. Rev. Letters **2**, 159 (1959); Ann. Phys. (N. Y.) **9**, 272 (1960).
[9] A. Bohr and B. R. Mottelson, Kgl. Danske Videnskab. Selskab, Mat-Fys. Medd. **30**, 1 (1955); C. A. Levinson, Phys. Rev. **132**, 2184 (1963).

can be used when $\hbar\omega$ vanishes. The nuclear cranking model[10] is shown here to be a special case of this method, which can be used to compute inertial parameters (moments of inertia and effective masses) of many-particle systems. Such parameters occur as eigenvalues of a system of linear equations. In general the linear equations implied by Eq. (1) are non-Hermitian, and the eigenvalue $\hbar\omega$ can be complex. A special procedure for dealing with this case is proposed here. The imaginary part of a complex excitation energy computed by this method corresponds to the width or lifetime of the corresponding nonstationary excitation.

Several examples are given, all using the random phase approximation (RPA),[11] and referring to an interacting electron–phonon system. The first example, plasma excitations of a purely electronic system, has been treated previously,[2] but the resulting intrinsic Hamiltonian, which contains a Coulomb interaction screened by the removal of plasma excitations, has not appeared explicitly in the literature. In the second example, the present method for zero excitation energy is used to derive the principal results of the adiabatic approximation for the phonon–electron system.[12] The third example considers the problem treated by Bardeen and Pines,[13] plasma and phonon excitations of the phonon–electron system, taking all interactions into account. The phonon excitations in this case (for a normal metal) have a natural lifetime, and provide an example of the method proposed here for complex excitation energies. A dispersion relation is derived that gives both plasma and phonon excitation energies, as well as the natural lifetime of the latter. The resulting intrinsic Hamiltonian contains a screened Coulomb interaction and a phonon-induced electronic interaction essentially the same as that considered in superconductivity theory. The present derivation does not use perturbation theory, and the validity of these results is limited only by the RPA. In all cases the intrinsic Hamiltonian is expressed in the original particle coordinates, eliminating any need for subsidiary conditions.[8]

When the operator equations considered here are truncated by use of the RPA, they become equivalent to the equations of the time dependent Hartree–Fock approximation, which can also be derived by a Green's function method. In the concluding section of this paper some of the formal properties of these truncated equations, discussed and derived by Thouless,[14] will be derived from the algebraic properties of excitation operators. This analysis demonstrates that the operator orthonormalization condition introduced here is consistent with the assumed equations for excitation operators, and generalizes the variational principle for $\hbar\omega$ proposed by Thouless.

Only the case of excitation operators that satisfy boson commutation rules will be considered here. Similar results can be derived for fermion operators,[7] but no intrinsic Hamiltonian of physical interest has been obtained by the use of such operators.

II. CONSTRUCTION OF AN INTRINSIC HAMILTONIAN

Suppose that Hermitian conjugate operators A, $A^\dagger$ are known such that for a given Hamiltonian H,

$$[H, A^\dagger] = \hbar\omega A^\dagger, \tag{3a}$$

$$[H, A] = -\hbar\omega A, \tag{3b}$$

$$[A, A^\dagger] = 1. \tag{3c}$$

Then an intrinsic Hamiltonian can be defined by[8]

$$H_0 = H - \hbar\omega A^\dagger A, \tag{4}$$

such that

$$[H_0, A^\dagger] = [H_0, A] = 0. \tag{5}$$

It is assumed here that ω is real (the complex case is discussed below). Equation (5) implies that H_0 is dynamically independent of $A^\dagger$ and A. These operators will be referred to as excitation and de-excitation operators, respectively.

Define

$$Q = (\hbar/2\omega)^{\frac{1}{2}}(A^\dagger + A),$$

$$P = i(\hbar\omega/2)^{\frac{1}{2}}(A^\dagger - A). \tag{6}$$

These operators are a collective coordinate and the corresponding conjugate momentum. By Eq. (5), H_0 commutes with both operators, and describes an intrinsic system not dynamically coupled to the collective motion described by P and Q. The effective Hamiltonian for collective motion is

$$\hbar\omega A^\dagger A = \tfrac{1}{2}(P^2 + \omega^2 Q^2 - \hbar\omega). \tag{7}$$

If there exists a function Ψ, not identically zero, that satisfies

$$H\Psi = E\Psi, \tag{8}$$

[10] D. R. Inglis, Phys. Rev. **96**, 1059 (1954); **97**, 701 (1955); D. J. Thouless, Nucl. Phys. **21**, 225 (1960); F. Villars, in Rendiconti della Scuola Internazionale di Fisica "Enrico Fermi," XXIII Corso: Fisica Nucleara (Academic Press, Inc., New York, 1963), pp. 1–47.

[11] D. Bohm and D. Pines, Phys. Rev. **92**, 626 (1953).

[12] J. M. Ziman, Proc. Cambridge Phil. Soc. **51**, 707 (1955); G. V. Chester, Adv. Phys. **10**, 357 (1961).

[13] J. Bardeen and D. Pines, Phys. Rev. **99**, 1140 (1955).

[14] D. J. Thouless, Nucl. Phys. **22**, 78 (1961).

then $A^\dagger\Psi$ and $A\Psi$ are also eigenfunctions of H, with eigenvalues $E + \hbar\omega$, $E - \hbar\omega$, respectively. Assuming that the energy spectrum of H is bounded below, when $\hbar\omega$ is positive, successive multiplications by A must eventually lead to a function Ψ_0, not identically zero, such that

$$A\Psi_0 \equiv 0. \tag{9}$$

Thus Ψ_0 describes the ground state of the collective motion, an eigenfunction of Eq. (7) with eigenvalue zero. Corresponding to each eigenfunction Ψ of H there is a ladder of eigenfunctions and equally spaced eigenvalues, terminating below when Eq. (9) holds, and terminating above if, for some nontrivial Ψ_f,

$$A^\dagger\Psi_f \equiv 0. \tag{10}$$

Equations (3) will not hold exactly for real systems unless such energy ladders occur, except for the trivial case

$$A\Psi_i \equiv 0, \qquad A^\dagger\Psi_i = \Psi_j, \qquad A^\dagger\Psi_j \equiv 0, \tag{11}$$

for any two eigenfunctions Ψ_i, Ψ_j of H.

As Eq. (11) shows, $\hbar\omega = E_j - E_i$ can be either positive or negative, since the excitation operator could equally well be defined by $A^\dagger\Psi_j = \Psi_i$. The normalization condition chosen here, Eq. (3c), is compatible with only one of these alternative definitions. By this convention, an excitation operator that has negative excitation energy corresponds to a negative definite collective Hamiltonian, with ground state characterized by Eq. (10).

Approximate solutions of Eqs. (3) are obtained by expressing $A^\dagger$ as a linear combination of known operators.[2,7] The commutators in these equations are simplified so that $[A, A^\dagger]$ is approximated by a pure number and $[H, A^\dagger]$ is approximated by a linear combination of operators from the given set. The coefficients in the assumed linear combination of operators then satisfy linear eigenvalue equations. The excitation energy $\hbar\omega$ is an eigenvalue of these equations, and Eq. (3c) can be used to normalize the eigenvector. This procedure breaks down if A and $A^\dagger$ commute, unless $\hbar\omega = 0$.

When $\hbar\omega$ vanishes, H commutes with any linear combination of A and $A^\dagger$. If $[A, A^\dagger]$ does not vanish, this simply means that H is dynamically independent of A and $A^\dagger$. A nontrivial special case occurs, however, when A is Hermitian. Then A and $A^\dagger$ are identical and their commutator vanishes. Such an operator can be denoted by P, where

$$[H, P] = 0. \tag{12a}$$

To construct an intrinsic Hamiltonian in this case, it suffices to find an operator Q such that

$$[H, Q] = (\hbar\Lambda/i)P, \tag{12b}$$

where both P and Q are Hermitian, Λ is a real number, and

$$[P, Q] = \hbar/i. \tag{12c}$$

Then the intrinsic Hamiltonian is defined by

$$H_0 = H - \tfrac{1}{2}\Lambda P^2, \tag{13}$$

such that

$$[H_0, P] = [H_0, Q] = 0. \tag{14}$$

The collective Hamiltonian is $\frac{1}{2}\Lambda P^2$.

Approximate solutions of Eqs. (12) can be obtained by expressing the operators P and Q as linear combinations of known operators, solving the resulting linear equations for the coefficients and for the eigenvalue Λ.

III. ZERO EXCITATION ENERGIES AND THE CRANKING MODEL

Given Eq. (12a), the cranking model[10] introduces a Lagrange multiplier λ, to define the modified Hamiltonian

$$H_\lambda = H - \lambda P, \tag{15}$$

with eigenvalues E_λ that correspond to eigenvalues P_λ of the operator P. The dependence of E_λ on P_λ is deduced by eliminating the independent variable λ, and this relationship can be used to construct an intrinsic Hamiltonian H_0, independent of P. If this procedure is carried out by means of a canonical transformation

$$H'_\lambda = \exp(-iS_\lambda/\hbar)H_\lambda \exp(iS_\lambda/\hbar), \tag{16}$$

it can be shown that the generating operator S_λ differs from the operator Q defined by Eq. (12b) only by a numerical factor. Hence the two methods are identical.

To see this in detail, treat λ as an expansion parameter in perturbation theory. S_λ is at least of first order in λ, and successive terms in Eq. (16) are

$$\begin{aligned} H'_\lambda = {} & H \\ & + (i/\hbar)[H, S_\lambda] - \lambda P \\ & + \tfrac{1}{2}(i/\hbar)^2[[H, S_\lambda], S_\lambda] - (i/\hbar)\lambda[P, S_\lambda] \\ & + \cdots . \end{aligned} \tag{17}$$

Choose S_λ so that the second line of Eq. (17) vanishes,

$$[H, S_\lambda] = (\hbar\lambda/i)P. \tag{18}$$

Then if the commutator $[P, S_\lambda]$ is a c number, of first order in λ, it follows from Eq. (18) that all higher-order commutators not indicated in Eq. (17) vanish. Thus if

$$[P, S_\lambda] = (\hbar/i)(\lambda/\Lambda), \tag{19}$$

where Λ is a constant, then without any approximation,

$$H'_\lambda = H - (i/2\hbar)\lambda[P, S_\lambda] = H - (\lambda^2/2\Lambda). \tag{20}$$

The transformed operator P' is

$$P' = P + (i/\hbar)[P, S_\lambda] = P + (\lambda/\Lambda). \tag{21}$$

The transformed operator H' is

$$H' = H'_\lambda + \lambda P'. \tag{22}$$

Eliminating λ between Eqs. (20) and (21) it follows that

$$H' - H = \tfrac{1}{2}\Lambda[(P')^2 - P^2], \tag{23}$$

or

$$H' - \tfrac{1}{2}\Lambda(P')^2 = H - \tfrac{1}{2}\Lambda P^2. \tag{24}$$

The operator defined by Eq. (24) is obviously independent of λ and hence of $P' - P$. It defines the intrinsic Hamiltonian of the cranking model.

If Eqs. (18), (19) are compared with Eqs. (12), above, it can be seen that they become identical if S_λ is set equal to $(\lambda/\Lambda)Q$. This comparison fails unless the eigenvalue of P varies linearly with λ.

IV. COMPLEX EXCITATION ENERGIES

Let $\Psi(t)$ be the time-dependent Schrödinger wavefunction for a stationary state of the Hamiltonian H. Then

$$\Psi(t) = \exp(Et/i\hbar)\Psi, \tag{25}$$

where Ψ is independent of t. Equation (3a) implies that

$$\Psi_a(t) = \exp[(E + \hbar\omega)t/i\hbar]A^\dagger\Psi \tag{26}$$

is the time-dependent wavefunction describing another state. If ω is complex this state is not stationary. If

$$[H, A^\dagger] = \hbar\omega A^\dagger, \tag{27}$$

and

$$\omega = \omega_0 - i\gamma, \qquad \gamma \geq 0, \tag{28}$$

this state decays spontaneously. Such a state would be realizable physically as the initial state of a closed system, if it were embedded in a continuum of other states to which transitions could occur in an essentially irreversible manner. The Hermitian conjugate operator A satisfies

$$[H, A] = -\hbar\omega^* A. \tag{29}$$

Since the imaginary part of $-\omega^*$ is negative, the state $A\Psi$ also decays spontaneously.

Because ω^* differs from ω, it is not possible to use Eq. (4) to define an intrinsic Hamiltonian that commutes with both $A^\dagger$ and A. It is possible, however, to generalize this procedure to the case of complex ω by using the time-inversion operator θ defined by Wigner.[15] Acting on an arbitrary wavefunction or operator, θ reverses the sign of all linear and angular momenta. The operator θ is *antilinear*, implying that for arbitrary wavefunctions Φ_0, Φ_1 and numerical coefficients α, β,

$$\theta(\alpha\Phi_0 + \beta\Phi_1) = \alpha^*\theta\Phi_0 + \beta^*\theta\Phi_1. \tag{30}$$

The time-inversion operator commutes with the Hamiltonian of a closed system. The operators $A^\dagger$ or A considered here act directly on wavefunctions, so the antilinear property must also hold for linear combinations of such operators. If Eq. (29) is transformed by the operation denoted by θ, it becomes

$$[H, \theta A\theta^{-1}] = -\hbar\omega\theta A\theta^{-1}. \tag{31}$$

From Eqs. (27) and (31) it can be seen that, for complex ω, $A^\dagger$ and $\theta A\theta^{-1}$ have an adjoint relationship which generalizes that between $A^\dagger$ and A. In particular, the operators can be normalized by requiring

$$[\theta A\theta^{-1}, A^\dagger] = 1, \tag{32}$$

and in general for independent excitations

$$[\theta A\theta^{-1}, B^\dagger] = \delta_{AB}. \tag{33}$$

The intrinsic Hamiltonian, which commutes with both $A^\dagger$ and $\theta A\theta^{-1}$, is

$$H_0 = H - \hbar\omega A^\dagger\theta A\theta^{-1}. \tag{34}$$

This operator is not Hermitian, as might be expected since the excitation described by $A^\dagger$ leads to a nonstationary state.

The Hermitian conjugate of Eq. (31) is

$$[H, (\theta A\theta^{-1})^\dagger] = \hbar\omega^*(\theta A\theta^{-1})^\dagger. \tag{35}$$

Comparison with Eq. (27) shows that $\theta A\theta^{-1}$ is just the Hermitian conjugate of a solution of Eq. (27) obtained by reversing the sign of the imaginary part of ω. This operator can be denoted by $A(\omega^*)$, and

[15] E. P. Wigner, *Group Theory* (Academic Press Inc., New York, 1959), pp. 325–348.

Eq. (34) becomes

$$H_0 = H - \hbar\omega A^\dagger(\omega)A(\omega^*), \tag{34'}$$

where

$$[A(\omega^*), A^\dagger(\omega)] = 1. \tag{32'}$$

Equation (35) shows that both $A(\omega^*)$ and its Hermitian conjugate refer to states that could not be realized as initial states of a closed system, since such states have amplitudes that increase spontaneously as functions of time. The normalization indicated by Eq. (32′) ensures that $A(\omega^*) \rightarrow A(\omega)$ when ω is real.

The special case of pure imaginary ω has been discussed by Sawada and Fukuda,[16] who show that this situation is characteristic of physical systems whose true ground state is qualitatively different from an assumed independent particle state, as in the BCS theory of superconductivity. Imaginary ω implies that overly restrictive approximations have been made in simplifying the expansion of an excitation operator as a linear combination of known operators. The equations for the coefficients in such an expansion must be modified by carrying out a canonical transformation of the assumed basis operators, in analogy to the Bogoliubov–Valatin transformation in superconductivity theory.

V. COLLECTIVE EXCITATIONS IN METALS

The electron–phonon Hamiltonian is assumed to be

$$H = H_e + H_n + H_{ne}, \tag{36}$$

where

$$H_e = \Sigma_{\sigma s}\epsilon_\sigma a^\dagger_{\sigma s}a_{\sigma s} + \tfrac{1}{2}\Sigma_\kappa M^2_\kappa \rho^\dagger_\kappa \rho_\kappa, \tag{37}$$

$$H_n = \tfrac{1}{2}\Sigma_\kappa (p^\dagger_\kappa p_\kappa + \Omega^2_\kappa q^\dagger_\kappa q_\kappa), \tag{38}$$

$$H_{ne} = \Sigma_\kappa v^i_\kappa \rho^\dagger_\kappa q_\kappa. \tag{39}$$

The electronic density operator is

$$\rho_\kappa = \Sigma_{\sigma s} a^\dagger_{\sigma s}a_{\sigma+\kappa,s} = \rho^\dagger_{-\kappa}, \tag{40}$$

where $a^\dagger_{\sigma s}$, $a_{\sigma s}$ are creation and annihilation operators, respectively, for electrons in Bloch waves of momentum $\hbar\sigma$ and spin m_s. Only longitudinal phonons are considered. The summation over phonon momentum κ in Eq. (38) is restricted to the first Brillouin zone, while the κ summations in the other equations are unrestricted. The traveling wave representation is used, implying

$$p^\dagger_\kappa = p_{-\kappa}; \qquad q^\dagger_\kappa = q_{-\kappa}. \tag{41}$$

[16] K. Sawada and N. Fukuda, Progr. Theoret. Phys. (Kyoto) **25**, 653 (1961).

The Hamiltonian of Eq. (36) is that considered by Bardeen and Pines,[13] who discuss the interpretation and derivation of each term. This Hamiltonian describes the unscreened interaction between Bloch electrons and ionic cores. If the latter are assumed to be point charges of charge Ze and mass M, of which there are N in total volume V, the constants in Eqs. (38) and (39) are approximated by

$$\Omega^2_\kappa = 4\pi Z^2 e^2 N/MV \tag{42}$$

and

$$v^i_\kappa = (v^i_{-\kappa})^* = -i\kappa(N/MV)^{\frac{1}{2}}(4\pi Ze^2/\kappa^2). \tag{43}$$

Thus Ω_κ is the classical plasma frequency for the ion cores, independent of κ, while v^i_κ, the Fourier coefficient of the gradient of an unscreened Coulombic interaction, is singular as $\kappa \rightarrow 0$. The electronic interaction is assumed to be the unscreened Coulomb interaction,

$$M^2_\kappa = 4\pi e^2/\kappa^2. \tag{44}$$

In Eq. (39), when κ is outside the first Brillouin zone, the operator q_κ refers to the phonon mode $\kappa - \mathbf{K}$ in the first zone, where $\mathbf{K}$ is a reciprocal lattice vector. The coefficients ϵ_σ in Eq. (37) are assumed to be independent of spin and to satisfy

$$\epsilon_\sigma = \epsilon_{-\sigma}. \tag{45}$$

They are the one-electron energies of Bloch waves for electrons moving in the periodic potential of the ionic cores in their equilibrium configuration. A divergent constant term, representing the potential due to a uniform distribution of charge in the assumed volume V, is canceled between H_e and H_n and is omitted from Eqs. (37) and (39).

The random-phase approximation (RPA) of Bohm and Pines[11,13] will be used to simplify the commutators given by Eqs. (3) for excitation operators. In general, this means that only terms connecting operators that represent the same momentum transfer κ are retained, that the electron number operators

$$f_{\sigma s} = a^\dagger_{\sigma s}a_{\sigma s} \tag{46}$$

are replaced by c numbers, their mean or statistical values in some specified state, that exchange terms in the electron–electron Coulomb interaction are ignored, and that in commutators like $[H_e, a^\dagger_i a_j]$ only those terms are retained that are of the form (const) $a^\dagger_k a_l$, where (const) is a function of the number operators. These simplifications lead to the following approximate expressions for the commutators:

$$[H_e, a^\dagger_{\sigma+\kappa,s}a_{\sigma s}] = (\epsilon_{\sigma+\kappa} - \epsilon_\sigma)a^\dagger_{\sigma+\kappa,s}a_{\sigma s} + M^2_\kappa(f_{\sigma s} - f_{\sigma+\kappa,s})\rho^\dagger_\kappa, \tag{47a}$$

$$[H_{ne}, a^{\dagger}_{\sigma+\kappa,s}a_{\sigma s}] = v^{i}_{-\kappa}(f_{\sigma s} - f_{\sigma+\kappa,s})q^{\dagger}_{\kappa}. \quad (47b)$$

Other commutators needed here, evaluated without approximation, are

$$[H_{ne}, p_{\kappa}] = \Sigma_{\mu} i\hbar v^{i}_{\kappa\mu}\rho^{\dagger}_{\kappa\mu}, \quad (48a)$$

$$[H_{n}, p_{\kappa}] = i\hbar\Omega^{2}_{\kappa}q^{\dagger}_{\kappa}, \quad (48b)$$

$$[H_{n}, q^{\dagger}_{\kappa}] = -i\hbar p_{\kappa}. \quad (48c)$$

In Eq. (48a), κ is restricted to the first Brillouin zone. The summation is over all momenta κ_{μ} such that

$$\kappa_{\mu} = \kappa + \mathbf{K}_{\mu}, \quad (49)$$

where $\mathbf{K}_{\mu}$ is a reciprocal lattice vector.

A. Electronic Plasma Excitations

In the RPA, Eqs. (3) for the electronic Hamiltonian H_e have solutions of the form

$$A^{\dagger}_{\kappa} = \Sigma_{\sigma s}\alpha(\kappa\sigma s)a^{\dagger}_{\sigma+\kappa,s}a_{\sigma s}. \quad (50)$$

The coefficients α satisfy the linear equations

$$(\hbar\omega - \epsilon_{\sigma+\kappa} + \epsilon_{\sigma})\alpha(\kappa\sigma s) = M^{2}_{k}\Sigma_{\sigma',s'}(f_{\sigma',s'} - f_{\sigma',+\kappa,s'})\alpha(\kappa\sigma' s') \quad (51)$$

subject to the normalization condition (assuming that ω is real)

$$\Sigma_{\sigma s}|\alpha(\kappa\sigma s)|^{2}(f_{\sigma s} - f_{\sigma+\kappa s}) = 1. \quad (52)$$

Equation (51) implies that

$$C_{\kappa}[1 - M^{2}_{\kappa}\Sigma_{\sigma s}(f_{\sigma s} - f_{\sigma+\kappa,s})/(\hbar\omega - \epsilon_{\sigma+\kappa} + \epsilon_{\sigma})] = 0, \quad (53)$$

where

$$C_{\kappa} = \Sigma_{\sigma s}(f_{\sigma s} - f_{\sigma+\kappa,s})\alpha(\kappa\sigma s). \quad (54)$$

Hence either $C_{\kappa} = 0$, which gives the continuum solutions of Eq. (51), or else the square bracket in Eq. (53) vanishes, giving the dispersion relation of Bohm and Pines[3] for the plasma excitation frequency ω_{κ},

$$1 = M^{2}_{\kappa}\Sigma_{\sigma s}(f_{\sigma s} - f_{\sigma+\kappa,s})/(\hbar\omega_{\kappa} - \epsilon_{\sigma+\kappa} + \epsilon_{\sigma}). \quad (55)$$

A stationary plasma solution exists only when the integrand in Eq. (55) is nonsingular, requiring $\hbar\omega_{\kappa}$ to exceed the maximum value of $\epsilon_{\sigma+\kappa} - \epsilon_{\sigma}$ for which $f_{\sigma s} - f_{\sigma+\kappa,s}$ does not vanish. If statistical values are given to the electron number operators f for finite temperatures, when the indicated sum is replaced by an integral a singularity occurs within the region of integration because $\epsilon_{\sigma+\kappa} - \epsilon_{\sigma}$ has no maximum value. Hence the integral in Eq. (55) is complex and Eq. (55) will in general have a complex root ω_{κ}, implying a nonstationary plasma excitation.

When the imaginary part of $\hbar\omega$ can be neglected, the normalization constant C_{κ} can be evaluated by differentiating the dispersion relation.[17] Equation (55) can be written as

$$1 = e^{2}F(\kappa, \hbar\omega_{\kappa}). \quad (56)$$

Differentiating with respect to e^2, this gives

$$e^{2}\frac{\partial F}{\partial(\hbar\omega_{\kappa})} = -\left[e^{2}\frac{d(\hbar\omega_{\kappa})}{de^{2}}\right]^{-1}. \quad (57)$$

But Eq. (52) is equivalent to

$$-M^{2}_{\kappa}C^{2}_{\kappa}e^{2}\,\partial F/\partial(\hbar\omega_{\kappa}) = 1, \quad (58)$$

which implies

$$C^{2}_{\kappa} = e^{2}[d(\hbar\omega_{\kappa})/de^{2}]M^{-2}_{\kappa}, \quad (59)$$

and

$$\alpha(\kappa\sigma s) = \{e^{2}[d(\hbar\omega_{\kappa})/de^{2}]M^{2}_{\kappa}\}^{\frac{1}{2}} \times [(\hbar\omega - \epsilon_{\sigma+\kappa} + \epsilon_{\sigma})]^{-1}. \quad (60)$$

If these results are substituted into Eq. (4), the residual electronic interaction in the intrinsic Hamiltonian is

$$\tfrac{1}{2}\Sigma_{\kappa}\Sigma_{\sigma s}\Sigma_{\sigma',s'}M^{2}_{\kappa} \times \left[1 - \frac{e^{2}d(\hbar\omega_{\kappa})^{2}/de^{2}}{(\hbar\omega_{\kappa} - \epsilon_{\sigma+\kappa} + \epsilon_{\sigma})(\hbar\omega_{\kappa} - \epsilon_{\sigma',+\kappa} + \epsilon_{\sigma'})}\right] \times a^{\dagger}_{\sigma+\kappa,s}a_{\sigma s}a^{\dagger}_{\sigma',s'}a_{\sigma',+\kappa,s'}. \quad (61)$$

Since ω_{κ} is given approximately by[3]

$$\omega^{2}_{\kappa} = \omega^{2}_{p} + \tfrac{3}{5}\kappa^{2}v^{2}_{0} + \cdots, \quad (62)$$

the coefficient in square brackets in Eq. (61) is approximately

$$1 - (\omega^{2}_{p}/\omega^{2}_{\kappa}) = \kappa^{2}/(\kappa^{2} + \lambda^{2}_{p}), \quad (63)$$

where

$$\lambda^{2}_{p} = 5\omega^{2}_{p}/3v^{2}_{0}. \quad (64)$$

The classical electronic plasma frequency ω_p is given by

$$\omega^{2}_{p} = 4\pi e^{2}NZ/mV, \quad (65)$$

since electronic density is Z times ion density in a neutral metal. The constant v_0 is electron velocity at the Fermi surface. Equations (61) and (63) show that the intrinsic Hamiltonian describes a screened Coulomb interaction, with M^{2}_{κ} replaced by

$$M^{2}_{\kappa}\kappa^{2}/(\kappa^{2} + \lambda^{2}_{p}) = 4\pi e^{2}/(\kappa^{2} + \lambda^{2}_{p}). \quad (66)$$

[17] K. Sawada, K. A. Brueckner, N. Fukuda, and R. Brout, Ref. 2.

B. Phonon–Electron Interaction in the Adiabatic Approximation

In the adiabatic approximation, electronic wavefunctions that depend parametrically on the lattice ion displacement normal coordinates q_κ are obtained as eigenfunctions of $H_e + H_{ne}$. The electronic energy eigenvalues are then functions of q_κ and are added into H_n to provide a Hamiltonian that describes lattice vibrations.

Since $H_1 = H_e + H_{ne}$ obviously commutes with q_κ, this provides an example of the use of Eqs. (12). It is convenient here to relabel the operators, and to solve the system of equations for both $\pm\kappa$,

$$[H_1, Q_\kappa^\dagger] = 0, \tag{67a}$$

$$[H_1, P_\kappa] = i\hbar\Lambda_\kappa Q_\kappa^\dagger, \tag{67b}$$

$$[P_\kappa, Q_\kappa] = \hbar/i, \tag{67c}$$

which lead to an intrinsic Hamiltonian

$$H_0 = H_1 - \tfrac{1}{2}\Sigma_\kappa \Lambda_\kappa Q_\kappa^\dagger Q_\kappa, \tag{68}$$

where $Q_\kappa^\dagger = Q_{-\kappa}$, $P_\kappa^\dagger = P_{-\kappa}$, and $\Lambda_\kappa = \Lambda_{-\kappa}$.

Let

$$Q_\kappa = q_\kappa, \tag{69}$$

and choose the coefficients in

$$P_\kappa = \Sigma_{\sigma s}\alpha(\kappa\dot{\sigma}\dot{s})a_{\sigma+\kappa,s}^\dagger a_{\sigma s} + \beta(\kappa)p_\kappa \tag{70}$$

so that Eqs. (67) are satisfied. Then

$$\beta(\kappa) = 1, \tag{71}$$

$$\Sigma_{\sigma s} v_{-\kappa}^i (f_{\sigma s} - f_{\sigma+\kappa,s})\alpha(\kappa\dot{\sigma}\dot{s}) = i\hbar\Lambda_\kappa, \tag{72}$$

and

$$\alpha(\kappa\dot{\sigma}\dot{s}) = (\hbar/i)v_\kappa[(\epsilon_{\sigma+\kappa} - \epsilon_\sigma)]^{-1}, \tag{73}$$

where

$$v_\kappa = v_\kappa^i + (M_\kappa^2/i\hbar)\Sigma_{\sigma s}(f_{\sigma s} - f_{\sigma+\kappa,s})\alpha(\kappa\dot{\sigma}\dot{s}). \tag{74}$$

A screening constant can be defined by

$$\lambda_\kappa^2 = \kappa^2 M_\kappa^2 \Sigma_{\sigma s}(f_{\sigma s} - f_{\sigma+\kappa,s})/(\epsilon_{\sigma+\kappa} - \epsilon_\sigma), \tag{75}$$

obviously positive and nonsingular as $\kappa \to 0$. In terms of λ_κ^2, Eq. (74) becomes

$$\begin{aligned} v_\kappa &= v_\kappa^i - (\lambda_\kappa^2/\kappa^2)v_\kappa \\ &= [\kappa^2/(\kappa^2 + \lambda_\kappa^2)]v_\kappa^i, \end{aligned} \tag{76}$$

and Eq. (72) becomes

$$\Lambda_\kappa = M_\kappa^{-2} v_{-\kappa}^i (v_\kappa - v_\kappa^i). \tag{77}$$

When $H_1 - H_0$ of Eq. (68) is combined with H_n, the effective phonon Hamiltonian is

$$H_n' = \tfrac{1}{2}\Sigma_\kappa (p_\kappa^\dagger p_\kappa + \omega_\kappa^2 q_\kappa^\dagger q_\kappa), \tag{78}$$

where

$$\omega_\kappa^2 = \Omega_\kappa^2 + \Lambda_\kappa. \tag{79}$$

Equations (42), (43), (76), (77), and (79) can be combined to give

$$\omega_\kappa^2 = \kappa^2\Omega_\kappa^2/(\kappa^2 + \lambda_\kappa^2). \tag{80}$$

When κ is small this is of the form $\omega_\kappa^2 = \kappa^2 s^2$, defining the longitudinal sound velocity s.

Since the intrinsic Hamiltonian H_0, defined by Eq. (68), commutes with both P_κ and $Q_\kappa = q_\kappa$, its energy spectrum is not affected by setting q_κ equal to zero, thus removing both H_{ne} of Eq. (39) and the explicit collective term in Eq. (68). The residual phonon–electron coupling arises solely from the presence of the operators p_κ rather than P_κ in H_n' of Eq. (78), since the former do not commute with H_0. From Eqs. (70), (71), and (73),

$$p_\kappa = P_\kappa + i\hbar v_\kappa \Sigma_{\sigma s}(\epsilon_{\sigma+\kappa} - \epsilon_\sigma)^{-1} a_{\sigma+\kappa,s}^\dagger a_{\sigma s}. \tag{81}$$

When this is substituted into Eq. (78), the three resulting terms are the modified phonon Hamiltonian

$$H_n'' = \tfrac{1}{2}\Sigma_\kappa (P_\kappa^\dagger P_\kappa + \omega_\kappa^2 Q_\kappa^\dagger Q_\kappa), \tag{82}$$

an explicit phonon–electron interaction

$$H_{ne}' = \tfrac{1}{2}\Sigma_\kappa \Sigma_{\sigma s}\{[i\hbar v_\kappa/(\epsilon_{\sigma+\kappa} - \epsilon_\kappa)]P_\kappa^\dagger a_{\sigma+\kappa,s}^\dagger a_{\sigma s} + \text{c.c.}\}, \tag{83}$$

and a correction to the original electron–electron interaction,

$$\Delta H_e = \tfrac{1}{2}\Sigma_\kappa \Sigma_{\sigma s}\Sigma_{\sigma',s'} \frac{\hbar^2 |v_\kappa|^2}{(\epsilon_{\sigma+\kappa} - \epsilon_\sigma)(\epsilon_{\sigma',+\kappa} - \epsilon_{\sigma'})} \times a_{\sigma+\kappa,s}^\dagger a_{\sigma s} a_{\sigma',s}^\dagger a_{\sigma',+\kappa,s'}. \tag{84}$$

These results are identical with those obtained by Chester,[12] who used a canonical transformation method to derive corrections to the usual Born–Oppenheimer approximation. Equation (84) does not represent the true phonon-induced electronic interaction, since it will be modified by a canonical transformation that removes the interaction term H_{ne}' of Eq. (83).[18] A more careful analysis of Eqs. (73) and (74) shows that the energy denominator $\epsilon_{\sigma+\kappa} - \epsilon_\sigma$ should contain a small imaginary part, which removes the apparent singularities in these equations.

The summation over reciprocal lattice vectors $\mathbf{K}_\mu$ in Eq. (48a) has been ignored in deriving the equations used here. If all terms were kept, these equations would couple all values of κ connected by reciprocal lattice vectors. This approximation is not

[18] T. R. Koehler and R. K. Nesbet, Phys. Rev. **135**, A638 (1964).

the same as neglecting all umklapp processes (κ outside the first Brillouin zone), since the value of κ defined by the electronic momentum transfer may or may not lie in the first zone. The phonon index κ here must be interpreted as the particular value of $\kappa - \mathbf{K}$ that is in the first zone. Neglecting all but a single value of κ is in the spirit of the RPA, but it is not required by the present formalism and is done here only to provide a simple example of the present method.

C. The Phonon-Induced Electronic Interaction

The methods described here can be used to eliminate both plasmon and phonon collective excitations from the full Hamiltonian H of Eq. (36). The intrinsic Hamiltonian then contains an electronic interaction modified by plasmon screening and by addition of a phonon-induced term, where interactions between the two types of excitations are fully taken into account.

Equations (3) for the Hamiltonian H, in the RPA, have solutions of the form

$$A_\kappa^\dagger = \Sigma_{\sigma s}\alpha(\kappa\sigma s)a_{\sigma+\kappa,s}^\dagger a_{\sigma s} + \beta(\kappa)p_\kappa + \gamma(\kappa)q_\kappa^\dagger, \tag{85}$$

if umklapp terms that connect different values of κ are ignored. Here κ is the electronic momentum transfer, and p_κ and $q_\kappa^\dagger$ refer to the appropriate phonon mode in the first Brillouin zone. The linear equations satisfied by the coefficients are, using Eqs. (47) and (48),

$$\begin{aligned}(\hbar\omega - \epsilon_{\sigma+\kappa} + \epsilon_\sigma)\alpha(\kappa\sigma s) &= M_\kappa^2\Sigma_{\sigma',s'}(f_{\sigma' s'} - f_{\sigma'+\kappa,s'})\alpha(\kappa\sigma' s') + i\hbar v_\kappa^i\beta(\kappa),\\ \hbar\omega\beta(\kappa) &= -i\hbar\gamma(\kappa),\\ \hbar\omega\gamma(\kappa) &= i\hbar\Omega_\kappa^2\beta(\kappa) + v_{-\kappa}^i\Sigma_{\sigma s}(f_{\sigma s} - f_{\sigma+\kappa,s})\alpha(\kappa\sigma s).\end{aligned} \tag{86}$$

If the auxiliary parameter C_κ is defined by Eq. (54), above, these equations reduce to

$$\alpha(\kappa\sigma s) = \left[M_\kappa^2 + \frac{|v_\kappa^i|^2}{\omega^2 - \Omega_\kappa^2}\right]\frac{C_\kappa}{\hbar\omega - \epsilon_{\sigma+\kappa} + \epsilon_\sigma}, \tag{87}$$

$$\beta(\kappa) = v_{-\kappa}^i C_\kappa/i\hbar(\omega^2 - \Omega_\kappa^2), \tag{88}$$

$$\gamma(\kappa) = i\omega\beta(\kappa). \tag{89}$$

Following the derivation of Eqs. (53) and (55), above, unless C_κ vanishes, ω is determined by the modified dispersion relation

$$1 = \left[M_\kappa^2 + \frac{|v_\kappa^i|^2}{\omega^2 - \Omega_\kappa^2}\right]\Sigma_{\sigma s}\frac{f_{\sigma s} - f_{\sigma+\kappa,s}}{\hbar\omega - \epsilon_{\sigma+\kappa} + \epsilon_\sigma}. \tag{90}$$

It can easily be seen that both Eqs. (55) and (80) can be derived from this equation when v_κ^i is small, the latter result requiring also that $\hbar\omega$ be neglected in the energy denominator. Thus in general, when the summation is replaced by an integral, Eq. (90) has two positive roots, $\omega_{a\kappa}$ for the modified plasma excitation, and $\omega_{b\kappa}$ for the modified phonon excitation, where $\omega_{a\kappa} \gg \omega_{b\kappa}$. Since $\omega_{b\kappa}$ lies within the range of $\epsilon_{\sigma+\kappa} - \epsilon_\sigma$, the integral in Eq. (90) is complex, and $\omega_{b\kappa}$ will be complex. Phonon excitations are expected to be nonstationary because they can decay into the degenerate continuum of electron excitations. The natural width or lifetime can be computed from Eq. (90).

When ω is the modified plasma frequency $\omega_{a\kappa}$, it follows from Eqs. (43) and (65) that the second term in the square bracket in Eq. (90) contains a factor m/M, and thus can be neglected for all practical purposes. Hence the plasma excitations are adequately described by the analysis given above (Sec. VA), and the resulting screened Coulomb interaction in the intrinsic electronic Hamiltonian is given by Eq. (61).

By making appropriate definitions, the derivation following Eq. (75), above, can be repeated almost *verbatim*. Define

$$v_\kappa = v_\kappa^i + (M_\kappa^2/i\hbar)\Sigma_{\sigma s}(f_{\sigma s} - f_{\sigma+\kappa,s})\alpha(\kappa\sigma s)/\beta(\kappa), \tag{91}$$

and

$$\lambda_\kappa^2 = \kappa^2 M_\kappa^2\Sigma_{\sigma s}(f_{\sigma s} - f_{\sigma+\kappa,s})/(\epsilon_{\sigma+\kappa} - \epsilon_\sigma - \hbar\omega_{b\kappa}). \tag{92}$$

It has been shown by Chester[12] that the integrals occurring in Eqs. (75) and (92) differ only by a term of order $m/M \sim 10^{-4}$. In the present case Eq. (76) follows from these definitions, exactly as written, as does Eq. (80) for the renormalized phonon frequency. The dispersion relation, Eq. (90), leads to

$$\omega_{b\kappa}^2 - \Omega_\kappa^2 = M_\kappa^2 v_{-\kappa}^i(v_\kappa - v_\kappa^i), \tag{93}$$

which is identical in form to Eqs. (77) and (79). Eq. (87) becomes

$$\alpha(\kappa\sigma s) = (\hbar/i)v_\kappa\beta(\kappa)/(\epsilon_{\sigma+\kappa} - \epsilon_\sigma - \hbar\omega_{b\kappa}), \tag{94}$$

which may be compared with Eq. (73), above.

Since $\omega_{b\kappa}$ is complex, the solution denoted by $A_\kappa^\dagger(\omega)$ should be chosen so that Im $\omega < 0$. The operator denoted by $A_\kappa(\omega^*)$ is then obtained by taking the Hermitian conjugate of the solution $A_\kappa^\dagger(\omega^*)$ with Im $\omega^* > 0$. Since all coefficients in Eq. (90) are real, it is obvious that complex roots occur in conjugate pairs. Thus $A_\kappa(\omega^*)$ is obtained from $A_\kappa^\dagger(\omega)$ by taking the Hermitian conjugate af all operators and the complex conjugate of all numerical coefficients, but always treating the parameter ω as if it were a real number. For traveling waves, the commutator in Eq. (32) is replaced by $[\theta A_{-\kappa}\theta^{-1}, A_\kappa^\dagger]$.

If $\alpha(\kappa\delta s)$ and $\gamma(\kappa)$ are expressed in terms of $\beta(\kappa)$, by Eqs. (89) and (94), the normalization condition, Eq. (32′), becomes

$$\beta^*(\omega^*)\beta(\omega) = \frac{1}{2\hbar\omega}\left[1 + \frac{\hbar^2 |v_\kappa|^2}{2\hbar\omega}\frac{\partial}{\partial(\hbar\omega)} \times \Sigma_{\sigma s}\frac{f_{\sigma s} - f_{\sigma+\kappa,s}}{\epsilon_{\sigma+\kappa} - \epsilon_\sigma - \hbar\omega}\right]^{-1}. \quad (95)$$

Since, by Chester's analysis,[12] the sum over δs in this equation depends on $\hbar\omega$ only to the extent of a correction of order 10^{-4}, the second term in square brackets is of this order and it can be neglected.

The phonon-induced electronic part of the final intrinsic Hamiltonian follows from Eqs. (34′), (94), and (95). It is

$$\Delta H_e = -\tfrac{1}{2}\Sigma_\kappa \Sigma_{\sigma s}\Sigma_{\sigma',s'} \times \frac{\hbar^2 |v_\kappa|^2}{(\epsilon_{\sigma+\kappa} - \epsilon_\sigma - \hbar\omega_{b\kappa})(\epsilon_{\sigma',+\kappa} - \epsilon_{\sigma'} - \hbar\omega_{b\kappa})} \times a^\dagger_{\sigma+\kappa,s}a_{\sigma s}a^\dagger_{\sigma',s'}a_{\sigma',+\kappa,s'}. \quad (96)$$

This interaction contains a term analogous to that considered in the theory of superconductivity, which is known to lead to a finite energy gap in the electronic excitation spectrum. When there is such a gap, it is to be expected that the Bogoliubov–Valatin canonical transformation, which modifies the electron annihilation and creation operators, will modify the energy denominators appearing here. In particular, in Eq. (90), $\hbar\omega$ should become complex only when its real part exceeds the electronic energy gap.

VI. FORMAL PROPERTIES OF EXCITATION OPERATORS

If $\hbar\omega$ is assumed to be complex, with negative imaginary part, the equations considered here are

$$[H, A^\dagger] = \hbar\omega_a A^\dagger, \quad [A^\theta, H] = \hbar\omega_a A^\theta, \quad (97)$$

where A^θ denotes $\theta A \theta^{-1}$. Consider some other excitation operator $B^\dagger$ with excitation energy $\hbar\omega_b$. Then

$$[B^\theta, [H, A^\dagger]] = \hbar\omega_a[B^\theta, A^\dagger], \quad [[B^\theta, H], A^\dagger] = \hbar\omega_b[B^\theta, A^\dagger]. \quad (98)$$

The difference between these two equations is

$$[H, [B^\theta, A^\dagger]] = \hbar(\omega_a - \omega_b)[B^\theta, A^\dagger]. \quad (99)$$

This equation shows that the orthonormalization condition given by Eq. (32) is consistent with the assumed operator equations. If $\omega_a \neq \omega_b$, Eq. (99) implies that $[B^\theta, A^\dagger]$ is itself an excitation operator. If this commutator is approximated by a c number, the left side of Eq. (99) vanishes, which requires that the number must be set equal to zero. If $\omega_a = \omega_b$, $[B^\theta, A^\dagger]$ is a constant of motion, and can be set equal to unity or zero as a normalization condition. Thus different eigenvalues of Eqs. (97) correspond to orthogonal operators, which can be normalized unless $[A^\theta, A^\dagger]$ vanishes. Degenerate solutions could be orthonormalized by the usual Schmidt process.

The variational principle proposed by Thouless[14] can be generalized by using the operator algebra that leads to Eq. (99). In particular, if an excitation operator that satisfies Eqs. (97) is modified by an infinitesimal variation δA that preserves normalization, so that

$$[\delta A^\theta, A^\dagger] + [A^\theta, \delta A^\dagger] = 0, \quad (100)$$

then it can easily be shown that the quantity

$$\hbar\omega = [A^\theta, [H, A^\dagger]]/[A^\theta, A^\dagger] \quad (101)$$

is stationary with respect to such a variation. This result is proved by replacing A by $A + \delta A$, and then using Eqs. (97) and (100) to show that all first-order terms vanish. The stationary value of Eq. (101) is equal to the excitation energy $\hbar\omega_a$.

The operator orthonormalization considered here can be used to discuss vector-space properties of a given set of operators, in analogy to the discussion of the completeness of the set of random phase approximation excitation operators given by Thouless,[14] who noted that solutions of equations analogous to Eqs. (12) were needed to complete the set of excitation operators. The proof given here in Sec. III, is a generalization of the remark made by Thouless that these equations are identical to those of the cranking model.

Approach to Stability of a Plasma

B. W. ABRAHAM
*Faculté des Sciences, Université Libre de Bruxelles, Brussels, Belgium**
(Received 3 January 1964; final manuscript received 21 July 1964)

The stabilization of a spatially homogeneous plasma by the collision term of the new kinetic equation of Balescu for a weakly unstable plasma is shown for a simplified model. The one-particle velocity distribution consists of two symmetrically displaced Lorentz distribution functions in which the displacement velocity and velocity spread are functions of time. For this distribution function, the collision term decreases the displacement velocity and increases the velocity spread until the plasma is quite stable. Physically this process can be viewed as conversion of relative energy to kinetic energy. This is demonstrated to be equivalent to the passage of the zero of the plus dielectric constant from the upper half complex plane, characterizing an unstable plasma, to the lower half complex plane, characterizing a stable plasma.

INTRODUCTION

A KINETIC equation for a spatially homogeneous plasma in the case of weak instabilities has recently been derived by Balescu.[1] It reduces in the case of an extremely stable plasma to the kinetic equation for the plasma previously derived by Lenard[2] and Balescu.[3] New collision terms of the Fokker–Planck form, highly nonlinear in the velocity distribution function, are found. The character of the collision term drives the system toward stability.

To demonstrate the stabilization of the system by collision terms we chose a particular model for the one-particle velocity distribution function, the symmetric Lorentz form. Physically this can correspond to electron or ion beams where the relative velocity of the two beams and the velocity spread of the beams are functions of time. It is shown explicitly for this model that the enhanced collisions do drive the system toward stability. As the relative velocity decreases and the velocity spread increases, the velocity distribution function of the system reaches a stable form. This is equivalent to the conversion of relative energy of the two beams to random kinetic energy of each beam. The model corresponds to two electron beams in the presence of a positive background. The Lorentz distribution function is shown to be consistent with the kinetic equation for the weakly unstable and just stable regions which are the regions of interest.

Mathematically, an unstable plasma can be related to the presence of zeros of the dielectric constant in the upper half complex plane. The stabilization of the system is characterized by the passage in time of γ, the imaginary part of the zero, into the lower half complex plane. One can, for our simple model, relate $d\gamma/dt$ to the time derivative of the relative energy where $d\gamma/dt$ is always decreasing. As γ eventually passes into the lower half complex plane or the plasma stabilizes, the mathematical and physical requirements for stabilization are connected.

First in Sec. I we consider the instability requirements and the unstable kinetic equation. Then in II, the particular model, Lorentz velocity distribution function, is discussed and total energy conservation expression evaluated. As the latter is inadequate to describe the stabilization, the conversion of relative energy of the beam to heat energy is utilized as a criteria in III.

I. BASIC EQUATIONS

The concept of unstable plasmas arose from the study of plasma oscillations.[1,4,5] A stable plasma is one which can support only damped oscillations in time, whereas an unstable plasma can support exponentially growing oscillations. These oscillations are the collective oscillations of the density or electric field of a slightly perturbed plasma which obeys the linearized Vlassov equation. The criterion for stability of a plasma is given by the dielectric constant $\epsilon^+(\nu)$, defined here for real frequency

$$\epsilon^+(\nu) = 1 - \frac{i\pi\omega_p^2}{l^2}\int d\mathbf{v}\ \delta(l\nu - \mathbf{l}\cdot\mathbf{v})\mathbf{l}\cdot\partial\varphi(\mathbf{v}) = \epsilon_1 + i\epsilon_2. \quad (1)$$

The ω_p is the plasma frequency, $\mathbf{l}$ the wave vector, and $\varphi(\mathbf{v})$ the one-particle velocity distribution. The $\delta_\pm$ functions are defined by

$$\delta_\pm = \delta(X) \pm i(P/\pi)(1/X). \quad (2)$$

* Present address: Space Physics Division, NASA–Ames Research Center, Moffett Field, California.

[1] R. Balescu, J. Math. Phys. **4**, 1009 (1963).
[2] A. Lenard, Ann. Phys. **3**, 390 (1963).
[3] R. Balescu, Phys. Fluids **3**, 52 (1960).
[4] L. Landau, J. Phys. (USSR) **10**, 25 (1946).
[5] J. D. Jackson, J. Nucl. Energy Pt. C, **1**, 171 (1960).

If $\epsilon^+(\nu)$ has one or more zeros in the upper half complex plane, the plasma is unstable. Here we restrict ourselves to the case of one zero ζ_+, or

$$\epsilon^+(\nu) = (\nu - \zeta_+)\sigma^+(\nu) \tag{3}$$

and ζ_+ is also a function of l. The existence of an unstable root depends upon the shape of the one-particle velocity distribution $\varphi(\mathbf{v})$ or more precisely the barred quantity $\bar{\varphi}(\nu)$ obtained from $\varphi(\mathbf{v})$ by integrating over velocities transverse to 1:

$$\bar{\varphi}(\nu) = \int \varphi(\mathbf{v})\delta\left(\nu - \frac{\mathbf{l}\cdot\mathbf{v}}{l}\right) d\mathbf{v}. \tag{4}$$

The $\bar{\varphi}(\nu)$ must be at least double-humped in ν space for an unstable zero to occur.

The plus and minus functions, $\bar{\varphi}'_\pm(\nu)$, which enter the expressions for $\epsilon^\pm(\nu)$, are defined in general by

$$A_\pm(X) = \pm\tfrac{1}{2}\int \delta_\pm(x - y)A(y)\, dy, \tag{5}$$

or Eq. (3) can be expressed by

$$\epsilon^+(\nu) = 1 - i\pi\omega_p^2\bar{\varphi}'_+(\nu)/l^2. \tag{6}$$

The magnitude of the imaginary part of the unstable root determines the degree of instability. If

$$\zeta_+ = \omega_0 + i\gamma, \tag{7}$$

then the requirement for weak instability is

$$l\bar{\gamma} \ll \omega_p, \tag{8}$$

where the maximum possible $l\gamma$ for a given plasma must be much smaller than the plasma frequency ω_p. If the zero ζ_+ is in the lower half plane but

$$\bar{l}\,|\bar{\gamma}| \ll \omega_p, \tag{9}$$

still holds, the system is weakly stable. This requirement is equivalent to a separation of the collision time scale ω_p^{-1} and the longer stability time $(l\gamma)^{-1}$.

Recently a kinetic equation for a spatially homogeneous plasma for $\varphi(\mathbf{v})$ has been derived in the unstable case. The characteristic feature of this equation, in the weakly unstable regime, is the occurrence of new "stabilization terms" which add to the normal collision term.

The kinetic equation for $\varphi(\mathbf{v})$ has the following form:

$$\partial_t\varphi(\mathbf{v};\, t) = -\omega_p^2\int d\mathbf{l}\,\frac{i\mathbf{l}}{l^2}\cdot\partial\mathfrak{F}_1(\mathbf{v};\, t) \tag{10}$$

in the unstable case,[6] where $\mathfrak{F}_1(\mathbf{v};\, t)$ is defined by the following integral:

$$\mathfrak{F}_1(\mathbf{v};\, t) = -\frac{1}{2\pi i}\int d\omega\,\frac{e^{-il\omega t}}{\omega}F_1(\mathbf{v};\omega). \tag{11}$$

The explicit form of $\mathfrak{F}_1(\mathbf{v};\, t)$ in the weakly unstable and stable case is then

$$\begin{aligned}\mathfrak{F}_1(\mathbf{v};\, t) = \mathrm{Im}\Bigg\{&\frac{q_1(\mathbf{v};0)}{\epsilon^-(\nu)} + \frac{\pi i}{l}d_1(\mathbf{v})\\ &\times\int\frac{d\nu_1\,\delta_-(\nu-\nu_1)\bar{R}(\nu_1;0)}{(\nu_1-\zeta_+)(\nu_1-\zeta_-)} + \frac{e^{-i(\nu-\zeta_-)lt}}{(\nu-\zeta_-)}\\ &\times\left[\frac{-q_1(\mathbf{v};\,\nu-\zeta_-)}{\sigma^-(\zeta_-)} - \frac{2\pi_i}{l}\frac{d_1(\mathbf{v})\bar{R}(\nu;\,\nu-\zeta_-)}{(\nu-\zeta_+)}\right]\\ &+ e^{-i(\zeta_+-\zeta_-)lt}\frac{2\pi i d_1(\mathbf{v})\bar{R}(\zeta_+;\,\zeta_+-\zeta_-)}{l(\zeta_+-\zeta_-)(\nu-\zeta_+)}\Bigg\},\end{aligned} \tag{12}$$

where it should be noted that only the imaginary part of the right-hand side of Eq. (12) contributes to the collision integral. The third term may also be rewritten as

$$\frac{e^{-i(\nu-\zeta_-)lt}}{(\nu-\zeta_-)}\left[\frac{-q_1(\mathbf{v};\,\nu-\zeta_-) + q_1^*(\mathbf{v};0)}{\sigma^-(\zeta_-)\epsilon^+(\nu)}\right]$$

as a form useful in finding explicit expressions for dU/dt and $d\mathcal{E}_T/dt$. The following quantities are defined for convenience:

$$\begin{aligned}q_1(\mathbf{v};\omega) = q_1 + iq_2 = \frac{1}{2}\left(\frac{i\omega_p^2}{4\pi^2l^2}\right)\int d\mathbf{v}\\ \times\,\delta_-(\mathbf{l}\cdot\mathbf{v} - l\omega - \mathbf{l}\cdot\mathbf{v}_1)l\cdot(\partial - \partial_1)\varphi\varphi_1\end{aligned} \tag{13}$$

for q_1 and q_2 real functions;

$$\begin{aligned}&\bar{R}(\nu;\omega)\\ &= \left(\frac{i\omega_p^2}{4\pi^2cl^2}\right)\frac{\bar{\varphi}'_-(\nu-\omega)\bar{\varphi}_+(\nu) - \bar{\varphi}_-(\nu-\omega)\bar{\varphi}'_+(\nu)}{\sigma^+(\nu)\sigma^-(\nu-\omega)}.\end{aligned} \tag{14}$$

The explicit forms of the kinetic equation for weakly stable and unstable plasmas are given in (6.6) and (8.7) of Ref. 1. The expressions for $\mathfrak{F}_1(\mathbf{v};\, t)$ for the two cases are different in appearance after the rapidly oscillating term $e^{il(\nu-\zeta_-)t}$ is smoothed out, but the two solutions are continuous. The two expressions for $\mathfrak{F}_1$ are, however, valid in many physical systems for only a limited range of 1 space. For a given system the zero of ϵ^+ can lie in any of the three regions defined in Fig. 1, depending on the value of 1. Since the choice of unstable, stable or normal form of $\mathfrak{F}_1$ depends upon the location of the zero of ϵ^+, then in general a collision integral is a sum of three terms, weakly unstable, weakly stable, and normal $\mathfrak{F}_1(\mathbf{v};\, t)$, each integrated over a portion of 1 space. These regions of 1 space are called U, for weakly unstable, S for weakly stable and N for normal stable cases.

[6] R. Balescu, *Statistical Mechanics of Charged Particles* (Interscience Publishers, Inc., New York, 1964).

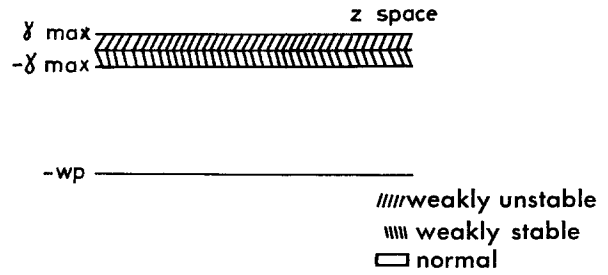

FIG. 1.

The regions of Fig. 1 in the complex plane correspond to integration regions of 1 space. In Fig. 1 the value of $\bar{\gamma}$ is the maximum acceptable value of γ, some fraction of ω_p/l. For a particular system ζ^+, the zero of ϵ^+, lies between $\bar{\gamma}$ and 0 for the weakly unstable case. The strip chosen for the weakly stable case is somewhat arbitrary but leads to simplicity in calculations and separates clearly the collision and unstable time scales. For $\gamma < \bar{\gamma}$, the collective transient is no longer sufficiently long lived to be considered separately from the other transients. It is dropped and the usual normal collision term is recovered for the normal stable case.

Many physical systems may be described by these equations (i.e., electron–ion beams, electron beams, or ion beams). Here we consider the particular model, applicable to electron beams in a positive background with a symmetrically displaced Lorentz velocity distribution function. The Lorentz velocity distribution function is defined to be

$$\varphi(\mathbf{v}) = \varphi^1(\mathbf{v}) + \varphi^2(\mathbf{v}), \tag{15}$$

where

$$\varphi^{1,2} = (u_0/2\pi)[(\mathbf{v} \mp \mathbf{v}_0)^2 + u_0^2]^{-2},$$

the upper sign referring to φ^1. Also the barred function $\bar{\varphi}(\nu)$ is

$$\bar{\varphi}(\nu) = \frac{u_0}{2\pi}\left[\frac{1}{(\nu - \nu_0)^2 + u_0^2} + \frac{1}{(\nu + \nu_0)^2 + u_0^2}\right]. \tag{16}$$

The velocity spread is given by u_0, and the mean velocity by $\pm\mathbf{v}_0$ for φ^1 and φ^2, respectively. Here

$$\nu_0 \equiv \mathbf{l}\cdot\mathbf{v}_0/l = v_0\mu, \tag{17}$$

where μ is the cosine of the angle between $\mathbf{v}_0$ and $\mathbf{l}$, the wave vector. The parameters u_0 and v_0 are functions of time or the shape of the velocity distribution can change. If the system is originally unstable, the frictional-diffusion action of the additional collision terms can be expected to reduce v_0 and increase u_0 to stabilize the system. The distribution function is a model similar in spirit to one used by Dreicer.[7] He demonstrated that the shape was retained by energy arguments. Here it is simpler to see that all quantities do not change greatly in the passage from weak instability to weak stability. Within the approximation in which all quantities the order of γ/ω_p are dropped, we see that the use of Lorentz distribution function is consistent with the kinetic equations. For example the maximum possible change of φ can be seen below to be

$$\Delta\varphi/\varphi = 0(\bar{\gamma}/u_0); \qquad \bar{\gamma}/u_0 \ll 1, \tag{18}$$

or it is quite small; therefore the fact that φ retains its shape during the stabilization process is not surprising. Once the system reaches the normal stable region, the Lorentz distribution function can no longer be assumed, but from previous results one expects the system to approach equilibrium.

The limits of integration in 1 space for the three cases are very model dependent, differing for displaced Maxwellian and Lorentz velocity distribution but having the same general appearance. For the Lorentz velocity distribution function, we have the expression for the zero of ϵ^+ in the upper half complex plane in the unstable case,[8]

$$l\gamma = -lu_0 \pm [(2l^2\mu^2v_0^2\omega_p^2 + \tfrac{1}{4}\omega_p^4)^{\frac{1}{2}} - l^2\mu^2v_0^2 - \tfrac{1}{2}\omega_p^2]^{\frac{1}{2}}. \tag{19}$$

The requirement that the system is weakly unstable gives

$$\gamma \ll u_0. \tag{20}$$

The above inequality must hold for all $\gamma(\mathbf{l})$ as a function of (l, μ). In the limit of small l, we have, using (19) and (20),

$$-\bar{\gamma} \leq v_0\mu - u_0 \leq \bar{\gamma}, \tag{21}$$

where $\bar{\gamma}$ is the maximum acceptable values of γ, some fraction of u_0 at equilibrium. From Eq. (21) we find, setting $\mu = 1$, the range of relative velocity for which the system will be in the entire range of l, at most, weakly unstable or stable.

Using this range and (21), the limits on μ follow for positive μ as the integrand is even in μ. The limits on l are found by solving (19) for l and then using the approximate range of μ.

Weakly unstable:

$$u_0/v_0 \leq \mu \leq 1, \tag{22}$$

$$0 \leq l \leq (v_0^2\mu^2 - u_0^2)^{\frac{1}{2}}/(v_0^2\mu^2 + u_0^2)\omega_p; \tag{23}$$

[7] H. Dreicer, Phys. Rev. 115, 2 (1959).
[8] Reference 6, p. 114.

Weakly stable:

$$(u_0 - \gamma)/v_0 \leq \mu \leq \min(1, u_0/v_0), \tag{24}$$

$$\max[0, (v_0^2\mu^2 - u_0^2)^{\frac{1}{2}}/(v_0^2\mu^2 + u_0^2)\omega_p] \leq l, \tag{25}$$

where the limits automatically insure that the new collision terms disappear for the normal case.

For approximate evaluation of integrals the simplified limits are helpful. The maximum possible range of μ for weakly unstable and stable cases is

$$1 - 2\epsilon \leq \mu \leq 1, \qquad \epsilon \equiv \tilde{\gamma}/u_0, \tag{26}$$

and l is

$$0 \leq l \leq (\epsilon)^{\frac{1}{2}}(\omega_p\lambda_D/u_0). \tag{27}$$

It should be noted that u_0 although a function of time varies only by a factor $(2)^{\frac{1}{2}}$.

The knowledge of the mean kinetic energy and correlation energy is insufficient to give the stabilization of a plasma as is seen below. For the normal term[6] $dU/dt = 0$ so that only the new collision terms contribute. If the mean kinetic energy U and the potential correlation $\mathcal{E}_{cor}$ are defined by

$$U = \int \tfrac{1}{2}mv^2\varphi(\mathbf{v}; t)\, d\mathbf{v},$$

$$\mathcal{E}_{cor} = \int \tfrac{1}{2}V(1)g_1(t)\, d1, \tag{28}$$

where g_1 is a two-particle correlation function, then the sum is conserved as can be seen by a simple derivation from the kinetic equations,

$$d[U + \mathcal{E}_{cor}]/dt = 0. \tag{29}$$

Using the kinetic equation (19), we have

$$\frac{dU}{dt} = 4\pi e^2 c \int d1 \int d\mathbf{v} \int_0^t d\tau \frac{i\nu}{l} \mathfrak{F}_1(\mathbf{v}; t). \tag{30}$$

The relation (12) for $\mathfrak{F}_1(\mathbf{v}; t)$ substituted in Eq. (30) gives

$$\begin{aligned}\frac{dU}{dt} &= 4\pi e^2 c \int_U \frac{d1}{l^2} \frac{(2\pi i)(i\zeta_+ l)}{(\zeta_+ - \zeta_-)} \\ &\times \left[\frac{8\pi^2 ice^2}{ml^2} \bar{\varphi}'_+(\zeta_+)\right] \bar{R}(\zeta_+; \zeta_+ - \zeta_-) e^{2\gamma l t} \\ &+ 4\pi e^2 c \int_S \frac{d1}{l^2} \frac{(2\pi i)(i\zeta_+ l)}{(\zeta_+ - \zeta_-)} \\ &\times \left[\frac{8\pi^2 ice^2}{ml^2} \bar{\varphi}'_-(\zeta_+)\right] \bar{R}(\zeta_+; \zeta_+ - \zeta_-) e^{2\gamma l t},\end{aligned} \tag{31}$$

where the integration is over the volume of 1 space as specified. The analytic continuation of $\bar{\varphi}_+(\zeta_+)$ gives

$$\bar{\varphi}'_+(\zeta_+) = \bar{\varphi}'_-(\zeta_+) + \bar{\varphi}(\zeta_+) \tag{32}$$

for the stable case where

$$(8\pi^2 ice^2/ml^2)\bar{\varphi}'_+(\zeta_+) = 1 \tag{33}$$

from the definition of the dielectric constant $\epsilon^+(\zeta_+)$ and the fact that ζ_+ is a zero. Also from Eq. (31) we have the relation

$$(8\pi^2 ice^2/ml^2)\bar{\varphi}'(\zeta_+) = \epsilon^-(\zeta_+). \tag{34}$$

From Eq. (14), again using relation (33), we find a simplified expression for $\bar{R}$:

$$\bar{R}(\zeta_+; \zeta_+ - \zeta_-) = \frac{-1}{8\pi^3 c} \frac{(\zeta_+ - \zeta_-)}{2\pi i\, |\sigma^+(\zeta_+)|^2} \int \frac{\bar{\varphi}(\nu)\, d\nu}{(\nu^2 + \gamma^2)}, \tag{35}$$

where for the Lorentz distribution we have

$$\int \frac{\bar{\varphi}(\nu)\, d\nu}{(\nu^2 + \gamma^2)} = \frac{u_0 + \gamma}{\gamma[\nu_0^2 + (u_0 + \gamma)^2]}. \tag{36}$$

Then substituting Eqs. (32) to (36) into (31) gives

$$\begin{aligned}\frac{dU}{dt} &= \frac{e^2}{2\pi^2} \int_{U+S} \frac{d1}{l} \frac{e^{2\gamma l t}}{|\sigma^+(\zeta_+)|^2} \frac{u_0 + \gamma}{[\nu_0^2 + (u_0 + \gamma)^2]} \\ &- \frac{e^2}{2\pi^2} \int_S \frac{d1}{l^3} \frac{\gamma e^{2\gamma l t}}{|\sigma^+(\zeta_+)|^2} \frac{u_0 + \gamma}{[\nu_0^2 + (u_0 + \gamma)^2]} \\ &\times 2\omega_p^2[3\nu_0^2 - (u_0 - \gamma)^2].\end{aligned} \tag{37}$$

From Eq. (37) we see that dU/dt has no definite sign or the stabilization cannot be deduced from this term. The time derivative of the kinetic energy of each beam is shown, however, to have a definite sign.

II. CONVERSION OF RELATIVE ENERGY TO HEAT ENERGY

In this section we demonstrate that the plasma stabilizes under the combined effect of enhanced diffusion and friction of the new collision terms. The relative energy is converted by collision into heat energy or random energy. (It should be noted that the kinetic energy of the system and the heat energy of each beam are not finite for the Lorentz velocity distribution function, but the time rate of change is. For a more realistic model these quantities are, of course, well defined. The essential condition for stability, that the relative velocity of the peaks decreases and the velocity spread increases is shown explicitly.) This is equivalent to a decrease in the relative velocity $\mathbf{v}_0$ and an increase in the thermal velocity u_0. These changes cause γ to decrease as seen from (19) until the system is stable. The plasma is assumed to consist of two beams described by

the symmetric Lorentz distribution function. As we discuss the stabilization of the plasma by means of certain macroscopic parameters, these are here defined as moments of the one-particle distribution function as follows:

Mean velocity

$$q_i = \int \mathbf{v}\varphi_i \, d\mathbf{v}/c_i, \tag{38}$$

where c_i is the density of particles for this example. The subscripts ($i = 1, 2$) refer to the beams; the absence of subscripts indicates that the quantity refers to the entire system.

Heat energy of beam

$$\mathcal{E}_{T,i} = \tfrac{3}{2}c_i kT_i = \tfrac{1}{2}\int m_i(\mathbf{v} - \mathbf{q}_i)^2\varphi_i \, d\mathbf{v}; \tag{39}$$

Relative energy of beam

$$\mathcal{E}_{r,i} = \tfrac{1}{2}mc(\mathbf{q}_i - \mathbf{q})^2; \tag{40}$$

Kinetic energy of system

$$U = \tfrac{1}{2}ckT + \tfrac{1}{2}mcq^2 = \tfrac{1}{2}m\int v^2\varphi \, d\mathbf{v}. \tag{41}$$

The kinetic energy of the system U in general is related to the energy of the beams

$$U = \sum_i (\mathcal{E}_{T,i} + \mathcal{E}_{r,i}) + \tfrac{1}{2}mcq^2. \tag{42}$$

For the special model chosen here, $\mathbf{q}$, the mean velocity of the system, vanishes. The vanishing of $d\mathbf{q}/dt$ for this model further simplifies the expression for the time derivative of the drift energy.[9] From Eq. (40) and the two above properties of $\mathbf{q}$ we have that

$$d\mathcal{E}_r/dt = mc\mathbf{v}_0 \cdot d\mathbf{v}_0/dt. \tag{43}$$

The kinetic equation for each beam can be found in a similar fashion to that found for the entire plasma. Only one beam need be considered as the two beams are symmetrical. The kinetic equation for a beam is

$$\partial_t\varphi^1(\mathbf{v}; t) = -\omega_p^2 \int d\mathbf{l} \, i\mathbf{l}\cdot\partial G_1(\mathbf{v}), \tag{44}$$

where G is defined by Eq. (12) and G is defined as

$$G_1(\mathbf{v}) = F_{11} + F_{12}.$$

F_{ij} is the correlation function of i and j particles integrated over one velocity. If one defines $F_1(\mathbf{v})$ to be the sum of all F_{ij},

$$F_1(\mathbf{v}) = \sum_{i,j}^{1} F_{ij}, \tag{45}$$

[9] This can easily be shown explicitly by integrating (30) over velocity.

then the integral equations for the barred quantities $\bar{G}_1(\nu)$ and $\bar{F}_1(\nu)$ are

$$\epsilon^-(\nu - \omega)\bar{F}_1(\nu) - i\epsilon_2(\nu)\int d\nu_1 \delta_-(\nu - \omega - \nu_1)\bar{F}_{-1}(-\nu_1) = \bar{q}(\nu);$$

$$\epsilon^-(\nu - \omega)\bar{G}_1(\nu) - i\epsilon_2^1(\nu)\int d\nu_1 \delta_-(\nu - \omega - \nu_1)\bar{F}_{-1}(-\nu_1) = \bar{q}^1(\nu). \tag{46}$$

The new quantities ϵ_2^1 and $\bar{q}^1$ are defined by

$$\epsilon_2^1(\nu) = (\omega_p^2/l^2)\pi\varphi_1^{1'};$$

$$\bar{q}^1(\nu) = \frac{1}{2}\left(\frac{i\omega_p^2}{4\pi^2 l^2}\right)\int d\mathbf{v}\, d\mathbf{v}_1 \, \delta_-(\mathbf{l}\cdot\mathbf{v} - l\omega - \mathbf{l}\cdot\mathbf{v}_1) \times \delta(\nu - \mathbf{l}\cdot\mathbf{v}/l)\mathbf{l}\cdot(\partial - \partial_1)\varphi^1(\mathbf{v})\varphi(\mathbf{v}_1). \tag{47}$$

The known solution for $\bar{F}_1(\nu)$ enables the solution for $\bar{G}_1(\nu)$ to be found immediately

$$\bar{G}_1(\nu) = \bar{q}^1(\nu)/\epsilon^-(\nu - \omega) + i\epsilon_2^1(\nu)\int \frac{d\nu_1 \, \delta_-(\nu - \nu_1)\bar{R}(\nu_1; \omega)}{(\nu_1 - \zeta_+)(\nu_1 - \omega - \zeta_-)}. \tag{48}$$

The calculations will be very similar to those previously done as $\bar{G}$ differs from $\bar{F}$ only in the $\bar{q}^1(\nu)$ and $\epsilon_2^1(\nu)$ terms. We see from Eq. (47) that for $\varphi(\mathbf{v})$ we have substituted $\varphi^1(\mathbf{v})$.

The equation for the time derivative of the relative energy from Eqs. (43), (44), and (45) becomes

$$\frac{d\mathcal{E}_r}{dt} = -4\pi^2 e^2 c\int \frac{d\mathbf{l}}{l^3}\nu_0 \int d\mathbf{v}\, \bar{q}_2(\nu; 0) + 4\pi^2 e^2 ci\int_{U+S}\frac{d\mathbf{l}}{l^3}\nu_0\int d\mathbf{v}\left\{\frac{e^{-i(\nu-\zeta_-)lt}}{(\nu - \zeta_-)} \times \left[\frac{-q_1(\mathbf{v}; \nu - \zeta_-) + q_1(\mathbf{v}; 0)}{\sigma^-(\zeta_-)\epsilon^+(\nu)}\right] + \frac{e^{-i(\zeta_+ - \zeta_-)lt}}{(\zeta_+ - \zeta_-)}\frac{2\pi i d_1(\mathbf{v})\bar{R}(\zeta_+; \zeta_+ - \zeta_-)}{l(\nu - \zeta_-)}\right\}. \tag{49}$$

It should be noted that the new collision terms are integrated only over the weakly unstable and weakly stable portion of l space. Taking the poles in ν and defining the barred quantities as previously, we find for Eq. (49)

$$\frac{d\mathcal{E}_r}{dt} = -4\pi e^2\int_{\nu-(U+S)}\frac{d\mathbf{l}}{l^3}\nu_0\int d\nu\,\frac{\bar{q}_2^1(\nu; 0)}{\epsilon^-(\nu)\epsilon^+(\nu)} - \left(\frac{e^2}{2\pi^2}\right)\omega_p^2\int_{U+S}\frac{d\mathbf{l}}{l^3}\frac{\nu_0^2 u_0^2}{\gamma[\nu_0^2 + u_0^2]^3\,|\sigma^+(i\gamma)|^2}(e^{2\gamma lt} - 1) + \left(\frac{e^2}{m\pi}\right)\omega_p^2\int_{U+S}\frac{d\mathbf{l}}{l^3}f(u_0, \nu_0; l; \omega_p), \tag{50}$$

where the real part of the normal collision term integrated over the unstable and just-stable part of l space has been removed from the normal collision term. Here $f(u_0, v_0; l; \omega_p)$ is a complicated function but for small l it is the order of $(l^4/\omega_p^4)Ku^4$. To find the order of magnitude of this term, we recall that the limits of μ and l are over the maximum ranges of order ϵ and $\epsilon^{\frac{1}{2}}$, respectively, ϵ having been defined in Eq. (26). Expanding the integral in power series in ϵ, keeping lowest-order terms, we find the maximum value of the last term of Eq. (50) to be $\epsilon^3 K/4$ (K is a constant the order of 1). Therefore the last term of Eq. (45) may be disregarded with respect to the first term. This is demonstrated in the Appendix explicitly for the case in which the first term may be expanded to give, as a first approximation, the Landau term. The Landau term will give for this model a similar contribution as found by Dreicer,[7] for dynamical friction. As the ratio of the dynamical friction term to the third term of Eq. (50) is the order, $\epsilon^{-3} \ln (L/\lambda_D)$ the latter may be neglected.

Having shown that the third term of Eq. (50) can be neglected with respect to the first term, the first term is always negative as it is the dynamical friction term, and noticing that the second term is always negative, we have

$$d\mathcal{E}_r/dt < 0. \tag{51}$$

The time behavior of $\mathcal{E}_T$ can be found from the relationship

$$d\mathcal{E}_T/dt = dU/dt - d\mathcal{E}_r/dt. \tag{52}$$

It can be seen from a comparison of Eqs. (37) and (50) that dU/dt is much smaller than $d\mathcal{E}_r/dt$ or we shall neglect it. This can be most easily seen if one expands the exponential $\exp(2\gamma lt)$ in a power series. The term of dU/dt independent explicitly of time is much smaller than the first term of Eq. (50). The time-independent terms can be compared term by term. The integrand of $d\mathcal{E}_T/dt$ is always γ^{-1} greater than that of dU/dt. As the integrands are positive, each term of $d\mathcal{E}_T/dt$ is much greater than of dU/dt. Thus

$$d\mathcal{E}_r/dt \approx -d\mathcal{E}_T/dt. \tag{53}$$

From Eq. (53) we have the result that the thermal energy of each beam is monotonically increasing. Equation (51) and (53) are equivalent to the statement that the relative velocity of each beam decreases and the thermal velocity increases.

Finally we relate the time behavior of the relative energy to the time behavior of γ. The time derivative of γ can be found from Eq. (19) where quantities the order of (γ/u_0) have been neglected.

$$\frac{d\gamma}{dt} = \frac{1}{mcu_0}\left[2\left(1 + \frac{8l^2v_0^2}{\omega_p^2}\right)^{-\frac{1}{2}}\frac{d\mathcal{E}_r}{dt} - \frac{dU}{dt}\right]. \tag{54}$$

From above we have that dU/dt is negligible with respect to $d\mathcal{E}_r/dt$ or

$$\frac{d\gamma}{dt} \approx \frac{1}{mcu_0}\left(1 + \frac{8l^2v_0^2}{\omega_p^2}\right)\frac{d\mathcal{E}_r}{dt}. \tag{55}$$

As $d\mathcal{E}_r/dt$ is always negative, it follows that

$$d\gamma/dt < 0. \tag{56}$$

From Eq. (50) for $d\mathcal{E}_r/dt$, one observes that there is never a value of γ for which $d\mathcal{E}_r/dt = 0$ in the weakly unstable or weakly stable range. Thus one can conclude that γ cannot approach some asymptote $\gamma = \gamma_0$ as $d\gamma/dt$ is never zero. Therefore γ continues to decrease until it becomes negative or the system is stable.

In conclusion, we have shown the physical requirement—that the system stabilizes, the relative energy decreases, and the thermal energy increases—is equivalent to the mathematical requirement that the imaginary part of the zero of the dielectric constant becomes negative.

This verifies the suggestion of Balescu that the new frictional and diffusional collision terms should stabilize the unstable plasma by decreasing the relative velocity and increasing the velocity spread. A variation of the model is being utilized to study electron runaway.

ACKNOWLEDGMENTS

We wish to thank Professor I. Prigogine for his interest and encouragement. Also we would like to thank Professor R. Balescu for many helpful discussions.

The support of the American Association of University Women is gratefully acknowledged. This work also has been partially supported by the U. S. Air Force under Contract No. AF 61(052)-179 and monitored by the European Office, Office of Aerospace Research.

APPENDIX

The first term of Eq. (53) may be expanded to give the Landau collision term on a first approximation under certain conditions. It is essentially the normal collision term with the divergent portion subtracted. We have shown that a necessary condition for the expansion of this normal term to give the Landau term for this model is that the relaxation

time $\omega_p^{-1}\Gamma^{-\frac{1}{3}}$ be much greater than some small fraction of the stability time $(\bar{\gamma}\bar{l})^{-1}$.

To expand the normal collision integral in the ring approximation for the symmetric model, we follow the procedure used by Balescu.[6] The normal collision term, $C(\varphi)$ can be rewritten:

$$C(\varphi) = A \int d\mathbf{v}'\, \partial_r R_{rs}(\partial_s - \partial_s')\varphi\varphi'. \tag{A1}$$

A is a constant and the tensor R_{rs} is given by

$$R_{rs}(\mathbf{v};\mathbf{v}') = \frac{1}{4\pi^4}\int d\mathbf{l}\,\frac{l_r l_s}{l^4}\,\frac{\delta[\mathbf{l}\cdot(\mathbf{v}-\mathbf{v}')]}{|1 + l^{-2}\Psi(\mathbf{l}\cdot\mathbf{v}/l)|^2} \tag{A2}$$

for Ψ defined by

$$\Psi = \left(\frac{i\omega_p^2}{\pi}\right)\int d\nu_1\,\delta_-\left(\frac{\mathbf{l}\cdot\mathbf{v}}{l} - \nu_1\right)\bar{\varphi}'(\nu_1)$$

$$= \Psi_1 + i\Psi_2, \tag{A3}$$

and Ψ_1, Ψ_2 are real. If Eq. (A2) is integrated over the direction cosine, $\cos\theta = \mathbf{g}\cdot\mathbf{l}/gl$, then R_{rs} simplifies to give

$$R_{rs} = \frac{1}{4\pi^4\,|\mathbf{v}-\mathbf{v}'|}\int_0^{2\pi} d\Phi \begin{Bmatrix}\cos\Phi\\ \sin\Phi\\ 0\end{Bmatrix}\begin{Bmatrix}\cos\Phi\\ \sin\Phi\\ 0\end{Bmatrix} J, \tag{A4}$$

where

$$J = \frac{1}{2}\ln\left[\frac{(\Psi_1 + L^2)^2 + \Psi_2^2}{\Psi_1^2 + \Psi_2^2}\right]$$

$$- \frac{\Psi_1}{\Psi_2}\tan^{-1}\left(\frac{L^2\Psi_2}{L^2\Psi_1 + \Psi_1^2 + \Psi_2^2}\right) - \frac{n\pi\Psi_1}{\Psi_2} \tag{A5}$$

and $n = 0$;

$$L^2\Psi_1 + |\Psi|^2 > 0, \qquad L^2\Psi_1 + |\Psi|^2 < 0.$$

Previously the logarithmic term of J was expanded using the condition that L the inverse impact parameter, is large compared to λ_D, the Debye length. Another equivalent condition is that $\Gamma^{\frac{1}{3}}$, the inverse number of particles in the Debye sphere, be small. The critical requirement was

$$|\psi| = |\Psi|/\beta e^2 c = O(1). \tag{A6}$$

From Balescu[6] we have that ψ_1 and ψ_2 for the Lorentz distribution are

$$\psi_1 = \frac{\omega_p^2}{2K_D^2}\left\{\frac{u_0^2 - (\nu-\nu_0)^2}{[(\nu-\nu_0)^2 + u_0^2]} + \frac{u_0^2 - (\nu+\nu_0)^2}{[(\nu+\nu_0)^2 + u_0^2]}\right\}, \tag{A7}$$

$$\psi_2 = \frac{u_0\omega_p^2}{K_D^2}\left\{\frac{(\nu-\nu_0)}{[(\nu-\nu_0)^2 + u_0^2]^2} + \frac{(\nu+\nu_0)}{[(\nu+\nu_0)^2 + u_0^2]^2}\right\}. \tag{A8}$$

Of course in the limit of extremely large ν, both ψ_1 and ψ_2 approach zero as before. By inspection we see that ψ_1 and ψ_2 also vanish for

$$\nu = 0 \quad \text{and} \quad \nu_0 = u_0. \tag{A9}$$

As the function ψ_2 vanishes only for $\nu = 0$, the resultant restriction on the expansion is that

$$|u_0 - \nu_0| > \delta u_0, \tag{A10}$$

where δ is some small number less than 1 to be determined. To find the value of the logarithm for $\nu = 0$, we note that

$$\psi = \psi_1 \approx \tfrac{1}{2}\delta \tag{A11}$$

on the logarithm expanded gives

$$\ln(1 + 8\Gamma^{-3}\delta^{-1}) \approx \ln 8\Gamma^{-3} + \ln\delta^{-1} \tag{A12}$$

To neglect the second term of Eq. (A12) with respect to the first we must require that

$$8\Gamma^{-3} \gg \delta^{-1}. \tag{A13}$$

The constant δ can be related to $\bar{\gamma}$, Eq. (22), since it was stated earlier that all terms for which

$$|u_0 - \nu_0| < \bar{\gamma} \tag{A14}$$

be excluded from integration of the normal collision term. Comparing Eqs. (A10) and (A14) gives

$$\delta = \bar{\gamma}/u_0. \tag{A15}$$

Substituting Eq. (A15) into Eq. (A13) gives finally

$$(8\Gamma^{-3})(\bar{\gamma}u_0^{-1}) \gg 1. \tag{A16}$$

While u_0 varies with time, reaching its maximum value at equilibrium it varies at the most for this model by a factor of $\sqrt{2}$. Setting u_0 equal $\omega_p\lambda_D$, we find

$$8(\Gamma^3\omega_p)^{-1} \gg (\bar{\gamma}K_D)^{-1}. \tag{A17}$$

The term $(\Gamma^3\omega_p)^{-1}$ is the relaxation time of the plasma. The stability time is less obvious in its definition, that the time of the largest time-dependent component to decay shall be called the stability time:

$$(\epsilon^{\frac{1}{2}}K_D\bar{\gamma})^{-1} \equiv t_s. \tag{A18}$$

Then Eq. (A16) can be rewritten

$$8t_r \gg \epsilon^{\frac{1}{2}}t_s. \tag{A19}$$

Thus the expansion of the normal collision integral for the unstable case, giving the Landau approximation as the first term, is valid if the relaxation and stability time obey the above inequality.

Upper Bound for the Double Spectral Function in Potential Scattering

D. BESSIS
Centre d'Etudes Nucléaires de Saclay, Service de Physique Théorique, Gif-sur-Yvette, S.-et-O., France
(Received 9 June 1964)

For potentials $V(z)$ holomorphic in $\mathrm{Re}\, z > 0$ and bounded by

$$|V(z)| < K/|z|^{\rho} \quad \rho < 2 \quad \text{for} \quad |z| \leq 1,$$
$$|V(z)| < K\,|z|^{-\gamma} e^{-\mu_0 \mathrm{Re}\, z} \quad \text{for} \quad |z| \geq 1, \quad \gamma > 7/4,$$

we show that the double spectral "function" $\rho(s, t)$ is a continuous function of s and t in $s > 0$, $t > 0$, and we obtain an upper bound for it. This upper bound shows clearly that the double integral of the Mandelstam representation in fact exists and defines an analytic function of s and t in two cut planes. We indicate how to generalize these results to the case when $\rho(s, t)$ is no longer a function but a distribution.

1. INTRODUCTION

THE Mandelstam representation for the amplitude in potential scattering has been written down by Blankenbecler, Goldberger, Khuri, and Treiman[1]:

$$f(s, t) = f_B(t) + \sum_{i=0}^{i=r} \frac{\Gamma_i(t)}{s + s_i} + \frac{t^{n+1}}{\pi^2} \int_0^{\infty} ds' \int_0^{\infty} dt' \frac{\rho(s', t')}{t'^{n+1}(t' - t)(s' - s)} + \sum_{k=0}^{k=n} \frac{t^k}{k!} \int_0^{\infty} \frac{g_k(s')\, ds'}{s' - s}. \tag{1.1}$$

Here we take $t = -2s(1 - \cos\theta)$; s = energy, $\cos\theta$ = scattering angle. In fact what one knows is that one can write from Khuri's work

$$f(s, t) = f_B(t) + \sum_{i=0}^{r} \frac{\Gamma_i(t)}{s + s_i} + \frac{1}{\pi} \int_0^{\infty} ds' \frac{\mathrm{Im}\, f(s', t)}{s' - s}, \tag{1.2}$$

and from Regge's work

$$\mathrm{Im}\, f(s', t) = \frac{t^{n+1}}{\pi} \int_0^{\infty} \frac{dt'}{t'^{n+1}(t' - t)} \rho(s', t') + \sum_{k=0}^{k=n} \frac{t^k}{k!} g_k(s'), \quad \text{for} \quad s' \text{ real} > 0. \tag{1.3}$$

The problem in which we are interested here is to show that the repeated integrals

$$\frac{1}{\pi} \int_0^{\infty} \frac{ds'}{s' - s} \left\{ \frac{t^{n+1}}{\pi} \int_0^{\infty} \frac{dt'}{t'^{n+1}} \frac{\rho(s', t')}{(t' - t)} \right.$$

converge and are in fact a double integral. We restrict ourselves in this paper to such potentials that give rise to a continuous $\rho(s, t)$. In this simple case, we just need to put a convenient upper bound on $\rho(s, t)$ to demonstrate the proposition. We indicate a general method for the case where $\rho(s, t)$ is a distribution. We do not consider the convergence of the one-dimensional integrals, (see, for instance, Ref. 2).

2. AN UPPER BOUND FOR THE AMPLITUDE $f(\nu, k)$

We consider a potential $V(z)$ holomorphic in $\mathrm{Re}\, z > 0$ and bounded by

$$\left.\begin{aligned} &|V(z)| < K/|z|^{\rho} \\ &|z| \leq 1 \end{aligned}\right\} \qquad \rho < 2, \tag{2.1}$$

$$\left.\begin{aligned} &|V(z)| < K\, e^{-\mu_0 \mathrm{Re} z}/|z|^{\gamma} \\ &|z| \geq 1 \end{aligned}\right\} \qquad \gamma > 0, \quad \mu_0 > 0. \tag{2.2}$$

It has been shown in Ref. 3 that, for $|\nu| \geq 1$, $0 \leq \arg\nu < \pi/2$, $-(\pi/2 - \arg\nu) < \arg k < +\pi/2$. The amplitude $f(\nu, k) \equiv e^{i\delta(\nu,k)} \sin\delta(\nu, k)$ is given by

$$f(\nu, k) = -A(\nu, k)/[1 + iB(\nu, k)] \tag{2.3}$$

with

$$|A(\nu, k)|, \quad |B(\nu, k)| < \tfrac{1}{2}\{e^{2C^2 I(\nu,k)} - 1\}, \tag{2.4}$$

with C a purely numerical constant ($C < 23$). $I(\nu, k)$ is given by the integral

$$I(\nu, k) = \frac{1}{|\nu|} \left|\frac{\nu}{k}\right|^2 \int_{\Gamma} \left| \frac{t\, dt}{(1 - t^2)^{\frac{1}{2}}} V\left(\frac{\nu t}{k}\right) \right|, \tag{2.5}$$

where Γ is a contour depending only on $\arg\nu$, made of two pieces: a rectilinear part from 0 to 1 and a curve Γ' ($\arg\nu$) shown in Fig. 1.

We need first to study $I(\nu, k)$. We suppose k real positive but most of the results extend to

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

[2] A. Martin, Institute for Advanced Study, Princeton, preprint (1964), p. 46.

[3] D. Bessis, Nuovo Cimento **33**, 797 (1964).

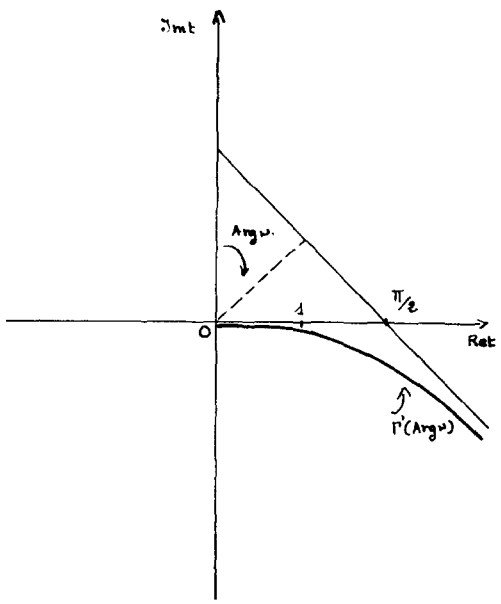

FIG. 1. The path of integration Γ.

$-(\pi/2 - \arg \nu) < \arg k < +\pi/2$. We consider two cases.

A. The Case where $|\nu/k| \geq 1$

We put

$$I_1(\nu, k) = \frac{1}{|\nu|}\left|\frac{\nu}{k}\right|^2 \int_0^1 \frac{t\,dt}{(1+t^2)^{\frac{1}{2}}}\left|V\left(\frac{\nu t}{k}\right)\right|;$$
$$I_2(\nu, k) = \frac{1}{|\nu|}\left|\frac{\nu}{k}\right|^2 \int_{\Gamma'}\left|\frac{t\,dt}{(1+t^2)^{\frac{1}{2}}}\,V\left(\frac{\nu t}{k}\right)\right|. \tag{2.6}$$

We have

$$I_1(\nu, k) \leq \frac{K}{|\nu|}\left|\frac{\nu}{k}\right|^2 \left\{\left|\frac{k}{\nu}\right|^\rho \int_0^{|k/\nu|} \frac{t\,dt}{(1-t^2)^{\frac{1}{2}}}\frac{1}{t^\rho} + \left|\frac{k}{\nu}\right|^\gamma \int_{|k/\nu|}^1 \frac{t\,dt}{(1-t^2)^{\frac{1}{2}}}\frac{e^{-\mu_0(\mathrm{Re}\,\nu/k)\,t}}{t^\gamma}\right\}. \tag{2.7}$$

Putting $|k/\nu| = \sin\alpha$, $0 \leq \alpha \leq +\pi/2$ and recalling that $2u/\pi \leq \sin u \leq u$ for $0 \leq u \leq +\pi/2$, we see that

$$\int_0^{|k/\nu|} \frac{t\,dt}{(1-t^2)^{\frac{1}{2}}}\frac{1}{t^\rho} = \int_0^\alpha \frac{du}{(\sin u)^{\rho-1}} \leq \left(\frac{\pi}{2}\right)^{3-\rho}\frac{1}{2-\rho}\left|\frac{k}{\nu}\right|^{2-\rho}, \quad \text{for } \rho < 2.$$

Putting $\theta = \arg \nu/k$, $\cos\theta \geq 0$, we have

$$\int_{|k/\nu|}^1 \frac{t\,dt}{(1-t^2)^{\frac{1}{2}}}\frac{e^{-\mu_0(\mathrm{Re}\,\nu/k)\,t}}{t^\gamma} = \int_\alpha^{\pi/2} \frac{du}{(\sin u)^{\gamma-1}} \exp\left(-\mu_0 \cos\theta \frac{\sin u}{\sin\alpha}\right)$$
$$\leq \left(\frac{\pi}{2}\right)^\gamma \int_{\sin\alpha}^{\pi/2} \frac{du}{u^{\gamma-1}} \exp\left(-\frac{2}{\pi}\mu_0 \frac{\cos\theta}{\sin\alpha}u\right)$$
$$\leq \left(\frac{\pi}{2}\right)^\gamma \left(\frac{\sin\alpha}{\cos\theta}\right)^{2-\gamma} \int_{\cos\theta}^{\pi\cos\theta/2\sin\alpha} \frac{dt}{t^{\gamma-1}} e^{-2\mu_0 t/\pi}$$
$$\leq \left(\frac{\pi}{2}\right)^\gamma \left(\frac{\sin\alpha}{\cos\theta}\right)^{2-\gamma} \int_{\cos\theta}^{+\infty} \frac{dt}{t^{\gamma-1}} e^{-2\mu_0 t/\pi}, \quad \gamma > 0.$$

If $0 < \gamma < 2$, this last integral is smaller than

$$\int_0^{+\infty} \frac{dt}{t^{\gamma-1}} e^{-2\mu_0 t/\pi} = \left(\frac{\pi}{2\mu_0}\right)^{2-\gamma}\Gamma(2-\gamma); \tag{2.8}$$

if $\gamma = 2$, this last integral is smaller than

$$\int_{\cos\theta}^{+\infty} \frac{dt}{t} e^{-2\mu_0 t/\pi} < \frac{\pi}{2\mu_0} e^{-2\mu_0/\pi} - \ln\cos\theta; \tag{2.9}$$

if $\gamma > 2$, this last integral is smaller than

$$\int_{\cos\theta}^{+\infty} \frac{dt}{t^{\gamma-1}} = \frac{(\cos\theta)^{2-\gamma}}{2-\gamma}. \tag{2.10}$$

And so,

for $0 < \gamma < 2$,

$$I_1(\nu, k) \leq \frac{K}{|\nu|}\left\{\left(\frac{\pi}{2}\right)^{3-\rho}\frac{1}{2-\rho} + \left(\frac{\pi}{2}\right)^2 \frac{\Gamma(2-\gamma)}{\mu_0^{2-\gamma}}\frac{1}{(\cos\theta)^{2-\gamma}}\right\};$$

for $\gamma = 2$,

$$I_1(\nu, k) \leq \frac{K}{|\nu|}\left\{\left(\frac{\pi}{2}\right)^{3-\rho}\frac{1}{2-\rho} + \left(\frac{\pi}{2}\right)^2\left[\frac{\pi}{2\mu_0}e^{-2\mu_0/\pi} - \ln\cos\theta\right]\right\};$$

for $\gamma > 2$,

$$I_1(\nu, k) \leq \frac{K}{|\nu|}\left\{\left(\frac{\pi}{2}\right)^{3-\rho}\frac{1}{2-\rho} + \left(\frac{\pi}{2}\right)^\gamma \frac{1}{2-\gamma}\right\}.$$

We have to obtain now an upper bound for $I_2(\nu, k)$. We write

$$I_2(\nu, k) \leq \frac{K}{|\nu|}\left|\frac{\nu}{k}\right|^{2-\gamma} \int_{\Gamma'} \frac{|t|^{1-\gamma}}{|1-t^2|^{\frac{1}{2}}} e^{-\mu_0 \mathrm{Re}\,\nu t/k}\,|dt|.$$

Taking into account $|\nu/k| \geq 1$ and that it has been shown in Appendix II of Ref. 3 that on Γ' (arg ν) we have $|t| \geq 1$ and $0 \leq \arg\nu(t-1) \leq \frac{1}{3}\arg\nu$. We get, k being real here for simplicity,

$$I_2(\nu, k) \leq \frac{K}{|\nu|}\left|\frac{\nu}{k}\right|^{2-\gamma} e^{-\mu_0 \mathrm{Re}\,\nu/k} \int_{\Gamma'} \frac{|t|}{|1-t^2|^{\frac{1}{2}}}\frac{|dt|}{|t|^\gamma} \times \exp\left\{-\mu_0\left|\frac{\nu}{k}\right||t-1|\,[\arg\nu(t-1)]\right\}$$
$$\leq \frac{K}{|\nu|}\left|\frac{\nu}{k}\right|^{2-\gamma} e^{-\mu_0 \mathrm{Re}\,\nu/k} \int_{\Gamma'} \frac{|t|}{|1-t^2|^{\frac{1}{2}}}\frac{|dt|}{|t|^\gamma} e^{-\mu_0|t-1|\cos\pi/6}, \quad \gamma > 0. \tag{2.10}$$

This last integral is always convergent, and being a continuous function of arg ν varying over the finite interval $(0, \pi/2)$, it has an upper bound $K_0(\mu_0)$ for $\gamma > 0$, and so

$$I_2(\nu, k) \leq KK_0(\mu_0)/|\nu|\ |\nu/k|^{2-\gamma}\ e^{-\mu_0 \mathrm{Re}\,\nu/k};$$

if $\gamma \geq 2$,

$$I_2(\nu, k) \leq KK_0(\mu_0)/|\nu|;$$

if $0 < \gamma < 2$,

$$I_2(\nu, k) \leq [KK_0(\mu_0)/\mu_0^{2-\gamma}\ |\nu|]1/(\cos\theta)^{2-\gamma}.$$

Summing up all results we find, calling $C(V)$ a constant depending *only* on potential properties,

$$I(\nu, k) \leq [C(V)/|\nu|]1/(\cos\theta)^{2-\gamma},$$

$$\text{for } 0 < \gamma < 2; \quad (2.11)$$

$$\leq C(V)/|\nu|, \quad \text{for } \gamma > 2; \quad (2.12)$$

$$\leq (K/|\nu|)[C(V) + (\pi/2)^2\ |\ln\cos\theta|],$$

$$\text{for } \gamma = 2. \quad (2.13)$$

B. The Case where $|\nu/k| \leq 1$

This case is harmless and following the same method, we show that

$$I(\nu, k) \leq [C(V)/|\nu|]|\nu/k| \leq C(V)/|\nu|. \quad (2.14)$$

Finally, we have proved that the Regge poles for k real > 0 are confined to the domain

(I) for $0 < \gamma \leq 2$ see Fig. 2 for the shape of the domain.

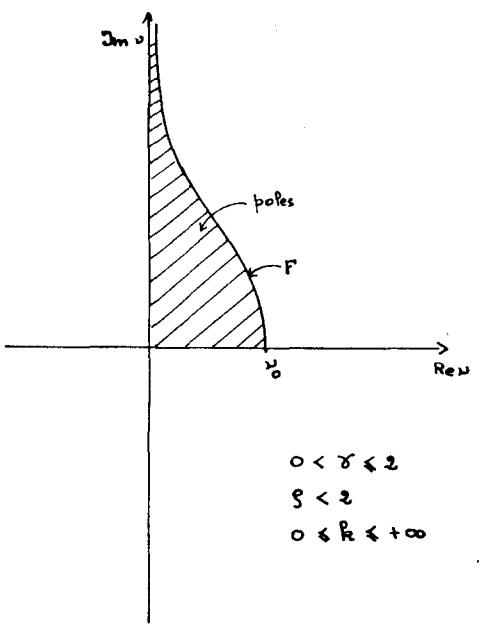

FIG. 2. Localization of Regge poles.

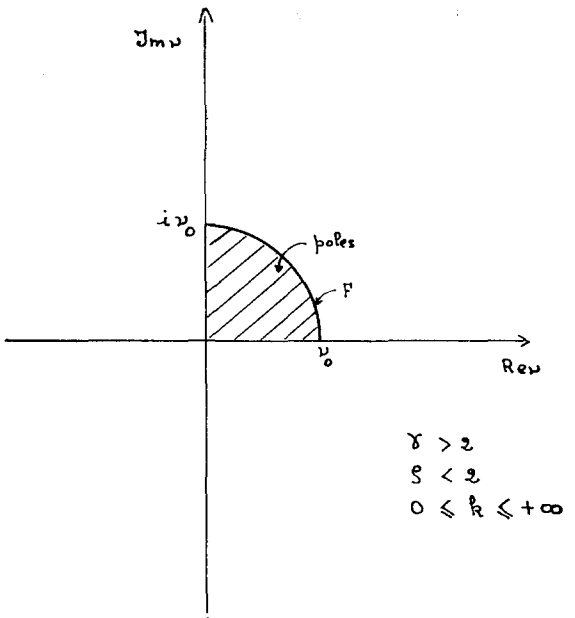

FIG. 3. Localization of Regge poles.

(II) for $\gamma > 2$ see Fig. 3 for the shape of the domain.

This results from the fact that there cannot be poles when

$$I(\nu, k) < \ln 3/C^2.$$

The point ν_0 (of Fig. 2 and Fig. 3) being independent of ν and k but depending on potential properties.

We want now to find an upper bound for $f(\nu, k)$ on a contour C defined as follows (see Fig. 4).

C is made of the line $(\epsilon - i\ \infty, P^*)$, the arc of circle P^*L_0P of center O, radius L_0 and the line $(P, \epsilon + i\infty)$. We impose $0 < \epsilon < \frac{1}{2}$ and L_0 is an integer. The line F containing all Regge poles must be completely inside the domain limited by the imaginary ν axis and C as shown in Fig. 4. It is always possible to choose a finite L_0 for the potential we are dealing with, as results from the preceding study. We suppose, as can always be done, $L_0 > 1$. On $(P, \epsilon + i\ \infty)$ we have

$$I(\nu, k) < C(V)/|\nu| \quad \text{for } \gamma > 2;$$

$$I(\nu, k) < [C(V)/\epsilon^{2-\gamma}]1/|\nu|^{\gamma-1} \quad \text{for } 0 < \gamma < 2;$$

$$I(\nu, k) < \frac{K}{|\nu|}\left[C(V) + \left(\frac{\pi}{2}\right)^2 \left|\log\frac{\epsilon}{|\nu|}\right|\right], \text{ for } \gamma = 2,$$

and because $f(\nu, k) = -A(\nu,k)\ /[1 + iB(\nu, k)]$ and $|B(\nu, k)| < 1$ on C, we see that

$$|f(\nu, k)| < C(V)/|\nu| \quad \text{for } \gamma > 2; \quad (2.15)$$

$$|f(\nu, k)| < [C(V)/\epsilon^{2-\gamma}](1/|\nu|^{\gamma-1})$$
$$\text{for } 0 < \gamma < 2; \qquad (2.16)$$

$$|f(\nu, k)| < \frac{K}{|\nu|}\left[C(V) + \left(\frac{\pi}{2}\right)^2 \left|\log \frac{\epsilon}{|\nu|}\right|\right]$$
$$\text{for } \gamma = 2. \qquad (2.17)$$

Through unitarity $S(\nu, k) \times S^*(\nu^*, k^*) = 1$ we see that these bounds hold on $(P^*, \epsilon - i\infty)$. On the arc of the circle, we need for $|A(\nu, k)|$ a better bound than $I(\nu, k)$ at least for $k \to +0$ in order to find the well known behavior $k^{2 \mathrm{Re}\, \nu}$. To do this we need to come back to the equations of Part III of Ref. 3, in which it has been shown that

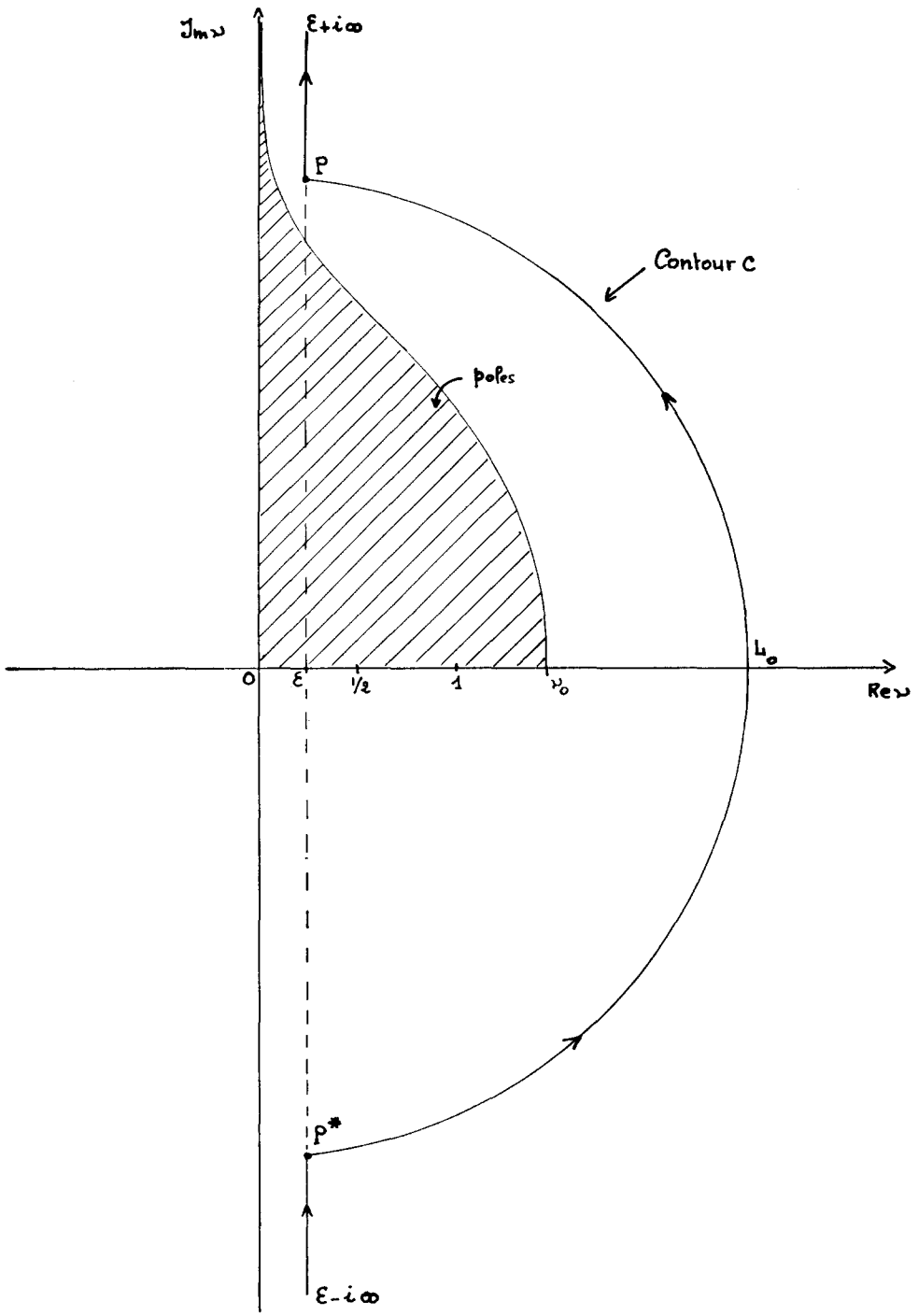

FIG. 4. Path of Integration.

$$|A(\nu, k)| \leq \frac{C^2}{|\nu|} e^{2C^2 I(\nu,k)} \left|\frac{\nu}{k}\right|^2 \times \left\{\int_0^{+\infty} e^{-2\,\mathrm{Re}\,\nu(\alpha-\tanh\alpha)} \left|V\left(\frac{\nu}{k}\frac{1}{\cosh\alpha}\right)\right| \frac{d\alpha}{\cosh^2\alpha} + \int_{\Gamma'(\arg\nu)} \frac{|t|}{|1-t^2|^{\frac{1}{2}}} \left|V\left(\frac{\nu}{k}t\right)\right| |dt|\right\}. \quad (2.18)$$

We shall suppose $|\nu/k| \geq 1$, because on the arc of the circle $|\nu|$ is fixed and equal to L_0, and we are only interested in $k < L_0$; for $k > L_0$ we use the previous bound.

So we get (putting $|\nu/k| = \cosh\beta$) that

$$\int_0^{\infty} e^{-2\,\mathrm{Re}\,\nu(\alpha-\tanh\alpha)} \left|V\left(\frac{\nu}{k}\frac{1}{\cosh\alpha}\right)\right| \frac{d\alpha}{\cosh^2\alpha} = \frac{1}{(\cosh\beta)^{\gamma}} \times \int_0^{\beta} e^{-2\,\mathrm{Re}\,\nu(\alpha-\tanh\alpha)} (\cosh\alpha)^{\gamma-2} e^{-\mu_0 \mathrm{Re}(\nu/k)(1/\tanh\alpha)}\, d\alpha + \frac{1}{(\cosh\beta)^{\rho}} \int_{\beta}^{+\infty} e^{-2\,\mathrm{Re}\,\nu(\alpha-\tanh\alpha)} (\cosh\alpha)^{\rho-2}\, d\alpha.$$

We have

$$\frac{1}{(\cosh\beta)^{\rho}} \int_{\beta}^{+\infty} e^{-2\,\mathrm{Re}\,\nu(\alpha-\tanh\alpha)} (\cosh\alpha)^{\rho-2}\, d\alpha \leq \frac{2^{2-\rho}\, e^{2\,\mathrm{Ri}\,\nu}}{2-\rho} \left|\frac{k}{\nu}\right|^{2\,\mathrm{Re}\,\nu+2}$$

where $\mathrm{Re}\,\nu > 0$ and $\rho < 2$,

and

$$J_2 = \int_0^{\beta} e^{-2\,\mathrm{Re}\,\nu(\alpha-\tanh\alpha)} (\cosh\alpha)^{\gamma-2} e^{-\mu_0\,\mathrm{Re}\,(\nu/k)(1/\cosh\alpha)}\, d\alpha \leq 2\, e^{2\,\mathrm{Re}\,\nu} \int_0^{\beta} e^{-\alpha(2\,\mathrm{Re}\,\nu+2-\gamma)-\mu_0(\mathrm{Re}\,\nu/|\nu|)}\, e^{\beta-\alpha}\, d\alpha.$$

Putting $t = \beta - \alpha$, we get

$$J_2 \leq 2\, e^{2\,\mathrm{Re}\,\nu-\beta(2\,\mathrm{Re}\,\nu+2-\gamma)} \times \int_0^{\beta} \exp\left[t(2\,\mathrm{Re}\,\nu+2-\gamma) - \frac{\mu_0}{2}\frac{\mathrm{Re}\,\nu}{|\nu|}e^{t}\right] dt \leq \frac{2\exp(2\,\mathrm{Re}\,\nu+|2\,\mathrm{Re}\,\nu+2-\gamma|\,\log 2)}{(\frac{1}{2}\mu_0\,\mathrm{Re}\,\nu/|\nu|)^{2\,\mathrm{Re}\,\nu+2-\gamma}} \left|\frac{k}{\nu}\right|^{2\,\mathrm{Re}\,\nu+2-\gamma} \times \int_{(\mu_0/2)\,\mathrm{Re}\,\nu/|\nu|}^{+\infty} e^{-u} u^{(2\,\mathrm{Re}\,\nu+2-\gamma)-1}\, du.$$

Recalling that $\epsilon \leq \mathrm{Re}\,\nu \leq L_0$ we see that

$$J_2 \leq C(V, \epsilon, L_0) \left|\frac{k}{\nu}\right|^{2\,\mathrm{Re}\,\nu+2-\gamma} \quad (2.19)$$

We consider now the integral

$$\int_{\Gamma'} \frac{|t|}{|1-t^2|^{\frac{1}{2}}} \left|V\left(\frac{\nu}{k}t\right)\right| |dt|$$

which is equal to

$$\left|\frac{k}{\nu}\right|^{\gamma} e^{-\mu_0\,\mathrm{Re}\,\nu t/k} \int_{\Gamma'} \frac{|t|}{|1-t^2|^{\frac{1}{2}}} \frac{|dt|}{|t|^{\gamma}} e^{-\mu_0\,\mathrm{Re}\,[\nu(t-1)]/k},$$

which is smaller than

$$K_0(\mu_0)\, |k/\nu|^{\gamma}\, e^{-\mu_0\,\mathrm{Re}\,\nu/k}. \quad (2.20)$$

Collecting results, we obtain

$$|A(\nu, k)| < (C(V, \epsilon)/|\nu|)\, |k/\nu|^{2\,\mathrm{Re}\,\nu} \quad (2.21)$$

and the amplitude $f(\nu, k)$ is, on the arc of the circle, and for $k < 1$, smaller than

$$|f(\nu, k)| \leq C(V, \epsilon) k^{2\,\mathrm{Re}\,\nu}, \qquad k < 1.$$

With all these bounds, we can now obtain a nice bound for the double spectral function, at least for $\gamma > \frac{7}{4}$. But we show that with much labor we can explore the structure of $\rho(s, t)$ even for $0 < \gamma < \frac{7}{4}$

3. THE DOUBLE SPECTRAL FUNCTION

For k real positive we can write

$$f(k, \cos\theta) = -\frac{1}{2k} \int_C \frac{f(\nu, k)}{\cos\pi\nu} P_{\nu-\frac{1}{2}}(-\cos\theta)\nu\, d\nu + \frac{1}{2ik} \sum_{l=0}^{l=L_0-1} (2l+1) f_l(k) P_l(\cos\theta) \quad (3.1)$$

with $f_l(k) \equiv f(l+\frac{1}{2}, k)$. This defines an analytic function of $\cos\theta$ in the $\cos\theta$ plane cut along $(1, +\infty)$.

Recalling $s = k^2$, $t = -2s(1-\cos\theta)$, we see that

$$2i\ \mathrm{Im}\, f(s, t) = f(k, t) - f^*(k, t^*) = \frac{-i}{k} \int_C \frac{f(\nu, k) - f^*(\nu^*, k)}{\cos\pi\nu} P_{\nu-\frac{1}{2}}\left(-1-\frac{t}{2s}\right)\nu\, d\nu + \frac{1}{k} \sum_{l=0}^{l=L_0-1} (2l+1)\ \mathrm{Im}\, s_l(k) P_l\left(1+\frac{t}{2s}\right). \quad (3.2)$$

Taking into account unitarity, that is

$$-i[f(\nu, k) - f^*(\nu^*, k)] = 2f(\nu, k) f^*(\nu^*, k), \quad (3.3)$$

we see that

$$2i\ \mathrm{Im}\, f(s, t) = \frac{2}{k} \int_C \frac{f(\nu, k) f^*(\nu^*, k)}{\cos\pi\nu} P_{\nu-\frac{1}{2}}\left(-1-\frac{t}{2s}\right)\nu\, d\nu + \frac{1}{k} \sum_{l=0}^{l=L_0-1} (2l+1)\ \mathrm{Im}\, f(k) P_l\left(1+\frac{t}{2s}\right). \quad (3.4)$$

This defines a function which is holomorphic in the t plane outside the cut $(0, +\infty)$. From the previous study, we see that, if $\gamma > \frac{7}{4}$,

$$|f(\nu, k)| < C/(|\nu|^{\frac{3}{4}} + n), \qquad n > 0, \quad (3.5)$$
$$|\mathrm{Im}\,\nu| \to \infty,$$

and using the bound

$$|P_{\epsilon-\frac{1}{2}+i\lambda}(\cosh\alpha)| \leq 3\cosh\alpha\epsilon/(\lambda\sinh\alpha)^{\frac{1}{2}} \text{ for } \lambda > 1$$
$$-\tfrac{1}{2} < \epsilon < +\tfrac{1}{2}, \qquad \alpha > 0, \qquad (3.6)$$

we see that the integral converges absolutely on both sides of the t cut.

The double spectral function is by definition

$$\rho(s, t) = (1/2i)\{\text{Im } f(s, t + i\epsilon) - \text{Im } f(s, t - i\epsilon)\}, \qquad (3.7)$$

and we find, all integrals being absolutely convergent,

$$\rho(s, t) = \frac{-i}{k}\int_C f(\nu, k)f^*(\nu^*, k)P_{\nu-\frac{1}{2}}\left(1+\frac{t}{2s}\right)\nu\, d\nu, \qquad (3.8)$$

where $s > 0$, $t > 0$.

4. UPPER BOUND FOR $\rho(s, t)$

It is easy now to obtain for $\gamma > \frac{7}{4}$ a nice upper bound for $\rho(s, t)$, from which the actual existence of the Mandelstam representation follows without difficulties. From the preceding study we see that the contribution to the integral giving $\rho(s, t)$, coming from $(P, \epsilon + i\infty)$ and $(P^*, \epsilon - i\infty)$ is, in modulus, less than

$$\rho^{\text{I}}(s, t) = \frac{2C(V, \epsilon)}{k}\int_1^{+\infty}\frac{1}{\lambda^{2\delta-1}}\left|P_{\epsilon-\frac{1}{2}+i\lambda}\left(1+\frac{t}{2s}\right)\right| dt$$
$$\text{if} \quad |f(\nu, k)| < \frac{C(\nu, \epsilon)}{|\nu|^{\delta}}. \qquad (4.1)$$

In the case $\gamma > \frac{7}{4}$, all the operations we have made in Ref. 3 are valid and even more the integral giving $\rho(s, t)$ is absolutely convergent and so defines a continuous function of s and t. This is so because

$$|P_{\epsilon-\frac{1}{2}+i\lambda}(\cosh\alpha)| \leq 3\cosh(\alpha\epsilon)/(\lambda\sinh\alpha)^{\frac{1}{2}}$$
$$\text{for } \lambda > 1, \quad -\tfrac{1}{2} < \epsilon < \tfrac{1}{2}, \quad \alpha \geq 0. \qquad (4.2)$$

So the previous integral is smaller than

$$C(V, \epsilon)\frac{\cosh\epsilon\alpha}{k(\sinh\alpha)^{\frac{1}{2}}}\int_1^{\infty}\frac{dt}{\lambda^{2\delta-\frac{1}{2}}}.$$

Putting $\cosh\alpha = 1 + t/2s$, the integral converges if $\delta > \frac{3}{4}$, that is, from the preceding study, if $\gamma > \frac{7}{4}$. So we have

$$\rho^{\text{I}}(s, t) \leq C(V, \epsilon)\frac{\cosh\epsilon\alpha}{s^{\frac{1}{2}}(\sinh\alpha)^{\frac{1}{2}}} \leq \frac{C(V, \epsilon)}{t^{\frac{1}{2}}(s + t/4)^{\frac{1}{2}}}\cosh\epsilon\alpha$$
$$\leq \frac{2^{\epsilon}C(V, \epsilon)}{t^{\frac{1}{2}}(s + t/4)^{\frac{1}{2}}}(\cosh\alpha)^{\epsilon} \leq C(V, \epsilon)\frac{(1 + t/2s)^{\epsilon}}{t^{\frac{1}{2}}(s + t/4)^{\frac{1}{2}}}.$$

Taking $\epsilon = \frac{1}{4}$, we get

$$\rho^{\text{I}}(s, t) \leq C(V)/s^{\frac{1}{2}}t^{\frac{1}{2}}. \qquad (4.3)$$

We are left with the integral on the arc of the circle. This integral is in modulus less than

$$\rho^{\text{II}}(s, t) \leq \frac{2C}{s^{\frac{1}{2}}}\int_0^{\pi/2-\epsilon_1}\frac{1}{L_0^{2\delta-1}}|P_{L_0 e^{i\phi}-\frac{1}{2}}(\cosh\alpha)|\, d\phi$$
$$\leq C(V)\, e^{\alpha(2\delta-\frac{1}{2})}/s^{\frac{1}{2}}$$

taking into account $|P_\nu(\cosh\alpha)| \leq e^{\alpha\,\text{Re}\,\nu}$

$$\leq [C(V)/s^{\frac{1}{2}}](1 + t/2s)^{L_0-\frac{1}{2}}; \qquad (4.4)$$

this bound is convenient for $s \geq 1$. For $s \leq 1$, instead of taking the bound $|f(\nu, k)| < C/(L_0)^{\nu}$, we take the more-refined bound obtained previously $|f(\nu, k)| < C(V)k^{2\text{Re}\,\nu}$, and get

$$\rho^{\text{II}}(s, t)$$
$$\leq \frac{C(V)}{s^{\frac{1}{2}}}\int_0^{\pi/2-\epsilon_1} s^{2L_0\cos\phi}|P_{L_0\cos\phi-\frac{1}{2}+iL_0\sin\phi}(\cosh\alpha)|\, d\phi$$
$$\leq \frac{C(V)}{s^{\frac{1}{2}}}e^{-\alpha/2}\int_0^{\pi/2-\epsilon_1} e^{H\cos\phi}\, d\phi, \quad H = L_0(\alpha + \log s^2).$$

But

$$H \leq L_0\log[1 + t/2 + (t + t^2/4)^{\frac{1}{2}}] \quad \text{for} \quad s \leq 1,$$

and so

$$\rho^{\text{II}}(s, t) < \frac{C(V)}{s^{\frac{1}{2}}}[1 + t/2 + (t + t^2/4)^{\frac{1}{2}}]^{L_0}, \quad s \geq 1. \qquad (4.5)$$

Adding up results we get

$$|\rho(s, t)| \leq C(V)\left\{\frac{1}{s^{\frac{1}{2}}t^{\frac{1}{2}}} + \frac{1}{s^{\frac{1}{2}}}\left(1 + \frac{t}{2s}\right)^{L_0-\frac{1}{2}}\right\} \qquad s \geq 1$$
$$\leq C(V)\{1/s^{\frac{1}{2}}t^{\frac{1}{2}} + (1/s^{\frac{1}{2}})(2 + t)^{L_0}\} \qquad s \leq 1. \qquad (4.6)$$

So in all cases

$$|\rho(s, t) \leq C(V)[1/s^{\frac{1}{2}}t^{\frac{1}{2}} + (2 + t)^{L_0}/s^{\frac{1}{2}}]. \qquad (4.7)$$

5. THE MANDELSTAM REPRESENTATION

We can now write for Im $f(s, t)$ the dispersion relation

$$\text{Im } f(s', t) = \frac{t^{L_0+1}}{\pi}\int_{4\mu_0^2}^{\infty} dt'\, \frac{\rho(s', t')}{t'^{L_0+1}(t' - t)}$$
$$+ \sum_{l=0}^{l=L_0}\frac{t^l}{l!}g_l(s'), \qquad s' \text{ real} > 0. \qquad (5.1)$$

From the result of Khuri, we know that

$$f(s, t) = f_B(t) + \sum_{i=0}^{r}\frac{\Gamma_i(t)}{s + s_i} + \frac{1}{\pi}\int_0^{\infty} ds'\, \frac{\text{Im } f(s', t')}{s' - s},$$
$$\text{Im } s \neq 0, \qquad (5.2)$$

which is valid for real $t \geq -4\mu_0^2$. Introducing the expression for Im $f(s', t)$, we see that

$$f(s, t) = f_B(t) + \sum_{i=0}^{r} \frac{\Gamma_i(t)}{s + s_i} + \frac{t^{L_0+1}}{\pi^2} \int_0^\infty \frac{ds'}{s' - s} \left\{ \int_{4\mu_0^2}^\infty dt' \frac{\rho(s', t')}{t'^{L_0+1}(t' - t)} \right\} + \sum_{l=0}^{l=L_0} \frac{t_l}{l!} \int_0^\infty \frac{ds'}{s' - s} g_l(s'). \tag{5.3}$$

Because of the bound obtained for $\rho(s, t)$ we see that the repeated integral is in fact a double integral and that the analytic continuation in t can be made because $f_B(t)$ is analytic in t outside the cut $\mu_0^2 + \infty$ and $\Gamma_i(t)$ are polynomials in t. For the convergence of the simple integrals see Martin (Ref. 2).

6. EXTENSION OF THE METHOD

To extend the method to the case $\gamma < \frac{7}{4}$ we need first a refinement of the method giving an upper bound on $f(\nu, k)$. This refinement consists essentially in extracting the Born approximation, and using the second iteration for the Volterra form of the Schrödinger radial equation. In that case, the kernel does not change but we improve the behavior of the inhomogeneous terms for $\nu \to i\infty$. This operation can be repeated if necessary, until the rest of the amplitude behaves better than $C/(|\nu| + \eta)$ for $|\text{Im } \nu| \to \infty$. Secondly, generally speaking, $\rho(s, t)$ will be a distribution; what kind of distributions shall we incounter in $\rho(s, t)$? Going back to

$$\rho(s, t) = -\frac{i}{k} \int_C f(\nu, k) f^*(\nu^*, k) P_{\nu-\frac{1}{2}}\left(1 + \frac{t}{2s}\right) \nu \, d\nu,$$

and taking into account that[4]

$$P_{-\frac{1}{2}+i\lambda}(\cosh \alpha) \sim \left(\frac{2}{\pi\lambda \sinh \alpha}\right)^{\frac{1}{2}} \left[\cos\left(\lambda\alpha - \frac{\pi}{4}\right) - \frac{1}{8\lambda} \coth \alpha \cos\left(\lambda\alpha + \frac{\pi}{4}\right) + O\left(\frac{1}{\lambda^2}\right), \right. \tag{6.1}$$

we can expect to find a finite number of distributions in $\rho(s, t)$. These distributions can be extracted from $\rho(s, t)$, the rest being continuous. It would be interesting to actually do this extraction and evaluate the "residues" of the distributions. The Mandelstam representation will then follow.

7. CONCLUSION

We have shown that, for potentials $V(z)$ holomorphic in Re $z > 0$, less singular than $1/z^2$ at the origin, decreasing exponentially on any ray of the right-hand plane (except on the imaginary axis, where they must decrease faster than $1/|z|^\gamma$; $\gamma > \frac{7}{4}$), the double spectral function is bounded by

$$|\rho(s, t)| \leq C(V) \left[\frac{1}{s^{\frac{1}{4}} t^{\frac{1}{4}}} + \frac{(2 + t)^{L_0}}{s^{\frac{1}{4}}} \right] \quad \text{for} \quad \begin{matrix} s > 0, \\ t > 0, \end{matrix}$$

where $C(V)$ depends only on potential properties as does L_0, and both are finite. This permitted us to show that the repeated integral of the Mandelstam representation actually exists and that it is in fact a double integral. Indications were also given to study the case $\gamma < \frac{7}{4}$.

ACKNOWLEDGMENT

It is a pleasure to thank Professor Froissart for many helpful discussions.

[4] L. Robin, *Fonctions sphériques de Legendre et fonctions sphéroïdales* (Gauthier-Villars, Paris, 1959), Vol. III, p. 157.

Uniqueness of Steady-State Solutions to the Fokker–Planck Equation*

A. H. Gray, Jr.
School of Engineering, University of California at Santa Barbara, Santa Barbara, California
(Received 19 May 1964)

The uniqueness of steady-state probability densities in certain continuous Markov processes is proven when the Fokker–Planck or Kolmogorov equation satisfied by the transition probability is of a variety called "steady." The tendency of other probability densities to approach such steady-state densities is formally demonstrated.

INTRODUCTION

SINCE the first treatments of Brownian motion as an example of a continuous Markov process, the applications of Markov processes in physical situations have extended over a wide range which includes such extremes as barometric pressure distributions and structural responses to earthquakes.

The analysis of such problems usually leads to a partial differential equation describing a "transition" probability density, which is called a Fokker–Planck or Kolmogorov equation. Too often such equations are too complex to solve. Sometimes a time-independent solution can be found, and it is called a steady-state probability density. It is then often tacitly assumed that such a steady-state probability density is unique and further that all solutions to the equation must approach this as time goes to infinity. Such assumptions are generally based on physical reasoning. Wang and Uhlenbeck[1] refer to a uniqueness proof for the steady-state probability density in a barometric distribution which was communicated to them by M. Dresden, but this proof does not appear to have been published.

It is proved in this paper that a certain class of Fokker–Planck or Kolmogorov equations do indeed have unique steady-state solutions (if they have any) and it is formally demonstrated that all solutions must "tend" towards the steady-state solutions.

THE FOKKER–PLANCK EQUATION

Let summation convention be implied unless otherwise specified, so that repeated subscripts imply a summation over all values of the subscript. The Fokker–Planck operator will be denoted by the symbol F, where

$$F(f) = -\partial(A_k f)/\partial y_k + \tfrac{1}{2}\partial^2(D_{ki}f)/\partial y_k\,\partial y_i, \qquad (1)$$

and the A_k and D_{ki} are each functions of y (that is, they are functions of the variables $y_1, y_2, \cdots, y_N$). The D_{ki} are elements of a positive-definite matrix so that the quadratic expression $D_{ki}x_ix_k$ is nonnegative. A function P of y and t will be said to satisfy the Fokker–Planck equation if

$$\partial P/\partial t = F(P). \qquad (2)$$

For simplicity, integrals denoted as single integrals in the variable y will be used as a shorthand notation for the N-dimensional improper Riemann integrals over the variables $y_1, y_2, \cdots, y_N$. That is,

$$\int f\,dy = \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f\,dy_1\,dy_2\cdots dy_N. \qquad (3)$$

As in all uniqueness proofs, there are certain restrictions that must be placed on the class of functions involved. To avoid continual repetition of these requirements in the pages to follow, these restrictions will be combined in a single definition. For lack of a better name, a function which satisfies these restrictions will be called "well behaved."

Definition 1. A probability density P is *well behaved* if and only if each of the following requirements are satisfied[2]:

(i) $\partial P/\partial t = F(P)$ for all $t > 0$ and all y.

(ii) $P > 0$ for all $t > 0$ and all finite y.

(iii) the multiple improper Riemann integrals over all y of each of the following functions are each absolutely and uniformly convergent for t in every positive closed interval: P, $A_k\partial P/\partial y_k$, $D_{ki}\partial^2 P/\partial y_i\partial y_k$, $P\partial A_k/\partial y_k$, $P\partial^2 D_{ki}/\partial y_i\partial y_k$, $(\partial P/\partial y_i)(\partial D_{ki}/\partial y_k)$, and $(D_{ki}/P)(\partial P/\partial y_k)(\partial P/\partial y_i)$. No summation convention is to be implied in these functions.

(iv) At $t = 0$, $\int P\,dy$ equals one.

* This paper is based on a portion of a thesis submitted to the California Institute of Technology in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Engineering Science.

[1] M. C. Wang and G. E. Uhlenbeck, Rev. Mod. Phys. **17**, 323 (1945). Reprinted in *Selected Papers on Noise and Stochastic Processes*, edited by N. Wax (Dover Publications, Inc., New York, 1954).

[2] Slight modifications are necessary in the theorems to follow when one is dealing with problems for which P is identically zero in portions of the N-dimensional phase space.

(v) $\partial P/\partial t$ is uniformly continuous in y and t for t in every positive closed interval, and each y_k in every closed interval.

(vi) The limit as y_k goes to plus or minus infinity of each of the following functions (summation convention is not implied here) exists and is zero: A_kP, $D_{ki}\partial P/\partial y_i$, and $P\partial D_{ki}/\partial y_i$.

It may be noted that well-behaved probability densities are not restricted to the so-called transition probabilities. In general the transistion probability is a well-behaved probability density with impulsive initial conditions.

Two more definitions will be useful. These are given here.

Definition 2. A *steady* Fokker–Planck equation is one for which the coefficients A_k and D_{ki} satisfy the following requirements:

(i) The N by N matrix made up of the elements D_{ki} contains an m by m strictly positive-definite matrix (where $1 \leq m \leq N$) such that the quadratic expression $D_{ki}x_ix_k$ is nonnegative, and can be zero if and only if each x_k, for $k \leq m$, is identically zero.

(ii) If $m < N$, then $D_{ki} = 0$ for $k > m$ or $i > m$, and further, the set of partial differential equations

$$\partial g/\partial y_k = 0 \quad \text{for} \quad k = m+1, m+2, \cdots N,$$

$$\partial g/\partial t = -A_k\, \partial g/\partial y_k,$$

has as its only solution g equal to a constant.

Definition 3. A *steady-state* probability density is any well-behaved probability density P, such that $\partial P/\partial t = 0$.

A MEASURE OF THE DIFFERENCE OF WELL-BEHAVED PROBABILITIES

Let P_1 and P_2 be two well-behaved probabilities. A measure of their difference will be denoted by $M(t)$, where

$$M(t) = \int \frac{(P_1 - P_2)^2}{P_1 + P_2}\, dy. \tag{4}$$

The integral defining $M(t)$ obviously exists, as $(P_1 - P_2)/(P_1 + P_2)$ must be a bounded continuous function. Thus, as P_1 and P_2 are each integrable, the products $P_1(P_1 - P_2)/(P_1 + P_2)$ and $P_2(P_1 - P_2)/(P_1 + P_2)$ must each be integrable, and the difference of the integrals is simply $M(t)$.

From the definition of $M(t)$, it is obvious that

$$M(t) \geq 0, \tag{5}$$

with equality holding if and only if $P_1 = P_2$. It will be shown here that under certain conditions $M(t)$ is a strictly monotonically decreasing function of time.

Theorem 1. Let P_1 and P_2 be well-behaved probability densities, and define $M(t)$ according to Eq. (4). Then the time derivative of $M(t)$ exists, is continuous, and

$$dM(t)/dt \leq 0. \tag{6}$$

Corollary. If the Fokker–Planck equation is steady, then equality can hold in Eq. (6) if and only if $P_1 = P_2$.

The proof of this theorem and its corollary can best be demonstrated by the use of some intermediate lemmas. The first of these is necessary if one is to even consider P_1 and P_2 as probability densities.

Lemma 1. If P is a well-behaved probability density, then

$$\int P\, dy = 1 \quad \text{for all} \quad t \geq 0.$$

The proof follows from Definition 1 for a well-behaved probability density. From items (i) and (iii) of Definition 1 it is seen that $\partial P/\partial t$ is integrable. From item (iii) it also follows that the order of integration is immaterial, so that

$$\int \frac{\partial P}{\partial t}\, dy = -\int \frac{\partial}{\partial y_k}(A_kP)\, dy + \frac{1}{2}\int \frac{\partial^2}{\partial y_k\, \partial y_i}(D_{ki}P)\, dy$$

may be evaluated by integrating each term with respect to y_k first. But from item (vi) of Definition 1, each term is zero so that one has $\int \partial P/\partial t\, dy = 0$. Further, from the uniform continuity of item (v), one finds that

$$\frac{d}{dt}\int P\, dy = \int \frac{\partial P}{\partial t}\, dy = 0.$$

Hence as the integral $\int P\, dy$ has a zero time derivative and is one at time zero, it must be one for all time.

Lemma 2. With $M(t)$ as defined by Eq. (4), one has that

$$\frac{dM(t)}{dt} = 4\int \frac{\partial}{\partial t}\left(\frac{P_1^2}{P_1 + P_2}\right) dy. \tag{7}$$

The proof of this follows by noting first that the integrand of Eq. (7) is equal to

$$\left[\frac{2P_1}{P_1 + P_2}\right]\frac{\partial P_1}{\partial t} - \left[\frac{P_1}{P_1 + P_2}\right]^2\left[\frac{\partial P_2}{\partial t} + \frac{\partial P_1}{\partial t}\right].$$

Thus, by noting that $P_1/(P_1 + P_2)$ is a bounded continuous function, and as in the proof of Lemma 1, $\partial P_1/\partial t$ and $\partial P_2/\partial t$ are each integrable and uniformly continuous, one has that the integral of Eq. (7) exists (and its integrand is uniformly continuous). As $M(t)$ is the integral of $(P_1 - P_2)^2/(P_1 + P_2)$ and

$$\frac{(P_1 - P_2)^2}{P_1 + P_2} = P_2 - 3P_1 + \frac{4P_1^2}{P_1 + P_2},$$

one can use Lemma 1 to show that

$$M(t) = -2 + 4\int \frac{P_1^2}{P_1 + P_2}\, dy.$$

By differentiating this expression, Eq. (7) of Lemma 2 follows.

Lemma 3. If P_1 and P_2 are well-behaved probability densities and the Fokker–Planck operator F is as defined by Eq. (1), then

$$\int F\left[\frac{P_1^2}{P_1 + P_2}\right] dy = 0.$$

This lemma can be proved in a manner similar to the preceding lemmas. First the definition of F given by Eq. (1) must be used to expand $F[P_1^2/(P_1 + P_2)]$ to give a sum of terms involving P_1, P_2, their derivatives, and the ratios $P_1/(P_1 + P_2)$ and $P_2/(P_1 + P_2)$. Each of these ratios is bounded and continuous. By using this fact along with the existence of the integrals implied by item (iii) of Definition 1, one can show that the integral of $F[P_1^2/(P_1 + P_2)]$ does indeed exist. As in the proof of Lemma 1, the integrals may be evaluated in any order, and by integrating each one with respect to y_k first, the boundedness of the ratios $P_1/(P_1 + P_2)$ and $P_2/(P_1 + P_2)$ coupled with item (vi) of Definition 1 leads to the fact that each of these integrals is zero. Thus the integral of $F[P_1^2/(P_1 + P_2)]$ is zero.

Lemma 4. If P_1 and P_2 are well-behaved probability densities and the Fokker–Planck operator F is as defined by Eq. (1), then

$$\frac{\partial}{\partial t}\left[\frac{P_1^2}{P_1 + P_2}\right] - F\left[\frac{P_1^2}{P_1 + P_2}\right]$$
$$= -D_{ki}(P_1 + P_2)\frac{\partial}{\partial y_k}\left[\frac{P_1}{P_1 + P_2}\right]\frac{\partial}{\partial y_i}\left[\frac{P_1}{P_1 + P_2}\right].$$

The proof of this lemma is algebraic; all that need be done is that the indicated derivatives be taken and the Fokker–Planck equation be applied for P_1 and P_2. Lemma 4 will follow directly.

The proof of the theorem now follows. For simplicity, let x be defined by the ratio

$$x = P_1/(P_1 + P_2). \tag{8}$$

Then, from Lemma 2, one has that

$$\frac{dM(t)}{dt} = 4\int \frac{\partial}{\partial t}(xP_1)\, dy.$$

But from Lemmas 3 and 4, one has that

$$\int \frac{\partial}{\partial t}(xP_1)\, dy = -\int D_{ki}(P_1 + P_2)\frac{\partial x}{\partial y_k}\frac{\partial x}{\partial y_i}\, dy,$$

so that

$$\frac{dM(t)}{dt} = -4\int D_{ki}(P_1 + P_2)\frac{\partial x}{\partial y_k}\frac{\partial x}{\partial y_i}\, dy. \tag{9}$$

As the D_{ki} are elements of a positive-definite matrix, the integral of Eq. (9) must be nonnegative, so that

$$dM(t)/dt \leq 0.$$

The continuity of $dM(t)/dt$ follows from the proof of Lemma 2.

To prove the corollary of Theorem 1, one assumes that the Fokker–Planck equation is steady. As the D_{ki} are positive-definite, and P_1 and P_2 are positive, the integrand in equation (9) must be positive. Thus $dM(t)/dt$ can be zero if and only if the integrand is identicaly zero,

$$D_{ki}(P_1 + P_2)\frac{\partial x}{\partial y_k}\frac{\partial x}{\partial y_i} = 0.$$

But as P_1 and P_2 themselves are positive, one can use item (i) of Definition 2 to show that from this, $dM(t)/dt$ can be zero if and only if

$$\partial x/\partial y_k = 0 \quad \text{for} \quad k = 1, 2, \cdots, m. \tag{10}$$

But, by using the fact that x is a ratio of two solutions to the Fokker–Planck equation along with Eq. (10), one has

$$\partial x/\partial t = -A_k\, \partial x/\partial y_k. \tag{11}$$

From Definition 2, it then follows that $dM(t)/dt$ can be zero if and only if x is a constant. If it is a constant then

$$\int x(P_1 + P_2)\, dy = x\int (P_1 + P_2)\, dy = 2x,$$

from Lemma 1. But one also has

$$\int x(P_1 + P_2)\, dy = \int P_1\, dy = 1,$$

so that the value of this constant must be $\frac{1}{2}$, which implies that $P_1 = P_2$. Thus, $dM(t)/dt$ can be zero if and only if P_1 and P_2 are equal.

UNIQUENESS

Let P_1 and P_2 each be well-behaved solutions to a Fokker–Planck equation, and as in Eq. (4), define $M(t)$. As pointed out earlier, $M(t)$ can be zero if and only if $P_1 = P_2$, and from Theorem 1, $M(t)$ cannot increase. Thus if P_1 and P_2 are equal at some initial time, $M(t)$ will be zero at that and all subsequent times. Thus $P_1 = P_2$ for all time following the initial time. This simply demonstrates the fact that well-behaved solutions to a Fokker–Planck equation, which satisfy a given set of initial conditions, are unique. This is not very profound, and of greater interest is uniqueness of the steady-state solutions.

Theorem 2. There is at most one steady-state solution to the steady Fokker–Planck equation.

The proof of this follows if one assumes that P_1 and P_2 are each steady-state solutions, so that their time derivatives are zero. Then from the definition of $M(t)$, one has $dM(t)/dt = 0$. However, if the equation is steady, from the corollary to Theorem 1, this implies that $P_1 = P_2$.

LARGE TIME BEHAVIOR

Let the Fokker–Planck equation be a steady equation, and let P_1 and P_2 be well-behaved probability densities. Then from Theorem 1 and its corollary one has that

$$M(t) \geq 0 \quad \text{with equality only for } P_1 = P_2, \tag{12}$$

$$dM(t)/dt \leq 0 \quad \text{with equality only for } P_1 = P_2. \tag{13}$$

As $M(t)$ is a decreasing function, bounded below, it must have a limit. Further as its derivative is negative and continuous, it too must have a limit, and this limit is zero. Thus there will be some number L, larger than or equal to 0, such that

$$\lim_{t\to\infty} [M(t)] = L,$$

$$\lim_{t\to\infty} [dM(t)/dt] = 0. \tag{14}$$

To form further conclusions, one must leave mathematical rigor behind. Formally, from Eqs. (13) and (14), one would suspect that as $t \to \infty$, P_1 would in some manner approach P_2, for $dM(t)/dt$ cannot be zero unless $P_1 = P_2$. Rigorously one cannot conclude such a thing without some sort of condition bounding $M(t)$ by some function of its derivative. However, until it is proven at a later date, one must be satisfied with the formal result that well-behaved solutions to a steady Fokker–Planck equation behave asymptotically the same.

Thus, if there exists a known steady-state solution, by letting it equal P_1, one can formally demonstrate that all well-behaved probability densities approach it as t goes to infinity.

Rigorously, all that can be concluded is that the function $M(t)$ is a decreasing function of time, and has a limit as t goes to infinity. Thus, in some sense P_1 and P_2 will tend towards each other.

ADDITIONAL COMMENT

Though it was not needed here for this, it might be of use for some purposes to use other definitions of $M(t)$. It can be shown[3] that if $f(x)$ is a function of x with a positive second derivative, the constants a and b are chosen so that $0 < a < b < 1$, and P_3 and P_4 are defined by

$$P_3 = aP_1 + (1 - a)P_2,$$

$$P_4 = bP_1 + (1 - b)P_2,$$

then an $M(t)$ could be defined by the integral

$$M(t) = \int P_4 f\left(\frac{P_3}{P_4}\right) dy.$$

In this case then, for a steady Fokker–Planck equation,

$$M(t) \geq f(1)$$

and

$$dM(t)/dt \leq 0,$$

with equality holding in both cases if and only if $P_1 = P_2$.

SUMMARY

It has been demonstrated that well-behaved probability densities (according to Definition 1)—which satisfy a given Fokker–Planck equation and initial conditions—are unique. Further, when this Fokker–Planck equation is steady (Definition 2) there can be at most one steady-state solution. Thus steady-state solutions are unique for the class of steady Fokker–Planck equations.

It has also been formally demonstrated that in the case of a steady Fokker–Planck equation, solutions tend towards each other asymptotically.

The theorems developed here are for the case where no portion of phase space is excluded. In the case where the probability of lying in some N-dimensional volume might be zero, certain modifications must be made, which are not discussed here.

[3] A. H. Gray, Jr., "Stability and Related Problems in Randomly Excited Systems," Doctoral thesis, California Institute of Technology (1964), pp. 1–20.

Property of the Vertex Function in Potential Theory*

R. ROCKMORE
Physics Department, Rutgers—The State University, New Brunswick, New Jersey and Brookhaven National Laboratory, Upton, New York
(Received 18 September 1964)

The vertex function of an s-wave bound state in potential theory is discussed in terms of a nonrelativistic Bethe–Salpeter formalism. It is shown that the Born series defining the off-shell amplitude, $\Gamma(k^2, P_0)$, has the property that $\Gamma(k^2, k^2/m) = \Gamma_0/D(k^2, \lambda)$, where $D(k^2, \lambda)$ is the Jost function for potential strength λ, and where the bare (Γ_0) and physical (Γ_c) couplings are related by $\Gamma_0 = \Gamma_c Z_1$, with Z_1, the vertex renormalization constant equal to $D(-\alpha^2, \lambda)$. The particular proof presented derives from a detailed comparison in momentum space of alternative, but equivalent, coordinate space definitions of the Jost function originating with Jost and Pais.

I. INTRODUCTION

WE have lately employed Bethe–Salpeter methods in constructing "a field theory of potential scattering,"[1] as it were, our purpose being to show explicitly how various aspects of the eigenvalue problem for bound states in potential theory may be characterized by assertions as to the vanishing of bound-state vertex and wavefunction renormalization constants. (It is perhaps almost unnecessary to remark here that the utility of discussing the eigenvalue problem from this point of view has been the subject of a number of communications[2] in recent times.) In this note we derive, rather than infer,[1] a property of the vertex function of a bound state in potential theory, which seems useful in justifying the formal manipulation of certain divergent expansions one encounters there. At the same time this derivation implies an alternative proof of a result obtained by Bertocchi *et al.*[3] in their recent discussion of vertex renormalization in the potential theory domain.

In Sec. II we introduce the nonrelativistic vertex function $\Gamma(k, p)$ of an s-wave bound state[4] via a Bethe–Salpeter formalism.[1] This device enables us to associate readily the successive terms of $\Gamma(k, p)$ in its perturbation expansion[5] with the corresponding simply-structured Feynman graphs. In Sec. III we prove the useful property of the vertex function, $\Gamma(k^2, p_0)$, namely that

$$\Gamma(k^2, k^2/m) = \Gamma_0/D(k^2, \lambda), \tag{1}$$

where $D(k^2, \lambda)$ is the Jost function for potential strength λ, and where the bare (Γ_0) and physical (Γ_c) couplings are related by

$$\Gamma_0 = \Gamma_c Z_1(-\alpha^2, \lambda), \tag{2}$$

with Z_1, the vertex renormalization constant, equal to $D(-\alpha^2, \lambda)$. The particular proof presented derives from a detailed comparison in momentum space of alternative, but equivalent, coordinate space definitions of the Jost function originating with Jost and Pais.[6]

II. BETHE–SALPETER FORMULATION OF VERTEX EQUATION

The bound-state vertex $\Gamma(x; y, z) \equiv \Gamma(x-y, x-z)$ satisfies the inhomogeneous integral equation

$$\Gamma(x; y, z) = \Gamma_0\, \delta(x-y)\, \delta(x-z) + \iiiint d^4y'\, d^4z'\, d^4y''\, d^4z'' \times \Gamma(x; y', z') K_0(y', y'') K_0(z', z'') G(y'', z''; y, z), \tag{3}$$

where, to recapitulate briefly,[1] the kernel function $K_0(y', y'')$ is given by

$$K_0(y', y'') = -i \int \frac{d^4k}{(2\pi)^4} \frac{e^{ik(y'-y'')}}{k_0 - \mathbf{k}^2/2m + i\epsilon}, \tag{4}$$

and

$$G(x_1, x_2; x_3, x_4) = -\int \frac{d^4q\, d^4q'\, d^4P}{(2\pi)^8} G(q, q'; P) \times \exp iP\left(\frac{x_1 + x_2}{2} - \frac{x_3 + x_4}{2}\right) \exp\,(iqx_{12} - iq'x_{34}); \tag{5}$$

---

* Work supported by the U. S. Atomic Energy Commission.

[1] R. Rockmore, Ann. Phys. (N. Y.) (to be published).

[2] M. Cini, CERN report 7508/Th 385; A. Salam, Nuovo Cimento **25**, 224 (1962); S. Weinberg, Phys. Rev. **130**, 776 (1963); A. Bastai, L. Bertocchi, S. Fubini, G. Furlan, and M. Tonin, Nuovo Cimento **30**, 512 (1963); R. Rockmore, Phys. Rev. **132**, 878 (1963).

[3] L. Bertocchi, M. McMillan, E. Predazzi, and M. Tonin, Nuovo Cimento **31**, 1352 (1964).

[4] The extension of this discussion to other values of l is straightforward.

[5] We refer to an expansion in powers of the potential strength λ.

[6] R. Jost and A. Pais, Phys. Rev. **82**, 840 (1952).

moreover, we restrict ourselves to the Yukawa kernel, so that

$$G(q, q'; P) = -i\lambda(4\pi^3)^{-1}[(\mathbf{q} - \mathbf{q}')^2 + \mu^2]^{-1}. \quad (6)$$

In momentum space, one finds, quite straightforwardly, the analogous integral equation

$$\Gamma(k, P) = \Gamma_0 + \int d^4k' \Gamma(k', P)\left[\left(\frac{1}{2}P_0 + k_0'\right) - \frac{(\frac{1}{2}\mathbf{P} + \mathbf{k}')^2}{2m} + i\epsilon\right]^{-1} \times [(\tfrac{1}{2}P_0 - k_0') - (\tfrac{1}{2}\mathbf{P} - \mathbf{k}')^2/2m + i\epsilon]^{-1} G(k', k; P), \quad (7)$$

where $\Gamma(k, P)$ is defined by

$$\Gamma(x; y, z) = \iint \frac{d^4k\, d^4P}{(2\pi)^8} e^{i[Px - (\frac{1}{2}P + k)y - (\frac{1}{2}P - k)z]} \Gamma(k, P). \quad (8)$$

The graphical structure of Eq. (7), as well as that of its perturbation expansion, is exhibited in Fig. 1. It is easy to see[7] that the vertex function $\Gamma(k, p)$ does not, in fact, depend on k_0, so that one may further reduce (2.5) to

$$\Gamma(\mathbf{k}, P) = \Gamma_0 + \int \frac{d\mathbf{k}'}{(2\pi)^3}\left[\frac{k'^2}{m} - \left(P_0 - \frac{\mathbf{P}^2}{4m}\right) - i\epsilon\right]^{-1} \times \frac{(4\pi\lambda)}{[(\mathbf{k} - \mathbf{k}')^2 + \mu^2]} \Gamma(\mathbf{k}', P). \quad (9)$$

Finally, we take note of the $l = 0$ character of the bound state, in writing $\Gamma(k^2, P_0)$ for the vertex function in the frame $\mathbf{P} = 0$. It may be that the vertex function rather than the wavefunction is a more natural subject for discussion in the nonrelativistic domain. However, the two are simply related in that domain by [8]

$$\varphi_0(k^2; P_0) = \Gamma(k^2; P_0)/(k^2 - mP_0 - i\epsilon). \quad (10)$$

It is not at all obvious, although it might be expected from previous work,[9] that the Born expansion,

$$\Gamma(k^2, P_0) = \Gamma_0\left\{1 + \int \frac{d\mathbf{k}'}{(2\pi)^3} \frac{[-m\tilde{V}(\mathbf{k}' - \mathbf{k})]}{(k'^2 - mP_0 - i\epsilon)} + \iint \frac{d\mathbf{k}'\, d\mathbf{k}''}{[(2\pi)^3]^2} \frac{[-m\tilde{V}(\mathbf{k}' - \mathbf{k}'')]}{(k'^2 - mP_0 - i\epsilon)} \times \frac{[-m\tilde{V}(\mathbf{k}'' - \mathbf{k})]}{(k''^2 - mP_0 - i\epsilon)} + \cdots\right\} \equiv \Gamma_0 M(k^2, P_0), \quad (11)$$

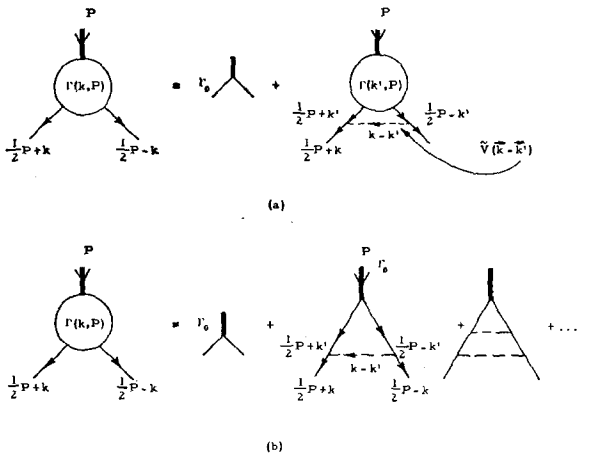

FIG. 1. Inhomogeneous Bethe–Salpeter equation for (a) the vertex function $\Gamma(k, P)$ and (b) its perturbation expansion in Feynman graphs.

has the property (1). Notice that since[3,10]

$$Z_1(-\alpha^2, \lambda) = D(-\alpha^2, \lambda), \quad (12)$$

it follows necessarily that

$$\Gamma(-\alpha^2, -\alpha^2/m) = \Gamma_c Z_1/D(-\alpha^2, \lambda) = \Gamma_c. \quad (13)$$

On the other hand, one could define a "renormalized" D function (D_c) such that

$$D_c(k^2, \lambda) = D(k^2, \lambda)/Z_1(-\alpha^2, \lambda)$$

with

$$\Gamma(k^2, k^2/m) = \Gamma_c/D_c(k^2, \lambda);$$

then[1]

$$D_c(k^2, \lambda) = \exp\left[\frac{(k^2 + \alpha^2)}{\pi}\int_0^\infty \frac{2k'\, dk'\, \delta^0(k'^2)}{(k'^2 + \alpha^2)(k'^2 - k^2 - i\epsilon)}\right].$$

In particular, we point to the apparent dissimilarity between the expansion of the symmetrical D function[6,11]

$$D(k^2, \lambda) = 1 - \frac{\lambda}{\pi}\int_0^\infty dk' \frac{1}{k'^2 - k^2 - i\epsilon} Q_0\left(1 + \frac{\mu^2}{2k'^2}\right) + \frac{\lambda^2}{2!}\left\{\left[\frac{1}{\pi}\int_0^\infty dk' \frac{1}{k'^2 - k^2 - i\epsilon} Q_0\left(1 + \frac{\mu^2}{2k'^2}\right)\right]^2 - \frac{1}{\pi^2}\int_0^\infty dk' \int_0^\infty dk'' \frac{1}{(k'^2 - k^2 - i\epsilon)(k''^2 - k^2 - i\epsilon)} \times \left[Q_0\left(\frac{k'^2 + k''^2 + \mu^2}{2k'k''}\right)\right]^2\right\} + \cdots \quad (14)$$

[7] See Eq. (2.4).
[8] We write the wavefunction of Eq. (10) off the mass shell ($P_0 = -\alpha^2/m$) of the composite.
[9] See the material in Ref. 1 after Eq. (39).
[10] Since the Jost function vanishes on the mass shell of the composite, we identify $Z_1 = 0$ with the eigenvalue condition.
[11] B. W. Lee, *Seminar on Theoretical Physics* (International Atomic Energy Agency, Trieste, 1963), p. 331.

and the inverse of the Born series $M(k^2, k^2/m)$ in (11),

$$M^{-1}(k^2, k^2/m)$$

$$= 1 - \frac{\lambda}{\pi}\int_0^\infty \frac{(k'/k)\,dk'}{k'^2 - k^2 - i\epsilon} Q_0\left(\frac{k'^2 + k^2 + \mu^2}{2k'k}\right)$$

$$+ \lambda^2\left\{\left[\frac{1}{\pi}\int_0^\infty \frac{(k'/k)\,dk'}{k'^2 - k^2 - i\epsilon} Q_0\left(\frac{k'^2 + k^2 + \mu^2}{2k'k}\right)\right]^2\right.$$

$$- \frac{1}{\pi^2}\int_0^\infty dk' \int_0^\infty dk'' \frac{(k'/k'')}{k'^2 - k^2 - i\epsilon} Q_0\left(\frac{k'^2 + k''^2 + \mu^2}{2k'k''}\right)$$

$$\left.\times \frac{(k''/k)}{k''^2 - k^2 - i\epsilon} Q_0\left(\frac{k''^2 + k^2 + \mu^2}{2k''k}\right)\right\} + \cdots . \quad (15)$$

[However, note that the terms $O(\lambda)$ *are* identical,

$$D_{(1)}(k^2, k^2/m)$$

$$= -\frac{1}{2\pi}\int_0^\infty \frac{ds'}{s'^{\frac{1}{2}}(s' - k^2 - i\epsilon)} Q_0\left(1 + \frac{\mu^2}{2s'}\right) \quad (16)$$

being merely $M_{(1)}^{-1}(k^2, k^2/m)$ expressed as a dispersion relation,

$$M_{(1)}^{-1}\left(k^2, \frac{k^2}{m}\right) = \frac{1}{\pi}\int_0^\infty ds' \frac{\operatorname{Im} M_{(1)}^{-1}(s', s'/m)}{s' - k^2 - i\epsilon}.\Big] \quad (17)$$

Note added in proof: The mode of proof suggested by Eqs. (16) and (17) has been extended to $O(\lambda^2)$ by the author (unpublished). It requires the use of the identity

$$P\frac{1}{k^2 - k_1^2} P\frac{1}{k^2 - k_2^2} = \pi^2\delta(k^2 - k_1^2)\delta(k^2 - k_2^2)$$

$$+ P\frac{1}{k_1^2 - k_2^2} P\frac{1}{k^2 - k_1^2} + P\frac{1}{k_2^2 - k_1^2} P\frac{1}{k^2 - k_2^2},$$

to put $M_{(2)}(k^2, k^2/m)$ into the symmetric dispersion form $D_{(2)}(k^2, k^2/m)$. Suitable generalization of the aforementioned identity for more than two variables is needed in the higher orders.

III. PROOF FROM THE COORDINATE SPACE FORMULATION OF POTENTIAL SCATTERING

We may resolve this seeming paradox through a detailed comparison of the Fourier analyses of the physical and Jost solutions to the radial Schrödinger equation. For scattering in the $l = 0$ partial wave, one has[12]

$$\psi_0(k, r) = \sin kr + \lambda\int_0^\infty dr' K_0(k; r, r')\psi_0(k, r'), \quad (18)$$

where the Green's function kernel, $K_0(k; r, r')$, is given by

$$K_0(k; r, r') = -k^{-1}\sin kr_< e^{ikr_>} V(r'); \quad (19)$$

$V(r')$ is defined by

$$V(r) = -e^{-\mu r}/r, \quad (20)$$

so as to exhibit explicitly the λ dependence. In terms of the Fourier sine transform of $\psi_0(k, r)$,

$$\tilde{\psi}_0(k, k') = \left(\frac{2}{\pi}\right)^{\frac{1}{2}}\int_0^\infty \psi_0(k, r)\sin k'r\,dr, \quad (21)$$

one finds analogously

$$\tilde{\psi}_0(k, k') = (\tfrac{1}{2}\pi)^{\frac{1}{2}}\delta(k - k') + \frac{1}{k'^2 - k^2 - i\epsilon}\frac{\lambda}{\pi}$$

$$\times \int_0^\infty dk'' Q_0\left(\frac{k'^2 + k''^2 + \mu^2}{2k'k''}\right)\psi_0(k, k''). \quad (22)$$

Now it is possible to rewrite[13] the integral equation (18) so that the factorized function, $\phi_0(k, r)$ defined by[14]

$$\psi_0(k, r) = M(k^2, k^2/m)\phi_0(k, r), \quad (23)$$

where[15]

$$M(k^2, k^2/m) = 1 - \lambda\int_0^\infty dr\, k^{-1} e^{ikr} V(r)\psi_0(k, r) \quad (24)$$

$$= 1 + \frac{\lambda}{\pi}\int_0^\infty \frac{(k'/k)\,dk'}{k'^2 - k^2 - i\epsilon}\int_0^\infty dk''$$

$$\times Q_0\left(\frac{k'^2 + k''^2 + \mu^2}{2k'k''}\right)\left(\frac{2}{\pi}\right)^{\frac{1}{2}}\tilde{\psi}_0(k, k'') \quad (25)$$

satisfies the Volterra equation

$$\phi_0(k, r) = \sin kr + \lambda\int_0^r dr' g_0(k; r, r')V(r')\phi_0(k, r'), \quad (26)$$

with

$$g_0(k; r, r') = k^{-1}\sin k(r - r'). \quad (27)$$

Since the momentum-space analog of (26) is

$$\tilde{\phi}_0(k, k') = \left(\frac{\pi}{2}\right)^{\frac{1}{2}}\delta(k - k')\left\{1 - \frac{\lambda}{\pi}\int_0^\infty dk''\right.$$

$$\times \int_0^\infty \frac{(k'''/k)\,dk'''}{k'''^2 - k^2 - i\epsilon} Q_0\left(\frac{k'''^2 + k''^2 + \mu^2}{2k''k'''}\right)$$

$$\left.\times \left(\frac{2}{\pi}\right)^{\frac{1}{2}}\tilde{\phi}_0(k, k'')\right\} + \frac{1}{k'^2 - k^2 - i\epsilon}\frac{\lambda}{\pi}$$

$$\times \int_0^\infty dk'' Q_0\left(\frac{k'^2 + k''^2 + \mu^2}{2k'k''}\right)\left(\frac{2}{\pi}\right)^{\frac{1}{2}}\tilde{\phi}_0(k, k''), \quad (28)$$

we have, by inspection,

$$M^{-1}\left(k^2, \frac{k^2}{m}\right) = 1 - \frac{\lambda}{\pi}\int_0^\infty dk''\int_0^\infty \frac{(k'''/k)\,dk'''}{k'''^2 - k^2 - i\epsilon}$$

$$\times Q_0\left(\frac{k'''^2 + k''^2 + \mu^2}{2k''k'''}\right)\left(\frac{2}{\pi}\right)^{\frac{1}{2}}\tilde{\phi}_0(k, k'') \quad (29)$$

[12] Our notation follows that of Gillespie [J. R. Gillespie, Ph.D. thesis, University of California, 1963 (available as UCRL-10762)] closely. However, for simplicity, we have taken our $\phi_l(k, r)$ [Eq. (26)] to be k^{l+1} times his. We have also corrected a misprint in his Eq. (48) defining $g_l(k; r, r')$.

[13] H. Brysk, J. Math. Phys. 4, 1536 (1963); N. Corngold, "Lectures on Neutron Physics" (unpublished).

[14] Our notation anticipates our result.

[15] The Born series [Eq. (11)] follows from the substitution into (25) of the iterative solution to Eq. (22).

$$= 1 + \lambda \int_0^\infty dr' k^{-1} e^{ikr'} V(r') \phi_0(k, r'). \tag{30}$$

Jost and Pais have shown rather elegantly in Appendix II of their paper[6] that the function $M^{-1}(k^2, k^2/m)$ so defined [Eqs. (29) and (30) above] is just the Fredholm determinant $D(k^2, \lambda)$. Indeed, defining the D function recursively, as in Brysk,[13] one has

$$D(k^2, \lambda) = \sum_{n=0}^{\infty} (-1)^n (n!)^{-1} c^{(n)}(k^2) \lambda^n, \tag{31}$$

with

$$c^{(0)} = 1, \tag{32}$$

and

$$c^{(n+1)}(k^2) = \int_0^\infty dk' \tilde{B}^{(n)}(k; k', k'), \tag{33}$$

where

$$\tilde{B}^{(n)}(k; k', k'') = \frac{(1/\pi)}{k'^2 - k^2 - i\epsilon} \left[c^{(n)}(k^2) Q_0\left(\frac{k'^2 + k''^2 + \mu^2}{2k'k''}\right) - n \int_0^\infty dk''' Q_0\left(\frac{k'''^2 + k'^2 + \mu^2}{2k'k'''}\right) \tilde{B}^{(n-1)}(k; k''', k'') \right], \tag{34}$$

so that

$$(-1)^{n+1}[(n+1)!]^{-1} c^{(n+1)}(k^2) = \int_0^\infty dr' k^{-1} e^{ikr'} V(r') \phi_0^{(n)}(k, r'), \tag{35}$$

where

$$\phi_0(k, r) = \sum_{n=0}^{\infty} \lambda^n \phi_0^{(n)}(k, r). \tag{36}$$

Finally, we note that our proof yields, as a by-product, the perturbation-theoretic representation for $(Z_1(-\alpha^2, \lambda))^{-1}$ obtained by Bertocchi *et al.*,[3,16]

$$(Z_1(-\alpha^2, \lambda))^{-1} = \lim_{k^2 \to -\alpha^2} \left\{ 1 + \frac{\lambda}{\pi} \int_0^\infty \frac{(k'/k)\, dk'}{k'^2 - k^2 - i\epsilon} Q_0\left(\frac{k'^2 + k^2 + \mu^2}{2k'k}\right) + \frac{\lambda^2}{\pi^2} \int_0^\infty dk' \int_0^\infty dk'' \frac{(k'/k'')}{k'^2 - k^2 - i\epsilon} Q_0\left(\frac{k'^2 + k''^2 + \mu^2}{2k'k''}\right) \times \frac{(k''/k)}{k''^2 - k^2 - i\epsilon} Q_0\left(\frac{k''^2 + k^2 + \mu^2}{2kk''}\right) + \cdots \right\}. \tag{37}$$

ACKNOWLEDGMENTS

We wish to express our gratitude to Dr. Jack Chernick and to the late Dr. Charles E. Porter for their hospitality during our stay at Brookhaven National Laboratory in the summer of 1964.

[16] However, they have chosen to consider a superposition of Yukawa kernels rather than the simple kernel (6).

Vertex Modification for Coulombic Interactions by Exact Summation of Ladder Diagrams

YOH-HAN PAO AND H. L. FRISCH
Bell Telephone Laboratories, Incorporated, Murray Hill, New Jersey
(Received 18 September 1964)

We have been able to sum a class of ladder diagrams for Coulombic interactions to infinite order exactly and are able therefore to replace a "bare" interaction with a self-consistent effective interaction or equivalently we are able to modify the vertex operator in a manner which may be of particular significance for systems of low particle density in analogy with what can be done for nuclear matter and liquid helium. The method is only valid for nearly free particles, the exact energy being negative but the zero-order description being that of free particles. Effective attractive interactions between two electrons or interaction with an external Coulombic field constitute perturbations which fall within the range of validity of the scheme. Despite the limitation of the specific results, the method is probably of more general use and interest and both the approach and specific results are described in this note.

FOR nuclear matter and liquid helium, the interaction between particles is so short ranged (effectively hard core), that matrix elements of these "bare" interactions actually diverge. However, in perturbation-theory studies, if the contributions of all those terms representing repeated interactions between two particles are summed to infinite order, it is found that the net contribution is finite and a well-behaved effective interaction may be defined in this manner.[1,2] This replacement of the original interaction by an effective interaction is often referred to as modification of the vertex and corresponds to replacement of the V matrix by the K matrix in the Brueckner theory or to evaluation of the reaction (R) matrix of scattering theory.

Although the obvious importance of vertex modification for studies of nuclear matter and liquid helium is well recognized, it has received little attention for Coulombic interactions. On the contrary, the results of investigations of the dense electron gas have focused attention on the "ring" diagrams which are important because their contributions are principally responsible for the description of collective motions in such gases.[3] In extremely low-density electron gases, the work of Wigner[4] and Carr[5] indicate that when kinetic energy is insufficient to overcome the repulsive interaction between electrons, the physical state is probably that of an orderly spatial array and interaction energies may be estimated on that basis.

It is interesting nevertheless that this replacement of a bare Coulombic interaction with an effective interaction can indeed be carried out for dilute systems where in the evaluation of a repeated interaction between two particles, all other particles may be considered to be in the unexcited state. From a pedagogical point of view, it would be of interest to show explicitly that the newer results yield a simple result which can be compared with the application of conventional theory.

We are able to sum to infinite order all contributions to the interaction energy of (a) two weakly bound electrons with paired momenta and paired spin interacting through a Coulomb potential (without exchange), (b) two weakly bound electrons with general (nonpaired) momenta and paired spin interacting again without exchange, (c) two weakly bound electrons with paired momenta and the same spin interaction with exchange, and (d) one free electron interacting with an external attractive Coulomb potential, but we do not treat particulars of this case. This summation is equivalent to replacing the original vertex operator with a modified vertex and since the Brillouin–Wigner perturbation expansion is used, the effective interaction so obtained is also a self-consistent one.

Unfortunately it seems that the results have a severely limited range of applicability being valid only for systems where the zero-order description is that of plane waves but the net effect of all the interactions is to produce a bound state. Presumably the results may be applied to yield the binding energy of a bound electron in a low-lying trap and to

[1] H. A. Bethe and J. Goldstone, Proc. Roy. Soc. (London) **A238**, 551 (1957).
[2] K. A. Brueckner and W. Wade, Phys. Rev. **103**, 1008 (1956).
[3] M. Gell-Mann and K. A. Brueckner, Phys. Rev. **106**, 364 (1957).
[4] E. P. Wigner, Phys. Rev. **46**, 1002 (1934); Trans. Faraday Soc. **34**, 678 (1938).
[5] W. J. Carr, Phys. Rev. **122**, 1437 (1961).

other problems of this nature. In the event that the final state is of positive energy, this procedure may be carried through to yield exact scattering cross sections for simple cases, but these latter results may be obtained equally well with other methods such as the partial wave analysis[6] and are not discussed further.

We consider two electrons with pair momenta and paired spins and let H_0, H, φ_0, and ψ represent the zero-order and exact Hamiltonians and wavefunctions, respectively. Using the Brillouin–Wigner expansion we proceed directly from the time-independent equations

$$H_0\varphi_0 = E_0\varphi_0, \tag{1}$$

$$H\psi = (H_0 + V)\psi = E\psi = (E_0 + \Delta E)\psi, \tag{2}$$

to obtain formally

$$\psi = \varphi_0 + [(1 - P_0)/(E - H_0)]V\psi, \tag{3}$$

which may be solved by iteration to yield

$$\psi = \varphi_0 + \frac{1 - P_0}{E - H_0} V\varphi_0 + \frac{1 - P_0}{E - H_0} V \frac{1 - P_0}{E - H_0} V\varphi_0 + \frac{1 - P_0}{E - H_0} V \frac{1 - P_0}{E - H_0} V \frac{1 - P_0}{E - H_0} V\varphi_0 + \cdots, \tag{4}$$

where P_0 is a projection operator onto the subspace spanned by φ_0.

A matrix element of the vertex operator $\langle\chi\,|V|\,\psi\rangle$ is then

$$\langle\chi|\ V\ |\psi\rangle = \langle\chi|\ V\ |\varphi_0\rangle + \langle\chi|\ V \frac{1 - P_0}{E - H_0} V\ |\varphi_0\rangle + \langle\chi|\ V \frac{1 - P_0}{E - H_0} V \frac{1 - P_0}{E - H_0} V\ |\varphi_0\rangle + \cdots, \tag{5}$$

or equivalently,

$$\langle\chi|\ V\ |\psi\rangle = \langle\chi|\ v\ |\varphi_0\rangle, \tag{6}$$

where

$$v = V + V[(1 - P_0)/(E - H_0)]v \tag{7}$$

and is the modified vertex operator.

Instead of solving the inhomogeneous Eq. (7) directly, we evaluate expression (6) if $\chi = |\mathbf{k}, -\mathbf{k}\rangle$ by summing all the terms. The Hamiltonian is $H = H_0 + V$ where $H_0 = \mathbf{p}_1^2/2m + \mathbf{p}_2^2/2m$ $V = -e^2/|\mathbf{r}_1 - \mathbf{r}_2|$ and is an effective attractive interaction between the weakly bound electrons due to (say) the presence of a polarizable center of attraction. The initial state is $\varphi_0 = |\mathbf{k}_0 - \mathbf{k}_0\rangle = v^{-1} \exp(-i\mathbf{k}_0\cdot\mathbf{r}_1 + i\mathbf{k}_0\cdot\mathbf{r}_2)$, where v is the volume of the system. Then it can be shown that in the limit as $v \to \infty$ the terms involving the projection operator P_0 are of a higher order in $1/v$ and can be neglected. In this approach, the contributions of the first two orders are treated as special cases, but all the other terms from third order onwards may be written so that the contribution of $(n + 2)$nd term is given by

$$\begin{aligned}
&\langle\chi|\ V \cdots (E - H_0)^{-1} V\ |\varphi_0\rangle^{(n)} \\
&= (1)^{n+1} \int d^3\mathbf{k}_n\, d^3\mathbf{k}_{n-1} \cdots d^3\mathbf{k}\, \langle\chi|\ V\ |\mathbf{k}_n\rangle \frac{1}{(\mathbf{k}_n^2 - E)} \\
&\quad \times \frac{1}{2\pi^2} \frac{1}{(\mathbf{k}_n - \mathbf{k}_{n-1})^2} \frac{1}{(\mathbf{k}_{n-1}^2 - E)} \\
&\quad \times \frac{1}{2\pi^2(\mathbf{k}_{n-1} - \mathbf{k}_{n-2})^2} \cdots \frac{1}{\mathbf{k}^2 - E} \langle\mathbf{k}|\ V\ |\varphi_0\rangle \\
&\equiv (-1)^{n+1} \int d^3\mathbf{k}_n \cdots d^3\mathbf{k}\, \langle\chi|\ V\ |\mathbf{k}_n\rangle \\
&\quad \times \frac{1}{(\mathbf{k}_n^2 - E)^{\frac{1}{2}}} G(\mathbf{k}_n, \mathbf{k}_{n-1}) \\
&\quad \times G(\mathbf{k}_{n-1}, \mathbf{k}_{n-2}) \cdots G(\mathbf{k}_1, \mathbf{k}) \frac{1}{(\mathbf{k}^2 - E)^{\frac{1}{2}}} \langle\mathbf{k}|\ V\ |\varphi_0\rangle,
\end{aligned}$$

where the function $G(\mathbf{k}, \mathbf{k}')$ is given by

$$G(\mathbf{k}, \mathbf{k}') = \left(\frac{1}{\mathbf{k}^2 - E}\right)^{\frac{1}{2}} \frac{1}{2\pi^2} \frac{1}{(\mathbf{k} - \mathbf{k}')^2} \left(\frac{1}{\mathbf{k}'^2 - E}\right)^{\frac{1}{2}}. \tag{9}$$

We expand $G(\mathbf{k}, \mathbf{k}')$ as a bilinear form so that

$$G(\mathbf{k}, \mathbf{k}') = \sum a_{\alpha\beta}\psi_\alpha(\mathbf{k})\psi_\beta^*(\mathbf{k}'), \tag{10}$$

where the as yet unknown functions $\psi_\alpha(\mathbf{k})$ constitutes a complete set of functions orthonormal in $\mathbf{k}$ space, that is

$$\int \psi_\alpha(\mathbf{k})\psi_\beta^*(k)\, d^3\mathbf{k} = \delta_{\alpha\beta}. \tag{11}$$

If the set ψ_α is the set of orthonormal characteristic functions of the integral equation

$$\psi_\alpha(k) = \lambda \int G(\mathbf{k}, \mathbf{k}')\psi_\alpha(\mathbf{k}')\, d^3\mathbf{k}', \tag{12}$$

then

$$a_{\alpha\beta} = \frac{1}{\lambda_\beta} \int \psi_\alpha^*(\mathbf{k})\psi_\beta(\mathbf{k})\, d^3\mathbf{k} = \frac{1}{\lambda_\beta}\, \delta_{\alpha\beta}. \tag{13}$$

Substituting this result back into Expression (8), we have

$$\begin{aligned}
&\langle\chi|\ V \frac{1}{E - H_0} V \cdots \frac{1}{E - H_0} V\ |\psi\rangle^{(n)} \\
&= (-1)^{n+1} \sum_\alpha \frac{1}{\lambda^n} \int d^3\mathbf{k}\, d^3\mathbf{k}'\, \langle\chi|\ V\ |\mathbf{k}\rangle \frac{1}{(\mathbf{k}^2 - E)^{\frac{1}{2}}} \\
&\quad \times \psi_\alpha(\mathbf{k})\psi_\alpha(\mathbf{k}') \frac{1}{(\mathbf{k}'^2 - E)^{\frac{1}{2}}} \langle\mathbf{k}'|\ V\ |\varphi_0\rangle.
\end{aligned} \tag{14}$$

[6] T. Y. Wu and T. Ohmura, *Quantum Theory of Scattering* (Prentice Hall, Inc., Englewood Cliffs, New Jersey, 1962).

Subsequent summation over n from $n = 1$ (the third-order term) to $n = \infty$, yields an expression for the sum of all the ladder diagrams exclusive of the first two contributions. Actually, as we shall see, solution of the integral equation is achieved in terms of $\bar{\varphi}_\alpha(\mathbf{s})$ which is the Fourier transform of the function $\psi_\alpha(\mathbf{k})/[\mathbf{k}^2-E]^{\frac{1}{2}}$ and we may therefore more compactly write the exact expression for $\langle\chi\,|v|\,\varphi_0\rangle$

$$\langle\chi|\ V\ |\varphi_0\rangle = \langle\chi|\ V + V\frac{1}{E-H_0}V + \sum_\alpha \frac{(N^m_{\alpha l})^2}{\lambda_\alpha+1}\int d^3\mathbf{k}\,d^3\mathbf{k}'\,d^3\mathbf{s}\,d^3\mathbf{s}' \times V\mathbf{k}\varphi_\alpha(\mathbf{s})e^{-i\mathbf{s}\cdot\mathbf{k}}\varphi^*_\alpha(\mathbf{s}')e^{i\mathbf{s}'\cdot\mathbf{k}'}\mathbf{k}'V\ |\varphi_0\rangle, \quad (15)$$

thus yielding the matrix element of the modified vertex operator in terms of the solutions of the homogeneous integral Eq. (12). In expression (15), $N^m_{\alpha l}$ is a normalization factor.

For weakly bound particles with paired momenta, and paired spin, the integral equation to be solved is

$$\psi_\alpha(\mathbf{k}) = \lambda_\alpha\int\frac{1}{[\mathbf{k}^2-E]^{\frac{1}{2}}}\frac{1}{2\pi^2(\mathbf{k}-\mathbf{k}')^2} \times\frac{1}{[\mathbf{k}'^2-E]^{\frac{1}{2}}}\psi_\alpha(\mathbf{k}')\,d^3\mathbf{k}'. \quad (16)$$

We define the function $\varphi_\alpha(\mathbf{k})$ by the relation

$$\varphi(\mathbf{k})[\mathbf{k}^2-E]^{\frac{1}{2}} = \psi(\mathbf{k}) \quad (17)$$

and find that the differential equation satisfied by the Fourier transform of $\varphi(\mathbf{k})$ is

$$\nabla^2_{\mathbf{s}}\varphi(\mathbf{s}) + (\lambda/|\mathbf{s}|)\varphi(\mathbf{s}) = -E\varphi(\mathbf{s}). \quad (18)$$

This may be compared with the Schrödinger equation for one electron in a spherically symmetric Coulomb potential and we know therefore that the solutions may be written in terms of products of radial parts and spherical harmonics. In particular we obtain the eigenfunctions to be

$$\varphi^m_{\alpha l}(\mathbf{s}) = -(2\gamma)^{\frac{3}{2}}\left\{\frac{(\alpha-l-1)!}{2\alpha[(\alpha+l)!]^3}\right\}^{\frac{1}{2}} \times e^{-\xi/2}\xi^l L^{2l+1}_{\alpha+l}(\xi)\,Y_{lm}(\theta,\varphi), \quad (19)$$

where $L^k_n(\xi)$ are the associated Laguerre functions and Y_{lm} are the spherical harmonics. The eigenvalues are

$$\lambda_\alpha = 2\gamma\alpha, \quad (20)$$

$$\gamma = (-E)^{\frac{1}{2}}, \quad (21)$$

and the variable ξ is related to s by the transformation $\xi = 2\gamma s$.

It is found that the normalization constant $N^m_{\alpha l}$ of the equation is independent of α, l and m and is simply given by

$$N = 1/2\gamma. \quad (22)$$

For two electrons with paired momenta but with the same spin, there is exchange and the integral equation corresponding to (16) is

$$\psi(\mathbf{k}) = \lambda\int\frac{1}{[\mathbf{k}^2-E]^{\frac{1}{2}}}\left\{\frac{1}{2\pi^2(\mathbf{k}-\mathbf{k}')^2} - \frac{1}{2\pi^2}\frac{1}{(\mathbf{k}+\mathbf{k}')^2}\right\}\frac{\psi(\mathbf{k}')}{[\mathbf{k}'^2-E]^{\frac{1}{2}}}\,d^3\mathbf{k}' \quad (23)$$

and we find that the solutions $\varphi(\mathbf{s})$ in this case are the odd members of those listed in expression (19) and the normalization of $\psi(\mathbf{k})$ is carried out as before with identical results.

For two electrons with paired spins and general momenta, a typical and general transition at a vertex would be from state $(-\mathbf{k}'-\mathbf{q},\,\mathbf{k}')$ to that of $(-\mathbf{k}-\mathbf{q},\,\mathbf{k})$ and the integral equation to be solved in order that we be able to modify this vertex is

$$\psi(\mathbf{k},\,-\mathbf{k}-\mathbf{q}) = \lambda\int[\tfrac{1}{2}\mathbf{k}^2+\tfrac{1}{2}(\mathbf{k}+\mathbf{q})^2-E]^{-\frac{1}{2}}[2\pi^2(\mathbf{k}-\mathbf{k}')^2]^{-1} \times[\tfrac{1}{2}\mathbf{k}'^2+\tfrac{1}{2}(\mathbf{k}'+\mathbf{q})^2-E]^{-\frac{1}{2}} \times\psi(\mathbf{k}',\,-\mathbf{k}'-\mathbf{q})\,d^3\mathbf{k}. \quad (24)$$

This equation may be solved in almost the same manner as that indicated in the first case, and (6) is

$$\langle\chi|\,v\,|\varphi_0\rangle = \langle\chi|\ V + V\frac{1-P_0}{E-H_0}V + (2\gamma)^2\sum_{\alpha,l,m}\frac{1}{\lambda_\alpha+1}\int d^3\mathbf{k}\,d^3\mathbf{k}'\,d^3\mathbf{s}\,d^3\mathbf{s}' \times V\mathbf{k}\bar{\varphi}^m_{\alpha l}(\mathbf{s})e^{-i\mathbf{s}\cdot\frac{1}{2}(\mathbf{k}-\mathbf{q})}\bar{\varphi}^{m*}_{\alpha l}(s')e^{i\mathbf{s}'\cdot\frac{1}{2}(\mathbf{k}-\mathbf{q})}k'V\ |\varphi_0\rangle, \quad (25)$$

where the eigenfunctions $\bar{\varphi}^m_{\alpha l}$ are formally the same as those exhibited in expression (14) except that the eigenvalues are now

$$\lambda_\alpha = 2(-E+\tfrac{1}{4}\mathbf{q}^2\alpha)^{\frac{1}{2}} \equiv 2\gamma s \quad (26)$$

and the transformation between ξ and s is similarly altered.

The self-consistency condition in all the cases cited is satisfied by using (6) to determine the true energy E corresponding to the exact wavefunction ψ and this is done by obtaining from Eq. (2) the

relation

$$E = \Delta E + E_0 = \frac{\langle\chi|\, H_0\,|\psi\rangle + \langle\chi|\, V\,|\psi\rangle}{\langle\chi\,|\,\psi\rangle} \equiv \frac{\langle\chi|\, H_0 v\,|\varphi_0\rangle + \langle\chi|\, v\,|\varphi\rangle}{\langle\chi|\, v\,|\varphi_0\rangle} \qquad (27)$$

and iterating the expression until the energy E so calculated is the same as that used in the energy denominators of the perturbation expansion.

It is clear that the vertex operator for one electron interacting with an external Coulomb potential may be modified in a similar manner.

We have found that the integrations and summations remaining in expression (15) can indeed be carried out analytically except for one remaining quadrature which may be evaluated numerically. However if the effect of the interaction V is to result in a binding energy it is clear that it needs be an attractive interaction which may be the sum of electron repulsion and attraction due an external Coulomb potential.

In the event that the final state is not a bound state, it is possible that our method can be used to find also explicitly (as an infinite sum) the S matrix for Coulomb scattering. The necessary eigenfunction solutions replacing (19) are then the continuum hydrogenic functions.[7]

[7] H. A. Bethe and E. E. Salpeter, in *Encyclopedia of Physics*, edited by S. Flügge (Springer-Verlag, Berlin, 1957), Vol. 35, p. 110.

Summerizing, we have demonstrated a new method of summing to infinite order, a class of contributions of infinite-order perturbation theory. Specific results are obtained for Coulombic interactions and a plane-wave representation and the effect is that of replacing the initial Coulombic interaction with a modified interaction representing the repeated interaction of two such particles. The matrix element of the modified interaction is given by expression (15), and for weakly bound particles with paired momenta and paired spin. The functions $\varphi_\alpha(s)$ are given explicitly by expression (19). In the remainder of the discussion, we discuss how the solutions to various other cases may also be recast in the form of the $\varphi_\alpha(s)$ functions and the final expressions for the matrix element are given explicitly.

In application of this method, we consider the different classes of diagrams representing the contributions of infinite-order perturbation theory, separate out those diagrams by using a modified vertex in the simplified diagrams. Finally, this method is not confined to plane-wave representations, and corresponding solutions for other representations would not be restricted to the discussion of weakly bound particles and would be interesting and useful.

ACKNOWLEDGMENT

The authors appreciate several informative conversations on the topic of this paper with J. McKenna.

Internal Symmetry and the Poincaré Group*

YUKIO TOMOZAWA
The Institute for Advanced Study, Princeton, New Jersey and Summer Institute for Theoretical Physics, University of Colorado, Boulder, Colorado
(Received 18 September 1964)

Theorems concerning the interrelation between the Poincaré group and the symmetry group of elementary-particle interactions have been proved under weaker conditions than that of McGlinn.

1.

THE possibility of combining the Poincaré group (≡ the inhomogeneous Lorentz group) and the symmetry group of elementary-particle interactions has been discussed by various authors.[1,2] Recently, McGlinn has proved[3] an interesting theorem concerning this problem: Assuming that

(I) A_j, $(j = 1, 2, \cdots, n)$, together with the generators of the Poincaré group, $M_{\mu\nu}$ and $P_\mu(\mu, \nu = 1, 2, 3, 4)$, form the bases of a Lie algebra L,

(II) Poincaré algebra P is a subalgebra of L, and that

(III) A_j's commute with all the $M_{\mu\nu}$,

it is concluded that (a) A_j's constitute a subalgebra A, and that (b) L is the direct sum of P and A, provided A is semisimple.

A motivation of the assumption (III) was that we wish to consider a symmetry group of the interactions, bases of which irreducible representation correspond to a multiplet with same spin parity. In order to acquire such a property of the symmetry group, however, it would be sufficient to assume that

(III′) A_j's commute with M_{12}, M_{23}, and M_{31}, the generators of the spacial rotations which correspond to the angular momentum operators.

In this note, we point out that a condition weaker than (III′),

(III″) A_j's commute with one of the generators $M_{\mu\nu}$,

leads to essentially the similar result as the McGlinn theorem.[4]

2.

According to the assumptions (II) and (III″), we have[5,6]

$$[M_{\lambda\mu}, M_{\nu\sigma}] = i(\delta_{\lambda\nu}M_{\mu\sigma} + \delta_{\mu\sigma}M_{\lambda\nu} - \delta_{\lambda\sigma}M_{\mu\nu} - \delta_{\mu\nu}M_{\lambda\sigma}), \quad (1)$$

$$[M_{\lambda\mu}, P_\nu] = i(\delta_{\lambda\nu}P_\mu - \delta_{\mu\nu}P_\lambda), \quad (2)$$

$$[P_\mu, P_\nu] = 0, \quad (3)$$

and

$$[A_j, M_{12}] = 0, \quad (4)$$

where we have chosen M_{12} as the generator which commutes with A_j's without loss of generality. Then the assumption (I) implies that

$$[A_j, M_{\lambda\mu}] = \tfrac{1}{2}a^{\nu\sigma}_{j,\lambda\mu}M_{\nu\sigma} + b^{\nu}_{j,\lambda\mu}P_\nu + c^{k}_{j,\lambda\mu}A_k \quad (5)$$

with

$$a^{\nu\sigma}_{j,12} = b^{\nu}_{j,12} = c^{k}_{j,12} = 0. \quad (6)$$

To have a consistent algebraic system, we must examine various Jacobi identities. Calculating the following

$$[[A_j, M_{\lambda\mu}], M_{\nu\sigma}] + [[M_{\lambda\mu}, M_{\nu\sigma}], A_j] + [[M_{\nu\sigma}, A_j], M_{\lambda\mu}] = 0,$$

we obtain

* Work supported by the National Science Foundation.

[1] F. Lurçat and L. Michel, Nuovo Cimento **21**, 574 (1961); L. Michel, Lecture note of the Istanbul Summer School, 1962 (to be published).

[2] A. O. Barut, Nuovo Cimento **32**, 234 (1964).

[3] W. D. McGlinn, Phys. Rev. Letters **12**, 469 (1964).

[4] For further developments of the McGlinn theorem, see F. Coester, M. Hamermesh, and W. D. McGlinn, Phys. Rev. **135**, B451 (1964); O. W. Greenberg, Phys. Rev. **135**, B1447 (1964); M. E. Mayer, H. J. Schnitzer, E. C. G. Sudarshan, R. Acharya, and M. Y. Han, "Concerning Space-Time and Symmetry Group" (preprint, 1964); A. Beskow and U. Ottoson, "On the problem of combining the inhomogeneous Lorentz group with a Lie group" (preprint, 1964); see also lectures given by L. Michel, M. Hamermesh, and M. E. Mayer at the Symposium on Lorentz group, Boulder, Colorado, 1964 (to be published).

[5] We use the pseudo-Euclidean form of the Minkowski metric such as used in W. Pauli, "Continuous Groups in Quantum Mechanics" (CERN preprint, 1956).

[6] Greek letters λ, μ, ν, and σ run over 1, 2, 3, and 4; Roman letters j, k, l, etc. run over 1, 2, $\cdots$, n; and the summation is assumed whenever repeated indices appear.

$$
\begin{aligned}
[A_i, M_{23}] &= a_i M_{31} + b_i M_{24} + c_i P_2, \\
[A_i, M_{31}] &= -a_i M_{23} - b_i M_{14} - c_i P_1, \\
[A_i, M_{14}] &= a_i M_{24} + b_i M_{31} + d_i P_1, \\
[A_i, M_{24}] &= -a_i M_{14} - b_i M_{23} + d_i P_2, \\
[A_i, M_{34}] &= d_i P_3 - c_i P_4.
\end{aligned} \tag{7}
$$

From these relations we get[7]

$$[A'_i, M_{\mu\nu}] = 0, \tag{8}$$

where

$$A'_i = A_i + ia_i M_{12} + ib_i M_{34} + ic_i P_3 + id_i P_4. \tag{9}$$

Then, the McGlinn theorem[3] leads to

$$[A'_i, P_\lambda] = e_i P_\lambda, \tag{10}$$

$$[A'_i, A'_k] = c^l_{ik} A'_l, \tag{11}$$

$$c^m_{ik} c^a_{ml} + c^m_{ki} c^s_{mj} + c^m_{li} c^s_{mk} = 0 \tag{12}$$

and

$$c^l_{ik} e_l = 0. \tag{13}$$

Thus we have proved the following theorem:

Theorem 1. Under the assumptions (I), (II), and (III''), we can find a set of generators A'_i, by suitable linear combinations of the given generators, which satisfies Eqs. (8)–(13).

The following theorems[8] can be proved readily.

Theorem 2. Under the assumptions (I), (II), and (III''), if $\{A_i\}$ forms a subalgebra A, then it has the same structure constants[9] c^l_{ik} as those of $\{A'_i\}$.

Proof. Using Eqs. (8)–(12), we obtain

$$
\begin{aligned}
[A_i, A_k] &= [A'_i, A'_k] \\
&\quad + i(c_i e_k - c_k e_i) P_3 + i(d_i e_k - d_k e_i) P_4 \\
&\quad + i(b_i d_k - b_k d_i) P_3 - i(b_i c_k - b_k c_i) P_4 \\
&= c^l_{ik} A_l + i c^l_{ik} a_l M_{12} + i c^l_{ik} b_l M_{34} \\
&\quad + i\{c^l_{ik} c_l + b_i d_k - b_k d_i + c_i e_k - c_k e_i\} P_3 \\
&\quad + i\{c^l_{ik} d_l - b_i c_k + b_k c_i + d_i e_k - d_k e_i\} P_4.
\end{aligned} \tag{14}
$$

Then, the assumption leads to

$$[A_i, A_k] = c^l_{ik} A_l \tag{15}$$

with

$$c^l_{ik} a_l = c^l_{ik} b_l = 0, \tag{16a}$$

$$c^l_{ik} c_l = -b_i d_k + b_k d_i - c_i e_k + c_k e_i \tag{16b}$$

and

$$c^l_{ik} d_l = b_i c_k - b_k c_i - d_i e_k + d_k e_i. \tag{16c}$$

Theorem 3. If the subalgebra A in the preceding theorem is semisimple, then $A_i = A'_i$ and $L = P \oplus A$.

Proof. According to the McGlinn theorem, Eqs. (13) and (16) imply that

$$a_i = b_i = c_i = d_i = e_i = 0$$

for the case where A is semisimple. From this and Eqs. (8)–(12), the theorem follows. (q.e.d.)

Note added in proof: Statement (a) of the first section of the present article has been assumed in the proof of the theorem in Ref. 3. For its proof, see the article of the author, "On the McGlinn Theorem" (Progr. Theoret. Phys., to be published).

ACKNOWLEDGMENTS

This work has been initiated by a stimulating conversation with Dr. A. DiGiacomo during the author's stay at Pisa and has been completed at the Summer Institute for Theoretical Physics, Boulder, Colorado. The author wishes to thank Professor L. A. Radicati (Pisa) and Professor W. E. Brittin (Boulder) for hospitality extended to him at both the Institutes, respectively. Thanks are also extended to Professor J. R. Oppenheimer for the hospitality at The Institute for Advanced Study.

[7] The author is indebted to Professor H. Morikawa for discussion on this point.

[8] Some related theorems have been discussed in a different manner by L. Michel at the Symposium on Lorentz group, Boulder, Colorado, 1964 (to be published).

[9] Physical systems classified by the representations of algebras $\{A_i\}$ or $\{A'_i\}$ are different, though they are mathematically isomorphic. However, the classification due to the algebra $\{A'_i\}$ is simpler than that due to $\{A_i\}$.

Necessary Condition on the Radial Distribution Function*

EUGENE FEENBERG
Wayman Crow Laboratory of Physics, Washington University, St. Louis, Missouri
(Received 11 September 1964)

Trial functions $g(r)$ may be used as radial distribution functions to study the ground-state properties of a uniform extended quantum fluid (liquid He, nuclear matter). A result obtained by Wigner and Seitz in the study of the charged electron gas imposes an integral inequality on any assumed $g(r)$. The relation is $\int_0^\infty [1 - g(r)]r\,dr/[\int_0^\infty [1 - g(r)]r^2\,dr]^{\frac{1}{2}} < 1.792/3^{\frac{1}{2}} = 1.243$. This inequality places an effective constraint on the location, magnitude, and width of the nearest-neighbor peak.

CONSIDER the standard many-particle problem of N identical interacting particles in a box of volume Ω (with $\rho = N/\Omega$ held constant and N allowed to increase without limit). Let $g(r)$ denote a trial function introduced to serve as radial distribution function for the ground state. The definition of a radial distribution function for a pure boson state,

$$\rho^2 g(\mathbf{r}_{12}) = N(N-1)\int |\psi(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \cdots, \mathbf{r}_N)|^2\, dv_{34\cdots N}, \quad (1)$$

requires the existence of a normalized symmetrical many-particle state function ψ capable of generating the proposed $g(r)$. For a pure fermion state the corresponding definition is

$$\rho^2 g(r_{12}) = N(N-1) \times \sum_{m_1,\cdots,m_N} \int |\psi(\mathbf{r}_1 m_1, \mathbf{r}_2 m_2, \mathbf{r}_3 m_3, \cdots, \mathbf{r}_N m_N)|^2\, dv_{34\cdots N} \quad (2)$$

implying the existence of a normalized antisymmetrical function ψ of the space and associated discrete coordinates of all the particles. The basic problem posed by Eqs. (1) and (2) is the determination of useful necessary conditions on any proposed $g(r)$ for the implied state functions to exist. The probability interpretation of $g(r)$ implicit in Eqs. (1) and (2) can be used to derive functional inequalities which must be satisfied by any function proposed for the role of radial distribution function. In general these inequalities limit the magnitude and width of the peaks and valleys exhibited by $g(r)$. Several necessary conditions of this type have been derived by Yamada[1] and by Garrod and Percus.[2]

An explicit inequality based on Eqs. (1) or (2) in a form illuminating and illustrating the arguments developed in Refs. 1 and 2 can be derived from the existence of an absolute potential-energy minimum in a classical physical problem. This problem is the charged gas in a box filled with a fixed uniform charge density of opposite sign and equal total amount of charge. Many years ago Wigner[3] called attention to the fact that the classical potential energy of the charged system attains an absolute minimum value when the particles are located at the lattice points of a body-centered cubic lattice. This statement is equivalent to the inequality

$$\rho^{-\frac{1}{2}} V_{\text{Coulomb}}(1, 2, \cdots, N) \geq -N(1.792e^2)/2r_s\rho^{\frac{1}{2}} = -Ne^2 0.896(4\pi/3)^{\frac{1}{2}}, \quad (3)$$

introducing the value at the minimum estimated by Wigner and Seitz[4] and by Fuchs.[5] In Eq. (3) r_s is the radius of a sphere of volume ρ^{-1}.

To obtain a functional inequality on $g(r)$ from Eqs. (1)–(3), I use the (presumed existing) normalized ψ to compute the expectation value of the Coulomb potential energy. The result is

$$\rho^{-\frac{1}{2}}\langle V_{\text{Coulomb}}\rangle = 2\pi N\rho^{\frac{1}{2}}\int_0^\infty [g(r) - 1]\frac{e^2}{r}r^2\,dr \geq -Ne^2 0.896(4\pi/3)^{\frac{1}{2}}, \quad (4)$$

or

$$\rho^{\frac{1}{2}}\int_0^\infty [1 - g(r)]r\,dr \leq \frac{1.792}{4\pi}\left(\frac{4\pi}{3}\right)^{\frac{1}{2}}. \quad (5)$$

The normalization condition on $g(r)$ implied by Eqs. (1)–(2),

$$\rho\int_0^\infty [1 - g(r)]r^2\,dr = 1/4\pi, \quad (6)$$

* Supported in part by the Office of Scientific Research Grant AFOSR-62-412, U. S. Air Force and the National Science Foundation Grant GP-3211.

[1] M. Yamada, Prog. Theoret. Phys. (Kyoto) **25**, 579 (1961).

[2] C. Garrod and J. K. Percus, J. Math. Phys. **5**, 1756 (1964).

[3] E. P. Wigner, Trans. Faraday Soc. **34**, 678 (1938).

[4] E. Wigner and F. Seitz, Phys. Rev. **46**, 509 (1934).

[5] K. Fuchs, Proc. Roy. Soc. (London) **A151**, 585 (1935).

may be combined with Eq. (5) to yield

$$\int_0^\infty [1 - g(r)]r\,dr \Big/ \left[\int_0^\infty [1 - g(r)]r^2\,dr\right]^{\frac{1}{2}} \leq \frac{1.792}{3^{\frac{1}{2}}} = 1.243; \quad (7)$$

Eq. (7) is the desired inequality.

[Equation (6) involves the asymptotic behavior $g(\infty) = 1 + O(N^{-\nu})$, $\nu > 1$, which can be shown to hold for the ground state when the interaction operator includes a strong short-range repulsive component (for any two particles approaching closely). Let X denote an arbitrary half-space (volume $= \frac{1}{2}\Omega$) in Ω; also $n(r) = 1$ in X, $n(r) = 0$ otherwise. $\mathfrak{N} = \sum n(r_l)$ is the particle number operator in X with mean value $\langle \mathfrak{N} \rangle = \frac{1}{2}N$. The result $\nu > 1$ follows from $\langle (\mathfrak{N} - \langle \mathfrak{N} \rangle)^2 \rangle \ll N$. The basic theorem on number fluctuations in X for the ground state is

$$\langle (\mathfrak{N} - \langle \mathfrak{N} \rangle)^2 \rangle = \rho^2 \int (1 - g(r_{12}))(1 - n(r_1))n(r_2)\,dv_1\,dv_2 \sim O(N^{2-\nu}) + \gamma N^{\frac{1}{2}},$$

where γ is proportional to the total surface area common to X and the complementary half-space (provided that the linear dimensions of the elements making up X are large compared to $\rho^{-1/3}$).]

Note added in proof: A more direct characterization of the asymptotic behavior of $g(r)$ follows from sum rules which require that the liquid structure function $S(k)$ vanish as $k \to 0$ [P. J. Price, Phys. Rev. **94**, 257 (1954)].

Trial functions in the form

$$g(r) = Z(\rho^{\frac{1}{2}}r/a; \mu_1, \cdots, \mu_p)$$

may be tested by Eq. (7). Failure of Eq. (7) for physically interesting values of the parameters μ is sufficient reason to discard any proposed $g(r)$. The plain implication of such failure is that no ψ function exists capable of generating the given $g(r)$ by Eqs. (1)–(2). Interesting values of μ are determined by using $g(r)$ to compute the expectation value of the physical Hamiltonian.[6-8] The minimum value of $\langle H \rangle$ in the μ space defines a point, $\mu = \mu^*(\rho)$, centrally located in the interesting region. In general Eq. (7) imposes a strong constraint on the form of $Z(s; \mu)$ and on the allowed range of μ values. Comparing integrands of the integrals occurring in Eqs. (5) and (6) we see that the former gives relatively more weight to the region below the first rising slope of $g(r)$ and relatively less weight to the nearest-neighbor peak. The two conditions together, as combined in Eq. (7), produce an effective constraint on the width of the region near the origin in which $g(r)$ is small and on the magnitude and width of the nearest-neighbor peak.[9]

I take pleasure in acknowledging instructive critiques by Dr. Henry Woodrow Jackson and Dr. Robert Griffiths. Thanks are also due Professor Wesley Brittin for the opportunity to participate in the summer program of theoretical physics at the University of Colorado.

[6] F. Y. Wu and E. Feenberg, Phys. Rev. **122**, 739 (1961).
[7] W. Massey, Phys. Rev. Letters **12**, 719 (1964).
[8] F. Y. Wu and E. Feenberg, Phys. Rev. **128**, 943 (1962).
[9] This simple qualitative characterization involves the implicit assumption that the oscillatory behavior of $g(r)$ beyond the nearest-neighbor peak is strongly damped.

High-Energy Behavior at Fixed Angle for the Five-Point Function in Perturbation Theory*

J. V. GREENMAN
Department of Applied Mathematics and Theoretical Physics, University of Cambridge, England
(Received 28 September 1964)

The validity of the d-line method used by Halliday in studying the high-energy behavior at fixed angle of the four-point function is examined in more detail. This is carried out within the context of determining the high-energy behavior at fixed angle for the five-point function. The leading asymptotic behavior of the sum of all planar five-point graphs in a ϕ^3 theory at fixed angle is s^{-2}.

INTRODUCTION

In studying the high-energy fixed-t behavior of ladder diagrams, Polkinghorne[1] showed that to obtain the leading asymptotic behavior one need only integrate over a small region around the zeros of the coefficient of s. Since the coefficient is in fact $\alpha_1 \alpha_2 \cdots \alpha_N$, the product of the α's of the rungs, a zero occurs whenever one of $\alpha_1, \cdots, \alpha_N$ is set zero. The integration procedure is then straightforward—one integrates successively the α's over a small region near the origin.

Halliday[2,3] has generalized the procedure for more complicated diagrams at fixed t and for planar graphs at fixed angle. The procedure is to find a minimum set of α's which set the coefficient zero when they themselves are set zero. One then integrates over a small region around the origin of these α's, making the approximation that only those terms of lowest order in the scaling variable are to be retained. This set is termed a d line since when traced on the graph it forms a connected line of length d, where length is given by the number of lines of the graph belonging to it. One then looks for another minimal set of α's which will set to zero the coefficient, now simplified by the omission of certain of the terms. It is not at all obvious whether this second set will be connected nor what length it will have. In fact in the fixed-angle case we can find disconnected sets. However, we show that, provided they do not form loops with the original d line, their length is bounded below by d. In determining our sequence of minimal sets we follow the procedure of Tiktopoulos.[4] We find that maximum sequence of minimal sets which do not form loops with one another. In Sec. II and the Appendix we prove that the length of these minimal sets can never be less than d. In the proof we need to introduce the notion of quasiconnected d line. We now integrate over all these minimal sets and find the asymptotic behavior as $s^{-d} \ln^{M-1} s$ where M is the number of sets in the sequence. If, however, the final integrand does not converge we need to study the singular configurations,[4] those sets of α's giving rise to a pole in the final integrand and to scale over these configurations as well as over the sequence of d lines. Section III is devoted to finding these configurations and Sec. IV to the rescaling procedure.

But even in graphs containing singular configurations the leading behavior is still $s^{-\alpha} \ln^{\beta} s$ where α, β are certain positive integers. The final crucial point is showing that for the subclass of planar graphs with minimal length d the power of the logarithm is bounded. Summing all planar graphs leads to an asymptotic series whose leading behavior is that of its first term, the Born term.

I. KINEMATICS AND THE TOPOLOGY OF FEYNMAN INTEGRALS

Let us first consider the kinematics of the two-to three-particle process in the center-of-mass frame. Figure 1 represents the 3-momentum of the particles. The 3-momenta of the three outgoing particles form a triangle and hence define a plane, the outgoing plane. Let E be the total energy of the system, and $\pi - \theta_4 (\pi - \theta_5)$ the angle the 3-momentum of particle 4 (particle 5) makes with that of particle 3. Thus the kinematics are completely

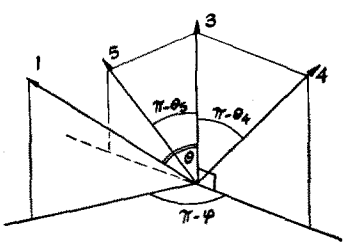

FIG. 1. Two to three-particle kinematics.

* The research reported in this document has been sponsored in part by the Air Force Office of Scientific Research, OAR, under Grant No. AF EOAR 63-79 with the European Office of Aerospace Research, U. S. Air Force.

[1] J. C. Polkinghorne, J. Math. Phys. **4**, 503, 1393 (1963).
[2] I. G. Halliday, Nuovo Cimento **30**, 177 (1963).
[3] I. G. Halliday, Ann. Phys. (N. Y.) (to be published).
[4] G. Tiktopoulos, Phys. Rev. **131**, 480 (1963).

determined if we are given E, θ_4, θ_5 and the angles θ, ϕ defining the line of the two incoming particles in relation to the outgoing plane. To be specific, let E_i, q_i be the energy and magnitude of the 3-momentum of the ith particle, then from the conservation laws and mass constraints we determine q_3, q_4, q_5 as

$$q_3 \simeq \frac{\sin(\theta_4 + \theta_5)E}{\sin\theta_4 + \sin\theta_5 + \sin(\theta_4 + \theta_5)} \equiv \lambda_3 E, \quad (1)$$

$$q_4 \simeq \frac{\sin\theta_5 E}{\sin\theta_4 + \sin\theta_5 + \sin(\theta_4 + \theta_5)} \equiv \lambda_4 E, \quad (2)$$

$$q_5 \simeq \frac{\sin\theta_4 E}{\sin\theta_4 + \sin\theta_5 + \sin(\theta_4 + \theta_5)} \equiv \lambda_5 E, \quad (3)$$

as E becomes large.

Hence, if p_i represents the ith particle's 4-momentum, we have

$$p_i p_{i+1} \simeq q_i q_{i+1}[1 - \cos(i, i+1)]$$
$$= \lambda_i \lambda_{i+1}[1 - \cos(i, i+1)]E^2 \equiv k_{i,i+1}E^2,$$

which tends to infinity with E. For the five-particle process we can take the five independent variables as $s_{i,i+1} \equiv (p_i + p_{i+1})^2$ $i = 1, \cdots, 5$. All these invariants tend to infinity as E.

In using perturbation theory for planar graphs we will be examining Feynman integrals in the form first derived by Chisholm.[5] This involves two functions C, D.

$C(\alpha)$ is a homogeneous polynomial in the α's of degree l where l is the number of loops in the diagram. There is a 1–1 correspondence between the terms of $C(\alpha)$ and the simply connected sets of lines of the graph. Each term in $C(\alpha)$ is the product of the α's of the lines that have to be removed to obtain the simply connected set associated with that term.

$D(\alpha) = s_{(p,q)}g_{(p,q)} + K$ where to each division of the $(p + q)$ external particles into two sets $(i_1 \cdots i_p; j_1 \cdots j_q)$ where $p \leq q$, $1 \leq p$ corresponds an invariant $s_{(p,q)} = (p_{i_1} + \cdots + p_{i_p})^2 = (p_{j_1} + \cdots + p_{j_q})^2$ in the center of mass. The $g_{(p,q)}$'s are homogeneous polynomials of order $l + 1$ in the α's. To each invariant corresponds a set of partitions of the graph which divide the graph into two connected parts; to one part are attached the $(i_1, \cdots, i_p)$ external particle lines and to the other part the remaining q external lines. (These partitions are not to cut lines of the graph more than once, enter loops more than once, nor to cut vertices.) The coefficient of $s_{(p,q)}$ is a sum of terms, each term corresponding to a particular partition of $s_{(p,q)}$ and is a product of the α's of the lines cut by partitions with the two C functions corresponding to the two connected parts separated by the partitions considered as separate graphs. $(-K)$ in the equal-mass case is $m^2C(\alpha)(\sum\alpha)$.

The general planar five-point graph gives an amplitude of the form,

$$K(l)\int_0^1 \cdots \int_0^1 \frac{C^l(\xi)\ \delta(\sum\xi - 1)\ \prod^r d\xi}{D^{l+2}(s, \xi)}, \quad (4)$$

where ξ is the generic symbol for the α's, $K(l)$ is a constant irrelevant for our discussion, and s denotes invariants. For the present kinematical problem we can express all invariants as proportional to s_{12} and consider the asymptotic form of the integral as s_{12} tends to infinity. In fact let us write $D \equiv sg(k_{i,i+1}, \alpha) + h$, where $s_{12} \equiv s$.

The procedure is this—we integrate only over those regions of the α's where $g \simeq 0$.[1] g is not always positive in the hypercontour and hence these regions can lie properly within the hypercontour. We can avoid g becoming zero here by distorting the contour except when this is prevented by the occurrence of a pinch.[1] Such pinches do occur in certain planar graphs. No rigorous argument has been given but examination of some specific graphs indicate that contributions from within the hypercontour do not dominate the contribution from the endpoint. We assume that this holds generally. Thus considering the endpoint contribution, suppose that, by setting p α's zero, g is zero. We integrate over a small neighborhood of the origin of the space of these p α's using for convenience the scaling transformation[6] which sweeps out the region of integration by planes parallel to the plane $\sum^p \alpha_i = 1$ rather than planes parallel to the axes. The leading asymptotic behavior is obtained when p is minimal[2] as can be verified by simple integration. The Feynman amplitude is transformed under the scaling transformation to the form

$$\int_0^1 \prod d\xi'\ \delta(\sum\xi' - 1)\int_0^1 \prod d\bar{\alpha}\ \delta(\sum\bar{\alpha} - 1)$$
$$\times \int_0^\epsilon \frac{d\rho\ \rho^{P-1}C'^l(\xi', \bar{\alpha})}{[\rho g' + h']^{l+2}}, \quad (5)$$

where ξ' is the generic symbol for the α's not scaled, C', h' are the C, h functions with only terms independent of ρ retained and g' is the g function with only linear terms in ρ retained. The variables $\bar{\alpha}_i$ arise in the scaling transformation as $\alpha_i = \rho\bar{\alpha}_i$ $(i = 1, \cdots, P)$ with the constraint $\sum_{i=1}^P \bar{\alpha}_i = 1$. ϵ is some small positive quantity. Now since the

[5] R. Chisholm, Proc. Cambridge Phil. Soc. **48**, 300 (1952).

[6] P. G. Federbush and M. T. Grisaru, Ann. Phys. (N. Y.) **22**, 263 (1963).

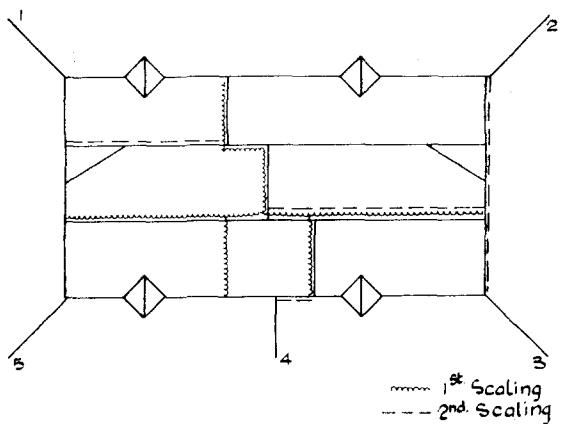

FIG. 2. A disconnected d line.

leading asymptotic behavior must be independent of ϵ we can take its value as large as we like with the same approximations made.[3] In fact it will be convenient to take $\epsilon = 1$.

We must now find these minimum sets which, following Halliday,[2] we call d lines, where $d = \min(p)$. Thus a d line is a minimal set of α's which reduce each term in g to zero when each of the (d) α's is zero. The solution is obvious when the problem is translated into a topological form. We need a minimum set of lines such that each partition cuts at least one of the lines. Hence a d line is a simply connected line to which belongs at least one vertex of each of the sections of the boundary.

II. DISCONNECTED d LINES

We now follow the method of Tiktopoulos[4] which requires us to scale in succession as many d lines as possible, provided

(i) in the set of d lines, the lines do not form loops with one another;

(ii) at any stage the α's in the new d line do not contain as a subset the α's contained in any one δ function.

In the first scaling, the d line is necessarily connected but this need not be so in any subsequent scaling. For consider the example in Fig. 2 where the first scaling is a connected d line of length 9. The second scaling set, a disconnected set of length 9, is an effective d line since it also reduces every term in g to zero when all the α's of the set are put equal to zero. On first looking at this graph this does not appear to be so, but we can check that if we find some partition associated with some invariant which does not cut any line of the second set then it cuts two or more lines of the first set. It is just these terms of the g function arising from these partitions that we ignore in our approximation. We call these disconnected sets, *disconnected d lines*. In future, reference to d lines implies both connected and disconnected d lines. It is an important question whether we can find a disconnected d line whose length is less than d.

Theorem 1. Disconnected d lines not forming loops with previous d lines can never have length less than d.

Proof. The proof is somewhat complicated and so we divide it into three sections.

(i) The case when we make a disconnected scaling after one connected scaling.

(ii) The case when we make a disconnected scaling after p connected d-line scalings not forming loops.

(iii) The case when we make a disconnected scaling after p d-line scalings (both connected and disconnected) not forming loops.

We study the first case here to introduce notation and to illustrate the method of argument in a simple case. The other two cases we consider in the Appendix.

(i) Suppose we have scaled a connected d line and in tracing it on the graph it divides the graph into five distinct areas. The d line itself is divided into seven sections which we now separate into two classes.

A section is a type I section if partitions of only one invariant cross it.

A section is a type II section if partitions of more than one invariant cross it.

The separation is meaningful since in our approximation only partitions which cross one section of the d line, hereafter referred to as "significant" partitions, need be considered. Hence each significant partition is associated with one section. A class I (class II) partition is one which is associated with a type I (type II) section. There are five sets of class I partitions each associated with an invariant. For each invariant $s_{i,i+1}$ the class I partitions define an area $A_{i,i+1}$ and an internal boundary line $x_{i,i+1}$. For instance, in Fig. 3, corresponding to s_{51}, $A_{51} = A_1 \cup A_5$ and $x_{51} = s_4 \cup s_1 \cup s_6$. To stop these partitions we need a continuous barrier between $x_{i,i+1}$ and the boundary b_i. The length of this barrier cannot be less than s_i by the minimal property of the d line. It is possible that such barriers can have a proper subset of α's in common. However, this does not alter our conclusion that the length of these barriers cannot be less than the sum of the

type I sections, otherwise we would contradict the minimal property of the d line.

There are two type II sections of our d line, s_6 and s_7. Consider first the class II partitions across s_6. If we do not put the barrier to stop these partitions along s_6 then the barrier must be in two sections, one in A_1, the other in A_4 as shown in Fig. 3. In this case the length of these barriers is not less than $s_4 + s_5 + s_6$, and class I partitions associated with s_4 and s_5 have also been stopped. Further in this case we are forced to stop the other set of class II partitions, namely those across s_7, with a barrier along s_7, otherwise we would form loops with the d line. Thus all class II partitions and the class I partitions of s_4, s_5 are stopped, and only the class I partitions of s_1, s_2, s_3 are not stopped. The length of the barbiers to stop these partitions is not less than $s_1 + s_2 + s_3$. Thus we have a lower bound of d on the length of the disconnected set. This conclusion is even simpler when barriers are put along both s_6 and s_7.

The argument is more straightforward and the conclusion unaltered when the original d line divides the graph into fewer than five areas.

In the Appendix we generalize the argument for case (ii). Case (iii) is essentially a corollary of (ii) when we realize that for each disconnected d line we can find a set of connected and quasiconnected d lines whose scaling eliminates the same partitions as the disconnected d line. A *quasiconnected d line* is a connected $(d + \alpha)$ line, in the rth type I section of which we have removed α_r lines where $\sum \alpha_r = \alpha$. It is only necessary to include a quasiconnected d line in the set in very special circumstances. See Fig. 18 and the discussion with Fig. 18. We now scale over as many such connected and disconnected d lines as possible and then integrate over all scaling parameters. If the maximum number of scalings is M using the integral relation[4]

$$\int_0^1 \cdots \int_0^1 \frac{\rho_1^{d-1} \cdots \rho_M^{d-1}\, d\rho_1 - d\rho_M}{[\rho_1 \cdots \rho_M g's + h']^{l+2}} \sim \frac{1}{s^d} (\ln s)^{M-1} \frac{1}{(h')^{l-d+2} g'^d} \quad \text{as} \quad s \to \infty, \qquad (6)$$

we have succeeded in separating out the s behavior from the final integral over the remaining variables. g', h' contain only the lowest-order terms from g, h. If the final integrand is nonconvergent this indicates that we are misled into believing this to be the leading asymptotic behavior. In this case we must study the poles of the final integrand.

The relation (6) is in fact true only if $l + 2 > d$,

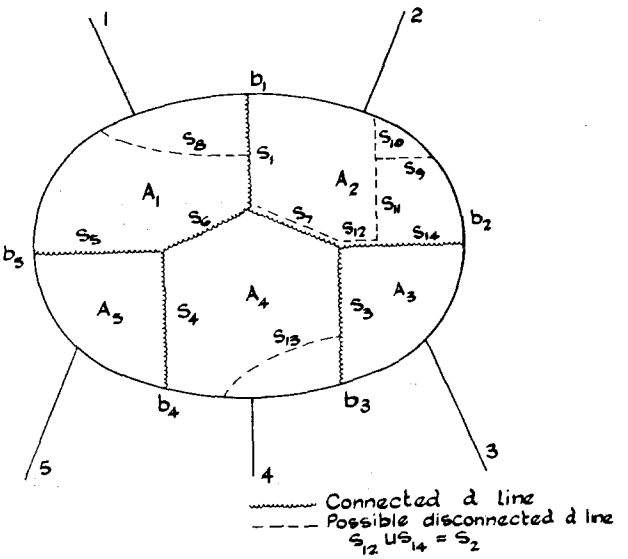

FIG. 3. A disconnected d line must have length not less than d.

but we can prove a lemma that $l + 2 \geq d$ for five-point planar graphs by a nontrivial extension of the argument used by Halliday[3] to prove a similar result for planar four-point graphs. We omit the proof. As in the four-point case, equality holds only for the ladder diagrams. These have been studied. A ladder with N rungs has asymptotic behavior, $s^{-(N+1)} \ln^{2N-3} s$.

III. SINGULAR CONFIGURATIONS

We wish to find those sets of α's which cause endpoint poles of the integrand. These sets of α's we term singular configurations.[4] Our study is conveniently divided into three sections.

1. Connected Singular Configurations

First, we consider only connected sets of α's forming singular configurations. Whatever α set we take C', h' will have the same order zero. Thus, suppose we take a set of L lines forming l_d loops with d lines of our sequence D and having l_r internal loops and then scale the α's, we find the power of the scaling parameter is $L - 1 + (l_d + l_r)(d - 2) - pd$ where p is the lowest power in g. Suppose $p = l_r + l_d$, that the number of vertices at which other lines of the graph meet the set is E', and only one d line is involved. Then the minimum value of L is $3l_r + E' + 2(l_d + 1) - 3$ giving rise to a pole of order $(2 - l_r - E')$. Consider those singular configurations involving only one section of the d line, then we generate the following family:

(A1) $p = l_r + l_d$; $l_r = 0$; $E' = 0$. We have the self-energy subgraph leading to a second-order pole (Fig. 4).

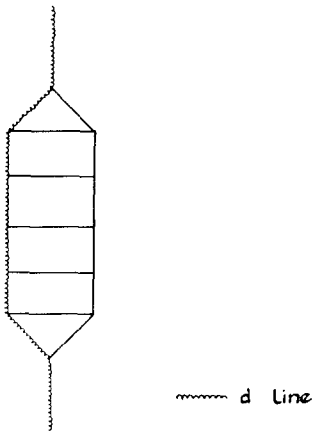

FIG. 4. A second-order pole.

(A2) $p = l_r + l_d$; $l_r = 0$; $E' = 1$. A first-order pole (Fig. 5).

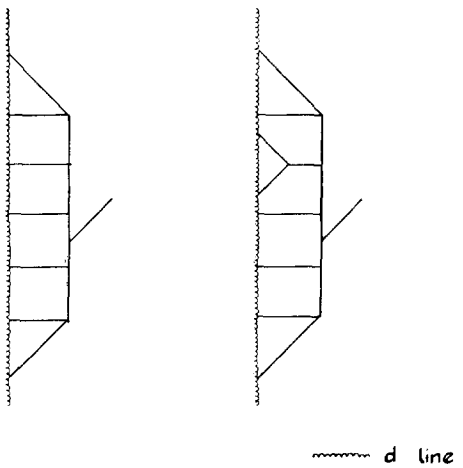

FIG. 5. A first-order pole.

(A3) $p = l_r + l_d$; $l_r = 1$; $E' = 0$. A first-order pole (Fig. 6).

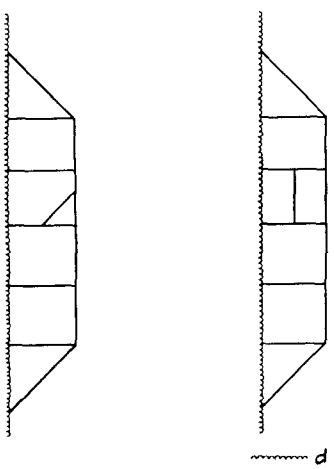

FIG. 6. A first-order pole.

When a junction point of the d line is involved, we can generate a much richer family of singular configurations.

(B1) $p = l_r + l_d$; $l_r = 0$; $E' = 0$ (Fig. 7).

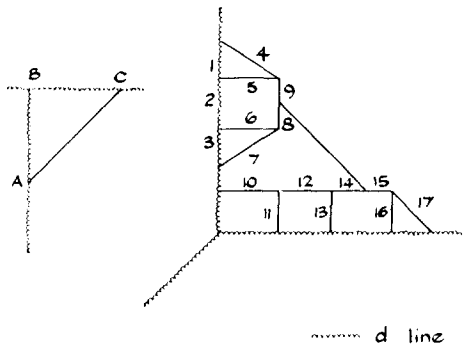

FIG. 7. Singular configurations peculiar to the fixed-angle limit.

As has been pointed out by Halliday[3] these singularities are peculiar to the fixed-angle limit problem. For instance the first singular configuration of Fig. 7 would imply a contradiction of the minimal property of the d line unless three sections of the d line met at B.

Class B2 (B3) is obtained by insertion of an external line vertex (internal loop) in class B1.

Similarly we can generate families (C1, C2, C3), (D1, D2, D3) when two, three junction points are involved. For instance in Fig. 8 we give an example

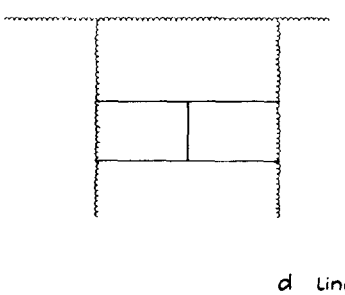

FIG. 8. A singular configuration involving two corners of a d line.

of a second-order singular configuration ($p = l_r + l_d$, $E' = 0$, $l_r = 0$) belonging to class C1.

Suppose now, still considering the case $p = l_r + l_d$, we look for connected singular configurations associated with more than one d line.

Theorem 2. Any connected singular configuration such that $p = l_r + l_d$ associated with more than one d line can be associated with not more than two connected or quasiconnected d lines.

Proof. We first note that (i) we can always find a connected or quasiconnected d line that contains as a

subset any section of a disconnected d line; (ii) for each d line and for each section of the boundary of the graph there corresponds a vertex which is the termination of a section of that d line.

As described before when we trace the d lines in our sequence on the graph, the graph is divided into areas, each area associated with some part of the external boundary. Any connected singular configuration must be wholly within one of these areas forming l_d loops with lines of the d lines forming the boundary of this area.

Let us suppose that the area we choose to study has no disconnected section of a quasiconnected d line along its boundary and that all the vertices referred to in (ii) lie properly within the section of the boundary.

We can check that if $p < l_r + l_d$ then the configuration is not singular. p is less than $l_r + l_d$ when a significant partition cuts two lines both of which (i) belong to the boundary of the area; (ii) belong to one of the l_d loops. Such lines are called "p lines." The configuration is singular only when there are no p lines. To examine the conditions when no p line exists let us examine Fig. 9. A is the area in which we look for a possible singular configuration. Let p_1, p_2 be two p lines. If p_1, p_2 belong to the same d line then (x_1) is not significant. Suppose not, then by (i) considering p_1, p_2 belonging to connected d lines α_1, α_2, we trace on the graph boundaries of regions B, C formed of the lines of α_1, α_2 only. In fact $c_2 \in d_1$, $c_2 \notin d_2$, $c_4 \in d_2$, $c_4 \notin d_1$. Because partitions (x_2), (x_3) do not intersect d_1, d_2 then between vertices

$$a_1,\ a_4(a_3,\ a_6)$$

on the external boundary there can either be at most one external particle line or four or more. Let the number be $l(m)$. By (ii), between a_1a_6 there must be at most two external lines. Let the number in fact be n. Then only the following possibilities exist:

	(a)	(b)	(c)	(d)
l	1	0	1	1
m	0	1	1	1
n	1	1	1	2 .

Case (*a*). Suppose there is an external particle vertex on b_1 then by (ii), there must be another external particle vertex on b_5. If instead the external particle vertex were on b_2 then (ii) can be satisfied.

Case (*b*). This is identical to case (a).

Case (*c*). The external particle vertex lies on b_3 and (ii) can be satisfied.

Case (*d*). Neither of the external particle vertices

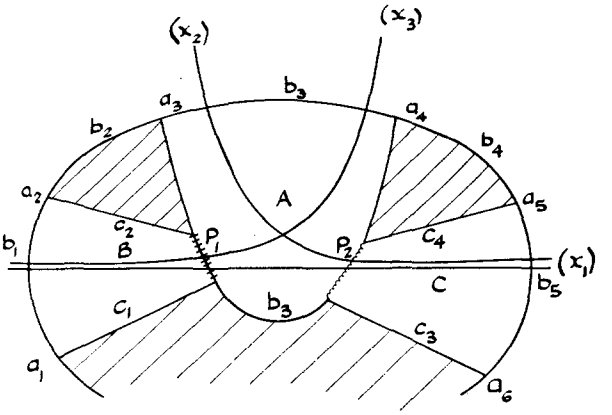

FIG. 9. A singular configuration is only associated with at most two d lines.

can be on b_3 for then one of the conditions $l \leq 1$, $m \leq 1$ is violated. Otherwise (ii) can be satisfied.

Now in case (c) and the second possibility of case (a) the partition (x_1) is not associated with a term in g. Thus in these cases there can be no significant partition cutting the area A from p_1 to p_2.

In the other cases either no p lines exist and the configuration is associated with one d line or there exists a significant partition which renders the configuration nonsingular. We have actually to prove that these partitions found in the above analysis are not eliminated when we trace all the other d lines in the regions B and C. Take the first possibility in case (a) where the external particle line on b_1 splits b_1 into a part b_1' adjacent to a_2 and b_1'' adjacent to b_2. Partition (x_1) must cut at least one line of all the other d lines. If it could not avoid cutting a d line twice then this implies that two sections of a d line terminate on the same section of the external boundary $(b_1''b_2b_3b_4b_5')$ contradicting (ii).

If the external particle lines lie on b_2, b_4 there must be a partition (x_1) otherwise the single connectedness of the tracing would be contradicted.

When $n = 1$ and the external particle vertex lies on $b_2b_3b_4$ let p_1, p_2 be the p lines adjacent to the two furthermost vertices of the boundary of A involved in the singular configuration. Because of (ii) $c_3 \in d_1$, $c_1 \in d_2$, and b_6 belongs to both d_1 and d_2 to avoid the formation of loops. We can define a new d line, d', containing p_1, b_6, p_2, c_2, c_4 except in the special configuration of d_1, d_2 in Fig. 10. In any case the singular configuration can be associated with at most two d lines.

We can remove the first restriction in the proof simply since quasiconnected d lines arise in very special circumstances. We examine the few ways quasiconnected d lines are incorporated in the com-

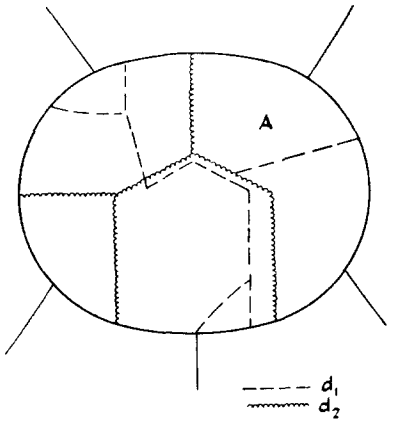

FIG. 10. The situation where the configuration is associated with two d lines.

plex of d lines and the theorem holds in these cases. An interesting case of a singular configuration is one associated with the disconnected section of a quasiconnected d line. An example is given in Fig. 11. Since a disconnected section of a quasiconnected d line necessarily lies along a section of a connected d line then this singular configuration is also associated with a connected d line.

Removing the second restriction is also straightforward.

2. Disconnected Singular Configurations

We can form "disconnected" singular configurations out of the classes A1–D3 in two basic ways—(a) we can string singular configurations on a d line (Fig. 12); (b) for each junction point we can fit at most two singular configurations in two corners (Fig. 13)—the final order of the pole being $\sum^N \alpha_r - (N-1)$ where there are N configurations involved, and α_r is the order of the rth configuration. We see here also how a $(d+\alpha)$ line $(\alpha > 0)$ can give rise to a singularity. If situation (a) is part of a $(d+2)$ line then this line is equivalent to a d line.

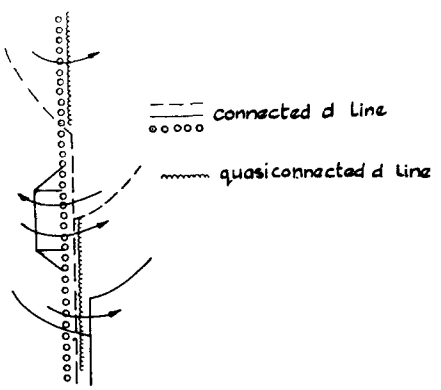

FIG. 11. A singular configuration associated with a quasiconnected d line.

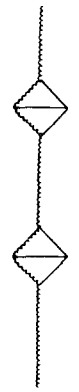

FIG. 12. Disconnected singular configurations.

3. d Lines Forming Loops with the Sequence D

We consider now the possibility $p > l_r + l_d$. In the fixed-t case, it is obvious that any connected d line forming loops with the sequence of d lines does not give rise to any new singular configuration if we scale the complete loops this d line forms with the other d lines.

In the fixed-angle case, the conclusion is not completely true. Consider a connected d line $(d_{l''})$ forming l'' loops with the sequence D, then there exists a significant partition such that $p \leq l''$. This is seen as follows. The partitions remaining after the sequence D can be divided into two classes:

Class α are those partitions which cross all the d lines on the same line θ of the graph.

Class β are those partitions which pass right across an A_i area, cutting all d-line sections to the external boundary defined by A_i.

Consider any class α partition and the two areas C, D defined by it. Suppose $d_{l''}$ forms l' loops with C and D then (i) the α partition cuts $l'+1$ sections of $d_{l''}$ implying that a section of $d_{l''}$ lies along θ; (ii) the α partition cuts l' sections of $d_{l''}$.

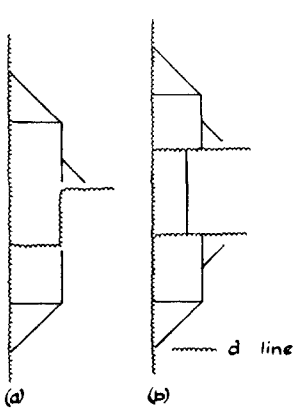

FIG. 13. Disconnected singular configurations peculiar to the fixed-angle case.

In the second case we have found a partition such that $p = l''$. In the first case $p = l'' + 1$. However, since $d_{1}{}''$ does not form loops with itself then in the first case we can see, by studying the class β partitions transversing the l'' loops, that there is a class β partition such that $p \leq l''$. If for $d_{1}{}''p$ is less than l'' there is no pole; if $p=l''$ there is no pole if $L\geq 2+1$, where L is the length of this d line. But this is a simple consequence of a ϕ^3 theory. However, if we scale over the d line and the lines of the loops, then $p = l'' + 1$ and we generate the singular configurations A1–D3 together with the "disconnected" configurations of (a) and (b) above. Similarly, if we had a $(d + 1)$ line then we generate the same family but the order of the pole is reduced by 1. The fact that a subconfiguration of the graph which would not be a singular configuration when d lines forming loops were not considered, but is when they are, is illustrated in Fig. 14.

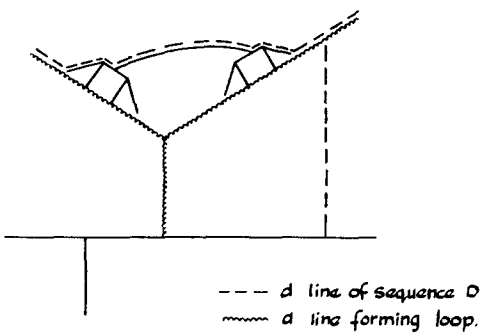

FIG. 14. d lines forming loops.

We now naturally ask whether disconnected sets giving rise to a zero of g' and forming l'' loops with the sequence D give poles. Here it is possible that $p = l'' + 1$ but then the length (L) of the set is greater than $d + 2l'' + 1$ and there is no pole. When $p = l''$ we have to show the inequality $L \geq 2l'' + 1$. In general, the length of such a disconnected set is less than d, but the fact that all partitions such that $p < l''$ must be excluded implies that with any loop formed, (i) we need two lines of the disconnected set to form it or (ii) there exists another line in the set not forming part of a loop.

Careful calculation gives the lower bound. If we include lines of the loop in our scaling set then the problem of finding singular configurations is immediately simplified since our set now consists of a disconnected d line not forming loops with the D sequence, together with loops. The problem then reduces to that of the connected-loop d line.

IV. THE RESCALING PROCEDURE

We must now go back to the original integrand and rescale. Each singularity is or can be associated with at most two connected or quasiconnected $(d + \alpha)$ lines where $\alpha \geq 0$.

First, for simplicity, we consider only those graphs where there is only one singularity on any $(d + \alpha)$ line. We have the following prescription: Examine whether any proper subset of the singular set is a singular set of the same order as the complete singular set.

(i) If there is no such subset, then scale the $(d + \alpha)$ line together with the complete singular set. For instance, in the example in Fig. 4, there is no singular subset of order 2. There is no such subset if the order of the pole is greater than 1.

(ii) If there are proper subsets of a first-order singular configuration also of the first order, then all such proper subsets together with the d line are to be scaled in succession. Because of the omission of terms in g, C, h at each scaling, new subsets leading to first-order poles are generated. Scale over these as for the other proper subsets. Any α's of the singular configuration not scaled are then scaled with the d line. Finally scale the d line. Menke[7] has studied certain special cases of singular configurations where the number of scalings is greater than that given by the prescription above due to the occurrence of disconnected scalings. An example of the prescription is given in the second diagram of Fig. 7. The subsets with the d line are scaled in the following sequence (4, 5), (6, 7), (10, 11), (16, 17), (12, 13), (8, 9, 14, 15, 18, 19).

(iii) If the singular configuration is associated with two $(d + \alpha)$ lines, scale one of the $(d + \alpha)$ lines first and then apply (i) or (ii) to the second $(d + \alpha)$ line and the singular configuration, depending on whether the configuration is of the first or second order. So effectively the singular configuration is associated with only one $(d + \alpha)$ line.

(iv) If the singulairty is associated with the disconnected part of a quasiconnected d line it is also associated with a connected d line. One scales the configuration with the connected d line first and then the configuration with the quasiconnected d line.

Since in any sequence of scalings only those scalings which have the minimum power of the scaling variable as a factor in the numerator contribute to the leading asymptotic behavior, we need only scale those configurations giving the highest-order pole.

In the general case, we find that subclass of $(d + \alpha)$ lines $(\alpha \geq 0)$ with which we have associated con-

[7] M. M. Menke, "High Energy Behaviour of Feynman Integrals Involving Singular Configurations," (Cambridge Preprint).

nected and "disconnected" singular configurations which give rise to the highest-order pole.

(i) If this order is greater than 1, then suppose there are q configurations of order greater than 1 associated with a particular d line and p the number of first-order poles. (a) If p is nonzero, then we use the procedure (ii) above where instead of the d line we read the d line together with the q singular configurations. (b) If p is zero, then we scale the d line together with the q singular configurations.

(ii) If the highest-order pole is one then we carry out the procedure in (ii) above.

Having performed these scalings, we now integrate according to Eq. (6) and the final integral must necessarily converge. The final asymptotic behavior is $s^{-d+\beta-1} \ln^{M-1} s$, where β is the order of the highest-order pole generated and M is the number of scalings on β-pole configurations.

By generating the highest-order poles for a d line one is easily convinced that β is bounded below by $d - 2$.

The argument is now as in the four-point case.[3] We show that for the subclass of graphs with minimal length d, the number of scalings is bounded, hence the power of the logarithm in the asymptotic behavior is bounded.

However, this is obvious since each d line must be connected to each of the five sections of the boundary and each singular configuration is tightly bound to the associated d lines.

Hence for a planar five-point graph in a ϕ^3 theory with minimal length d, the leading asymptotic behavior at fixed angle is proportional to $s^{-d-1+\beta} \ln^N s$ where N is bounded with d.

Summing all planar graphs we have a genuine asymptotic series whose leading asymptotic behavior is the leading asymptotic behavior of its leading term, namely the Born term which, as s becomes large, becomes proportional to s^{-2}.

ACKNOWLEDGMENTS

I would like to thank Dr. J. C. Polkinghorne for suggesting the problem and for many stimulating discussions. I would also like to thank the Department of Scientific and Industrial Research for the award of a maintenance grant.

APPENDIX

Theorem 1. We now give the proof of (ii).

We have made $p(\geq 2)$ connected d-line scalings, the first d line having divided the graph into five distinct regions. The d lines, when traced on the graph, define certain areas. Because no loops have been formed by the d lines, each area is associated with part of the external boundary line. Divide these areas into two classes:

A type A area is one which is associated with part of the external boundary not containing the vertex of an external particle line.

A type B area is one associated with part of the external boundary containing the vertex of an external particle line. Take any section of the boundary line and define an area A_i (where the suffix refers to the particular section b_i) as the union of all those type A areas associated with the section. Let the type B area associated with the jth particle line be labeled B_j.

We define these particular areas A_i, B_j because they make transparent the partition structure of the graph after p d-line scalings.

There are a finite number of ways of combining the ten areas A_i, B_j in the graph but avoiding loops. We must consider each possibility in turn and examine which partitions have been eliminated. We only consider one configuration of A_i, B_j here. Those configurations likely to give rise to new phenomena have also been studied and in these cases the theorem still holds. Consider the configuration in Fig. 15. Note that lines x_5, x_7, x_8, x_3 belong to all d lines. Let us examine the structure of two areas A_i, A_j which have part of their boundaries in common. Consider, for instance, A_3 and A_4. Now a d line can either send two separate barriers down from x_7, x_8 to b_4, b_3 (a d' line) or it can send one line down dividing into two branches to b_4, b_3 (a d'' line).

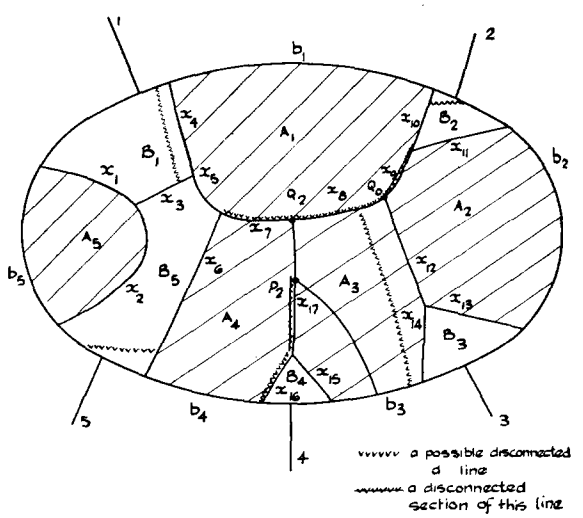

FIG. 15. A disconnected scaling set chosen after p connected d-line scalings.

Lemma 1. Given two d' lines (d_1', d_2'), y_1' the barrier from x_7x_8 to b_3 of d_1' and y_2' the barrier from x_7x_8 to b_4 of d_2'. Then y_1', y_2' cannot have a line in common.

Proof. By using the minimal properties of the two d lines, (Fig. 16).

If there is more than one d'' line, then (except for possibly one of them) we can divide the d'' lines into two classes.

Class L are those d'' lines whose barrier to $b_4(y_4'')$ branches off x_{17} before its barrier to $b_3(y_3'')$.

Class R are those d'' lines whose barrier to $b_3(y_3'')$ branches off x_{17} before its barrier to $b_4(y_4'')$.

Lemma 2. Class L and Class R d'' lines cannot overlap. We cannot have the sequences $y_4''^i\ y_3''^i$, $y_4''^i,\ y_3''^i$ and $y_4''^i,\ y_3''^i,\ y_3''^i,\ y_4''^i$ along x_{17} where $i \neq j$; $i, j = 1, 2$ indexing the two d'' lines.

Proof. By simple application of the minimal property of the two d lines (Fig. 17).

Corollary 1. Through every line of x_{17} there passes a partition, either of s_{45} or s_{34} or both, unless there is a d' line containing some lines of x_{17}.

Corollary 2. A d' line and a d'' line cannot overlap. Hence Corollary 1 is true for every line of x_{17}.

Suppose there is in fact a d' line lying along x_{17} with a barrier Y_1 to b_3. As we have seen from Lemma 2, only across certain lines of that part (x_{17}') of x_{17} lying between Y_1 and the apex of B_4 can partitions of s_{45} pass. Label these lines $a_1 \cdots a_r$; let the remaining lines of x_{17}' be $z_1 \cdots z_s$. Partitions of s_{43} cross the z lines and in general some of the a lines. Let the length of x_{16} be p; the length of $x_{17} - x_{17}' \equiv x_{17}''$ be q; the length of Y_1 be r and the length of the two barriers of the d' line containing Y_1 to b_3, b_4 be $x + q + r$, $x + y$.

Lemma 3. $p + \sum^r a \geq x$.

Proof. By a simple application of the minimal property of the d lines involved. Suppose there exists a d' line having its b_3 barrier (Y_2) partly along x_{12}. There are several ways of erecting a disconnected set of barriers stopping the partitions of invariants s_{34}, s_{45}. One way is to erect a continuous barrier

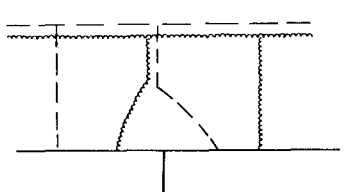

FIG. 16. Overlapping d lines.

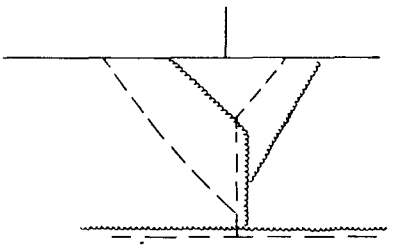

FIG. 17. Overlapping d lines.

from x_8 to b_3 and to put a barrier along the a lines of x_{17} and along x_{16}. By Lemma 3 the length of these barriers is not less than the length of a d'' line from Q_2 to b_3, b_4, namely $x + y$. Similarly, by the minimal property of the d lines and Lemma 3, we can show all other barriers stopping s_{34}, s_{45} partitions have this lower bound.

Similarly one studies the structure of A_1, A_2 and the same phenomena occurs. It is now simple to see the possible disconnected sets and to obtain a lower bound of d on their length. A typical example is shown in Fig. 15.

We now prove Theorem 1 (iii).

Consider Fig. 3 where we have performed one connected d-line scaling and then a disconnected d-line scaling.

Define the connected d lines

(1) (8, 1, 7, 3, 13, 12, 14),

(2) (4, 5, 6, 7, 3, 12, 11, 10, 9),

(3) (4, 5, 6, 1, 7, 12, 14, 15),

(4) (4, 5, 6, 7, 1, 3, 12, 14),

where the numbers refer to the indices of the s_i labels to the lines of the connected and disconnected lines. Every line of the original d line occurs three times in these sets (type a) except when the disconnected d line lies along the original d line when this line occurs four times in the sets (type b). Hence if we cross any type a line of the original d line we can

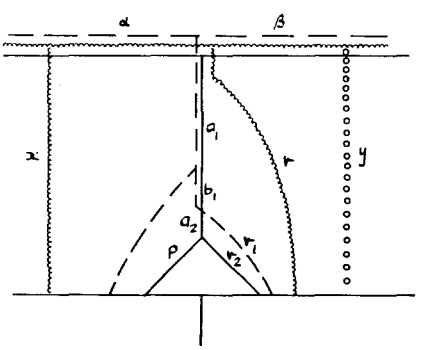

FIG. 18. Formation of a quasiconnected d line.

only cross that part of the fourth d line which is not common to the other three, namely across the disconnected set. If we cross a type b line then we cannot cross any other. Thus only those partitions are allowed where two lines of the original d line are not crossed and the disconnected set is crossed only once. These are precisely the partitions remaining after the connected and disconnected d-line scalings. In the general case we have seen that we can always find a connected d line containing any section of a disconnected d line unless we have placed a disconnected barrier along the common boundary of two adjacent A_i areas. In this case we need to define a quasiconnected d line as being a connected $(d + \alpha)$ line in the rth type I section, of which we have removed α_r lines where $\sum \alpha_r = \alpha$. Suppose, in particular, to stop partitions of invariants s_{34}, s_{45} we had erected the barriers described above, then the quasiconnected d line contains x_5, x_7, x_8, x_3 and these barriers. An example is given in Fig. 18. In the notation of Lemma 3 $p + \sum' a = x$ and $r = \sum' z + s$. In future scalings, we can never eliminate those partitions of s_{45} which cross x_{17} between P_2 and P_6 (where the continuous barrier from x_{17} to b_4 leaves x_{17}), since this would transform a lines into b lines. We could then rederive the inequality of the lemma as $r \geq \sum'' z + s$ where $s' > s$ leading to a contradiction with the equality above. The structure of this part of the graph therefore becomes fixed. Hence after more connected d-line scalings we have exactly the same choice of erecting barriers to b_3, b_4 as before the first disconnected set.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 6, NUMBER 4 APRIL 1965

ERRATA

The Representations of the Inhomogeneous Lorentz Group in Terms of an Angular Momentum Basis, J. S. LOMONT AND H. E. MOSES [J. Math. Phys. **5**, 294 (1964)]; **Representation of the Inhomogeneous Lorentz Group in Terms of an Angular Momentum Basis: Derivation for the Cases of Nonzero Mass and Zero Mass, Discrete Spin,** J. S. LOMONT AND H. E. MOSES [J. Math. Phys. **5**, 1438 (1964)]. The mass μ in the formulas of the above papers should be replaced by $\lambda\mu$, where λ is the sign of the energy.

only cross that part of the fourth d line which is not common to the other three, namely across the disconnected set. If we cross a type b line then we cannot cross any other. Thus only those partitions are allowed where two lines of the original d line are not crossed and the disconnected set is crossed only once. These are precisely the partitions remaining after the connected and disconnected d-line scalings. In the general case we have seen that we can always find a connected d line containing any section of a disconnected d line unless we have placed a disconnected barrier along the common boundary of two adjacent A_i areas. In this case we need to define a quasiconnected d line as being a connected $(d + \alpha)$ line in the rth type I section, of which we have removed α_r lines where $\sum \alpha_r = \alpha$. Suppose, in particular, to stop partitions of invariants s_{34}, s_{45} we had erected the barriers described above, then the quasiconnected d line contains x_5, x_7, x_8, x_3 and these barriers. An example is given in Fig. 18. In the notation of Lemma 3 $p + \sum' a = x$ and $r = \sum' z + s$. In future scalings, we can never eliminate those partitions of s_{45} which cross x_{17} between P_2 and P_6 (where the continuous barrier from x_{17} to b_4 leaves x_{17}), since this would transform a lines into b lines. We could then rederive the inequality of the lemma as $r \geq \sum'' z + s$ where $s' > s$ leading to a contradiction with the equality above. The structure of this part of the graph therefore becomes fixed. Hence after more connected d-line scalings we have exactly the same choice of erecting barriers to b_3, b_4 as before the first disconnected set.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 6, NUMBER 4 APRIL 1965

ERRATA

The Representations of the Inhomogeneous Lorentz Group in Terms of an Angular Momentum Basis, J. S. LOMONT AND H. E. MOSES [J. Math. Phys. **5**, 294 (1964)]; **Representation of the Inhomogeneous Lorentz Group in Terms of an Angular Momentum Basis: Derivation for the Cases of Nonzero Mass and Zero Mass, Discrete Spin,** J. S. LOMONT AND H. E. MOSES [J. Math. Phys. **5**, 1438 (1964)]. The mass μ in the formulas of the above papers should be replaced by $\lambda\mu$, where λ is the sign of the energy.