JOURNAL OF
MATHEMATICAL
PHYSICS

VOLUME 7, NUMBER 6 JUNE 1966

Unified Unitary Representation of the Poincaré Group for Particles of Zero and Positive Rest Mass

A. CHAKRABARTI
Centre de Physique Théorique de l'Ecole Polytechnique 17, rue Descartes, Paris V, France
(Received 1 November 1965)

As is well known, Wigner's construction of the unitary representations in terms of little groups gives quite dissimilar forms for the cases $m = 0$ and $m > 0$, while the spinor representation which does give a unified description of the above two cases is not unitary. We point out that it is quite possible to give a unified unitary representation for both cases (Sec. 1). This is achieved simply by noting that, while the factorization of $U(\Lambda)$ corresponding the choice of the little group of (1, 0) is not valid for $m = 0$, the explicit final expression for the Wigner rotation R_W and the corresponding infinitesimal generators remain perfectly well defined for $m = 0$, and continue to furnish a unitary (discrete spin) representation for this case, which is compatible with the restriction of helicity to a particular fixed value. Moreover, the representation thus obtained has a very simple and direct geometrical significance. The relation of our formulation with that of Wigner is studied (in Sec. 2) and the comparison with the spinor representations is given (Sec. 3). We rederive our representation, starting from a particular simple condition (4.1) (in Sec. 4) which holds for both cases ($m = 0$, $m > 0$). We then consider (Sec. 5) the application of our unified formulation to the reduction of direct products, involving particles of positive and zero rest mass, comparing the result with that of helicity coupling. In the Appendix we make certain remarks concerning the Hermiticity of the generators and the possibility of defining a position operator for $m = 0$. Comparison with Foldy's representation is given (in the Appendix and Sec. 3), explaining why his representation cannot be considered to be strictly unitary, though the relation with the unitary case is quite a simple one.

I. INTRODUCTION

LET us start by fixing our notations and conventions for the well-known case of Wigner's unitary or canonical representation for positive rest mass.[1-3] Elsewhere,[4,5] we have already discussed at length the various aspects of this case. Here we simply write down the results for the irreducible representation $[m, S]$. The infinitesimal generators are

$$P^0[=(\mathbf{P}^2+m^2)^{\frac{1}{2}}],\ \mathbf{P},$$
$$\mathbf{N} = -iP^0\frac{\partial}{\partial \mathbf{P}} - \frac{\mathbf{P}\times\mathbf{S}}{P^0+m}, \tag{1.1}$$
$$\mathbf{M} = -i\mathbf{P}\times\frac{\partial}{\partial \mathbf{P}} + \mathbf{S}.$$

For a finite transformation we have

$$U(a, \Lambda)\,|p\rangle = e^{i\Lambda\cdot p\cdot a}\,|\Lambda\cdot p\rangle\cdot D^{(s)}(R_W). \tag{1.2}$$

Here, when Λ is a pure rotation R, R_W coincides with R; and when Λ is a pure Lorentz transformation corresponding to the 4-velocity u, R_W again represents a rotation about the axis $\mathbf{p}\times\mathbf{u}$ through an angle ω, such that[4d,5]

[1] E. P. Wigner, Ann. Math. 40, 149 (1939).
[2] Yu. M. Shirokov, Zh. Eksperim. i Teor. Fiz. 33, 1196 (1957) [English transl.: Soviet Phys.—JETP 6, 919 (1958)].
[3] H. Joos, Fortschr. Phys. 10, 65 (1962).
[4] A. Chakrabarti, (a) J. Math. Phys. 4, 1215 (1963); (b) *ibid.* 4, 1223 (1963); (c) *ibid.* 5, 922 (1964), (d) *ibid.* 5, 1747 (1964.
[5] A. Chakrabarti, thesis, University of Paris (1965).

$$\sin\omega = \frac{2b}{b^2 + |\mathbf{p}\times\mathbf{u}|^2}\,|\mathbf{p}\times\mathbf{u}|, \tag{1.3}$$

where

$$b = (u^0 + 1)(p^0 + m) + \mathbf{u}\cdot\mathbf{p}.$$

In both cases $\langle\mathbf{S}\rangle$, the expectation value of the spin operator, undergoes the same rotation R_W,

$$\langle\mathbf{S}\rangle' = R_W\cdot\langle\mathbf{S}\rangle. \tag{1.2'}$$

As is well known, for $m > 0$, R_W can be written as

$$R_W = (\Lambda_{(\Lambda\cdot p)}\cdot\Lambda\cdot\Lambda^{-1}_{(p)}), \tag{1.4}$$

where

$$\Lambda_{(p)}\cdot p = (m, \mathbf{0}). \tag{1.4'}$$

This factorization corresponds to Wigner's choice of the little group leaving (1, **0**) invariant. Since the transformations $\Lambda_{(p)}$ are not well defined for $m = 0$, such a factorization is not valid for this case and so Wigner chose (for $m = 0$) another little group, leaving (1, 0, 0, 1) invariant.

All this is well known. Let us, however, note one simple but crucial fact: *Though the factorization* (1.4) *is not valid for* $m = 0$, *the total rotation* R_W *remains perfectly determinate if one substitutes directly* $m = 0$ *in the result* (1.3). In fact, as may be verified, this gives

$$\sin\omega = |\mathbf{p}\times\mathbf{p}'|/|\mathbf{p}|\,|\mathbf{p}'| \qquad (p' = \Lambda\cdot p). \tag{1.5}$$

(The joint effect of two not properly defined factors $\Lambda^{-1}_{(p)}$ and $\Lambda_{(\Lambda\cdot p)}$ leaves us with something well defined.)

Thus R_W now has the simple geometrical significance of the angle turned through by $\mathbf{p}$ as a consequence of the transformation (whether Λ be a pure rotation or a pure Lorentz transformation). This fact is also not unknown. But this fact gives us the welcome opportunity (which does not seem to have been utilized before) of constructing unitary representations for both cases ($m > 0$ and $m = 0$) in a unified fashion. This leads to interesting applications (see Sec. 5). It is to be noted that no particularly delicate limiting process is required. We may just put $m = 0$ in (1.1) and (1.3), where no discontinuity is to be found.

In fact, if in (1.1) we put $m = 0$, and write

$$\begin{aligned} &P^0[= (\mathbf{P}^2)^{\frac{1}{2}}], \mathbf{P}, \\ &\mathbf{N} = -iP^0\frac{\partial}{\partial\mathbf{P}} - \frac{\mathbf{P}\times\mathbf{S}}{P^0}, \\ &\mathbf{M} = -i\mathbf{P}\times\frac{\partial}{\partial\mathbf{P}} + \mathbf{S}, \end{aligned} \tag{1.6}$$

all the commutation relations continue to be satisfied and there is no additional difficulty, since the point $p^0 = 0$, i.e., the vector (0, 0, 0, 0) is not in the orbit of lightlike 4-momenta ($p^2 = 0$, $p^0 \neq 0$). Also, in view of (1.5) we may write now for finite transformations

$$\begin{aligned} &U(a, \Lambda)\,|p\rangle \\ &= e^{ip'\cdot a}\,|p'\rangle\cdot D^{(s)}[R(\hat{p}, \hat{p}')] \qquad (p' = \Lambda\cdot p), \end{aligned} \tag{1.7}$$

where $R(\hat{p}, \hat{p}')$ represents in all cases the rotation undergone by $\mathbf{p}$ (irrespective of the change in $\mathbf{p}^2$ for pure Lorentz transformations). It is to be understood that we are always talking about the case of finite and discrete spin. The singling out of a particular helicity value presents, of course, no problem at all. Since all the generators (1.6) now commute with the "helicity" operator

$$(\mathbf{S}\cdot\mathbf{P})/|\mathbf{P}|, \tag{1.8}$$

we may multiply (without disturbing the commutation relations) each of the generators P, $\mathbf{N}$, $\mathbf{M}$ by a projection operator for a definite possible value of helicity, to obtain a representation effectively equivalent to an one-dimensional one. Or equivalently and more simply we may just define that they act on the invariant space formed by the kets

$$\frac{\mathbf{S}\cdot\mathbf{P}}{|\mathbf{P}|}\,|p, \sigma\rangle = \sigma\,|p, \sigma\rangle \qquad (\sigma = -s, \cdots, s). \tag{1.9}$$

Also, directly from (1.6), we have

$$\omega = (\mathbf{P}\cdot\mathbf{M}, P^0\mathbf{M} - \mathbf{P}\times\mathbf{N}) = \frac{\mathbf{S}\cdot\mathbf{P}}{|\mathbf{P}|}P. \tag{1.10}$$

The explicit equivalence and transformation to one-dimensional form and comparison with the $(2S + 1)$-dimensional spinor representations are carried out in the following two sections (2 and 3). In Sec. 4, we indicate how one can derive the representation (1.6)–(1.7) by starting with the one chosen restriction

$$U(\Lambda)\,|p\rangle = |p'\rangle \qquad (p' = \Lambda\cdot p) \tag{1.11}$$

when $\mathbf{p}\times\mathbf{p}' = 0$ (for arbitrary spin and helicity). It will be noted that the restriction imposed by Wigner for $m > 0$, namely

$$U(\Lambda_{(p)})\,|p\rangle = |(m, \mathbf{0})\rangle, \tag{1.12}$$

automatically ensures (1.11) (for any spin projection). *But while* (1.12) *cannot be applied for* $m = 0$, *the condition* (1.11) *works perfectly well for both cases, and thus leads us to a unified description of the two.*

Practical application of this unification is briefly discussed in Sec. 5. Here we may note that quite apart from possible utility (1.7) has an agreeably simple and direct geometrical significance. Invariance of $\mathbf{S}\cdot\mathbf{P}/|\mathbf{P}|$ (for $m = 0$) demands that, if $|\Lambda\cdot p\rangle$ in (1.7) is multiplied by a rotation matrix, the rotation in question must be $R(\hat{p}, \hat{p}')$ [as in (1.7)] so that the polarization $\langle\mathbf{S}\rangle$ may follow p without the "lag" typical of the $m > 0$ case. The transformation (1.7) corresponds exactly to

$$\langle\mathbf{S}\rangle' = R(\hat{p}, \hat{p}')\cdot\langle\mathbf{S}\rangle. \tag{1.7'}$$

A few points concerning the Hermicity of the generators (1.6) and the possible definition of a position operator are discussed in the Appendix.

II. A UNITARY TRANSFORMATION TO ONE-DIMENSIONAL FORM

Wigner[1] chose (for $m = 0$) the little group of $p_{(0)} = (1, 0, 0, 1)$ and a representation such that

$$U(\Lambda_{p_{(0)}\leftarrow p})\,|p\rangle = |p_{(0)}\rangle. \tag{2.1}$$

Hence in order to pass to this form from (1.6) and (1.7), we have to transform by the *operator* corresponding to

$$D^{(s)}(\hat{p}, \hat{p}_{(0)}) \tag{2.2}$$

(i.e., the one obtained by replacing $\mathbf{p}$ by the operators $\mathbf{P}$). This gives us the unitary transforming operator

$$U = \exp\left\{i\,\frac{P^1S^2 - P^2S^1}{[(P_1)^2 + (P_2)^2]^{\frac{1}{2}}}\,\theta\right\}, \tag{2.3}$$

where

$$\theta = \tan^{-1}[(P_1)^2 + (P_2)^2]^{\frac{1}{2}}/P^3.$$

We note that we have the same $D^{(S)}$ and the same U, if we consider instead of $\Lambda_{p_{(0)}\leftarrow p}$, *the pure rotation*,

$$R(\hat{p}, \hat{\pi}), \quad \text{where} \quad \pi \equiv (|\mathbf{p}|, 0, 0, |\mathbf{p}|) \tag{2.4}$$

(this is to be compared with situation in Sec. 3).

We have, (along with $P_{tr} = P$)

$$\mathbf{M}_{tr} = U\mathbf{M}U^{-1} = -i\mathbf{P} \times \frac{\partial}{\partial\mathbf{P}} + \boldsymbol{\zeta}, \tag{2.5}$$

$$\mathbf{N}_{tr} = U\mathbf{N}U^{-1} = -iP^0\,\frac{\partial}{\partial\mathbf{P}} + \mathbf{n}, \tag{2.6}$$

where

$$\boldsymbol{\zeta} \equiv (P^1/P^0 + P^3, P^2/P^0 + P^3, 1)S^3, \tag{2.7}$$

$$\mathbf{n} \equiv (-P^2/P^0 + P^3, P^1/P^0 + P^3, 0)S^3 = \mathbf{p}_{(0)} \times \boldsymbol{\zeta}. \tag{2.8}$$

Also,

$$U(\mathbf{S}\cdot\mathbf{P}/|\mathbf{P}|)U^{-1} = S^3 \tag{2.9}$$

so that

$$W_{tr} = S^3P, \tag{2.10}$$

and corresponding to (1.9) we have

$$S^3\,|p, \sigma\rangle_{tr} = \sigma\,|p, \sigma\rangle_{tr}. \tag{2.11}$$

Thus we have diagonalized the helicity operator [as shown by (2.9)] and corresponding to each eigenvalue, we have one one-dimensional representation. [The operator U was used in a somewhat different context elsewhere.[6]]

The generators (2.5), (2.6) have been given by Lomont and Moses,[7] who have also given, for the more general case, including continuous spin, their relations with the two-dimensional Euclidean group introduced by Wigner [1]for $m = 0$. Here we establish their equivalence with (1.6), constructing the required transformation explicitly.

We may note that Shirokov[2] has given a different form of the generators for the one-dimensional case, namely

$$\begin{aligned}\mathbf{M} &= -i\mathbf{P} \times \frac{\partial}{\partial\mathbf{P}} \\ &\quad \pm \Sigma\left[\frac{P^0P^1}{(P_1)^2 + (P_2)^2}, \frac{P^0P^2}{(P_1)^2 + (P_2)^2}, 0\right], \\ \mathbf{N} &= -iP^0\,\frac{\partial}{\partial\mathbf{P}} \\ &\quad \pm \Sigma\left[\frac{-P^2P^3}{(P_1)^2 + (P_2)^2}, \frac{P^3P^1}{(P_1)^2 + (P_2)^2}, 0\right].\end{aligned} \tag{2.12}$$

However, acting on an eigenstate of momentum $|p\rangle$, terms like

$$\frac{P^1}{(P_1)^2 + (P_2)^2}, \frac{P^2}{(P_1)^2 + (P_2)^2} \tag{2.12'}$$

give coefficients which tend to infinity as $1/\epsilon$ when p^1 and p^2 tend to zero as ϵ (and $p^3 \to p^\circ \neq 0$). Since the states $|p^0, 0, 0, p^\circ\rangle$ certainly cannot be excluded and since all the group-theoretical requirements are satisfied by the well-behaved forms (2.5), (2.6), it seems hardly desirable to introduce (2.12).

III. COMPARISON WITH THE SPINOR REPRESENTATION

The generators $\mathbf{N}$, $\mathbf{M}$ for the two fundamental inequivalent spinor representations of $(2S + 1)$ dimensions are

[6] A. Chakrabarti, Nuovo Cimento **18**, 617 (1960).
[7] J. S. Lomont and H. E. Moses, J. Math. Phys. **3**, 405 (1962).

$$\mathbf{N}_s = -iP^0 \frac{\partial}{\partial \mathbf{P}} \pm i\mathbf{S}, \qquad \mathbf{M}_s = -i\mathbf{P} \times \frac{\partial}{\partial \mathbf{P}} + \mathbf{S}. \tag{3.1}$$

As is well known, (3.1) does provide a unified description for both cases, $m > 0$ and $m = 0$, though the representations are not unitary. (The two representations are generally coupled for $m > 0$ for invariance under parity.)

We have discussed elsewhere[4,5] the question of equivalence of (1.1) and (3.1) for $m > 0$. Also in Ref. 5, the explicit reduction of (3.1) to the one-dimensional form (for $m = 0$) has been carried out with the help of exactly the same operator U (2.13) of Sec. 2. (As explained later on, the significance of U does not remain quite the same in the two cases.)

As it is quite interesting to compare the results with those of Sec. 2, we briefly recapitulate some of them. For $\mathbf{M}$ we have exactly the same form as in (2.5). However, for $\mathbf{N}$ we have

$$U\mathbf{N}_s U^{-1} = \left(-iP^0 \frac{\partial}{\partial \mathbf{P}} + \mathbf{n}\right) \pm i \frac{\mathbf{P}}{P^0} S^3 + \mathbf{n}_\pm, \tag{3.2}$$

where

$$\mathbf{n}_\pm = \left[\hat{\eta}_\pm - \frac{\pm iP^1 + P^2}{P^0(P^0 + P^3)}(S^1 \pm iS^2)(\mathbf{P} + \boldsymbol{\pi})\right] \tag{3.3}$$

and

$$\hat{\eta}_\pm = (\pm i, 0, 0)(S^1 \pm iS^2) \tag{3.3'}$$

$$= R(\hat{p}, \hat{\pi}) \cdot \mathbf{n}_\pm \quad [(2.4) \text{ gives } \boldsymbol{\pi} \text{ and } R(\hat{p}, \hat{\pi})]. \tag{3.3''}$$

Also,

$$U\omega_s U^{-1} = S^3 P \mp iP^0(0, \mathbf{n}_\pm). \tag{3.4}$$

Thus we see that the transformation of $(\mathbf{N}_S, \mathbf{M}_S)$ by the same U gives not only all the terms of (2.5), (2.6) but also the extra terms

$$\left(\mathbf{n}_\pm \pm \frac{i\mathbf{P}}{P^0} S^3\right). \tag{3.5}$$

The nature of these extra terms are discussed at length in Ref. 5. Here we may mention that, when along with the generators (3.1) we impose the helicity restrictions

$$\frac{\mathbf{S}\cdot\mathbf{P}}{|\mathbf{P}|} |p, \pm s\rangle = \pm s\, |p, \pm s\rangle \tag{3.6}$$

and hence obtain in the transformed representation

$$S^3 |p, \pm s\rangle_{tr} = \pm s\, |p, \pm s\rangle_{tr}, \tag{3.7}$$

we can simply suppress the terms $\mathbf{n}_\pm$, since

$$\mathbf{n}_+ |p, +s\rangle = 0 = \mathbf{n}_- |p, -s\rangle. \tag{3.8}$$

It is to be noted that (3.8) *is not true for any other helicity* σ *such that* $\sigma < S$ *for* $\mathbf{n}_+$ *and* $\sigma > -S$ *for* $\mathbf{n}_-$. *This is related to the fact that the generators* (3.1) *leave invariant only the subspaces* $|+S\rangle$ *and* $|-S\rangle$, *respectively, whereas the generators* (1.6) *leave invariant each of the subspaces* $\sigma(-S \leq \sigma \leq S)$ *separately.* This is a fundamental difference.

Even if we get rid of the terms $\mathbf{n}_\pm$ by considering only the spaces $|+S\rangle$ and $|-S\rangle$ [corresponding respectively to $\pm i\mathbf{S}$ in (3.1)], there still remains the term

$$(i\mathbf{P}/P^0)S. \tag{3.9}$$

The addition of such a term amounts to a transformation of $\mathbf{N}$ and $\mathbf{M}$ by $(P^0)^S$, so that we have

$$U\mathbf{N}_s U^{-1} = (P^0)^S\left(-iP^0 \frac{\partial}{\partial \mathbf{P}} + \mathbf{n}\right)(P^0)^{-S} \tag{3.10}$$

[where we have now put $\mathbf{n}_\pm \simeq 0$ for the restricted case (3.6)–(3.8)].

Such a transformation, though of a simple form, is, of course, not unitary. Thus even in the case of fully polarized particles (2.6) and (3.2) are not strictly equivalent and hence neither are (1.6) and (3.1).

The change in the scalar product due to such a nonunitary transformation [by $(P^0)^\lambda$, say] is sometimes absorbed by a change of normalization of the states $|p\rangle$ by a factor $(p^0)^\lambda$. Such a prescription is acceptable, but is equivalent to a nonunitary transformation.

[As pointed out in Ref. 5, and also in the Appendix, a similar feature is involved in the transformation of Foldy.[8] The following comments are also relevant in this correction.]

The lack of strict unitary equivalence may also be viewed as follows.

If the starting point is a nonunitary representation such as (3.1) and we transform it with a unitary matrix U [such as (2.3) or the transformation of Foldy for the Dirac equation] then the resulting representation is naturally also a nonunitary one. That is why, in the case $m > 0$, we transformed the spinor representation with the help of a matrix (4.5) which is not unitary in the usual sense, though leaving invariant a spinor scalar product defined with the metric Γ^0 in the general case (Appendix A, of Refs. 4d and 5). Thus a physical equivalence could be established with the unitary canonical

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

form (1.1). The relation of this transformation to the pure Lorentz transformation $\Lambda_{(p)}[\Lambda_{(p)}\cdot p=(m,\mathbf{0})]$ was also noted very explicitly. This operator is, of course, no longer available for $m = 0$.

For the case $m = 0$, it is interesting to note that *while for the unitary representation* (1.6), U (2.3) *may be considered as the "operator" form of the matrix corresponding to both the transformations* (2.1) *and* (2.4), *for the spinor* (3.1) *only the definition* (2.4) *of* U *holds.*

IV. DERIVATION OF THE REPRESENTATION FROM THE CONDITION

(4.1): We now propose to show how one can derive the explicit forms of the generators and, from them, the formulas corresponding to finite transformations, by starting with the one fundamental criterion that *we must have, for any spin*

$$U(\Lambda)\,|p\rangle = |\Lambda\cdot p\rangle, \tag{4.1}$$

where Λ *is a pure Lorentz transformation parallel to* $\mathbf{p}$.

For a $(2S + 1)$-dimensional representation for a particle of zero rest mass, we may at once write the generators of translation and rotations in the usual forms

$$P^0[=(\mathbf{P}^2)^{\frac{1}{2}}],\ \mathbf{P}, \qquad \mathbf{M} = -i\mathbf{P}\times\frac{\partial}{\partial\mathbf{P}} + \mathbf{S}. \tag{4.2}$$

Our problem is to find the form of $\mathbf{N}$ that satisfies (4.1) along with the group commutation relations.

We see that (4.1) is satisfied automatically if $\mathbf{N}$ has the form

$$\mathbf{N} = -iP^0\frac{\partial}{\partial\mathbf{P}} + \mathbf{0}\times\mathbf{P} \tag{4.3}$$

since $(\mathbf{0}\times\mathbf{p})\cdot\hat{n} = 0$ when $\hat{n}\,||\,\mathbf{p}$ and hence in $e^{i\chi\mathbf{N}\cdot\hat{n}}\,|p\rangle$, the term containing the effect of spin drops out.

We now have to find the explicit form of $\mathbf{0}$ with the help of the commutation relations.

We suppose first that $\mathbf{0}$ commutes with $\mathbf{P}$ and the orbital part of $\mathbf{M}$ and proceed to examine if a solution can be found in this case. We find that the commutation relations

$$[N^i, M^i] = i\epsilon_{ijk}N^k \tag{4.4}$$

are all satisfied if

$$[0^i, S^i] = i\epsilon_{ijk}0^k. \tag{4.5}$$

So let us now put

$$\mathbf{0} = \lambda\mathbf{S}, \tag{4.6}$$

where we suppose λ (commuting with $\mathbf{Mor}$ and $\mathbf{P}$) to be a function of P^0 only and see if such a form is compatible with the relations

$$[N^i, N^j] = -i\epsilon_{ijk}M^k. \tag{4.7}$$

We find that these relations are satisfied if

$$\frac{1}{\lambda^2}\frac{\partial\lambda}{\partial P^i} = -\frac{P^i}{P^0} \qquad (i = 1, 2, 3) \tag{4.8}$$

and

$$2\lambda P^0 - \lambda^2(P^0)^2 = 1. \tag{4.9}$$

Hence

$$\lambda = 1/P^0. \tag{4.10}$$

The commutation relations with $\mathbf{P}$ are, of course, satisfied so long as $\mathbf{0}$ commutes with $\mathbf{P}$.

With this form of $\mathbf{N}$, we have for an *infinitesimal* pure Lorentz transformation $[d\chi = th^{-1}(dv)]$

$$e^{i\mathbf{N}\cdot\hat{n}d\chi}\,|p\rangle = e^{-i\frac{(\mathbf{p}\times\mathbf{S})}{p^0}\cdot\hat{n}\,dx}\,|p''\rangle = e^{id\chi\frac{(\mathbf{p}\times\hat{n})}{p^0}\cdot\mathbf{S}}\,|p''\rangle, \tag{4.11}$$

where $\mathbf{p}'' = \mathbf{p} + p^0 d\chi\hat{n}$ ($\hat{n}$ being an arbitrary direction). Hence $|p'\rangle$ is multiplied by a rotation matrix $D^{(S)}$ corresponding to the infinitesimal rotation

$$d\theta = \sin(d\theta) = \frac{d\chi}{p^0}\,|\mathbf{p}\times n| = \frac{|\mathbf{p}\times\mathbf{p}''|}{|\mathbf{p}|\,|\mathbf{p}''|} \tag{4.12}$$

(up to first order in $d\chi$)
about the axis $\mathbf{p}\times\mathbf{p}''$.

Now if we continue to add infinitesimal Lorentz transformation always parallel to $\hat{n}$ (*in which case the parameters* $d\chi$ *and* $d\theta$ *are additive*), for a finite transformation Λ we just have a rotation

$$\sin\theta = |\mathbf{p}\times\mathbf{p}'|/|\mathbf{p}|\,|p'| \qquad (p' = \Lambda\cdot p) \tag{4.13}$$

about the axis $\mathbf{p}\times\hat{n}$. This is the required result.

Here we have considered specifically the case $m = 0$. The corresponding derivations for the case $m > 0$ are quite similar, though somewhat less simple.

V. APPLICATIONS

Elsewhere[4c,d,5] we have discussed a relatively convenient method for reducing direct products of irreducible representations of the Poincaré group, when $m > 0$ for each component representation.

Our preceding unified unitary representation allows us to extend this technique to include zero

mass particles. Let us consider the simplest non-trivial case consisting of three representations $[m_i, S_i]$ $(i = 1, 2, 3)$, such that

$$m_1, m_2 > 0 \quad \text{and} \quad m_3 = 0.$$

An examination of the technique employed shows that we can write the relativistic Clebsch–Gordan coefficients for this case (using similar notations) as

$$\left\langle p_1\sigma_1; p_2\sigma_2; p_3\sigma_3 \,\middle|\, \begin{matrix} 0, \Sigma[m, S] \\ S'S''l'l''l \end{matrix} \right\rangle$$
$$= \rho_{(12)3}\left(\frac{2l''+1}{4\pi}\right)^{\frac{1}{2}}$$
$$\times \langle \sigma_1\sigma_2\sigma_3 m'0 \mid (S'S''l'l''l)S\Sigma\rangle y_{l'm'}(\hat{\pi}_{12}), \quad (5.1)$$

where

$$\mathbf{p}_1 + \mathbf{p}_2 + \mathbf{p}_3 = 0, \qquad p_3^x = p_3^y = 0,$$
$$\pi_{12} = \Lambda(p_1 + p_2)\cdot(p_1 + p_2),$$
$$m_{12}^2 = (p_1 + p_2)^2, \qquad m^2 = (p_1 + p_2 + p_3)^2 > 0,$$
$$\rho_{(12)3} = 2m_{12}^{\frac{1}{2}}\lambda^{-\frac{1}{4}}(m_{12}^2, m_1^2, m_2^2)\cdot 2m^{\frac{1}{2}}(m^2 - m_{12})^{-\frac{1}{2}},$$

and

$$\langle \sigma_1\sigma_2\sigma_3 m'0 \mid (S'S''l'l''l)S\Sigma\rangle$$
$$\equiv (S_1\sigma_1, S_2\sigma_2 \mid S'\sigma')(S'\sigma', S_3\sigma_3 \mid S''\sigma'')$$
$$\times (l'm'l''0 \mid lm')(lm'S''\sigma'' \mid S\Sigma). \quad (5.2)$$

The reduction of S-matrix elements and calculation of the density matrix and multipole parameters may be carried out as before.[4c,d,5]

In contrast with helicity coupling for three-particle states,[9] our formula (5.1) contains no $D^{(S)}$ matrices. In fact this feature persists if we continue to add more and more particles of positive rest mass. While the Clebsch–Gordan coefficients (5.2) can directly be substituted from the tables and provide no serious complications, not only are our Y_{lm}'s intrinsically simpler than the D matrices, but, in addition, their arguments are determined by pure Lorentz transformations instead of more complicated Wigner rotations which intervene for the D matrices. This is an advantage, even apart from the fact that the explicit considerations of the orbital contributions may have an intrinsic interest in many cases.

If there are two zero mass particles in the system considered, then it is more convenient to couple them separately and then couple the irreducible components to the particles of positive rest mass.

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

APPENDIX

Here we discuss certain points concerning the Hermiticity of the generators (1.6) with a chosen scalar product and the problem of the definition of a position operator for $m = 0$.

We define the normalization of the kets and the corresponding scalar product as

$$\langle p \mid p'\rangle = 2p^0\delta(\mathbf{p} - \mathbf{p}') \quad (A1)$$

and

$$\langle \Psi\Phi\rangle = \int \frac{d\mathbf{p}}{2p^0}\langle \Psi p\rangle\langle p\Phi\rangle = \int \frac{d\mathbf{p}}{2p^0}\Psi^*(p)\Phi(p). \quad (A2)$$

[We suppress the helicity indices, since they will be the same for all the kets in the space defined by (1.9).]

We now note the relations

$$\langle p\mid i\frac{\partial}{\partial\mathbf{P}}\mid p'\rangle = i2p^0\delta'(\mathbf{p} - \mathbf{p}') \quad (A3)$$

and

$$\langle p'\mid i\frac{\partial}{\partial\mathbf{P}}\mid p\rangle^* = i2p^{0'}\delta'(\mathbf{p} - \mathbf{p}') \quad (A4)$$
$$= i2\left[p^0\delta'(\mathbf{p} - \mathbf{p}') + \frac{\mathbf{p}}{p^0}\delta(\mathbf{p} - \mathbf{p}')\right]. \quad (A5)$$

(A5) follows on using the identity

$$[f(\xi') - f(\xi)]\delta'(\xi - \xi') = f'(\xi)\delta(\xi - \xi'). \quad (A6)$$

(A4) and (A6) have for consequence, as may easily be verified, that given (A1) the operator $i\partial/\partial\mathbf{P}$ is not Hermitian but only the combination

$$\left[i\frac{\partial}{\partial\mathbf{P}} - \frac{i\mathbf{P}}{2(P^0)^2}\right] \equiv \mathbf{X}, \quad \text{say.} \quad (A7)$$

The same conclusion naturally also follows from (A3) by noting that

$$\int \psi^*\left(i\frac{\partial}{\partial\mathbf{p}}\varphi\right)\frac{d\mathbf{p}}{2p^0}$$
$$= \int \left\{\left[i\frac{\partial}{\partial\mathbf{p}} - \frac{i\mathbf{p}}{(p^0)^2}\right]\psi\right\}^*\varphi\frac{d\mathbf{p}}{2p^0}, \quad (A8)$$

(assuming as usual that $\psi^*\varphi/p^0$ vanishes at the limits). Hence, since

$$-\tfrac{1}{2}i(P^0\mathbf{X} + \mathbf{X}P^0) = iP^0\partial/\partial\mathbf{P} \quad (A9)$$

we find that $\mathbf{N}$ (as well as $\mathbf{M}$) in (1.6) is already Hermitian and it would be wrong to symmetrize the first term into

$$-\frac{i}{2}\left(P^0\frac{\partial}{\partial\mathbf{P}} + \frac{\partial}{\partial\mathbf{P}}P^0\right) = -iP^0\left[(P^0)^{-\frac{1}{4}}\frac{\partial}{\partial\mathbf{P}}(P^0)^{\frac{1}{4}}\right]. \quad (A10)$$

The same considerations hold, of course, for the case $m > 0$ and the generators (1.1). We take the trouble to draw attention to this fact since the form (A10) appears in the formulation of Foldy[8] and, following him, in those of many other authors. In fact, the symmetrization (A10) corresponds just to a nonunitary transformation such as considered in (3.10) (now with $S = -\frac{1}{2}$) and the comments in that section may be noted in this connection.

Next, we come to the related question of definition of a relativistic position operator. We have noted elsewhere (4.4) that **X** (A7) gives a satisfactory definition of a Hermitian relativistic position operator for $m > 0$.

If we try to adopt the same definition for the representation (1.6), (1.9) we run into a difficulty since (1.9) is not invariant under the action of **X**. (This difficulty is not present for $m > 0$.) On the other hand, if we use the representation defined by (2.5), (2.6), and (2.11) and adopt again the definition **X**, the above difficulty no longer exists since (1.9) has now been replaced by (2.11) and S_3 commutes with **X**. It should be noted, however, that though we now have

$$[X^i, P^j] = i\delta_{ij}, \qquad i[P^0, \mathbf{X}] = \mathbf{P}/P^0 \qquad \text{(A11)}$$

we no longer have

$$[X^i, M^j] = i\epsilon_{ijk}X^k \qquad \text{(except for } \sigma = 0\text{)}. \qquad \text{(A12)}$$

Hence, if (A12) is also imposed as well as (A11) [along with the requirement of Hermiticity and the invariance of the space (2.11)], then **X** is no longer a satisfactory definition.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 7, NUMBER 6 JUNE 1966

Generalized Discrete–Continuum Radial Integrals with Coulomb Functions*

H. Barry Bebb
Institute of Optics, University of Rochester, Rochester, New York
(Received 9 November 1965)

Simple closed-form computational formulas involving only elementary functions are obtained for the hydrogenic, discrete–continuum matrix elements $<k, l'|r^m|n, l>$. The results allow for different effective charge parameters Z and Z' for the discrete and continuum functions. The central radial integral computed is sufficiently general to allow the evaluation of matrix elements of $r^m \exp(-r/\epsilon)$ between the Coulomb continuum functions and any discrete function formed from $r^k \exp(-r/A)$. Results are also given for the free-particle limit as $Z' \to 0$.

1. INTRODUCTION

THE elegance and generality of Gordon's formulas[1] for (discrete–discrete, discrete–continuum, and continuum–continuum) dipolar matrix elements in some ways obviate their utility. Given a particular set of quantum numbers, a numerical result is realized only after considerable algebra. Stobbe[2] has carried out the computation for a few of the discrete-continuum matrix elements. More recently, Menzel[3] has devised a simple method of obtaining generalized radial integrals with hydrogenic functions. However, his final results are of the same form as Gordon's (though more general) and do not provide convenient formulas for the discrete–continuum matrix elements.

In the present work, we obtain new closed-form computational formulas for the discrete–continuum Coulomb matrix elements, generalized to include the (multipole) matrix elements r^m assuming different effective charges Z and Z' for the discrete and continuum functions, respectively. In Sec. 2 we record the discrete and continuum functions in rather diverse representations. We then give computational formulas for the discrete–continuum matrix elements in Sec. 4. In Sec. 5 we consider the free-particle limit as Z' goes to zero. Selected numerical results are presented in Sec. 6. The mathematical details of evaluating the central integral arising in the matrix elements are relegated to the Appendix.

* This research was supported in part under contract with the Army Research Office, Durham.

[1] W. Gordon, Ann. Physik. (5), 2, 1031 (1929).
[2] M. Stobbe, Ann. Physik (5) 7, 661 (1930).
[3] D. Menzel, Rev. Mod. Phys. 36, 613 (1964). See also A. Burgess, Monthly Notices of Roy. Astron. Soc. 118, 477 (1958) for an approximate calculation and additional references.

The same considerations hold, of course, for the case $m > 0$ and the generators (1.1). We take the trouble to draw attention to this fact since the form (A10) appears in the formulation of Foldy[8] and, following him, in those of many other authors. In fact, the symmetrization (A10) corresponds just to a nonunitary transformation such as considered in (3.10) (now with $S = -\frac{1}{2}$) and the comments in that section may be noted in this connection.

Next, we come to the related question of definition of a relativistic position operator. We have noted elsewhere (4.4) that **X** (A7) gives a satisfactory definition of a Hermitian relativistic position operator for $m > 0$.

If we try to adopt the same definition for the representation (1.6), (1.9) we run into a difficulty since (1.9) is not invariant under the action of **X**. (This difficulty is not present for $m > 0$.) On the other hand, if we use the representation defined by (2.5), (2.6), and (2.11) and adopt again the definition **X**, the above difficulty no longer exists since (1.9) has now been replaced by (2.11) and S_3 commutes with **X**. It should be noted, however, that though we now have

$$[X^i, P^j] = i\delta_{ij}, \qquad i[P^0, \mathbf{X}] = \mathbf{P}/P^0 \qquad \text{(A11)}$$

we no longer have

$$[X^i, M^j] = i\epsilon_{ijk}X^k \qquad \text{(except for } \sigma = 0\text{)}. \qquad \text{(A12)}$$

Hence, if (A12) is also imposed as well as (A11) [along with the requirement of Hermiticity and the invariance of the space (2.11)], then **X** is no longer a satisfactory definition.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 7, NUMBER 6 JUNE 1966

Generalized Discrete–Continuum Radial Integrals with Coulomb Functions*

H. Barry Bebb
Institute of Optics, University of Rochester, Rochester, New York
(Received 9 November 1965)

Simple closed-form computational formulas involving only elementary functions are obtained for the hydrogenic, discrete–continuum matrix elements $<k, l'|r^m|n, l>$. The results allow for different effective charge parameters Z and Z' for the discrete and continuum functions. The central radial integral computed is sufficiently general to allow the evaluation of matrix elements of $r^m \exp(-r/\epsilon)$ between the Coulomb continuum functions and any discrete function formed from $r^k \exp(-r/A)$. Results are also given for the free-particle limit as $Z' \to 0$.

1. INTRODUCTION

THE elegance and generality of Gordon's formulas[1] for (discrete–discrete, discrete–continuum, and continuum–continuum) dipolar matrix elements in some ways obviate their utility. Given a particular set of quantum numbers, a numerical result is realized only after considerable algebra. Stobbe[2] has carried out the computation for a few of the discrete-continuum matrix elements. More recently, Menzel[3] has devised a simple method of obtaining generalized radial integrals with hydrogenic functions. However, his final results are of the same form as Gordon's (though more general) and do not provide convenient formulas for the discrete–continuum matrix elements.

In the present work, we obtain new closed-form computational formulas for the discrete–continuum Coulomb matrix elements, generalized to include the (multipole) matrix elements r^m assuming different effective charges Z and Z' for the discrete and continuum functions, respectively. In Sec. 2 we record the discrete and continuum functions in rather diverse representations. We then give computational formulas for the discrete–continuum matrix elements in Sec. 4. In Sec. 5 we consider the free-particle limit as Z' goes to zero. Selected numerical results are presented in Sec. 6. The mathematical details of evaluating the central integral arising in the matrix elements are relegated to the Appendix.

* This research was supported in part under contract with the Army Research Office, Durham.

[1] W. Gordon, Ann. Physik. (5), 2, 1031 (1929).
[2] M. Stobbe, Ann. Physik (5) 7, 661 (1930).
[3] D. Menzel, Rev. Mod. Phys. 36, 613 (1964). See also A. Burgess, Monthly Notices of Roy. Astron. Soc. 118, 477 (1958) for an approximate calculation and additional references.

2. COULOMB FUNCTIONS

The behavior and form of the solutions of the one-electron wave equation $H\psi = (\mathbf{p}^2/2m - Ze^2/r)\psi = E\psi$ is rather different for negative and positive energies. The negative energy (discrete) solutions common to H, $|\mathbf{l}|^2$, and l_z are $R_{n,l}(\rho)Y_l^m(\theta,\phi)$ with the radial functions defined by

$$R_{n,l}(\rho) = N_{n,l}\, e^{-\rho/2} \sum_{t=0}^{n-l-1} (-1)^t\, d_t \rho^{t+l}, \tag{1a}$$

where

$$N_{n,l} = [(2Z/n)^3(n+l)!\,(n-l-1)!/2n]^{\frac{1}{2}}, \tag{1b}$$

and

$$d_t = [(n-l-1-t)!\,(2l+1+t)!\,t!]^{-1}. \tag{1c}$$

In Eqs. (1), $\rho = 2Zr/n$ and r is measured in units of a_0.

The positive energy (continuum) solutions are most usefully written as a partial wave expansion in terms of the eigenfunctions $R_{l'}(\gamma, kr)Y_{l'}^{m'}(\theta,\phi)$ common to H, $|\mathbf{l}|^2$, and l_z,[4]

$$|\pm\mathbf{k}\rangle = 4\pi \sum_{l'=0}^{\infty} \sum_{m'=-l'}^{l'} i^{l'}\, e^{i\eta_{l'}^{\pm}} \times R_{l'}(\gamma, kr)Y_{l'}^{m'}(\theta,\phi)Y_{l'}^{m'*}(\theta_k,\phi_k), \tag{2}$$

where $\eta_{l'}^{\pm} = \arg\Gamma(l'+1\mp i\gamma)$, $\gamma = Z'/k$, and θ_k, ϕ_k specify the direction of k. Here k is measured in units of a_0^{-1} and r in units of a_0. We have, for additional generality, assumed a different effective charge Z' for the continuum states (though the continuum functions are not orthogonal to the discrete functions for $Z' \neq Z$). The asymptotic expansions of $|+\mathbf{k}\rangle$ and $|-\mathbf{k}\rangle$, respectively, represent a Coulomb-modified plane wave together with an outgoing spherical wave and a Coulomb-modified plane wave plus an incoming spherical wave.[4]

The radial continuum functions can be expressed in a variety of ways, the most common being in terms of the confluent hypergeometric function,

$$R_{l'}(\gamma, kr) = N_{l'}(\gamma)(2kr)^{l'}[(2l'+1)!]^{-1} \times e^{-ikr}F(l'+1+i\gamma\,|2l'+2|\,2ikr), \tag{3a}$$

with

$$N_{l'}(\gamma) = |\Gamma(l'+1+i\gamma)|\, e^{\pi\gamma/2}. \tag{3b}$$

The normalization is here chosen so that $R_{l'}(\gamma, kr)$ asymptotically approaches

$$R_{l'}(\gamma, kr) \underset{kr\to\infty}{\sim} (kr)^{-1} \times \sin(kr + \gamma\ln(2kr) - \pi l'/2 + \eta_{l'}^{\pm}).$$

With this normalization, the radial Coulomb functions go over to spherical Bessel functions as $Z' \to 0$[4] (or as $\gamma = Z'/k \to 0$),

$$\lim_{\gamma\to 0} R_{l'}(\gamma, kr) = j_{l'}(kr).$$

Thus, the partial wave expansion (2) goes over to the partial wave expansion appropriate to a plane wave $\exp(i\mathbf{k}\cdot\mathbf{r})$ as the nuclear charge Z' goes to zero.

For our purpose, it is advantageous to employ the integral representation[4,5]

$$R_{l'}(\gamma, kr) = N_{l'}(\gamma)(-2kr)^{-l'-1}(1/2\pi) \oint e^{-i2krz}G(z)\, dz, \tag{4a}$$

where

$$G(z) = (z+\tfrac{1}{2})^{\alpha^*}(z-\tfrac{1}{2})^{\alpha} \text{ and } \alpha = i\gamma - l' - 1. \tag{4b}$$

The contour is taken counterclockwise around the two branch points of $G(z)$.

3. COULOMB MATRIX ELEMENTS

Employing these rather diverse representations for the discrete and continuum functions, we form the matrix element

$$\langle k, l'|\, r^m\, |n, l\rangle = \int_0^{\infty} R_{l'}(\gamma, kr)r^m R_{n,l}(r)r^2\, dr$$

$$= N_{l'}(\gamma)N_{n,l}(n/2Z)^{m+3}(-K_n)^{-l'-1} \sum_{t=0}^{n-l-1} (-1)^t\, d_t$$

$$\times \left\{\int_0^{\infty} d\rho\, e^{-\rho/2}\rho^{m+1+l-l'+t}\, \frac{1}{2\pi}\oint e^{-iK_n\rho z}G(z)\, dz\right\}, \tag{5}$$

where d_t is given by Eq. (1c) and $K_n = nk/Z$. Generalizing slightly the double integral appearing in (5), we define

$$I(A, B, N) = \left\{\int_0^{\infty} d\rho\, e^{-\rho/A}\rho^N\, \frac{1}{2\pi}\oint e^{-iB\rho z}G(z)\, dz\right\}. \tag{6a}$$

The integral over ρ is obtained immediately, giving

$$I(A, B, N) = \frac{N!}{2\pi}\oint \frac{G(z)}{(1/A + iBz)^{N+1}}\, dz. \tag{6b}$$

The evaluation of this integral is central to this work. Nevertheless, for convenience of presentation, the mathematical details are deferred to the Appendix. There, we find,

[4] W. Gordon, Z. Physik **48**, 180 (1928). For a more recent account, see A. Messiah, *Quantum Mechanics* (John Wiley & Sons, Inc., New York, 1962) Vols. I and II.

[5] A. Sommerfeld and G. Schur, Ann. Physik (5) **4**, 409 (1930).

$$I(A, B, N) = (-1)^{l'}(-B)^{-N-1}(AB/|\beta|)^{N+2l'+2} \times e^{-2\gamma \tan^{-1}(AB/2)} S(N), \quad (7a)$$

where,

$$S(N) = \sum_{s=0}^{N} \binom{N}{s} |\zeta(N-s, s)| \cos \psi(N, s), \quad (7b)$$

$$\zeta(\lambda, \sigma) = \frac{\Gamma(\alpha^* + 1)\Gamma(\alpha + 1)}{\Gamma(\alpha^* + 1 - \lambda)\Gamma(\alpha + 1 - \sigma)}, \quad (7c)$$

$$\psi(N, s) = \arg \zeta(N-s, s) + (s - N/2) \arg(\beta/-\beta^*), \quad (7d)$$

and

$$\beta = AB/2 + i. \quad (7e)$$

The slight generalization incorporated in (6a) facilitates the use of this result for the other discrete–continuum matrix elements involving discrete functions of the form $r^n \exp(-r/A)$, e.g., Slater functions with integral indices. Further, it allows evaluation of matrix elements for more general operators, say $\langle k, l' | r^m \exp(-r/\epsilon) | n, l\rangle$.[6]

Utilizing (7a) in (5) we write the Coulomb matrix element of r^m

$$\langle k, l' | r^m | n, l\rangle = N_{l'}(\gamma) N_{n,l}(n/2Z)^{m+3}(-K_n)^{-l'-1} \times \sum_{t=0}^{n-l-1} (-1)^t d_t I(2, K_n, m+1+l-l'+t), \quad (8)$$

where we have set $A = 2$, $B = K_n$ and $N = m + 1 + l - l' + t$. With the aid of the identity $|\Gamma(i\gamma)| = (\pi/\gamma \sinh \pi\gamma)$ we can reduce (8) to a convenient computational form,

$$\langle k, l' | r^m | n, l\rangle = \eta (K_n)^{l'} e^{-2\gamma \tan^{-1} K_n} \times \sum_{t=0}^{n-l-1} d_t [2/(1 + K_n^2)^{\frac{1}{2}}]^{N+2l'+2} S(N), \quad (9a)$$

where

$$\eta = (-1)^{m+1+l-l'}(n/2Z)^{m+\frac{1}{2}} \times \left[\frac{2\pi\gamma(n+l)!\,(n-l-1)!}{2n(1-e^{-2\pi\gamma})} \prod_{r=1}^{l'} (r^2 + \gamma^2)\right]^{\frac{1}{2}}, \quad (9b)$$

and

$$S(N) = \sum_{s=0}^{N} \binom{N}{s} |\zeta(N-s, s)| \cos \psi(N, s), \quad (9c)$$

$$|\zeta(N-s, s)| = \prod_{p=1}^{N-s} [\gamma^2 + (l' + p)^2] \times \prod_{q=1}^{s} [\gamma^2 + (l' + q)^2], \quad (9d)$$

$$\psi(N, s) = N\pi + \sum_{p=1}^{N-s} \phi(p) - \sum_{q=1}^{s} \phi(q) + (N - 2s)\theta, \quad (9e)$$

with

$$\phi(j) = \tan^{-1}(\gamma/(l' + j)) \quad \text{and} \quad \theta = \tan^{-1} K_n. \quad (9f)$$

Here, as before, $\gamma = Z'/k$, $K_n = nk/Z$, and $N = m+1+l-l'+t$; m is restricted by $m \geq -(1+l-l')$. In (9), we have also reduced several of the defining relations given in Eqs. (7) to a more convenient computational form. The matrix element is in units of $a_0^{m+\frac{1}{2}}$ (i.e., replace $n/2Z$ by $na_0/2Z$).

While our result does not possess the elegance of Gordon's, it has significant advantages for computations (especially with the aid of a digital computer).

4. PLANE-WAVE LIMIT

As previously pointed out, the Coulomb continuum functions approach the spherical Bessel functions as the effective nuclear charge Z' approaches zero. Making use of this, we can derive the radial matrix elements $\langle j_{l'}(kr) | r^m | R_{n,l}(r)\rangle$ directly from the Coulomb matrix elements given in Eqs. (9). Letting $Z'/k = \gamma \to 0$, we readily find

$$\langle j_{l'}(kr) | r^m | R_{n,l}(r)\rangle = \eta_{\gamma\to 0}(K_n)^{l'} \sum_{t=0}^{n-l-1} d_t \times [2/(1 + K_n^2)^{\frac{1}{2}}]^{N+2l'+2} S_{\gamma\to 0}(N), \quad (10a)$$

where

$$\eta_{\gamma\to 0} = (-1)^{m+1+l-l'}(n/2Z)^{m+\frac{1}{2}} \times [(n+l)!\,(n-l-1)!/2n]^{\frac{1}{2}} l'!, \quad (10b)$$

and

$$S_{\gamma\to 0}(N) = \sum_{s=0}^{N} \binom{N}{s} [(l' + N - s)!\,(l' + s)!/(l'!)^2] \times \cos \psi_{\gamma\to 0}(N, s), \quad (10c)$$

$$\psi_{\gamma\to 0}(N, s) = N\pi + (N - 2s) \tan^{-1} K_n. \quad (10d)$$

The plane-wave matrix elements follow from the partial wave expansion (2) taken in the limit as $\gamma \to 0$.

5. NUMERICAL RESULTS

The presentation of numerical results in a sense contradicts the main purpose of this work, namely

[6] Had we formed the hydrogenic matrix element of $r^m \exp(-r/\epsilon)$ in analogy with Eq. (5), then we would have found $A = (\frac{1}{2} + n/2Z\epsilon)^{-1}$.

TABLE I. Discrete–continuum dipolar hydrogenic matrix elements $\langle k, p\ |r|\ 1, s\rangle$ and $\langle k, l \pm 1\ |r|\ 2, l\rangle$ as a function of k (measured in a_0^{-1}).

k	$k, p; 1, s$	$k, p; 2, s$	$k, s; 2, p$	$k, d; 2, p$
0.05	12.08	36.600	10.560	42.053
0.1	8.427	24.735	7.106	28.007
0.2	5.647	14.754	4.177	15.818
0.3	4.229	9.369	2.590	9.276
0.4	3.264	6.002	1.609	5.412
0.5	2.541	3.875	1.001	3.164
0.6	1.981	2.536	0.627	1.874
0.7	1.545	1.689	0.399	1.133
0.8	1.205	1.147	0.258	0.702
0.9	0.941	0.794	0.170	0.445
1.0	0.738	0.561	0.114	0.290
1.2	0.458	0.296	0.055	0.131
1.4	0.291	0.166	0.028	0.65
1.6	0.189	0.099	0.015	0.034
1.8	0.126	0.061	0.0086	0.019
2.0	0.086	0.040	0.0051	0.011
2.5	0.036	0.015	0.0016	0.0035
3.0	0.017	0.0068	0.0006	0.0013

the development of simple computational formulas. We give only selected exemplifying results as a convenient check and slight extension to the results given by Stobbe,[2,7] A FORTRAN computer program was written to evaluate the discrete–continuum matrix elements using Eqs. (9). Some results are given in Tables I and II. Table I (essentially duplicating Stobbe's tables) gives the dipolar radial matrix element $\langle k, l \pm 1|\ r\ |n, l\rangle$ for $n = 1$ and 2 as a function of the propagation vector k (measured in a_0^{-1}). Table II extends the corresponding results for $n = 3$. In both tables, Z and Z' are taken as unity.

TABLE II. Discrete–continuum dipolar hydrogenic matrix elements $\langle k, l \pm 1\ |r|\ 3, l\rangle$ as a function of k.

k	$k, p; 3, s$	$k, s; 3, p$	$k, d; 3, p$	$k, p; 3, d$	$k, f; 3, d$
0.05	70.37	28.34	84.89	12.60	75.08
0.1	45.14	18.00	53.72	7.854	46.03
0.2	22.72	8.733	25.69	3.556	19.71
0.3	11.87	4.330	12.45	1.592	8.279
0.4	6.386	2.193	6.130	0.714	3.510
0.5	3.585	1.154	3.129	0.330	1.552
0.6	2.107	0.636	1.670	0.159	0.726
0.7	1.293	0.366	0.933	0.081	0.359
0.8	0.825	0.219	0.544	0.043	0.187
0.9	0.545	0.136	0.330	0.024	0.102
1.0	0.371	0.087	0.207	0.014	0.058
1.2	0.186	0.039	0.089	0.0051	0.021
1.4	0.101	0.019	0.042	0.0021	0.0088
1.6	0.058	0.010	0.022	0.0010	0.0040
1.8	0.036	0.0056	0.012	0.0005	0.0019
2.0	0.023	0.0033	0.0069	0.00025	0.0010
2.5	0.0086	0.0010	0.0021	0.00006	0.00025
3.0	0.0038	0.0004	0.0008	0.00002	0.00008

[7] Since the explicit numerical results depend on the normalization, we emphasize that we have normalized the continuum functions so that $\langle k', l'|k, l\rangle = k^{-2}(\pi/2)\ \delta(k'-k)\ \delta_{l',l}$ in contrast to Stobbe's normalization of $\langle k',\ l'\ |\ k,\ l\rangle = \delta(k'-k)\ \delta_{l',l}$. To bring our results into accord with Stobbe's quantity $C(k)^2$, multiply the square of our matrix element by $k^2(2/\pi)$. We caution also that Stobbe has recorded the matrix elements as a function of $X(\equiv K_n) = nk/Z'$ rather than the wave number of k.

ACKNOWLEDGMENTS

The author wishes to thank Professor A. Gold for reading the manuscript and several helpful suggestions and J. O'Brien for several helpful discussions and corroborating the results by working through much of the problem.

APPENDIX

We wish to evaluate the integral given in Eqs. (6),

$$I(A, B, N) = \frac{N!}{2\pi}\oint \frac{G(z')}{(1/A + iBz')^{N+1}}\,dz', \qquad \text{(A1)}$$

where $G(z') = (z' + \frac{1}{2})^{\alpha^*}(z' - \frac{1}{2})^{\alpha}$ and $\alpha = i\gamma - l' - 1$.

The contour is taken counterclockwise around the two branch points and the singularity at $z' = -i/AB$.[8] Changing variables to $z' = (z + i)/AB$, we find

$$I(A, B, N) = (1/iB)^{N+1}(AB)^{N+2l'+2}J, \qquad \text{(A2a)}$$

with

$$J = (N!/2\pi)\oint (F(z)/z^{N+1})\,dz, \qquad \text{(A2b)}$$

and

$$F(z) = (z + \beta)^{\alpha^*}(z - \beta^*)^{\alpha}, \quad \beta = AB/2 + i. \qquad \text{(A2c)}$$

Extending the contour [which encloses the branch points and the $(N + 1)$th-order pole at $z = 0$] to infinity, the integration of (A2b) yields simply the residue at the pole $z = 0$,

$$J = \frac{N!}{2\pi}\frac{2\pi i}{N!}F^{(N)}(0) = iF^{(N)}(0), \qquad \text{(A3)}$$

provided $N + 1 > 0$. Since $N = m + 1 + l - l' + t$, where t is a summation index starting at $t = 0$, m is restricted to $m \geq -(1 + l - l')$.

In order to compute the Nth derivative of $F(z)$, define two new functions

$$f(z) = (z + \beta)^{\alpha^*} \quad \text{and} \quad g(z) = (z - \beta^*)^{\alpha},$$

then

$$F(z) = (z + \beta)^{\alpha^*}(z - \beta^*)^{\alpha} = f(z)g(z).$$

[8] H. A. Bethe and E. Salpeter, *Quantum Theory of One and Two Electron Atoms* (Academic Press, Inc., New York, 1957), pp. 21, 304. See also the original work of Sommerfeld and Schur quoted in Ref. 5.

Now by applying Leibnitz' rule for the differentiation of a product (in a symmetrized form), we get

$$F^{(N)}(0) = (d/dz)^N f(z)g(z)\,|_{z=0}$$
$$= \frac{1}{2}\sum_{s=0}^{N}\binom{N}{s}[f^{(N-s)}(0)g^{(s)}(0) + f^{(s)}(0)g^{(N-s)}(0)]. \tag{A4}$$

The symmetrization of Leibnitz' rule is crucial. As we shall see below, $F^{(N)}(0)$ is either real or imaginary depending on N. By using a symmetrized form of Leibnitz' rule we have put $F^{(N)}(0)$ in the form

$$F^{(N)}(0) = \tfrac{1}{2}(F^{(N)}(0) \pm F^{(N)}(0)^*),$$

which yields either a real or an imaginary function explicitly. If Leibnitz' rule is applied in its usual form, it is very difficult to demonstrate this essential character of $F^{(N)}(0)$.

From here, the evaluation of $F^{(N)}(0)$ is straightforward. First we compute the derivatives

$$f^{(\lambda)}(0) = \left(\frac{d}{dz}\right)^{\lambda}(z+\beta)^{\alpha^*}\Big|_{z=0}$$
$$= \frac{\Gamma(\alpha^*+1)}{\Gamma(\alpha^*+1-\lambda)}\beta^{(\alpha^*-\lambda)},$$
$$g^{(\sigma)}(0) = \left(\frac{d}{dz}\right)^{\sigma}(z-\beta^*)^{\alpha}\Big|_{z=0}$$
$$= \frac{\Gamma(\alpha+1)}{\Gamma(\alpha+1-\sigma)}(-\beta^*)^{(\alpha-\sigma)},$$

and then defining $\zeta(\lambda, \sigma)$,

$$\zeta(\lambda,\sigma) = \frac{\Gamma(\alpha^*+1)\Gamma(\alpha+1)}{\Gamma(\alpha^*+1-\lambda)\Gamma(\alpha+1-\sigma)} = \prod_{p=0}^{\lambda-1}(\alpha^*-p)\prod_{q=0}^{\sigma-1}(\alpha-q), \tag{A5}$$

and observing that $\zeta(\lambda, \sigma) = \zeta^*(\sigma, \lambda)$, we get

$$f^{(\lambda)}(0)g^{(\sigma)}(0) + f^{(\sigma)}(0)g^{(\lambda)}(0)$$
$$= \zeta(\lambda,\sigma)\beta^{\alpha^*-\lambda}(-\beta^*)^{\alpha-\sigma} + \zeta^*(\lambda,\sigma)\beta^{\alpha^*-\sigma}(-\beta^*)^{\alpha-\lambda}.$$

After a straightforward reduction, letting $\lambda = N - s$, $\sigma = s$, and $\alpha = i\gamma - l' - 1$, we arrive at

$$f^{(N-s)}(0)g^{(s)}(0) + f^{(s)}(0)g^{(N-s)}(0)$$
$$= \frac{(-\beta^*/\beta)^{i\gamma}}{(-\beta\beta^*)^{N/2+l'+1}}\,|\zeta(N-s,s)|$$
$$\times (\exp\{i[\arg\zeta(N-s,s) + (s-N/2)\arg(\beta/-\beta^*)]\} + \text{complex conjugate})$$
$$= \frac{(-i)^N(-1)^{l'+1}\,e^{-2\gamma\tan^{-1}(AB/2)}}{|\beta|^{2l'+2+N}}\,2\,|\zeta(N-s,s)| \times \cos\psi(N,s), \tag{A6a}$$

where

$$\psi(N,s) = \arg\zeta(N-s,s) + (s-N/2)\arg(\beta/-\beta^*). \tag{A6b}$$

Putting this result into Eq. (A4) completes the computation of the integral J,

$$J = i[(-i)^N(-1)^{l'+1}\,e^{-2\gamma\tan^{-1}(AB/2)}/|\beta|^{2l'+2+N}]S(N), \tag{A7a}$$

where

$$S(N) = \sum_{s=0}^{N}\binom{N}{s}|\zeta(N-s,s)|\cos\psi(N,s). \tag{A7b}$$

Combining (A7a) with (A2a) yields the final result quoted in Eqs. (7) of the text.

Nonlinear Theory of Elastic Directed Surfaces*

H. COHEN AND C. N. DESILVA
University of Minnesota, Minneapolis, Minnesota
(Received 15 November 1965)

The present paper develops a nonlinear theory for the deformation of an elastic directed surface by assuming the existence of a strain energy function and postulating a principle of virtual work which governs its mechanical behavior. The equations of equilibrium and the boundary conditions are shown to involve both the classical stress as well as the double stress. Constitutive equations are derived which give the stress and double stress as functions of a complete set of strain measures which describe the deformation of directed surfaces.

1. INTRODUCTION

EARLY investigations in the theory of elasticity were mainly concerned with establishing special theories for thin bodies. Once, however, the general three-dimensional equations of classical elasticity had been formulated, theories of thin bodies were derived as special cases of this general theory.[1] One attempt to develop a special theory of thin bodies independently of the equations of classical elasticity was carried out by E. and F. Cosserat.[2] The Cosserats following an idea of Duhem[3] introduced the directed or oriented curve and surface in order to construct theories of rods and shells. In their work, they assigned to every point of the surface a set of rigid vectors called directors. Ericksen and Truesdell[4] generalized the concepts of the Cosserats by requiring the set of directors to be deformable. In this reference, Ericksen and Truesdell gave a complete exposition of the kinematics of an oriented surface at each point of which is assigned a triad of deformable directors. They showed that a complete description of the strain of such a surface was specified by a set of deformation measures which they defined. Moreover, by postulating a stress principle, they obtained a set of differential equations of equilibrium for the directed surface; they did not, however, treat the problem of constitutive relations.

The recent work in developing a consistent nonlinear theory of shells[5-9] all have, as their starting point, the equations of classical, three-dimensional elasticity. However, Cohen and DeSilva[10] formulated a nonlinear theory of elastic shells by a direct treatment of the deformation of a restricted class of directed surfaces. In this paper, we obtain a general theory of hyperelasticity governing the nonlinear behavior of a directed surface with a triad of deformable directors.

In Sec. 2, the kinematics of the deformation of a directed surface are reviewed, the basic strain variables of Ericksen and Truesdell[4] are defined and their relation with the two Love–Kirchhoff deformation tensors is established. In Sec. 3, a principle of virtual work is postulated as governing the mechanical behavior of an elastic directed surface. In this principle, we assume the existence of a strain energy function which depends on the classical deformation gradient as well as on the directors and their gradients. This function is required to satisfy the principle of material indifference. The principle of virtual work is applied in Sec. 4 to a directed surface under the action of arbitrary virtual displacements. It is shown that in addition to the classical Cauchy stress tensor, there arises from the interaction of the director forces a general Mindlin double stress.[11] A part of this latter stress tensor gives rise to the conventional Cosserat couple stress, the other part to a momentless double stress. As a

* This research was supported by the National Aeronautics and Space Administration under Research Grant NGR-24-005-059.

[1] A. E. H. Love, *The Mathematical Theory of Elasticity*, Historical Introduction, (Cambridge University Press, Cambridge, England, 1927), pp. 1–31.

[2] E. and F. Cosserat, *Théorie des Corps Déformables*, (Hermann et Fils, Paris, 1909).

[3] P. Duhem, Ann. École Norm, (3) **10**, 187 (1893).

[4] J. L. Ericksen and C. Truesdell, Arch. Rat. Mech. Anal. **1**, 295 (1958).

[5] R. W. Leonard, "Nonlinear First Approximation Thin Shell and Membrane Theory," Ph.D. thesis, Virginia Polytechnic Institute (1961).

[6] J. L. Sanders, Quart. Appl. Math. **21**, 21 (1963).

[7] P. M. Naghdi and R. P. Nordgren, Quart. Appl. Math. **21**, 49 (1963).

[8] W. L. Wainwright, Int. J. Eng. Sci. **1**, 339 (1963).

[9] C. B. Sensenig, IMM-NYU 313, New York University (1963).

[10] H. Cohen and C. N. DeSilva, J. Math. Phys. **7**, 246 (1966).

[11] R. D. Mindlin, Arch. Ratl. Mech. Anal. **16**, 51 (1964).

consequence, for equilibrium, in addition to the equations recorded by Ericksen and Truesdell[4] there are six additional equations involving the 12 components of the momentless double stress. It is also shown that the stress tensor and the double stress tensor depend on the Ericksen–Truesdell strain measures and the general nonlinear constitutive relations are derived.

Finally in Sec. 5 we obtain special cases of the general theory when we restrict ourselves to a homogeneous deformation together with a particular choice of the directors. In particular, the relation of the present theory to that of Ref. 10 is clarified.

2. DEFORMATION OF A DIRECTED SURFACE

A directed surface S_D is defined as a surface S at each point $\mathbf{X}$ of which there is associated a noncoplanar triad field of vectors D_a^K termed the directors. Thus S_D is given by

$$X^K = X^K(U^\Delta); \qquad D_a^K = D_a^K(U^\Delta), \tag{2.1}$$

where the X^K are rectangular Cartesian coordinates and the U^Δ are curvilinear coordinates on S. Here, as in the rest of the paper, Latin and German indices take the values 1, 2, and 3 while Greek indices take the values 1 and 2.

We also define the reciprocal directors D_K^a which satisfy:

$$D_a^K D_M^a = \delta_M^K; \qquad D_a^K D_K^b = \delta_a^b. \tag{2.2}$$

We recall that the surface given by Eq. (2.1) is described by the fundamental surface tensors $A_{\Delta\Sigma}$ and $B_{\Delta\Sigma}$ given by

$$A_{\Delta\Sigma} = X^K_{;\Delta} X^K_{;\Sigma}; \qquad B_{\Delta\Sigma} = N^K X^K_{;\Delta\Sigma}, \tag{2.3}$$

where N^K is the unit normal to S, and where the semicolon denotes the total covariant derivative (Ericksen[12]).

The tensors $\mathbf{A}$ and $\mathbf{B}$ satisfy the usual equations of surface theory associated with the names of Gauss, Weingarten, and Mainardi-Codazzi. However, Ericksen and Tuesdell[4] have shown that, for a differential description of a directed surface S_D, the basic measures are X_D^a, G_{ab}, and $W_{b\Delta}^a$ given by

$$X_\Delta^a = D_K^a X^K_{;\Delta}; \quad G_{ab} = D_a^K D_b^K; \quad W_{b\Delta}^a = D_K^a D^K_{b;\Delta}. \tag{2.4}$$

Specification of the above quantities, subject to certain compatability conditions, are sufficient to determine completely a directed surface S_D to within a rigid-body motion combined with a reflection. In particular, $A_{\Delta\Sigma}$ and $B_{\Delta\Sigma}$ are given in terms of X_Δ^a, D_a^K, G_{ab}, and $W_{b\Delta}^a$ by

$$\begin{aligned} A_{\Delta\Sigma} &= G_{ab} X_\Delta^a X_\Sigma^b, \\ B_{\Delta\Sigma} &= N_K D_a^K (X^a_{\Delta;\Sigma} + W_{b\Sigma}^a X_\Delta^b). \end{aligned} \tag{2.5}$$

We now assume S_D to deform into a directed surface s_d given by

$$\begin{aligned} x^i &= x^i(u^\delta), \\ d_a^k &= d_a^k(u^\delta). \end{aligned} \tag{2.6}$$

Majuscule and miniscule Latin and Greek letters and indices will be associated with S_D, and s_d, respectively. We assume $\mathbf{X}$ in S maps into $\mathbf{x}$ in s, D_a at $\mathbf{X}$ maps into d_a at $\mathbf{x}$. If we specify the mapping

$$u^\delta = u^\delta(U^\Delta), \tag{2.7}$$

then the deformation of S_D into s_d is given by Eqs. (2.1) and (2.6) augmented by Eq. (2.7).

We introduce at each $\mathbf{x}$ of s_d reciprocal directors d_k^a satisfying the dual of Eqs. (2.2) and define, for a differential description of the strain, the quantities

$$x_\Delta^a = d_i^a x^i_{;\Delta}; \tag{2.8a}$$

$$C_{ab} = d_a^i d_b^i; \tag{2.8b}$$

$$F_{b\Delta}^a = d_i^a d^i_{b;\Delta}. \tag{2.8c}$$

When u^δ, x_Δ^a, C_{ab}, and $F_{b\Delta}^a$ are given as functions of U^Δ on S_D, subject to a prescribed set of compatability conditions, they determine s_d to within a rigid-body motion combined with a reflection. In particular, the Love–Kirchhoff strain tensors $L_{\Delta\Sigma}$ and $K_{\Delta\Sigma}$ defined by Cohen and DeSilva[10] are given in terms of the directors d_a^k and the Ericksen–Truesdell fundamental forms, Eq. (2.8), by

$$\begin{aligned} L_{\Delta\Sigma} &= x^i_{;\Delta} x^i_{;\Sigma} = C_{ab} x_\Delta^a x_\Sigma^b, \\ K_{\Delta\Sigma} &= n^i x^i_{;\Delta\Sigma} = n^i d_a^i [F_{b\Sigma}^a x_\Delta^b + x^a_{\Delta;\Sigma}]. \end{aligned} \tag{2.9}$$

3. A PRINCIPLE OF VIRTUAL WORK FOR ELASTIC DIRECTED SURFACES

We shall postulate a principle of virtual work which is assumed to govern the mechanical behavior of a directed surface. This principle assumes the existence of a strain energy function thus endowing the surface with hyperelastic material properties. The concept of a "variation" as applied to the deformed state is exactly that defined by Truesdell and Toupin.[13]

[12] J. L. Ericksen, "Tensor Fields" in *Handbuch der Physik*, edited by S. Flügge (Springer-Verlag, Berlin, 1960) Vol. III/1.

[13] C. Truesdell and R. Toupin, "Classical Field Theories," *Handbuch der Physik*, edited by S. Flügge (Springer-Verlag, Berlin, 1960), Vol. III/1.

The virtual work $\mathcal{C}$ associated with an arbitrary virtual displacement δx^i, δd_a^i of s_d is defined by:

$$\mathcal{C} = \oint_c [s_k\, \delta x^k + p_k^a\, \delta d_a^k]\, dc + \int_\sigma \gamma[f_k\, \delta x^k + q_k^a\, \delta d_a^k]\, d\sigma, \qquad (3.1)$$

where c is a circuit enclosing a region σ in s, and where γ is the surface mass density of s_d. The quantities s_k and f_k correspond to the usual stress and body-force vectors respectively, while p_k^a and q_k^a correspond to generalized stress and body-force vectors associated with the directors d_a^k.

If we define the variation δd_a^k in terms of a set of infinitesimal quantities $\delta\lambda_i^k$ according to

$$\delta d_a^k = \delta\lambda_i^k\, d_a^i, \qquad (3.2)$$

then the variation $\delta\lambda_i^k$ has an obvious physical interpretation. Introducing Eq. (3.2) into Eq. (3.1) results in

$$\mathcal{C} = \oint_c [s_k\, \delta x^k + m_{ik}\, \delta\lambda^{ik}]\, dc + \int_\sigma \gamma[f_k\, \delta x^k + l_{ik}\, \delta\lambda^{ik}]\, d\sigma, \qquad (3.3)$$

where

$$m_k^i = d_a^i p_k^a; \qquad l_k^i = d_a^i q_k^a. \qquad (3.4)$$

We thus see that the system of forces p_k^a, q_k^a associated with the directors may be replaced by an energetically equivalent system of couples $m^{[ik]}$, $l^{[ik]}$ and double stresses without moment $m^{(ik)}$, $l^{(ik)}$ acting on s. Here we employ the convention of Ericksen[12] regarding the meaning of parentheses and brackets about two indices.

We assume the existence of a strain energy function such that

$$\epsilon = \epsilon(x^i_{;\Delta};\, d_a^i;\, d^i_{a;\Delta}). \qquad (3.5)$$

The energy W stored during deformation is then given by

$$W = \int_\sigma \gamma\epsilon\, d\sigma. \qquad (3.6)$$

We now postulate a principle of virtual work which requires that

$$\mathcal{C} = \delta W \qquad (3.7)$$

be a necessary condition for equilibrium of a directed surface for arbitrary virtual displacements δx^k, δd_a^k. The variation δW is subject to the requirement imposed by the principle of mass conservation expressed by

$$\delta(\gamma\, d\sigma) = 0. \qquad (3.8)$$

The strain energy function ϵ is required to satisfy the condition of material indifference, i.e., ϵ is invariant under rigid motions of s_d. Thus we must have

$$\delta\epsilon = 0 \quad \text{for} \quad \begin{cases} \delta x^k = c^k, \\ \delta d_a^k = 0, \end{cases} \qquad (3.9)$$

and

$$\delta\epsilon = 0 \quad \text{for} \quad \begin{cases} \delta x^k = c^{ik}\, x^i, \\ \delta d_a^k = c^{ik}\, d_a^i, \end{cases} \qquad (3.10)$$

where c^k and c^{ik} are an arbitrary constant vector and an infinitesimal skew-symmetric tensor respectively. From the form of Eq. (3.5) it follows that both conditions (3.9) and (3.10) are satisfied if we require ϵ to be a solution of the system of partial differential equations:

$$\frac{\partial\epsilon}{\partial x^{[i}_{;\Delta}}\, x^{k]}_{;\Delta} + \frac{\partial\epsilon}{\partial d_a^{[i}}\, d_a^{k]} + \frac{\partial\epsilon}{\partial d^{[i}_{a;\Delta}}\, d^{k]}_{a;\Delta} = 0. \qquad (3.11)$$

On applying the principle of virtual work subject to the variations (3.9) and (3.10) we obtain the following compatability equations with respect to the applied loading:

$$\oint_c s_k\, dc + \int_\sigma \gamma f_k\, d\sigma = 0 \qquad (3.12)$$

$$\oint_c (x^{[i}s^{k]} + m^{[ik]})\, dc + \int_\sigma \gamma(x^{[i}f^{k]} + l^{[ik]})\, d\sigma = 0. \qquad (3.13)$$

Equations (3.12) and (3.13) are the equations of applied force and applied moment equilibrium, respectively.

4. BASIC EQUATIONS OF AN ELASTIC DIRECTED SURFACE

In this section we shall obtain, by application of the principle of virtual work, the basic equilibrium equations, boundary conditions and constitutive equations for a hyperelastic directed surface. In order to satisfy the principle of material indifference, the assumed form of the strain energy function is shown to depend only on the Ericksen–Truesdell fundamental quantities, defined in Sec. 2, which serve as measures of strain for a directed surface.

Application of the principle of virtual work, expressed by Eq. (3.7), for arbitrary virtual displacements δx^k, δd_a^k requires, when Eqs. (3.1), (3.5), (3.6), and (3.8) are used, that

$$t^{k\delta}_{;\delta} + \gamma f^k = 0 \qquad \text{in} \quad \sigma, \qquad (4.1)$$

$$s^k = t^{k\delta}\nu_\delta \quad \text{on} \quad c, \tag{4.2}$$

$$\mu^{\alpha k\delta}_{;\delta} - \phi^{\alpha k} + \gamma q^{\alpha k} = 0 \quad \text{in} \quad \sigma, \tag{4.3}$$

$$p^{\alpha k} = \mu^{\alpha k\delta}\nu_\delta \quad \text{on} \quad c, \tag{4.4}$$

where ν_α are the components of the unit outward normal vector to c which is tangential to s. In order to obtain Eqs. (4.1) and (4.3), we have defined:

$$t^\delta_k = \gamma(\partial\epsilon/\partial x^k_{;\Delta})u^\delta_{;\Delta}, \tag{4.5a}$$

$$\phi^\alpha_k = \gamma\,\partial\epsilon/\partial d^k_\alpha, \tag{4.5b}$$

$$\mu^{\alpha\delta}_k = \gamma(\partial\epsilon/\partial d^k_{\alpha;\Delta})u^\delta_{;\Delta}, \tag{4.5c}$$

and have made use of Green's theorem in the form

$$\int_\sigma c^{k\delta}_{;\delta}\,d\sigma = \oint_c c^{k\delta}\nu_\delta\,dc, \tag{4.6}$$

where $c^{k\delta}$ is an arbitrary double tensor field defined in s. Equations (4.1) are the equations of force equilibrium and equation (4.2) represents the corresponding boundary conditions. The quantities $t^{k\delta}$ are the components of the Cauchy stress tensor and are given in terms of the strain energy function by the constitutive relations (4.5a). Equations (4.3) are the equations of equilibrium for the director forces; Eqs. (4.4) are the boundary conditions for the director forces. The quantity $\boldsymbol{\mu}$ with components $\mu^{\alpha k\delta}$ given by the constitutive relations (4.5c) is the stress tensor corresponding to the director forces and may be regarded as expressing the effect of the interaction between the directors. The quantity $\boldsymbol{\phi}$ with components $\phi^{\alpha k}$ given by the constitutive relations (4.5b) is not transmitted across a curve c in s and thus has the character of a body force which may be thought of as arising from internal resistance of the directors to deformation.

If we make use of Eq. (3.2) in defining an arbitrary virtual displacement, the equilibrium equations (4.3) and the boundary conditions (4.4) may be cast into the equivalent form:

$$\mu^{ik\delta}_{;\delta} - \mu^{\alpha k\delta}\,d^i_{\alpha;\delta} - \phi^{ik} + \gamma l^{ik} = 0, \tag{4.7}$$

$$m^{ik} = \mu^{ik\delta}\nu_\delta, \tag{4.8}$$

where

$$\phi^{ik} = d^i_\alpha\phi^{\alpha k}; \qquad \mu^{ik\delta} = d^i_\alpha\mu^{\alpha k\delta}. \tag{4.9}$$

From Eqs. (3.3), (3.4), and (4.8), we see that $\mu^{ik\delta}$ correspond to the components of a double stress distribution arising from the director forces p^α_k. We shall refer to this general double stress distribution as the Mindlin stress.[11] This stress may be decomposed into a double stress with moment—which we will call the couple stress—and into a double stress without moment, which we will call simply the double stress. Similarly, from Eq. (4.7) we see that ϕ^{ik} may be regarded as a body double force. The parts of Eqs. (4.7) and (4.8) antisymmetric in the free indices are the equations of moment equilibrium and the boundary conditions, respectively. The part of these equations symmetric in the free indices will give the equations for the distribution of the double stress.

The principle of material indifference, Eq. (3.11), together with the use of Eqs. (4.5) and (4.9), yields

$$t^{[i\delta}x^{k]}_{;\delta} - \phi^{[ik]} - \mu^{\alpha[k\delta}d^{i]}_{\alpha;\delta} = 0. \tag{4.10}$$

When we apply Eq. (4.10) to the antisymmetric part of Eq. (4.7) we obtain:

$$\mu^{[ik]\delta}_{;\delta} - t^{[i\delta}x^{k]}_{;\delta} + \gamma l^{[ik]} = 0. \tag{4.11}$$

If we set

$$F^b_{a\delta} = F^b_{a\Delta}U^\Delta_{;\delta}, \tag{4.12}$$

then it follows from Eq. (2.8c) that

$$d^i_{a;\delta} = d^i_b F^b_{a\delta}. \tag{4.13}$$

The symmetric part of Eq. (4.7) then becomes, with the use of Eq. (4.13),

$$\mu^{(ik)\delta}_{;\delta} - d^{(i}_b F^b_{a\delta}\mu^{ak)\delta} - \phi^{(ik)} + \gamma l^{(ik)} = 0. \tag{4.14}$$

Equations (4.1) and (4.11) are well known and have been recorded by Ericksen and Truesdell.[4] It is believed that for the case of a directed surface Eq. (4.14) is new.

The condition of material indifference provides a set of three linear homogeneous partial differential Eqs. (3.11) in the 33 independent variables $x^k_{;\Delta}$, d^k_a, and $d^k_{a;\Delta}$. From the theory of such equations it follows that there are 30 independent solutions in terms of which ϵ may be expressed. These solutions of Eq. (3.11) are

$$x^a_\Delta = d^a_k x^k_{;\Delta}; \quad C_{ab} = d^k_a d^k_b; \quad F^a_{b\Delta} = d^a_k d^k_{b;\Delta}, \tag{4.15}$$

which are the Ericksen–Truesdell basic strain measures for a directed surface as recorded in Sec. 2.

In terms of these variables, the constitutive relations (4.5) take the form

$$t^\delta_k = \gamma\,d^a_k(\partial\epsilon/\partial x^a_\Delta)u^\delta_{;\Delta}, \tag{4.16a}$$

$$\phi^i_k = -\gamma\left[\frac{\partial\epsilon}{\partial x^a_\Delta}x^i_{;\Delta}d^a_k - 2\,\frac{\partial\epsilon}{\partial C_{ab}}d^i_a d_{kb} + \frac{\partial\epsilon}{\partial F^b_{c\Delta}}F^a_{c\Delta}d^i_a d^b_k\right], \tag{4.16b}$$

$$\mu^{i\delta}_k = \gamma\,d^i_a d^c_k(\partial\epsilon/\partial F^c_{a\Delta})\mu^\delta_{;\Delta}, \tag{4.16c}$$

where we have used the condition that $\partial\epsilon/\partial C_{ab} =$

$\partial\epsilon/\partial C_{ba}$. We point out that these relations must satisfy specified conditions in the undeformed or reference state.

If r^i, r^{ik}, and $r^{ik\delta}$ are the components of arbitrary mixed tensors, then they may be represented in terms of components tangent and normal to the surface by

$$\begin{aligned} r^i &= r^\alpha x^i_{;\alpha} + rn^i, \\ r^{ik} &= r^{\alpha\beta}x^i_{;\alpha}x^k_{;\beta} + r^\alpha n^{(i}x^{k)}_{;\alpha} + \bar{r}^\alpha n^{[i}x^{k]}_{;\alpha} + rn^in^k, \\ r^{ik\delta} &= r^{\alpha\beta\delta}x^i_{;\alpha}x^k_{;\beta} + r^{\alpha\delta}n^{(i}x^{k)}_{;\alpha} + \bar{r}^{\alpha\delta}n^{[i}x^{k]}_{;\alpha} + r^\delta n^in^k. \end{aligned} \tag{4.17}$$

By making use of equations of the type (4.17), the basic equations may be given in terms of their surface decomposition. The equations of force equilibrium (4.1) and the boundary conditions (4.2) may consequently be decomposed into forms parallel and normal to s given by

$$\begin{aligned} t^{\beta\alpha}_{;\alpha} - b^\beta_\alpha t^\alpha + \gamma f^\beta &= 0, \\ t^\alpha_{;\alpha} + b_{\beta\alpha}t^{\beta\alpha} + \gamma f &= 0, \end{aligned} \tag{4.18}$$

and

$$s^\alpha = t^{\alpha\beta}\nu_\beta; \qquad s = t^\alpha\nu_\alpha. \tag{4.19}$$

The constitutive relations (4.16) for t^δ_k may be brought into the equivalent form:

$$\begin{aligned} t^{\alpha\delta} &= \gamma\, d^{\alpha\mathfrak{a}}(\partial\epsilon/\partial x^{\mathfrak{a}}_\Delta)u^\delta_{;\Delta}, \\ t^\delta &= \gamma\, d^{\mathfrak{a}}(\partial\epsilon/\partial x^{\mathfrak{a}}_\Delta)u^\delta_{;\Delta}, \end{aligned} \tag{4.20}$$

where, following Eq. (4.17), we have set

$$\begin{aligned} t^{k\delta} &= t^{\alpha\delta}x^k_{;\delta} + t^\delta n^k, \\ d^k_{\mathfrak{a}} &= d^\alpha_{\mathfrak{a}}x^k_{;\alpha} + d_{\mathfrak{a}}n^k. \end{aligned}$$

The conditions for moment equilibrium of Eq. (4.11) may similarly, by using Eq. (4.17), be reduced to equations parallel and normal to s given respectively by:

$$\begin{aligned} \bar{\mu}^{\beta\delta}_{;\delta} + 2b_{\alpha\delta}\mu^{[\alpha\beta]\delta} - t^\beta + \gamma\bar{l}^\beta &= 0, \\ \mu^{[\alpha\beta]\delta}_{;\delta} - b^{[\alpha}_\delta\bar{\mu}^{\beta]\delta} - t^{[\alpha\beta]} + \gamma l^{[\alpha\beta]} &= 0. \end{aligned} \tag{4.21}$$

The skew-symmetric part of Eq. (4.8) yields the boundary conditions for couple stress parallel to s

$$\bar{m}^\alpha = \bar{\mu}^{\alpha\delta}\nu_\delta \tag{4.22a}$$

and for couple stress normal to s

$$m^{[\alpha\beta]} = \mu^{[\alpha\beta]\delta}\nu_\delta. \tag{4.22b}$$

We note that the $\bar{\mu}^{\alpha\beta}$ are the four "in-surface" components of the couple stress tensor while $\mu^{[\alpha\beta]\delta}$ are the two normal components of the couple stress tensor. These components give rise on the boundary c to a couple vector with two tangential components $\bar{m}^\alpha$ and a normal component $m^{[\alpha\beta]}$. The constitutive equations for these components of the couple stress tensor may, from Eq. (4.16c), be shown to be given by:

$$\begin{aligned} \bar{\mu}^{\alpha\delta} &= 2\gamma(\partial\epsilon/\partial F^c_{\mathfrak{a}\Delta})d^\alpha_{[\mathfrak{a}}d^{c]}u^\delta_{;\Delta}, \\ \mu^{[\alpha\beta]\delta} &= \gamma(\partial\epsilon/\partial F^c_{\mathfrak{a}\Delta})d^{[\alpha}_{\mathfrak{a}}d^{\beta]c}u^\delta_{;\Delta}. \end{aligned} \tag{4.23}$$

The same procedure is next applied to Eq. (4.14), again using the surface decomposition forms of Eq. (4.17). We simply record the equations of the double force distribution in their final form:

$$\mu^{(\alpha\beta)\delta}_{;\delta} - b^{(\alpha}_\delta\mu^{\beta)\delta} - d^{(\alpha}_b F^b_{\mathfrak{a}\delta}\mu^{\mathfrak{a}\beta)\delta} - \phi^{(\alpha\beta)} + \gamma l^{(\alpha\beta)} = 0, \tag{4.24}$$

$$\mu^{\beta\delta}_{;\delta} + 2b_{\alpha\delta}(\mu^{(\alpha\beta)\delta} - a^{\alpha\beta}\mu^\delta) - F^b_{\mathfrak{a}\delta}(d^\beta_b\mu^{\mathfrak{a}\delta} + d_b\mu^{\mathfrak{a}\beta\delta}) - \phi^\beta + \gamma l^\beta = 0, \tag{4.25}$$

$$\mu^\delta_{;\delta} + b_{\alpha\delta}\mu^{\alpha\delta} - d_bF^b_{\mathfrak{a}\delta}\mu^{\mathfrak{a}\delta} - \phi + \gamma l = 0. \tag{4.26}$$

The boundary conditions for the double stress, by considering the symmetric part of Eq. (4.8), take the form

$$m^{(\alpha\beta)} = \mu^{(\alpha\beta)\delta}\nu_\delta; \quad m^\alpha = \mu^{\alpha\delta}\nu_\delta; \quad m = \mu^\delta\nu_\delta. \tag{4.27}$$

By considering these equations and the procedure of their derivation, we interpret $\mu^{(\alpha\beta)\delta}$ as the six components of a membrane double stress, $\mu^{\beta\delta}$ as the four components of a double transverse shear stress, and μ^δ as the two components of a double transverse pressure stress. These components give rise on the boundary c to a membrane double force $m^{(\alpha\beta)}$, to a double transverse shear force m^α, and to a double normal pressure m, respectively. In an analogous manner, we may interpret $\phi^{(\alpha\beta)}$ as the three components of a body "in-surface" or membrane double force, ϕ^α as the two components of a body double shear force transverse to the surface, and ϕ as a body double normal pressure or "squeeze" force. Similarly, $\phi^{[\alpha\beta]}$ is the component of a body couple vector normal to s, and $\bar{\phi}^\alpha$ are the surface or membrane components of this body couple vector.

Equation (4.24) represents therefore a set of three equations for the equilibrium of membrane double stress parallel to s_d, Eq. (4.25) is a set of two equations for the distribution of the double shear stress, while Eq. (4.26) is the equation for double pressure stress normal to s_d.

Finally, we record the constitutive equations for the relevant components of $\mathbf{\mu}$ and ϕ when we suitably modify Eqs. (4.16):

$$\mu^{(\alpha\beta)\delta} = \gamma(\partial\epsilon/\partial F^c_{a\Delta})d^{(\alpha}_a d^{\beta)}{}_c u^\delta_{;\Delta},$$
$$\mu^{\alpha\delta} = 2\gamma(\partial\epsilon/\partial F^c_{a\Delta})d^\alpha_{(a} d^{c)} u^\delta_{;\Delta}, \tag{4.28}$$
$$\mu^\delta = \gamma(\partial\epsilon/\partial F^c_{a\Delta})d_a d^c u^\delta_{;\Delta},$$

and

$$\phi^{\alpha\beta} = -\gamma\Big[\frac{\partial\epsilon}{\partial x^b_\Delta}u^\alpha_{;\Delta}d^{\beta b} - 2\frac{\partial\epsilon}{\partial C_{ab}}d^\alpha_a d^\beta_b + \frac{\partial\epsilon}{\partial F^b_{c\Delta}}F^a_{c\Delta}d^\alpha_a d^{\beta b}\Big],$$
$$\phi^\alpha = -\gamma\Big[\frac{\partial\epsilon}{\partial x^b_\Delta}u^\alpha_{;\Delta}d^b - 4\frac{\partial\epsilon}{\partial C_{ab}}d^\alpha_{(a}d_{b)} + 2\frac{\partial\epsilon}{\partial F^b_{c\Delta}}F^a_{c\Delta}d^\alpha_{(a}d^{b)}\Big], \tag{4.29}$$
$$\bar\phi^\alpha = -\gamma\Big[\frac{\partial\epsilon}{\partial x^b_\Delta}u^\alpha_{;\Delta}d^b + 2\frac{\partial\epsilon}{\partial F^b_{c\Delta}}F^a_{c\Delta}d^\alpha_{[a}d^{b]}\Big],$$
$$\phi = \gamma\Big[2\frac{\partial\epsilon}{\partial C_{ab}}d_a d_b - \frac{\partial\epsilon}{\partial F^b_{c\Delta}}F^a_{c\Delta}d_a d^b\Big].$$

In summary, the equations of equilibrium are given by Eqs. (4.18), (4.21), (4.24), (4.25), and (4.26). The 33 constitutive relations are given by Eqs. (4.20), (4.23), (4.28), and (4.29). Substitution of the latter set of equations into the equilibrium equations will yield a set of 12 equations involving the 12 unknown functions x^k and d^k_a.

5. SOME SPECIAL CASES OF THE THEORY

We shall consider in this section certain special cases which depend not only on a particular choice of directors, but also on a restriction in the class of permissible deformations. For our purpose it is convenient to subdivide the director triad $\mathbf{d}_a$ into a set of two directors $\mathbf{d}_{\hat a}$ ($\hat a = 1, 2$) and a single director $\mathbf{d}_3$; the reciprocal set is then denoted by $\mathbf{d}^{\hat a}$, $\mathbf{d}^3$. It is, moreover, also convenient to replace the Ericksen–Truesdell measures given in Eq. (4.15) by an equivalent set

$$L_{\Delta\Sigma} = x^k_{;\Delta}x^k_{;\Sigma}, \qquad x^a_\Delta = x^k_{;\Delta}d^a_k,$$
$$C_{a3} = d^k_a d^k_3, \qquad F^a_{b\Delta} = d^k_a d^k_{b;\Delta}. \tag{5.1}$$

As a consequence, the first two of the constitutive equations (4.16) take the form:

$$t^\delta_k = \gamma\Big[\frac{\partial\epsilon}{\partial L_{\Delta\Sigma}}x^k_{;\sigma}u^\sigma_{;\Sigma} + \frac{\partial\epsilon}{\partial x^a_\Delta}d^a_k\Big]u^\delta_{;\Delta},$$
$$\phi^i_k = -\gamma\Big[\frac{\partial\epsilon}{\partial x^b_\Delta}x^i_{;\Delta}d^b_k - 2\frac{\partial\epsilon}{\partial C_{b3}}d^{(i}_b d_{k)3} + \frac{\partial\epsilon}{\partial F^b_{c\Delta}}F^a_{c\Delta}d^i_a d^b_k\Big]. \tag{5.2}$$

The third equation (4.16c) remains unchanged.

Case 1

We first consider the case in which the strain energy function depends only on two directors and their gradients (in addition to $x^i_{;\Delta}$), i.e.,

$$\epsilon = \epsilon(x^i_{;\Delta}; d^i_{\hat a}; d^i_{\hat a;\Delta}). \tag{5.3}$$

Since there are 24 independent arguments appearing in Eq. (5.3), there will be 21 independent solutions satisfying the condition of material indifference, Eq. (3.11). This set of solutions is a subset of Eq. (5.1) and is given by:

$$L_{\Delta\Sigma} = x^i_{;\Delta}x^i_{;\Sigma}; \quad x^a_\Delta = d^a_k x^k_{;\Delta}; \quad F^a_\Delta = d^a_k d^k_{b;\Delta}. \tag{5.4}$$

If the two directors $d^k_{\hat a}$ are chosen as the surface vectors $x^k_{;\Delta}$, then

$$d^k_{\hat a} = \delta^\Delta_{\hat a}x^k_{;\Delta} = \delta^\Delta_{\hat a}x^k_{;\delta}u^\delta_{;\Delta}, \tag{5.5}$$

with the reciprocal set given by

$$d^a_k = a^{\sigma\delta}U^\Delta_{;\sigma}\delta^a_\Delta x_{k;\delta} + \delta^a_3 n_k. \tag{5.6}$$

On substituting Eqs. (5.5) and (5.6) into Eq. (5.4) we find that the basic strain variables are:

$$L_{\Delta\Sigma} = x^i_{;\Delta}x^i_{;\Sigma},$$
$$F^3_{b\Delta} = F^3_{\Sigma\Delta} = K_{\Sigma\Delta} = n_i x^i_{\Delta\Sigma}, \tag{5.7}$$
$$F^{\hat a}_{b\Delta} = F^\Sigma_{\Gamma\Delta} = U^\Sigma_{;\sigma}u^\sigma_{;\Gamma\Delta}.$$

From Eqs. (5.5) and (5.6) the directors and their reciprocal set have components tangent and normal to the surface, as defined in the previous section, given by

$$d^\alpha_a = \delta^\Delta_a u^\alpha_{;\Delta}; \qquad d_a = 0,$$
$$d^{\alpha a} = a^{\sigma\alpha}U^\Delta_{;\sigma}\delta^a_\Delta; \qquad d^a = \delta^a_3. \tag{5.8}$$

If we now restrict the deformation to be homogeneous, i.e., if we set

$$u^\sigma_{;\Delta\Sigma} = 0, \tag{5.9}$$

then the nonzero components of the constitutive Eqs. (5.2) and (4.16c) are

$$t^{\alpha\beta} = 2\gamma(\partial\epsilon/\partial L_{\Delta\Sigma})u^\alpha_{;\Delta}u^\beta_{;\Sigma}, \tag{5.10}$$
$$\mu^{\alpha\beta} = \bar\mu^{\alpha\beta} = \gamma(\partial\epsilon/\partial K_{\Delta\Sigma})u^\alpha_{;\Delta}u^\beta_{;\Sigma}, \tag{5.11}$$

where ϵ is a function of $L_{\Delta\Sigma}$ and $K_{\Delta\Sigma}$ regarded as eight independent quantities subject to the restriction

$$\epsilon_{\Delta\Sigma}(\partial\epsilon/\partial L_{\Delta\Sigma}) = \epsilon_{\Delta\Sigma}(\partial\epsilon/\partial K_{\Delta\Sigma}) = 0.$$

We note that $t^{[\alpha\beta]} = \bar\mu^{[\alpha\beta]} = 0$. It is straightforward to show that the equations of equilibrium of the previous section now reduce to

$$\left.\begin{aligned} t^{\beta\alpha}_{;\alpha} + \gamma f^{\beta} &= 0, \\ b_{\beta\alpha} t^{\beta\alpha} + f &= 0, \end{aligned}\right\} \tag{5.12}$$

$$\left.\begin{aligned} \bar{\mu}^{\beta\delta}_{;\delta} + \gamma \bar{l}^{\beta} &= 0, \\ b^{\alpha}_{\delta} \bar{\mu}^{\beta\delta} - \gamma l^{\alpha\beta} &= 0, \end{aligned}\right\} \tag{5.13}$$

$$b_{\alpha\beta} \bar{\mu}^{\alpha\beta} + \gamma l = 0, \tag{5.14}$$

and we must also have $\bar{l}^{\alpha} = l^{\alpha}$.

The stress boundary conditions become

$$\begin{aligned} &s^{\alpha} = t^{\alpha\beta}\nu_{\beta}; \quad \bar{m}^{\alpha} = \bar{\mu}^{\alpha\delta}\nu_{\delta}; \quad m^{\alpha} = \mu^{\alpha\delta}\nu_{\delta}, \\ &s = m^{\alpha\beta} = m = 0. \end{aligned} \tag{5.15}$$

Equations (5.10), (5.11), (5.12), (5.13), and (5.15) are identical with Eqs. (5.25), (5.26), (5.19), and (5.12) of Cohen and DeSilva[10] provided we put the undetermined constants appearing in the above equations of this reference equal to zero. We point out that Eq. (5.14) is an additional equation pertaining to the double pressure normal to the surface. It arises here because of the more general treatment used in the present paper.

Case 2

We now choose the directors $d^k_{\bar{a}}$ and d^k_3 to be the surface vectors $x^k_{;\Delta}$ and the normal vector n^k, respectively. The strain energy function is then given by

$$\epsilon = \epsilon(x^i_{;\Delta}; n^i; x^i_{;\Delta\Sigma}; n^i_{;\Delta}).$$

We may now write

$$\begin{aligned} d^k_{\bar{a}} &= \delta^{\Delta}_{\bar{a}} u^{\delta}_{;\Delta} x^k_{;\delta} + \delta^3_{\bar{a}} n^k, \\ d^{\bar{a}}_k &= a^{\sigma\delta} U^{\Delta}_{;\sigma} \delta^{\bar{a}}_{\Delta} x_{k;\delta} + \delta^{\bar{a}}_3 n_k. \end{aligned} \tag{5.16}$$

The strain variables given by Eq. (5.1) reduce to the set of Eq. (5.7) plus the additional quantity

$$F^{\bar{a}}_{3\Delta} = F^{\Sigma}_{3\Delta} = -b^{\beta}_{\alpha} u^{\alpha}_{;\Delta} U^{\Sigma}_{;\beta} = -K^{\Sigma}_{\Delta}, \tag{5.17}$$

where

$$K_{\Delta\Sigma} = L_{\Delta\Gamma} K^{\Gamma}_{\Sigma}. \tag{5.18}$$

If we now impose the condition of a homogeneous deformation expressed by Eq. (5.9), the constitutive equations for the nonzero components of the stress and body force are given by

$$\begin{aligned} t^{\alpha\beta} &= 2\gamma(\partial\epsilon/\partial L_{\Delta\Sigma}) u^{\alpha}_{;\Delta} u^{\beta}_{;\Sigma}, \\ \bar{\mu}^{\alpha\beta} &= 2\gamma(\partial\epsilon/\partial K_{\Delta\Sigma}) u^{\alpha}_{;\Delta} u^{\beta}_{;\Sigma}, \\ \phi^{\alpha\beta} &= \tfrac{1}{2} b^{\alpha}_{\sigma} \bar{\mu}^{\beta\sigma}, \qquad \phi = \tfrac{1}{2} b_{\sigma\delta} \bar{\mu}^{\sigma\delta}. \end{aligned} \tag{5.19}$$

The equilibrium equations are exactly those given by Eqs. (5.12), (5.13), and (5.14), with the proviso that $l^{\alpha} = 0$. The boundary conditions are identical with Eq. (5.15) with $m^{\alpha} = 0$. We note that $\mu^{\alpha\beta} = 0$ and that $\bar{\mu}^{\alpha\beta}$ differs by a factor of 2 from the form given in Eq. (5.11). Note also that the quantity ϕ does not vanish and will give rise to a body-couple and a body-double force.

q-Equivalent Particle Hamiltonians. I. The Classical One-Dimensional Case*

DOUGLAS G. CURRIE AND EUGENE J. SALETAN
Physics Department, Northeastern University, Boston, Massachusetts
(Received 28 September 1965)

The classes of equivalent Lagrangians in one-dimensional particle dynamics are found. These classes contain not only Lagrangians yielding the same equations of motion (Lagrangians differing by a total time derivative), but also those implying each other's equations of motion. The corresponding classes of Hamiltonians, all of which give the same orbits in configuration space, but in general different orbits in phase space, are also found. Some specific examples are presented.

I. INTRODUCTION

THERE has, of late, been some interest shown in establishing a formulation of dynamics in which the momentum variable plays only a secondary role, the position variable being distinguished as of particular importance.[1-6] For example, it is shown by the no-interaction theorem that, for a canonical representation of the Lorentz group in which the position transforms properly, the only possible Hamiltonians are those which give straight world lines.[1] Further, every such Hamiltonian is related by a rather trivial canonical transformation (one which transforms the momentum, but leaves the physical position unaltered) to another in which the momentum is a constant. Thus in this theorem the momentum is a sort of dummy variable. If one takes the point of view that the momentum may be treated as a dummy, so that in the Hamiltonian formalism, as in the Lagrangian, it is q and $\dot{q}$ rather than q and p which are of physical significance, then several new possibilities open up. In particular it becomes interesting to treat two Hamiltonians as equivalent if only the equations for the position variables are the same. In this paper we consider the question of finding all Hamiltonians equivalent in this sense.

Let us state the problem more carefully. Let $H(q, p, t)$ be a Hamiltonian for a dynamical system with one degree of freedom. (We treat only the case of one degree of freedom. The problem for more than one becomes significantly more complicated.) We write down Hamilton's canonical equations $\dot{q} = \partial H/\partial p$ and $-\dot{p} = \partial H/\partial q$, and by differentiating the first with respect to the time and then eliminating p and $\dot{p}$ by use of both equations, we arrive finally at an equation of the form

$$\ddot{q} = \rho(q, \dot{q}, t). \tag{1}$$

The problem we set is to find all Hamiltonians which, when subjected to this procedure, yield precisely the same Eq. (1).[7] We call these the set of all *q-equivalent* Hamiltonians. (They are not, of course, entirely equivalent in the usual sense, since they may very well give different second order equations for p and thus different orbits in phase space.)

Some cautionary remarks: Note that we are leaving aside the somewhat wider problem of finding a Hamiltonian when only the q equation of motion is known.[8] We assume always that at least one Hamiltonian is given. Note also that there is no question here of transformations, point, contact, or otherwise. We are looking for the set of all functions $H(q, p, t)$ which will give the same Eq. (1) when the dummy variable p is handled as in the procedure outlined above. If, perchance, there exists some transformation that will carry us from one such Hamiltonian to another, it is for the present entirely

* This work was supported in part by a grant from the Air Force Office of Scientific Research.

[1] D. G. Currie, J. Math. Phys. 4, 1470 (1963).

[2] D. G. Currie, T. F. Jordan and E. C. G. Sudarshan, Rev. Mod. Phys. 35, 350 (1963).

[3] H. Van Dam and E. P. Wigner, Phys. Rev. 138, 1576 (1965).

[4] J. F. Kennedy and E. H. Kerner, Am. J. Phys. 33, 463 (1965).

[5] P. Havas and J. Plebanski, Bull. Am. Phys. Soc. 5, 433 (1960).

[6] The recent discussions in quantum field theory which emphasize observables that are local in Minkowski space rather than those that have special properties in momentum space.

[7] More restricted cases of this problem have been previously considered. D. G. Currie, Ann. Phys. 36, 104 (1966), considers the case in which the force vanishes. The preprint of Ref. 4, received during the course of this work, considers the case in which the force is independent of velocity and time. Havas[8] has also implicitly touched on some of these questions and has brought to our attention a related problem treated in a different form by S. Lie, Arch. Math. Naturvidenskab 1877, p. 129.

[8] P. Havas, Nuovo Cimento, 5, 363 (1957), (see p. 371), who shows that in our one-dimensional case the problem is in fact no wider. Futher, one of the results in that paper is equivalent for one dimension to our Theorem 1b.

irrelevant. Note finally that at least one subset of the set we are looking for is well known. This is because it is well known that two Lagrangians differing only by a total time derivative of some function of q and t will yield the same equations of motion. Consequently so will the Hamiltonians constructed from them in the usual way.

We turn aside momentarily to see how two such Hamiltonians are related. Let $L(q, \dot{q}, t)$ be a Lagrangian, and consider the Lagrangian $L' = L + d\Psi/dt$, where $\Psi = \Psi(q, t)$ and $d\Psi/dt \equiv \dot{q}\partial\Psi/\partial q + \partial\Psi/\partial t$. Let $H(q, p, t)$ to the Hamiltonian obtained in the usual way from L, and $H'(q, p', t)$ be the one obtained in the usual way from L' (we write p' instead of p because p' and p will be different functions of $q, \dot{q}, t$). Let the solution for $\dot{q}$ of the equation

$$p' = L'_{\dot{q}} = L_{\dot{q}} + \Psi_q$$

be $\dot{q} = \eta(p', q, t)$, and let the solution for $\dot{q}$ of the equation

$$p = L_{\dot{q}}$$

be $\dot{q} = \xi(p, q, t)$. (We use subscripts to denote partial derivatives.) It is clear that

$$\eta(p', q, t) = \xi(p' - \Psi_q, q, t).$$

Then in the usual way,

$$\begin{aligned} H'(q, p', t) &= [p'\dot{q} - L']\,|_{\dot{q}=\eta} \\ &= p'\eta - L\,|_{\dot{q}=\eta} - \eta\Psi_q - \Psi_t \\ &= (p' - \Psi_q)\xi(p' - \Psi_q, q, t) \\ &\qquad - L[q, \xi(p' - \Psi_q, q, t), t] - \Psi_t, \end{aligned}$$

or, using the definition of H in terms of L,

$$H'(q, p', t) = H(q, p' - \Psi_q, t) - \Psi_t.$$

We say, for obvious reasons, that H' is obtained from H by a gauge transformation, or briefly by *gauging* H. Now this, as mentioned in the previous paragraph, is only a small class of an entire set of q-equivalent Hamiltonians. It is, incidentally, useful to note that a gauge transformation is canonical.

To give an example of what we mean, we turn to the simple harmonic oscillator. The conventional Hamiltonian is $H = \frac{1}{2}q^2 + \frac{1}{2}p^2$, and the equation of motion is $\ddot{q} = -q$. Now let $\Psi(q, t)$ be any differentiable function, and form the Hamiltonian

$$H = \tfrac{1}{2}q^2 + \tfrac{1}{2}(p + \partial\Psi/\partial q)^2 + \partial\Psi/\partial t. \qquad (2)$$

This Hamiltonian will yield the same equation of motion for q, but it will not yield the same canonical equations, for although the orbits in configuration space are the same, the phase-space orbits differ. The Hamiltonian of (2) can be obtained from the conventional one by the canonical (gauge) transformation[1]

$$\bar{q} = q,$$

$$\bar{p} = p + \partial\Psi/\partial q.$$

Consider, however, the Hamiltonian

$$H = 2 \ln |q \sec \tfrac{1}{2}qp|. \qquad (3)$$

It is easily verified that this one also yields $\ddot{q} = -q$, but it can also be shown that there is no canonical transformation (i.e., none that preserves Poisson brackets) that will carry one from the standard Hamiltonian to (3). It should perhaps be mentioned that the p and q of (3) will, of course, satisfy the canonical equations. But although there exists a transformation that will produce the Hamiltonian of (3), this transformation is not canonical. This is because a canonical transformation is one that carries *every* Hamiltonian system into a Hamiltonian system.[9] The transformation that yields (3) works for the harmonic oscillator, but not in general for other Hamiltonians (see Sec. IV, Example 1a). The standard Hamiltonian is, according to our definition, q-equivalent to (3). We wish to establish a procedure for finding all Hamiltonians q-equivalent to a given Hamiltonian and to parametrize the set of them.

Before attacking this problem, however, we wish to turn to an equivalent one for the Lagrangian formalism. This we do briefly in Sec. II. The Hamiltonian problem is treated in Sec. III, and some examples are given in Sec. IV.

II. THE LAGRANGIAN PROBLEM

Consider a Hamiltonian $H(q, p, t)$ derived in the usual way from a Lagrangian $L(q, \dot{q}, t)$. It is well known that essentially all the orbits in configuration space obtained from the solutions of Hamilton's canonical equations are also obtained from the Euler–Lagrange equations

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}} - \frac{\partial L}{\partial q} \equiv \mathcal{E}L = 0. \qquad (4)$$

For such orbits we may thus say that (4) implies (1), and vice versa. Note that (4) and (1) are not identically the same equation, for in (4) $\ddot{q}$ appears multi-

[9] More accurately, let $Q = Q(p, q, t)$ and $P = P(p, q, t)$ be a transformation from p, q to P, Q. Then there is a theorem that says that this transformation preserves brackets only if for *every* Hamiltonian $H(p, q, t)$ there exists a $K(P, Q, t)$ such that the canonical equations $\dot{q} = \partial H/\partial p$, $\dot{p} = -\partial H/\partial q$ are carried over into $\dot{Q} = \partial K/\partial P$, $\dot{P} = -\partial K/\partial Q$.

plied by a coefficient, while in (1) it stands alone. Now suppose we have two q-equivalent Hamiltonian H and $\bar{H}$, and let H be derived from the Lagrangian L, and $\bar{H}$ from $\bar{L}$. Then Eq. (1) implies and is implied by both $\mathcal{E}L = 0$ and $\mathcal{E}\bar{L} = 0$.

Thus we may conclude that H and $\bar{H}$ are q-equivalent Hamiltonians if and only if their Lagrangians L and $\bar{L}$ are *equivalent* in the sense that the Euler–Lagrange equations obtained from L imply and are implied by those obtained from $\bar{L}$. We wish first to study such equivalent Lagrangians, but we start, as will be seen, from a somewhat weaker requirement.

Let $L(q, \dot{q}, t)$ and $\bar{L}(q, \dot{q}, t)$ be two Lagrangians for an unconstrained system with one degree of freedom. Assume that the mass "tensors" $L_{\dot{q}\dot{q}}$ and $\bar{L}_{\dot{q}\dot{q}}$ fail to vanish or to become infinite in some region R of $(q, \dot{q}, t)$ space.[10] We call $\bar{L}$ *subordinate to* L *in* R if $\mathcal{E}L = 0$ implies that $\mathcal{E}\bar{L} = 0$. With this definition we prove the following.

Theorem 1. (a) If $\bar{L}$ is subordinate to L in R, then L is subordinate to $\bar{L}$ in R, i.e., L and $\bar{L}$ are equivalent in R. Further

$$\mathcal{E}\bar{L} = f(q, \dot{q}, t)\,\mathcal{E}L, \tag{5}$$

where $f(q, \dot{q}, t)$ is a constant of the motion (which is the ratio of the mass tensors).

(b) Consider a Lagrangian L and a constant of the motion $f(q, \dot{q}, t)$, such that neither f nor $L_{\dot{q}\dot{q}}$ vanish or become infinite in a region R of $(q, \dot{q}, t)$ space. Then there exists an $\bar{L}$ equivalent to L in R and satisfying (5). Furthermore, $\bar{L}$ is unique up to the total time derivative of some function of q and t.[8] [A general expression for $\bar{L}$ is given in Eq. (18).]

Proof. The Euler-Lagrange equation for L may be written

$$-\ddot{q}L_{\dot{q}\dot{q}} = \dot{q}L_{\dot{q}q} + L_{\dot{q}t} - L_q. \tag{6}$$

The Euler-Lagrange equation for $\bar{L}$ is similar; it may be solved for $\ddot{q}$ to yield

$$-\ddot{q} = (\dot{q}\bar{L}_{\dot{q}q} + \bar{L}_{\dot{q}t} - \bar{L}_q)/\bar{L}_{\dot{q}\dot{q}}. \tag{7}$$

By the assumption of part (a) of the theorem, (6) implies (7), so that (7) may be inserted into (6). Then some simple algebra leads to

$$\begin{aligned}\mathcal{E}L &= \ddot{q}L_{\dot{q}\dot{q}} + \dot{q}L_{\dot{q}q} + L_{\dot{q}t} - L_q \\ &= \frac{L_{\dot{q}\dot{q}}}{\bar{L}_{\dot{q}\dot{q}}}(\ddot{q}\bar{L}_{\dot{q}\dot{q}} + \dot{q}\bar{L}_{\dot{q}q} + \bar{L}_{\dot{q}t} - \bar{L}_q) = \frac{1}{f}\mathcal{E}\bar{L},\end{aligned} \tag{8}$$

where $P - \equiv \bar{L}_{\dot{q}\dot{q}}/L_{\dot{q}\dot{q}}$. Since by assumption neither $L_{\dot{q}\dot{q}}$ nor $\bar{L}_{\dot{q}\dot{q}}$ is zero or infinite in R, neither is f or $1/f$. It follows that $\mathcal{E}\bar{L} = 0$ implies $\mathcal{E}L = 0$, so that L is subordinate to $\bar{L}$. Two such Lagrangians are what we have called *equivalent.*

We now show that f is a constant of the motion. Let L and $\bar{L}$ be equivalent, thus satisfying (5) and (8) for some f. By subtracting $\ddot{q}\bar{L}_{\dot{q}\dot{q}} = \ddot{q}fL_{\dot{q}\dot{q}}$ from both sides of (5), we obtain

$$\dot{q}fL_{q\dot{q}} + fL_{\dot{q}t} - fL_q = \dot{q}\bar{L}_{\dot{q}q} + \bar{L}_{\dot{q}t} - \bar{L}_q. \tag{9}$$

Take the derivative of both sides of this equation with respect to $\dot{q}$ and use (8), arriving eventually at

$$L_{\dot{q}\dot{q}}(f_q\dot{q} + f_t) = f_{\dot{q}}(\dot{q}L_{\dot{q}q} + L_{\dot{q}t} - L_q) = 0. \tag{10}$$

This differential equation, which must be satisfied by f, imples that f is a constant of the motion. To see this, let $q(t)$ be a solution of the equation of motion (6). Then (10) becomes

$$L_{\dot{q}\dot{q}}(f_q\dot{q} + f_t + f_{\dot{q}}\ddot{q}) - L_{\dot{q}\dot{q}}(df/dt) = 0. \tag{11}$$

Since $L_{\dot{q}\dot{q}} \not\equiv 0$, this means that $df/dt = 0$, or that f is a constant of the motion. This completes the proof of part (a) of the theorem.

The remaining problem is to show that once f is chosen $\bar{L}$ exists and is unique up to the total time derivative of a function of q and t. We do this by first writing down the general solution of $\bar{L}_{\dot{q}\dot{q}} = fL_{\dot{q}\dot{q}}$. The expression so obtained is then subjected to (9), which guarantees that it will yield the same equation of motion as does L. Since this differential equation involves only derivatives with respect to $\dot{q}$, it is an ordinary differential equation, so that its solution is obtained almost trivially. Parenthetically, this is no longer true for the case of more than one dimension, when the analog of this equation becomes a set of partial differential equation.

The solution may be written

$$\begin{aligned}\bar{L}(q, \dot{q}, t) = \int_W^{\dot{q}} d\dot{q}' \int_W^{\dot{q}'} d\dot{q}''\, f(q, \dot{q}'', t)L_{\dot{q}\dot{q}}(q, \dot{q}'', t) \\ + \dot{q}A(q, t) + B(q, t),\end{aligned} \tag{12}$$

where A, B, and W are arbitrary functions of q and t, except that $\dot{q}A + B$ must be a sufficiently smooth function in R and that the surface $\dot{q} = W(q, t)$ lie in R, so that the region of integration may be contained in R. Actually the two lower limits need not be the same, but setting them equal leads to no loss of generality, since adjustment of A and B can take up any change in the lower limits. It is often convenient to pick $W = 0$ which is possible when $\dot{q} = 0$ lies in R, but we leave the choice open. The problem, once W is chosen, is to find A and B such that Eq. (9) is satisfied.

[10] For simplicity we assume that all functions have continuous first and second derivatives in R.

In proceeding we write the integral in Eq. (12) without the primes and double primes and omit the upper limit from the integral signs. It is helpful to use the following formula, obtained by integration by parts. For any function $Z(q, \dot{q}, t)$ we have

$$\int_W d\dot{q} \int_W d\dot{q}\, Z = \dot{q} \int_W d\dot{q}\, Z - \int_W d\dot{q}\, \dot{q} Z.$$

After some elementary operations, Eq. (9) may be written

$$\dot{q} f L_{\dot{q}q} + f L_{\dot{q}t} - f L_q = \int_W d\dot{q} [\dot{q} f L_{\dot{q}q} + f L_{\dot{q}t} - f L_q]_{\dot{q}}$$

$$+ \int_W d\dot{q} [L_{\dot{q}\dot{q}}(f_q \dot{q} + f_t) - f_{\dot{q}}(\dot{q} L_{\dot{q}q} + L_{\dot{q}t} - L_q)]$$

$$- f(q, W, t) L_{\dot{q}\dot{q}}(q, W, t)[W_t + W W_q] + A_t - B_q.$$

The first integral is obtained immediately; the integrand in the second integral vanishes by Eq. (10). The result is

$$f(q, W, t)[L_q - L_{\dot{q}t} - \dot{q} L_{\dot{q}q} - (W_t + W W_q) L_{\dot{q}\dot{q}}]|_{\dot{q}=W(q,t)} + A_t - B_q = 0.$$

Let the first term on the left-hand side be called $G(q, t)$. Then we must find A and B such that

$$G(q, t) + A_t - B_q = 0. \tag{13}$$

This equation is easily integrated. Two convenient forms of the solution are

$$\begin{aligned} A(q, t) &= \Psi_q - \int G\, dt, \\ B(q, t) &= \Psi_t; \end{aligned} \tag{14a}$$

and

$$\begin{aligned} A(q, t) &= \Phi_q, \\ B(q, t) &= \Phi_t + \int G\, dq. \end{aligned} \tag{14b}$$

Here Ψ and Φ are arbitrary functions of q and t, and the integrals of G are taken with respect to the one variable indicated and from an arbitrary lower limit. That is, for instance,

$$\int G\, dt \equiv \int_{\alpha(q)}^{t} G(q, t')\, dt',$$

where $\alpha(q)$ is an arbitrary function and q is treated as a parameter in the integral. Finally, if (14a) is used, (12) becomes

$$\bar{L} = \int_W d\dot{q} \int_W d\dot{q}\, f L_{\dot{q}\dot{q}} - \dot{q} \int G\, dt + \frac{d\Psi}{dt}. \tag{15a}$$

If (14b) is used, (12) becomes

$$\bar{L} = \int_W d\dot{q} \int_W d\dot{q}\, f L_{\dot{q}\dot{q}} + \int G\, dq + \frac{d\Phi}{dt}. \tag{15b}$$

Equations (15a) and (15b) differ by a total time derivative.

To recapitulate, given a Lagrangian L, one finds a constant of the motion f and then constructs a new Lagrangian $\bar{L}$ according to (15) with arbitrary W and Ψ (or Φ) (note that G is determined by the choice of f and W). It is clear that Ψ contributes a gauge transformation. We say that $\bar{L}$ is obtained by *fouling* L with f.[11] Thus, fouling L with a given constant of the motion f determines $\bar{L}$ up to gauge.

One may ask how this result is related to the folk theorem according to which two Lagrangians are equivalent if and only if they differ in gauge, i.e., by the total time derivative of a function of q and t. The difference is that for fouling we are demanding a weaker form of equivalence. For our purposes it is enough that $\mathcal{E}\bar{L} = 0$ if and only if $\mathcal{E}L = 0$. In the folk theorem, equivalence means that $\mathcal{E}\bar{L} = \mathcal{E}L$. As a particularly simple example, the folk theorem does not apply to the Lagrangians that yield $\ddot{q} = -q$ and $\alpha\ddot{q} = -\alpha q$. In fact the second Lagrangian is obtained from the first by trivially fouling it with the constant α.

III. THE HAMILTONIAN PROBLEM

Our original problem, to find the set of q-equivalent Hamiltonians, could now be solved by constructing them from the set of equivalent Lagrangians. We proceed, however, more directly.

First we make the following remark. Let $H(q, p, t)$ and $\bar{H}(q, \bar{p}, t)$ be q-equivalent. Then it follows that[12]

$$H_p(p, q, t) = \dot{q} = \bar{H}_{\bar{p}}(\bar{p}, q, t) \tag{16}$$

is an equation connecting p and $\bar{p}$. Let its solution for p be written as

$$p = P(\bar{p})$$

(we suppress the q, t dependence). Now, it is ordinarily assumed that the first equality of (16) can be solved for p in some region V of extended phase space (the direct product of time and phase space). This requires that H_{pp} fails to vanish or become infinite in V or, as we shall say, that H is good in V. Similarly $\bar{H}$ is good in a region $\bar{V}$ of extended phase space. Then we take as the domain $\bar{V}^*$ of $P(\bar{p})$

[11] This terminology is adapted from a suggestion by F. A. E. Pirani.

[12] As before, we assume that all the functions we deal with have first and second derivatives.

that part of $\bar{V}$ on which $P_{\bar{p}}$ is neither zero nor infinite. Let the range of P be $V^* = P(\bar{V}^*)$.

We now proceed to the problem. Let H and $\bar{H}$ be two Hamiltonians good in V and $\bar{V}$, respectively, and consider the regions V^* and $\bar{V}^*$, namely the range and domain of $P(\bar{p})$. We restrict our consideration for the time being to these regions. The second of Hamilton's canonical equations obtained from H is

$$H_q = -\dot{p} = -P_q\bar{H}_{\bar{p}} + P_{\bar{p}}\bar{H}_q - P_t. \tag{17}$$

We write $H_q(q, p, t)$ as a function of $q, \bar{p}, t$ by inserting $p = P(\bar{p})$, and we then take the derivative of both sides of (17) with respect to $\bar{p}$. The left-hand side becomes

$$(H_q)_{\bar{p}} = H_{qp}P_{\bar{p}}. \tag{18a}$$

Now $H_{qp} = H_{pq}$, and from (16) it follows that

$$H_{pq} = \bar{H}_{\bar{p}q} + \bar{H}_{\bar{p}\bar{p}}[\partial\bar{p}(q, p)/\partial q], \tag{18b}$$

where we have suppressed the t dependence in writing $\bar{p}$ as a function of q, p, t. Finally we use this and $P_q = -(\partial\bar{p}/\partial q)P_{\bar{p}}$ to write the derivative of (17) in the form

$$\bar{H}_{\bar{p}q}P_{\bar{p}} - \bar{H}_{\bar{p}\bar{p}}P_q = -P_{\bar{p}q}\bar{H}_{\bar{p}} - P_q\bar{H}_{\bar{p}\bar{p}} + P_{\bar{p}\bar{p}}\bar{H}_q + P_{\bar{p}}\bar{H}_{q\bar{p}} - P_{\bar{p}t}$$

or

$$dP_{\bar{p}}/dt = 0.$$

Thus $P_{\bar{p}}$ and hence also

$$f = 1/P_{\bar{p}} \tag{19}$$

is a constant of the motion.

It can be shown that, if H is obtained in the usual way from L, and $\bar{H}$ from $\bar{L}$, then $\bar{L}$ can be obtained by fouling L with the constant of the motion appearing in (19). This is seen to be reasonable when one notes that

$$\bar{H}_{\bar{p}\bar{p}} = (H_p)_{\bar{p}} = H_{pp}P_{\bar{p}} = H_{pp}/f \tag{20}$$

and recalls that, in general, if H and L are the Hamiltonian and the Lagrangian for a given physical system, then $H_{pp}L_{\dot{q}\dot{q}} = 1$. Comparison with (8) then yields the desired result immediately.

Once f and H are known, $\bar{H}$ is determined up to gauge. This may be seen by constructing both $\bar{p}$ and $\bar{H}$ by using (19) and (16). Let H be known, and let $f(q, p, t)$ be some constant of the motion. In Eq. (19) q and t are only parameters, so that (19) is easily integrated. We obtain

$$\bar{p} = \int_C^p f\,dp + R(q, t), \tag{21}$$

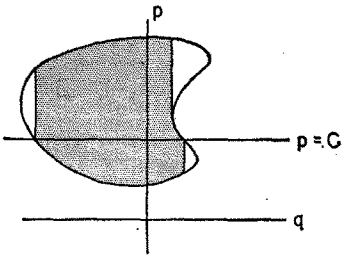


FIG. 1. Typical region in which $\bar{p}(p)$ is defined by a single integral.

where R is an arbitrary function and C is some fixed value of p within V^*. Note, incidentally that, in terms of f, V^* may now be defined as that part of V in which $0 < |f| < \infty$. It may not always be possible to construct $\bar{p}(p)$ throughout V^* by using only a single integral. The region in which (21) will work is of the kind indicated by shading in Fig. 1, where, ignoring the time dimension of extended phase space, we have let the free-form shape represent V^*. Then the region in which $\bar{p}(p)$ is defined by (21) is seen to depend on the choice of C. In general more than one such integral is needed.[13] It is also possible to choose the lower limit on the integral to be a function of q and t, but usually this is not necessary. In any case, the general solution of (19) is (21), and R can be adjusted to account for any change in the lower limit.

When (21) is inverted to give $p = P(\bar{p})$, Eq. (16) can be integrated. This yields

$$\bar{H} = \int^{\bar{p}} H_p\,d\bar{p} + S(q, t) = \int^{P(\bar{p})} H_p\frac{\partial\bar{p}}{\partial p}\,dp + S(q, t) = \int_C^{P(\bar{p})} fH_p\,dp + S(q, t). \tag{22}$$

The lower limit on the integral need not be chosen the same as in (21), but it is convenient to do so. We have thus far used three of the four canonical equations of Hamilton to obtain (21) and (22). The fourth is

$$\bar{H}_q = -\dot{\bar{p}} = \bar{p}_pH_q - \bar{p}_qH_p - \bar{p}_t.$$

With (21) and (22) this becomes

$$0 = -\int [f_qH_p + fH_{pq}]\,dp + P_{\bar{p}}\bar{p}_qfH_p - S_q + fH_q - \bar{p}_qH_p - \int [f_pH_q - f_qH_p]\,dp - R_t,$$

where we have used $P_q = -\bar{p}_qP_{\bar{p}}$ and the fact that f is a constant of the motion. Now with (19) this

[13] See Example 1a of Sect. IV.

becomes

$$-\int [fH_q]_p\, dp - S_q + fH_q - R_t = fH_q\,|_{p=C} - S_q - R_t = 0.$$

Let us write

$$fH_q\,|_{p=C} = -\Gamma(q, t). \tag{23}$$

[It can be shown that Γ is equal to the function G of Eq. (13) if $W(q, t)$ is chosen from the condition $p = L_{\dot q} = C$.] Thus the fourth canonical equation requires that R and S be chosen so that[14]

$$-\Gamma - S_q - R_t = 0. \tag{24}$$

Two convenient forms for the general solution of these equations are

$$\begin{aligned} S &= \Psi_t - \int \Gamma\, dq, \\ R &= -\Psi_q; \end{aligned} \tag{25a}$$

and

$$\begin{aligned} S &= \Phi_t, \\ R &= -\Phi_q - \int \Gamma\, dt. \end{aligned} \tag{25b}$$

Here Ψ and Φ are arbitrary functions of q and t, and the integrals are explained in Sec. II [see Eq. (14)].

Finally, if (25a) is used, (21) and (22) become

$$\begin{aligned} \bar p &= \int_C^p f\, dp - \Psi_q, \\ \bar H &= \int_C^{P(\bar p)} fH_p\, dp - \int \Gamma\, dq + \Psi_t. \end{aligned} \tag{26a}$$

If (25b) is used, (21) and (22) become

$$\begin{aligned} \bar p &= \int_C^p f\, dp - \int \Gamma\, dt - \Phi_q, \\ \bar H &= \int_C^{P(\bar p)} fH_p\, dp + \Phi_t. \end{aligned} \tag{26b}$$

Equations (26a) and (26b) differ by a gauge transformation.

What we have found then, is the following. Given a Hamiltonian H, one finds a constant of the motion and then constructs a new Hamiltonian according to (26) with arbitrary C and Ψ (or Φ) (note that Γ is determined by the choice of f and C). It is clear that Ψ contributes a gauge transformation. As in the case of the Lagrangian, we say that $\bar H$ is obtained by *fouling* H with f. Then our results can be stated in the following form.

Theorem 2. (a) Let $H(q, p, t)$ and $\bar H(q, \bar p, t)$ be q-equivalent Hamiltonians good in regions V and $\bar V$, respectively. Let $\dot q = H_p = \bar H_{\bar p}$ define the function $P(\bar p)$ through the solution $p = P(\bar p)$. Consider the subregion $\bar V^* \subset \bar V$ in which $P_{\bar p}$ is neither zero nor infinite. Then $P_{\bar p} = \bar H_{\bar p\bar p}/H_{pp}$ is a constant of the motion, and $\bar H$ can be obtained by fouling H with $f = 1/P_{\bar p}$.

(b) Consider a Hamiltonian H, good in a region V of extended phase space, and a constant of the motion $f(q, p, t)$ which fails to vanish or become infinite in V. Then fouling H with f yields a Hamiltonian $\bar H$ which is q-equivalent to H. Further, $\bar H$ is unique up to gauge.

IV. EXAMPLES

In this section we given some examples of equivalent Lagrangians and q-equivalent Hamiltonians.

Example 1. Simple Harmonic Oscillator (SHO)

Consider first the SHO, whose usual unfouled Lagrangian and Hamiltonian are, respectively,

$$\begin{aligned} L &= \tfrac{1}{2}\dot q^2 - \tfrac{1}{2}q^2, \\ H &= \tfrac{1}{2}q^2 + \tfrac{1}{2}p^2. \end{aligned}$$

1a.

Choose $W = 0$ and $f = (\frac{1}{2}q^2 + \frac{1}{2}\dot q^2)^{-1}$. Then straightforward calculation yields (up to gauge)

$$\bar L = \frac{\dot q}{q} \arctan \frac{\dot q}{q} - \ln (q^2 + \dot q^2).$$

With $C = 0$ and the same fouling constant now written $f = (\frac{1}{2}q^2 + \frac{1}{2}p^2)^{-1} = H^{-1}$, the new Hamiltonian becomes [see Eq. (3)]

$$\bar H = 2 \ln |q \sec \tfrac{1}{2}q\bar p|. \tag{27}$$

It is a simple matter to obtain $\bar H$ from $\bar L$ or vice versa.

Note that although H is good throughout extended phase space (since $H_{pp} = 1$), $\bar H$ is not good on the surface $q = 0$. This is because $P_{\bar p} = 1/f$ vanishes at $q = p = 0$. Then we see that $\bar V^*$ does not include the origin, and (referring to Fig. 1) the region in which our construction defines $\bar H$ is shown in Fig. 2. By choosing C positive, $\bar H$ can be

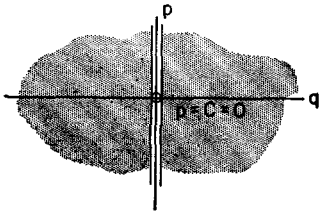


FIG. 2. Region $\bar V^*$ for SHO Hamiltonian of Eq. (27).

[14] More generally, if the lower limit on the integral in (21) is taken as $C_1(q, t)$, and the lower limit on the integral in (22) is taken as $C_2(q, t)$, we set

$$-\Gamma = [f(H_pC_{2q} + H_q)]\,|_{p=C_2} + [fC_{1t}]|_{p=C_1}.$$

defined everywhere above the q axis, or by choosing C negative it can be defined below.
The more general result is

$$\bar{H}' = 2 \ln |q \sec [\tfrac{1}{2}\bar{p}q - \arctan (C/q)]|. \quad (28)$$

That (28) and (27) differ only in gauge where they are both good is seen when (28) is written in the form

$$\bar{H}' = 2 \ln |q \sec \{\tfrac{1}{2}q[\bar{p} - (2/q) \arctan (C/q)]\}|.$$

In general if

$$H'(q, p, t) = H(q, p - h, t),$$

where h is a function of q alone, then H and H' differ in gauge alone.

1*b*. It is possible also to foul H so that $\bar{H}$ depends on time. All that is needed is a time-dependent constant of the motion f (i.e. one for which $\partial f/\partial t \neq 0$, although, of course, $df/dt = 0$). Choose $C = 0$ and $f = (p \cos t + q \sin t)^{-1}$. Then by somewhat tedious calculation one obtains

$$\bar{H} = q \tan t \times \sec t[(e^{\bar{p}} \cot \tfrac{1}{2}t)^{\cos t} - \ln (e^{\bar{p}} \cot \tfrac{1}{2}t)^{\cos t} - 1],$$

which is obviously well fouled.

Example 2. If a Hamiltonian is of the form $H = T(p) + V(q)$, that is, if it is the sum of a part depending only on p and another depending only on q, we shall call it *separated*. Given an unseparated H, it is sometimes possible to separate it by fouling. When this is possible, we say that H is *separable*. An example of an unseparated but separable Hamiltonian is given in Eq. (2). It is separated by the gauge transformation shown there (of course, a gauge transformation is a special case of fouling).

Under what conditions can H be separated? Let H and $\bar{H}$ be q-equivalent Hamiltonians, and let $\bar{H}$ be separated: $\bar{H}_{\bar{p}q} = 0$. Further, let $\bar{H}$ be obtained by fouling H with f. Then by combining (18b), (19), and (20), we find that $H_{pq} = H_{pp}\bar{p}_q/f$. Now we use (21) to write $\bar{p}_q$, which yields

$$H_{pq} = H_{pp} \frac{1}{f}\left[\int f_q \, dp + R_q(q, t)\right].$$

Rearrange the terms and take the derivative with respect to p, obtaining

$$f_q = (fH_{pq}/H_{pp})_p = f_p(H_{pq}/H_{pp}) + f(H_{pq}/H_{pp})_p.$$

Since f is a constant of the motion, $f_qH_p = f_pH_q - f_t$. Then if we multiply both sides of the above equation by H_p, some simple algebra will lead to

$$[(f/H_{pp})\{H, H_p\}]_p = f_t, \quad (29)$$

where $\{\ \}$ is the Poisson bracket (with respect to q and p). Thus if H is separable, there must exist a constant of the motion $f(q, p, t)$ such that (29) can be satisfied. It is helpful to note that since $H_p = \dot{q}$, the Poisson bracket in (29) is just equal to $-\ddot{q}$.

We make use of this separation procedure in the following problem. Consider the Lagrangian

$$L = \tfrac{1}{2}(1 + \dot{q}^2) \exp \{-2V(q)\}. \quad (30)$$

This Lagrangian is of interest because it has certain quasi-relativistic properties which will be discussed more fully in a later paper. In particular, its equation of motion is

$$\ddot{q} = -(1 - \dot{q}^2)V',$$

and it is easily seen that if the velocity $\dot{q}$ is less than 1 it will never surpass 1. The Hamiltonian obtained in the usual way from this L is

$$H = \tfrac{1}{2}p^2 \exp (2V) - \tfrac{1}{2} \exp (-2V), \quad (31)$$

which is not separated, but, as we shall show, is separable.

Let us try to find a time-independent constant of the motion satisfying (29). For the Hamiltonian of (31) this means that we are looking for an f such that $fV'H$ is independent of p. Obviously the p dependence of this product can be eliminated by setting $f = aH^{-1}$, where a is any function of q. In particular, since H itself is a constant of the motion (it does not depend explicitly on the time), a may be any number. Thus by fouling H with such an f we may be able to separate it.

Let us start by using (26a) to calculate $\bar{p}$. We have

$$\bar{p} = -2a[\tanh^{-1} (pe^{2V}) - \tanh^{-1} (Ce^{2V})],$$

where a and C are yet to be chosen. Clearly it is most convenient to choose $a = -\frac{1}{2}$, $C = 0$. Then we have

$$\bar{p} = \tanh^{-1} (pe^{2V}). \quad (32)$$

It is now a simple matter to use the rest of (26a) to calculate $\bar{H}$. The result is

$$\bar{H}(q, \bar{p}) = \ln [\cosh \bar{p}] + V. \quad (33)$$

This separated Hamiltonian has several interesting relativistic properties which will be discussed in the subsequent publication already mentioned.

It should be pointed out that not all Hamiltonians are separable. This may be seen as follows. Let $H = T(p) + V(q)$ be separated. Then it is easily shown that the equation of motion obtained from H is

$$\ddot{q} = -S(\dot{q})V'(q), \quad (34)$$

where $S(\dot{q})$ is equal to $T''(p)$ evaluated for p given

by the solution of $T'(p) = \dot{q}$ (the primes indicate differentiation with respect to p). Thus if a Hamiltonian is to be separable, the equation of motion obtained from it must be of the form of (34), that is the "force" must be a product of a function of q with a function of $\dot{q}$. The equation of motion of our example is of this form, which is why we were able to separate the Hamiltonian.

V. CONCLUSION

The principal motivation for the present investigation lies in its relevance to the study of equations of motion which are relativistically invariant.[15] When such equations are found, it is desirable to imbed the dynamics in a Hamiltonian formalism. But as we have seen, the Hamiltonian is far from uniquely determined by the equations. It is chosen usually with other considerations in mind, such as that it equal the energy, that it possess certain cluster decomposition properties, or that it transform in some special way under the Poincaré group. To attempt to satisfy as many such properties as possible it is helpful to classify the set of all Hamiltonians which yield a given equation. It is, moreover, useful to know what properties cannot be satisfied by a Hamiltonian. (For instance, there is no separable Hamiltonian that will yield $\ddot{q} = q + \dot{q}^2$.)

A second motivation is related to the representations of the Poincaré group (or more generally any relativity group). A particular Hamiltonian generates a representation on phase space (see the paper cited in Ref. 7), and thus by studying all the Hamiltonians associated with a given equation of motion we can consider the class of representations all of which give the same dynamics. This class includes representations with widely varying properties. For instance, $H = (p^2 + m^2)^{\frac{1}{2}}$ and $H' = p^2/2m$ both yield the free-particle equations, but H generates the usual canonical representation of the Poincaré group, while H' generates a noncanonical one.

[15] D. G. Currie, Phys. Rev. 142, 817 (1966).

The general problem of finding all Hamiltonians equivalent to a given one has been solved here only for a very special case, namely for one-dimensional particle motion. It is interesting to extend these results to the case of n degrees of freedom, $n > 1$. Especially interesting is the case of two particles, for then the study of relativistic interactions becomes nontrivial. However the extension is not trivial, for all the ordinary differential equations in one dimension become partial differential equations in n dimensions. We are at present studying this problem. Other extensions of these considerations are in the direction of quantum mechanics and field theory. The fouled Hamiltonian of Eq. (27), for instance, will obviously lead to a treatment of the harmonic oscillator differing greatly from the usual quantum mechanical treatment, for when p (we have dropped the bar from $\bar{p}$) is replaced as always by $i\,\partial/\partial q$, Schrödinger's equation becomes a linear differential equation of infinite order. It is thus obvious that fouling can be used to complicate a simple problem, but it is also perhaps possible to use it to simplify complicated problems. It should be noted also that the Hamiltonian is actually changed to a new dynamical variable by fouling; it is not as though we were merely changing the functional form of a variable by going to a new coordinate system. Thus its spectrum will change, sometimes radically. In the classical case, for instance, we saw for the simple harmonic oscillator that the values that the Hamiltonian can take on may shift from the positive real line to the entire real line. Finally, gauging of particle Hamiltonians is the analog of the gauging normally encountered in field theory. In the same way, fouling, which is a generalization of gauging, has an analog in field theory, but this analog is even more difficult to find than it is to extend the particle case to more than one dimension. Although there are many formal analogies between our work and field theory, we are not prepared at this time to go into these considerations.

On the Development of the Covariant Formulation of the Conservation Laws of General Relativity

W. R. DAVIS AND M. K. MOSS
Department of Physics, North Carolina State University, Raleigh, North Carolina
(Received 20 September 1965)

A method of obtaining Komar's covariant formulation of the conservation laws of general relativity directly from the variation of the scalar curvature density is presented. The procedure of obtaining this expression is free of the addition of arbitrary elements and only tensorial terms and operations are employed in the development.

USING the results of Bergmann[1] and Møller,[2] Komar[3] first obtained a covariant formulation for the conservation laws of the general theory of relativity. This formulation is given by the expression

$$\{[(-g)^{\frac{1}{2}}(\xi^{i;j} - \xi^{j;i})]_{;j}\}_{;i} \equiv \{[(-g)^{\frac{1}{2}}(\xi^{i;j} - \xi^{j;i})]_{;j}\}_{,i} \equiv 0. \quad (1)$$

While expression (1) is evidently the simplest covariant expression of its type that is identically satisfied, it is nevertheless of fundamental interest to investigate the considerations which lead to its development.

Here we present a method of directly obtaining expression (1) from the equations guaranteeing the general covariance of the field equations of general relativity. Aside from the simplicity of this method, it is seen, in contrast to other formulations, that every step of this development involves only tensorial terms.

Starting with the invariant Lagrangian density $(-g)^{\frac{1}{2}} R$ of the general theory of relativity,[4] we consider its variation under the infinitesimal coordinate transformation $\bar{x}^i = x^i + \epsilon\xi^i$. This variation can be immediately written as

$$\delta[(-g)^{\frac{1}{2}}R] = \delta[(-g)^{\frac{1}{2}}g^{ij}R_{ij}] = -(-g)^{\frac{1}{2}}G^{ij}\,\delta g_{ij} + (-g)^{\frac{1}{2}}g^{ij}\,\delta R_{ij}, \quad (2)$$

where $G^{ij} \equiv R^{ij} - (\frac{1}{2})Rg^{ij}$.

Using the following well-known results:

$$\delta R_{ij} = \delta\left\{{k \atop ik}\right\}_{;j} - \delta\left\{{k \atop ij}\right\}_{;k},$$

$$\delta\left\{{i \atop jk}\right\} = \tfrac{1}{2}g^{il}[\delta g_{jl;k} + \delta g_{kl;j} - \delta g_{jk;l}],$$

$$\delta g_{ij} = -\epsilon(\xi_{i;j} + \xi_{j;i}),$$

where the $\left\{{i \atop jk}\right\}$ are the usual Christoffel symbols of the second kind, Eq. (2) can be written in the form

$$\delta[(-g)^{\frac{1}{2}}R] = \epsilon\,(-g)^{\frac{1}{2}}[2G^{ij}\xi_{i} - (g^{kl}g^{ij} - g^{ki}g^{jl})(\xi_{k;l} + \xi_{l;k})_{;i}]_{;j}. \quad (3)$$

Substituting the Ricci identity in the form

$$\xi_{k;l;i} = \xi_{k;i;l} + \xi_{m}R^{m}_{kli}$$

into the second term of (3) then yields,

$$\delta[(-g)^{\frac{1}{2}}R] = \epsilon\,(-g)^{\frac{1}{2}}[2G^{ij}\xi_{i} - g^{kl}g^{ji}R^{m}_{kli}\xi_{m} + g^{ki}g^{jl}R^{m}_{kli}\xi_{m} - \tfrac{1}{2}(g^{kl}g^{ij} - g^{il}g^{jk})\xi_{k;i;l}]_{;j}. \quad (4)$$

Consider now the following terms:

$$g^{kl}g^{ji}R^{m}_{kli} = g^{kl}g^{ij}g^{mh}R_{hkli} = g^{ij}g^{mh}R^{l}_{hil} = g^{ij}R^{m}_{i} = R^{jm};$$

$$g^{ki}g^{jl}R^{m}_{kli} = g^{ki}g^{jl}g^{mh}R_{hkli} = -g^{jl}g^{mh}R^{i}_{hli} = -R^{jm}.$$

Thus, using these results, Eq. (4) becomes

[1] P. G. Bergmann, Phys. Rev. **112**, 287 (1958).
[2] C. Møller, Ann. Phys. **4**, 347 (1958).
[3] A. Komar, Phys. Rev. **113**, 934 (1959).
[4] C. Møller, (Kgl. Danske Videnskab, Selskab Mat. Fys. Medd. **31**, No. 14, 1959), and B. Laurent, Nuovo Cimento **11**, 740 (1959) have shown that the "superpotential" underlying the Møller pseudo tensor can be constructed directly from the differential identities which are obtained by varying the invariant Lagrangian $(-g)^{\frac{1}{2}} R$ under a general infinitesimal coordinate transformation $\bar{x}^i = x^i + \epsilon\xi^i$. In particular, the terms of this variational expression are grouped as the nontensorial coefficients of the ξ^i, $\xi^i{}_{,j}$, $\xi^i{}_{,j,k}$, $\xi^i{}_{,j,k,l}$ and then, after appropriate symmetrization, set identically equal to zero. [In this manner one obtains the most general set of identities following in consequence of the demand of the general coordinate invariance of the variational principle underlying the theory. See, for example, W. R. Davis, Z. Physik **148**, 1 (1957)]. Thus, it is in this connection that Møllers superpotential follows from the variation of $(-g)^{\frac{1}{2}} R$. Using the results of Møller, Komar showed that one could obtain a tensor formulation which included Møller's nontensorial expression as a special case. By way of contrast, it is noted in this paper that Komar's tensor formulation is directly obtained from the variation of the scalar curvature density $(-g)^{\frac{1}{2}} R$ in a manner which is free of the addition of arbitrary elements.

$$\delta[(-g)^{\frac{1}{2}}R] = \epsilon\,(-g)^{\frac{1}{2}}[2G^{ij}\xi_i - 2R^{ij}\xi_i - \tfrac{1}{2}(g^{kl}g^{ij} - g^{il}g^{jk})\xi_{k;i;l}]_{;j},$$

which may be immediately rewritten in the form

$$\delta[(-g)^{\frac{1}{2}}R] = \epsilon\{(-g)^{\frac{1}{2}}[-R\xi^j - \tfrac{1}{2}(\xi^{l;j} - \xi^{j;l})_{;l}]\}_{;j}. \quad (5)$$

Since the transformation law of the invariant scalar density Lagrangian $(-g)^{\frac{1}{2}}R$ under an infinitesimal coordinate transformation is given by $\delta[(-g)^{\frac{1}{2}}R] = -\epsilon[(-g)^{\frac{1}{2}}R\xi^i]_{,i}$, Eq. (5) becomes

$$-[(-g)^{\frac{1}{2}}R\xi^i]_{,i} = \{(-g)^{\frac{1}{2}}[-R\xi^j - \tfrac{1}{2}(\xi^{l;j} - \xi^{j;l})_{;l}]\}_{;j}. \quad (6)$$

Then noting that

$$[(-g)^{\frac{1}{2}}R\xi^i]_{;i} = [(-g)^{\frac{1}{2}}R\xi^i_{;i}]_{;i},$$

Eq. (6) results in the simple expression

$$[(-g)^{\frac{1}{2}}(\xi^{l;j} - \xi^{j;l})_{;l}]_{;j} \equiv 0. \quad (7)$$

Expression (7) is precisely the tensor expression that was first presented in the literature by Komar. In terms of the above development, it has been shown that (7) can be obtained from the infinitesimal transformation law of the invariant Lagrangian $(-g)^{\frac{1}{2}}R$ in a manner which is free of the addition of arbitrary elements and where only tensor quantities and tensor operations are used in the development.[5]

Clearly, the expression (7) is arbitrary to the extent that one can still add to it any quantity C^j where $C^j_{,j} = 0$. However, additive terms of this type will in general destroy the tensor properties and/or the simple form of the resulting expression.

---

[5] For a discussion of this expression and its application, see W. R. Davis and M. K. Moss, Nuovo Cimento **27**, 1492 (1963); *ibid.* **38**, 1531, 1558 (1965); also see A. Trautman, *Lectures on General Relativity* (Prentice-Hall, Inc., Englewood Cliffs, New Jersey 1964), pp. 185–188.

Upper and Lower Bounds for Canonical Ensemble Averages*

HAROLD FALK
Department of Physics, University of Pittsburgh, Pittsburgh, Pennsylvania
(Received 1 October 1965)

Upper and lower bounds are obtained for canonical ensemble averages. The bounds are expressed entirely in terms of averages with respect to an arbitary unperturbed canonical ensemble density operator. A weak form of the derived bounds is used to show that, for magnetic fields exceeding a given critical value, the magnetization of the antiferromagnetic linear chain approaches ferromagnetic saturation as the temperature approaches absolute zero.

I. INTRODUCTION

UPPER and lower bounds for the canonical ensemble average $\langle Q\rangle$ of an operator Q are formulated by combining certain convexity properties of the free energy[1,2] with a recently published inequality of Golden.[3] The bounds are expressed entirely in terms of averages with respect to an arbitrary unperturbed canonical ensemble.

Essential to the present formulation is Fisher's useful observation[1] that upper and lower bounds for a convex function lead directly to upper and lower bounds for the first derivative of the function.

As an application of a weak form of the derived bounds, it is shown that the magnetization of the antiferromagnetic linear chain approaches ferromagnetic saturation[4] as the temperature approaches absolute zero for values of the magnetic field satisfying $g\mu\mathfrak{H}/2J > 2$.

II. DEFINITIONS AND DERIVATIONS

Let H and H_0 denote Hermitian Hamiltonians and write

$$H = H_0 + (H - H_0) \equiv H_0 + H_1.$$

In terms of a Hermitian operator Q define

$$F(\gamma) \equiv (-\beta)^{-1} \ln \mathrm{Tr} \exp [-\beta(H + \gamma Q)],$$

where γ is an arbitrary real parameter to which we assign the dimension of energy, and $\beta \equiv 1/\kappa T$ denotes (Boltzmann's constant times the absolute temperature)$^{-1}$. In this paper any operator whose trace is indicated is required to be bounded. We do not explicitly note the dependence of $F(\gamma)$ on β, H, and Q, which are to be regarded as fixed.

In a canonical ensemble with a density

$$\rho(H, \beta) \equiv \exp(-\beta H)/\mathrm{Tr} \exp(-\beta H),$$

the Helmholtz free energy is

$$F(H, \beta) = F(0),$$

and the ensemble average of Q is

$$\langle Q\rangle \equiv \mathrm{Tr}\,[\rho(H, \beta)Q] = F'(0).$$

The unperturbed quantities $F(H_0, \beta)$ and $\langle Q\rangle_0$ are similarly defined by replacing H by H_0 in the above definitions.

The bounds to be derived are

$$\langle Q\rangle_0 - g(\gamma_1) \leq \langle Q\rangle \leq \langle Q\rangle_0 - g(\gamma_2), \quad \gamma_2 < 0 < \gamma_1, \tag{1}$$

where

$$g(\gamma) \equiv \langle Q\rangle_0 + \gamma^{-1}[F^+(0) - F^-(\gamma)] =$$
$$= (\gamma\beta)^{-1} \ln \langle \exp\{-\beta[(H_1 - \langle H_1\rangle_0) + \gamma(Q - \langle Q\rangle_0)]\}\rangle_0,$$

and $F^\pm$ will be given subsequently.

The derivation follows.

Since $F(\gamma)$ is a convex (upward) function[2] of γ, the function is everywhere below its tangent; thus, provided that the indicated derivative exists,

$$F(\gamma) \leq F(0) + \gamma F'(0) = F(H, \beta) + \gamma\langle Q\rangle.$$

The weak form of Peierls' inequality[5] is

$$F(H, \beta) \leq F(H_0, \beta) + \langle H_1\rangle_0,$$

so that

$$F(\gamma) \leq F^+(\gamma) \equiv F(H_0, \beta) + \langle H_1\rangle_0 + \gamma\langle Q\rangle. \tag{2}$$

Now Golden has shown[3] that

$$\mathrm{Tr} \exp [-\beta(H_0 + H_1 + \gamma Q)]$$
$$\leq \mathrm{Tr}\,\{\exp(-\beta H_0) \exp [-\beta(H_1 + \gamma Q)]\};$$

therefore

$$F(\gamma) \geq F^-(\gamma) \equiv$$
$$\equiv F(H_0, \beta) + \langle H_1\rangle_0 + \gamma\langle Q\rangle_0 - \gamma g(\gamma). \tag{3}$$

* This work was supported in part by the Air Force Office of Scientific Research under Grant No. AF 196-63.

[1] M. E. Fisher, J. Chem. Phys. **42**, 3852 (1965).
[2] R. B. Griffiths, J. Math. Phys. **5**, 1215 (1964).
[3] S. Golden, Phys. Rev. **137**, B1127 (1965); C. J. Thompson, J. Math. Phys. **6**, 1812 (1965).
[4] R. B. Griffiths, Phys. Rev. **133**, A768 (1964).
[5] See, for example, H. Falk, Physica **29**, 1114 (1963).

The inequality (2) may be written

$$F(\gamma) \leq F^+(0) + \gamma\langle Q\rangle,$$

which, in combination with (3), gives for $\gamma_2 < 0 < \gamma_1$,

$$[F^-(\gamma_1) - F^+(0)]/\gamma_1 \leq \langle Q\rangle \leq [F^-(\gamma_2) - F^+(0)]/\gamma_2,$$

thus completing the derivation of (1).

It should be noted that the bounds contain unperturbed averages only, and that the optimization conditions presented by Michael Fisher[1] are certainly applicable, in principle, to (1). One also observes that if $\langle Q\rangle$ and $\langle Q\rangle_0$ are extensive variables in terms of N, the number of particles in the system, then a sufficient condition for $\langle Q\rangle/N = \langle Q\rangle_0/N$ in the thermodynamic limit[6] is that γ_1 and γ_2 can be found for which the following conditions obtain for all temperatures:

$$\lim \text{therm}\ \frac{g(\gamma_{1,2})/N}{\langle Q\rangle_0/N} = 0, \qquad (\gamma_2 < 0 < \gamma_1),$$

where $\langle Q\rangle_0/N \neq 0$.

It is possible to weaken (1) in order to gain tractability. For example, one can use the fact that

$$|\mathrm{Tr}\ A^\dagger B| \leq (\mathrm{Tr}\ A^\dagger A)^{\frac{1}{2}}\,(\mathrm{Tr}\ B^\dagger B)^{\frac{1}{2}},$$

which is an expression of the Cauchy–Schwarz inequality as applied to square matrices A and B. An immediate consequence is

$$\ln \mathrm{Tr}\,\{\exp(-\beta H_0)\exp[-\beta(H_1+\gamma Q)]\} \leq \tfrac{1}{2}\ln \mathrm{Tr}\exp(-2\beta H_0) + \tfrac{1}{2}\ln \mathrm{Tr}\exp[-2\beta(H_1+\gamma Q)];$$

thus

$$F(\gamma) \geq F^-(\gamma) \geq \bar{F}^-(\gamma) \equiv \equiv F(H_0, 2\beta) + F(H_1+\gamma Q, 2\beta).$$

Use has been made of the positive semidefiniteness of the trace of the product of two positive semidefinite Hermitian matrices. By introducing

$$\bar{g}(\gamma) \equiv \langle Q\rangle_0 + \gamma^{-1}[F^+(0) - \bar{F}^-(\gamma)] = \gamma^{-1}\{[(\langle H_1\rangle_0 + \gamma\langle Q\rangle_0) - F(H_1+\gamma Q, 2\beta)] + [F(H_0,\beta) - F(H_0, 2\beta)]\},$$

and applying the above inequality, it is easily verified that $g(\gamma_1) \leq \bar{g}(\gamma_1)$ and $g(\gamma_2) \geq \bar{g}(\gamma_2)$. Hence (1) yields the weaker bounds,[7]

$$\langle Q\rangle_0 - \bar{g}(\gamma_1) \leq \langle Q\rangle \leq \langle Q\rangle_0 - \bar{g}(\gamma_2), \quad (\gamma_2 < 0 < \gamma_1). \tag{4}$$

In summary,

$$\langle Q\rangle_0 - \bar{g}(\gamma_1) \leq \langle Q\rangle_0 - g(\gamma_1) \leq \langle Q\rangle \leq \leq \langle Q\rangle_0 - g(\gamma_2) \leq \langle Q\rangle_0 - \bar{g}(\gamma_2).$$

III. AN APPLICATION

As an application of (4) consider the problem of obtaining bounds for the magnetization associated with the N-spin model known as the antiferromagnetic linear chain. The appropriate quantities are

$$\frac{H_0}{2J} = \sum_{f=1}^{N}(S_f^x S_{f+1}^x + S_f^y S_{f+1}^y) + \frac{h}{2}\sum_{f=1}^{N} S_f^z,$$

$$\frac{H_1}{2J} = \sum_{f=1}^{N} S_f^z S_{f+1}^z + \frac{h}{2}\sum_{f=1}^{N} S_f^z,$$

and

$$Q = \sum_{f=1}^{N} S_f^z.$$

Here $\mathbf{S}_f$ denotes the spin operator for the spin ($s = \frac{1}{2}$) at lattice site f on the chain which is closed so that $\mathbf{S}_{f+N} \equiv \mathbf{S}_f$. With g, μ, $\mathcal{H}$, and J symbolizing the Landé factor, the Bohr magneton, the external magnetic field, and the exchange energy, respectively, define $h \equiv g\mu\mathcal{H}/2J$. It is also convenient to introduce

$$h^* \equiv \gamma/2J + h/2$$

so that

$$\frac{(H_1+\gamma Q)}{2J} = \sum_{f=1}^{N} S_f^z S_{f+1}^z + h^*\sum_{f=1}^{N} S_f^z.$$

The total Hamiltonian $H = H_0 + H_1$ was split into two separately soluble parts,[7] and the Zeeman term was equally divided between the two parts. Explicitly, H_0 contains the "x–y" model Hamiltonian[8] and $H_1 + \gamma Q$ contains the antiferromagnetic Ising-model Hamiltonian[9]—both exactly soluble in the presence of a magnetic field. The details of transform-

[6] The notation lim therm is used in the sense of M. D. Girardeau, J. Math. Phys. **6**, 1083 (1965).

[7] Dr. R. B. Griffiths has kindly pointed out that the weak bounds can alternatively be derived by using, instead of (3), the lower bound for $F(\gamma)$ obtained from the "chord property" of a convex function (see Fig. 1 of Ref. 2). It is interesting to note that our application of the weak bounds implicitly utilizes the free energy lower bound mentioned in Eq. (13) of R. B. Griffiths, Phys. Rev. **136**, A751 (1964).

[8] See, for example, H. Falk and Th. W. Ruijgrok, Phys. Rev. **139**, A1203 (1965); references to earlier work are contained therein.

[9] The ferromagnetic Ising chain is discussed by K. Huang, *Statistical Mechanics* (John Wiley & Sons, Inc., New York, 1963), p. 346. To transform to the notation of the present antiferromagnetic model, replace his ϵ and $|B|$ by $-J/2$ and $g\mu\ |H|/2$, respectively; his λ_+ given by (16.74) then corresponds to our $\lambda_{\max}$. However, to proceed further one should rederive his expressions (16.76) and (16.77) which contain errors.

ing H_0 to a diagonal, bilinear, Fermi operator form and calculating $F(H_0, \beta)$ and $\langle H_1\rangle_0$ are readily accessible in references cited.[8] Only the required results will be written below. The notation is consistent with that in Ref. 8 unless otherwise specified.

For the x–y model, including the Zeeman term,

$$\frac{F(H_0, \frac{1}{2}J\theta)}{2JN} = \frac{\theta}{N}\sum_k \ln(1 - n_k),$$

where

$$n_k = \{\exp[(\tfrac{1}{2}h - \cos k)/\theta] + 1\}^{-1}, \quad (\theta = \kappa T/2J).$$

In terms of these quantities

$$\frac{\langle Q\rangle_0}{N} = \frac{1}{N}\sum_k (n_k - \tfrac{1}{2}) \equiv -\frac{\sigma_0}{2},$$

$$\frac{\langle H_1\rangle_0}{2JN} = \frac{\sigma_0^2}{4} - \eta_0^2 - \frac{h\sigma_0}{4},$$

where

$$\eta_0 \equiv \frac{1}{N}\sum_k n_k \cos k.$$

For the Ising model[9] with $N \gg 1$,

$$\frac{F(H_1 + \gamma Q, \frac{1}{2}J\theta)}{2JN} = -\theta \ln \lambda_{\max},$$

where

$$\lambda_{\max}(h^*, \theta) = e^{-\frac{1}{2}\theta}\{\cosh(h^*/2\theta) + [\sinh^2(h^*/2\theta) + e^{1/\theta}]^{\frac{1}{2}}\}.$$

It is seen that as $\theta \to 0$ for $h > 2$; $\sigma_0 \to 1$, $\eta_0 \to 0$, $\langle H_1\rangle_0/2JN \to \frac{1}{4} - h/4$, $\gamma\langle Q\rangle_0/2JN \to -\tilde{\gamma}/2$, where $\tilde{\gamma} \equiv \gamma/2J$.

Asymptotically as $\theta \to 0$ for $h^* > 2$ and $|\tilde{\gamma}| \ll 1$,

$$\frac{F(H_1 + \gamma Q, \frac{1}{2}J\theta)}{2JN} \sim \frac{1 - |h|}{4} - \frac{\tilde{\gamma}}{2}\operatorname{sgn}(h) + 0(\tilde{\gamma}^2) = \frac{1 - h}{4} - \frac{\tilde{\gamma}}{2} + 0(\tilde{\gamma}^2).$$

Clearly then, $|\bar{g}|/N \sim |0(\tilde{\gamma})|$ which we may choose to be as small a positive value as desired so that $\langle Q\rangle/N \to \langle Q\rangle_0/N = -\frac{1}{2}$.

This demonstrates that the magnetization of the antiferromagnetic linear chain approaches ferromagnetic saturation as the temperature approaches absolute zero for $g\mu\mathfrak{H}/2J > 2$.

ACKNOWLEDGMENTS

I would like to thank Dr. Robert Griffiths for his comments, which have been very helpful. A conversation with Dr. Levere Hostler is also gratefully acknowledged.

Finite-Dimensional Representations of Some Non-Semisimple Lie Algebras

CLAUDE GEORGE
Institut Elie Cartan, Nancy, France
AND
MONIQUE LÉVY-NAHAS
Centre de Physique Théorique de l'Ecole Polytechnique, Paris, France
(Received 6 December 1965)

We study the finite-dimensional representations of non-semisimple Lie algebras. We give some general properties, and apply them to the case of the motion group and of the inhomogeneous Lorentz group.

INTRODUCTION

THE finite-dimensional representations of non-semisimple Lie groups (or Lie algebras) have until now not been extensively used in physics, because they are not, in general, unitary. However, the study of the representations of the inhomogeneous Lorentz group by Bargmann and Wigner in 1948, furnishes an example of construction of infinite-dimensional unitary representations starting from finite-dimensional nonunitary representations of the homogeneous Lorentz group. Likewise, the finite-dimensional representations of the "homogeneous" Galilei group (which is isomorphic to the motion group $T \times R$) are of interest in physics.[1] The aim of this paper is to study the finite-dimensional representations of any Lie algebra by using the well-known fact that any Lie algebra is an extension of a semisimple Lie algebra.

In Sec. I, we give some general features of the representations. Then we compute in Sec. II, the finite-dimensional representations of the motion group and in Secs. III, IV, those of the inhomogeneous Lorentz group (in fact of their Lie algebras). In Appendix II we give the connection between this last calculation and the derivation of invariant wave equations.

In the following, K is a field of characteristic 0 (in fact **R** or **C**). Lie algebras are of finite dimension and by representation we always mean finite-dimensional representation.

I. SOME DEFINITIONS AND STANDARD THEOREMS[2,3]

A. Extension of a Semisimple Lie Algebra by a Lie Algebra

Let $\mathfrak{G}$ and $\mathfrak{a}$ be any Lie algebras. $\tilde{\mathfrak{G}}$ is called the extension of $\mathfrak{G}$ by $\mathfrak{a}$ if $\mathfrak{a}$ is isomorphic to an ideal $\mathfrak{a}'$ of $\tilde{\mathfrak{G}}$ and if the quotient $\tilde{\mathfrak{G}}/\mathfrak{a}'$ is isomorphic to $\mathfrak{G}$. The extension is called inessential if $\mathfrak{G}$ is isomorphic to a semidirect product of $\mathfrak{a}$ by $\mathfrak{G}$.[4]

Theorem: An extension of a semisimple algebra is always inessential.

In the following, $\mathfrak{G}$ will be a semisimple Lie algebra, so we are concerned only with inessential extensions. Let us consider

$$0 \to \mathfrak{a} \xrightarrow{i} \tilde{\mathfrak{G}} \underset{k}{\overset{\Delta}{\rightleftarrows}} \mathfrak{G} \to 0.$$

We identify $\mathfrak{a}$ to an ideal of $\tilde{\mathfrak{G}}$ by i, and $\mathfrak{G}$ to a subalgebra of $\tilde{\mathfrak{G}}$ by k. $\mathfrak{a}$ and $\mathfrak{G}$ are supplementary in $\tilde{\mathfrak{G}}$, which we note $\tilde{\mathfrak{G}} = \mathfrak{a} \oplus \mathfrak{G}$. Recalling that, for a Lie algebra $\mathfrak{G}$, and $X \in \mathfrak{G}$, ad X is the linear mapping of $\mathfrak{G}$ into $\mathfrak{G}$, defined by: $Y \in \mathfrak{G} \to [X, Y] \in \mathfrak{G}$; ad X is a derivation of $\mathfrak{G}$, that is:

$$\text{ad } X([Y, Z]) = [\text{ad } X(Y), Z] + [Y, \text{ad } X(Z)].$$

Let J be an ideal in $\mathfrak{G}$; $\text{ad}_J X$ is the restriction of ad X on J.

There is an homomorphism θ of the Lie algebra $\mathfrak{G}$ into the Lie algebra of derivations on $\mathfrak{a}$. θ is defined by

$$\theta(X) = \text{ad}_{\mathfrak{a}} X, \qquad \forall X \in \mathfrak{G}.$$

In $\tilde{\mathfrak{G}}$ we have, when $A \in \mathfrak{a}$, $B \in \mathfrak{a}$, $X \in \mathfrak{G}$, and $Y \in \mathfrak{G}$:

$$[A + X, B + Y] = [A, B] + \theta(X)B - \theta(Y)A + [X, Y].$$

By the mapping θ, $\mathfrak{a}$ becomes a $\mathfrak{G}$-module. Let us recall that a $\mathfrak{G}$-module is a couple $(V; \rho)$, where V is a vector space and ρ a representation: $X \to \rho(X)$ of the Lie algebra $\mathfrak{G}$ into the Lie algebra $\mathfrak{L}(V)$. Then we write

$$X_V \cdot v = \rho(X) \cdot v, \quad \text{where} \quad v \in V, \qquad X \in \mathfrak{G}.$$

[1] J. M. Lévy-Leblond (to be published).
[2] All these results can be found in J. Dixmier, *Algèbres de Lie* (C. D. U. 1962).
[3] See also: N. Jacobson, *Lie Algebras* (John Wiley & Sons, Inc., New York, 1962) and J. P. Serre, "Lie Algebras and Lie Groups," lectures given at Harvard University, 1964.
[4] That is, if $\mathfrak{G}$ is isomorphic to a subalgebra of $\tilde{\mathfrak{G}}$.

(When no confusion would arise, we write $X\cdot v$ instead of $X_V\cdot v$.) So, in $\mathfrak{a}$ we have

$$X\cdot A = \theta(X)A = [X, A],$$

where the last term designates the bracket of A and X, considered as elements of $\tilde{\mathfrak{G}}$.

B. Some Complements about $\mathfrak{G}$-Modules

Let V_1, V_2, and V be any $\mathfrak{G}$-modules. We define a $\mathfrak{G}$-module structure on the following spaces by

$\underline{V_1 \otimes V_2}$: $X\cdot(v_1 \otimes v_2) = X\cdot v_1 \otimes v_2 + v_1 \otimes X\cdot v_2$

if $v_1 \in V_1$, $v_2 \in V_2$, and $X \in \mathfrak{G}$;

$\underline{\mathfrak{L}(V_1, V_2)}$: $(X\cdot f)(v_1) = X\cdot f(v_1) - f(X\cdot v_1)$

if $f \in \mathfrak{L}(V_1, V_2)$ and $v_1 \in V_1$. In particular, we can take for V_2 the trivial $\mathfrak{G}$-module (i.e., $V_2 = K$ and $X\cdot v_2 = 0$ $\forall v_2 \in V_2$, $X \in \mathfrak{G}$), thus recovering the definition of the contragradient (or dual) representation $X^* = -{}^tX$.

Remark: It is easy to see that these two definitions agree with the canonical identification of $V_1^* \otimes V_2$ with $\mathfrak{L}(V_1, V_2)$, which associates to each element $v_1^* \otimes v_2$ the operator

$$v_1 \longrightarrow \langle v_1, v_1^*\rangle v_2.$$

$\underline{B(V_1 \times V_2; V)}$ (i.e., the space of the bilinear mappings of $V_1 \times V_2$ into V).

$$(X\cdot\Phi)(v_1, v_2) = X\cdot\Phi(v_1, v_2) - \Phi(X\cdot v_1, v_2) - \Phi(v_1, X\cdot v_2),\ \Phi \in B(V_1 \times V_2; V).$$

In particular, if V is the trivial $\mathfrak{G}$-module, Φ is a bilinear form and

$$(X\cdot\Phi)(v_1, v_2) = -\Phi(X\cdot v_1, v_2) - \Phi(v_1, X\cdot v_2).$$

Invariant elements. Let V be a $\mathfrak{G}$-module. An element $v \in V$ is called $\mathfrak{G}$-invariant if $X\cdot v = 0$, $\forall X \in \mathfrak{G}$. [from the group viewpoint, by taking $g = \exp({}^tX)$, we have $gv = v$]. $V^{\mathfrak{h}}$ will be the set formed by the invariant elements of V. If $\mathfrak{G}$ is semisimple, $V = \oplus_i V_i$, where the V_i are simple modules (i.e., are irreducible components of the representation of $\mathfrak{G}$ in V), and $V^{\mathfrak{h}}$ is nothing other than the trivial part of the representation, that is the sum of the V_i which are isomorphic to the trivial $\mathfrak{G}$-module (i.e., the trivial representation). From the preceding definitions we have

$$\mathfrak{L}(V_1, V_2)^{\mathfrak{h}} = \{f;\ f \in \mathfrak{L}(V_1, V_2),\ X\cdot f(v_1) = f(X\cdot v_1),\ \forall v_1 \in V_1;\ \text{and}\ \forall X \in \mathfrak{G}\}.$$

It is the set of $\mathfrak{G}$ homomorphisms of V_1 into V_2. We have also

$$B(V_1, V_2; V)^{\mathfrak{h}} = \{\Phi;\ \Phi \in B(V_1, V_2; V);\ X\cdot\Phi(v_1, v_2) = \Phi(X\cdot v_1, v_2) + \Phi(v_1, X\cdot v_2)\}.$$

In particular, if V is the trivial $\mathfrak{G}$-module, Φ satisfies $\Phi(X\cdot v_1, v_2) + \Phi(v_1, X\cdot v_2) = 0$. [From the group viewpoint, this relation is equivalent to $\Phi(g\cdot v_1, gv_2) = \Phi(v_1, v_2)$.]

C. Structure of the Finite-Dimensional Representations of Semisimple Lie Algebras Extensions

Let ρ be a $\tilde{\mathfrak{G}}$ representation in a finite-dimensional vector space V, and let f be the restriction of ρ to $\mathfrak{a}$. From the identifications made in I.A, we may write

$$\rho(A + X) = f(A) + \rho(X); \qquad A \in \mathfrak{a}, \qquad X \in \mathfrak{G}.$$

V is considered as a $\mathfrak{G}$-module (by restriction of ρ), and we write $X\cdot v$ instead of $\rho(X)\cdot v$ ($X \in \mathfrak{G}$, $v \in V$). $\mathfrak{a}$ is also a $\mathfrak{G}$-module for the representation θ defined in Sec. I.A and, in the same way, we write $X\cdot A$, instead of $\theta(X)\cdot A$ ($X \in \mathfrak{G}$, $A \in \mathfrak{a}$). From the fact that ρ is a representation, we found the following two conditions for f:

$$f[A, B] = [f(A), f(B)], \qquad \forall A, B \in \mathfrak{a}, \tag{1a}$$

$$f(X\cdot A) = X\cdot f(A), \qquad f(A) \in \mathfrak{L}(V) \tag{2a}$$

(by defining on $\mathfrak{L}(V)$, the $\mathfrak{G}$-module structure given in Sec. I.B. *Therefore condition (2a) is equivalent to*

$$f \in \mathfrak{L}[\mathfrak{a}, \mathfrak{L}(V)]^{\mathfrak{h}}. \tag{2b}$$

We put the preceding condition in more explicit form. As a $\mathfrak{G}$-module, V may be written

$$V = \bigoplus_i (D_i \otimes N_i), \tag{3}$$

where the N_i are $\mathfrak{G}$-modules such that $N_i^{\mathfrak{h}} = N_i$ and the D_i are simple modules such that $D_i \chi D_j$ if $i \neq j$.

We have

$$\mathfrak{L}[\mathfrak{a}, \mathfrak{L}(V)] \simeq \mathfrak{a}^* \otimes V^* \otimes V \simeq \bigoplus_{i,j} (\mathfrak{a}^* \otimes D_j^* \otimes D_i \otimes N_j^* \otimes N_i)$$

and

$$\mathfrak{L}[\mathfrak{a}, \mathfrak{L}(V)]^{\mathfrak{h}} \cong \bigoplus_{i,j} \{(\mathfrak{a}^* \otimes D_j^* \otimes D_i)^{\mathfrak{h}} \otimes N_j^* \otimes N_i\}.$$

So, corresponding to the decomposition (3), $f(A)$ can be put in the matrix form $(f_{ij}(A))$, where

$$f_{ij}(A) = T_{ij}(A) \otimes M_{ij} \tag{4}$$

and

$$T_{ij} \in (\mathfrak{a}^* \otimes D_j^* \otimes D_i)^{\mathfrak{g}}, \tag{2.c}$$

$$M_{ij} \in N_j^* \otimes N_i \simeq \mathfrak{L}(N_j, N_i). \tag{5}$$

Computing f_{ij} is equivalent to extracting the trivial representation from the product of three representations. Now we also assume that $\mathfrak{a}$ *is a simple* $\mathfrak{G}$-*module* (for the representation θ).

The dimension of $(\mathfrak{a}^* \otimes D_j^* \otimes D_i)^{\mathfrak{g}}$ is equal to the number of times that $\mathfrak{a}$ appears in $D_j^* \otimes D_i$. If, in the product of two irreducible representations of $\mathfrak{G}$, each irreducible representation appears at most once (we say that $\mathfrak{G}$ is "simply reducible"), the dimension of $(\mathfrak{a}^* \otimes D_j^* \otimes D_i)^{\mathfrak{g}}$ is equal to one or zero, according as $\mathfrak{a}$ does or does not appear in the decomposition of $D_j^* \otimes D_i$.

We restrict ourselves to that case. So we found that the T_{ij} are known except for a multiplicative constant.

The condition (1a) must still be satisfied. For this purpose, the following remark is useful. Let V_1, V_2, V_3 be $\mathfrak{G}$-modules, and $\mathfrak{a}$ the $\mathfrak{G}$-module defined in Sec. I.A. Let us consider

$$f_1 \in \mathfrak{L}(\mathfrak{a}; \mathfrak{L}(V_1, V_2))^{\mathfrak{g}},$$

$$f_2 \in \mathfrak{L}(\mathfrak{a}; \mathfrak{L}(V_2, V_3))^{\mathfrak{g}},$$

$$f_3 \in \mathfrak{L}(\mathfrak{a}; \mathfrak{L}(V_1, V_3))^{\mathfrak{g}}.$$

We defined

$$\Phi(A, B) = f_2(A)f_1(B) - f_2(B)f_1(A) - f_3([A, B]).$$

Then

$$\Phi \in B(\mathfrak{a} \wedge \mathfrak{a}; \mathfrak{L}(V_1, V_3))^{\mathfrak{g}}.$$

If $B(\mathfrak{a} \wedge \mathfrak{a}; \mathfrak{L}(D_j, D_i))^{\mathfrak{g}}$ is not a zero-dimensional space, let (Φ_λ^{ij}) be a basis of this space and

$$\Phi_{ikj}(A, B) = T_{ik}(A)T_{kj}(B) - T_{ik}(B)\cdot T_{kj}(A) = \sum_\lambda \gamma_{ijk}^\lambda \Phi_\lambda^{i,j}(A, B),$$

$$T_{ij}([A, B]) = \sum_\lambda \delta_{ij}^\lambda \Phi_\lambda^{ij}(A, B).$$

The condition (1a) may be written

$$\sum_k f_{ik}(A)f_{kj}(B) - f_{ik}(B)\cdot f_{kj}(A) = f_{ij}([A, B]),$$

and, from (4)

$$\sum_{\lambda,k} \gamma_{ikj}^\lambda \Phi_\lambda^{ij}(A, B) \otimes M_{ik}M_{kj} = \sum_\lambda \delta_{ij}^\lambda \Phi_\lambda^{ij}(A, B) \otimes M_{ij}.$$

So we have the following conditions:

$$\delta^\lambda M_{ij} = \sum_k \gamma_{ikj}^\lambda M_{ik}M_{kj}. \tag{1.b}$$

for each pair (i, j) for which $\mathfrak{L}(\mathfrak{a} \wedge \mathfrak{a}; \mathfrak{L}(D_j, D_i))^{\mathfrak{g}}$ is not a zero-dimensional space.

D. Application to the Finite-Dimensional Representations of any Lie Algebra

The following well-known result is used: An arbitrary Lie algebra $\tilde{\mathfrak{G}}$ is the semidirect product of its radical R and a semisimple algebra $\mathfrak{G}$. (R is the greatest solvable ideal of $\mathfrak{G}$.) The preceding results can be applied. There are now some other properties, due to the fact that R is solvable.

Let (ρ, V) be an *irreducible* representation of $\tilde{\mathfrak{G}}$.

(a) ρ *is zero on* $N = [R, \tilde{\mathfrak{G}}]$. In fact $N = [R, \tilde{\mathfrak{G}}]$ is the nilpotent radical of $\tilde{\mathfrak{G}}$; but it is well known that N is also the set of $X \in \tilde{\mathfrak{G}}$ such that $\rho(X) = 0$ for each irreducible representation.

(b) *For each* $A \in R$, $\rho(A) = \lambda \cdot 1_V$. It results from Schur's lemma (if the field is algebraically closed: i.e. $K = \mathbf{C}$). In particular, if $R = N$, $\rho(R)$ is zero for each irreducible representation. It is the case when $R^{\mathfrak{g}} = 0$, because in that case $[R, \mathfrak{G}] = R$ and *a fortiori* $[R, \tilde{\mathfrak{G}}] = R$.

Consequence: Each irreducible representation of $\tilde{\mathfrak{G}}$ restricted to $\mathfrak{G}$ is still an irreducible representation of $\mathfrak{G}$.

E. Application to the Semidirect Product of an Abelian Lie Algebra by a Semisimple One

We suppose that $\tilde{\mathfrak{G}} = \mathfrak{a} \times \mathfrak{G}$, where A is Abelian and is a simple $\mathfrak{G}$-module. We have $R = \mathfrak{a}$ and $N = [\mathfrak{a}, \tilde{\mathfrak{G}}] = [\mathfrak{a}, \mathfrak{G}] = \mathfrak{a}$. Let us set

$$\begin{cases} V_p = \{v \in V;\ f(A)v \in V_{p-1}, \quad \forall A \in \mathfrak{a}\}, \\ V_{-1} = \{0\}. \end{cases}$$

Then

$$\{0\} \subset V_0 \subset V_1 \subset \cdots, \subset V_n = V,$$

and, from the preceding sections

$$V_p \neq V_{p-1}, \quad \text{if} \quad p = 0, \cdots n.$$

It is obvious that the V_p are sub-$\mathfrak{G}$-module of V, so we have

$$V_p = \sum_i (D_i \otimes N_{i,p}),$$

where $N_{i,p} \subset N_i$, and it is easy to see that

$$\begin{cases} N_{i,p} = \{n; n \in N_i, M_{ij}n \\ \qquad\qquad \in N_{i,p-1} \forall i \quad \text{such that} \quad T_{ij} \neq 0\}, \\ N_{i,-1} = \{0\}. \end{cases}$$

It is then possible to find subspaces $\tilde{N}_{i,p}$, $p=0, 1 \cdots n$ of N_i such that

$$N_{i,p} = \tilde{N}_{i,0} \oplus \tilde{N}_{i,1} \oplus \cdots \oplus \tilde{N}_{i,p}.$$

(But it is to be noticed that this choice is by no mean canonical.) We can put the matrices M_{ij} in the form

$$M_{ij} = (M_{ij,pq})_{p,q},$$

where $M_{ij,pq}$ are $n_{i,p} \times n_{j,q}$ matrices and $n_{i,p} = \dim N_{i,p}$. It is clear that

$$M_{ij,pq} = 0 \quad \text{if} \quad p \geq q.$$

We can now resume the situation in the following way. To get an explicit form for a finite-dimensional representation of $\tilde{\mathcal{G}} = \mathfrak{a} \times \mathcal{G}$, we have to take a family $(n_{i,p})_{0 \leq p \leq n}\ i \in I$ of integers, where I is the set of classes of equivalent finite-dimensional representation of $\mathcal{G}$, and a family $(M_{ij,pq})_{0 \leq p,q \leq n}\ i, j \in I$ of $n_{i,p} \times n_{j,q}$ matrices, where

$$M_{ij,pq} = 0 \quad \text{if} \quad \dim (\mathfrak{a}^* \otimes D_i^* \otimes D_i)^{\mathcal{G}} = 0$$

$$\text{or if} \quad p \geq q$$

and satisfying the following conditions[5]:

$$\sum_k \gamma_{ikj}^{\lambda} M_{ik} M_{kj} = 0, \tag{6}$$

with

$$M_{ij} = (M_{ij,pq})_{p,q}.$$

Defining $\tilde{V}_p = \bigoplus_i (D_i \otimes \tilde{N}_{i,p})$; to the decomposition $V = \bigoplus_p \tilde{V}_p$ corresponds a matrix form by "blocks" for ρ:

$$\rho(X + A) = \begin{bmatrix} \rho_1(X) & \tilde{f}_{12}(A) & \cdots & \tilde{f}_{1n}(A) \\ 0 & \rho_2(X) & \cdots & \tilde{f}_{2n}(A) \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \rho_n(X) \end{bmatrix},$$

where

$$\rho_p(X) = \bigoplus_i (D_i(X) \otimes \mathbb{1}_{n_{i,p}}),$$

$$\tilde{f}_{p,q}(A) = (f_{ij,pq}(A))_{i,j} \in I,$$

$\tilde{f}_{ij,pq}(A)$ being a matrix which maps $D_i \otimes \tilde{N}_{i,q}$ into $D_i \otimes \tilde{N}_{i,p}$ and which is defined by

$$\tilde{f}_{ij,pq}(A) = T_{ij}(A) \otimes M_{ij,pq}.$$

Remark 1: We see from the preceding results (4)–(5) that, if the representation ρ restricted to $\mathfrak{a}$ is nontrivial, then ρ is reducible, but not completely reducible.[4]

Remark 2. Equivalence of representations.
Let us consider two representations $\rho = \sum_i D_i \otimes N_i$ and $\rho' = \sum_i D_i \otimes N_i$ (where we identify space and representation). The intertwinning operators have the following form:

$$A = \sum_i \mathbb{1}_{D_i} \otimes A_i, \qquad A_i \in \mathfrak{L}(N_i),$$

where the A_i are such that

$$A_i M_{ij} = M'_{ij} A_j.$$

Then, the necessary and sufficient condition for the equivalence of two representations is the existence of invertible matrices A_i such that

$$M'_{ij} = A_i M_{ij} A_i^{-1}.$$

II. APPLICATION TO THE FINITE-DIMENSIONAL REPRESENTATION OF THE MOTION GROUP OF THREE-DIMENSIONAL SPACE

This group is the semidirect product $T \times R$ of the translation group and the rotation group in a three-dimensional Euclidian space. We want to find the continuous finite-dimensional representations, so it is sufficient to determine the representations of the corresponding Lie algebra. This one is also a semidirect product $\tilde{\mathcal{G}} = \mathfrak{a} \times \mathcal{G}$, where $\mathcal{G} \simeq SU(2)$ and $\mathfrak{a}$ is an Abelian three-dimensional Lie algebra, which as a $\mathcal{G}$-module, is equivalent to D_1. (As usual, the irreducible representations of $\mathcal{G}$ are designed by D_l, where $2l$ is an integer.)

A. T_{ij}'s Derivation

Let X_+, X_-, and H be the canonical elements of $\mathcal{G}$ such that

$$[X_+, X_-] = 2H,$$

$$[H, X_+] = X_+,$$

$$[H, X_-] = -X_-.$$

The T_{ij} are more easily computed in the basis (e_h^l) of D_l defined by

$$\begin{cases} X_+ e_h^l = (l - h)\, e_{h+1}^l, \\ X_- e_h^l = (l + h)\, e_{h-1}^l, \\ H e_h^l = h e_h^l. \end{cases}$$

From

$$(\mathfrak{a}^* \otimes D_j^* \otimes D_i)^{\mathcal{G}} = (D_1 \otimes D_j \otimes D_i)^{\mathcal{G}}$$

it is clear that this space is not zero-dimensional only if i and j are not simultaneously equal to 0.

[5] Remark: the matrices $M_{ji,\ pq}$ are also submitted to the condition $\bigcap_j \ker M_{ji,\ p-1,\ p} = \{0\}$.

So we have only to compute

$$T_{i,i} \quad (i \neq 0), \qquad T_{i+1,i} \quad \text{and} \quad T_{i-1,i}$$

(*remark*: i is either always an integer or always a half-integer).

We search T_{ij} under the form

$$\sum_{\lambda,\mu} C^{ij}_{\lambda\mu}\, e^1_{-\lambda} \otimes e^j_{\mu} \otimes e^i_{-\mu+\lambda},$$

where the $C^{ij}_{\lambda\mu}$ are Clebsch–Gordan coefficients. Then the corresponding operators are given by

$$[T_{i,j}(e^1_\lambda)]\cdot e^j_\mu = (e^i_{-\mu}, e^j_\mu)(e^1_{-\lambda}, e^1_\lambda) C^{ij}_{-\lambda,-\mu}\, e^i_{\lambda+\mu},$$

where $(\cdot\ ,\ \cdot)$ is the scalar product defined by

$$(e^i_\mu, e^i_{\mu'}) = \delta_{\mu,-\mu'}(-)^\mu (i-\mu)!\,(i+\mu)!$$

After computation and a suitable choice of the arbitrary constant in each case, we find

$$[T_{i,i}(e^1_\lambda)]\, e^i_\mu = (\mu - i\lambda)\, e^i_{\mu+\lambda},$$

$$[T_{i+1,i}(e^1_\lambda)]\, e^i_\mu = e^{i+1}_{\mu+\lambda},$$

$$\begin{cases} [T_{i-1,i}(e^1_1)]\, e^i_\mu = -(i-\mu)(i-\mu-1)\, e^{i-1}_{\mu+1}, \\ [T_{i-1,i}(e^1_0)]\, e^i_\mu = +(i+\mu)(i-\mu)\, e^{i-1}_{\mu}, \\ [T_{i-1,i}(e^1_{-1})]\, e^i_\mu = -(i+\mu)(i+\mu-1)\, e^{i-1}_{\mu-1}. \end{cases}$$

This result can be expressed in the ordinary (and orthogonal) basis ξ^l_m by the following transformation.[6]

$$e^l_m = \gamma^l_m \xi^l_m$$

with

$$\gamma^l_l = 1 \quad \text{and} \quad \gamma^l_m = [(l+m)!\,(l-m)!/(2l)!]^{\frac{1}{2}}.$$

We have (also with a choice of the arbitrary constant in each case):

$$\begin{cases} [T_{i,i}(e^1_1)]\xi^i_m = [(i-m)(i+m+1)]^{\frac{1}{2}}\xi^i_{m+1}, \\ [T_{i,i}(e^1_0)]\xi^i_m = m\xi^i_m, \\ [T_{i,i}(e^1_{-1})]\xi^i_m = [(i+m)(i+1-m)]^{\frac{1}{2}}\xi^i_{m-1}, \end{cases}$$

$$\begin{cases} [T_{i+1,i}(e^1_1)]\xi^i_m = [(i+m+2)(i+m+1)]^{\frac{1}{2}}\xi^{i+1}_{m+1}, \\ [T_{i+1,i}(e^1_0)]\xi^i_m = [(i+1)^2 - m^2]^{\frac{1}{2}}\xi^{i+1}_m, \\ [T_{i+1,i}(e^1_{-1})]\xi^i_m = [(i+2-m)(i+1-m)]^{\frac{1}{2}}\xi^{i+1}_{m-1}, \end{cases}$$

$$\begin{cases} [T_{i-1,i}(e^1_1)]\xi^i_m = -[(i-m)(i-1-m)]^{\frac{1}{2}}\xi^{i-1}_{m+1}, \\ [T_{i-1,i}(e^1_0)]\xi^i_m = (i^2 - m^2)^{\frac{1}{2}}\xi^{i-1}_m, \\ [T_{i-1,i}(e^1_{-1})]\xi^i_m = -[(i+m)(i-1+m)]^{\frac{1}{2}}\xi^{i-1}_{m-1}. \end{cases}$$

[6] This result was already given under a more general form (concerning every convex Lie group) by D. Zelobenko, Tr. Mosk. Mat. Obshchestva 12, 62 (1963).

B. The Coefficients γ_{ikj}

We have (cf. Sec. I.C):

$$(\mathfrak{a} \wedge \mathfrak{a}) \otimes D^*_j \otimes D_i = D_1 \otimes D_j \otimes D_i$$

because it is well known that $D_1 \wedge D_1 = D_1$. This space is then at most one-dimensional, and we have only to compute

$$\gamma_{i,i+1,i}, \qquad \gamma_{i,i-1,i}, \qquad \gamma_{i,i,i}, \qquad \gamma_{i+1,i,i},$$
$$\gamma_{i+1,i+1,i}, \qquad \gamma_{i-1,i,i}, \qquad \gamma_{i-1,i-1,i}.$$

By a suitable choice of arbitrary constants we find

$$\begin{cases} \gamma_{i,i-1,i} = 2i-1, \\ \gamma_{i,i,i} = 1, \\ \gamma_{i,i+1,i} = -(2i+3), \end{cases}$$

$$\begin{cases} \gamma_{i+1,i,i} = i, \\ \gamma_{i+1,i+1,i} = -(i+2), \end{cases}$$

$$\begin{cases} \gamma_{i-1,i,i} = i+1, \\ \gamma_{i-1,i-1,i} = -(i-1). \end{cases}$$

C. Diagram Associated with a Representation

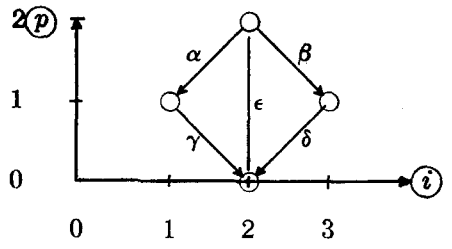
1. Description

Let us recall first (see Sec. I.E) that the representations space V has a decomposition $V = \bigoplus \tilde{V}_p$, where $\tilde{V}_p = \bigoplus_i (D_i \otimes \tilde{N}_{i,p})$; for which the representation matrix has a triangular form. Then to each representation we associate the diagram formed by the points of coordinates (i, p). (i is an integer or an half-integer, but we can always restrict ourselves to the case where all i are of the same kind.) Now each point (i, p) is labeled by a number $\tilde{n}_{i,p} = \dim \tilde{N}_{i,p}$ = multiplicity of D_i in $\tilde{V}_p$. [We consider only the points (i, p) for which $\tilde{n}_{i,p}$ is different from zero.] Such a diagram does not determine completely the representation. It is also necessary to know the matrix $M_{ij,pq}$, which performs the transition from the point (j, q) to the point (i, p); i.e., there is a transition between points such that $q > p$ and $|i-1| \leq j \leq i+1$ ($j, i \neq 0$ simultaneously). We represent one matrix on the diagram by an arrow joining the points.

We do not consider nonconnected diagrams which certainly correspond to representations which are reducible in a direct sum of representations.

2. Example

One case is particularly simple: When all the $\tilde{n}_{i,p}$ are equal to one or zero. Then the matrix $M_{ij,pq}$ reduces to a scalar $\alpha_{ij,pq}$. Let us consider, for example, the following diagram:



where we have dropped the coefficient $\tilde{n}_{i,p}$ always equal to one. It means

$$\tilde{V}_0 = D_2, \qquad \tilde{V}_1 = D_1 \oplus D_3, \quad \text{and} \quad \tilde{V}_2 = D_2.$$

To this diagram corresponds the representation

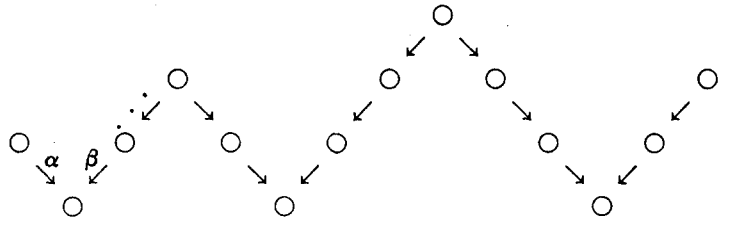
$$\rho(X+A) = \begin{array}{|c|c|c|c|} \hline D_2(X) & \gamma T_{21}(A) & \delta T_{23}(X) & \epsilon T_{22}(A) \\ \hline 0 & D_1(X) & 0 & \alpha T_{21}(A) \\ \hline 0 & 0 & D_3(X) & \beta \cdot T_{32}(A) \\ \hline 0 & 0 & 0 & D_2(X) \\ \hline \end{array}$$

The coefficients $\alpha, \beta \cdots$ are not independent. They must satisfy the conditions (6). In this case (6) reduces to

$$\alpha\gamma/\beta\delta = \tfrac{1}{3}.$$

3. *Particular Representations*

Let us consider the particular representations which do not contain two identical D_i. Then it is easy to see that the associated diagrams must be of the following form:



and in this case, the form of the diagram completely determines the representation. More precisely, to each diagram of this sort corresponds a class of equivalent representations.

In fact, by changing the basis in each D_i, it is always possible to reduce any coefficient β to 1. So the representations corresponding to different coefficients are equivalent.

D. Tensor Product of Representations

In Sec. I.B, we have already given the $\mathfrak{G}$-module structure for the tensor product of $\mathfrak{G}$-modules V and V'. It corresponds to

$$[(\rho \otimes \rho')(\tilde{X})](e \otimes e') = \rho(\tilde{X})e \otimes e' + e \otimes \rho'(\tilde{X})e',$$

where

$$e \in V, \quad e' \in V', \quad \tilde{X} \in \tilde{\mathfrak{G}}, \quad \text{i.e,} \quad \tilde{X} = X + A$$

with $X \in \mathfrak{G}$, $A \in \mathfrak{a}$.

We first have to study the structure of the representation space $V \otimes V'$, or equivalently, the structure of the diagram associated with the tensor product of the two representations.

With the notations defined in Sec. II.C, it is possible to show that

$$(V \otimes V')_r = \sum_{p+q \leq r} \tilde{V}_p \otimes \tilde{V}'_q,$$

which means that it is possible to take

$$(V \otimes V')_r^- = \sum_{p+q=r} \tilde{V}_p \otimes \tilde{V}'_q;$$

so we obtain the rth stage of the diagram associated with the representation $\rho \otimes \rho'$, in terms of the pth and qth stages of the diagrams associated respectively with ρ or ρ'.

Then we have to find the explicit form of the matrices. After some Clebsch–Gordan calculations, we find

$$[T_{i'i} \otimes 1](A) = \sum_{\lambda,\lambda'} \gamma_{i'i}^{\lambda'\lambda} T_{\lambda'\lambda}(A), \qquad A \in \mathfrak{a},$$

where $T_{i'i} \otimes 1$ is a map: $D_i \otimes D_j \to D_{i'} \otimes D_j$, the coefficient $\gamma_{i'i}^{\lambda'\lambda}$ being proportional to the (6–j) coefficient

$$\begin{Bmatrix} 1 & \lambda' & \lambda \\ j & i & i' \end{Bmatrix}.$$

III. FINITE-DIMENSIONAL REPRESENTATION OF $\tilde{\mathcal{G}} = M \times SU(2)$ WITH $M \simeq D^{1/2}$

We give here the results concerning this Lie algebra, because, as we see in the next section, it is then very easy to obtain the structure of the finite representations of the inhomogeneous Lorentz group.

Now we have simpler results in this case than in the preceding one. This comes from the following fact:

$$T_{ij} \in (D^{\frac{1}{2}} \otimes D^{j} \otimes D^{i})^{\frac{1}{2}},$$

i.e., there is only a transition matrix T_{ij} for

$$|i - j| = \tfrac{1}{2}.$$

We get

$$[T_{i+\frac{1}{2},i}(e^{\frac{1}{2}}_{\epsilon})]\, e^{i}_{\lambda} = e^{i+\frac{1}{2}}_{\lambda+\epsilon},$$

$$[T_{i-\frac{1}{2},i}(e^{\frac{1}{2}}_{\epsilon})]\, e^{i}_{\lambda} = (2\epsilon i - \lambda)\, e^{i-\frac{1}{2}}_{\lambda+\epsilon}.$$

We come now to the condition (1a). We have to consider the space

$$[(D^{\frac{1}{2}} \wedge D^{\frac{1}{2}}) \otimes D^{i} \otimes D^{j}]^{\frac{1}{2}} = (D^{0} \otimes D^{i} \otimes D^{j})^{\frac{1}{2}}.$$

This space is zero or one dimensional according as $i \neq j$ or $i = j$. An easy calculation shows that

$$\left.\begin{aligned} \Phi_{i,i+\frac{1}{2},i} &= (i+1)\Phi_i \\ \Phi_{i,i-\frac{1}{2},i} &= -i\Phi_i \end{aligned}\right\} \text{ or } \Phi_{i,i+\epsilon,i} = 2\epsilon(i + \tfrac{1}{2} + \epsilon)\Phi_i,$$

$$\Phi_{i,j,k} = 0, \qquad \text{in the other cases,}$$

where the $\Phi_{i,k,j}$ have been defined (Sec. I.C), and Φ_i is a nonzero element of $(D^{0} \otimes D^{i} \otimes D^{j})^{\frac{1}{2}}$. In other terms we have

$$\gamma_{i,i+\frac{1}{2},i} = i + 1,$$

$$\gamma_{i,i-\frac{1}{2},i} = -i,$$

the other $\gamma_{i,k,j}$ being zero.

Finally, the preceding remarks concerning the diagram associated with a representation, and the tensor product of two representations are always valid in this case.

IV. FINITE-DIMENSIONAL REPRESENTATIONS OF THE INHOMOGENEOUS LORENTZ GROUP

In the same way as the preceding section, we consider only the Lie algebra.

A. Notations and Basic Facts

1. The homogeneous Lorentz group

A basis of the real Lie algebra $\mathcal{G}_0$ of the homogeneous Lorentz group is constituted by the elements

$$\text{(a)} \quad \begin{aligned} A_1 &= \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right), & B_1 &= \left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right), & C_1 &= \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right), \\ A_2 &= \left(\begin{smallmatrix} 0 & i \\ 0 & 0 \end{smallmatrix}\right), & B_2 &= \left(\begin{smallmatrix} 0 & 0 \\ i & 0 \end{smallmatrix}\right), & C_2 &= \left(\begin{smallmatrix} i & 0 \\ 0 & -i \end{smallmatrix}\right). \end{aligned}$$

Considering the corresponding complex Lie algebra $\mathcal{G}$, we can take the basis

$$\text{(b)} \quad \begin{aligned} X_1 &= A_1 + iA_2, & X_2 &= A_1 - iA_2, \\ Y_1 &= B_1 + iB_2, & Y_2 &= B_1 - iB_2, \\ H_1 &= \tfrac{1}{2}(C_1 + iC_2), & H_2 &= \tfrac{1}{2}(C_1 - iC_2). \end{aligned}$$

It should be kept in mind that the i which appear in (a) and in (b) have different meaning, thus obtaining the following multiplication table:

$$[H_1, X_1] = X_1, \qquad [H_2, X_2] = X_2,$$

$$[H_1, Y_1] = -Y_1, \qquad [H_2, Y_2] = -Y_2,$$

$$[X_1, Y_1] = 2H_1, \qquad [X_2, Y_2] = 2H_2,$$

and the other commutators are zero.

The finite-dimensional representations of $\mathcal{G}$ are labeled by numbers n and m, which can be either integer or half-integer $D_{n,m}$. First we choose the noncanonical basis which is formed in an obvious way from the tensor product of basis e^{l}_{h} defined in Sec. II. Denoting the elements of this basis by $e^{n,m}_{h,k}$, we have

$$X_1\, e^{n,m}_{h,k} = (n - h)\, e^{m,n}_{h+1,k}, \qquad X_2\, e^{n,m}_{h,k} = (m - k)\, e^{n,m}_{h,k+1},$$

$$Y_1\, e^{n,m}_{h,k} = (n + h)\, e^{n,m}_{h-1,k}, \qquad Y_2\, e^{n,m}_{h,k} = (m + k)\, e^{n,m}_{h,k-1},$$

$$H_1\, e^{n,m}_{h,k} = h \cdot e^{n,m}_{h,k}, \qquad H_2\, e^{n,m}_{h,k} = k\, e^{n,m}_{h,k}.$$

2. The inhomogeneous case

The Lie algebra of the inhomogeneous Lorentz group is a semidirect product $\tilde{\mathcal{G}} = M \times \mathcal{G}$, where M is an Abelian four-dimensional Lie algebra, isomorphic to $D^{\frac{1}{2},\frac{1}{2}}$. An element of M is denoted by $e^{\frac{1}{2},\frac{1}{2}}_{i,\eta}$.

B. The Finite-Dimensional Representations

1. $T_{ii'}$, Derivation

Applying the results of Sec. I, we have to compute the matrices T_{ij}, where i and j are now a set of two integers or half-integers: $i = (n, m)$; $j = (n', m')$.

$$T_{ij} \in (D^{\frac{1}{2},\frac{1}{2}} \otimes D^{n,m} \otimes D^{n',m'})^{\frac{1}{2}}.$$

So there are only transitions $T_{n'm';nm}$ for $|n' - n| = \frac{1}{2}$ and $|m' - m| = \frac{1}{2}$. But from the following identification:

$$(D^{\frac{1}{2},\frac{1}{2}} \otimes D^{n,m} \otimes D^{n',m'})^{\frac{1}{2}} = (D^{\frac{1}{2}} \otimes D^{n} \otimes D^{n'})^{\frac{1}{2}} \otimes (D^{\frac{1}{2}} \otimes D^{m} \otimes D^{m'})^{\frac{1}{2}},$$

we get

$$T_{nm,n'm'}(x \otimes y) = T_{n,n'}(x) \otimes T_{m,m'}(y),$$

where

$$x, y \in D^{\frac{1}{2}}; \qquad x \otimes y \in D^{\frac{1}{2},\frac{1}{2}}.$$

It means that with our choice of basis, we have only to make a tensor product of the results given in Sec. III, thus obtaining

$$[T_{n+\frac{1}{2},m+\frac{1}{2};n,m}(\epsilon,\eta)]\, e^{n,m}_{h,k} = e^{n+\frac{1}{2},m+\frac{1}{2}}_{h+\epsilon,k+\eta},$$

$$[T_{n-\frac{1}{2},m+\frac{1}{2};n,m}(\epsilon,\eta)]\, e^{n,m}_{h,k} = (2\epsilon n - h)\, e^{n-\frac{1}{2},m+\frac{1}{2}}_{h+\epsilon,k+\eta},$$

$$[T_{n+\frac{1}{2},m-\frac{1}{2};n,m}(\epsilon,\eta)]\, e^{n,m}_{h,k} = (2\eta m - k)\, e^{n+\frac{1}{2},m-\frac{1}{2}}_{h+\epsilon,k+\eta},$$

$$[T_{n-\frac{1}{2},m-\frac{1}{2};n,m}(\epsilon,\eta)]\, e^{n,m}_{h,k} = (2\epsilon n - h)(2\eta m - k)\, e^{n-\frac{1}{2},m-\frac{1}{2}}_{h+\epsilon,k+\eta},$$

where $T_{i,i}(\epsilon,\eta)$ means $T_{i,i}(e^{\frac{1}{2},\frac{1}{2}}_{\epsilon,\eta})$.

Now, from the physicist's viewpoint, it may be useful to know the results in the canonical and more physical basis ξ^{τ}_{lm}, as described in Ref. 5, p. 188. The results are given in Appendix I.

2. Relations between the matrices M_{ij}

Let us recall that

$$\Phi_{n,\nu,n'}(x, x') = T_{n\nu}(x)\cdot T_{\nu n'}(x') - T_{n\nu}(x')\cdot T_{\nu n'}(x) \qquad x, x' \in D^{\frac{1}{2}}.$$

Now we define also a symmetric bilinear form

$$\Psi_{n,\nu,n'}(x, x') = \tfrac{1}{2}\,[T_{n\nu}(x)T_{\nu n'}(x') + T_{n\nu}(x')\cdot T_{\nu n'}(x)].$$

It is easy to prove exactly in the same way that

$$\Psi_{n,\nu,n'} \in [(D^{\frac{1}{2}} \underset{s}{\otimes} D^{\frac{1}{2}}) \otimes D^{n} \otimes D^{n'}]^{\frac{1}{2}},$$

with

$$D^{\frac{1}{2}} \underset{s}{\otimes} D^{\frac{1}{2}} = D^{1}.$$

We have to consider the bilinear form $\Phi_{nm;\nu\mu;\,n'm'}$ but it is clear that

$$\Phi_{nm;\nu\mu;n'm'}(x \otimes y, x' \otimes y') = \Phi_{n,\nu,n'}(x, x') \otimes \Psi_{m,\mu,m'}(y, y') + \Psi_{n,\nu,n'}(x, x') \otimes \Phi_{m,\mu,m'}(y, y').$$

The only nonzero $\Phi_{nm;\nu\mu;n'm'}$ are those corresponding to

$$|n - \nu| = |m - \mu| = |n' - \nu| = |m' - \mu| = \tfrac{1}{2},$$

$$|n - n'| = 0, 1; \qquad |m - m'| = 0, 1.$$

Now, the only nonzero Ψ are

$$\Psi_{m,m+\frac{1}{2},m} = \Psi_{m,m-\frac{1}{2},m} = \Psi^{0}_{m},$$

$$\Psi_{m+1,m+\frac{1}{2},m} = \Psi^{1}_{m},$$

$$\Psi_{m-1,m-\frac{1}{2},m} = \Psi^{-1}_{m},$$

where the first equality can be easily verified. We have to distinguish three cases:

I. $\begin{cases} n' = n + 2\epsilon \\ m' = m \end{cases} \quad \begin{cases} \nu = n + \epsilon \\ \mu = m + \eta \end{cases}$ with $\epsilon, \eta = \pm\frac{1}{2}$,

$$\Phi_{n',m';\nu,\mu;n,m} = 2\eta(m + \eta + \tfrac{1}{2})\Psi^{2\epsilon}_{\eta} \otimes \Phi_{m}.$$

II. $\begin{cases} n' = n \\ m' = m + 2\eta \end{cases} \quad \begin{cases} \nu = n + \epsilon \\ \mu = m + \eta \end{cases}$ with $\epsilon, \eta = \pm\frac{1}{2}$,

$$\Phi_{n',m';\nu,\mu;n,m} = 2\epsilon(n + \epsilon + \tfrac{1}{2})\Phi_{n} \otimes \Psi^{2\eta}_{m}.$$

III. $\begin{cases} n' = n \\ m' \quad m \end{cases} \quad \begin{cases} \nu = n + \epsilon \\ \mu = m + \eta \end{cases}$

$$\Phi_{n,m;\nu\mu;n,m} = 2\eta(m + \eta + \tfrac{1}{2})\Psi^{0}_{n} \otimes \Phi_{m} + 2\epsilon(n + \epsilon + \tfrac{1}{2})\Phi_{n} \otimes \Psi^{0}_{m}.$$

From there, we get the following relations between the matrices M_{ij} [the relations (6)]:

I. $$\sum_{\eta} 2\eta(m + \eta + \tfrac{1}{2}) \times M_{n+2\epsilon,m;n+\epsilon,m+\eta}\cdot M_{n+\epsilon,m+\eta;n,m} = 0,$$

II. $$\sum_{\epsilon} 2\epsilon(n + \epsilon + \tfrac{1}{2}) \times M_{n,m+2\eta;n+\epsilon,m+\eta}\cdot M_{n+\epsilon,m+\eta;n,m} = 0,$$

III. gives two relations:

$$\sum_{\epsilon,\eta} 2\epsilon(n + \epsilon + \tfrac{1}{2})M_{n,m;n+\epsilon,m+\eta}\cdot M_{n+\epsilon,m+\eta;n,m} = 0$$

and

$$\sum_{\epsilon,\eta} 2\eta(m + \eta + \tfrac{1}{2})M_{n,m;n+\epsilon,m+\eta}\cdot M_{n+\epsilon,m+\eta;n,m} = 0.$$

ACKNOWLEDGMENTS

One of us (M.L.N.) is grateful to J. Lascoux and P. Renouard for several interesting comments and to L. Boutay de Monvel for helpful discussions about Lie algebras.

APPENDIX I. THE MATRICES T_{ij} IN THE "CANONICAL" BASIS $\xi_{lm}{}^{\tau}$

We adopt here the notations of Ref. 7, i.e.,

$$\tau = (l_0, l_1) \quad \text{with} \quad \begin{cases} l_0 = n - m, \\ l_1 = n + m + 1. \end{cases}$$

In this case, we found that there is only transition matrices $T^{\tau'\tau}$ between $\tau = (l_0, l_1)$ and $\tau' = (l_0, l_1 \pm 1)$ or $\tau' = (l_0 \pm 1, l_1)$. After some straightforward calculations, we get, using the notation $\xi^{(l_0,l_1)}_{l,m}$:

[7] J. M. Gel'fand-R. A. Minlos and Z. Ya Shapiro, *Representations of the Rotation and Lorentz Groups and their Applications* (Pergamon Press Inc., New York, 1963), p. 145.

$$[T^{\tau'\tau}(\xi^{(0,2)}_{1,1})]\xi^{\tau}_{l,m} = \alpha^{\tau'\tau}_{l}[(l+m+2)(l+m+1)]^{\frac{1}{2}}\xi^{\tau'}_{l+1,m+1} + \beta^{\tau'\tau}_{l}[(l-m)(l+m+2)]^{\frac{1}{2}}\xi^{\tau'}_{l,m+1} + \gamma^{\tau'\tau}_{l}[(l-m)(l-m-1)]^{\frac{1}{2}}\xi^{\tau'}_{l-1,m+1},$$

$$[T^{\tau'\tau}(\xi^{(0,2)}_{1,0})]\xi^{\tau}_{l,m} = \alpha^{\tau'\tau}_{l}[(l+1)^2 - m^2]^{\frac{1}{2}}\xi^{\tau'}_{l+1,m} + \beta^{\tau'\tau}_{l}\cdot m\cdot\xi^{\tau'}_{l,m} - \gamma^{\tau'\tau}_{l}(l^2-m^2)^{\frac{1}{2}}\xi^{\tau'}_{l-1,m},$$

$$[T^{\tau'\tau}(\xi^{(0,2)}_{1,-1})]\xi^{\tau}_{l,m} = \alpha^{\tau'\tau}_{l}[(l+2-m)(l+1-m)]^{\frac{1}{2}}\xi^{\tau'}_{l+1,m-1} + \beta^{\tau'\tau}_{l}[(l+m)(l+1-m)]^{\frac{1}{2}}\xi^{\tau'}_{l,m-1} + \gamma^{\tau'\tau}_{l}[(l+m)(l-1+m)]^{\frac{1}{2}}\xi^{\tau'}_{l-1,m-1},$$

$$[T^{\tau'\tau}(\xi^{(0,2)}_{0,0})]\xi^{\tau}_{l,m} = c^{\tau'\tau}_{l}\xi^{\tau'}_{l,m},$$

where the coefficients α, β, γ are given by the following formulas:

$$\alpha^{\tau'\tau}_{l} = c^{\tau'\tau}_{l}C^{\tau'}_{l+1} - c^{\tau'\tau}_{l+1}C^{\tau}_{l+1},$$

$$\beta^{\tau'\tau}_{l} = c^{\tau'\tau}_{l}[A^{\tau'}_{l} - A^{\tau}_{l}],$$

$$\gamma^{\tau'\tau}_{l} = c^{\tau'\tau}_{l-1}C^{\tau}_{l} - c^{\tau'\tau}_{l}C^{\tau'}_{l}.$$

where (cf. Ref. 7, p. 189)

$$C^{\tau}_{l} = \frac{i}{l}\left[\frac{(l^2 - l_0^2)(l^2 - l_1^2)}{4l^2 - 1}\right]^{\frac{1}{2}}$$

and

$$A^{\tau}_{l} = \frac{il_0l_1}{l(l+1)}.$$

First case: $\tau' = (l_0, l_1+1)$.

$$c^{\tau'\tau}_{l} = c^{\tau'\tau}_{1}[(l-l_1)(l+l_1+1)]^{\frac{1}{2}},$$

$$\alpha^{\tau'\tau}_{l} = \frac{1}{l+1}\left\{\frac{[(l+1)^2-l_0^2][l+l_1+2][l+l_1+1]}{4(l+1)^2-1}\right\}^{\frac{1}{2}}c^{\tau'\tau}_{1},$$

$$\beta^{\tau'\tau}_{l} = \frac{l_0}{l(l+1)}[(l-l_1)(l+l_1+1)]^{\frac{1}{2}}c^{\tau'\tau}_{1},$$

$$\gamma^{\tau'\tau}_{l} = -\frac{1}{l}\left\{\frac{(l-l_1)(l-l_1-1)(l^2-l_0^2)}{4l^2-1}\right\}^{\frac{1}{2}}c^{\tau'\tau}_{1},$$

where $c^{\tau'\tau}_{1}$ is a constant independent of l.

Second case: $\tau' = (l_0+1, l_1)$.

$$c^{\tau'\tau}_{l} = c^{\tau'\tau}_{2}[(l-l_0)(l+l_{0+1})]^{\frac{1}{2}},$$

and α, β, γ are given by the same formulas l_0 and l_1 being interchanged.

The two other cases $\tau' = (l_0, l_1 - 1)$ and $\tau' = (l_0 - 1, l_1)$ correspond to the same formulas, with different constants: $c^{\tau\tau'}_{1}$ and $c^{\tau\tau'}_{2}$.

APPENDIX II. INVARIANT WAVE EQUATIONS[1,7]

Let us consider the first-order equation:

$$\sum_i L_i\frac{\partial\psi}{\partial\chi_i} + i\chi\Psi = 0, \tag{A1}$$

where $\chi \in \mathrm{R}$, L_i are certain matrices which act on a space V, $\psi_{(x)}$ is a vector of V.

Following Gel'fand, Minlos, and Shapiro,[7] we say that Eq. (A1) is invariant under a group of (transformations G, if its form is unchanged by the simultaneous transformations

$$x'_i = \sum_k g_{ki}x_k, \qquad g \in G, \tag{A2}$$

$$\Psi'(x') = T_g\Psi(x), \tag{A3}$$

where $g \to T_g$ is a representation of the group G.

Now it is easily seen[6] that this assumption implies the following condition for the L_i:

$$T_gL_iT_g^{-1} = \sum_k g_{ki}L_k. \tag{A4}$$

For $G = R$ the rotation group, $i, k = 1, 2, 3$. For $G = \mathfrak{L}$ the Lorentz group, $i, k = 0, 1, 2, 3$. Then we can consider the condition equivalent to (4) on the Lie algebra, by putting

$$T_g = 1 + \epsilon A_g,$$

$$g(\epsilon) = 1 + \epsilon\,\delta, \qquad \delta_{ii} = 0,$$

we have

$$[A_g, L_i] = \sum_k \delta_{ki}L_k. \tag{A5}$$

But (A5) means that, in case $G = R$, the set of L_i ($i = 1, 2, 3$) is isomorphic to D^1 as a G-module, and, in case $G = \mathfrak{L}$, (A5) means that this set ($i = 0, 1, 2, 3$) is isomorphic to $D^{\frac{1}{2},\frac{1}{2}}$.

In other words, finding the matrices L_i is equivalent to solving Eq. (1a) of the representation problem. We then recover the values of the possible L_i: they are given by the matrices T.

For example, the correspondence with the Gel'fand, Shapiro, and Minlos results is in the case $G = R$:

$$(L_+)_{ij} = T_{ij}(e^1_1),$$

$$(L_-)_{ij} = T_{ij}(e^1_{-1}),$$

$$(L_3)_{ij} = -T_{ij}(e^1_0),$$

thus explaining the "transition rules" for the matrices L_i, and the necessity of using a sum of irreducible representations of the group G, in order to have nonzero matrices L_i, i.e., a nontrivial invariant wave equation (1) must involve wavefunctions ψ which are defined on a sum of irreducible spaces.

Higher-Order Poles in the S-Matrix*

C. R. HAGEN
Department of Physics and Astronomy, University of Rochester, Rochester, New York
(Received 15 November 1965)

The question of the possible existence of multiple poles in the S-matrix is discussed. It is shown by means of a simple model that such poles are not generally incompatible with Lagrangian field theory and can only be excluded in axiomatic formulations by the assumption of a unique correspondence between stable particles and the poles of the S-matrix on the physical sheet.

THE condition that the poles of the S-matrix be simple has for some time been accepted as an essential ingredient of any consistent theory of elementary particles. Although a considerable effort[1-3] has been recently made to demonstrate the fact that multiple poles of the S-matrix may in fact occur on unphysical sheets[4] (thus giving rise to deviations from purely exponential decay for unstable systems), the possibility that such poles could occur on the physical sheet has largely been disregarded.

The apparent basis for this exclusion of higher-order poles consists in Streater's assertion[5] that a necessary and sufficient condition for a field to satisfy the asymptotic condition is that its retarded vacuum expectation values have the usual one-particle singularities. Thus, one need only accept the assumption of the existence of an asymptotic field operator corresponding to each pole of the S-matrix to immediately exclude the possibility of higher-order poles. While it is not our intention to suggest here that this view is necessarily incorrect in the context of strong interactions, we wish to remark that Lagrangian field theory is rich enough to provide some examples of trivial theories for which the poles of the S-matrix have no particle interpretation and can consequently be of higher order.

For the sake of definiteness we choose to illustrate this point by reference to the case of the scattering of two spinless particles having masses m_b and m_c. It is further assumed that these particles interact with a third meson of mass m_a through the virtual transition

$$a \rightleftarrows b + c.$$

The scattering amplitude

$$f = e^{i\delta} \sin \delta/\rho$$

with

$$\rho = \{[s - (m_b + m_c)^2][s - (m_b - m_c)^2]/4s^2\}^{\frac{1}{2}}$$

has the well-known representation

$$f = \bar{f} + (g^2/8\pi)\Gamma\, \Delta\Gamma, \tag{1}$$

the convenience of which is derived from the fact that f and $\bar{f}$ are each unitary amplitudes, i.e., in the elastic region

$$f(s + i\epsilon) - f(s - i\epsilon) = 2i\rho f f^*,$$
$$\bar{f}(s + i\epsilon) - \bar{f}(s - i\epsilon) = 2i\rho \bar{f}\bar{f}^*.$$

In Eq. (1), $\Delta(s)$ is the renormalized propagator associated with the field $\phi_a(x)$ and $\Gamma(s)$ is the proper vertex function normalized to unity for all three particles on the mass shell.

The basic observation consists in noting that a pole in the vertex function in certain cases may give rise to a double pole in the scattering amplitude. It is to be emphasized that this in no way contradicts the recent discussion by Goebel and Sakita[6] on the significance of the poles of the vertex function. This can be illustrated most clearly by displaying the analytic properties of the propagator and vertex function on the second sheet[7] (in obvious notation) in the form

$$\Delta_{II} = \Delta_I S^{-1} \bar{S},$$

* This research was supported in part by the U. S. Atomic Energy Commission.

[1] M. L. Goldberger and K. M. Watson, Phys. Rev. **136**, B1472 (1964).

[2] R. J. Eden and P. V. Landshoff, Phys. Rev. **136**, B1817 (1964).

[3] J. S. Bell and C. J. Goebel, Phys. Rev. **138**, B1198 (1965).

[4] It is perhaps instructive to note here that such double poles always occur for at least one value of the coupling constant when an S-wave virtual or bound state becomes unstable. While this is in no way relevant to the question of the exponential decay of a resonance, it does provide a strikingly simple example of the coincidence of poles on an unphysical sheet.

[5] R. F. Streater, Nuovo Cimento **25**, 274 (1962). See also W. Zimmermann Nuovo Cimento **13**, 503 (1959); *ibid.* **16**, 690 (1960); K. Symanzik, J. Math. Phys. **1**, 249 (1960).

[6] C. J. Goebel and B. Sakita, Phys. Rev. Letters **11**, 293 (1963). See also Y. S. Jin and S. W. MacDowell, Phys. Rev. **137**, B688 (1965).

[7] We here identify the second sheet with that obtained by passing through the elastic portion of the unitarity cut.

$$\Gamma_{II} = \Gamma_I \bar{S}^{-1},$$

where

$$S = 1 + 2i\rho f_I,$$
$$\bar{S} = 1 + 2i\rho \bar{f}_I.$$

On the basis of this result the authors of Ref. 6 have pointed out that a zero of $\bar{S}$ which gives rise to a pole in Γ_{II} is also responsible for producing a corresponding zero in Δ_{II}. Since an appropriate increase in the coupling will generally make it possible for such a pole to pass through the elastic threshold and appear on the physical sheet, $\Gamma_I(s)$ can itself acquire a pole. However, because $\Delta\Gamma$ is regular at this point, the term $\Gamma\Delta\Gamma$ has only a simple pole which is subsequently canceled by a corresponding singularity in $\bar{f}$.[6]

It is to be noted that this discussion of the significance of vertex function poles is valid only if these poles are associated with a singularity in $\bar{S}$ or $\bar{S}^{-1}$. However, there is no compelling reason to exclude the possibility of poles simultaneously appearing in Γ_I and Γ_{II} at a regular point of $\bar{S}$. Such poles need not coincide with a zero of $\Delta_I(s)$ and would consequently lead to a double pole in f_I.

In order to illustrate this possibility we select a model for which $\bar{f} \equiv 0$, a choice which entirely excludes the situation discussed by Goebel and Sakita. In this case $\Gamma_I = \Gamma_{II}$ and all of the above remarks can be realized in a Zachariasen-type model[8] which has a "built-in" vertex pole. In keeping with the assertion that this possibility is realizable in Lagrangian field theory we write for this model the Lagrangian[9]

$$\mathfrak{L} = \tfrac{1}{2}(\phi\, \partial_\mu\phi^\mu - \phi^\mu\, \partial_\mu\phi) - \tfrac{1}{2}m_0^2\phi^2 + \tfrac{1}{2}\phi^\mu\phi_\mu$$
$$+ \frac{1}{2}\int_{(m_b+m_c)^2}^{\infty} ds$$
$$\times \{\phi(s)\, \partial_\mu\phi^\mu(s) - \phi^\mu(s)\, \partial_\mu\phi(s) - s\phi^2(s) + \phi^\mu(s)\phi_\mu(s)\}$$
$$+ g_0 \frac{1}{(8\pi^2)^{\frac{1}{4}}}\phi \int_{(m_b+m_c)^2}^{\infty} ds\, \phi(s)\rho^{\frac{1}{2}}(s)\Gamma(s).$$

The propagators

$$Z_3\, \Delta(x - x') = i\langle 0|\, (\phi(x)\phi(x'))_+\, |0\rangle,$$
$$\Delta(x - x'; s, s') = i\langle 0|\, (\phi(x, s)\phi(x', s'))_+\, |0\rangle$$

have the momentum-space representation

$$Z_3\, \Delta(p) = \left[p^2 + m_0^2 - \frac{g_0^2}{8\pi^2}\int_{(m_b+m_c)^2}^{\infty} ds \frac{\rho(s)\Gamma^2(s)}{p^2 + s}\right]^{-1},$$

$$\Delta(p, s, s') = \frac{\delta(s - s')}{p^2 + s} + \frac{g^2}{8\pi^2}\frac{\rho^{\frac{1}{2}}(s)\Gamma(s)}{p^2 + s}\, \Delta(p)\, \frac{\rho^{\frac{1}{2}}(s')\Gamma(s')}{p^2 + s'},$$

where we have introduced $g^2 = Z_3 g_0^2$.

By making use of the fact that $\phi(x, s)$ creates (b, c) pairs, one finds for the scattering amplitude the required form

$$f = (g^2/8\pi)\Gamma\, \Delta\Gamma$$

which is clearly unitary if Γ has no right-hand cut. Since $\Gamma(s)$ is furthermore to be a real analytic function of s with no left-hand cut, it follows immediately that $\Gamma(s)$ is meromorphic in the entire s plane. As the simplest assumption for $\Gamma(s)$ (other than a constant) we may take the single-pole structure

$$\Gamma(s) = (m_a^2 - s_0)/(s - s_0)$$

$[s_0 < (m_b + m_c)^2, s_0 \neq m^2]$, a form which immediately leads to a double pole in the scattering amplitude at $s = s_0$. One might further mention here that it is possible in this model to put higher-order poles in the vertex function and so generate poles of arbitrarily high order in the S-matrix. Thus it is clear from this result that there is no essential connection between zeros of the propagator and poles of the vertex function. In general it is just the absence of such a correlation which leads to higher-order poles in the S-matrix.

The singularities which we have built into the vertex function have been seen to appear in the analytic structure of the S-matrix despite the fact that they have no direct significance in terms of particles. While the model given here is not truly representative of real field theories in view of the highly arbitrary form of the vertex function, it is of interest that Lagrangian field theory is apparently less restrictive on the form of the scattering amplitude than is the S-matrix approach. Thus the nonapplicability of Streater's work to the model considered here, while clearly a consequence of the absence of a complete set of asymptotic fields, must cast some doubt on the desirability of insisting on a strict identification of S-matrix poles with physical particles.

[8] F. Zachariasen, Phys. Rev. **121**, 1851 (1961).
[9] W. Thirring, Phys. Rev. **126**, 1209 (1962). C. R. Hagen, Ann. Phys. (N. Y.) **31**, 185 (1965).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 7, NUMBER 6 JUNE 1966

Dynamics of Certain Spherical Charge Distributions

RICHARD L. LIBOFF
School of Electrical Engineering, Cornell University, Ithaca, New York
(Received 23 August 1965)

Two initial-value problems are solved in the Lagrange representation. The initial configuration of the first is an isolated, negative, finite spherical gas. In the second a fixed, infinitely thin, positive, permeable, spherical shell is placed concentric and exterior to the negative gas of the first configuration. The charge of the total system is zero. The first system is totally soluble. Application to the special case of uniform initial density gives that the charge density at a point far removed from the initial sphere decays as the inverse cube of time, and is independent of radius. In the second configuration if a is the radius of the positive shell and b the initial radius of the negative gas, then for uniform initial distribution, all charge in the shell $b(b/a)^{\frac{1}{2}} \leq r \leq b$ escapes. The remaining charge oscillates and is randomized by phase mixing.

1. INTRODUCTION

IN this paper we consider the dynamics of certain charged spherical systems.[1] Understanding the motion of such models is of importance in three distinct areas of study. In cosmology it is relevant to the Lyttleton–Bondi[2] universe which is hypothesized to carry a net charge. If the initial distribution is assumed spherical, then to within classical estimates, knowledge of the subsequent dynamics lends to our understanding of the evolution of the universe, and may lead to inductive support of the theory. Secondly, in the study of nuclear detonations,[3] one is concerned with the effects which spherically expanding ionizing radiation has on the natural medium of transport. Such radiation tends to leave a positive "residue" behind the radial "primary" electrons. The question arises as to the nature of the oscillations that such an initial configuration will enduce.

With regard to plasma physics, the motion of charged spherical systems affords an opportunity for obtaining exact nonlinear solutions. It is particularly relevant to the theory of "large amplitude" oscillations[4] where it is very possible to find large local accumulations of net charge. The question arises as to the influence which the natural "Coulomb dispersion" has on such charge accumulations, so that the problem of obtaining characteristic diffusion times of such aggregates becomes important. If these times are sufficiently short, then they may well accompany an over-all destruction in the wave form, and the wave will damp. This is a form of "phase mixing" phenomena[5,6] which has been suggested to be the essential mechanism which underlies Landau damping.[7] Another explanation by Jackson[8] attributes this "collisionless" damping to a "trapping mechanism" in which the wave expends energy in accelerating particles. Other mechanisms attribute the damping to resonant energy absorption.[9-11] Weitzner,[12] in an analysis of the Landau problem, finds that the Vlasov equation, in the region of large wavenumber, does not describe decaying plane waves, but rather a pure relaxation of the given perturbation to equilibrium. This, of course, casts strong doubt on the conclusions of the original Landau analysis. The relaxation which Weitzner finds goes as the reciprocal power of the time, the exact behavior depending critically on the nature of the original distribution. (Long-wavelength plasma oscillations remain undamped.) Such inverse power decay is, of course, strongly suggestive of a diffusive mechanism.

Still another area related to the included analysis lies within the realm of nonequilibrium statistical mechanics. The configurations considered are of particular interest because they naturally evolve into unbounded systems. The procedure of dropping the usual "surface terms" which accompany most

[1] Similar problems are treated by: J. Dawson, Phys. Rev. **113**, 383, (1959) and H. Derfler, Tech. Rept. No. 104–7, Stanford Electronics Lab., (1961).
[2] R. A. Lyttleton and H. Bondi, Proc. Roy. Soc. (London) **A252**, 313 (1959).
[3] W. Karzas and R. Latter, Phys. Rev. **137**, B1369 (1965); M. Johnson and B. Lippman, *ibid.* **119**, 827 (1960).
[4] P. A. Sturock, J. Nucl. Energy **C2**, 158 (1961); Proc. Roy. Soc., (London) **A242**, 277 (1957); D. A. Tidman and G. H. Weiss, Phys. Fluids **4**, 866 (1961).
[5] M. Kruskal in *The Theory of Neutral and Ionized Gases* (Hermann & Cie, Paris, 1960), pp. 287–291.
[6] J. Linhart in *Plasma Physics* (North-Holland Publishing Company, Amsterdam, 1960), pp. 111–112.
[7] L. Landau, J. Phys. (USSR) **10**, 25 (1946).
[8] J. D. Jackson, J. Nucl. Energy **C1**, 171 (1960).
[9] J. Dawson, Phys. Fluids **4**, 869 (1961).
[10] A. Kidal, Nuovo Cimento **20**, 104 (1961).
[11] C. S. Wu, Phys. Rev. **127**, 1419 (1962).
[12] H. Weitzner, Phys. Fluids **6**, 1123 (1963); *ibid.* **7**, 476 (1964); Commun. Pure Appl. Math. **18**, 307 (1965).

H-theorem discussions[13] becomes incorrect. The fact that the systems are unbounded also precludes Poincaré recurrence arguments. The final state attained is a strongly "preferred" state. Such an occurrence is of interest because it may be described precisely by a "collisionless" (Vlasov) Boltzmann equation, and affords an example of this type equation generating an equilibrium configuration wholly distinct from the initial state.

The included analysis is similar to previous work on the thermalization of charged sheets due to crossing.[14,15]

2. DESCRIPTION OF CONFIGURATIONS AND RESULTS

Two spherically symmetric initial distributions are considered. The first is an isolated, negatively charged, finite spherical gas of uniform density.[16]

In the second, a fixed, infinitely thin, positive, permeable, spherical shell is placed concentric and exterior to the negative gas of the first configuration. If the radius of the negative gas is b and that of the positive shell is a, then in the initial configuration, $a > b$. The charge of the negative gas is $-Q_0$ while that of the positive shell is $+Q_0$.

The dynamics of the first configuration is exactly soluble. The technique employed to solve the problem centers about writing equations of motion for shells of differential thickness. Each shell carries with it an "s number." The variable s is the initial radius value of a shell. This method of writing an equation of motion for the radius r of a shell is indicative of a Lagrangian representation, as opposed to the more standard Eulerian[17] one. It is better suited for this type problem as long as the shells do not intersect. In this event the problem takes on an added element of complexity, and the Eulerian representation becomes more appropriate.

For the first problem of an initial uniformly charged sphere of finite radius, one readily concludes that shells do not intersect. As an immediate consequence, a given shell "sees" a constant amount of charge concentrated at the origin. The problem is forthwith reduced to quadratures. An asymptotic estimate for radius large compared to the initial s value gives that the charge density decays as the inverse cube of time, with a characteristic decay time which is the inverse of the classical plasma frequency (associated with the initial charge distribution). This result is independent of the radial position r.

In the second initial configuration, a positive, fixed, permeable, infinitely thin shell (charge $+Q_0$) is concentric and exterior to the negative gas considered in the first problem. The radius of the positive shell is a, while that of the initial spherical negative gas is b. In the initial configuration $a > b$. A vacuous shell of thickness $a - b$ separates the positive outer shell from the interior negative gas.

The dynamics of this problem is only partially reducible to quadratures. Namely, it is so for those shells of the negative expanding gas which do not intersect.

The motion which unfolds is as follows. When the negative gas is released from its initial spherical configuration of radius b, it expands under Coulomb repulsion. As the first infinitesimally thin shell passes through the positive shell at $r = a$, it leaves behind an equal and opposite charge, which it sees concentrated at the origin and which attracts it back toward the origin. However, by the time it has reached $r = a$, it has a finite velocity, so that it escapes. Indeed, all charge with s number, $s \geq b(b/a)^{\frac{1}{2}}$ escapes. Furthermore the shells which comprise this escaping motion, do not intersect and this portion of the problem is reduced to quadratures.

The charge, which is bound, oscillates. However the analytic content of the motion is obscured by intersecting shells. The dynamics is formulated within a Lagrangian representation. An expression is obtained for the frequency of oscillation—in a nonintersecting shell approximation. The frequency of oscillation vanishes at the origin and at $s = b(b/a)^{\frac{1}{2}}$. In the remaining interval it is of the order of the plasma frequency of the initial distribution. The vanishing frequency for particles at the origin is consistent with the observation that the particle at $r = 0$, remains there inasmuch as it feels no force. For small s number the (least) time it takes for a particle to reach $r = a$ is

$$t_T \approx a/\omega_p s.$$

It is these very slow moving inner shells which must ultimately "collide" (i.e., "pass through") with the oscillating outer shells. When such an event occurs the force on the slowly expanding shell increases, whence the parenthetical "least" in the above statement.

The system is characterized by its partially un-

[13] Such as, for instance, a proof by W. Newcomb on the constancy of H implied by the Vlasov equation, which appears in I. Bernstein, Phys. Rev. 109, 10 (1958).
[14] O. Buneman, Phys. Rev. 115, 503 (1959).
[15] J. Dawson, Phys. Fluids 7, 419 (1964).
[16] The results obtained are valid for all initial charge densities which do not increase faster than a constant, with decreasing radius.
[17] In the Eulerian representation one asks: what is the (say) density at r and t. In the Lagrangian representation one asks what is r at t (of a particle, fluid element, etc.).

bound–expanding component and partially bound–oscillating component. The ultimate state of the system is drastically distinct from the initial very well ordered state. The bound–oscillating component undergoes a gross phase mixing procedure, which is additionally catostrophic since the system violates the hypothesis of the Poincaré recurrence theorem and need never necessarily return to its initial configuration. In addition, owing to the intersection of orbits at such radii, there is an increase in temperature, i.e., a spread in the local velocity. Furthermore, calculation of the Boltzmann H function, employing a Vlasov equation, indicates that the function decreases, this owing to the particles which ultimately cross the surface at infinity.[18]

The cosmological significance of this second model is that (to within classical estimates) such an initial configuration would lead to a partially expanding, partially oscillating universe.

3. DISPERSION OF A CLOUD OF NEGATIVE CHARGE

We consider a spherically symmetric single-species charge distribution which at time $t = 0$ is released from a stationary position with a prescribed distribution. One wishes to uncover the charge density at time $t > 0$.

The initial charge distribution is given by

$$Q(t = 0) = Q_0(s) = \int_0^s q_0(s) 4\pi s^2 \, ds, \quad 0 \leq s \leq b, \qquad (1)$$
$$q_0 = 0, \qquad s > b.$$

The function Q represents the total charge contained within a sphere of radius s. The initial charge density is q_0. At subsequent time, $Q = Q(r, t)$ and the related charge density is given by

$$\partial Q/\partial r = 4\pi r^2 q(r). \qquad (2)$$

The equation of motion for a shell of charge Q, of thickness dr at r is

$$[4\pi r^2 q \, dr]\left(\frac{m}{e}\right)\ddot{r} = \frac{[4\pi r^2 q \, dr]}{4\pi\epsilon_0 r^2} \int_0^r 4\pi\eta^2 q(\eta, t) \, d\eta, \qquad (3.1)$$

or equivalently,

$$\ddot{r} = eQ/4\pi\epsilon_0 m r^2. \qquad (3.2)$$

The permittivity of free space is ϵ_0. The mass-to-charge ratio of a gas particle is (e/m). In the event that q_0 is such that shells do not collide in the subsequent motion (i.e., an inner "s shell" never overtakes an outer "s shell"), the charge contained in any "s sphere" (an "s sphere" is the sphere interior to an "s shell") is constant, and Q obeys the equation

$$Q(r_s - \dot{r}_s t) = \text{const} = Q_0(s). \qquad (4)$$

Here $\dot{r}_s$ is again the velocity of a specific s shell. Equation (4) states that the charge contained within any expanding s sphere remains constant. Such development is ensured if the initial charge density does not increase faster than a constant, with decreasing radius. If, for instance, q_0 is independent of s, then the acceleration of any s shell varies as s, and outer shells expand outward faster than interior ones. If q_0 obeys this criterion, then a first integral of Eq. (3.1) is

$$\dot{r}^2 + (\alpha^2/r) = \alpha^2/s, \qquad (5)$$

where

$$\alpha^2 = eQ(s)/2\pi\epsilon_0 m. \qquad (6)$$

In the event that $q_0(s)$ is uniform, then

$$\alpha^2 = \alpha_0^2(s/b)^3 = \omega_p^2 s^3, \qquad (7.1)$$

where ω_p is the plasma frequency of the initial configuration,

$$\omega_p^2 = 2eQ_0/3\epsilon_0 b^3 m = (8\pi e/9\epsilon_0 m) q_0. \qquad (7.2)$$

More generally the solution to Eq. (5) is

$$\alpha t/s^{\frac{3}{2}} = (r/s)^{\frac{1}{2}}[(r/s) - 1]^{\frac{1}{2}} + \cosh^{-1}(r/s)^{\frac{1}{2}}. \qquad (8)$$

Equation (8) may be rewritten in terms of the charge Q,

$$Q^{\frac{1}{2}} = t^{-1}(2\pi\epsilon_0 m s^3/e)^{\frac{1}{2}} \times \{(r/s)^{\frac{1}{2}}[(r/s) - 1]^{\frac{1}{2}} + \cosh^{-1}(r/s)^{\frac{1}{2}}\}. \qquad (9)$$

The value which Q assumes on the left-hand side relates to the s value on the right-hand side. These two variables are, of course, equivalent Lagrange parameters and are uniquely related through the initial data, $Q_0 = Q_0(s)$. It is this value of Q which is carried along in a manner described by the right-hand side of Eq. (9).

In order to obtain the charge density q, as given by Eq. (2), the initial data equation must be inverted [$s = s(Q)$] and then substituted into the right-hand side of Eq. (9) to obtain the desired form $Q = Q(r, t)$. Although this recipe renders the problem solved, in principal, the transcendental nature of the resulting algebraic equations suggests that final solution must be numerically evaluated. However, as usual, analytic forms may be obtained in certain asymptotic limits. The case in point relates to the domain, $r \gg b \geq s$. In this region Eq. (9) becomes

[18] R. L. Liboff (unpublished).

$$Q^{\frac{1}{2}} \sim t^{-1}(2\pi\epsilon_0 ms/e)^{\frac{1}{2}}[r + \tfrac{1}{2}s \ln (r/s)] \sim (r/t)(2\pi\epsilon_0 ms/e)^{\frac{1}{2}}. \quad (10)$$

If the initial charge distribution is uniform, then one obtains [from Eqs. (2), (7.1), and (10)]

$$q/q_0 \sim (t\omega_p)^{-3}. \quad (11)$$

Thus we obtain the interesting result that the charge density at a given position r, decays as the inverse cube of time, but is independent of r, granted that $r \gg b$. The decay time is proportional to the inverse classical plasma frequency. Again, the inverse decay in time is indicative of classical diffusion.

4. DYNAMICS OF A SPHERE OF NEGATIVE GAS INITIALLY INTERIOR TO A POSITIVE SHELL

In this second case we consider that the initial negative uniform spherical charge of radius b, and net charge $-Q_0$, is surrounded by a positive fixed, perfectly permeable, infinitely thin shell of net charge $+Q_0$, and radius a, $a > b$.

While in the above case one is most concerned with obtaining an expression for the time behavior of the density at a fixed position, the properties of the present configuration are more numerous, and it suffices to obtain information regarding the more fundamental of these properties.

First it is important to discern which solutions are still obtainable within a noninteracting orbit formalism. If, for instance $a \leq b$, such solution is impossible. In this case ($a < b$) the finite negative shell which lies exterior to $r = b$ is attracted inward, while the sphere interior to $r = b$ expands outward, and shells in the neighborhood of the positive fixed shell intersect. If the initial configuration is such that the negative gas just fills the cavity defined by the positive surface (i.e., $a = b$), then interior shells again overtake the more lethargic ones at $r = a$. In these cases solution by noninteresting orbits is impossible.

If, on the other hand, the inner sphere is initially of radius $b < a$, then a portion of the negative gas will pass through the fixed positive shell and escape. The motion of these escaping shells is completely describable in terms of nonintersecting orbit theory, and is reducible to quadratures.

In the subsequent analysis we obtain the following information: (1) the critical s-number which separates the escaping outer shells from the bound interior sphere, (2) a general expression for the turning points of all shells, (3) a qualitative description of the motion of all shells, (4) frequency of oscillation of bound shells, in the nonintersecting orbit approximation, (5) the least time for a small s-number shell to reach the positive shell. The salient features of the analysis appear in the form of four theorems.

At any instant prior to two orbits crossing, all s shells satisfy the energy equation

$$\dot{r}^2 + (\alpha^2/r)[1 + (Q_0/Q)([r/a] - 1)H([r/a] - 1)] = \alpha^2/s. \quad (12)$$

The augmented positive-layer charge, $Q_0([r/a] - 1)$, guarantees a continuous potential across $r = a$. Also, recall that all shells are such that $s \leq b$. The step function $H(x)$ is such that $H = 1$ for $x \geq 0$ and vanishes for all other x.

As in the proceeding analysis, $Q(s)$ is the charge initially contained within the corresponding s sphere, while Q_0 is the absolute value of the charge initially within the sphere $r = b$. The above equation gives the quadrature

$$\alpha t = \int_s^r \frac{g\,dz}{[(g^2/s) - (H/a) + (H - g^2)/z]^{\frac{1}{2}}}, \quad (13)$$

$$g^2 = Q/Q_0, \quad (14)$$

where α is as defined in Eq. (6). For shells which have passed through the positive surface at $r = a$, this integral becomes

$$\alpha t/2s^{\frac{1}{2}} = \int_1^{(a/s)^{\frac{1}{2}}} \frac{u^2\,du}{(u^2 - 1)^{\frac{1}{2}}} + (g/h) \int_{(a/s)^{\frac{1}{2}}}^{(r/s)^{\frac{1}{2}}} \frac{u^2\,du}{(u^2 + \beta^2)^{\frac{1}{2}}}. \quad (15)$$

The dummy variable u^2 replaces (r/s). The variables β and h are

$$\beta^2 \equiv [1 - g^2]/h^2, \quad (16.1)$$

$$h^2 \equiv [g^2 - (s/a)]. \quad (16.2)$$

These variables are sketched in Fig. 1. For a uniform initial charge distribution, if $s < b(b/a)^{\frac{1}{2}}$ then h^2 is a negative number, whence β^2 is also negative. In this domain we set

$$y^2 \equiv -h^2 \geq 0; \quad \gamma^2 \equiv -\beta^2 \geq 0, \text{ for } s \leq b(b/a)^{\frac{1}{2}} \quad (17)$$

and Eq. (15) becomes

$$\alpha t/s^{\frac{1}{2}} = \int \frac{u^2\,du}{(u^2 - 1)^{\frac{1}{2}}} + (g/y) \int \frac{u^2\,du}{(r^2 - u^2)^{\frac{1}{2}}}. \quad (18)$$

The limits on the integrals are the same as in Eq. (15). The cutoff s number $b(b/a)^{\frac{1}{2}}$ is a very relevant parameter. Consider for instance the problem of evaluating the turning points of shells. This may be obtained quite simply from Eq. (12) by merely setting $\dot{r} = 0$, and realizing that the first turning

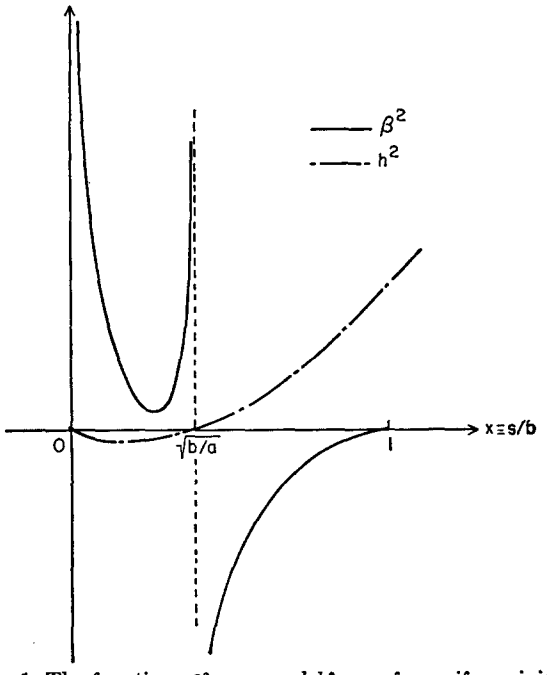


FIG. 1. The functions β^2 vs x, and h^2 vs x for uniform initial charge distribution.

value must be at $r \geq a$, since there is no attraction toward the origin for $r < a$. There follows, for arbitrary $Q(s)$,

$$r_T/s = \frac{1 - [Q(s)/Q_0]}{(s/a) - [Q(s)/Q_0]}. \tag{19}$$

In this formula r_T is the turning radial value of the corresponding s shell. Inasmuch as r_T is necessarily a positive number, s shells with s numbers which satisfy

$$\frac{1 - Q/Q_0}{s/a - Q/Q_0} < 0, \tag{20}$$

escape. If, on the other hand, this ratio is positive then corresponding s shells are turned back toward the origin, i.e., they oscillate. The following theorem results if the denominator in Eq. (19) is expanded about $s = 0$.

Theorem I: If a negatively charged gas which initially contains an amount of charge Q in a sphere of radius $s(s \leq b)$, and is concentric with an infinitely thin, permeable, shell of positive charge equal and opposite to the entire charge contained within the negative gas, and of radius $a > b$, is such that $(Q/Q_0)(a/s) \to 0$ as $s \to 0$, then there is a neighborhood about $s = 0$ which contains oscillating orbits. The turning value, r_T, of these orbits in the limit as $s \to 0$, (excluding the point at $s = 0$) is at least a.

The relevance of the critical s number, $b(b/a)^{\frac{1}{2}}$ pertains to the special case of uniform initial charge distribution. In this event (r_T/b) becomes

$$r_T/b = \frac{1 - x^3}{(b/a) - x^2} \tag{21}$$

This curve is sketched in Fig. 2, from which it is clear that all s shells with

$$s \geq b(b/a)^{\frac{1}{2}}$$

escape. This observation yields the following:

Theorem II: A uniformly negatively charged sphere initially of radius b which is concentric with an infinitely thin, fixed, permeable shell of positive charge, equal and opposite to the entire charge of the negative gas and of radius $a > b$, will lose all charge which is initially exterior to and including the surface of the sphere of radius $b(b/a)^{\frac{1}{2}}$.

Furthermore, since larger shells with larger s value feel a stronger repulsion and a weaker attraction (back toward the origin) than do smaller s-value shells, those shells which escape do not collide. In this domain [viz., $s \geq b(b/a)^{\frac{1}{2}}$] the integral, Eq. (15), gives the precise solution

$$\alpha t/s^{\frac{1}{2}} = \{(a/s)^{\frac{1}{2}}[(a/s) - 1]^{\frac{1}{2}} + \cosh^{-1}(a/s)^{\frac{1}{2}}\} + (g/h)\Big((r/s)^{\frac{1}{2}}[(r/s) + \beta^2]^{\frac{1}{2}} - (a/s)^{\frac{1}{2}}[(a/s) + \beta^2]^{\frac{1}{2}} - \beta^2 \ln \left\{\frac{(r/s)^{\frac{1}{2}} + [(r/s) + \beta^2]^{\frac{1}{2}}}{(a/s)^{\frac{1}{2}} + [(a/s) + \beta^2]^{\frac{1}{2}}}\right\}\Big). \tag{23}$$

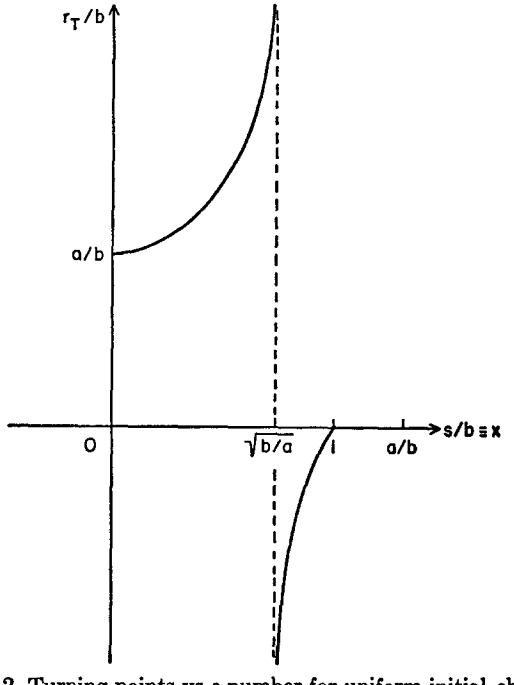


FIG. 2. Turning points vs s number for uniform initial charge distribution.

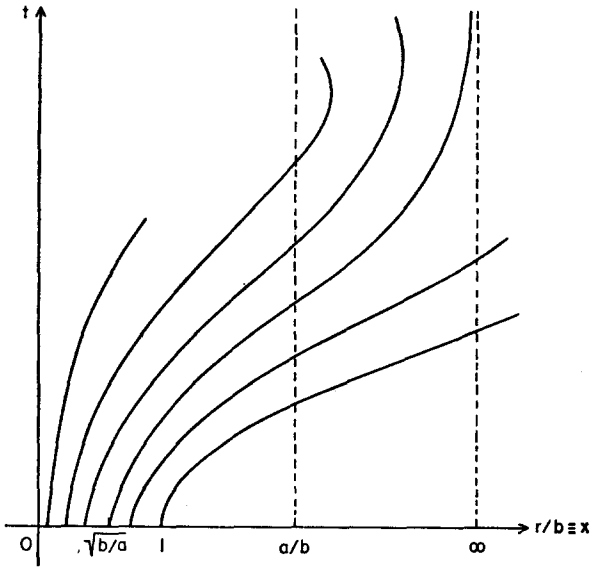


FIG. 3. s-shell characteristics.

The outermost s shell ($s = b$, $\beta^2 = 0$) gives linear motion. More generally, the motion of shells is best depicted on a $t - r$ diagram such as in Fig. 3.

That the remaining shells suffer collisions may be seen by the following argument. Suppose that shells in the bound domain $s < b(b/a)^{\frac{1}{2}}$ do not intersect. Then their precise motion is given by the integral [Eq. (18)].

$$\alpha t/s^{\frac{1}{2}} = \{(a/s)^{\frac{1}{2}}[(a/s) - 1]^{\frac{1}{2}} + \cosh^{-1}(a/s)^{\frac{1}{2}}\} + (g/y)\Big((a/s)^{\frac{1}{2}}[(a/s) - 1]^{\frac{1}{2}} - (r/s)^{\frac{1}{2}}[(r/s) - 1]^{\frac{1}{2}} + \gamma^2\Big\{\sin^{-1}\left[\frac{(r/s)^{\frac{1}{2}}}{\gamma}\right] - \sin^{-1}\left[\frac{(a/s)^{\frac{1}{2}}}{\gamma}\right]\Big\}\Big). \quad (24)$$

Theorem I indicates that for $s \approx 0$, $r_T \approx a$. The corresponding "turning time" t_T is obtained by evaluating the first integral in Eq. (15).

$$\omega_p t_T \sim [(a/s) + (1/2)\ln(a/s)], \quad (25)$$

$$t_T \sim a/\omega_p s.$$

Theorem III: Under the same conditions as Theorem II, the least time it takes for shells with vanishingly small s number to reach the radius $r = a$, varies as s^{-1}, the coefficient of proportionality being a/ω_p.

Again the adjective "least" incorporates the possibility of intersecting shells.

On the other hand, the time t_0 it takes for the $s = b$ shell to reach the $r = a$ surface is precisely given by inserting $r = a$, $s = b$ into Eq. (8). There results

$$\omega_p t_0 = (a/b)^{\frac{1}{2}}[(a/b) - 1]^{\frac{1}{2}} + \cosh^{-1}(a/b)^{\frac{1}{2}} \quad (26)$$

a finite well-defined value. Combining this result with Theorem III indicates that for all time, $t \geq t_0$, outwardly expanding shells reach the surface $r = a$.

Now let us return to the bound trajectories which follow from a nonintersecting orbit assumption, i.e., Eq. (24). The oscillating part of the expression implies the frequency of vibration ω

$$(\omega/\omega_p)^2 = (b/a - x^2)^3/(1 - x^3)^2 \quad (27)$$

where $x = s/b$. This expression pertains to shells [with s numbers: $0 < s \leq b(b/a)^{\frac{1}{2}}$] during the oscillating part of their history, viz., after they pass through the $r = a$ surface. It gives an estimate of the frequency of oscillation of such shells, which for all values excluding the interval $x^2 \geq b/a$ is a finite number.

Now such oscillating shells, after turning at r_T, given by Eq. (21), continue to approach the origin at all radii, $r > a$. Independent of the motion of the system, a shell exterior to $r = a$ sees a positive charge at the origin and is attracted thereto. It follows that all shells which return reach the value $r = a$. Equation (27) indicates such shells return in a finite time ($\sim\omega^{-1} > t_0$). But one of our former conclusions was that outwardly expanding shells continue to reach $r = a$ for all time $t > t_0$ [cf. Eq. (26)]. It follows that these outwardly expanding shells must, in a finite time, collide with at least a finite set of those shells which have begun to oscillate. But this violates the hypothesis that shells do not collide. We conclude the following.

Theorem IV: For the system described in Theorem II, at a finite time, two s shells with distinct s numbers are at a common radius.

Although orbits ultimately intersect, Eq. (27) is a reasonable estimate for the frequency ω in the early times, before phase mixing totally obscures the dynamics.

4. LAGRANGIAN FORMULATION FOR INTERSECTING ORBITS

If shells intersect it is no longer true that s shells see a constant amount of charge at the origin. As a consequence the energy equation, Eq. (12), must be rewritten,

$$\dot{r}^2 + [\alpha^2(r, t)/r]\left[1 + \frac{Q_0}{Q(r, t)}([r/a] - 1)H([r/a] - 1)\right] = \alpha^2(r, 0)/s. \quad (28)$$

The charge contained within the surface of the sphere of radius r, at the time t is $Q(r, t)$. It is related to the charge density q, through

$$Q(r, t) = \int_0^r q(z, t)4\pi z^2\, dz. \quad (29)$$

In order to construct the function $q(r, t)$, we calculate the charge in the volume element $4\pi r^2 dr$. Suppose only two shells with s numbers s_1 and s_2 intersect at the time t. Then the charge in the said volume element is

$$4\pi r^2\, dr\, q(r, t) = 4\pi s_1^2\, ds_1\, q(s_1) + 4\pi s_2^2\, ds_2\, q(s_2) - 4\pi r^2 a^2\, dr\, Q_0\, \delta(r - a) \quad (30)$$

or equivalently,

$$r^2 q(r, t) = [s^2\, \partial s/\partial r\, q_0(s)]_{s_1} + [s^2\, \partial s/\partial r\, q_0(s)]_{s_2} - Q_0 a^2\, \delta(r - a). \quad (31)$$

The initial charge distribution is q_0. The values s_1 and s_2 are the roots of the equation

$$r - r(s, t) = 0. \quad (32)$$

More generally, if this equation has, say, N roots then Eq. (31) appears as

$$q(r, t) = \sum_{i=1}^{n} (s_i/r)^2(\partial s/\partial r)s_i q(s_i) - Q_0(a/r)^2\, \delta(r - a). \quad (33)$$

Equations (28), (29), (32), and (33) are the desired equations of motion.

ACKNOWLEDGMENTS

The author is indebted to Dr. C. S. Gardner and Professor H. Weitzner for many helpful suggestions rendered during the early development of the paper. He is also grateful to S. Goldblatt of Nuclear Research Associates for suggesting relevant applications of these results, and feels especially indebted to Dr. John M. Dawson for his invaluable criticism of the final manuscript.

This research was supported in part by the Courant Institute of Mathematical Sciences, New York University, under Contract AT(30-1)-1480 with the U. S. Atomic Energy Commission; and in part by the Mechanics Branch of the Office of Naval Research under Contract Nonr 401(25) with the Graduate School of Aerospace Engineering, Cornell University.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 7, NUMBER 6 JUNE 1966

The Most General Clebsch–Gordan Coefficients of the Universal Covering Group of the Inhomogeneous Lorentz Group*

MARTIN KUMMER
The University of Michigan, Ann Arbor, Michigan
(Received 7 October 1964)

The most general Clebsch–Gordan coefficients to reduce the physically most important n-fold product representations of the groups $\tilde{P}_+^\uparrow$ and $\tilde{P}$ (Universal covering groups of the restricted and full inhomogemeous Lorentz groups) are derived. They are used to answer the question: What can be said about the S-matrix if only Lorentz invariance is postulated?

1. INTRODUCTION

SEVERAL authors[1-5] have constructed relativistic angular momentum states to achieve a general "partial-wave decomposition" of the S-matrix. In this work, we intend to approach this problem from a quite general point of view. Namely, we want to determine the most general Clebsch–Gordan coefficients for an n-fold product of irreducible representations of $\tilde{P}_+^\uparrow$ and $\tilde{P}$ (i.e., of the universal covering groups of the restricted and the full inhomogeneous Lorentz group), if only representations of $\tilde{P}_+^\uparrow$ are considered with character $(m, \epsilon = +1, s)(m =$

* This work was supported by the National Science Foundation.
[1] A. J. Macfarlane, Rev. Mod. Phys. 34, 41 (1962).
[2] M. Jacob and G. C. Wick, Ann. Phys. 7, 404 (1959).
[3] J. S. Lomont, J. Math. Phys. 1, 237 (1960).
[4] J. Wehrle, Nucl. Phys. 44, 579, 637 (1963).
[5] H. Joos, Forschr. Physik 10, 65 (1962).

$$\dot{r}^2 + [\alpha^2(r, t)/r]\left[1 + \frac{Q_0}{Q(r, t)}([r/a] - 1)H([r/a] - 1)\right] = \alpha^2(r, 0)/s. \quad (28)$$

The charge contained within the surface of the sphere of radius r, at the time t is $Q(r, t)$. It is related to the charge density q, through

$$Q(r, t) = \int_0^r q(z, t)4\pi z^2\, dz. \quad (29)$$

In order to construct the function $q(r, t)$, we calculate the charge in the volume element $4\pi r^2 dr$. Suppose only two shells with s numbers s_1 and s_2 intersect at the time t. Then the charge in the said volume element is

$$4\pi r^2\, dr\, q(r, t) = 4\pi s_1^2\, ds_1\, q(s_1) + 4\pi s_2^2\, ds_2\, q(s_2) - 4\pi r^2 a^2\, dr\, Q_0\, \delta(r - a) \quad (30)$$

or equivalently,

$$r^2 q(r, t) = [s^2\, \partial s/\partial r\, q_0(s)]_{s_1} + [s^2\, \partial s/\partial r\, q_0(s)]_{s_2} - Q_0 a^2\, \delta(r - a). \quad (31)$$

The initial charge distribution is q_0. The values s_1 and s_2 are the roots of the equation

$$r - r(s, t) = 0. \quad (32)$$

More generally, if this equation has, say, N roots then Eq. (31) appears as

$$q(r, t) = \sum_{i=1}^{n} (s_i/r)^2(\partial s/\partial r)s_i q(s_i) - Q_0(a/r)^2\, \delta(r - a). \quad (33)$$

Equations (28), (29), (32), and (33) are the desired equations of motion.

ACKNOWLEDGMENTS

The author is indebted to Dr. C. S. Gardner and Professor H. Weitzner for many helpful suggestions rendered during the early development of the paper. He is also grateful to S. Goldblatt of Nuclear Research Associates for suggesting relevant applications of these results, and feels especially indebted to Dr. John M. Dawson for his invaluable criticism of the final manuscript.

This research was supported in part by the Courant Institute of Mathematical Sciences, New York University, under Contract AT(30-1)-1480 with the U. S. Atomic Energy Commission; and in part by the Mechanics Branch of the Office of Naval Research under Contract Nonr 401(25) with the Graduate School of Aerospace Engineering, Cornell University.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 7, NUMBER 6 JUNE 1966

The Most General Clebsch–Gordan Coefficients of the Universal Covering Group of the Inhomogeneous Lorentz Group*

MARTIN KUMMER
The University of Michigan, Ann Arbor, Michigan
(Received 7 October 1964)

The most general Clebsch–Gordan coefficients to reduce the physically most important n-fold product representations of the groups $\tilde{P}_+^\uparrow$ and $\tilde{P}$ (Universal covering groups of the restricted and full inhomogemeous Lorentz groups) are derived. They are used to answer the question: What can be said about the S-matrix if only Lorentz invariance is postulated?

1. INTRODUCTION

SEVERAL authors[1-5] have constructed relativistic angular momentum states to achieve a general "partial-wave decomposition" of the S-matrix. In this work, we intend to approach this problem from a quite general point of view. Namely, we want to determine the most general Clebsch–Gordan coefficients for an n-fold product of irreducible representations of $\tilde{P}_+^\uparrow$ and $\tilde{P}$ (i.e., of the universal covering groups of the restricted and the full inhomogeneous Lorentz group), if only representations of $\tilde{P}_+^\uparrow$ are considered with character $(m, \epsilon = +1, s)(m =$

* This work was supported by the National Science Foundation.
[1] A. J. Macfarlane, Rev. Mod. Phys. 34, 41 (1962).
[2] M. Jacob and G. C. Wick, Ann. Phys. 7, 404 (1959).
[3] J. S. Lomont, J. Math. Phys. 1, 237 (1960).
[4] J. Wehrle, Nucl. Phys. 44, 579, 637 (1963).
[5] H. Joos, Forschr. Physik 10, 65 (1962).

mass, ϵ = sign of the energy, $m \geq 0$), and if we confine ourselves to representations of $\tilde{P}$ with the same character, and for which, in addition, the square of the time-inversion and the square of the space–time inversion are represented by

$$(-1)^{2s}.$$

For the explicit construction of these representations see Refs. 6–10.

We start from states which are represented by square-integrable functions of the four-momentum p on the mass-hyperboloid $p^2 = m^2$, and which at the same time are eigenfunctions of the spin operator

$$S_3(p).$$

[For its definition see (2.18).] It is well known that the spin operator of a particle in a coordinate system in which the four-momentum p is measured is defined only up to a transformation induced by an element of the little group of p.

However, we want to start from states for which the spin of the lth particle is defined quite generally. The only assumption we make is that the spin operators corresponding to physically indistinguishable particles are equally defined. For the convenience of the reader, we rederive many results which can already be found in literature at different places, e.g., the decomposition rules of the product representation of $\tilde{P}_+^\uparrow$ can be found in Refs. 3, 5, and 10. The corresponding rules for $\tilde{P}$ which are also given in our paper, as far as the author knows, cannot be found elsewhere. The works of Jacob and Wick[2] and Wehrle[4] served as a guide for the author. The present work may be looked upon as a generalization of their work, in which the helicity formalism is exclusively used.

We apply our methods also to the S-matrix which we first represent as a sum of covariants of $\tilde{P}_+^\uparrow$ multiplied by invariant amplitudes. These may then be decomposed into partial-wave amplitudes. Finally, the connection between this approach and the M-functions of Stapp[11,12] is given (see also Ref. 12).

2. SOME FACTS ABOUT $SL(2,C)$

It is well known that there exists a one-to-one correspondence between the Hermitian two-by-two matrices and the vectors of Minkowski space p,

$$p = (p^0, p^1, p^2, p^3) \leftrightarrow p_\sim = \sigma^\mu p_\mu, \tag{2.1}$$

where

$$\begin{aligned} \sigma^0 &= \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right), = \mathbb{1}, \quad \sigma^1 = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right), \\ \sigma^2 &= \left(\begin{smallmatrix} 0 & -i \\ i & 0 \end{smallmatrix}\right), \quad \text{and} \quad \sigma^3 = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right) \end{aligned} \tag{2.2}$$

are the Pauli-matrices. It follows that

$$\det p_\sim = (p^0)^2 - \mathfrak{p}^2 = m^2. \tag{2.3}$$

We shall only consider the case where p lies in the future cone.
By

$$Ap_\sim A^\dagger = [\Lambda(A)p]_\sim, \tag{2.4}$$

a two–one homomorphism between $SL(2, C)$ (Group of complex unimodular two-by-two matrices) and the restricted homogeneous Lorentz group is established. If we introduce the caret operation for any two-by-two matrix by the definition

$$\hat{A} = \epsilon \bar{A} \epsilon^{-1}, \qquad \epsilon = -i\sigma^2 \tag{2.5}$$

(the bar indicates the complex conjugate), it follows for $A \in SL(2, C)$

$$\hat{A} = (A^\dagger)^{-1}. \tag{2.6}$$

The caret operation is an outer automorphism of $SL(2, C)$, which, at the same time, is an involution. Because

$$\hat{p}_\sim = [\Pi p]_\sim, \tag{2.7}$$

its application to (4) yields

$$\hat{A}[\Pi p]_\sim \hat{A}^\dagger = [\Pi \Lambda(A) p]_\sim,$$

i.e.,

$$\Pi \Lambda(A) \Pi = \Lambda(\hat{A}), \tag{2.8}$$

where Π denotes the space inversion. Note that the caret operation leaves the elements of $SU(2, C)$ (Group of the unitary unimodular matrices) invariant. Let $V^+(m)$ be the mass hyperboloid belonging to the mass m in the future cone. The relation (4) then shows that to each $p \in V^+(m)$ there exists a transformation $\alpha(p) \in SL(2, C)$ with the property

$$\begin{aligned} m\alpha(p)\alpha^\dagger(p) &= p_\sim \qquad (m > 0); \\ \tfrac{1}{2}\alpha(p)(\sigma_0 + \sigma_3)\alpha^\dagger(p) &= p_\sim \qquad (m = 0). \end{aligned} \tag{2.9}$$

The sets $Q(p)$ of the transformations $\alpha(p)$, which possess these properties form left cosets of the little

---

[6] E. Wigner, Ann. Math. **40**, No. 1 (1939).
[7] V. Bargmann, Gruppentheoretische Analyse der Lorentzinvarianz, lecture given at the Federal Technical Institute, Zurich, 1963.
[8] V. Bargmann and E. Wigner, Proc. Natl. Acad. Sci. U.S. **34**, (1948) 211.
[9] Fierz, Einführung in die Quantenfeldtheorie, lecture given at the Federal Technical Institute, Zurich 1963.
[10] A. S. Wightman in *Relation de dispersion et particules elementaires* (1960), p. 161.
[11] H. P. Stapp, Phys. Rev. **125**, 2139 (1962).
[12] K. Hepp, Helv. Phys. Acta **37**, 55 (1964).

or isotropy groups of the matrices $\mathbb{1}$ ($m > 0$) and $\frac{1}{2}(\sigma_0 + \sigma_3)$ ($m = 0$), respectively, which corresponds to the momentum $(m, \mathbf{0})$ and $(\frac{1}{2}, 0, 0, \frac{1}{2})$ in Minkowski space. The isotropy groups are

$$L(m) = \begin{cases} SU(2, C) \subset SL(2, C), & m > 0, \\ L(0) \subset SL(2, C), & m = 0, \end{cases} \tag{2.10}$$

where $L(0)$ means the group of transformations

$$\tau(a)\rho(\varphi), \tag{2.11}$$

$$\tau(a) = \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}, \tag{2.12}$$

$$\rho(\varphi) = \begin{pmatrix} e^{i\varphi/2} & 0 \\ 0 & e^{-i\varphi/2} \end{pmatrix}. \tag{2.13}$$

Specially important choices of the transformations $\alpha(p)$ are [we reserve parentheses for a general element in $Q(p)$, applying brackets and braces to the special elements quoted in the following]

$m > 0$

$$\alpha\{p\} = \left(\frac{p_\sim}{m}\right)^{\frac{1}{2}} = \frac{m + p_\sim}{[2m(m + \omega(\mathfrak{p}))]^{\frac{1}{2}}} \tag{2.14}$$

with

$$\omega(\mathfrak{p}) = +(m^2 + \mathfrak{p}^2)^{\frac{1}{2}},$$

and, for $m \geq 0$,

$$\alpha[p] = \rho[p]\, h[p]. \tag{2.15}$$

Here

$$h[p] = \begin{cases} \begin{pmatrix} [m^{-1}(\omega(\mathfrak{p}) + \mathfrak{p})]^{\frac{1}{2}} & 0 \\ 0 & [m^{-1}(\omega(\mathfrak{p}) - \mathfrak{p})]^{\frac{1}{2}} \end{pmatrix}, & m > 0, \\ \begin{pmatrix} (2\mathfrak{p})^{\frac{1}{2}} & 0 \\ 0 & (2\mathfrak{p})^{-\frac{1}{2}} \end{pmatrix}, & m = 0, \end{cases} \tag{2.16}$$

where $\mathfrak{p}$ is the Euclidian length of $\mathfrak{p}$,
and

$$\rho[p] = \begin{cases} e^{-i\varphi/2\sigma^3} e^{-i\vartheta/2\sigma^2} \text{ for } (0 < \vartheta < \pi; 0 \leq \varphi < 2\pi), \\ \mathbb{1} \quad \text{for} \quad \vartheta = 0, \\ i\sigma^1 \quad \text{for} \quad \vartheta = \pi. \end{cases} \tag{2.17}$$

φ, ϑ are defined by

$$\mathfrak{p} = \mathfrak{p}(\sin\vartheta\cos\varphi, \sin\vartheta\sin\varphi, \cos\vartheta). \tag{2.18}$$

In the case $m > 0$ one has

$$h[p] = \alpha\{p(3)\} \tag{2.19}$$

with $p(3) = (\omega(\mathfrak{p}), 0, 0, \mathfrak{p})$. In addition one derives from (14) for $v \in SU(2, C)$

$$v\alpha\{p\}v^\dagger = \alpha\{\Lambda(v)p\}. \tag{2.20}$$

From (19) and (20) we conclude that

$$\alpha[p] = \rho[p]\, h[p] = \alpha\{p\}\, \rho[p]. \tag{2.21}$$

3. SURVEY ABOUT THE PHYSICALLY MOST IMPORTANT REPRESENTATIONS OF THE INHOMOGENEOUS LORENTZ GROUP[1-4]

Let $A = \left(\begin{smallmatrix} \alpha & \beta \\ \gamma & \delta \end{smallmatrix}\right)$ be a matrix of $SL(2, C)$. It is well known that, by

$$\frac{(\alpha x + \gamma y)^{s+\lambda}(\beta x + \delta y)^{s-\lambda}}{[(s + \lambda)!\,(s - \lambda)!]^{\frac{1}{2}}} = \sum_{\lambda'} \frac{x^{s+\lambda'} y^{s-\lambda'}}{[(s + \lambda')!\,(s - \lambda')!]^{\frac{1}{2}}} \mathfrak{D}^s_{\lambda'\lambda}(A)_+, \tag{3.1}$$

where λ is one of the numbers of the set

$$I_+(s) = \{-s, -s + 1, \cdots 0, 1 \cdots s - 1, s\}$$

and the summation over λ' is to be carried out over the same set, the representation $(s, 0)$

$$A \to \mathfrak{D}^s(A)_+$$

and the representation $(0, s)$

$$A \to \mathfrak{D}^s(\hat{A})_+$$

of $SL(2, C)$ is defined. If $A \in SU(2, C)$ the two representations coincide and become unitary.

Let K be that subgroup of $SL(2, C)$, the elements of which are either of the form

$$\begin{pmatrix} \alpha & 0 \\ 0 & \alpha^{-1} \end{pmatrix}$$

or of the form

$$\begin{pmatrix} 0 & \gamma^{-1} \\ \gamma & 0 \end{pmatrix}.$$

Equation (1) then gives an irreducible representation of K if we restrict λ, λ' to the set of the two numbers

$$I_0(s) = \{-s, s\}.$$

We denote the matrix which represents the element $A \in K$ in this representation by

$$\mathfrak{D}^s(A)_0.$$

We find

$$\mathfrak{D}^s(A)_0 = \begin{cases} \begin{pmatrix} \alpha^{2s} & 0 \\ 0 & \alpha^{-2s} \end{pmatrix} & \text{for } A = \begin{pmatrix} \alpha & 0 \\ 0 & \alpha^{-1} \end{pmatrix}, \\ \begin{pmatrix} 0 & (-\gamma)^{-2s} \\ \gamma^{2s} & 0 \end{pmatrix} & \text{for } A = \begin{pmatrix} 0 & \gamma^{-1} \\ \gamma & 0 \end{pmatrix}. \end{cases} \tag{3.2}$$

We notice that the subgroup $K_U = K \cap SU(2, C)$ ($\cap$ means intersection) is thus represented unitarily. $\mathfrak{D}^s(A)_\tau (\tau = 0, +1)[A \in SL(2, C)$ for $\tau = +1$,

$A \in K$ for $\tau = 0$] defined in (1), and (2) is a matrix operating in a $N_\tau(s)$-dimensional vector space, where

$$N_\tau(s) = \begin{cases} 2s+1, & \tau = +1, \\ 2, & \tau = 0. \end{cases} \tag{3.3}$$

Let $\mathfrak{H}(m, s)$ be the Hilbert space of the matrices with one column and $N_\tau(s)$ rows, the elements of which are square-integrable functions defined on the mass-hyberboloid $V^+(m)$. We define the scalar product in $\mathfrak{H}(m, s)$ by

$$(f, g) = \int_{V^+(m)} \frac{d^3p}{2\omega(\mathfrak{p})} f^\dagger(p) g(p) \quad f, g \in \mathfrak{H}(m, s), \tag{3.4}$$

where the integration extends over $V^+(m)$ and the symbol † indicates the change to the adjoint (transposed and complex-conjugate) matrix.

The unitary irreducible representations of the orthochronous quantum mechanical Poincaré group $\tilde{P}^\uparrow$, the universal covering group of the orthochronous inhomogeneous Lorentz-group $P^\uparrow$ is defined in $\mathfrak{H}(m, s)$ by

$$(U(a)f)(p) = e^{ipa} f(p), \tag{3.5a}$$

$$(U(A)f)(p) = \mathfrak{D}^s(\rho(p, A))_\tau f(\Lambda^{-1}(A)p), \tag{3.5b}$$

$$(U(\Pi)f)(p) = \eta_\Pi \, \mathfrak{D}^s(\rho_\Pi(p))_\tau f(\Pi p). \tag{3.5c}$$

The unimodular factor η_Π is the parity. $\mathfrak{D}^s(\ )_\tau$ is defined in (1) and (2). [In future we drop the index τ, keeping in mind that, in the case $m > 0$ ($m = 0$), we have to take the matrix $\mathfrak{D}^s(\ )_+$, $\mathfrak{D}^s_0(\ )$, respectively.]

The matrices ρ are defined as follows

$$\rho(p, A) = \alpha^{-1}(p) A \alpha(\Lambda^{-1}(A)\rho), \tag{3.6}$$

$$\rho_\Pi(p) = \alpha^\dagger(p) \alpha(\Pi p), \tag{3.7}$$

with arbitrary $\alpha(p) \in Q(p)$. [The physical significance of our choice of $\alpha(p)$ is seen below.] $\rho_\Pi(p)$ has the properties

$$\epsilon^{-1} \rho_\Pi(p) \in L(m), \qquad \rho_\Pi(p) \rho_\Pi(\Pi p) = 1. \tag{3.8}$$

Note that, for $m = 0$, $\rho_\Pi(p) \notin L(0)$.

From (5c) one derives

$$U^2(\Pi) = \eta_\Pi^2. \tag{3.9}$$

These formulas are valid for cases $m > 0$ and $m = 0$, if one determines any matrix belonging to the group $L(0)$ modulus the invariant subgroup of "the translations" $\tau(a)$ as element of the group $DU(2, C)$ which is defined as the group of the matrices $\rho(\varphi)$ given in (2.13). [It is well known that only those representations of $L(0)$ have physical significance, in which the invariant subgroup of the "translations" is represented by the unity.]

By doing this, $\rho(p, A)$ as well as $\rho_\Pi(p)$ in the case $m = 0$ become elements of K_U, so that $\mathfrak{D}^s(\ )_0$ is defined for both of them. For $\alpha(p) \to \alpha[p]$, we get

$$\epsilon^{-1}\rho_\Pi(p) = \begin{cases} i\sigma^3 \text{ for } \begin{cases} 0 \le \varphi < \pi \\ 0 < \vartheta < \pi \end{cases} \text{ and } \vartheta = 0, \\ -i\sigma^3 \text{ for } \begin{cases} \pi \le \varphi < 2\pi \\ 0 < \vartheta < \pi \end{cases} \text{ and } \vartheta = \pi. \end{cases} \tag{3.10}$$

For $\alpha(p) \to \alpha\{p\}$, we get

$$\rho_\Pi\{p\} = 1. \tag{3.11}$$

If we confine ourselves to the restricted group $\tilde{P}^\uparrow_+$, the space $\mathfrak{H}(m, s)$ remains irreducible in the case $m > 0$, decomposing into two subspaces with the elements

$$\begin{pmatrix} f_{+s}(p) \\ 0 \end{pmatrix}, \qquad \begin{pmatrix} 0 \\ f_{-s}(p) \end{pmatrix},$$

respectively, in the case $m = 0$.

Finally, we define the time-inversion antiunitary by

$$\begin{aligned} (V(T)f)(p) &= \mathfrak{D}^s(\epsilon)(\overline{U(\Pi)f})(p) \cdot \eta \\ &= \mathfrak{D}^s(\rho_\Pi(p)\epsilon)\bar{f}(\Pi p) \cdot \eta_T \end{aligned} \tag{3.12}$$

with

$$\eta = \frac{\eta_T}{\eta_\Pi} \qquad |\eta|^2 = |\eta_T|^2 = 1. \tag{3.13}$$

Note that

$$\mathfrak{D}^s_{\lambda'\lambda}(\epsilon) = (-1)^{s-\lambda} \delta_{\lambda', -\lambda} \tag{3.14}$$

so that it follows

$$V^2(T) = (-1)^{2s}. \tag{3.15}$$

Defining $V(T)$ by (12), it is not necessary to enlarge the space to represent the time inversion. We do not consider representations with different values of $U^2(\Pi)$ and $V^2(T)$. These are found, for example, in Ref. 7.

If we define $V(\Pi T)$ by

$$V(\Pi T) = U(\Pi) V(T) = \frac{\eta_\Pi}{\bar{\eta}_\Pi} V(T) U(\Pi), \tag{3.16}$$

we get

$$V^2(\Pi T) = (-1)^{2s} \cdot \frac{\eta_\Pi}{\bar{\eta}_\Pi}. \tag{3.17}$$

The freedom in the choice of $\alpha(p) \in Q(p)$ is the mathematical expression of the fact that the spin vector of a particle in a coordinate system, in which

the momentum p is measured is only determined up to a transformation with an element out of the isotropy group of p. If $\mathfrak{J}$ is the vector of the total angular momentum in our representation (vector of the Hermitian represented infinitesimal rotations), we get for the spin operator in a coordinate system in which the particle has the momentum p,

$$\begin{aligned} S_k(p) &= U(\alpha(p))W_k(p(m))U^{-1}(\alpha(p)) \\ &= W_r(p)\Lambda^r_k(\alpha^{-1}(p)). \end{aligned} \tag{3.18}$$

Here

$$\begin{aligned} p(m) &= (m, \mathbf{0}) \quad \text{for} \quad m > 0, \\ p(0) &= (\tfrac{1}{2}, 0, 0, \tfrac{1}{2}), \\ W_r(p) &= [p^0(m)]^{-1}(\mathfrak{p}\cdot\mathfrak{J}, p^0\mathfrak{J} + \mathfrak{p}\times\mathfrak{N}), \end{aligned} \tag{3.19}$$

and $\mathfrak{N}$ is the vector of the Hermitian Operators representing the infinitesimal special Lorentz transformations in the three directions of our coordinate-system.

For $\alpha(p) \to \alpha[p]$ we get

$$\begin{aligned} S_3[p] &= U(\alpha[p])J_3U^{-1}(\alpha[p]) \\ &= U^{-1}(\rho[p])J_3U^{-1}(\rho[p]) = \mathfrak{J}\cdot\mathfrak{p}/\mathfrak{p}, \end{aligned} \tag{3.20}$$

i.e., the projection of the total angular momentum on the direction of the momentum, the so-called helicity.

For $\alpha(p) \to \alpha\{p\}$ it follows

$$\begin{aligned} S_3\{p\} &= \frac{1}{m}\left[-(\mathfrak{J}\cdot\mathfrak{p})\frac{p_3}{m+\omega(\mathfrak{p})}\right. \\ &\quad \left. + p_0J_3 + p_1N_2 - p_2N_1\right] \\ &= J_3 - \frac{1}{m}\left(N_1 - \frac{p_2J_3 - p_3J_2}{m+\omega(\mathfrak{p})}\right)p_2 \\ &\quad + \frac{1}{m}\left(N_2 - \frac{p_3J_1 - p_1J_3}{m+\omega(\mathfrak{p})}\right)p_1 \\ &= J_3 - (\mathfrak{X}\times\mathfrak{p})_3, \end{aligned} \tag{3.21}$$

where

$$\mathfrak{X} = \frac{\mathfrak{N}}{m} - \frac{\mathfrak{p}\times\mathfrak{J}}{m(m+\omega(\mathfrak{p}))} + \mathfrak{p}f(\mathfrak{N}, \mathfrak{p}) \tag{3.22}$$

with suitable chosen $f(\mathfrak{N},\mathfrak{p})$ turns out to be the (Newton–Wigner) position operator.

Therefore, this choice of the transformation $\alpha(p)$ is equivalent to the definition of the spin (in a coordinate system in which the particle has momentum p) as difference between the total and the angular momentum.

Concluding this section, we mention that, in the case $m > 0$, the transformation

$$f'(p) = \mathfrak{D}^s(\alpha(p))f(p) \tag{3.23}$$

represents an isomorphism of the Hilbert space $\mathfrak{H}(m, s)$ and a Hilbert space equipped with the scalar product

$$(f', g') = \int \frac{d^3p}{2\omega(\mathfrak{p})} f'^{\dagger}(p)\ \mathfrak{D}^s\left(\frac{\hat{p}}{m}\right)g'(p) = (f, g). \tag{3.24}$$

In this Hilbert space the unitary representation of $\tilde{P}^{\uparrow}_{+}$ takes the well-known "$(s, 0)$-Spinorfield-form"

$$\begin{aligned} (U(a)f')(p) &= e^{ipa}f'(p), \\ (U(A)f')(p) &= D^s(A)f'(\Lambda^{-1}(A)p), \\ (U(\Pi)f')(p) &= \mathfrak{D}^s\left(\frac{p}{m}\right)f'(\Pi p)\cdot\eta_{\Pi}, \\ (V(T)f')(p) &= \mathfrak{D}^s(\epsilon)\bar{f}'(\Pi p)\cdot\eta_T. \end{aligned} \tag{3.25}$$

Equations (23)–(25) remain valid, of course, if a caret is put over each matrix occuring in $\mathfrak{D}^s(\ )$ leading to the "$(0, s)$-Spinorfield form" of the unitary representation of $\tilde{P}^{\uparrow}_{+}$. The space inversion maps the so-defined representation spaces of $\tilde{P}^{\uparrow}_{+}$ onto each other.

4. PREPARATIONS FOR THE DECOMPOSITION OF THE GENERAL PRODUCT REPRESENTATION OF $\tilde{P}_{+}{}^{\uparrow}$

Let us first consider the set $\mathfrak{P}_n(\mathbf{m})(\mathbf{m} = (m_1, \cdots m_n)$ $|\mathbf{m}| = \sum_{i=1}^{n} m_i > 0)$ of the oriented closed polygons in three-dimensional space, which have n edges labeled with "weights" $m_1 \cdots m_n$. Strictly given the set of "weights" $m_1 \cdots m_n$ such a polygon is still characterized by n vectors $\mathfrak{k}_1 \cdots \mathfrak{k}_n$ which satisfy the subsidiary condition

$$\sum_{i=1}^{n} \mathfrak{k}_i = 0. \tag{4.1}$$

We define the total weight of a polygon by

$$M = \sum_{i=1}^{n} (\mathfrak{k}_i^2 + m_i)^{\frac{1}{2}}. \tag{4.2}$$

Instead of characterizing the polygon by $\mathbf{m}$ and $(\mathfrak{k}_1 \cdots \mathfrak{k}_n)$, we may also characterize it by n four-vectors

$$\mathbf{k} = (k_1 \cdots k_n), \qquad k_i = [(\mathfrak{k}_i^2 + m_i^2)^{\frac{1}{2}}, \mathfrak{k}_i], \tag{4.3}$$

which fulfill the subsidiary condition

$$\sum_{i=1}^{n} k_i = (M, \mathbf{0}) \tag{4.4}$$

or applying the mapping (2.1) by n Hermitian matrices

$$\mathbf{k}_{\sim} = (k_{\sim 1} \cdots k_{\sim n}) \tag{4.5}$$

satisfying the subsidiary condition

$$\sum_{i=1}^{n} k_{\sim i} = M \cdot \mathbb{1}. \tag{4.6}$$

Consequently, the weights of the edges are then given by the roots of the determinants of these matrices.

We call equivalent two of the elements of $\mathcal{P}_n(\mathfrak{m})$ if and only if there exists a rotation $R_{\varphi\vartheta\chi}(\varphi, \vartheta, \chi =$ Euler angles) which transforms one of the elements into the other. Hence $\mathcal{P}_n(\mathfrak{m})$ decomposes into equivalence classes. Let us imagine that a system of representatives of these classes is chosen. Assume that it is given by polygons characterized by the vectors

$$\mathfrak{b}_i = \mathfrak{b}_i(\kappa) \quad (i = 1 \cdots n), \qquad \sum_{i=1}^{n} \mathfrak{b}_i = 0, \tag{4.7}$$

where $\kappa = (\kappa_1 \cdots \kappa_{3n-6})$ are $(3n - 6)$ inner (i.e., rotation-invariant) parameters which parametrize the system of representatives. We may choose as κ_1 (the square of) the total weight M. In general, M is a function of κ:

$$M = M(\kappa). \tag{4.8}$$

A possible choice of the parameters is

$$\mathfrak{b}_1\mathfrak{b}_2 \cdots \mathfrak{b}_{n-1}, \qquad \vartheta_2, \vartheta_3\varphi_3 \cdots \vartheta_{n-1}\varphi_{n-1}, \tag{4.9}$$

where φ_k, ϑ_k are the polar angles of $\mathfrak{b}_k$ relative to $\mathfrak{b}_1(\varphi$ is measured starting from the plane put up by $\mathfrak{b}_1$ and $\mathfrak{b}_2$; $\mathfrak{b}_i$ denotes the length of $\mathfrak{b}_i$).

It is possible to choose the Lorentz-invariant parametrization

$$\begin{gathered}(b, b_1) \cdots (b, b_{n-1}), \quad (b_1, b_2) \cdots (b_1, b_{n-1}), \\ [b_1, b_2, b_3, b_4] \cdots [b_{n-3}, b_{n-2}, b_{n-1}, b_n],\end{gathered} \tag{4.10}$$

where

$$b_k = ([m_k^2 + \mathfrak{b}_k^2]^{\frac{1}{2}}, \mathfrak{b}_k), \qquad b = \sum_{i=1}^{n} b_i,$$

(b_1, b_2) is the Minkowski scalar product of the vectors b_1 and b_2, and

$$[b_1, b_2, b_3, b_4] = \epsilon^{\nu\mu\tau\sigma} b_{1\nu} b_{2\mu} b_{3\tau} b_{4\sigma} \tag{4.11}$$

stands for the volume in Minkowski space defined by the vectors $b_1 \cdots b_4$.

In future, we use a general parametrization if we do not explicitly refer to a special one. The domain in κ-space, the points of which correspond to real polygons, we denote by C_n, and the submanifold of C_n defined by (4.8) we denote by $C_n(M)$.

In the case $n = 2$, C_n is the positive real axis. One has

$$\mathfrak{b}_1 = -\mathfrak{b}_2 = \mathfrak{b}$$

and one may choose as inner parameter $\mathfrak{b}$ or M. The two quantities are connected by

$$M = (\mathfrak{b}^2 + m_1^2)^{\frac{1}{2}} + (\mathfrak{b}^2 + m_2^2)^{\frac{1}{2}}, \tag{4.12}$$

i.e.,

$$\mathfrak{b} = \Delta^{\frac{1}{2}}(m_1, m_2, M), \tag{4.13}$$

where

$$\begin{aligned}\Delta(m_1, m_2, M) = (1/4M^2)(&m_1^4 + m_2^4 + M^4 \\ &- 2m_1^2M^2 - 2m_2^2M^2 - 2m_1^2m_2^2).\end{aligned} \tag{4.14}$$

It is no loss of generality to assume that $\mathfrak{b}_1$ points in the direction of the 3-axis in three-dimensional space, so that

$$\begin{aligned}b_{\sim 1} &= \begin{pmatrix} \omega_1(\mathfrak{b}) + \mathfrak{b} & 0 \\ 0 & \omega_1(\mathfrak{b}) - \mathfrak{b} \end{pmatrix}, \\ b_{\sim 2} &= \begin{pmatrix} \omega_2(\mathfrak{b}) - \mathfrak{b} & 0 \\ 0 & \omega_2(\mathfrak{b}) + \mathfrak{b} \end{pmatrix}.\end{aligned} \tag{4.15}$$

We now want to parametrize one of the equivalence-classes in $\mathcal{P}_n(\mathfrak{m})$. It is evident that, in the case $n \geq 3$, the three-dimensional rotation group (and in the case $n = 2$, the two-dimensional sphere) yields a parametrization. However, the parametrization adapted to our purposes is the following. We consider the isotropy-group $G_n \subset SL(2, C)$ of the n-vector of Hermitian two-by-two matrices

$$\mathfrak{b}_{\sim} = (b_{\sim 1} \cdots b_{\sim n}).$$

We find

$$G_n = \begin{cases} DU(2, C), & n = 2, \\ \text{cyclic group of the} \\ \text{two elements } (\mathbb{1}, -\mathbb{1}), & n \geq 3. \end{cases} \tag{4.16}$$

Here we have assumed that $b_{\sim 1}$, $b_{\sim 2}$ (in the case $n = 2$) have the form (4.15) so that $DU(2, C)$ is, as above, the group of the matrices $\rho(\varphi)$ defined in (2.13). We denote by

$$SU(2, C)/G_n \tag{4.17}$$

the manifold of the left cosets of the group G_n with respect to the group $SU(2, C)$. In the case $n \geq 3$, G_n is an invariant subgroup; hence (4.17) is a group (isomorphic to the three-dimensional rotation group).

A parametrization of the manifold (4.17) is now

seen to be equivalent to one of an equivalence class of $\mathcal{P}_n(\mathbf{m})$, i.e., if $\{\rho\}$ is a system of representatives of the left cosets, $\{\rho\}$ yields a parametrization of an equivalence class of $\mathcal{P}_n(\mathbf{m})$. Let $\mathbf{p} = (p_1, \cdots p_n)$ be an element of the product set

$$V_n^+(\mathbf{m}) = V^+(m_1) \otimes V^+(m_2) \otimes \cdots \otimes V^+(m_n), \tag{4.18}$$

and let q denote

$$q = \sum_{i=1}^{n} p_i. \tag{4.19}$$

Then we have

$$q^2 \geq |\mathbf{m}|^2 > 0.$$

Further we assume

$$\alpha(q) \in Q(q).$$

By the relation

$$\alpha^{-1}(q)\mathbf{p}_\sim(\alpha^\dagger(q))^{-1} = \mathbf{k}_\sim,$$

we attach to each $\mathbf{p} \in V^+(\mathbf{m})$ a polygon and therefore a representative $\rho \in \{\rho\}$ and a point $\kappa \in C_n([q^2]^{\frac{1}{2}})$. On the other hand, if $\kappa \in C_n([q^2]^{\frac{1}{2}})$ and $\rho \in SU(2, C)/G_n$, $q^2 \geq |\mathbf{m}|^2$, and $\mathbf{b}(\kappa)$ is a representative of a class in $\mathcal{P}_n(\mathbf{m})$, we get a vector $\mathbf{p} \in V^+(\mathbf{m})$ by the relation

$$\mathbf{p}_\sim = \alpha(q)\rho\mathbf{b}_\sim(\kappa)\rho^\dagger\alpha^\dagger(q). \tag{4.20}$$

In other words, there exists a parametrization of the manifold $V^+(\mathbf{m})$ by the one-to-one correspondence

$$\mathbf{p} \leftrightarrow q, \rho, \kappa, \tag{4.21}$$

where $q^2 \geq |\mathbf{m}|^2 > 0$, $\rho \in SU(2, C)/G_n$, $\kappa \in C_n([q^2]^{\frac{1}{2}})$. By $\kappa \in C_n([q^2]^{\frac{1}{2}})$, a equivalence class of polygons in $\mathcal{P}_n(\mathbf{m})$ with total weight $[q^2]^{\frac{1}{2}}$ is fixed. We call it the equivalence-class κ. We consider now equivalence-classes, which are related to the given one:

(1) The equivalence class of the mirror polygons, which we denote by $\Pi\kappa$. (Note that in the cases $n = 2, 3$: $\Pi\kappa = \kappa$.)

(2) Let γ_n' be the subgroup of the symmetric group γ_n of n elements which contains all the elements permutating those edges of a polygon in $\mathcal{P}_n(\mathbf{m})$ which have the same "weights", so that $\mathcal{P}_n(\mathbf{m})$ is mapped onto itself by each $\pi \in \gamma_n'$. The permuted equivalence-class κ we denote by $\pi\kappa$.

Because the polygon characterized by $\Pi\mathbf{b}(\kappa)$ must lie in the equivalence class $\Pi\kappa$, and, similarly, the polygon characterized by $\pi\mathbf{b}(\kappa)$ in the equivalence class $\pi\kappa$, there exists transformations $v_\Pi(\kappa)$, $v_\pi(\kappa) \in SU(2, C)$ with the property

$$\mathbf{b}_\sim(\Pi\kappa) = v_\Pi(\kappa)[\Pi\mathbf{b}(\kappa)]_\sim v_\Pi^\dagger(\kappa) \tag{4.22}$$

and

$$\mathbf{b}_\sim(\pi\kappa) = v_\pi(\kappa)[\pi\mathbf{b}(\kappa)]_\sim v_\pi^\dagger(\kappa), \tag{4.23}$$

respectively. They are only determined up to an element in G_n. Furthermore, because of

$$\Pi(\Pi\kappa) = \kappa, \quad \pi_1(\pi_2)\kappa = (\pi_1\pi_2)\kappa, \tag{4.24}$$

we have

$$v_\Pi(\kappa)\, v_\Pi(\Pi\kappa) = \rho(\chi_\Pi) \in G_n, \tag{4.25}$$

i.e.,

$$\begin{aligned} &0 \leq \chi_\Pi < 4\pi \quad \text{for} \quad n = 2, \\ &\chi_\Pi = 0, 2\pi \qquad \text{for} \quad n \geq 3, \end{aligned} \tag{4.26}$$

and

$$v_{\pi_2}(\pi_1\kappa)\, v_{\pi_1}(\kappa) = \rho(\chi_{\pi_2\pi_1})\, v_{\pi_2\pi_1}(\kappa), \tag{4.27}$$

where

$$\rho(\chi_{\pi_2\pi_1}) \in G_n. \tag{4.28}$$

Before continuing the general discussion, we give the most convenient choice of $\mathbf{b}(\kappa)$, $v_\pi(\kappa)$ and $v_\Pi(\kappa)$. For $n = 2, 3$, we have $\mathbf{b}(\Pi\kappa) = \mathbf{b}(\kappa)$; hence we may choose $\mathbf{b}$ to be real and

$$v_\pi(\kappa) = \epsilon \quad \text{for all} \quad \kappa \in C_n. \tag{4.29}$$

For $n \geq 4$, we could in principle choose $v_\Pi(\kappa) = 1$ by choosing $\Pi\mathbf{b}(\kappa)$ as representative of the class $\Pi\kappa$. However, we shall see that the choice (29) is the most convenient one for all n*. It corresponds to the choice $Y\mathbf{b}(\kappa)$ for the class $\Pi\kappa$, where Y denotes the reflection about the xz-plane.

For $n = 2$, one has

$$\mathbf{b}(\pi\kappa) = \mathbf{b}(\kappa), \tag{4.30}$$

where $\mathbf{b}(\kappa)$ is given by (15), so that we may set

$$v_\pi(\kappa) = \epsilon. \tag{4.31}$$

If $n \geq 3$, and if we disregard all cases where either

$$\begin{aligned} &\Pi\kappa = \kappa \\ \text{or} \quad &\\ &\pi\kappa = \Pi\kappa, \end{aligned} \tag{4.32}$$

by recognizing that the set of the points $\kappa \in C_n$ with one of the properties (32) has zero (Lebesgue) measure in κ-space (we shall see the reason why we may neglect sets of zero measure in κ-space later on), we may choose

$$v_\pi(\kappa) = \mathbb{1}. \tag{4.33}$$

* At least if one considers space and time inversion simultaneously.

We now claim that, simultaneously with the correspondence (4.21), one has

$$\Lambda^{-1}(A)\mathbf{p} \leftrightarrow \Lambda^{-1}(A)q,\ \rho^{-1}(q,\ A)\rho,\kappa, \tag{4.34a}$$

$$\Pi\mathbf{p} \leftrightarrow \Pi q,\ \rho_\Pi^{-1}(q)\rho v^{-1}(\kappa),\ \Pi\kappa, \tag{4.34b}$$

$$\pi\mathbf{p} \leftrightarrow q,\ \rho v_\pi^{-1}(\kappa),\ \pi\kappa. \tag{4.34c}$$

We show the validity of these correspondences in the case of the space-inversion and leave the proof of the other ones to the reader.

Application of the caret operation to Eq. (4.20) yields

$$\hat{\mathbf{p}}_\sim = \hat{a}(q)\rho\hat{\mathbf{b}}_\sim(\kappa)\rho^\dagger\hat{a}^\dagger(q),$$

i.e.,

$$\begin{aligned}[\Pi\mathbf{p}]_\sim &= \hat{a}(q)\rho v_\Pi^\dagger(\kappa)\mathbf{b}_\sim(\Pi\kappa)v_\Pi(\kappa)\hat{a}^\dagger(q)\\ &= \alpha(\Pi q)\rho_\Pi^{-1}(q)\rho v_\Pi^{-1}(\kappa)\mathbf{b}_\sim(\Pi q)v_\Pi(\kappa)\rho_\Pi(q)\alpha^\dagger(\Pi q).\end{aligned}$$

The last equation shows the statement (4.34b).

We now define the transformation

$$a(\mathbf{p}) = \alpha(q)\rho \quad \text{if} \quad \mathbf{p} \leftrightarrow q,\ \rho,\kappa. \tag{4.35}$$

With the help of (34), one easily derives

$$\begin{aligned}a(\Lambda^{-1}(A)\mathbf{p}) &= A^{-1}a(\mathbf{p}),\\ a(\Pi\mathbf{p}) &= \hat{a}(\mathbf{p})v_\Pi^{-1}(\kappa),\\ a(\pi\mathbf{p}) &= a(\mathbf{p})v_\pi^{-1}(\kappa).\end{aligned} \tag{4.36}$$

Now, let $u(b_l)$ be an element of $L(m_l)$ which may depend on b_l [$L(m)$ was defined in (2.10)], and let

$$\alpha_l(p_l) \in Q(p_l)$$

[$Q(p)$ is defined after Eq. (2.9)]. Then we define

$$\alpha_l^u(b_l) = \alpha_l(b_l)u(b_l) \in Q(b_l) \tag{4.37}$$

and

$$\begin{aligned}r_l(\mathbf{p}) &= u^{-1}(b_l)\rho_l(a^{-1}(\mathbf{p}),\ \Lambda^{-1}[a(\mathbf{p})]p_l)[\rightarrow(3.6)]\\ &= [\alpha_l^u(b_l)]^{-1}a^{-1}(\mathbf{p})\alpha_l(p_l) \in L(m_l).\end{aligned} \tag{4.38}$$

[As always, we have to determine the transformations of $L(0)$ as diagonal matrices belonging to $DU(2, C)$, calculating modulus "the translations."]

By a change of the representative ρ of the left coset of G_n,

$$\rho \rightarrow \rho\rho(\chi)\ \begin{cases}0 < \chi < 4\pi, & n = 2,\\ \chi = 2\pi, & n \geq 3,\end{cases}$$

one is led to the substitutions

$$\begin{aligned}a(\mathbf{p}) &\rightarrow a(\mathbf{p})\rho(\chi),\\ r_l(\mathbf{p}) &\rightarrow \begin{cases}y(b_l, \chi)r_l(\mathbf{p}), & n = 2,\\ -r_l(\mathbf{p}), & n \geq 3,\end{cases}\end{aligned} \tag{4.39}$$

where

$$y(b_l, \chi) = [\alpha_l^u(b_l)]^{-1}\rho(-\chi)\alpha_l^u(b_l) = \hat{y}(b_l, \chi). \tag{4.40}$$

Furthermore, one easily derives

$$r_l(\Lambda^{-1}(A)\mathbf{p}) = r_l(\mathbf{p})\rho_l(p_l,\ A), \tag{4.41a}$$

$$r_l(\Pi\mathbf{p}) = w_\Pi^l(\kappa)r_l(\mathbf{p})\rho_\Pi^l(p_l), \tag{4.41b}$$

$$r_l(\pi\mathbf{p}) = w_\pi^l(\kappa)r_{\pi l}(\mathbf{p}), \quad \pi \in \gamma_n', \tag{4.41c}$$

where

$$\begin{aligned}w_\Pi^l(\kappa) &= [\alpha_l^u(b_l^\tau)]^{-1}v_\Pi(\kappa)\hat{\alpha}_l^u(b_l), \quad b_l^\Pi \equiv b_l(\Pi\kappa),\\ w_\pi^l(\kappa) &= [\alpha_l^u(b_l^\tau)]^{-1}v_\pi(\kappa)\alpha_l^u(b_{\pi l}), \quad b_l^\pi \equiv b_l(\pi\kappa).\end{aligned} \tag{4.42}$$

We give the proofs of (41b,c).

$$\begin{aligned}r_l(\Pi\mathbf{p}) &= [\alpha_l^u(b_l^\Pi)]^{-1}v_\Pi(\kappa)a^{-1}(\mathbf{p})\alpha_l(\Pi p_l)\\ &= [\alpha_l^u(b_l^\Pi)]^{-1}v_\Pi(\kappa)\hat{\alpha}_l^u(b_l)\hat{r}_l(\Pi\mathbf{p})\rho_\Pi^l(p_l);\\ r_l(\pi\mathbf{p}) &= [\alpha_l^u(b_l^\tau)]^{-1}v_\pi(\kappa)a^{-1}(\mathbf{p})\alpha_l(p_{\pi l})\\ &= [\alpha_l^u(b_l^\tau)]^{-1}v_\pi(\kappa)\alpha_l^u(b_{\pi l})r_{\pi l}(\mathbf{p})\end{aligned}$$

if and only if

$$\alpha_l(p) = \alpha_{\pi l}(p) \quad \text{for} \quad \pi \in \gamma_n'. \tag{4.43}$$

The validity of Eq. (4.41c) is founded on the assumption (4.43). Note that, in the case $m_l = 0$, $w_\Pi^l(\kappa) \notin L(0)$, but $w_\Pi^l(\kappa)\epsilon \in L(0)$ because of Eq. (3.8), so that $w_\Pi^l(\kappa)\epsilon$ is to be determined as an element of $DU(2, C)$.

We now remark that, for arbitrary fixed $\alpha_l(p_l) \in Q(p_l)$, we may choose $u(b_l) \in L(m)_l$ always in such a way that, for all l,

$$\alpha_l^u(b_l) = \alpha[b_l], \tag{4.44}$$

where $\alpha[p]$ is the transformation defined in Eq. (2.15). With this choice of $u(b_l)$ and the choices of v_Π, v_π as quoted in Eqs. (4.29), (4.31), and (4.33), we get

$$\begin{aligned}w_\Pi^l(\kappa) &= \alpha[Yb_l]^{-1}\epsilon\hat{\alpha}[b_l] = \pm\bar{\alpha}[b_l]^{-1}\epsilon\hat{\alpha}[b_l]\\ &= \pm\epsilon\begin{cases}-\text{sign, if } \mathfrak{b}_l \text{ points in the direction of the}\\ \quad\text{negative 3-axis,}\\ +\text{sign otherwise.}\end{cases}\end{aligned}$$

With the restrictions, we have already put on the choice of b_l in the case $n \geq 3$, it is always possible to choose these vectors in such a way that no one points into the direction of the negative 3-axis. However, this is not possible in the case $n = 2$ in which we have chosen $\mathfrak{b}_1$ pointing in the direction of the positive 3-axis and $\mathfrak{b}_2 = -\mathfrak{b}_1$. That is why we get

$$w_\Pi^l(\kappa) = \begin{cases}-\epsilon & n = 2, \quad l = 2,\\ +\epsilon & \text{otherwise.}\end{cases} \tag{4.45}$$

Furthermore, we get in the case $n = 2$

$$w_\pi^l(\kappa) = \alpha[b_l]^{-1}\epsilon\alpha[\Pi b_l] = \alpha[b_l]^{-1}\epsilon\bar{\alpha}[b_l]\rho_\Pi[b_l] = \alpha^{-1}[b_l]\bar{\alpha}[b_l]\epsilon\rho_\Pi[b_l].$$

According to Eqs. (2.16), (2.17), and (3.10), it follows that

$$w_\pi^l(\kappa) = -i\sigma_3, \qquad l = 1, 2, \tag{4.46}$$

and (4.40) yields

$$\begin{aligned} y(b_1, \chi) &= \alpha^{-1}[b_1]\rho(-\chi)\alpha[b_1] = \rho(-\chi), \\ y(b_2, \chi) &= (i\sigma_1)^{-1}\rho(-\chi)i\sigma_1 = \rho(\chi). \end{aligned} \tag{4.47}$$

Finally, we get in the case $n \geq 3$ with the choice (33),

$$w_\pi^l(\kappa) = \mathbb{1}. \tag{4.48}$$

This choice given by (4.29), (4.31), (4.33), and (4.44) corresponds to the construction of angular momentum states by Wick, Jacob, and Wehrle. We refer to it as the WJW choice.

We have not yet finished the necessary preparations for the decomposition of the direct product

$$\mathfrak{H}_n(\mathbf{m}, \mathbf{s}) = \mathfrak{H}(m_1, s_1) \otimes \mathfrak{H}(m_2, s_2) \otimes \cdots \otimes \mathfrak{H}(m_n, s_n), \tag{4.49}$$

$\mathbf{m} = (m_1, \cdots m_n)\mathbf{s} = (s_1 \cdots s_n)$. We have to study the change of the relativistic invariant volume element on $V^+(\mathbf{m})$ if we introduce the parameters standing on the right side of (4.21). We start from the invariant element of volume in p-space and write

$$d^{4n}p = d^{4(n-1)}p\, d^4q = d^{4(n-1)}k\, d^4q; \qquad q = \sum_{k=1}^{n} p_k,$$

where $\mathbf{k} = \Lambda^{-1}(\alpha(p))\mathbf{p}$. From

$$p^0\, dp^0 = m\, dm,$$

it follows

$$\frac{d^{3n}p}{\omega_1(\mathfrak{p}_1)\cdots\omega_n(\mathfrak{p}_n)}\, m_n\, dm_n, = \frac{d^{3(n-1)}k}{\omega_1(\mathfrak{b}_1)\cdots\omega_{n-1}(\mathfrak{b}_{n-1})}\, M\, \frac{d^3q}{\omega(\mathfrak{q})}\, dM,$$

where

$$\omega(\mathfrak{p}) = +[m^2 + \mathfrak{p}^2]^{\frac{1}{2}},$$

or, because of

$$M = \sum_{i=1}^{n} [m_i^2 + \mathfrak{b}_i^2]^{\frac{1}{2}},$$

we get

$$\frac{d^{3n}p}{\omega_1(\mathfrak{p}_1)\cdots\omega_n(\mathfrak{p}_n)} = \frac{d^{3(n-1)}k}{\omega_1(\mathfrak{b}_1)\cdots\omega_{n-1}(\mathfrak{b}_{n-1})\omega_n(\mathfrak{b}_n)}\, M\, \frac{d^3q}{\omega(\mathfrak{q})}.$$

The invariant volume element of the unitary group $SU(2, C)$ is

$$d\rho = [16\pi^2]^{-1}\, d\cos\vartheta\, d\varphi\, d\chi. \tag{4.50}$$

We have normed it in such a way that the volume of $SU(2, C)$ is one. It coincides (up to a factor) with the invariant volume element of the three-dimensional rotation group. We therefore get ($n \geq 3$)

$$\frac{1}{2^n}\cdot\frac{d^{3n}p}{\omega_1(\mathfrak{p}_1)\cdots\omega_n(\mathfrak{p}_n)} = A_n(\kappa)\, d\rho\, \frac{d^3q}{2\omega(\mathfrak{q})}\, d^{3n-6}\kappa, \tag{4.51}$$

where $\omega(\mathfrak{q}) = [\mathfrak{q}^2 + M^2(\kappa)]^{\frac{1}{2}}$ also depends on κ and $A_n(\kappa)$ stands for the expression

$$A_n(\kappa) = \frac{1}{2^{n-1}}\frac{\partial^{3(n-1)}k}{\partial\rho\,\partial\kappa}\cdot\frac{M(\kappa)}{\omega_1(\mathfrak{b}_1)\cdots\omega_n(\mathfrak{b}_n)}\cdot \tag{4.52}$$

Note that $A_n(\kappa)$ is not dependent on ρ because $d^{3(n-1)}k$ is rotation-invariant, i.e., it is constant on a equivalence class of $\mathfrak{O}_n(\mathbf{m})$. If we now choose that parametrization of the equivalence classes of $\mathfrak{O}_n(\mathbf{m})$ which is given in (9) and which we denote by κ_0, we find

$$d^{3(n-1)}k = 16\pi^2 \prod_{i=1}^{n-1} \mathfrak{b}_i^2\, d\cos\vartheta_2\, d\Omega(\mathbf{e}_3) \times \cdots d\Omega(\mathbf{e}_{n-1})\, d\rho\, d\mathfrak{b}_1 \cdots d\mathfrak{b}_{n-1} \tag{4.53}$$

(Compare also Ref. 4), where $d\Omega(\mathbf{e}) = d\cos\vartheta d\varphi$ is the element of volume of the sphere in three-dimensional space. Hence, the result is

$$A_n(\kappa_0) = \frac{\pi^2}{2^{n-5}} \prod_{i=1}^{n-1} \frac{\mathfrak{b}_i^2}{\omega_i(\mathfrak{b}_i)}\, M(\kappa_0)\cdot\frac{1}{\omega_n(\mathfrak{b}_n)}, \tag{4.54}$$

where we have to set

$$\mathfrak{b}_n = -\sum_{i=1}^{n-1} \mathfrak{b}_i; \qquad M(\kappa_0) = \sum_{i=1}^{n} \omega_i(\mathfrak{b}_i). \tag{4.55}$$

We notice that, by choosing this parametrization, $A_n(\kappa)$ has the same value on the equivalence class of the mirror polygons as on the original one

$$A_n(\Pi\kappa_0) = A_n(\kappa_0), \tag{4.56}$$

and, on the permuted class, it has the value

$$A_n(\pi\kappa_0) = \mathfrak{b}_n^2\, \mathfrak{b}_{\pi^{-1}n}^{-2}\, A_n(\kappa_0). \tag{4.57}$$

In the case $n = 2$, we get

$$\frac{1}{4}\frac{d^3p_1\, d^3p_2}{\omega_1(\mathfrak{p}_1)\omega_2(\mathfrak{p}_2)} = \frac{d^3q}{2\omega(\mathfrak{q})}\, A_2(\mathfrak{b})\, d\Omega(\mathbf{e})\, d\mathfrak{b} = \frac{d^3q}{2\omega(\mathfrak{q})}\, A_2(M)\, d\Omega(\mathbf{e})\, dM, \tag{4.58}$$

with $A(|\mathfrak{b}|)$ and $A(M)$ given by

$$A_2(|\mathfrak{b}|) = \frac{\mathfrak{b}^2}{2\omega_1(\mathfrak{b})\omega_2(\mathfrak{b})} M(\mathfrak{b}), \qquad A_2(M) = \left(\frac{dM}{d\mathfrak{b}}\right)^{-1} A_2(\mathfrak{b}) = \frac{\mathfrak{b}}{2}. \tag{4.59}$$

The last equation is a consequence of

$$\frac{dM}{d\mathfrak{b}} = \left(\frac{1}{\omega_1(\mathfrak{b})} + \frac{1}{\omega_2(\mathfrak{b})}\right)\mathfrak{b} = \frac{M}{\omega_1(\mathfrak{b})\omega_2(\mathfrak{b})}\,\mathfrak{b}.$$

In the case $n = 2$, the mirror and the permuted equivalence class of polygons coincide with the original one, on which A is a constant.

5. DECOMPOSITION OF AN N-FOLD PRODUCT OF SPACES

We now consider the space (4.49), where the assumption $|\mathbf{m}| = \sum_{i=1}^{n} m_i > 0$ is made. The elements of $\mathfrak{H}_n(\mathbf{m}, \mathbf{s})$ are column matrices f, g, with $N_{\tau_1}(s_1)\cdot N_{\tau_2}(s_2)\cdot \cdots N_{\tau_n}(s_n)$ ($\tau_k = \operatorname{sgn} m_k$) rows, which depend on $\mathbf{p}$: $f(\mathbf{p})$, $g(\mathbf{p})$, and the scalar product is defined by

$$(f, g) = \int_{V^+(\mathbf{m})} \frac{d^{3n}p}{\omega_1(\mathbf{p}_1)\cdots\omega_n(\mathbf{p}_n)} \cdot \frac{1}{2^n} f^\dagger(\mathbf{p}) g(\mathbf{p}). \tag{5.1}$$

The representation of $\tilde{P}$ in $\mathfrak{H}_n(\mathbf{m}, \mathbf{s})$ is, according to (3.5) and (3.12), given by

$$\begin{aligned} (U(a)f)(\mathbf{p}) &= e^{iqa} f(\mathbf{p}), \\ (U(A)f)(\mathbf{p}) &= \mathfrak{D}^{\mathbf{s}}(\rho(\mathbf{p}, A)) f(\Lambda^{-1}(A)\mathbf{p}), \\ (U(\Pi)f)(\mathbf{p}) &= \mathfrak{D}^{\mathbf{s}}(\rho_\Pi(\mathbf{p})) f(\Pi\mathbf{p})\cdot \mathbf{n}_\Pi, \\ (V(T)f)(\mathbf{p}) &= \mathfrak{D}^{\mathbf{s}}(\epsilon)(\overline{U(\Pi)f})(\mathbf{p})\cdot \mathbf{n}, \end{aligned} \tag{5.2}$$

where $\mathfrak{D}^{\mathbf{s}}(\rho_\Pi(\mathbf{p}))$ is defined by

$$\mathfrak{D}^{\mathbf{s}}(\rho_\Pi(\mathbf{p})) = \mathfrak{D}^{s_1}(\rho_{1\Pi}(p_1)) \otimes \cdots \otimes \mathfrak{D}^{s_n}(\rho_{n\Pi}(p_n)) \tag{5.3}$$

and $\mathfrak{D}^{\mathbf{s}}(\rho(\mathbf{p}, A))$ in a corresponding way. ($\otimes$ means Kronecker product.) Moreover,

$$\mathbf{n}_\Pi = \eta_{1\Pi} \cdots \eta_{n\Pi}, \qquad \mathbf{n} = \eta_1 \cdots \eta_n. \tag{5.4}$$

Now we define a representation of γ_n' in $\mathfrak{H}_n(\mathbf{m}, \mathbf{s})$, where γ_n' is that subgroup of the symmetric group of n elements, which permutes equal particles among themselves by

$$(U(\pi^{-1})f)(\mathbf{p}) = C(\pi^{-1}) f(\pi\mathbf{p}), \tag{5.5}$$

or, more explicitly,

$$\begin{aligned} (U(\pi^{-1})f)_{\mu_1\cdots\mu_n}(\mathbf{p}) &= C^{\nu_1\cdots\nu_n}_{\mu_1\cdots\mu_n}(\pi^{-1}) f_{\nu_1\cdots\nu_n}(\pi\mathbf{p}) \\ &= f_{\pi\mu_1\cdots\pi\mu_n}(\pi\mathbf{p}). \end{aligned}$$

By this equation the matrix C is defined. The subspace of $\mathfrak{H}_n(\mathbf{m}, \mathbf{s})$, which carries the physically correct representation of the group γ_n', we denote by $\mathfrak{H}_n^\gamma(\mathbf{m}, \mathbf{s})$. The elements belonging to this subspace satisfy the relation

$$(U(\pi^{-1})f)(\mathbf{p}) = f(\mathbf{p})\cdot \eta(\pi), \tag{5.6}$$

where $\eta(\pi)$ is the signature of the fermion permutation under π. We now define

$$\mathfrak{D}^{\mathbf{s}}(\mathbf{p}) \equiv \mathfrak{D}^{s_1}(r_1(\mathbf{p})) \otimes \cdots \otimes \mathfrak{D}^{s_n}(r_n(\mathbf{p})), \tag{5.7}$$

where $r_l(\mathbf{p})$ was defined in (4.38).

Moreover we define

$$F(\mathbf{p}) = \mathfrak{D}^{\mathbf{s}}(\mathbf{p}) f(\mathbf{p}). \tag{5.8}$$

Hence, we find according to (4.41)

$$\begin{aligned} (U(A)F)(\mathbf{p}) &= \mathfrak{D}^{\mathbf{s}}(\mathbf{p})(U(A)f)(\mathbf{p}) = F(\Lambda^{-1}(A)\mathbf{p}), \\ (U(\Pi)F)(\mathbf{p}) &= \mathfrak{D}^{\mathbf{s}}(\mathbf{p})(U(\Pi)f)(\mathbf{p}) \\ &= \mathfrak{D}^{\mathbf{s}}(w_\Pi^{-1}(\kappa))F(\Pi\mathbf{p})\cdot \mathbf{n}_\Pi, \\ (U(\pi^{-1})F)(\mathbf{p}) &= \mathfrak{D}^{\mathbf{s}}(\mathbf{p}) C(\pi^{-1}) f(\pi\mathbf{p}), \end{aligned} \tag{5.9}$$

or, because of

$$\mathfrak{D}^{\mathbf{s}}(\mathbf{p}) C(\pi^{-1}) = C(\pi^{-1}) \mathfrak{D}^{\mathbf{s}}_\pi(\mathbf{p})$$

[where

$$\mathfrak{D}^{\mathbf{s}}_\pi(\mathbf{p}) = \mathfrak{D}^{s_{\pi_1}}(r_{\pi_1}(\mathbf{p})) \otimes \cdots \otimes \mathfrak{D}^{s_{\pi_n}}(r_{\pi_n}(\mathbf{p}))]$$

and (4.41c), we have

$$(U(\pi^{-1})F)(\mathbf{p}) = C(\pi^{-1}) \mathfrak{D}^{\mathbf{s}}(w_\pi^{-1}(\kappa)) F(\pi\mathbf{p}). \tag{5.10}$$

Besides the space $\mathfrak{H}_n(\mathbf{m}, \mathbf{s})$, we now consider the space of the column matrices $\tilde{\varphi}$ with $N_{\tau_1}(s_1)\cdot \cdots N_{\tau_n}(s_n)$ rows, which are square-integrable functions on the set

$$\begin{gathered} \{q, \rho, \kappa\} \quad q^2 \geq |\mathbf{m}|^2; \\ \rho \in SU(2, C)/G_n; \quad \kappa \in C_n([q^2]^{\frac{1}{2}}), \end{gathered} \tag{5.11}$$

and which, in addition, have the following properties [in the case $n = 2$, we may drop the dependence of $\tilde{\varphi}$ on κ because $\kappa = \kappa_1$ is determined by $M = (q^2)^{\frac{1}{2}}$]:

$$\begin{aligned} n = 2 &: \rho(\chi) \in DU(2, C), \\ \tilde{\varphi}(q, \rho\rho(\chi)) &= \mathfrak{D}^{\mathbf{s}}(y(\mathfrak{b}, \chi))\tilde{\varphi}(q, \rho), \end{aligned} \tag{5.12}$$

with $y(b_l, \chi)$ defined in (4.40);

for $n \geq 3$

$$\tilde{\varphi}(q, -\rho, \kappa) = (-1)^{2|s|}\tilde{\varphi}(q, \rho, \kappa), \tag{5.13}$$

where $|s| = \sum_{i=1}^{n} s_i$ and

$$\mathfrak{D}^{\mathbf{s}}(y(\mathbf{b}, \chi)) = \mathfrak{D}^{s_1}(y(b_1, \chi)) \otimes \mathfrak{D}^{s_2}(y(b_2, \chi)). \tag{5.14}$$

We denote the space of these functions by $\tilde{\mathfrak{K}}_n(\mathbf{m}, \mathbf{s})$. The scalar product in it is defined by

$$(\tilde{\varphi}, \tilde{\psi}) = \int d\rho \int d^{3n-6}\kappa \times \int_{V^+{}_{M(\kappa)}} \frac{d^3q}{2\omega(\mathbf{q})} \tilde{\varphi}^\dagger(q, \rho, \kappa)\tilde{\psi}(q, \rho, \kappa)$$

$$= \int_{q^2 \geq |\mathbf{m}|^2} d^4q \int d\rho \int d^{3n-6}\kappa\theta(q)\, \delta(q^2 - M^2(\kappa)) \times \tilde{\varphi}^\dagger(q, \rho, \kappa)\tilde{\psi}(q, \rho, \kappa), \qquad (5.15)$$

where

$$\theta(q) = \begin{cases} 1 & q_0 \geq 0, \\ 0 & q_0 < 0. \end{cases}$$

It is invariant under the substitution

$$\tilde{\varphi} \to \mathfrak{D}^{\mathbf{s}}(y(\mathbf{b}, \chi))\tilde{\varphi}$$

because $y(b_l, \chi)$ is unitary.

According to (4.39), the definitions (5.8), and (3.1), the properties (5.12) and (5.13) are exactly the ones possessed by the functions $F(\mathbf{p})$, if they are comprehended as functions of q, ρ, and κ. Hence, there exists a one-to-one correspondence between the elements of $\mathfrak{H}_n(\mathbf{m}, \mathbf{s})$ and that of $\tilde{\mathfrak{K}}_n(\mathbf{m}, \mathbf{s})$. To make it an isomorphism between the Hilbert spaces,

$$\mathfrak{H}_n(\mathbf{m}, \mathbf{s}) \cong \tilde{\mathfrak{K}}_n(\mathbf{m}, \mathbf{s}),$$

we have to set

$$\tilde{\varphi}(q, \rho, \kappa) = A_n^{\frac{1}{2}}(\kappa)F(\mathbf{p}). \qquad (5.16)$$

with $A_n(\kappa)$ defined in (4.52) and (4.59). Indeed, we now get

$$(\tilde{\varphi}, \tilde{\psi}) = (F, g) = (f, g).$$

According to (4.34), Eq. (5.16) induces the following representation of $\tilde{P}$ and γ_n' in $\tilde{\mathfrak{K}}_n(\mathbf{m}, \mathbf{s})$:

$$\begin{aligned}
(U(a)\tilde{\varphi})(q, \rho, \kappa) &= e^{iqa}\tilde{\varphi}(q, \rho, \kappa), \\
(U(A)\tilde{\varphi})(q, \rho, \kappa) &= \tilde{\varphi}(\Lambda^{-1}(A)q, \rho^{-1}(q, A)\rho, \kappa), \\
(U(\Pi)\tilde{\varphi})(q, \rho, \kappa) &= \mathfrak{n}_\Pi B_\Pi(\kappa) \\
&\quad \times \mathfrak{D}^{\mathbf{s}}(w_\Pi^{-1}(\kappa))\tilde{\varphi}(\Pi q, \rho_\Pi^{-1}(q)\rho v_\Pi^{-1}(\kappa), \Pi\kappa), \\
(U(\pi^{-1})\tilde{\varphi})(q, \rho, \kappa) &= B_\pi(\kappa)C(\pi^{-1}) \\
&\quad \times \mathfrak{D}^{\mathbf{s}}(w_\pi^{-1}(\kappa))\tilde{\varphi}(q, \rho v_\pi^{-1}(\kappa), \pi\kappa), \qquad (5.17)
\end{aligned}$$

where

$$B_\Pi(\kappa) = \left[\frac{A_n(\kappa)}{A_n(\Pi\kappa)}\right]^{\frac{1}{2}}; \qquad B_\pi(\kappa) = \left[\frac{A_n(\kappa)}{A_n(\pi\kappa)}\right]^{\frac{1}{2}}.$$

If we now make the WJW choice of the adjustable quantities given in (4.29), (4.31), (4.33), (4.44) and assume a parametrization of the equivalence classes of $\mathcal{O}_n(\mathbf{m})$ for which

$$B_\Pi(\kappa) = 1$$

[this condition is for example satisfied for the parametrization, given in (4.9)], we get

$$\begin{aligned}
(U(a)\tilde{\varphi})(q, \rho, \kappa) &= e^{iqa}\tilde{\varphi}(q, \rho, \kappa), \\
(U(A)\tilde{\varphi})(q, \rho, \kappa) &= \tilde{\varphi}(\Lambda^{-1}(A)q, \rho^{-1}(q, A)\rho, \kappa), \qquad (5.18) \\
(U(\Pi)\tilde{\varphi})_\lambda(q, \rho, \kappa) &= (-1)^{|s|-|\lambda|}\tilde{\varphi}_{-\lambda}(\Pi q, \rho_\Pi^{-1}(q)\rho\epsilon^{-1}, \Pi\kappa)\cdot\mathfrak{n}_\Pi,
\end{aligned}$$

where

$$|s| = \sum_{i=1}^n s_i; \qquad |\lambda| = \begin{cases} \sum_{i=1}^n \lambda_i, & n \geq 3, \\ \lambda_1 - \lambda_2, & n = 2; \end{cases}$$

and, for $n = 2$,

$$(U(\pi^{-1})\tilde{\varphi})_{\lambda_1\lambda_2}(q, \rho) = (-1)^{\lambda_1+\lambda_2}\tilde{\varphi}_{\lambda_2\lambda_1}(q, \rho\epsilon^{-1}), \qquad (5.19)$$

for $n \geq 3$,

$$(U(\pi^{-1})\tilde{\varphi})(q, \rho, \kappa) = B_\pi(\kappa)C(\pi^{-1})\tilde{\varphi}(q, \rho, \pi\kappa); \ \pi \in \gamma_n'.$$

This means that the states belonging to the physically correct representation of the group have the property

$$n \geq 3 \quad B_\pi(\kappa)C(\pi^{-1})\tilde{\varphi}(q, \rho, \pi\kappa) = \eta(\pi)\tilde{\varphi}(q, \rho, \kappa). \qquad (5.20)$$

In the case $n = 2$, if the two particles are equal,

$$\tilde{\varphi}_{\lambda_1\lambda_2}(q, \rho) = \tilde{\varphi}_{\lambda_2\lambda_1}(q, \rho\epsilon^{-1})(-1)^{\lambda_1+\lambda_2}\cdot\eta(\pi). \qquad (5.21)$$

Because of (4.47), the property (5.12) of φ takes the form

$$\tilde{\varphi}_{\lambda_1\lambda_2}(q, \rho\rho(\chi)) = e^{-i(\lambda_1-\lambda_2)\chi}\tilde{\varphi}_{\lambda_1\lambda_2}(q, \rho). \qquad (5.22)$$

We now execute the transformation

$$\mathring{\varphi}(q; J\lambda\sigma; \kappa) = (2J+1)^{\frac{1}{2}} \int d\rho\, \tilde{\varphi}(q, \rho, \kappa)\, \mathfrak{D}^J_{\lambda\sigma}(\rho), \qquad (5.23)$$

which possesses the inverse

$$\tilde{\varphi}(q, \rho, \kappa) = \sum_{J,\sigma,\lambda} (2J+1)^{\frac{1}{2}}\, \mathring{\varphi}(q; J\lambda\sigma; \kappa)\overline{\mathfrak{D}}{}^J_{\lambda\sigma}(\rho), \qquad (5.24)$$

and which defines an isomorphism between $\tilde{\mathfrak{K}}_n(\mathbf{m}, \mathbf{s})$ and a new Hilbert space, with a scalar product defined as

$$(\mathring{\varphi}, \mathring{\psi}) = \sum_{J\sigma\lambda} \int d^{3n-6}\kappa \int_{V^+{}_{M(\kappa)}} \frac{d^3q}{2\omega(\mathbf{q})} \times \mathring{\varphi}^\dagger(q; J\lambda\sigma; \kappa)\, \mathring{\psi}(q; J\lambda\sigma; \kappa)$$

$$= \sum_{J\sigma\lambda} \int_{C_n} d^{3n-6}\kappa \int_{q^2 \geq |\mathbf{m}|^2} d^4q\, \theta(q)\, \delta(q^2 - M^2(\kappa)) \times \mathring{\varphi}^\dagger(q; J\lambda\sigma; \kappa)\, \mathring{\psi}(q; J\lambda\sigma; \kappa). \qquad (5.25)$$

The transformed relations (5.13), (5.22), and (5.21) are

$$n \geq 3 : \mathring{\varphi}(q; J\lambda\sigma;\kappa)[1 - (-1)^{2(J-|s|)}] = 0, \quad (5.26)$$

$$n = 2 : \mathring{\varphi}_{\lambda_1\lambda_2}(q; J\lambda\sigma)(1 - e^{i[\sigma-(\lambda_1-\lambda_2)]\chi}) = 0, \quad (5.27)$$

$$\mathring{\varphi}(q; J\lambda\sigma) = C(\pi)\mathring{\varphi}(q; J, \lambda, -\sigma)(-1)^{J-2s}\cdot\eta(\pi), \quad (5.28)$$

and the representation of $\tilde{P}^{\uparrow}_{+}$ gets the demanded form in which it decomposes into irreducible parts,

$$(U(a)\mathring{\varphi})(q; J\lambda\sigma;\kappa) = e^{iqa}\mathring{\varphi}(q; J\lambda\sigma;\kappa), \quad (5.29a)$$

$$(U(A)\mathring{\varphi})(q; J\lambda\sigma;\kappa) = \mathfrak{D}^{J}_{\lambda\lambda'}(\rho(q, A))\mathring{\varphi}(\Lambda^{-1}(A)q; J\lambda'\sigma;\kappa), \quad (5.29b)$$

$$(U(\Pi)\mathring{\varphi})_{\lambda}(q; J\lambda\sigma;\kappa) = \mathfrak{n}_{\Pi}(-1)^{J-\sigma-|s|+|\lambda|}\,\mathfrak{D}^{J}_{\lambda\lambda'}(\rho_{\Pi}(q))\mathring{\varphi}_{-\lambda}(\Pi q; J\lambda', -\sigma;\Pi\kappa), \quad (5.29c)$$

$$(V(T)\mathring{\varphi})(q; J\lambda\sigma;\kappa) = (2J+1)^{\frac{1}{2}}\int d\rho\, \mathfrak{D}^{s}(\epsilon)(\overline{U(\Pi)\bar{\varphi}})(q, \rho, \kappa)\, \mathfrak{D}^{J}_{\lambda\sigma}(\rho)\cdot\mathfrak{n} = \mathfrak{D}^{J}_{\lambda\lambda'}(\rho_{\Pi}(q)\epsilon)\bar{\mathring{\varphi}}(\Pi q; J\lambda'\sigma;\Pi\kappa)\cdot\mathfrak{n}_{T}. \quad (5.29d)$$

If we denote by the symbol $\mathring{\vartheta}_n(x)$ the set of points in x-space on which the function

$$\mathring{\varphi}_{\lambda}(q; J\lambda\sigma;\kappa)$$

may be arbitrarily defined [defining φ on the following sets is, of course, equivalent with defining $f_{\lambda}(\mathbf{p})$ on the sets: $\mathbf{p} \in V^{+}(\mathbf{m})$; $\lambda \in \Pi^{n}_{k=1} \otimes I_{\tau_k}(s_k)$]—with the restriction that $(\mathring{\varphi}, \mathring{\varphi})$ is finite—we get according to (5.26) and (5.27)

$$\mathring{\vartheta}_n(q) = \{q; q^2 \geq |\mathrm{m}|^2\},$$

$$\mathring{\vartheta}_n(J) = \{J; 2J \equiv 2|s| \mod 2\}, \quad (5.30)$$

$$\mathring{\vartheta}_n(\sigma, \lambda) = \begin{cases} I_{+}(J) \otimes \prod_{k=1}^{n} \otimes I_{\tau_k}(s_k); & n \geq 3, \\ [\sigma = \lambda_1 - \lambda_2; \lambda \in I_{\tau_1}(s_1) \otimes I_{\tau_2}(s_2); |\lambda_1 - \lambda_2| \leq J]; & n = 2 \end{cases}$$

$$\mathring{\vartheta}_n(\kappa) = C_n[(q^2)^{\frac{1}{2}}].$$

The set $I_{\tau_1}(s_1) \otimes I_{\tau_2}(s_2)$ forms a rectangle ($s_1 \neq s_2$) or a square ($s_1 = s_2$) in the $\lambda_1\lambda_2$-plane.

$\mathring{\vartheta}_2(\lambda)$ consists of those points of this set simultaneously lying between or on the straight lines given by the equations

$$\lambda_1 - \lambda_2 = J, \qquad \lambda_2 - \lambda_1 = J.$$

[In the case $n = 2$, we may drop the dependence of $\mathring{\varphi}$ on κ and σ because κ is determined by $M = (q^2)^{\frac{1}{2}}$ and σ by λ.]

The space of these functions $\mathring{\varphi}$ equipped with the scalar product (5.25) we denote by $\mathring{\mathfrak{K}}_n(\mathbf{m}, \mathbf{s})$.

If two or more of the particles are equal, and if γ'_n is, as above, the group of those permutations which interchange equal particles among themselves, the functions $\mathring{\varphi}$, in addition, have the property

$$B_{\pi}(\kappa)C(\pi^{-1})\mathring{\varphi}(q; J\lambda\sigma;\pi\kappa) = \eta(\pi)\mathring{\varphi}(q; J\lambda\sigma;\kappa) \quad (5.31)$$

for $n \geq 3$, and

$$C(\pi^{-1})\mathring{\varphi}(q; J\lambda)(-1)^{J-2s}\cdot\eta(\pi) = \mathring{\varphi}(q; J\lambda) \quad (5.32)$$

for $n = 2$.

The space of these functions equipped with the scalar product (5.25) we denote by $\mathring{\mathfrak{K}}^{\gamma}_{n}(\mathbf{m}, \mathbf{s})$ and the set on which the function comprehended as a function of x may be defined arbitrarily [up to the restriction $(\mathring{\varphi}, \mathring{\varphi}) < \infty$] we call $\mathring{\vartheta}^{\gamma}_{n}(x)$.

Let τ be a set of points in λ-space. We introduce a measure on these sets by the definition

$$\mu(\tau) = \text{number of points contained in } \tau.$$

We call a set τ_1 smaller than a set τ_2 if and only if

$$\mu(\tau_1) < \mu(\tau_2).$$

We do the same in κ-space except we take as measure the Lebesgue measure $\mu(\kappa)$. $\mathring{\vartheta}^{\gamma}_{n}(\lambda)$ belongs to the set of the smallest subsets τ of $\mathring{\vartheta}_n(\lambda)$ which have the property

$$\bigcup_{\pi\in\gamma_{n'}} \pi\tau = \mathring{\vartheta}_n(\lambda)$$

(the symbol $\bigcup$ means "union"). An analogous statement holds for $\mathring{\vartheta}_n(\kappa)$.

In the case $n = 2$, we only have to consider $\mathring{\vartheta}^{\gamma}_{2}(\lambda)$ which may be chosen as that subset of the square formed by $\mathring{\vartheta}_2(\lambda)$ lying on the upper side of the diagonal given by the equation $\lambda_1 = \lambda_2$, the diagonal included or not depending on whether the upper or lower sign holds in the equation

$$\eta(\pi)(-1)^{J-2s} = \pm 1. \quad (5.33)$$

Comparing (5.29) with (3.5) and (3.12) we remark that the space $\mathring{\mathfrak{K}}_n(\mathbf{m}, \mathbf{s})$, although reduced with respect to the group $\tilde{P}^{\uparrow}_{+}$, is still not reduced with respect to $\tilde{P}$. But, because

$$V^2(T) = (-1)^{2J}, \qquad V^2(\Pi T) = (-1)^{2J} \quad \text{(if all the } \eta_{\Pi}\text{'s are real)}, \quad (5.34)$$

it should be possible to perform an additional transformation in such a way that Π and T are represented in agreement with (3.5c) and (3.12) (see Ref. 7).

The transformation we are looking for is given by

$$\varphi(q, J\lambda; \sigma\lambda\kappa\xi) = \tfrac{1}{4}\{\mathring{\varphi}_\lambda(q; J\lambda\sigma; \kappa) + \xi_2\mathring{\varphi}_\lambda(q; J\lambda\sigma; \Pi\kappa) + \xi_1\xi_2(-1)^{J-\sigma-|s|+|\lambda|}[\mathring{\varphi}_{-\lambda}(q; J\lambda, -\sigma; \kappa) + \xi_2\mathring{\varphi}_{-\lambda}(q; J\lambda, -\sigma; \Pi\kappa)]\}. \tag{5.35}$$

Here $\xi = (\xi_1, \xi_2)$ are two new parameters both of which, in general, can take the values $(+1, -1)$.

φ has again the properties (5.26), (5.27), and, in addition, it satisfies the relations

$$\varphi(q, J\lambda; \sigma\lambda, \Pi\kappa, \xi) = \xi_2\varphi(q, J\lambda; \sigma\lambda\kappa\xi),$$
$$\varphi(q, J\lambda; -\sigma, -\lambda\kappa\xi) = \xi_1\xi_2(-1)^{J-\sigma-|s|+|\lambda|}\varphi(q, J\lambda; \sigma\lambda\kappa\xi). \tag{5.36}$$

By $\mathfrak{R}_n(\mathbf{m}, \mathbf{s})$, we denote the space of the functions φ with all these properties equipped with the scalar product

$$(\varphi, \psi) = \sum_{J\lambda}\sum_{\sigma\lambda\xi}\int d^{3n-6}\kappa \int_{V+M(\kappa)} \frac{d^3q}{2\omega(q)} \times \bar{\varphi}(q, J\lambda; \sigma\lambda\kappa\xi)\psi(q, J\lambda; \sigma\lambda\kappa\xi). \tag{5.37}$$

Here summation and integration extends over the domain

$$\{(J\lambda; \sigma\lambda\kappa) \in \mathring{\vartheta}_n(J\lambda; \sigma\lambda\kappa); \xi_1 = \pm 1; \xi_2 = \pm 1\}.$$

We now see that

$$(\varphi, \psi) = (\mathring{\varphi}, \mathring{\psi})$$

holds. Hence, in sum,we get the following sequence of isomorphisms:

$$\mathfrak{H}_n(\mathbf{m}, \mathbf{s}) \cong \tilde{\mathfrak{R}}_n(\mathbf{m}, \mathbf{s}) \cong \mathring{\mathfrak{R}}_n(\mathbf{m}, \mathbf{s}) \cong \mathfrak{R}_n(\mathbf{m}, \mathbf{s}). \tag{5.38}$$

We are therefore allowed to identify all these spaces. Because of (5.29c, d), we deduce from (5.36)

$$(U(\Pi)\varphi)(q, J\lambda; \sigma\lambda\kappa\xi) = \mathfrak{n}_\Pi\xi_1\, \mathfrak{D}^J_{\lambda\lambda'}(\rho_\Pi(q))\varphi(\Pi q, J\lambda'; \sigma\lambda\kappa\xi),$$
$$(V(T)\varphi)(q, J\lambda; \sigma\lambda\kappa\xi) = \mathfrak{n}_T\xi_2\, \mathfrak{D}^J_{\lambda\lambda'}(\rho_\Pi(q)\epsilon)\tilde{\varphi}(\Pi q, J\lambda'; \sigma\lambda\kappa\xi),$$

which indeed is in agreement with (3.5c) and (3.12), if we make the substitutions

$$\eta_\Pi \to \mathfrak{n}_\Pi\xi_1, \qquad f_\lambda(q) \to \varphi(q, J\lambda; \sigma\lambda\kappa\xi).$$
$$\eta_T \to \mathfrak{n}_T\xi_2,$$

The subspace of $\mathfrak{R}_n(\mathbf{m}, \mathbf{s})$ which carries the correct representation of the permutation group we denote by $\mathfrak{R}^\gamma_n(\mathbf{m}, \mathbf{s})$. Because space inversion and a permutation of the edges of a (three-dimensional) polygon are interchangeable, we get

$$\pi\Pi\kappa = \Pi\pi\kappa,$$

and, therefore

$$B_\pi(\Pi\kappa) = B_\pi(\kappa),$$

if we choose the parametrization κ in such a way that $A_n(\Pi\kappa) = A_n(\kappa)$. Hence the space $\mathfrak{R}^\gamma_n(\mathbf{m}, \mathbf{s})$ is, according to (5.31) and (5.32), characterized as a subspace of $\mathfrak{R}_n(\mathbf{m}, \mathbf{s})$ by the following additional properties of φ:

$$n = 2 : \eta(\pi)(-1)^{J-2s}\varphi(q, J\lambda; \lambda_2\lambda_1\xi) = \varphi(q, J\lambda; \lambda_1\lambda_2\xi);$$
$$n \geq 3 : B_\pi(\kappa)\varphi(q, J\lambda; \sigma, \pi\lambda, \pi\kappa\xi) = \eta(\pi)\varphi(q, J\lambda; \sigma\lambda\kappa\xi). \tag{5.39}$$

If we denote the set in x-space on which the function φ may be arbitrarily defined [up to the restriction $(\varphi, \varphi) < \infty$] by $\vartheta_n(x)$, we find

$$\vartheta_n(q) = \{q; q^2 \geq |\mathbf{m}|^2\},$$
$$\vartheta_n(J) = \{J; 2J \equiv 2|s| \bmod 2\},$$
$$\vartheta_n(\lambda) = I_+(J).$$

$\vartheta_n(\sigma, \lambda)$ belongs to the set of the smallest subsets τ of $\mathring{\vartheta}_n(\sigma_1\lambda)$ with the property

$$\tau \cup \Pi\tau = \mathring{\vartheta}_n(\sigma, \lambda).$$

(Note that the effect of Π on a "point in spin-space" is defined by $\Pi(\sigma, \lambda) = (-\sigma, -\lambda)$.)

$\vartheta_n(\kappa)$ belongs to the set of the smallest subsets τ of $\mathring{\vartheta}_n(\kappa)$ with the property

$$\tau \cup \Pi\tau = \mathring{\vartheta}_n(\kappa) = C_n([q^2]^{\frac{1}{2}}).$$

Because of our definition of $\mathfrak{R}_n(\mathbf{m}, \mathbf{s})$, two functions φ, which differ only on a set of zero measure in κ-space must be considered as the same element of $\mathfrak{R}_n(\mathbf{m}, \mathbf{s})$. That's why $\vartheta_n(\kappa)$ is only determined up to a set of zero measure, and, consequently in the case $n \geq 4$, we may assume that φ is zero on all points with the property

$$\Pi\kappa = \kappa \tag{5.40}$$

so that $\vartheta_n(\kappa)$contains no such points. In the case $n = 3$, all points κ have the property (5.40), or, expressed differently,

$$\vartheta_3(\kappa) = \mathring{\vartheta}_3(\kappa) = C_3([q^2]^{\frac{1}{2}})$$

holds. In addition, we have

$$\begin{aligned}\vartheta_n(\xi_2) &= \{\xi_2;\ \xi_2 = \pm 1 \quad \text{for} \quad n \geq 4;\\ &\qquad\ \xi_2 = 1 \quad \text{for} \quad n = 2, 3\};\\ \vartheta_n(\xi_1) &= \{\xi_1;\ \xi_1 = \pm 1 \quad \text{for} \quad (\sigma, \lambda) \neq 0;\\ &\qquad\ \xi_1 = \xi_2(-1)^{J-|s|} \quad \text{for} \quad (\sigma, \lambda) = 0\}.\end{aligned} \tag{5.41}$$

$\vartheta_2(\lambda)$ may be chosen as the set of points of $\mathring{\vartheta}_2(\lambda)$ lying on the upper side of the diagonal ($\lambda_1 = -\lambda_2$) of the rectangle which is formed by the points of

$$I_{\tau_1}(s_1) \otimes I_{\tau_2}(s_2)$$

and on one half of the diagonal itself [the point (0, 0) included, if it belongs to $\mathring{\vartheta}_2(\lambda)$, which is exactly the case if s_1 and s_2 are integers].

In analogy to the sets $\mathring{\vartheta}_n^\gamma(x)$ we introduce the sets $\vartheta_n^\gamma(x)$. $\vartheta_2^\gamma(\lambda)$ may be chosen as $\vartheta_2(\lambda) \cap \mathring{\vartheta}_2^\gamma(\lambda)$. ($\cap$ means intersection), where the last two domains are chosen as indicated above. From (5.36a) and (5.36b) we deduce

$$\eta(\pi)\varphi(q, J\lambda; \lambda_2\lambda_1\xi) = \xi_1\varphi(q, J\lambda; -\lambda_1, -\lambda_2\xi),$$

and consequently

$$\begin{aligned}\vartheta_2^\gamma(\xi_1) &= \{\xi_1;\ \xi_1 = \pm 1 \quad \text{for} \quad \lambda_1 \neq -\lambda_2;\\ &\qquad\ \xi_1 = \eta(\pi) \quad \text{for} \quad \lambda_1 = -\lambda_2\}.\end{aligned} \tag{5.42}$$

$\vartheta_n^\gamma(\sigma, \lambda)$ for $n \geq 3$ belongs to the set of the smallest subsets τ of $\vartheta_n(\sigma, \lambda)$ with the property

$$\bigcup_{g \in \gamma_n'(\Pi)} g\tau = \mathring{\vartheta}_n(\sigma, \lambda),$$

where $\gamma_n'(\Pi)$ stands for the group γ_n' supplied by the element Π which commutes with the elements π of γ_n'. Note that the effect of π on "a point in spin-space" is defined by

$$\pi(\sigma, \lambda) = (\sigma, \pi\lambda).$$

Analogous to $\vartheta_n^\gamma(\sigma, \lambda)$, $\vartheta_n^\gamma(\kappa)$ is defined. By the same reason as in the case of $\vartheta_n(\kappa)$, $\vartheta_n^\gamma(\kappa)$ is only determined up to a set of zero measure. That is why we may assume that no point of the submanifolds in κ-space defined by

$$\begin{aligned}\pi\kappa &= \kappa, \quad \pi \in \gamma_n', \quad n = 3,\\ g\kappa &= \kappa, \quad g \in \gamma_n'(\Pi), \quad n \geq 3\end{aligned}$$

occur in $\vartheta_n^\gamma(\kappa)$.

It is now easy to derive the decomposition rules of the space with respect to the groups $\tilde{P}_+^\uparrow$, $\tilde{P}$, respectively.

Case $n = 2$: The number $n^{(\gamma)}(M, J)$ of representations (M, J) of the group $\tilde{P}_+^\uparrow$ equals the number of points contained in $\mathring{\vartheta}_2(\lambda)$ and $\mathring{\vartheta}_2^\gamma(\lambda)$, respectively, depending on whether the particles are of different kind or not. In formulas,

$$n^{(\gamma)}(M,J) = \begin{cases}\mu(\mathring{\vartheta}_2(\lambda)) & \text{if the particles are different,}\\ \mu(\mathring{\vartheta}_2^\gamma(\lambda)) & \text{if the particles are equal.}\end{cases}$$

If the particle are different and both have masses different from zero, this number is ($s_1 \geq s_2$):

$$n(M, J) = \begin{cases}(2s_1 + 1)(2s_2 + 1), & J \geq s_1 + s_2,\\ (2s_1 + 1)(2s_2 + 1) & \\ \quad - (s_1 + s_2 - J)(s_1 + s_2 - J + 1), & \\ & s_1 - s_2 \leq J \leq s_1 + s_2,\\ (2J + 1)(2s_2 + 1), & J \leq s_1 - s_2.\end{cases} \tag{5.43}$$

(Compare this with Refs. 3, 5, and 10.)

If it is even (which is always the case unless both of the particles have integral spin) half of it belongs to $\xi_1 = +1$ and half of it belongs to $\xi_1 = -1$. In the exceptional case, one of the representations (M, J) belongs to

$$\xi_1 = (-1)^{J-s_1-s_2} \qquad \text{(corresponding to } \lambda = 0),$$

and $\frac{1}{2}(n(M, J) - 1)$ to $\xi_1 = +1$, $\xi_1 = -1$, respectively. If both particles are equal with mass different from zero, the numbers $n^\gamma(M, J)$ are

	$J \geq 2s$	$J \leq 2s$
$\eta(\pi)(-1)^{J-2s} = +1$	$(2s + 1)(s + 1)$	$(J + 1)(2s + 1) - \frac{1}{2}[J(J + 1)]$
$\eta(\pi)(-1)^{J-2s} = -1$	$(2s + 1)s$	$2Js - \frac{1}{2}[J(J - 1)]$.

(5.44)

If we denote by $d(J)$ the number of points in $\vartheta_2^\gamma(\lambda)$ simultaneously lying on the straight line given by the equation $\lambda_1 = -\lambda_2$ we see that

$$\tfrac{1}{2}(n^\gamma(M, J) - d(J))$$

of the representations (M, J) belongs to $\xi_1 = +1$ and the same amount belongs to $\xi_1 = -1$. The rest, namely $d(J)$, belongs to

$$\xi_1 = \eta(\pi).$$

The numbers $d(J)$ are given in the following table.

$2s$	J	$J > 2s$	$J \leq 2s$
even	even	$\frac{1}{2}(2s + 1 \pm 1)$	$\frac{1}{2}(J + 1 \pm 1)$
even	odd	$\frac{1}{2}(2s + 1 \pm 1)$	$\frac{1}{2}(J \pm 1)$
odd	even	$\frac{1}{2}(2s + 1)$	$\frac{1}{2}J$
odd	odd	$\frac{1}{2}(2s + 1)$	$\frac{1}{2}(J + 1)$

(5.45)

The plus and minus sign in the first two lines simultaneously holds with the corresponding sign in the Eq. (5.33).

If we assume the correct connection between spin and statistics, i.e.,

$$\eta(\pi) = (-1)^{2s}, \tag{5.46}$$

then the numbers of representations (M, J) belonging to $\xi_1 = +1$, $\xi_1 = -1$, respectively, are

	$2s$	even	even	odd	odd
	J	even	odd	even	odd
$\xi_1 = +1$	$J \leq 2s$	$(J+1)(s+1) - \frac{J}{2}\left(\frac{J}{2}+1\right)$	$Js - \frac{(J-1)^2}{4}$	$\frac{1}{2}(2s+1)(J+1) - \frac{J(J+2)}{4}$	$J(s+\frac{1}{2}) - \frac{(J+1)^2}{4}$
	$J > 2s$	$(s+1)^2$	$s(s+1)$	$\frac{1}{4}(2s+1)^2$	$s^2 - \frac{1}{4}$
$\xi_1 = -1$	$J \leq 2s$	$(J+1)s - \frac{J^2}{4}$	$Js - \frac{J^2-1}{4}$	$\frac{1}{2}(2s+1)(J+1) - \frac{J^2}{4}$	$J(s+\frac{1}{2}) - \frac{J^2-1}{4}$
	$J > 2s$	$s(s+1)$	s^2	$\frac{1}{4}(2s+1)(2s+3)$	$(s+\frac{1}{2})^2$

For $n \geq 3$, to each point $\kappa \in \vartheta_n^{(\gamma)}(\kappa)$ there exists

$$\mu_n^{(\gamma)}, \quad n = 3,$$
$$2\mu_n^{(\gamma)}, \quad n \geq 4$$

representations (M, J) of $\tilde{P}_+^\uparrow$, where

$$\mu_n^{(\gamma)} = \mu(\mathring{\vartheta}_n^{(\gamma)}(\sigma, \lambda)).$$

If not all of the numbers s_k are integral,

$$n \geq 4 \begin{cases} \frac{1}{2}\mu_n^{(\gamma)} \text{ belongs to } \xi = (1, 1) \\ \frac{1}{2}\mu_n^{(\gamma)} \text{ belongs to } \xi = (-1, 1) \end{cases} \quad n = 3,$$
$$\frac{1}{2}\mu_n^{(\gamma)} \text{ belongs to } \xi = (1, -1),$$
$$\frac{1}{2}\mu_n^{(\gamma)} \text{ belongs to } \xi = (-1, 1).$$

If all the numbers s_k are integers, $\mu_n^{(\gamma)}$ in the table above has to be replaced by

$$\mu_n^{(\gamma)} - 1,$$

and we have to complete it by the following statements:

$$n \geq 4 \begin{cases} 1 \text{ belongs to } \xi = ((-1)^{J-|s|}, 1)\} & n = 3, \\ 1 \text{ belongs to } \xi = (-(-1)^{J-|s|}, -1). \end{cases}$$

Note especially that

$$\mu(\mathring{\vartheta}_n(\sigma, \lambda)) = (2J + 1) \prod_{k=1}^{n} (2s_k + 1). \tag{5.47}$$

Now let us consider the space $\mathfrak{H}(\mathring{\vartheta}_n)$ of the functions

$$f(\sigma, \lambda, \kappa)$$

defined on the domain

$$\mathring{\vartheta}_n(\sigma, \lambda, \kappa), \tag{5.48}$$

and assume that $\mathfrak{H}(\mathring{\vartheta}_n)$ is equipped with the scalar product

$$\sum_{\lambda, \sigma} \int d^{3n-6}\kappa \; \delta(q^2 - M^2(\kappa)) \bar{f}(\sigma, \lambda, \kappa) g(\sigma, \lambda, \kappa),$$

where the summation and integration are extended over the domain (5.48). We also consider that subspace $\mathfrak{H}(\mathring{\vartheta}_n^{(\gamma)})$ of $\mathfrak{H}(\mathring{\vartheta}_n)$ whose elements satisfy the relation

$$B_\pi(\kappa) f(\sigma, \pi\lambda, \pi\kappa) = \eta(\pi) f(\sigma, \lambda, \kappa), \quad n \geq 3,$$
$$\eta(\pi)(-1)^{J-2s} f(\pi\lambda) = f(\lambda), \quad u = 2.$$

Let us denote by

$$\langle \nu | \, H^{(\gamma)}(\mathbf{m}, \mathbf{s}, JM) \, | \sigma, \lambda, \kappa \rangle \tag{5.49}$$

a complete orthonormal system of functions of the space $\mathfrak{H}(\mathring{\vartheta}_n^{(\gamma)})$ which may depend also on the variables $\mathbf{m}$, $\mathbf{s}$, M, J. In the future we shall suppress this dependence to make the formulas look less cumbersome. Completeness and orthogonality are expressed by the relations

$$\sum_{\lambda\sigma} \int d^{3n-6}\kappa \; \delta(q^2 - M^2(\kappa)) \overline{\langle \nu | \, H^{(\gamma)} \, | \sigma, \lambda, \kappa \rangle} \times \langle \nu' | \, H^{(\gamma)} \, | \sigma, \lambda, \kappa \rangle = \delta_{\nu\nu'}, \tag{5.50a}$$

$$\sum_{\gamma} \delta(q^2 - M^2(\kappa))\langle \nu|\, \overline{H^{(\gamma)}}\, |\sigma, \lambda, \kappa\rangle \times \langle \nu|\, H^{(\gamma)}\, |\sigma', \lambda', \kappa'\rangle = \delta^{3n-6}(\kappa - \kappa')\, \delta_{\lambda\lambda'}\, \delta_{\sigma\sigma'}. \tag{5.50b}$$

In (5.49) the parameters ν "counting the members of the system" may partially be discreet, partially continuous. Accordingly we have to replace the sum in (5.50b) partially by an integral and the Kronecker-symbol on the right-hand side of (5.50a) by δ-functions. If we start from $\mathring{\mathfrak{R}}_n(\mathbf{m}, \mathbf{s})$ and perform the transformation

$$\hat{\varphi}(q; J\lambda; \nu) = \sum_{\lambda\sigma} \int d^{3n-6}\kappa\, \delta(q^2 - M^2(\kappa)) \times \langle \mu|\, H\, |\sigma, \lambda, \kappa\rangle \mathring{\varphi}_\lambda(q; J\lambda\sigma; \kappa), \tag{5.51}$$

where summation and integration again extend over the domain (5.48), then the space $\hat{\mathfrak{R}}_n(\mathbf{m}, \mathbf{s})$ of the functions $\hat{\varphi}$ equipped with the scalar product

$$(\hat{\varphi}, \hat{\psi}) = \sum_{J,\lambda} \sum_{\nu} \int d^4q\, \bar{\hat{\varphi}}(q; J\lambda; \nu)\hat{\psi}(q; J\lambda; \nu) \tag{5.52}$$

is operator-isomorphic to the original one with respect to the group $\tilde{P}_+^\uparrow$ and therefore splits into irreducible spaces with character (M, J) in the same way as $\mathring{\mathfrak{R}}_n(\mathbf{m}, \mathbf{s})$.

It is easily seen that (5.51) is the most general transformation with this property. The most general C–G coefficient of the group $\tilde{P}_+^\uparrow$ is therefore defined by

$$\hat{\varphi}(q; J\lambda; \nu) = \sum_{\lambda} \int \frac{1}{2^n} \frac{d^{3n}p}{\omega_1(\mathbf{p}_1) \cdots \omega_n(\mathbf{p}_n)} \left(\begin{matrix} \nu & MJ \\ & q\lambda \end{matrix} \middle|\, \begin{matrix} \mathbf{ms} \\ \mathbf{p}\lambda \end{matrix}\right)_{(\gamma)} f_\lambda(\mathbf{p}) \tag{5.53}$$

and is consequently given by the expression

$$\left(\begin{matrix} \nu & MJ \\ & q\lambda \end{matrix} \middle|\, \begin{matrix} \mathbf{ms} \\ \mathbf{p}\lambda \end{matrix}\right) = \delta^4(q - q')[(2J+1)A_n^{-1}(\kappa)]^{\frac{1}{2}} \times \sum_{\lambda',\sigma} \mathfrak{D}^s_{\lambda'\lambda}(\mathbf{p})\, \mathfrak{D}^J_{\lambda\sigma}(\rho)\langle \nu|\, H^{(\gamma)}\, |\lambda', \sigma, \kappa\rangle. \tag{5.54}$$

Here q', ρ, and κ have the values determined by $\mathbf{p}$ through the relation (4.21). In an analogous way we determine the C–G coefficients of $\tilde{P}$ considering only representations for which

$$V^2(T) = (-1)^{2s} \quad \text{and} \quad V^2(\Pi T) = (-1)^{2s} \text{(if } \eta_\Pi \text{ is real)}.$$

We only have to replace the functions

$$\langle \nu|\, H(\gamma)\, |\lambda', \sigma, \kappa\rangle$$

in (5.54) by the following linear combinations of these functions:

$$\langle \nu|\, H^{(\gamma)}(\xi)\, |\sigma\lambda\kappa\rangle = \tfrac{1}{4}\{[\langle \nu|\, H^{(\gamma)}\, |\sigma, \lambda, \kappa\rangle + \xi_2\langle \nu|\, H^{(\gamma)}\, |\sigma, \lambda, \Pi\kappa\rangle] + \xi_1\xi_2(-1)^{|s|-|\lambda|-J+\sigma}[\langle \nu|\, H^{(\gamma)}\, |-\sigma, -\lambda, \kappa\rangle + \xi_2\langle \nu|\, H^{(\gamma)}\, |-\sigma, -\lambda, \Pi\kappa\rangle]\}. \tag{5.55}$$

If we make the identification of all the spaces isomorphic to each other [see (5.38)], the transformations which reduce the space $\mathfrak{H}_n(\mathbf{m}, \mathbf{s})$ are seen to be orthogonal. That's why we know without direct calculation that the C–G coefficients satisfy the relations

$$\sum_J \sum_{\gamma,\lambda} \int d^4q \overline{\left(\begin{matrix} \nu & MJ \\ & q\lambda \end{matrix} \middle|\, \begin{matrix} \mathbf{ms} \\ \mathbf{p}^{(1)}\lambda^{(1)} \end{matrix}\right)} \left(\begin{matrix} \nu & MJ \\ & q\lambda \end{matrix} \middle|\, \begin{matrix} \mathbf{ms} \\ \mathbf{p}^{(2)}\lambda^{(2)} \end{matrix}\right) = 2^n \prod_{k=1}^{n} \omega(p_k^{(1)})\, \delta^3(\mathfrak{p}_k^{(1)} - \mathfrak{p}_k^{(2)}),$$

$$\sum_{\lambda} \int \frac{1}{2^n} \frac{d^{3n}p}{\omega_1(\mathfrak{p}_1) \cdots \omega_n(\mathfrak{p}_n)} \times \overline{\left(\begin{matrix} \nu_1 & M_1J_1 \\ & q_1\lambda_1 \end{matrix} \middle|\, \begin{matrix} \mathbf{ms} \\ \mathbf{p}\lambda \end{matrix}\right)} \left(\begin{matrix} \nu_2 & M_2J_2 \\ & q_2\lambda_2 \end{matrix} \middle|\, \begin{matrix} \mathbf{ms} \\ \mathbf{p}\lambda \end{matrix}\right) = \delta_{\nu_1\nu_2}\, \delta^4(q_1 - q_2) \cdot \delta_{j_1j_2} \cdot \delta_{\lambda_1\lambda_2}.$$

The analogous relations are satisfied by the C–G coefficients of $\tilde{P}$.

6. THE PHYSICAL MEANING OF THE QUANTITIES ON WHICH THE FUNCTIONS $\mathring{\varphi}$ DEPEND.

In this section, we want to find the physical meaning of the quantities λ and σ on which the functions $\mathring{\vartheta}_\lambda(q; j\lambda\sigma; \kappa)$ depend. For this purpose we consider in $\mathfrak{H}_n(\mathbf{m}, \mathbf{s})$ the representation of $\tilde{P}_+^\uparrow$ which is isomorphic to that one taking place in $\mathfrak{H}(m_l, s_l)$, namely

$$\begin{aligned} (U_l(a)f)(\mathbf{p}) &= e^{ip_la} f(\mathbf{p}), \\ (U_l(A)f)(\mathbf{p}) &= \mathfrak{D}^{s_l}(\rho_l(p_l, A))f(\Lambda^{-1}(A)_l\mathbf{p}), \end{aligned} \tag{6.1}$$

where $\Lambda^{-1}(A)_l\mathbf{p}$ is defined by

$$\Lambda^{-1}(A)_l\mathbf{p} = (p_1 \cdots p_{l-1}, \Lambda^{-1}(A)p_l, p_{l+1} \cdots p_n).$$

It follows immediately

$$U(A) = \prod_{l=1}^{n} U_l(A).$$

The infinitesimal translations in the representation U_l we call P_l.

Now let

$$v(p_l) = \alpha_l(p_l)v\alpha^{-1}(p_l)$$

be an element of the little group of $p_{\sim l}$.

It follows

$$\rho(p_l, v(p_l)) = v,$$

and therefore

$$(U_l(v(p_l))F)(\mathbf{p}) = \mathfrak{D}^{s_l}(r_l(\mathbf{p})vr_l^{-1}(\mathbf{p}))F(\mathbf{p}),$$

especially

$$(U_l(\alpha_l(p_l)r_l^{-1}(\mathbf{p})\rho(\chi)r_l(\mathbf{p})\alpha^{-1}(p_l))F)(\mathbf{p}) = \mathfrak{D}^{s_l}(\rho(\chi))F(\mathbf{p}). \quad (6.2)$$

If we go to infinitesimal transformations $\rho(\chi)$, we get from the last equation

$$W_\gamma^{(l)}\Lambda_3^\gamma(r_l(\mathbf{p})\alpha^{-1}(p_l))F(\mathbf{p}) = \lambda_l F(\mathbf{p}), \quad (6.3)$$

where $W_\gamma^{(l)}$ is the operator defined in (3.19) with the difference that $\mathfrak{J}$ and $\mathfrak{N}$ have to be replaced by $\mathfrak{J}^{(l)}$ and $\mathfrak{N}^{(l)}$, the infinitesimal Lorentz transformations in the representation U_l of $\tilde{P}_+^\dagger$.

From (6.3) we deduce immediately: The functions

$$\mathring{\varphi}_\lambda(q; J\lambda\sigma; \varkappa)$$

are eigenfunctions of the operator

$$W_\gamma^{(l)}\Lambda_3^\gamma(r_l(\mathbf{P})\alpha_l^{-1}(P_l)) \quad (6.4)$$

to the eigenvalue λ_l. Here the replacement

$$p_k \to P_k$$

in the expression $r_l(p)\alpha_l^{-1}(p_l)$ is defined by corresponding power series. No difficulties arise from the indefiniteness of the order of the different operators because the P_l commute. In general the operator (4) depends on all the momenta operators $P_1 \cdots P_n$.

However, if we make the WJW choice

$$\alpha^w(b_l) = \alpha[b_l],$$

and if we confine ourselves to the center-of-mass system, this operator becomes very simple. In this case we get

$$\alpha_l(p_l)r_l^{-1}(\mathbf{p}) = \rho\alpha[b_l] = \alpha[p_l]\rho(\chi(\mathbf{p})),$$

where $\chi(\mathbf{p})$ in general depends on all the momenta p_k. But if we introduce this expression in (6.2), the $\rho(\chi(\mathbf{p}))$ drops out and we get

$$(U_l(\alpha[p_l]\rho(\chi)\alpha^{-1}[p_l]F)(\mathbf{p}) = \mathfrak{D}^{s_l}(\rho(\chi))F(\mathbf{p}),$$

i.e., $\mathring{\vartheta}_\lambda(q_k; j\lambda\sigma; \varkappa)$ [where $q_k = (M, \mathbf{0})$] is eigenfunction to the value λ_l of the operator

$$W_\gamma^{(l)}\Lambda_3^\gamma(\alpha^{-1}[P_l]),$$

which is the helicity operator as we have seen in Sec. 3. Now let's start from the equation

$$(U(\rho_0\rho(\chi)\rho_0^{-1})\tilde{\varphi})(q_k, \rho_0\rho(\chi'), \varkappa) = \tilde{\varphi}(q_k, \rho_0\rho(\chi' + \chi), \varkappa).$$

Here

$$\rho_0 = \rho\rho(-\chi') \equiv \rho_0(\mathbf{p}),$$

where ρ is determined by $\mathbf{p}$ through the relation (4.21). If we set by definition

$$\tilde{\varphi}(q_k, \rho_0, \sigma, \varkappa) = \frac{1}{4\pi}\int_0^{4\pi} d\chi'\, \tilde{\varphi}(q_k, \rho_0\rho(\chi'), \varkappa),$$

we get

$$(U(\rho_0\rho(\chi)\rho_0^{-1})\tilde{\varphi})(q_k, \rho_0, \sigma, \varkappa) = e^{i\sigma\chi}\tilde{\varphi}(q_k, \rho_0, \sigma, \varkappa),$$

or, if we go to infinitesimal transformations $\rho(\chi)$,

$$\mathfrak{J}n(\mathbf{p})\tilde{\varphi}(q_k, \rho_0, \sigma, \varkappa) = \sigma\tilde{\varphi}(q_k, \rho_0, \sigma, \varkappa). \quad (6.5)$$

Here $n(\mathbf{p})$ is the vector we get if we apply the rotation

$$\Lambda(\rho_0^{-1}(\mathbf{p}))$$

to the unit vector pointing in the direction of the positive 3-axis. If we again replace the momentum p_l in $n(\mathbf{p})$ by the corresponding operator in the representation U_l, we may complete the transformation (5.23) in (6.5) without difficulties. We then observe that σ in $\mathring{\vartheta}(q; j\lambda\sigma; \varkappa)$ has the meaning of the total angular momentum in the center-of-mass system relative to an axis which is connected with the momenta of the n particles in a definite way. The situation is similar to the case of a top: σ corresponds to the angular momentum relative to the figure axis and λ to the one relative to an axis fixed in space.

7. APPLICATION TO THE S-MATRIX

We consider the S-matrix of a process

$$a_1^{(1)} + a_2^{(1)} + \cdots + a_{n_1}^{(1)} \to a_1^{(2)} + a_2^{(2)} + \cdots + a_{n_2}^{(2)}. \quad (7.1)$$

Let $m_k^{(i)}$ be the mass, $s_k^{(i)}$ the spin of the particle, $a_k^{(i)}$ $p_k^{(i)}$ its momentum, and $\lambda_k^{(i)}$ the third component of its spin (for definition of this quantity in a coordinate-system in which the momentum $p_k^{(i)}$ is measured, see Sec. 3). The S-matrix maps the $\Pi_{k=1}^{n_1}(2s_k^{(1)} + 1)$-dimensional spin space of the incoming particles on the $\Pi_{k=1}^{n_2}(2s_k^{(2)} + 1)$-dimensional spin space of the outgoing particles, and is, at the same time, a Lorentz-invariant generalized function of the variables

$$\mathbf{p}^{(i)} = (p_1^{(i)} \cdots p_{n_i}^{(i)})\ (i = 1, 2) : S(\mathbf{p}^{(1)}, \mathbf{p}^{(2)}).$$

From translation invariance, one derives that the support of S is the submanifold on $V^+(\mathbf{m}^{(1)}) \otimes$

$V^+(\mathbf{m}^{(2)})$ defined by

$$q^{(1)} = q^{(2)}; \qquad q^{(i)} = \sum_{k=1}^{n_i} p_k^{(i)}.$$

Lorentz invariance implies

$$\mathfrak{D}^{s^{(1)}}(\rho(\mathbf{p}^{(1)}, A))S(\Lambda^{-1}(A)\mathbf{p}^{(1)}, \Lambda^{-1}(A)\mathbf{p}^{(2)}) \times \mathfrak{D}^{s^{(2)}}(\rho^{\dagger}(\mathbf{p}^{(2)}, A)) = S(\mathbf{p}^{(1)}, \mathbf{p}^{(2)}).$$

If we now define

$$T(\mathbf{p}^{(1)}, \mathbf{p}^{(2)})\ \delta^4(q^{(1)} - q^{(2)}) = \mathfrak{D}^{s^{(1)}}(\mathbf{p}^{(1)})S(\mathbf{p}^{(1)}\mathbf{p}^{(2)})\ \mathfrak{D}^{s^{(2)}\dagger}(\mathbf{p}^{(2)}), \qquad (7.2)$$

where we subtract a $\delta^4(q^{(1)} - q^{(2)})$ function on the right side, if the process under consideration is an elastic scattering, Lorentz invariance reads

$$T(\mathbf{p}^{(1)}, \mathbf{p}^{(2)}) = T(\Lambda^{-1}(A)\mathbf{p}^{(1)}, \Lambda^{-1}(A)\mathbf{p}^{(2)}).$$

Hence the inverse of (2) is

$$S(\mathbf{p}^{(1)}, \mathbf{p}^{(2)}) = \delta^4(q^{(1)} - q^{(2)})\hat{T}(\mathbf{p}^{(1)}, \mathbf{p}^{(2)}), \qquad (7.3)$$

where

$$\hat{T}(\mathbf{p}^{(1)}, \mathbf{p}^{(2)}) = \mathfrak{D}^{s^{(1)}\dagger}(\mathbf{p}^{(1)})T(\mathbf{p}^{(1)}, \mathbf{p}^{(2)})\ \mathfrak{D}^{s^{(2)}}(\mathbf{p}^{(2)}) \qquad (7.4)$$

gives a representation of the S-matrix as a sum of the covariants of $\tilde{P}_+^\uparrow$, each multiplied by an invariant amplitude $T(\mathbf{p}^{(1)}, \mathbf{p}^{(2)})$.

Because of $q - q^{(1)} = q^{(2)}$, we may write, according to (5.16),

$$T(\mathbf{p}^{(1)}, \mathbf{p}^{(2)}) = \tau_1(q; \rho^{(1)}, \rho^{(2)}; \boldsymbol{\kappa}^{(1)}\boldsymbol{\kappa}^{(2)}) \times A_{n_1}^{-\frac{1}{2}}(\boldsymbol{\kappa}^{(1)})A_{n_2}^{-\frac{1}{2}}(\boldsymbol{\kappa}^{(2)}).$$

(From the rigorous mathematical point of view, this transformation would need a justification depending on the test-function space, which is used to define the general functions.) Lorentz invariance now reads

$$\tau_1(\Lambda^{-1}(A)q; \rho^{-1}(q, A)\rho^{(1)}, \rho^{-1}(q, A)\rho^{(2)}; \boldsymbol{\kappa}^{(1)}\boldsymbol{\kappa}^{(2)}) = \tau_1(q; \rho^{(1)}, \rho^{(2)}; \boldsymbol{\kappa}^{(1)}, \boldsymbol{\kappa}^{(2)}).$$

Choosing

$$A = \alpha(q)\rho^{(2)},$$

we get

$$\rho(q, A) = \rho^{(2)}, \qquad \Lambda^{-1}(A)q = (M, \mathbf{0}) = q_k,$$

and therefore

$$\tau_1(q; \rho^{(1)}, \rho^{(2)}; \boldsymbol{\kappa}^{(1)}, \boldsymbol{\kappa}^{(2)}) = \tau_1(q_k; \rho, \mathbb{1}; \boldsymbol{\kappa}^{(1)}, \boldsymbol{\kappa}^{(2)}) \equiv \bar{\tau}(M, \rho; \boldsymbol{\kappa}^{(1)}\boldsymbol{\kappa}^{(2)}),$$

where $\rho = (\rho^{(2)})^{-1}\rho^{(1)}$ and the sign $\equiv$ means "equal per definition". It is clear that

$$M(\boldsymbol{\kappa}^{(1)}) = M(\boldsymbol{\kappa}^{(2)}) = M = (q^2)^{\frac{1}{2}}.$$

Performing the transformation

$$\tau(MJ; \sigma_1\boldsymbol{\kappa}^{(1)}; \sigma_2\boldsymbol{\kappa}^{(2)}) = (2J+1)^{\frac{1}{2}} \int d\rho\ \mathfrak{D}^J_{\sigma_2\sigma_1}(\rho)\bar{\tau}(M, \rho; \boldsymbol{\kappa}^{(1)}\boldsymbol{\kappa}^{(2)}),$$

we finally get "the generalized partial wave decomposition"

$$T(\mathbf{p}^{(1)}\mathbf{p}^{(2)}) = A_{n_1}^{-\frac{1}{2}}(\boldsymbol{\kappa}^{(1)})A_{n_2}^{-\frac{1}{2}}(\boldsymbol{\kappa}^{(2)}) \times \sum_{J\sigma_1\sigma_2} (2J+1)^{\frac{1}{2}}\tau(MJ; \sigma_1\boldsymbol{\kappa}^{(1)}, \sigma_2\boldsymbol{\kappa}^{(2)})\ \bar{\mathfrak{D}}^J_{\sigma_2\sigma_1}(\rho). \qquad (7.5)$$

In the case $n_1 = n_2 = 2$, the substitution

$$\rho_1 \to \rho_1\rho(\chi_1), \qquad \rho_2 \to \rho_2\rho(\chi_2),$$

which leads to

$$\rho \to \rho(-\chi_2)\rho\rho(\chi_1)$$

must leave all the equations invariant. It is easily seen that this leads to the relations

$$\lambda_1^{(1)} - \lambda_2^{(1)} = \sigma_1, \qquad \lambda_1^{(2)} - \lambda_2^{(2)} = \sigma_2;$$

if $n_1 = 2$, $n_2 \geq 3$, only the first of these relations is valid, if $n_1 \geq 3$, $n_2 = 2$, only the second one. In any case (also in the case $n_1 \geq 3$, $n_2 \geq 3$), the summation in (7.5) has to be carried out over those J, for which

$$2J \equiv 2\,|s^{(1)}| \equiv 2\,|s^{(2)}| \mod 2.$$

According to (5.29c), invariance with respect to space inversion has the consequence

$$\tau_{\lambda^{(1)}\lambda^{(2)}}(MJ; \sigma_1\boldsymbol{\kappa}^{(1)}, \sigma_2\boldsymbol{\kappa}^{(2)}) = \eta(\Pi)\tau_{-\lambda^{(1)},-\lambda^{(2)}}(M, J; -\sigma_1\Pi\boldsymbol{\kappa}^{(1)}, -\sigma_2\Pi\boldsymbol{\kappa}^{(2)}),$$

where

$$\eta(\Pi) = \mathfrak{n}_\Pi^{(1)}\bar{\mathfrak{n}}_\Pi^{(2)}(-1)^{|s^{(1)}|-|\lambda^{(1)}|+|s^{(2)}|-|\lambda^{(2)}|-2J+\sigma_1+\sigma_2}.$$

All the results of this section can also be derived using the fact that, according to group-theoretical considerations, the S-matrix, being an operator-isomorphism with respect to the group $\tilde{P}_+^\uparrow$ which maps $\mathfrak{H}_{n_1}(\mathbf{m}^{(1)}, \mathbf{s}^{(1)})$ onto $\mathfrak{H}_{n_2}(\mathbf{m}^{(2)}, \mathbf{s}^{(2)})$, must have the representation

$$\int d^4q \sum_{J,\lambda} \sum_{\nu_1\nu_2} \overline{\left(\begin{matrix}\nu_1 & MJ \\ & q\lambda\end{matrix}\middle|\begin{matrix}\mathbf{m}^{(1)}\mathbf{s}^{(1)} \\ \mathbf{p}^{(1)}\lambda^{(1)}\end{matrix}\right)}\left(\begin{matrix}\nu_2 & MJ \\ & q\lambda\end{matrix}\middle|\begin{matrix}\mathbf{m}^{(2)}\mathbf{s}^{(2)} \\ \mathbf{p}^{(2)}\lambda^{(2)}\end{matrix}\right) \times \langle\nu_1|\ \tau(JM)\ |\nu_2\rangle, \qquad (7.6)$$

where the C–G coefficients

$$\left(\begin{matrix}\nu & MJ \\ & q\lambda\end{matrix}\middle|\begin{matrix}\mathbf{ms} \\ \mathbf{p}\lambda\end{matrix}\right)$$

are defined in (5.54).

Indeed if we set

$$\tau_{\lambda^{(1)}\lambda^{(2)}}(MJ;\sigma_1\kappa^{(1)},\sigma_2\kappa^{(2)}) = \sum_{\nu_1,\nu_2} \overline{\langle \nu_1|\, H\,|\sigma_1\lambda^{(1)},\kappa^{(1)}\rangle} \times \langle \nu_1|\,\tau(JM)\,|\nu_2\rangle\langle \nu_2|\, H\,|\sigma_2\lambda^{(2)},\kappa^{(2)}\rangle,$$

(7.6) is in complete agreement with the relation we obtain if we combine (7.5) and (7.2).

In an analogous way, we may define a decomposition of the S-matrix with respect to the group $\tilde{P}$ using the corresponding C–G coefficients given in (5.54) and (5.55).

Finally, we remark that, according to (3.23) in the case $m_k^{(i)} > 0$, the connection between our $\hat{T}(\mathbf{p}^{(1)}, \mathbf{p}^{(2)})$-matrices and the $\sum_{k,i} \oplus (s_k^{(i)}, 0)$-tensors (with respect to $L_+^\uparrow$) of Stapp[11] and Hepp[12] is given by

$$\mathfrak{D}^s(\alpha(\mathbf{p}^{(1)}))\hat{T}(\mathbf{p}^{(1)},\mathbf{p}^{(2)})\,\mathfrak{D}^s(\epsilon\alpha^T(\mathbf{p}^{(2)})) = M(\mathbf{p}^{(1)},\mathbf{p}^{(2)})$$

$[M(\mathbf{p}^{(1)},\mathbf{p}^{(2)}) = \sum_{k,i} \oplus (s_k^{(i)}, 0)$-tensor].

Notes added before print:

A. After having finished the preprint of this paper it was learned by the author that two other authors have written a paper on a similar subject: P. Moussa and R. Stora, "Some Remarks on the Product of Irreducible Representations of the Inhomogeneous Lorentz Group (Preprint, edited at Centre d'Etudes Nucléaire de Saclay, Gif-sur-Yvettes, France.)

R. Stora has called to this author's attention their paper and their method of reducing the product representation applied in their work: the method of induced representations, presented in Mackey's book. [G. W. Mackey, *The Theory of Group-Representations* (The University of Chicago Press, Chicago, 1955)] It turned out that the method applied in the present paper is strongly related to Mackey's general scheme.

I wish to thank Dr. Stora for the correspondence concerning this point.

B. A reader looking through the present paper without studying it in detail might get the impression that, besides the helicity coupling scheme, no other scheme has been treated, which wouldn't justify the term "most general" in the title. That's why the author wants to emphasize that the C–G coefficients (5.54) are quite general and especially that it includes also the l–s coupling scheme of Joos[5] and Macfarlane.[1]

To show this in some detail, assume $m_i > 0$ $(i = 1 \cdots n)$ and replace (4.44) by

$$\alpha_i^u(b_i) = \alpha\{b_i\}. \tag{B1}$$

Then (5.54) is formally unchanged, the only difference being that $\mathfrak{D}^s_{\lambda',\lambda}(\mathbf{p})$ is defined with the help of

$$r_i\{\mathbf{p}\} = \alpha^{-1}\{b_i\}a^{-1}(\mathbf{p})\alpha_i(p_i) \tag{B2}$$

instead of

$$r_i[\mathbf{p}] = \alpha^{-1}[b_i]a^{-1}(\mathbf{p})\alpha_i(p_i). \tag{B3}$$

In addition, we choose $\langle \nu|\, H\,|\lambda', \sigma, \kappa\rangle$ in (5.54) to be

$$\left(\begin{matrix}s & l \\ \lambda' & \tau\end{matrix}\middle|\begin{matrix}J \\ \sigma\end{matrix}\,\zeta\right), \qquad \nu = (l, \tau, \zeta), \tag{B4}$$

i.e., the C–G coefficient of $SU(2)$ which composes $s_1 \cdots s_n$, l to J. $\tau = \sigma - \sum_{i=1}^n \lambda'_i$; ζ is conveniently chosen to be $\zeta = (s_{12}, s_{123\ldots}s_{12\ldots n})$, where $s_{12\ldots k}$ is obtained by composing $s_1 \cdots s_k$. Because of the relation

$$r_i\{\mathbf{p}\} = \rho[b_i]r_i[\mathbf{p}]$$

following from (B2), (B3), and (2.21), the procedure described so far is equivalent to directly choosing

$$\langle \nu|\, H\,|\lambda', \sigma, \kappa\rangle$$

in (5.54) to be

$$\sum_{\substack{\lambda'' \\ \sum_{i=1}^n \lambda_i'' + \tau = \sigma}} \mathfrak{D}^s_{\lambda'',\lambda'}(\rho[(\mathbf{b})])\left(\begin{matrix}s & l \\ \lambda'' & \tau\end{matrix}\middle|\begin{matrix}J \\ \sigma\end{matrix}\,\zeta\right).$$

It is easy to see that this choice leads to the C–G coefficients of Joos[5] and Macfarlane[1] in the case $n = 2$. Notice that the choice (1) implies $y(b_1, \chi) = y(b_2, \chi) = \rho(-\chi)$ [Compare (4.40)], such that in the case $n = 2$: $\lambda_1 + \lambda_2 = \sigma$ and therefore $\tau = 0$. For $n \geq 3$, however, it leads to a simpler expression than the one obtained by Macfarlane who essentially constructs the C–G coefficients for $n \geq 3$ according to the "Dalitz-scheme" (1), (2) $\rightarrow$ (12); (12)(3) $\rightarrow$ (123), etc. using only C–G coefficients for $n = 2$.

ACKNOWLEDGMENTS

I wish to thank Professor Williamson for the critical study of this paper. Also, some clarifying discussions with Dr. K. Hepp and S. Albeverio at an earlier stage of this work are gratefully acknowledged.

Quantum-Mechanical Extension of the Lebowitz–Penrose Theorem on the Van Der Waals Theory*

ELLIOTT LIEB†
Belfer Graduate School of Science, Yeshiva University, New York, New York
(Received 1 November 1965)

Recently Lebowitz and Penrose gave a rigorous derivation of the van der Waals–Maxwell theory of the liquid–vapor transition, and showed how the Maxwell equal area-rule could be obtained from a proper statistical mechanical calculation. Their results are quite general—being valid in any number of dimensions and for a broad class of pair potentials—but they were proved only for classical mechanics. In the present work we extend the proof to quantum systems with any statistics—Boltzmann, Bose, or Fermi. One corollary of this extended theorem is a model of a Bose gas with a first-order phase transition.

INTRODUCTION

IN 1959 Kac[1] showed how to calculate exactly the partition function of a one-dimensional system[2] of particles interacting via the pair potential

$$\begin{aligned} v(x) &= \infty && \text{for} \quad |x| < r_0 \\ &= \tfrac{1}{2}\alpha\gamma \exp(-\gamma|x|) && \text{for} \quad |x| > r_0, \end{aligned} \tag{1.1}$$

where γ and α are arbitrary constants, but with $\gamma > 0$. He also proved the expected result that such a system has no phase transition when α and γ are finite.

Shortly thereafter Baker,[3] who independently considered the same type of interaction for an Ising spin system, pointed out that while there is no phase transition when γ is finite, one does obtain a phase transition (if $\alpha < 0$) by taking the limit $\gamma \to 0$ *after* taking the bulk limit $N \to \infty$. (Plainly, taking the limit $\gamma \to 0$ before the bulk limit would not be interesting because only the hard-core potential would remain.) Physically, this means considering a sequence of systems with ever increasing range of interaction. The range of interaction is always vanishingly small compared to the size of the system, and in each case the isotherms are continuously differentiable (no phase transition). Nevertheless, the limiting isotherms have flat portions characteristic of a first-order vapor–liquid transition. The properties of these isotherms, especially near the critical point, as well as the manner in which the limiting isotherms are approached as $\gamma \to 0$ was investigated in detail by several authors.[4]

The free energy of the limiting system turns out to be merely that of the familiar van der Waals theory, namely

$$a(\rho) = \text{C.E. } \{f(\rho)\}, \tag{1.2a}$$

where

$$f(\rho) = kT\rho[\ln(\rho/1-\rho r_0) - \tfrac{1}{2}\ln(mkT/2\pi\hbar^2) - 1] + \tfrac{1}{2}\alpha\rho^2, \tag{1.2b}$$

and where $a(\rho)$ is the free energy *per unit volume* and ρ is the particle density, which is necessarily restricted to be less than r_0^{-1}. The symbol C.E. $\{f(\rho)\}$ in (1.2a) means the following: For a physically stable system $a(\rho)$ must be a convex function of ρ in order that the pressure, P, defined by

$$P = \rho^2(\partial/\partial\rho)[\rho^{-1}a(\rho)], \tag{1.3}$$

be a nondecreasing function of ρ. $f(\rho)$ in (1.2b) may not have this convexity property, in which case C.E. $\{f(\rho)\}$ (or *convex envelope*) is defined to be the greatest convex function which is everywhere less than or equal to $f(\rho)$. In other words, the graph of C.E. $\{f(\rho)\}$ consists partly of convex segments of $f(\rho)$ and partly of straight lines, each of which is tangent to two points of the $f(\rho)$ graph. In terms of the pressure, (1.3), this double tangent construction is exactly equivalent to drawing horizontal straight lines through any loops in the $P(\rho^{-1})$ curve, using the equal-area rule of Maxwell. A first-order phase transition is the result.

There are two noteworthy aspects of Eq. (1.2). The first is that the convex envelope, or double tangent construction is not imposed *ad hoc*, but is

* This work was supported by the U. S. Air Force Office of Scientific Research at Yeshiva University, Grant No. AF–AFOSR–713–64 and AFOSR–508–66.

† Present address: Physics Dept., Northeastern University, Boston, Mass.

[1] M. Kac, Phys. Fluids 2, 8 (1959). Actually, Kac only considered the attractive case, $\alpha < 0$. For a discussion of the repulsive case see D. Newman, J. Math. Phys. 5, 1153 (1964).

[2] For a review of exactly soluble one-dimensional statistical mechanical problems see E. Lieb and D. Mattis, *Mathematical Physics in One-Dimension* (Academic Press Inc., New York, 1966).

[3] G. Baker, Phys. Rev. 122, 1477 (1961); 126, 2071 (1962).

[4] M. Kac, G. Uhlenbeck, and P. Hemmer, J. Math. Phys. 4, 216, 229 (1963); 5, 60 (1964); P. Hemmer, *ibid.*, p. 75; E. Helfand, *ibid.*, p. 127.

derived directly from the pair potential (1.1) without any additional assumptions. It is in this sense that one has a rigorous derivation of the van der Waals–Maxwell theory, although it must be admitted that the system in question is only a limiting system and is, moreover, one dimensional.

The second point concerns the form of $f(\rho)$ itself, Eq. (1.2b). As we remarked before, $f(\rho)$ was originally calculated by first finding the partition function for finite γ, and then passing to the limit $\gamma \to 0$. While this procedure requires a good deal of sophisticated mathematical analysis, the final result, (1.2b) could have been easily foreseen on the basis of molecular field theory. The first term in (1.2b) is the well-known configurational free energy of a pure hard-core gas. The third term $\frac{1}{2}\alpha\rho^2$, could be obtained by observing that if γ is very small (long-range forces) then each particle "feels" only the average value of the long-range potential, assuming that the fluctuations in particle density have a range small compared to γ^{-1}. Because of this homogeneity assumption, molecular field theory would give $a(\rho) = f(\rho)$ and not C.E. $\{f(\rho)\}$, i.e., molecular field theory would give a van der Waals loop. Where molecular field theory goes wrong is that in the transition region fluctuations are indeed macroscopically large.

Nevertheless, molecular field theory is plainly on the right track and it would seem that it should be possible to prove the analog of (1.2) for any system in any number of dimensions given that the pair potential consists of a suitable short range part, whose free energy is assumed known, and a long range part whose range tends to infinity. In other words, it ought to be possible to avoid the detailed analysis for finite γ, which can be carried out only for the special potential (1.1) in one dimension. That this can be done in a completely rigorous manner is the content of a recent paper by Lebowitz and Penrose,[5] although van Kampen[6] had earlier given a heuristic derivation of what we will here refer to as the Lebowitz–Penrose theorem. Unfortunately LP could prove their theorem only for classical systems, and it is the aim of the present work to extend the theorem to quantum systems with any statistics—Boltzmann, Bose, or Fermi.

In Sec. II we give a precise statement of the theorem and point out what additional lemmas are needed in order for the LP proof to be valid for quantum systems. At this point the reader may justifiably ask what relevance this theorem has to the statistical mechanics of real systems. On the one hand the theorem is of little real use because the fundamental problem of statistical mechanics is to understand why short-range forces behave like ultra-long-range forces, in the sense that they give rise to phase transitions. The analysis of systems with ultra-long-range forces sheds little light on this central problem. On the positive side, however, there are two things that can be said in favor of the theorem. The first is that in a subject as complicated as statistical mechanics, the number of nontrivial, rigorous statements that can be proved is small indeed, and the addition of one more can possibly be beneficial. The second is that one can conceivably use the $\gamma \to 0$ theorem as the first step in an expansion in γ for real systems with nonzero, but small γ.

II. STATEMENT OF THEOREM AND OUTLINE OF PROOF

Consider a system of N, ν-dimensional particles ($\nu = 1, 2, 3, \cdots$) each constrained to lie in a cubic region of volume Ω in ν-dimensional space. The particles each have mass m and interact via a pair potential given by (following the notation of LP)

$$v(\mathbf{r}) = q(\mathbf{r}) + \gamma^{\nu}\varphi(\gamma\mathbf{r}), \tag{2.1}$$

where $\mathbf{r} = \mathbf{r}_i - \mathbf{r}_j$ is the relative coordinate of any two particles. Let $Z(N, \Omega, \gamma)$ be the partition function of the system (classical or quantum with any statistics), whence

$$A(N, \Omega, \gamma) \equiv -kT \ln Z(N, \Omega, \gamma) \tag{2.2}$$

is the free energy of the system. We pass to the bulk (or thermodynamic) limit by considering a sequence of cubes of increasing volume together with an increasing number of particles such that $\rho = N\Omega^{-1}$ is constant, and define the limiting free energy per unit volume by

$$a(\rho, \gamma) \equiv \lim_{\Omega\to\infty} \Omega^{-1}A(\rho\Omega, \Omega, \gamma). \tag{2.3}$$

The potentials q and φ (the short- and long-range parts, respectively) are subject to certain conditions to be stated below, and these guarantee that the limit in (2.3) exists for any finite, positive γ, and, in the quantum mechanical case, for vanishing boundary conditions at least. In addition to these conditions, it is assumed that φ is Riemann-integrable,

$$\int \varphi(\mathbf{r})\, d\mathbf{r} = \alpha. \tag{2.4}$$

[5] J. Lebowitz and O. Penrose, "A Rigorous Treatment of the Van der Waals-Maxwell Theory of the Vapor-Liquid Transition," J. Math. Phys. 7, 98 (1966). References to this paper will be simply designated LP.

[6] N. van Kampen, Phys. Rev. 135, A362 (1964).

We also assume that the corresponding free energies [denoted by $A^0(N, \Omega)$ and $a^0(\rho)$, respectively] are *known* when $\varphi(\mathbf{r}) \equiv 0$ (the base system) and that $a^0(\rho)$ is independent of boundary conditions in the quantum-mechanical case [see (q3) below].

We now take the limit $\gamma \to 0$ (ultra-long-range forces) and define

$$a(\rho, 0+) = \lim_{\gamma \to 0+} a(\rho, \gamma). \tag{2.5}$$

Our aim is to prove the following

Theorem: Subject to the conditions on q and φ stated below:

(1) For all statistics and all boundary conditions $a(\rho, 0+)$ exists and is equal to

$$a(\rho, 0+) = \text{C.E.}\ \{a^0(\rho) + \tfrac{1}{2}\alpha\rho^2\}, \tag{2.6}$$

where $a^0(\rho)$ is the bulk limiting free energy per unit volume for the base system with the appropriate statistics—Boltzmann, Bose, or Fermi.

(2) For all statistics and for *vanishing* boundary conditions, the limiting pressure is equal to the pressure derived from the limiting free energy, namely

$$\begin{aligned} \lim_{\gamma \to 0+} P(\gamma) &\equiv \lim_{\gamma \to 0+} \rho^2(\partial/\partial\rho)[\rho^{-1}a(\rho, \gamma)] \\ &= \rho^2(\partial/\partial\rho)[\rho^{-1}a(\rho, 0+)]. \end{aligned} \tag{2.7}$$

We list first the conditions [in addition to (2.4)] on the *long-range* part of the potential. These are the same as in LP.

(φ1). There exist positive constants D_3 and ϵ such that

$$|\varphi(\mathbf{r})| < D_3 r^{-\nu-\epsilon} \quad \text{for all} \quad \mathbf{r}. \tag{2.8}$$

(φ2). $\varphi(\mathbf{r})$ is continuous at $\mathbf{r} = 0$.

In addition φ satisfies one of the following three conditions:

(φ3a). $$\Phi(\mathbf{p}) \geq 0 \quad \text{all} \quad \mathbf{p}, \tag{2.9}$$

where $\Phi(\mathbf{p})$ is the ν-dimensional Fourier transform of φ.

(φ3b). $$\varphi(\mathbf{r}) \leq 0 \quad \text{for all} \quad \mathbf{r}. \tag{2.10}$$

(φ3c). This condition is discussed in detail in Sec. V of LP. Of the several possibilities raised there, only two are sufficient to establish the theorem unambiguously. One is a repetition of (φ3a) above, while the other is

$$\min_{\mathbf{p}} \Phi(\mathbf{p}) = \Phi(0). \tag{2.11}$$

It will be noted that (φ3b) is really a special case of (φ3c), since (2.10) implies (2.11).

The classical conditions on the *short-range* part of the potential, as given in LP, are easier to state:

(q1). (Hard core) $q(\mathbf{r}) = +\infty$ for $r < r_0$. (2.12)

(q2). There exist positive constants D_2 and ϵ such that

$$|q(\mathbf{r})| < D_2 r^{-\nu-\epsilon} \quad \text{for} \quad r > r_0. \tag{2.13}$$

To these we append a purely quantum-mechanical condition which, most likely, is already implied by (q1) and (q2) above, although no proof of this implication has yet been found. In order to specify the quantum-mechanical problem it is necessary to impose homogeneous boundary conditions on the walls of the container, Ω. There are three which are customarily used: (1) vanishing ($\Psi = 0$) boundary conditions (VBC); (2) normal ($\mathbf{n}\cdot\nabla\Psi = 0$) boundary conditions (NBC); (3) periodic boundary conditions (PBC). We shall designate the finite volume free energy in the three cases by $A_1(N, \Omega)$, $A_2(N, \Omega)$ and $A_3(N, \Omega)$, respectively. It is easy to show that (see Lemma 4 below),

$$A_1(N, \Omega) \geq A_3(N, \Omega) \geq A_2(N, \Omega). \tag{2.14}$$

Now, Fisher[7] has shown that conditions (q1) and (q2) ensure that a_1^0 the bulk limit of A_1^0 exists for any statistics, but no proof has been given that the bulk limits of A_2^0 or even A_3^0 exist. We shall assume that these limits not only exist but that they are independent of boundary conditions, namely

(q3). $$\lim_{\Omega \to \infty} A_2^0(\rho\Omega, \Omega) = a_1^0(\rho), \tag{2.15}$$

for all statistics—Boltzmann, Bose, or Fermi.

If condition (q3) were not satisfied the base system would indeed be pathological, for it would mean that the so-called surface effects would have a macroscopic effect. When the base system is the one-dimensional hard-core gas (1.1), as was originally used by Kac and Baker, a direct calculation shows that (q3) is satisfied. We shall return to a discussion of the quantum mechanical version of (1.2) later. Suffice it to say that condition (q3) is a condition imposed, along with other conditions, on the *base system alone* whose free energy we assume can be calculated. It is not a condition imposed on the combined system. Indeed, part of the theorem is to show that (2.6) is true irrespective of boundary conditions.

Owing to the inequality (2.14), the theorem will be true for PBC once it is established for VBC

[7] M. Fisher, Arch. Ratl. Mech. Anal. 17, 377 (1964).

and NBC. We shall therefore ignore PBC henceforth, except in Lemma 5 below.

In order to prove the theorem we need several quantum-mechanical lemmas whose proof will be given in Sec. III. In this section we shall merely state the lemmas and then show how they, together with the analysis in LP, prove the theorem. We shall be concerned with the Hamiltonian

$$H = -\frac{\hbar^2}{2m}\sum_1^N \nabla_i^2 + U(\mathbf{r}_1, \cdots, \mathbf{r}_N), \qquad (2.16)$$

where ∇ is the ν-dimensional Laplacian and U is an arbitrary potential. The $N\nu$-dimensional domain in which H operates will be denoted by D.

Lemma 1. Let $E_1 \leq E_2 \leq \cdots$ be the eigenvalue spectrum of H with VBC on D. Let $D_1, \cdots, D_\alpha$ be connected, disjoint, open subdomains of D. Let e_i^j (with $e_i^j \leq e_{i+i}^j$) be the ordered eigenvalues of H in D_j with VBC. Let e_i be the ordered union of all the e_i^j (with $e_i < e_{i+i}$). Then $e_i \geq E_i$ for all i.

As an immediate corollary of Lemma 1 we have the following:

Corollary 1.

$$Z_1 \geq \sum_{j=1}^{\alpha} Z_{1j} \qquad (2.17)$$

where Z_1 is the partition function in D and Z_{1j} is the partition function in subdomain D_j, all with VBC.

As it stands, Lemma 1 and its corollary is too general to permit us to discuss quantum statistics, because the potential U and the domains D, D_1, $\cdots$, D_α have no particular symmetry (permutation) properties. We list now the requisite symmetry conditions to which the succeeding lemmas will refer:

(S1). $U(\mathbf{r}_1, \cdots, \mathbf{r}_N)$ is a symmetric function of the $\mathbf{r}_i$.

(S2). The domain D is permutation invariant, i.e., if the $N\nu$-dimensional point $\mathbf{R} = (\mathbf{r}_1, \cdots \mathbf{r}_N)$ is in D, then so is any permutation of this point, $P\mathbf{R} = (\mathbf{r}_{p1}, \cdots, \mathbf{r}_{pN})$ where $p1, \cdots, pN$ is a permutation, P, of $1, \cdots, N$.

One way to achieve this property, but by no means the only way, is if $D = \Omega_1 \times \Omega_2 \times \cdots \times \Omega_N$, where Ω_i is the ν-dimensional region $\mathbf{r}_i \in \Omega$, and Ω is any region (possibly a cube, but in any event independent of i) in ν-space. We shall refer to this latter possibility by writing $D = \Omega^N$.

(S3). The subdomains $D_1, \cdots, D_\alpha$ are permutation copies of each other.

By this is meant that for every permutation P, if $\mathbf{R} \in D_i$, then for some j, $P\mathbf{R} \in D_j$ (with j possibly equal to i). Moreover if $P\mathbf{R} \in D_j$ for some $\mathbf{R} \in D_i$ and some j, then $P\mathbf{R} \in D_j$ for *all* $\mathbf{R} \in D_i$. As a result of this definition we can write $PD_i = D_j$. A few examples with $N = 2$ might make this definition clearer. Suppose $R = \Omega_1 \times \Omega_2$, where Ω is the sum of two parts ω_1 and ω_2. If D_1 is defined as $\mathbf{r}_1 \in \omega_1$, $\mathbf{r}_2 \in \omega_1$ then $PD_1 = D_1$ for all P and it would then not be necessary to have any subdomain other than D_1. We could, however, also include the region $D_2 : \mathbf{r}_1 \in \omega_2$, $\mathbf{r}_2 \in \omega_2$ so that $PD_2 = D_2$ for all P. A second possibility might be $D_3 : \mathbf{r}_1 \in \omega_1$, $\mathbf{r}_2 \in \omega_1 + \omega_2$. This is not allowed because for some points $\mathbf{R}$, (namely, $\mathbf{r}_1 \in \omega_1, \mathbf{r}_2 \in \omega_1$) $P\mathbf{R} \in D_3$ for all P, while for other points (namely, $\mathbf{r}_1 \in \omega_1$, $\mathbf{r}_2 \in \omega_2$) $T\mathbf{R}$ is not in D_3 (here T is the transposition permutation). Still a third possibility is $D_4 : \mathbf{r}_1 \in \omega_1$, $\mathbf{r}_2 \in \omega_2$. For this it is *necessary* to include the region $D_5 : \mathbf{r}_1 \in \omega_2$, $\mathbf{r}_2 \in \omega_1$, so that $TD_4 = D_5$. In addition to D_4 and D_5 we could also include the domain D_1, defined above, whence $TD_1 = D_1$.

As a consequence of condition (S3) we observe that the domains $D_1, \cdots, D_\alpha$ provide a representation of the permutation (or symmetric) group S_N. That is to say, with each permutation P, we can associate an $\alpha \times \alpha$ matrix $\bar{P}$, whose (i, j) element is 1 if $PD_j = D_i$ and is zero otherwise. Consequently, $\bar{P}\,\bar{Q} = (\bar{P}\bar{Q})$. Furthermore, we can partition the domains into equivalence classes according to the rule:

D_i and D_j are in the same class C if and only if there exists a permutation P such that $PD_i = D_j$. The symbol $C(i)$ will denote the class of D_i. If we denote by f_{ij} the number of distinct permutations such that $PD_i = D_j$, and denote $s_i = f_{ii} \geq 1$, then the f_{ij} have the following properties:

$$\sum_{j=1}^{\alpha} f_{ij} = N! \qquad (2.18)$$

$$\begin{aligned} f_{ij} &= s_i = s_j \quad \text{if} \quad C(i) = C(j), \\ &= 0 \qquad\quad \text{if} \quad C(i) \neq C(j). \end{aligned} \qquad (2.19)$$

From (2.18) and (2.19) we have

$$n_i s_i = N!, \qquad (2.20)$$

where n_i = number of domains in $C(i)$.

It is also clear that the set of permutations, G_i, which leave D_i invariant (i.e., $P \in G_i$ if $PD_i = D_i$) is a subgroup of S_N. Note that G_i is not necessarily an obvious subgroup, i.e., the full symmetric on a

subset of the indices $1, \cdots, N$. It could be any subgroup of S_N. Also, if D_i and D_j belong to the same class, then $G_j = PG_iP^{-1}$, where P is any permutation such that $PD_i = D_j$.

Turning now to the definition of the partition functions, we will use the superscripts $*$, $+$, and $-$, respectively, to denote Boltzmann, Bose, or Fermi statistics. As before, the subscript 1 or 2, respectively, will denote vanishing or normal boundary conditions, but we shall omit these subscripts on statements which do not depend upon boundary condition.

The partition function is defined in terms of energy eigenvalues. For the full domain D the ordered spectra will be denoted by E_i^*, E_i^+, and E_i^- (with $E_i \leq E_{i+1}$). They are defined in the usual way: E^* is the full spectrum, E^+ is the spectrum of symmetric eigenfunctions, and E^- is the spectrum of eigenfunctions which are antisymmetric under permutations of space *and* spin. For each subdomain, D_j, we likewise define ordered spectra e_i^{j*}, e_i^{j+}, and e_i^{j-}, the only difference being that the symmetry or antisymmetry requirements on the eigenfunctions refer only to the subgroup G_j and not to the full group S_N. Because of the symmetry, if D_k and D_j belong to the same class then the three spectra will be the same for the two domains. For each of the three statistics, if e_i^j is an eigenvalue in D_j with eigenfunction Ψ_i^j, then D_k will have the same eigenvalue with eigenfunction $P\Psi_i^j$, where $PD_j = D_k$.

The three partition functions for the full domain D are defined by

$$Z^* = (N!)^{-1} \sum_{i=1}^{\infty} \exp(-\beta E_i^*), \tag{2.21}$$

$$Z^{\pm} = \sum_{i=1}^{\infty} \exp(-\beta E_i^{\pm}). \tag{2.22}$$

Correspondingly, we can define the partition functions for each domain, D_j:

$$Z_j^* = (s_j)^{-1} \sum_{i=1}^{\infty} \exp(-\beta e_i^{j*}), \tag{2.23}$$

$$Z_j^{\pm} = \sum_{i=1}^{\infty} \exp(-\beta e_i^{j\pm}). \tag{2.24}$$

We can also define the combined subdomain (CSD) partition functions, designated by $\sim$, as:

$$\tilde{Z}^* = \sum_C Z_C^* = (N!)^{-1} \sum_{i=1}^{\alpha} \sum_{i=1}^{\infty} \exp(-\beta e_i^{j*}), \tag{2.25}$$

$$\tilde{Z}^{\pm} = \sum_C Z_C^{\pm}. \tag{2.26}$$

Here $\sum_C$ means sum on classes and Z_C is the common value of the partition function of any domain in class C.

Using the definition (2.25), Corollary 1 now reads

Corollary 2. If conditions (S1), (S2), and (S3) hold, then

$$Z_1^* \geq \tilde{Z}_1^*. \tag{2.27}$$

The analogous statement for $Z^{\pm}$ is not a corollary of Lemma 1, but is a separate lemma:

Lemma 2. If conditions (S1), (S2), and (S3), hold, then

$$Z_1^{\pm} \geq \tilde{Z}_1^{\pm}. \tag{2.28}$$

The above lemmas refer to VBC. The natural question of what effect altering the boundary conditions has can be answered provided we add one more condition—which will be referred to as the *filling condition* (F)—namely:

(F). The full domain, D, is the union of the subdomains, D_i.

Using (F), the corresponding lemmas for NBC are contained in

Lemma 3. Let the filling condition (F) be satisfied. Then, in Lemmas 1 and 2 and in Corollaries 1 and 2, if the boundary conditions be changed to NBC instead of VBC—all other conditions remaining unaltered—the conclusions remain true if all the inequalities be reversed.

We shall refer to these altered lemmas and corollaries as Lemma 1′ and 2′ and Corollary 1′ and 2′, respectively. In particular (under appropriate conditions), we have

Corollary 2′. $$Z_2^* \leq \tilde{Z}_2^*, \tag{2.29}$$

Lemma 2′. $$Z_2^{\pm} \leq \tilde{Z}_2^{\pm}. \tag{2.30}$$

Finally, we need three more lemmas.

Lemma 4. Let H be given as in (2.16), in a domain D, but with $U = U_1 + U_2$. Let m_- and m_+ be two numbers such that

$$m_- \leq U_2(\mathbf{R}) \leq m_+ \tag{2.31}$$

for all $\mathbf{R} \in D$. Let E_i be the ordered spectrum (for any boundary conditions) of H and let E_i' be the corresponding spectrum with $U = U_1$. Then for all i

$$E_i' + m_- \leq E_i \leq E_i' + m_+. \tag{2.32}$$

Moreover, if conditions (S2) and (S3) hold, then the inequality (2.32) is true for any statistics and for any subdomain D_i, assuming that U_1 and U_2

are separately invariant with respect to the subgroup, G_i, of that subdomain.

Lemma 5. [Precise statement of (2.14)]. Let (S1), (S2), and (S3) hold and, for any statistics, let E_{1i} and E_{2i} be the ordered spectra with VBC and NBC, respectively. Then, for all i,

$$E_{1i} \geq E_{2i}. \tag{2.33}$$

Moreover, if $D = \Omega^N$, where Ω is a cube, let E_{3i} be the ordered spectrum with PBC. Then, for all i,

$$E_{i\,1} \geq E_{3i} \geq E_{2i}. \tag{2.34}$$

Lemma 6. Let (S1), (S2), and (S3) hold and let $D_i' \subset D_i$ be two subdomains invariant under the same subgroup G_i. For any statistics, but for VBC, let e_i' and e_i be the ordered spectra in the two domains, respectively. Then, for all i,

$$e_i' \geq e_i. \tag{2.35}$$

Using these lemmas we can now prove the theorem. The basic strategy of the LP proof is to find upper and lower bounds on the free energy and then to show that, in the appropriate limit, they both agree with (2.6). The first step is to divide the cube Ω into M smaller cubes (which completely fill Ω), each of side $s + t$ ($s > 0, t > 0$). The ith cube is designated by ω_i, and its volume is $(s + t)^\nu \equiv \omega$, with $\Omega = M\omega$. Concentric with, and interior to ω_i one places the smaller cube ω_i' of side s and volume $\omega' = s^\nu$ so that the distance between any two of the smaller cubes is $\geq 2t$. Just as Ω is the union of all the ω_i, we shall define Ω' to be the union of all the ω_i'. Ω' consists of M disconnected pieces. Corresponding to the domain $D = \Omega^N$, we define $D' = (\Omega')^N$.

We now introduce subdomains of D, each of which is characterized by having certain *specified* particles in each of the M domains ω_i. There are M^N such subdomains, and plainly, for the problem at hand, conditions (S1), (S2), (S3), and (F) are satisfied. Likewise, for each subdomain D_k of D we can define the subdomain D_k' of D' in which the same particles are restricted to ω_i' instead of merely to ω_i. Furthermore, according to the previous definition, all subdomains which have the same *number* of particles in each of the ω_i, and which, therefore, differ only in the identity of the particles, belong to the same equivalence class. Hence, the classes are specified by the nonnegative integers $N_1, \cdots, N_M$ (whose sum is N) and

$$s(C) = s(N_1, \cdots, N_M) = N_1!\, N_2! \cdots N_M!. \tag{2.36}$$

Moreover, the subgroup G_i, for

$$D_i \in C = (N_1, \cdots, N_M),$$

is the product of the symmetric groups $S_{N_1} \times S_{N_2} \times \cdots \times S_{N_M}$, where S_{N_i} is the full symmetric group on the variables in the cube ω_i.

The total potential U is the sum over pairs of (2.1) and, for any subdomain D_i or D_i' it is convenient to write it as

$$U = V' + V'', \tag{2.37}$$

where V' is the sum of *intra*-cell interactions and V'' is the sum of *inter*-cell interactions. Clearly both V' and V'' are invariant under G_i. Were V'' absent, we would have (for any statistics, and any boundary conditions) that Z_i, the partition function of D_i, would be equal to $\prod_{i=1}^M Z(N_i, \omega, \gamma)$. Nevertheless, the effect of V'' can be bounded using Lemma 4. The maximum value of V'' in D_i clearly depends only on the class C of D_i and not on the particular value of i. It will be denoted by $V_+(C)$, and by $V_+'(C)$ if we are concerned with the subdomain D_i'.

We also require another division of the potential, U. For any subdomain D_i we write[8]

$$U = Q' + W, \tag{2.38}$$

where Q' is the *intra*-cell contribution to U from q alone, the short range potential. The minimum value of W (which, again, depends only on the class of D_i) will be denoted by $W_-(C)$.

Now, combining all the foregoing diverse lemmas and corollaries, we have the following bounds which apply for each of the three statistics:

$$\begin{aligned} &\sum_C \left\{ \prod_{i=1}^M Z_2^0(N_i, \omega) \exp\left[-\beta W_-(C)\right] \right\} \\ &\geq Z_2(N, \Omega, \gamma) \geq Z_1(N, \Omega, \gamma) \\ &\geq \sum_C \left\{ \prod_{i=1}^M Z_1(N_i, \omega', \gamma) \exp\left[-\beta V_+'(C)\right] \right\}, \end{aligned} \tag{2.39}$$

where the superscript 0 refers, of course, to the base system. These three inequalities, which are trivial to derive classically, are in fact the *starting* point of the LP analysis, which consists almost entirely in estimating the comtributions from the exponential factors in (2.39). We may therefore follow their analysis, except that we must carefully distinguish between Z_1 and Z_2, a distinction which has no classical counterpart. In this connection, it is important to recall condition (q3). It is also important to recall that Fisher[7] has proved that the bulk limit exists for VBC and finite γ, and for all statistics. Namely

[8] Our W is referred to in LP as $Q + W$.

$$\lim_{\Omega\to\infty} A_1(\rho\Omega, \Omega, \gamma) = a_1(\rho, \gamma) \tag{2.40}$$

and that $a_1(\rho, \gamma)$ is a convex function of ρ.

Using (2.40), together with the LP analysis and the inequalities (2.39), we have (for all statistics)

$$\begin{aligned}
&\text{C.E. } \{a^0(\rho) + \tfrac{1}{2}\alpha\rho^2\} \\
&\leq \liminf_{\gamma\to 0} \liminf_{\Omega\to\infty} \Omega^{-1}A_2(\rho\Omega, \Omega, \gamma) \\
&\leq \limsup_{\gamma\to 0} \limsup_{\Omega\to\infty} \Omega^{-1}A_2(\rho\Omega, \Omega, \gamma) \\
&\leq \limsup_{\gamma\to 0} a_1(\rho, \gamma) \leq \text{C.E. } \{a^0(\rho) + \tfrac{1}{2}\alpha\rho^2\}.
\end{aligned} \tag{2.41}$$

This series of inequalities, (2.41), plainly prove part 1 of the theorem, Eq. (2.6).

The second part of the theorem, Eq. (2.7), which refers to the pressure, is quite independent of the previous inequalities. As LP point out in section VI, part 2 follows merely from the observation that for nonzero γ the bulk limit of the free energy, $a(\rho, \gamma)$, exists and is a convex function of ρ. For VBC we know this to be true by Fisher's theorem.[7] For NBC, on the other hand, no such theorem exists and the inequalities in (2.41) shed no light on this question. Therefore, part 2, unlike part 1, must remain qualified as to boundary conditions.

III. PROOF OF THE LEMMAS

Of the six lemmas we wish to prove, only Lemma 3 is really new. The others have been known for many years and it is difficult to say who first thought of them. We will mention, however, that Corollaries 1 and 2 and Lemma 2, all of which refer to partition functions, are to be found in the aforementioned work of Fisher.[7] The reason we wish to prove these again is to make it clear that they are really only a consequence of deeper theorems about individual energy levels.

Our starting point is the variational expression

$$E(\Psi) \equiv \left\{\int_D |\nabla\Psi|^2 + \int_D U\,|\Psi|^2\right\} \Big/ \int_D |\Psi|^2, \tag{3.1}$$

from which quantum-mechanics flows.[9] The eigenvalues are defined by the minimum principle:

$$E_1 = \min_{\Psi} E(\Psi), \quad E_2 = \min_{\Psi,\,\Psi\perp\Psi_1} E(\Psi), \text{ etc.,} \tag{3.2}$$

where $\Psi\perp\Psi_1$ means that the admissible Ψ's are restricted to those which are orthogonal to Ψ_1 = the Ψ which gives ϵ_1. The admissible Ψ's also satisfy the appropriate boundary conditions.

It is important to note, however, that the admissible Ψ's need only be continuous and *piecewise* differentiable. This point is sometimes a cause of confusion. The calculus of variations tells us that a minimizing Ψ satisfies the Schrödinger equation and, therefore, that its first derivative is continuous. The continuity of the first derivative is thus a condition *derived* from the variational principle, rather than a postulated condition.

Since $i = (-1)^{\frac{1}{2}}$ does not appear explicitly in (3.1), all the eigenfunctions can be taken to be real, even for PBC. This is time-reversal invariance. We also have two, important, well-known theorems about the eigenvalues, E_i, as follows:

Maxi–Min Principle[10]: Let $\varphi^1 \cdots \varphi^n$ be any set of n admissible functions and let $e(\varphi^1, \cdots, \varphi^n)$ be the minimum of $E(\Psi)$ such that Ψ is orthogonal to $\varphi^1, \cdots, \varphi^n$. Then $e(\varphi^1, \cdots, \varphi^n) \leq e(\Psi^1, \cdots, \Psi^n) = E_{n+1}$, where the Ψ^i are the true eigenfunctions.

Rayleigh–Ritz Principle: Let $\varphi^1, \cdots, \varphi^n$ be a set of n orthonormal admissible functions (called trial functions) and A be the $n \times n$ matrix whose (i, j) element is

$$\int_D [(\nabla\varphi^i)^* \cdot (\nabla\varphi^j) + U\varphi^{i*}\varphi^j].$$

Since A is obviously Hermitian, let $e_1 \leq e_2 \leq \cdots \leq e_n$ be its real ordered eigenvalues. Then $e_i \geq E_i$ for $i = 1, \cdots, n$.

The proof of Lemma 1 is now trivial. In each subdomain D_i let Ψ_i^i be the normalized ith eigenfunction with VBC in D_i. Let φ^{ii} be a set of trial functions defined in D as follows: $\varphi^{ii}(\mathbf{R}) = \Psi_i^i(\mathbf{R})$ if $\mathbf{R}$ is in D_i, otherwise it is zero. The set φ^{ii} are clearly admissible and orthonormal, hence Lemma 1 is merely a consequence of the Rayleigh–Ritz principle.

Lemma 2 follows from the same sort of argument plus elementary symmetry considerations. Instead of defining a set of trial functions for each subdomain, D_i, we define a set of trial functions, φ^{Ci}, for each class as follows: For each class, C, pick a definite member of that class, D_C. Let Ψ_i^C be the eigenfunctions with VBC in D_C as above. Then

$$\varphi_i^C(\mathbf{R}) \equiv \sum_P \epsilon_P \Psi_i^C(P\mathbf{R}), \tag{3.3}$$

[9] A genuinely rigorous statement would contain smoothness conditions on U as well as on the boundary of the domain. Such conditions are never mentioned in the physics literature, and we shall not attempt to delineate them here.

[10] R. Courant and D. Hilbert, *Methods of Mathematical Physics* (Interscience Publishers, Inc., New York, 1953), Vol. I, p. 132.

where the sum in (3.3) is over all permutations and where $\epsilon_P = 1$ for Bose statistics or $\epsilon_P = \pm 1$ for Fermi statistics. For the latter it is understood that spins, as well as coordinates, are permuted in (3.3). Lemma 2 is thus proved.

Likewise, Lemma 6 can be proved by defining trial functions in D_i to be eigenfunctions in D_i' with VBC if $\mathbf{R}$ is in D_i', while if $\mathbf{R}$ is in $D_i - D_i'$ they are defined to be zero.

Turning to Lemma 5 we see that part of it, at least, namely $E_{1i} \geq E_{3i}$ (or else $E_{1i} \geq E_{2i}$ if $D \neq \Omega^N$) follows from the above ideas. It is only necessary to consider a domain D_ϵ which lies in D and whose boundary is a distance ϵ from the boundary of D. Let $E_{1i}(\epsilon)$ be the appropriate spectrum with VBC in D_ϵ. As $\epsilon \to 0$, $E_{1i}(\epsilon) \to E_{1i}$, but for any $\epsilon \neq 0$ the eigenfunctions in D_ϵ define, by the above trivial extension into D, a set of trial functions which are admissible for *all* of the three boundary conditions—VBC, NBC, and PBC.

The proof that $E_{3i} \geq E_{2i}$ requires a new idea which is contained in the well-known statement of the calculus of variations:

Free Boundary Conditions: In the variational principle (3.1) and (3.2), let no boundary conditions at all be imposed on the admissible functions. Then the resulting eigenfunctions and eigenvalues will be the same as if we had imposed NBC.

We shall refer to this theorem as FBC. In order to prove that $E_{3i} \geq E_{2i}$, we denote by Ψ_{3i} and Ψ_{2i}, respectively, the corresponding eigenfunctions, and observe that $E_{3,i+1} = \min E(\Psi)$ [such that $\Psi \perp (\Psi_{31}, \cdots, \Psi_{3i})$ and Ψ obeys PBC] $\geq \min E(\Psi)$ [such that $\Psi \perp (\Psi_{21}, \cdots, \Psi_{2i})$ and Ψ obeys PBC] $\geq \min E(\Psi)$ [such that $\Psi \perp (\Psi_{21}, \cdots, \Psi_{2i})$] $= E_{2i}$.

Lemma 4 is proved as follows: let Ψ_i and Ψ_i' be the eigenfunctions for $U = U_1 + U_2$ and $U = U_1$, respectively, and let $E(\Psi)$ and $E'(\Psi)$ be the respective variational expressions (3.1). Then, for any boundary conditions and any statistics, $E_{i+1} = \min E(\Psi)$ [such that $\Psi \perp (\Psi_1, \cdots, \Psi_i)$] $\geq \min E'(\Psi)$ [such that $\Psi \perp (\Psi_1', \cdots, \Psi_i')$] $+ m_- = E_i' + m_-$. The other inequality, $E_i \leq E_i' + m_+$ follows *mutatis mutandis*.

We come now to the heart of the matter—Lemma 3! It is a generalization of Lemmas 1 and 2 to the case of NBC *provided* the filling condition (F) is satisfied. We shall prove Lemma 1′ only, because as in the VBC case, Lemma 2′ requires merely the addition of trivial symmetry considerations. For simplicity we shall suppose that there are only two subdomains, so that $D = D_1 + D_2$, and we shall let S denote the surface common to D_1 and D_2.

We recall that e_i^j ($j = 1, 2$) was defined to be the ith eigenvalue in the subdomain D_j, with NBC on the boundary of that subdomain, while Ψ_i^j was the corresponding eigenfunction in the same subdomain. The set e was defined to be the ordered union of the two sets, e^j. For each i we now define the function φ_i (in D) as follows: If e_i come from the set e^j, so that $e_i = e_k^j$ (for some k), then $\varphi_i \equiv \varphi_k^j$ in D^j and $\varphi_i \equiv 0$ in D_l for $l \neq j$.

Another convenient definition is

$$\alpha_1(\Psi) \equiv \int_{D_1} |\nabla\Psi|^2 + \int_{D_1} U\,|\Psi|^2, \qquad \beta_1(\Psi) \equiv \int_{D_1} |\Psi|^2, \tag{3.4}$$

with similar definitions for α_2 and β_2. From these we have:

$$E(\Psi) = (\alpha_1 + \alpha_2)/(\beta_1 + \beta_2), \tag{3.5}$$

which (if neither β_1 nor β_2 is zero) satisfies the simple inequality:

$$\min\left(\frac{\alpha_1}{\beta_1}, \frac{\alpha_2}{\beta_2}\right) \leq E \leq \max\left(\frac{\alpha_1}{\beta_1}, \frac{\alpha_2}{\beta_2}\right). \tag{3.6}$$

Now consider the variational problem, (3.1) and (3.2), to be modified as follows: *functions which are discontinuous across the surface S (which separates D_1 and D_2) are admissible.* This modified problem has, in fact, a solution, and the reader can easily verify with the help of FBC and (3.6) that the eigenvalues are the e_i, while the eigenfunctions are the φ_i.

Lemma 1′ is a simple consequence of the foregoing observation, together with the maxi–min principle. Defining the symbol $(\min)_m$ to mean minimization as modified above, we have:

$$\begin{aligned} E_{i+1} &= \min E(\Psi) \text{ [such that } \Psi \perp (\Psi_1, \cdots \Psi_i)] \\ &\geq \min E(\Psi) \text{ [such that } \Psi \perp (\varphi_1, \cdots \varphi_i)] \\ &\geq (\min)_m E(\Psi) \text{ [such that } \Psi \perp (\varphi_1, \cdots \varphi_i)] \\ &= e_{i+1}. \end{aligned} \tag{3.7}$$

IV. APPLICATION TO THE ONE-DIMENSIONAL MODEL

With the theorem (2.6) proved, we can go on to present the analog of (1.26) for the quantum-mechanical case. No attempt will be made to discuss

the physical properties or significance of this quantum-mechanical model.[11]

It is a simple matter[2] to calculate the energy levels for a hard-core gas in a one-dimensional "box" of length L. The bulk limiting free energy (2.3) is not only independent of boundary conditions, it is independent of statistics as well (except for a trivial spin entropy). Moreover, if VBC are used, the free energy (2.2) is rigorously independent of statistics even for a finite system, and we shall therefore discuss VBC.

Each state is characterized by a set of *distinct* numbers $k_1, \cdots, k_N$, each of which is of the form

$$k_i = n_i\pi[L - (N - 1)r_0]^{-1}, \tag{4.1}$$

where n_i is any nonzero *positive* integer. The total energy of the state is

$$E = \frac{\hbar^2}{2m}\sum_{i=1}^{N} k_i^2. \tag{4.2}$$

Unlike the classical system, the quantum system will have the zero-point energy or ground-state energy (any statistics):

$$E_0 = [L - (N - 1)r_0]^{-2}\frac{\hbar^2}{2m}\sum_{n=1}^{N} n^2 \rightarrow \frac{\pi^2\hbar^2}{6m}\frac{\rho^2}{(1 - \rho r_0)^2}N, \tag{4.3}$$

where $\rightarrow$ means passage to the bulk limit. This ground-state energy does not vanish even if $r_0 \rightarrow 0$.

The partition function is given by

$$Z(N, L) = \sum_n \exp\left(-\beta B\sum_{i=1}^{N} n_i^2\right), \tag{4.4}$$

where

$$B = \pi^2(\hbar^2/2m)[L - (N - 1)r_0]^{-2}, \tag{4.5}$$

and $\sum_n$ signifies a summation over the N integers such that $0 < n_1 < n_2 < \cdots < n_N$. $Z(N, L)$ is the partition function for Boltzmann, Bose, or one-component Fermi (i.e., fermions without spin) statistics. For conventional Fermi statistics, all that is necessary is to multiply (4.4) by the spin degeneracy factor, 2^N.

The computation of (4.4) in the bulk limit is facilitated by the introduction[12] of a fictitious chemical potential, μ.

[The quantity μ is merely an aid in evaluating (4.4), because this sum is that of a one-component Fermi gas in a box of length $L' = L[1 - \rho r_0]$. μ is not the true chemical potential because L' is not fixed, but depends on N. The true chemical potential is, as always, $\partial a^0(\rho)/\partial\rho$. It is a pleasure to thank J. Lebowitz for bringing this matter to my attention, as well as for many interesting conversations.]

In the standard manner, one obtains

$$a^0(\rho) = kT\rho\left\{\beta\mu - \left(\frac{1 - \rho r_0}{\pi\rho}\right)\left(\frac{2mkT}{\hbar^2}\right)^{\frac{1}{2}} \times \int_0^\infty dy \ln [1 + z\,e^{-y^2}]\right\}, \tag{4.6}$$

where $z = \exp(\beta\mu)$ satisfies

$$1 = \left(\frac{1 - \rho r_0}{\pi\rho}\right)\left(\frac{2mkT}{\hbar^2}\right)^{\frac{1}{2}}\int_0^\infty dy\,\frac{z\,e^{-y^2}}{1 + z\,e^{-y^2}}. \tag{4.7}$$

[11] These properties will be elucidated in a forthcoming paper by E. A. Burke, J. Lebowitz, and E. Lieb (to be published).

Stationary Dust-Filled Cosmological Solution with $\Lambda = 0$ and without Closed Timelike Lines

S. C. MAITRA
University of Maryland, Department of Physics and Astronomy, College Park, Maryland
(Received 6 October 1965)

An analytic and complete solution of Einstein's field equations without the Λ term is presented for a dust-filled universe ($p = 0$). The solution is stationary and inhomogeneous and does not contain any closed time-like lines. Also some of the properties of the solution are studied.

INTRODUCTION

THE purpose of this paper is to present an analytic and complete solution of Einstein's field equations for a dust-filled universe ($\rho > 0$, $p = 0$) without the cosmological Λ term. The solution is inhomogeneous and stationary, with cylindrical symmetry, so it will not be found appropriate in the discussions of observational cosmology; but its existence may give reason to hope that there may also exist nonstationary solutions, to avoid the singular epochs found in the Friedman solutions and other related cosmological models. The solution presented here has the further merit that it does not contain any closed timelike lines. All known solutions of Einstein's field equations for a dust-filled universe seem to suffer from some undesirable features. Consider first the stationary solution. The spatially homogeneous solutions of Einstein, Gödel,[1] Ozsvath and Schücking,[2] and Ozsvath[3] all require a nonvanishing cosmological Λ term. The inhomogeneous solution of Lanczos[4] and van Stockum[5] (which has recently been rediscovered by Wright[6]) although it does not require a cosmological term, nevertheless contains closed timelike lines as in Gödel's universe. Further, as pointed out by Shepley,[7] the solution has a singularity at a finite proper distance from the axis of symmetry where the matter density and scalar curvature become infinite. Ehlers[8] has shown how one can construct all solutions with $\Lambda = 0$ for distributions of matter in rigid rotation from static vacuum metrics, but the global properties regarding the presence of singularities and closed time-like lines have not been investigated. Of the nonstationary solutions with vanishing cosmological constant, the singular epochs in the Friedman homogeneous and isotropic solutions are familiar, and Shepley[7] has shown that a large class of closed, homogeneous nonisotropic solutions also involve singular epochs, while Hawking[9] shows that all solutions which at some epoch differ from the open Friedman model in sufficiently small but otherwise arbitrary ways have, like the Friedman model itself, evolved from a singular beginning. We leave as a problem for further investigations to decide whether the present example is entirely exceptional, or whether it is an especially simple limiting case for some significant class of nonsingular cosmologies.

I. STATEMENT OF RESULTS

The solution given in Eqs. (I.1)–(I.5) below has the following properties:

1. It is cylindrically symmetric.
2. It is stationary but the time-like Killing vector is not the velocity vector of matter.
3. Defined with respect to velocity vector of matter, shear and rotation do not vanish but the expansion vanishes. Thus unlike the case of Lanczos–van Stockum and Ehlers solutions the motion is nonrigid.
4. It does not contain any closed timelike line.
5. The space is complete.
6. The solution is open in all spatial directions, i.e. it extends to infinite proper distance in all directions.
7. Matter everywhere moves in circles about the axis of symmetry.
8. The solution is spatially inhomogeneous and the density as well as the kinematic quantities rotation and shear tend to zero as one goes to arbitrarily large distances from the axis of symmetry.

[1] K. Gödel, Rev. Mod. Phys. **21**, 447 (1949).
[2] I. Ozsvath and E. Schücking, Nature **193**, 1168 (1962).
[3] I. Ozsvath, J. Math. Phys. **6**, 591 (1965).
[4] C. Lanczos, Z. Physik, **21**, 73 (1924).
[5] W. J. van Stockum, Proc. Roy. Soc. Edinburgh **57**, 135 (1937).
[6] J. P. Wright, J. Math. Phys. **6**, 103 (1965).
[7] L. Shepley, Proc. Natl. Acad. Sci. U. S. **52**, 1403 (1964).
[8] J. Ehlers, *Recent Developments in General Relativity* (Pergamon Press, Inc., New York, 1962), p. 201.
[9] S. W. Hawking, (preprint). See also S. Hawking and G. F. R. Ellis Phys. Letters **17**, 246 (1965).

Proof of the nonobvious statements will be given later.

The line element

$$-ds^2 = g_{\mu\nu}\,dx^\mu\,dx^\nu = -dt^2 + e^{2\Psi}(dr^2 + dz^2) + (r^2 - m^2)\,d\varphi^2 + 2m\,d\varphi\,dt \qquad (I.1)$$

is a solution of the Einstein field equations

$$R_{\mu\nu} - \tfrac{1}{2}Rg_{\mu\nu} = -KT_{\mu\nu} \qquad (I.2)$$

with

$$T_{\mu\nu} = \rho v_\mu v_\nu, \qquad (I.3)$$

where v^μ is a unit time-like vector and constitutes a geodesic congruence and ρ is the matter density with following values for Ψ and m:

$$\Psi = -(1/4x^2)[(1 + x^2)^{\frac{1}{2}} - 1] + \tfrac{1}{8} - \tfrac{1}{4}\ln\tfrac{1}{2}[(1 + x^2)^{\frac{1}{2}} + 1] \qquad (I.4)$$

$$m = \tfrac{1}{2}a\{[(1 + x^2)^{\frac{1}{2}} - 1 - \ln\tfrac{1}{2}[(1 + x^2)^{\frac{1}{2}} + 1]\}. \qquad (I.5)$$

Here we have introduced a new variable x through

$$r = \tfrac{1}{2}a\cdot x,$$

where a is a constant.

It is to be noted that by introducing the coordinates $\chi = (2/a)r$, $\tau = (2/a)t$, $\xi = (2/a)z$, and $\varphi = \varphi$, we can write

$$ds^2 = \tfrac{1}{4}a^2(ds^0)^2,$$

where $(ds^0)^2$ depends only on $\chi\tau\xi\varphi$ and does not contain a. Hence this constant can actually be reduced to a scale factor by a coordinate transformation and does not play any further role.

The nonvanishing components of v^μ are

$$v^\varphi = -y/r(1 - y^2)^{\frac{1}{2}} \qquad (I.6)$$

$$v^t = (r - my)/r(1 - y^2)^{\frac{1}{2}}, \qquad (I.7)$$

where

$$y = m' = dm/dr. \qquad (I.8)$$

The density is given by

$$K\rho = (4/a^2)\,e^{-2\Psi}\cdot[x^4(1 + x^2)^{\frac{1}{2}}]^{-1}[(1 + x^2)^{\frac{1}{2}} - 1]^2. \qquad (I.9)$$

II. COMPUTED PROPERTIES

A number of properties of the solution stated above follows from straightforward computations, the results of which will be given here.

It can easily be shown from the expression (I.9) of density that the total amount of matter (calculated per unit proper length along z direction) is finite whereas the total proper volume (per unit proper length in z direction) is infinite, so that the matter is distributed with zero average density or in infinite dilution.

The vorticity vector corresponding to the above velocity is defined by

$$\omega^\mu = (1/2(-g)^{\frac{1}{2}})\epsilon^{\mu\alpha\beta\gamma}v_\alpha(\partial v_\beta/\partial x^\gamma) \qquad (II.1)$$

$$\epsilon^{0123} = +1.$$

It has only a component in z direction, and the magnitude of angular velocity is given by

$$\omega^2 = g_{zz}\omega^z\omega^z = \tfrac{1}{4}\cdot(1/a^2)\,e^{-2\Psi}\cdot 1/(1 + x^2). \qquad (II.2)$$

The shear tensor

$$\varphi_{\mu\nu} = \tfrac{1}{2}(v_{\mu;\nu} + v_{\nu;\mu}) - \tfrac{1}{3}(g_{\mu\nu} + v_\mu v_\nu)v^\rho_{;\rho} \qquad (II.3)$$

has only nonvanishing components φ^{tr} and $\varphi^{\varphi r}$ given by

$$\varphi^{r\varphi} = \frac{1}{a^2}\left(\frac{x}{2y}\right)^{\frac{1}{2}} e^{-2\Psi}\,\frac{[(1 + x^2)^{\frac{1}{2}} - 1]^2}{x^3(1 + x^2)^{\frac{1}{2}}} \qquad (II.4)$$

$$\varphi^{tr} = -\tfrac{1}{2}a\ln\tfrac{1}{2}[(1 + x^2)^{\frac{1}{2}} + 1]\varphi^{r\varphi}. \qquad (II.5)$$

From this one gets

$$\varphi^2 = \varphi_{\alpha\beta}\varphi^{\alpha\beta} = \frac{1}{2a^2}\,e^{-2\Psi}\,\frac{[(1 + x^2)^{\frac{1}{2}} - 1]^4}{x^4(1 + x^2)}. \qquad (II.6)$$

From the above expression it follows that

$$\tfrac{1}{2}K\rho + \varphi^2 = 2\omega^2. \qquad (II.7)$$

It may be noted that, as $r \to \infty$, ρ vanishes more rapidly than φ^2 and ω^2.

We give below the components of Riemann tensor computed in an orthonormal frame. We write

$$ds^2 = -(\omega^0)^2 + (\omega^1)^2 + (\omega^2)^2 + (\omega^3)^2 = g_{\mu\nu}\omega^\mu\omega^\nu,$$

where

$$g_{\mu\nu} = \text{diag}\,(+1, +1, +1, -1) \qquad (II.8)$$

and

$$\omega^1 = e^\Psi\cdot dr, \quad \omega^2 = e^\Psi\cdot dz, \quad \omega^3 = r\cdot d\varphi, \quad \omega^0 = dt - m\,d\varphi. \qquad (II.9)$$

We compute the curvature tensor using the methods described by Misner.[10] Referred to the above orthonormal frame, it has the following independent nonvanishing components:

$$R_{1212} = -e^{-2\Psi}\cdot\Psi'',$$

$$R_{1310} = (e^{-2\Psi}/r)[(y/2r) + \tfrac{1}{2}\Psi' y - \tfrac{1}{2}y')], \qquad (II.10)$$

$$R_{1313} = e^{-2\Psi}[\tfrac{3}{4}(y^2/r^2) + \Psi'/r],$$

$$R_{1010} = e^{-2\Psi}\tfrac{1}{4}(y^2/r^2), \quad R_{2323} = -e^{-2\Psi}(\Psi'/r),$$

$$R_{2320} = -e^{-2\Psi}(y/2r)\Psi', \quad R_{3030} = \tfrac{1}{4}(y^2/r^2)\,e^{-2\Psi}.$$

[10] C. W., Misner, J. Math. Phys. 4, 924 (1963).

It will be noted that, as $r \to \infty$, all components go to zero as $r^{-\frac{1}{2}}$. The proper radial distance r_p out to a radius r goes as $\int^r e^{\Psi}\, dr = Kr^{\frac{1}{2}}$ for $r \to \infty$, K being a constant. Hence we see that as $r_p \to \infty$, all components fall off like r_p^{-2}. This is significant in this frame since, with each $g_{\mu\nu} = \pm 1$, one sees that every invariant polynomial in $R_{\mu\nu\alpha\beta}$ will also vanish as $r \to \infty$. In other frames where $g_{\mu\nu}$ depends on r, the behavior of curvature invariants is not easily deduced from that of curvature components.

Finally we compute the c-energy scalar as introduced by Thorne[11] for our system. At a point $(trz\varphi)$ it is defined as

$$U = \frac{1}{8}\left(1 + \frac{A_{,\mu}A^{,\mu}}{4\pi^2\,|\vec{\xi}z|^2}\right), \qquad \text{(II.11)}$$

where $A_{,\mu}$ is the space–time gradient of the area A of the invariant surface passing through the point and consisting of the points $(t,\ r,\ z + \alpha,\ \varphi + \beta)$ where $0 \leq \alpha \leq 1$, $0 \leq \beta \leq 2\pi$, and $|\vec{\xi}_z|$ is the length of the standard "translation" Killing vector at that point. For our metric, we have

$$U = \tfrac{1}{8}\{1 - e^{-2\Psi}\cdot q[\Psi' + (q'/2q)]^2\}, \qquad \text{(II.12)}$$

where we have introduced

$$q = r^2 - m^2 \qquad (r \to \infty,\quad U \to \tfrac{1}{8}). \qquad \text{(II.13)}$$

III. ANALYTICITY

Using the metric components from Eq. (I.1) one computes

$$(-g)^{\frac{1}{2}} = r\, e^{2\Psi}. \qquad \text{(III.1)}$$

This metric is consequently singular at $r = 0$ where $(-g)^{\frac{1}{2}} = 0$, but is analytic for $r > 0$ where each of the metric components $r^2 - m^2$, m, and $e^{2\Psi}$ is analytic and where $(-g)^{\frac{1}{2}} > 0$. We show then that this $r = 0$ singularity is spurious (removable) by interpreting $tr\varphi z$ as cylindrical coordinates; that is, we introduce new coordinates $tXYZ$ by the transformation

$$t = t,\quad Z = z,\quad X = r\cos\varphi,\quad Y = r\sin\varphi, \qquad \text{(III.2)}$$

and discuss analyticity in the new coordinates. The Jacobian of the transformation is just r, so the metric remains analytic in the region $r^2 = X^2 + Y^2 > 0$ and we need consider in detail only the neighborhood of $r = 0$. To transform Eq. (I.1) to these "rectangular" coordinates, it is most convenient to write

$$\begin{aligned} ds^2 = {} & [-dt^2 + dr^2 + dz^2 + r^2\, d\varphi^2] \\ & + 2\, dt(m\, d\varphi) - (m\, d\varphi)^2 \\ & + (e^{2\Psi} - 1)(dr^2 + dz^2), \end{aligned} \qquad \text{(III.3)}$$

where the quantity in square bracket is, by a familiar computation, analytic (even flat) in $tXYZ$ coordinates. We show now that the remaining terms contribute analytic functions to the metric components, as is obvious for the term $(e^{2\Psi} - 1)dz^2$ [which contributes $(e^{2\Psi} - 1)$ to g_{zz}] since $e^{2\Psi}$ is an analytic function of $r^2 = X^2 + Y^2$, and hence of X, Y. We next note that

$$r\, dr = X\, dX + Y\, dY$$

and

$$r^2\, d\varphi = -Y\, dX + X\, dY$$

are analytic differential forms, so the analyticity of the contributions from

$$2\, dt(m\, d\varphi) = 2(m/r^2)\, dt(r^2\, d\varphi)$$

and $(md\varphi)^2 = (m/r^2)^2(r^2 d\varphi)^2$ follows from that of m/r^2 (which it is very easy to show from the expression for m). Similarly, from

$$(e^{2\Psi} - 1)\, dr^2 = [(e^{2\Psi} - 1)/r^2](r\, dr)^2,$$

one gets analytic contributions since $(e^{2\Psi} - 1)/r^2$ is an analytic function of X and Y for all X, Y.

The determinant of the transformed metric is just $-e^{4\Psi} \neq 0$ so that contravariant components are also everywhere analytic. The velocities in $(XYZt)$ coordinates are given by

$$\begin{aligned} V^X &= Y(y/r)/(1 - y^2)^{\frac{1}{2}}, \\ V^Y &= X(y/r)/(1 - y^2)^{\frac{1}{2}}, \\ V^t &= [1 - m(y/r)]/(1 - y^2)^{\frac{1}{2}}. \end{aligned} \qquad \text{(III.4)}$$

Since (y/r) is an analytic function of X, Y, and $x^2 = (4/a^2)(X^2 + Y^2)$, and

$$y = dm/dr = x/[(1 + x^2)^{\frac{1}{2}} + 1][x = (2/a)r]$$

is less than unity for all finite x and hence all finite X, Y, we find from Eqs. I.6 and I.7 that v^μ remain analytic for all finite X, Y.

Since $dm/dr < 1$ we have $m < r$ and hence $r^2 - m^2 > 0$ for all r, a fact which will be needed later.

The preceding calculations are given in such meticulous detail because such computations do not appear in most texts, and the results [that m/r^2 and $(e^{2\Psi} - 1)/r^2$ need to be analytic function of r^2] are not obvious without computations. There is no

[11] K. S. Thorne, Ph. D. Thesis Princeton University (1965). It should be noted that this definition is slightly different from and is superior to that given in K. S. Thorne, Phys. Rev. **138**, B251 (1965).

general method for asserting differentiability of a metric except to display it in a coordinate system where the components are differentiable and where $(-g)^{\frac{1}{2}} > 0$.

IV. COMPLETENESS

In this section we show that our space is complete, i.e. every geodesic has infinite length in both directions; for null geodesics we have to measure the length by means of an affine parameter. Because of the high symmetry involved, the problem is essentially easy, so that we have a large number of constants of motion. If we take, for the general equations, the Lagrangian

$$L = \tfrac{1}{2}g_{\mu\nu}(dx^\mu/d\lambda)(dx^\nu/d\lambda), \qquad \text{(IV.1)}$$

where λ is a parameter, which can be taken as path length s for non-null lines, we get the following constants of motion

$$\begin{aligned} E &= -P_t = dt/d\lambda - m(d\varphi/d\lambda), \qquad P_z = e^{2\Psi}\, dz/d\lambda, \\ l &= P_\varphi = (r^2 - m^2)(d\varphi/d\lambda) + m(dt/d\lambda), \\ \epsilon &= -(dt/d\lambda)^2 + e^{2\Psi}[(dr/d\lambda)^2 + (dz/d\lambda)^2] \\ &\quad + (r^2 - m^2)(d\varphi/d\lambda)^2 + 2m(d\varphi/d\lambda)(dt/d\lambda), \qquad \text{(IV.2)} \end{aligned}$$

where E, l, P_z can be interpreted as energy, angular momentum, and momentum along z direction for a particle of unit mass. We have

$$\epsilon = 0, \pm 1$$

depending on whether the geodesic is null, spacelike, or timelike.

We can rewrite the above relations as follows:

$$\begin{aligned} dt/d\lambda &= E + (ml - m^2E)/r^2 \\ &= [lm + E(r^2 - m^2)]/r^2, \\ d\varphi/d\lambda &= (l - mE)/r^2, \quad dz/d\lambda = e^{-2\Psi}\cdot P_z, \\ (dr/d\lambda)^2 &+ e^{-2\Psi}[-E^2 - \epsilon + (l - Em)^2/r^2] + e^{-4\Psi}P_z^2 = 0. \end{aligned} \qquad \text{(IV.3)}$$

The last equation can be rewritten as

$$\tfrac{1}{2}(dr/d\lambda)^2 + V_{eff}(r) = 0 \qquad \text{(IV.4)}$$

with a suitable definition of $V_{eff}(r)$.

It is to be noted that we must have $l = 0$ for a particle passing through origin.

Equation (IV.4) is particularly easy to understand, since it resembles the motion of a particle in a potential well. If we look at the behavior of Ψ we see that for large r

$$e^{-\Psi} \sim r^{\frac{1}{2}}.$$

Hence, if $P_z \neq 0$, $V_{eff}(r)$ becomes positive for sufficiently large r. Hence motion along r coordinate will be bounded so that there will be a value of r corresponding to all values of λ. Again, since motion of r is bounded, we find that

$$dz/d\lambda < A,$$

where A is a constant depending on the particular geodesic. Hence in z direction the particle can also escape to infinity only at an infinite value of λ so that (with respect to z coordinate) the geodesics can be continued for all values of λ. Similar arguments hold for other coordinates. It is to be noted that even for particles passing through the origin there is no singularity involved, since here $l = 0$ and m/r^2 is finite at the origin.

Next we take up the case $P_z = 0$. In this case as $r \to \infty$, $V_{eff}(r) \sim -r^{\frac{1}{2}}$, so that $dr/d\lambda \sim r^{\frac{1}{2}}$.

This shows that infinite value of r is reached only when $\lambda \to \infty$. The other equations show that as λ and hence $r \to \infty$, $dt/d\lambda$ and $d\varphi/d\lambda \to 0$ so that with respect to these coordinates λ can also be continued to infinite value. Hence we see that in all cases the geodesics can be continued for all values of the path parameter. Hence the space is complete.

V. ABSENCE OF CLOSED TIMELIKE LINES

If the space contains closed timelike lines, then t will be a periodic function of the parameter λ describing the line.

In such a case we must have maxima and minima of t as function of λ so that there will be points where $(dt/d\lambda)$ will be zero. At such a point we will have

$$\begin{aligned} -(ds/d\lambda)^2 &= g_{\mu\nu}(dx^\mu/d\lambda)(dx^\nu/d\lambda) \\ &= e^{2\Psi}[(dr/d\lambda)^2 + (dz/d\lambda)^2] \\ &\quad + (r^2 - m^2)(d\varphi/d\lambda)^2 > 0 \end{aligned}$$

as both $e^{2\Psi}$ and $(r^2 - m^2)$ are positive for all values of r. Hence at such points the line is no longer timelike, and our space does not have closed timelike lines.

VI. CHARACTERIZATION OF THE METRIC

In this section we shall try to characterize our solution by its Killing vectors. In the following, we do not distinguish between vectors X^μ and their corresponding differential operators related by

$$X = X^\mu(\partial/\partial X^\mu). \qquad \text{(VI.1)}$$

A basis for Killing vectors for our space are the following three vectors:

$$T = \partial/\partial t, \qquad Z = \partial/\partial z, \qquad \Phi = \partial/\partial\varphi. \qquad \text{(VI.2)}$$

Each of them satisfies Killing's equations

$$\xi_{\mu;\nu} + \xi_{\nu;\mu} = 0, \tag{IV.3}$$

which reduces to

$$(\partial/\partial x^{\beta})g_{\mu\nu} = 0 \tag{VI.4}$$

if ξ is the vector $\partial/\partial x^{\beta}$. If we now try to look at the problem from a more general point of view and try to investigate infinitely long cylindrical systems which are stationary, the above are the natural Killing vectors for such a system. Since the Killing vectors commute with each other, we can choose coordinate axes $t\varphi z$ so that the Killing vectors point along them. In such a situation the metric components will depend only on the fourth coordinate r. If we impose the following reflection symmetries which are appropriate to an infinitely long cylindrically symmetric system, which is rotating, we can, with one further restriction on g_{tt}, arrive at our form of line element.

We impose two reflection symmetries. The first is $z \to -z$. The second is the simultaneous reflection $t \to -t$ and $\varphi \to -\varphi$. The first one eliminates all cross terms in z. Consider the term g_{rz} and make the transformation $z' = -z$, other coordinates remaining same. This gives

$$g_{rz'} = -g_{rz}.$$

But from the reflection symmetry

$$g_{rz'} = g_{rz}.$$

Hence $g_{rz} = 0$. Similarly using the other symmetry, all cross terms except the one in $\varphi - t$ are eliminated. Hence our metric takes the form

$$g_{tt}\,dt^2 + g_{rr}\,dr^2 + g_{zz}\,dz^2 + g_{\varphi\varphi}\,d\varphi^2 + 2g_{\varphi t}\,d\varphi\,dt, \tag{VI.5}$$

where each of the components depends only on r. Now by a simple scale transformation for r, we can make $g_{rr} = g_{zz}$. This does not change any of the symmetries and hence leaves the above form unchanged. Next we make the simplifying assumption that $g_{tt} = -1$.

Hence, calling $g_{rr} = g_{zz} = e^{2\Psi}$, $g_{t\varphi} = m$, we have

$$-ds^2 = e^{2\Psi}(dr^2 + dz^2) + g_{\varphi\varphi}\,d\varphi^2 + 2m\,d\varphi\,dt - dt^2. \tag{VI.6}$$

Now the field equations give

$$g_{\varphi\varphi} = r^2 - m^2.$$

Hence we get

$$-ds^2 = e^{2\Psi}(dr^2 + dz^2) + (r^2 - m^2)\,d\varphi^2 + 2m\,d\varphi\,dt - dt^2. \tag{VI.7}$$

Even with this form of line element, there are two solutions of the field equations. If we use a co-moving frame we get van Stockum's solution which represent matter in rigid rotation. If we employ our form of energy–momentum tensor we get solutions for matter in nonrigid rotation invariantly distinguished from van Stochum's solution by the presence of shear.

An invariant way of restating our special condition $g_{tt} = -1$ is to demand that the timelike Killing vector T has a constant magnitude.

$$T \cdot T = -1. \tag{VI.8}$$

A constant value of $T \cdot T$ implies that that congruence of curves to which the time-like Killing vector is tangent is a geodesic congruence.

It may be remarked that, if we relax the condition that g_{tt} be a constant, we will get a family of solutions in general—two solutions for a given choice of the function $g_{tt}(r)$.

We sum up the contents of this section by giving below the conditions that uniquely lead to our form of the metric.

1. There exist three commuting Killing vectors T, Φ, Z.

2. Our system has reflection symmetries appropriate to a rotating cylinder of infinite length—i.e., it is invariant under the following conditions: $(t, r, \varphi, z) \to (t, r, \varphi, -z)$, $(t, r, \varphi, z) \to (-t, r, -\varphi, z)$.

3. $g_{tt} = -1$.

4. The system is in nonrigid rotation, i.e., shear is present.

VII. CO-MOVING COORDINATES AND COSMOLOGY

One can transform the line element (I.1) to a form in which the velocity vector is $V^{\mu} = \delta^{\mu}_{0}$ (co-moving frame). In this system, the metric explicitly involves time and is no longer stationary. It has the following form in the co-moving system:

$$\begin{aligned} -ds^2 = {} & -dt^2 + \left\{e^{2\Psi} + \frac{t^2}{a^2}\cdot\frac{1}{1+x^2}\,y^4\right\} dr^2 \\ & + e^{2\Psi}\,dz^2 + (r^2 - m^2)\,d\varphi^2 \\ & + \frac{2}{(1+x^2)^{\frac{1}{2}}}\left(\frac{y}{a}\right)^{\frac{1}{2}}(r - my)r^{\frac{1}{2}}t\,dr\,d\varphi \\ & + 2(r/ay)^{\frac{1}{2}}(ry - m)\,d\varphi\,dt. \end{aligned} \tag{VII.1}$$

Although the expansion vanishes, the nonvanish-

ing shear would in this solution give rise to a Doppler shift in the frequency of light emitted by a particle and received by another. Put in another way, the nonstationary nature of the metric in the co-moving system would cause a spectral shift. However, in general, this Doppler shift would be strongly anisotropic (unlike the actually observed, more or less isotropic, Hubble red-shift). We do not therefore propose the solution as a model of the observed universe, but as noted earlier we can hope to build a singularity-free dynamical model from this. This solution further emphasizes that one can construct anti-Mach metrics without taking recourse to the Λ term or introducing unphysical situations.

ACKNOWLEDGMENTS

I wish to express my deep gratitude to Professor A. K. Raychaudhuri for suggesting the problem and for numerous helpful discussions. I also wish to thank Professor C. W. Misner for many helpful suggestions, and Professor J. Ehlers for reading part of the manuscript and suggesting improvements. The research was supported by NASA Grant NsG 436.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 7, NUMBER 6 JUNE 1966

Nonhomogeneous Differential Equation with a Second-Order Turning Point

D. J. McGuinness
Department of Mathematics, University of Maryland, College Park, Maryland
(Received 5 October 1965)

The asymptotic behavior of solutions of a class of parameter-dependent second-order nonhomogeneous linear ordinary differential equations with a second-order turning point is investigated. It is shown that, under certain conditions, particular solutions can be represented asymptotically by expansions involving certain special functions. Properties of these special functions are studied.

1. INTRODUCTION

THE asymptotic behavior of solutions of second-order linear ordinary differential equations of the form

$$d^2u/dt^2 - [\lambda^2P(t) + \lambda Q(t, \lambda)]u = \lambda^2F(t, \lambda), \qquad (1.1)$$

where t belongs to some finite closed interval I of the real axis and λ is a large complex parameter, has been studied by several authors. In the homogeneous case, the asymptotic nature of the solutions of (1.1) was investigated by G. D. Birkhoff[1] when $P(t) \neq 0$, $t \in I$. R. E. Langer suggested that, when $P(t)$ vanishes at isolated points of I, solutions of (1.1) could be represented asymptotically by expansions involving solutions of differential equations which are easily solved and which possess the predominant features of the original equation. Each point at which $P(t)$ vanishes is called a *turning point.* In particular, when $P(t)$ has a single zero of order one at $t_0 \in I$ (a simple first-order turning point), Langer[2] has described solutions of (1.1) in terms of Bessel functions oj order $\pm\frac{1}{3}$ (or Airy functions) and R. W. McKelvey[3] has used Whittaker functions (or parabolic cylinder functions) to study the case where $P(t)$ has a single zero of order two on I (a simple second-order turning point).

Recently, interest has arisen in the nonhomogeneous case. This interest is partially motivated by the occurrence of nonhomogeneous differential equations of the form (1.1) in the theory of thin elastic toroidal shells. For a summary of the applications to shell theory, see R. A. Clark.[4] For $F(t, \lambda) \not\equiv 0$ and $P(t) \neq 0$ on I, the theory is well known. R. A. Clark[5] has described particular solutions (that remain bounded for $|x| \geq \delta > 0$ as $\lambda \to \infty$ along straight lines in the λ plane which pass through the origin) of the nonhomogeneous equation with a simple first-order turning point in terms of Lommel functions. In view of the simplicity of the expansions obtained, it seems advantageous to study nonhomogeneous turn-

[1] G. D. Birkhoff, Trans. Am. Math. Soc. 9, 219 (1908).
[2] R. E. Langer, Trans. Am. Math. Soc. 67, 461 (1949).
[3] R. W. McKelvey, Trans. Am. Math. Soc. 79, 103 (1955).
[4] R. A. Clark, *Asymptotic Solutions of Differential Equations*, edited by C. H. Wilcox (John Wiley & Sons, Inc., New York, 1964), pp. 185–209.
[5] R. A. Clark, Arc. Ratl. Mech. Anal. 12, 34 (1963).

ing shear would in this solution give rise to a Doppler shift in the frequency of light emitted by a particle and received by another. Put in another way, the nonstationary nature of the metric in the co-moving system would cause a spectral shift. However, in general, this Doppler shift would be strongly anisotropic (unlike the actually observed, more or less isotropic, Hubble red-shift). We do not therefore propose the solution as a model of the observed universe, but as noted earlier we can hope to build a singularity-free dynamical model from this. This solution further emphasizes that one can construct anti-Mach metrics without taking recourse to the Λ term or introducing unphysical situations.

ACKNOWLEDGMENTS

I wish to express my deep gratitude to Professor A. K. Raychaudhuri for suggesting the problem and for numerous helpful discussions. I also wish to thank Professor C. W. Misner for many helpful suggestions, and Professor J. Ehlers for reading part of the manuscript and suggesting improvements. The research was supported by NASA Grant NsG 436.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 7, NUMBER 6 JUNE 1966

Nonhomogeneous Differential Equation with a Second-Order Turning Point

D. J. McGuinness
Department of Mathematics, University of Maryland, College Park, Maryland
(Received 5 October 1965)

The asymptotic behavior of solutions of a class of parameter-dependent second-order nonhomogeneous linear ordinary differential equations with a second-order turning point is investigated. It is shown that, under certain conditions, particular solutions can be represented asymptotically by expansions involving certain special functions. Properties of these special functions are studied.

1. INTRODUCTION

THE asymptotic behavior of solutions of second-order linear ordinary differential equations of the form

$$d^2u/dt^2 - [\lambda^2P(t) + \lambda Q(t, \lambda)]u = \lambda^2F(t, \lambda), \qquad (1.1)$$

where t belongs to some finite closed interval I of the real axis and λ is a large complex parameter, has been studied by several authors. In the homogeneous case, the asymptotic nature of the solutions of (1.1) was investigated by G. D. Birkhoff[1] when $P(t) \neq 0$, $t \in I$. R. E. Langer suggested that, when $P(t)$ vanishes at isolated points of I, solutions of (1.1) could be represented asymptotically by expansions involving solutions of differential equations which are easily solved and which possess the predominant features of the original equation. Each point at which $P(t)$ vanishes is called a *turning point.* In particular, when $P(t)$ has a single zero of order one at $t_0 \in I$ (a simple first-order turning point), Langer[2] has described solutions of (1.1) in terms of Bessel functions oj order $\pm\frac{1}{3}$ (or Airy functions) and R. W. McKelvey[3] has used Whittaker functions (or parabolic cylinder functions) to study the case where $P(t)$ has a single zero of order two on I (a simple second-order turning point).

Recently, interest has arisen in the nonhomogeneous case. This interest is partially motivated by the occurrence of nonhomogeneous differential equations of the form (1.1) in the theory of thin elastic toroidal shells. For a summary of the applications to shell theory, see R. A. Clark.[4] For $F(t, \lambda) \not\equiv 0$ and $P(t) \neq 0$ on I, the theory is well known. R. A. Clark[5] has described particular solutions (that remain bounded for $|x| \geq \delta > 0$ as $\lambda \to \infty$ along straight lines in the λ plane which pass through the origin) of the nonhomogeneous equation with a simple first-order turning point in terms of Lommel functions. In view of the simplicity of the expansions obtained, it seems advantageous to study nonhomogeneous turn-

[1] G. D. Birkhoff, Trans. Am. Math. Soc. 9, 219 (1908).
[2] R. E. Langer, Trans. Am. Math. Soc. 67, 461 (1949).
[3] R. W. McKelvey, Trans. Am. Math. Soc. 79, 103 (1955).
[4] R. A. Clark, *Asymptotic Solutions of Differential Equations*, edited by C. H. Wilcox (John Wiley & Sons, Inc., New York, 1964), pp. 185–209.
[5] R. A. Clark, Arc. Ratl. Mech. Anal. 12, 34 (1963).

ing-point problems by the direct approach of Clark, rather than considering a Green's function associated with the corresponding homogeneous equation.

It will be shown here that, in the nonhomogeneous second-order turning-point case, certain solutions of (1.1) which remain bounded for $|x| \geq \delta > 0$ as $\lambda \to \infty$ along straight lines through the origin can be represented asymptotically by expansions of the form

$$\lambda \sum_{n=0}^{\infty} \lambda^{-n}[A_n(t)T_{0,K}(z) + \lambda^{-\frac{1}{2}}B_n(t)T_{1,\hat{K}}(z) + \lambda^{-\frac{1}{2}}C_n(t)T'_{0,K}(z) + \lambda^{-1}D_n(t)T'_{1,\hat{K}}(z) + \lambda^{-1}E_n(t)],$$

where

$$z = \lambda^{\frac{1}{2}}x,\ x(t) = \operatorname{sgn} t \left|2\int_{t_0}^{t} P^{\frac{1}{2}}(s)\,ds\right|^{\frac{1}{2}},$$

K, $\hat{K}$ are certain functions of λ, and $T_{0,K}(z)$, $T_{1,\hat{K}}(z)$ are certain solutions of

$$T''_{0,K}(z) + [K - z^2]T_{0,K}(z) = -1,$$

$$T''_{1,K}(z) + [\hat{K} - z^2]T_{1,\hat{K}}(z) = -z.$$

Other research in this direction has been done by Sanders and Liepins,[6] and by Jordan and Shelley.[7] Sanders and Liepins give, in a special case of (1.1), the first term of such an expansion, while Jordan and Shelley construct a series which formally satisfies the differential equations. For λ real, it appears that the series given here is the same series given by the construction of Jordan and Shelley. However in this paper, the terms of this series are grouped in such a way that the order of the error, which arises when the actual solution is approximated by a truncation of this series, can be given. That is, the series constructed here are asymptotic expansions of certain solutions uniformly on any given closed finite interval which contains the turning point in its interior.

2. ASSUMPTIONS

For notational convenience, we will denote the statement "$|\mu|$ sufficiently large" by $|\mu| \geq |\mu_0|$ where μ may be either a variable or a parameter. When a multivalued function is encountered, its principal value is intended and primes or (superscripts) denote differentiation.

It is assumed that $P(t)$ is real-valued, is of class $C^\infty(I)$, and has a second-order zero at some point t_0 in the interior of I. We assume that, if necessary, a translation has been performed and a constant absorbed into λ^2 to bring $P(t)$ into the form $P(t) = t^2 + O(t^3)$.

Although we allow $Q(t, \lambda)$, $F(t, \lambda)$ to be complex-valued, we require that they possess asymptotic expansions

$$Q(t, \lambda) = \sum_{i=0}^{N} Q_i(t)\lambda^{-i} + O(\lambda^{-N-1}),$$

$$F(t, \lambda) = \sum_{i=0}^{N} F_i(t)\lambda^{-i} + O(\lambda^{-N-1})$$

for $|\lambda| \geq |\lambda_0|$, $t \in I$, and N arbitrary. All the coefficients $Q_i(t)$, $F_i(t)$ are assumed to be of class $C^\infty(I)$ and in general λ is complex.

3. A NORMALIZING TRANSFORMATION AND A NORMAL FORM

It is convenient to transform the differential equation under consideration into a normal form. Under the transformation

$$u(t) = \alpha(t)y[x(t)],$$

where

$$x(t) = \operatorname{sgn} t \left|2\int_0^t P^{\frac{1}{2}}(s)\,ds\right|^{\frac{1}{2}}, \qquad \alpha(t) = [x^2(t)/P(t)]^{\frac{1}{4}},$$

Eq. (1.1) becomes

$$d^2y/dx^2 - [\lambda^2x^2 + \lambda p(x) + q(x, \lambda)]y = \lambda^2 f(x, \lambda), \tag{3.1}$$

where the coefficients here are related to those in (1.1) by the expressions

$$\begin{aligned} p[x(t)] &= Q_0(t)\alpha^4(t), \\ f[x(t), \lambda] &= F(t, \lambda)\alpha^3(t), \\ q[x(t), \lambda] &= \lambda[Q(t, \lambda) - Q_0(t)]\alpha^4(t) - \alpha''(t)\alpha^3(t). \end{aligned} \tag{3.2}$$

Note that $\alpha(t)$ and its reciprocal are bounded on I and that $\alpha(t)$ has as many continuous derivatives as $P(t)$. Moreover, observe that the interval I is mapped in a one-to-one fashion onto some closed finite interval $[a, b]$ of the x axis which contains $x = 0$ in its interior.

In view of the assumptions of Sec. 2 and the relations (3.2), it can be established that $q(x, \lambda)$ and $f(x, \lambda)$ possess asymptotic expansions

$$q(x, \lambda) = \sum_{i=0}^{N} q_i(x)\lambda^{-i} + O(\lambda^{-N-1}),$$

$$f(x, \lambda) = \sum_{i=0}^{N} f_i(x)\lambda^{-i} + O(\lambda^{-N-1})$$

[6] J. L. Sanders, Jr. and A. A. Liepins, AIAA, 1, 2105 (1963).
[7] P. F. Jordan and P. E. Shelley, J. Math. Phys. 6, 118 (1965).

for $|\lambda| \geq |\lambda_0|$, $x \in [a, b]$, where N is arbitrary and all $p(x)$, $q_i(x)$, $f_i(x)$ are of class $C^\infty[a, b]$. We now restrict our analysis to Eq. (3.1)

4. A FORMAL SOLUTION

We will now construct an expansion which formally satisfies (3.1), the differential equation in normal form. In general, the series will not converge. However, in a later section it will be shown that there exists a solution of (3.1) for which the formal series solution furnishes an asymptotic expansion.

With the following definitions,

$$M(y) \equiv d^2y/dz^2 + [K(\lambda) - z^2]y, \quad K(\lambda) = \sum_{i=0}^{\infty} k_i\lambda^{-i},$$

$$\hat{M}(y) \equiv d^2y/dz^2 + [\hat{K}(\lambda) - z^2]y, \quad \hat{K}(\lambda) = \sum_{i=0}^{\infty} \hat{k}_i\lambda^{-i},$$

where $z = \lambda^{\frac{1}{2}}x$ and k_i, $\hat{k}_i$, $j = 0, 1, 2, \cdots$, are constants (independent of λ) which will be specified later, observe that

$$\begin{aligned} d^2y/dx^2 - \lambda^2x^2y &= \lambda M(y) - \lambda K(\lambda)y \\ &= \lambda\hat{M}(y) - \lambda\hat{K}(\lambda)y. \end{aligned} \tag{4.1}$$

Denoting particular solutions (which will be specified later) of

$$M(y) = -1, \qquad \hat{M}(y) = -z, \tag{4.2}$$

by $T_{0,K}(z)$, $T_{1,\hat{K}}(z)$, respectively, we can formulate the following theorem.

Theorem 1: There exists a formal series solution of Equation (3.1) of the form

$$\begin{aligned} \tilde{y}(x, \lambda) = \lambda \sum_{n=0}^{\infty} \lambda^{-n}[&a_n(x)T_{0,K}(z) + \lambda^{-\frac{1}{2}}b_n(x)T_{1,\hat{K}}(z) \\ &+ \lambda^{-\frac{1}{2}}c_n(x)T'_{0,K}(z) + \lambda^{-1}\, d_n(x)T'_{1,\hat{K}}(z) + \lambda^{-1}e_n(x)], \end{aligned} \tag{4.3}$$

where the coefficient functions $a_n(x)$, $b_n(x)$, $\cdots$, $e_n(x)$, $n = 0, 1, 2, \cdots$ are all of class $C^\infty[a, b]$.

Proof. We wish to show that all the functions $a_n(x)$, $\cdots$, $e_n(x)$, can be calculated and are of class $C^\infty[a, b]$. Substituting (4.3) into (3.1), using (4.1), (4.2), and rearranging terms, we have

$$\begin{aligned} \lambda^2 \sum_{n=0}^{\infty} \lambda^{-n}\Big\{ &-\Big[a_n + xb_n + 2c'_{n-1} + d_{n-1} + 2x\, d'_{n-1} \\ &+ x^2e_n - e''_{n-2} + p(x)e_{n-1} + \sum_{i=0}^{n-2} e_iq_{n-i-2}(x) \Big] \\ &+ \Big[a''_{n-1} - \sum_{i=0}^{n} a_ik_{n-i} - p(x)a_n - \sum_{i=0}^{n-1} a_iq_{n-i-1}(x) \\ &+ 2xc_n + 2x^2c'_n - 2\sum_{i=0}^{n-1} c'_ik_{n-i-1} \Big]T_{0,K} \\ &+ \Big[b''_{n-1} - \sum_{i=0}^{n} b_i\hat{k}_{n-i} - p(x)b_n - \sum_{i=0}^{n-1} b_iq_{n-i-1}(x) \\ &+ 2x\, d_n + 2x^2\, d'_n - 2\sum_{i=0}^{n-1} d'_i\hat{k}_{n-i-1} \Big]\lambda^{-\frac{1}{2}}T_{1,\hat{K}} \\ &+ \Big[2a'_n + c''_{n-1} - \sum_{i=0}^{n} c_ik_{n-i} \\ &- p(x)c_n - \sum_{i=0}^{n-1} c_iq_{n-i-1}(x) \Big]\lambda^{-\frac{1}{2}}T'_{0,K} \\ &+ \Big[2b'_n + d''_{n-1} - \sum_{i=0}^{n} d_i\hat{k}_{n-i} - p(x)\, d_n \\ &- \sum_{i=0}^{n-1} d_iq_{n-i-1}(x) \Big]\lambda^{-1}T'_{1,\hat{K}} \Big\} = \lambda^2 \sum_{i=0}^{n} \lambda^{-n}f_i(x), \end{aligned} \tag{4.4}$$

where we agree that all terms with negative subscripts are defined to be zero. Note the similarity between the coefficients of $T_{0,K}$ and $\lambda^{-\frac{1}{2}}T_{1,\hat{K}}$ and between those of $\lambda^{-\frac{1}{2}}T'_{0,K}$ and $\lambda^{-1}T'_{1,\hat{K}}$ (i.e. interchange a_n, c_n, k_i with b_n, d_n, $\hat{k}_i$ for all n, j). Equation (4.4) will become an identity provided the following relationships are satisfied for $n = 0, 1, 2, \cdots$:

$$2(x^2c'_n + xc_n) - [p(x) + k_0]a_n = h_{1,n}(x), \tag{4.5}$$

$$2a'_n - [p(x) + k_0]c_n = h_{2,n}(x), \tag{4.6}$$

$$2(x^2\, d'_n + x\, d_n) - [p(x) + \hat{k}_0]b_n = \hat{h}_{1,n}(x), \tag{4.7}$$

$$2b'_n - [p(x) + \hat{k}_0]\, d_n = \hat{h}_{2,n}(x), \tag{4.8}$$

$$\begin{aligned} x^2e_n = -\Big[&f_n(x) + a_n + xb_n + 2c'_{n-1} + d_{n-1} + 2x\, d'_{n-1} \\ &- e''_{n-2} + p(x)e_{n-1} + \sum_{i=0}^{n-2} e_iq_{n-i-2}(x) \Big], \end{aligned} \tag{4.9}$$

where

$$\begin{aligned} h_{1,n}(x) &= \sum_{i=0}^{n-1} \{2c'_ik_{n-i-1} \\ &\qquad + [q_{n-i-1}(x) + k_{n-i}]a_i\} - a''_{n-1}, \\ \hat{h}_{1,n}(x) &= \sum_{i=0}^{n-1} \{2\, d'_i\hat{k}_{n-i-1} \\ &\qquad + [q_{n-i-1}(x) + \hat{k}_{n-i}]b_i\} - b''_{n-1}, \\ h_{2,n}(x) &= \sum_{i=0}^{n-1} [q_{n-i-1}(x) + k_{n-i}]c_i - c''_{n-1}, \\ \hat{h}_{2,n}(x) &= \sum_{i=0}^{n-1} [q_{n-i-1}(x) + \hat{k}_{n-i}]\, d_i - d''_{n-1}. \end{aligned}$$

From these equations, we calculate the coefficient functions $a_n(x)$, $b_n(x)$, $\cdots$, $e_n(x)$, $n = 0, 1, 2, \cdots$. Recalling that all quantities with negative subscripts are defined to be zero, we have $h_{1,0}(x) \equiv$

$h_{2,0}(x) \equiv 0$. We now solve the system of differential equations (4.5) and (4.6) with $n = 0$ and obtain

$$a_0(x) = a_0(0) \cosh Q(x), \quad c_0(x) = a_0(0)[\sinh Q(x)]/x,$$

where

$$Q(x) = \tfrac{1}{2} \int_0^x (1/s)[p(s) + k_0]\, ds.$$

Clearly, it is necessary to set $k_0 = -p(0)$ to ensure the existence of the above integral. From (4.9) we have

$$x^2 e_0(x) = -[f_0(x) + a_0(x) + x b_0(x)],$$

and consequently, in order that $e_0(x)$ be bounded on $[a, b]$, we must select $a_0(0) = -f_0(0)$ and $b_0(0) = -f_0'(0)$. Similarly, on solving (4.7) and (4.8) with $n = 0$, we find it also necessary to set $\hat{k}_0 = -p(0)$. Finally, replacing k_0 by $-p(0)$ in the definition of $Q(x)$, we have

$$a_0(x) = -f_0(0) \cosh Q(x), \quad c_0(x) = -f_0(0)[\sinh Q(x)]/x,$$

$$b_0(x) = -f_0'(0) \cosh Q(x), \quad d_0(x) = -f_0'(0)[\sinh Q(x)]/x,$$

$$e_0(x) = -\{f_0(x) - [f_0(0) + x f_0'(0)] \cosh Q(x)\}/x^2.$$

Since $p(x)$ and $f_0(x)$ are of class $C^\infty[a, b]$, so are the coefficient functions $a_0(x), b_0(x), \cdots, e_0(x)$.

We can proceed in a recursive fashion using (4.9) to determine $a_n(0)$, $b_n(0)$, and $e_n(x)$ and solving the systems (4.5), (4.6) and (4.7), (4.8) to obtain $a_n(x)$, $b_n(x)$, $c_n(x)$, and $d_n(x)$. In the process of solving these systems of differential equations, it will be necessary to specify the constants k_j, $\hat{k}_j$, $j = 1, 2, 3, \cdots$, in a unique manner which will be outlined below.

Solving the system (4.5), (4.6) we obtain

$$a_n(x) = (\tfrac{1}{2}) \int_0^x [h_{1,n}(s)/s] \sinh [Q(x) - Q(s)]\, ds + (\tfrac{1}{2}) \int_0^x h_{2,n}(s) \cosh [Q(x) - Q(s)]\, ds + a_n(0) \cosh Q(x),$$

$$c_n(x) = (\tfrac{1}{2}x) \int_0^x [h_{1,n}(s)/s] \cosh [Q(x) - Q(s)]\, ds + (\tfrac{1}{2}x) \int_0^x h_{2,n}(s) \sinh [Q(x) - Q(s)]\, ds + a_n(0)[\sinh Q(x)]/x.$$

In order to guarantee the existence of the above integrals, for each n we must make $h_{1,n}(x)$ vanish to at least first order at $x = 0$. If $f_0(0) \neq 0$, and hence $a_0(x) \neq 0$, we select k_n so that $h_{1,n}(0) = 0$ for each n. Since $h_{1,n}$ is linear in k_n, this choice is always possible and, furthermore, the differentiability requirements on all the functions $p(x)$, $q_j(x)$, $f_j(x)$ imply that $h_{1,n}(x)$ vanishes to at least first order. If $f_0(0) = 0$, and consequently $a_0(0) = 0$, then $a_0(x) \equiv c_0(x) \equiv h_{1,1}(x) \equiv 0$. If, in addition, $p(x)$, $q(x, \lambda)$, $f(x, \lambda)$ are such that $a_1(0) \neq 0$, we select k_j such that $h_{1,1+j}(0) = 0$, $j = 1, 2, \cdots$. In general, if $p(x)$, $q(x, \lambda)$, $f(x, \lambda)$ are such that $a_0(0) = a_1(0) = \cdots = a_{m_1-1}(0) = 0$, $a_{m_1}(0) \neq 0$, we select k_j such that $h_{1,m_1+j}(0) = 0$, $j = 1, 2, \cdots$ and observe that $a_{n-1}(x) \equiv c_{n-1}(x) \equiv h_{1,n}(x) \equiv 0$ for $n = 1, \cdots, m_1$.

Similarly, we solve the system (4.7), (4.8) and obtain

$$b_n(x) = (\tfrac{1}{2}) \int_0^x [\hat{h}_{1,n}(s)/s] \sinh [Q(x) - Q(s)]\, ds + (\tfrac{1}{2}) \int_0^x \hat{h}_{2,n}(s) \cosh [Q(x) - Q(s)]\, ds + b_n(0) \cosh Q(x),$$

$$d_n(x) = (\tfrac{1}{2}x) \int_0^x [\hat{h}_{1,n}(s)/s] \cosh [Q(x) - Q(s)]\, ds + (\tfrac{1}{2}x) \int_0^x \hat{h}_{2,n}(s) \sinh [Q(x) - Q(s)]\, ds + b_n(0)[\sinh Q(x)]/x.$$

If $p(x)$, $q(x, \lambda)$, $f(x, \lambda)$ are such that $b_0(0) = b_1(0) = \cdots = b_{m_2-1}(0) = 0$, $b_{m_2}(0) \neq 0$, we observe that $b_{n-1}(x) \equiv d_{n-1}(x) \equiv \hat{h}_{1,n}(x) \equiv 0$, $n = 1, \cdots, m_2$, and consequently select $\hat{k}_j$ such that $\hat{h}_{1,m_2+j}(0) = 0$, $j = 1, 2, \cdots$. It can be shown by induction that all of the coefficient functions are of class $C^\infty[a, b]$.

5. PROPERTIES OF SOLUTIONS OF $T_{0,K}'' + [K - z^2]T_{0,K} = -1$ AND $T_{1,\hat{K}}'' + [\hat{K} - z^2]T_{1,\hat{K}} = -z$

Solutions of the associated homogeneous differential equation $M(y) = 0$ can be expressed in terms of parabolic cylinder functions (see Ref. 8, pp. 91, 92). The parabolic cylinder function of order ν, $D_\nu(z)$, is that solution of

$$D_\nu''(z) + (\nu + \tfrac{1}{2} - \tfrac{1}{4}z^2)\, D_\nu(z) = 0 \qquad (5.1)$$

which satisfies the initial conditions

$$D_\nu(0) = \Gamma(\tfrac{1}{2})2^{\frac{1}{2}\nu}/\Gamma[\tfrac{1}{2}(1 - \nu)], \quad D_\nu'(0) = \Gamma(-\tfrac{1}{2})2^{\frac{1}{2}(\nu-1)}/\Gamma(-\tfrac{1}{2}\nu). \qquad (5.2)$$

[8] W. Magnus and F. Oberhettinger, *Formulas and Theorems for the Special Functions of Mathematical Physics* (Chelsea Publishing Company, New York, 1949).

For $|z| \geq |z_0|$, $|z| \gg |\nu|$, it has the asymptotic representation

$$D_\nu(z) \sim \exp(-\tfrac{1}{4}z^2)\cdot z^\nu[1 + O(z^{-2})] \quad \text{for} \quad -\tfrac{3}{4}\pi < \arg z < \tfrac{3}{4}\pi$$

$$\sim \exp(-\tfrac{1}{4}z^2)\cdot z^\nu[1 + O(z^{-2})] - \frac{(2\pi)^{\frac{1}{2}}}{\Gamma(-\nu)} e^{\nu\pi i} \exp(\tfrac{1}{4}z^2)\cdot z^{-\nu-1}[1 + O(z^{-2})] \quad \text{for} \quad \tfrac{1}{4}\pi < \arg z < \tfrac{5}{4}\pi$$

$$\sim \exp(-\tfrac{1}{4}z^2)\cdot z^\nu[1 + O(z^{-2})] - \frac{(2\pi)^{\frac{1}{2}}}{\Gamma(-\nu)} e^{-\nu\pi i} \exp(\tfrac{1}{4}z^2)\cdot z^{-\nu-1}[1 + O(z^{-2})] \quad \text{for} \quad -\tfrac{5}{4}\pi < \arg z < -\tfrac{1}{4}\pi. \tag{5.3}$$

Furthermore, the functions $D_\nu(-z)$, $D_{-\nu-1}(iz)$, $D_{-\nu-1}(-iz)$ are also solutions of (5.1) and if β is a non-negative integer, $D_\beta(z) \to 0$ as $z \to \infty$ with $|\arg z| < \frac{1}{4}\pi$ or $|\arg z - \pi| < \frac{1}{4}\pi$.

Until now it has only been required that $T_{0,K}(z)$, $T_{1,\hat{K}}(z)$ be solutions of $M(y) = -1$, $\hat{M}(y) = -z$, respectively. Since the particular solutions of (3.1) which are convenient in the applications are those which, outside of a neighborhood of the turning point, remain bounded as $\lambda \to \infty$ along straight lines through $\lambda = 0$, we select the particular solutions of $M(y) = -1$, $\hat{M}(y) = -z$, accordingly. Consequently, since $z = \lambda^{\frac{1}{4}}x$, it is desirable to choose T-functions which remain bounded as $z \to \infty$ along straight lines through the origin in the complex z-plane.

We will now show that there are solutions $T_{0,K}(z)$, $T_{1,\hat{K}}(z)$ of $M(y) = -1$, $\hat{M}(y) = -z$, respectively, having the following properties. With

$$I(t, z, K) = \frac{1}{2}\frac{(1+t)^{\frac{1}{4}K}(1-t)^{-\frac{1}{4}K}}{(1-t^2)^{\frac{1}{2}}} \exp(-\tfrac{1}{2}z^2t)$$

and θ fixed such that $0 < \theta < \Delta \ll 1$, we have

$$T_{0,K}(z) = \int_0^1 I(t, z, K)\, dt \quad \text{for } |z| < \infty, \ \operatorname{Re} K \leq 1 - \delta < 1 \tag{5.4a}$$

$$= \int_0^\infty I(te^{i\theta}, z, K)\, e^{i\theta}\, dt + C(K)\, D_\nu(2^{\frac{1}{2}}z)$$

for $|z| > 0$, $|\arg z| \leq \frac{1}{4}\pi - \Delta$, $|K| < \infty$ (5.4b)

$$= \int_0^\infty I(te^{i\theta}, z, K)\, e^{i\theta}\, dt + C(K)\, D_\nu(-2^{\frac{1}{2}}z)$$

for $|z| > 0$, $|\arg z - \pi| \leq \frac{1}{4}\pi - \Delta$, $|K| < \infty$, (5.4c)

where

$$\nu = \tfrac{1}{2}(K-1),$$

$$C(K) = -\frac{2^{-\frac{1}{2}}}{D'_\nu(0)} \lim \frac{d}{dz}\int_0^\infty I(te^{i\theta}, z, K)\, e^{i\theta}\, dt,$$

the limit being taken as $z \to 0$ with $|\arg z| \leq \frac{1}{4}\pi - \Delta$. Also, with

$$\hat{I}(t, z, \hat{K}) = \frac{z}{2}\frac{(1+t)^{\frac{1}{4}\hat{K}}(1-t)^{-\frac{1}{4}\hat{K}}}{(1-t^2)^{\frac{1}{2}}} \exp(-\tfrac{1}{2}z^2t)$$

and θ as above, we have

$$T_{1,\hat{K}}(z) = \int_0^1 \hat{I}(t, z, \hat{K})\, dt \quad \text{for } |z| < \infty, \ \operatorname{Re}\hat{K} \leq 3 - \delta \tag{5.5a}$$

$$= \int_0^\infty \hat{I}(te^{i\theta}, z, \hat{K})\, e^{i\theta}\, dt + \hat{C}(\hat{K})\, D_{\hat{\nu}}(2^{\frac{1}{2}}z)$$

for $|z| > 0$, $|\arg z| \leq \frac{1}{4}\pi - \Delta$, $|K| < \infty$ (5.5b)

$$= \int_0^\infty \hat{I}(te^{i\theta}, z, \hat{K})\, e^{i\theta}\, dt - \hat{C}(\hat{K})\, D_{\hat{\nu}}(-2^{\frac{1}{2}}z)$$

for $|z| > 0$, $|\arg z - \pi| \leq \frac{1}{4}\pi - \Delta$, $|K| < \infty$, (5.5c)

where

$$\hat{\nu} = \tfrac{1}{2}(\hat{K}-1),$$

$$\hat{C}(\hat{K}) = -\frac{1}{D_{\hat{\nu}}(0)} \lim \int_0^\infty \hat{I}(te^{i\theta}, z, \hat{K})\, e^{i\theta}\, dt,$$

the limit being taken as $z \to 0$ with $|\arg z| \leq \frac{1}{4}\pi - \Delta$.

Except for those values of K, $\hat{K}$ which lie in arbitrarily small neighborhoods of the eigenvalues corresponding to solutions of $M(y) = 0$ which approach 0 as $z \to \pm\infty$ for real z, they have asymptotic expansions

$$T_{0,K}(z) = \sum_{n=0}^{N} A_n(K) z^{-2(n+1)} + O(z^{-2(N+2)}),$$

$$T_{1,\hat{K}}(z) = \sum_{n=0}^{N} B_n(K) z^{-(2n+1)} + O(z^{-(2N+3)}),$$

where

$$A_0(K) = 1,$$

$$A_n(K) = KA_{n-1}(K) + 2(n-1)(2n-1)A_{n-2}(K) \quad \text{for } n \geq 1,$$

$$B_0(\hat{K}) = 1,$$

$$B_n(\hat{K}) = \hat{K}B_{n-1}(\hat{K}) - 2(n-1)(2n-3)B_{n-2}(\hat{K}) \quad \text{for } n \geq 1,$$

$$|\arg z| \leq \tfrac{1}{4}\pi - \Delta, \qquad |\arg z - \pi| \leq \tfrac{1}{4}\pi - \Delta,$$
$$|z| \geq |z_0|, \qquad 0 < \Delta \ll 1,$$

N arbitrary. For all finite K, $\hat{K}$, z; $T_{0,K}(z)$, $T_{1,\hat{K}}(z)$ have power series representations

$$T_{0,K}(z) = \sum_{n=0}^{\infty} C_n(K) z^{2n}, \qquad T_{1,\hat{K}}(z) = \sum_{n=0}^{\infty} D_n(\hat{K}) z^{2n+1},$$

where

$$C_0(K) = T_{0,K}(0), \qquad D_0(K) = T'_{1,\hat{K}}(0),$$
$$C_1(K) = -[KC_0(K) + 1]/2,$$
$$C_n(K) = [C_{n-2}(K) - KC_{n-1}(K)]/[2n(2n-1)] \quad \text{for } n \geq 2,$$
$$D_1(\hat{K}) = -[\hat{K}D_0(\hat{K}) + 1]/6,$$
$$D_n(\hat{K}) = [D_{n-2}(\hat{K}) - \hat{K}D_{n-1}(\hat{K})]/[2n(2n+1)] \quad \text{for } n \geq 2.$$

By substituting directly and integrating by parts, it is seen that the representations (5.4a–c) satisfy $M(y) = -1$ in the indicated regions of the z-plane. With the given value of $C(K)$, (5.4b, c) satisfy the same initial conditions (in the limit) at $z = 0$ and, consequently, represent the same solution. This solution approaches zero as $z \to \pm\infty$, z real. However, so does the function (5.4a) and since there are no solutions of the associated homogeneous equation with that property (except for those corresponding to a discrete set of eigenvalues), (5.4a–c) all represent the same solution in the appropriate regions of the z-plane. The representations (5.5a–c) follow analogously.

On the other hand, the functions defined by

$$T^*_{0,K}(z) \equiv -T_{0,-K}(iz), \qquad T^*_{1,\hat{K}}(z) \equiv iT_{1,-\hat{K}}(iz)$$

are also solutions of $M(y) = -1$, $M(y) = -z$, respectively. Their properties follow immediately from those of $T_{0,K}(z)$ and $T_{1,\hat{K}}(z)$. Recalling that $z = \lambda^{\frac{1}{3}}x$, note that $T_{0,K}(z)$, $\lambda^{-\frac{1}{3}}T_{1,\hat{K}}(z)$, $\lambda^{-\frac{1}{3}}T'_{0,K}(z)$, $\lambda^{-1}T'_{1,\hat{K}}(z)$ are uniformly bounded in the sectors defined by $|\arg z| \leq \frac{1}{4}\pi$ and $|\arg z - \pi| \leq \frac{1}{4}\pi$ provided $\operatorname{Re} K \leq 1 - \delta$. Also, the functions $T^*_{0,K}(z)$, $\lambda^{-\frac{1}{3}}T^*_{0,\hat{K}}(z)$, $\lambda^{-\frac{1}{3}}T^{*\prime}_{0,K}(z)$, $\lambda^{-1}T^*_{1,\hat{K}}(z)$ are uniformly bounded in the sectors defined by

$$|\arg z - \tfrac{1}{2}\pi| \leq \tfrac{1}{4}\pi \quad \text{and} \quad |\arg z + \tfrac{1}{2}\pi| \leq \tfrac{1}{4}\pi$$

provided $\operatorname{Re} K \geq -1 + \delta$.

Values of $T_{0,0}(z)$, $T_{1,0}(z)$ (i.e., $K = \hat{K} = 0$) and their first derivatives have been tabulated in Refs. 6 and 7 for real positive z. In the latter, the relationship between these functions and Lommel functions is discussed while in the former, a contour integral representation for $T_{1,0}(z)$ is given in the form of an inverse Mellin transform.

6. ASYMPTOTIC BEHAVIOR OF THE FORMAL SOLUTION

In this section, the asymptotic nature of the formal series solution will be discussed. Consider finite truncations of the formal expansion denoted by

$$\tilde{y}_N(x, \lambda) = \lambda \sum_{n=0}^{N} \lambda^{-n}[a_n(x)T_{0,K_N}(z) + \lambda^{-\frac{1}{3}}b_n(x)T_{1,\hat{K}_N}(z) + \lambda^{-\frac{1}{3}}c_n(x)T'_{0,K_N}(z) + \lambda^{-1}d_n(x)T'_{1,\hat{K}_N}(z) + \lambda^{-1}e_n(x)] \tag{6.1}$$

and $\tilde{y}^*_N(x, \lambda)$, where $\tilde{y}^*_N$ is formed from $\tilde{y}_N$ by replacing the T functions with the corresponding T^* functions. Here K_N, $\hat{K}_N$ are given by

$$K_N = \sum_{i=0}^{N-m_1} k_i\lambda^{-i}, \qquad \hat{K}_N = \sum_{i=0}^{N-m_2} \hat{k}_i\lambda^{-i},$$

assuming quantities with negative subscripts to be identically zero, and all the quantities $a_n(x), \cdots, e_n(x)$, k_i, $\hat{k}_i$ satisfy the recursion relations of Sec. 4. The numbers m_1, m_2 are such that

$$a_0(0) = a_1(0) = \cdots = a_{m_1-1}(0) = 0, \qquad a_{m_1}(0) \neq 0$$

and

$$b_0(0) = b_1(0) = \cdots = b_{m_2-1}(0) = 0, \qquad b_{m_2}(0) \neq 0.$$

Finally, for $0 < \rho \ll 1$, let $R = \bigcup_{j\in J} N(j, \rho)$ where J is the set of odd positive integers and $N(j, \rho)$ is a ρ-neighborhood of j. Then:

Theorem 2: If $|\lambda| \geq |\lambda_0|$, $-(\operatorname{sgn} \operatorname{Re} \lambda)p(0) \notin R$ and (a) if $|\arg \lambda| \leq \frac{1}{2}\pi - 2\Delta$, then there exists a solution $y(x, \lambda)$ of

$$Ly \equiv d^2y/dx^2 - [\lambda^2x^2 + \lambda p(x) + q(x, \lambda)]y = \lambda^2 f(x, \lambda) \tag{6.2}$$

such that for arbitrary N,

$$y(x, \lambda) = \tilde{y}_N(x, \lambda) + O(\lambda^{-N}) \tag{6.3}$$

uniformly in x for $x \in [a, b]$. (b) If $|\arg \lambda - \pi| \leq \frac{1}{2}\pi - 2\Delta$, then there exists a solution $y^*(x, \lambda)$ of (6.2) such that for arbitrary N,

$$y^*(x, \lambda) = \tilde{y}^*_N(x, \lambda) + O(\lambda^{-N})$$

uniformly in x for $x \in [a, b]$.

Theorem 3: If $|\lambda| \geq |\lambda_0|$, conclusion (a) of Theorem 2 is valid with $\Delta = 0$ provided $\operatorname{Re} p(0) \geq +\delta$ and (b) is valid with $\Delta = 0$ provided $\operatorname{Re} p(0) \leq -\delta$.

Remarks: The above results show that, using the appropriate special functions, the formal solution is an asymptotic expansion of some solution of the nonhomogeneous differential equation (6.2) in the following sense. If we terminate our calculations after computing the first N terms (i.e., after computing $a_n, b_n, \cdots, e_n$, $n = 1, \cdots, N$ and the $k_i, \hat{k}_i$'s needed in the process), the error encountered is of the same order in λ as the first term omitted [i.e. $O(\lambda^{-N})$].

If $p(x)$, $q_j(x)$ or $f_j(x)$ for some j is not of class $C^\infty[a, b]$, but possesses only a finite number of continuous derivatives, the above theorems are true for $N \leq N^*$ for some suitable N^*.

Note that in Theorem 3 we can use the simple integral representations (5.4a), (5.5a) for $T_{0,K}(z)$, $T_{1,\hat{K}}(z)$ as well as the corresponding ones for the T^* functions.

Proof of Theorem 2. R. W. McKelvey[3] has shown that, by use of transformations, it is possible to construct a "related" differential equation with explicitly known solutions which approximates the homogeneous differential equation associated with (6.2) to terms of order λ^{-N} for arbitrary N. From McKelvey's work it follows that functions

$$Y_n(x, \lambda) = \mu_0(x, \lambda) V_n(x, \lambda) + \mu_1(x, \lambda) V_n'(x, \lambda)/\lambda, \quad n = 0, 1, 2, 3, \tag{6.4}$$

can be constructed which are solutions of

$$d^2Y/dx^2 - [\lambda^2 x^2 + \lambda p(x) + T(x, \lambda)]Y = 0$$

such that the Wronskians (with respect to x) satisfy

$$W_x(Y_n, Y_i) = W_x(V_n, V_i).$$

Here

$$V_n(x, \lambda) = D_\nu(2^{\frac{1}{2}} e^{-(\frac{1}{2})n\pi i} \lambda^{\frac{1}{2}} x),$$

$$\nu = 2(-1)^n k - \tfrac{1}{2}, \quad k = -p(0)/4 + k^*/\lambda. \tag{6.5}$$

k^* is a constant and $\mu_0(x, \lambda)$, $\mu_1(x, \lambda)$, $T(x, \lambda)$ are well defined functions which are uniformly bounded in x and λ for $|\lambda| \geq |\lambda_0|$, $x \in [a, b]$.

Using the initial values (5.2) of the parabolic cylinder function, we see that

$$W_x(Y_0, Y_2) = -\frac{\lambda^{\frac{1}{2}}\Gamma(-\frac{1}{2})}{2\Gamma[2(\frac{1}{4} - k)]},$$

$$W_x(Y_1, Y_2) = i\frac{\lambda^{\frac{1}{2}}\Gamma(-\frac{1}{2})}{2\Gamma[2(\frac{1}{4} + k)]}.$$

First, assume $|\arg \lambda| \leq \frac{1}{4}\pi - 2\Delta$. Then, $T_{0,K}(z)$, $\lambda^{-\frac{1}{2}} T_{1,\hat{K}}(z)$, $\lambda^{-\frac{1}{2}} T_{0,K}'(z)$, $\lambda^{-1} T_{1,\hat{K}}'(z)$ and all the coefficient functions $a_n, b_n, \cdots, e_n$, $n = 0, 1, 2, \cdots$, are bounded. Consequently, applying the operator L defined by (6.2) to $\tilde{y}_N(x, \lambda)$ leads to a result which may be written

$$L\tilde{y}_N(x, \lambda) = \lambda^2 f(x, \lambda) - \lambda^{-N+1}\alpha_N(x, \lambda),$$

where $\alpha_N(x, \lambda)$ is bounded independent of x and λ for $|\lambda| \geq \lambda_0|$, $x \in [a, b]$. Defining $y(x, \lambda)$ by

$$y(x, \lambda) = \tilde{y}_N(x, \lambda) + w(x, \lambda), \tag{6.6}$$

where $w(x, \lambda)$ satisfies

$$Lw = \lambda^{-N+1}\alpha_N(x, \lambda), \tag{6.7}$$

we conclude that $y(x, \lambda)$ is a solution of the original differential equation $Ly = \lambda^2 f(x, \lambda)$. Therefore, in view of (6.6), we need to show that there exists a solution $w(x, \lambda)$ of (6.7) which is of a suitable order in λ for $|\lambda| \geq |\lambda_0|$, $x \in [a, b]$.

Let us rewrite (6.7) in a form which allows a Green's function approach:

$$L^*w \equiv d^2w/dx^2 - [\lambda^2 x^2 + \lambda p(x) + T(x, \lambda)]w = \lambda^{-N+1}\alpha_N(x, \lambda) + [q(x, \lambda) - T(x, \lambda)]w.$$

With the construction of a Green's function $G(x, s; \lambda)$ associated with the operator L^*, we write

$$w(x, \lambda) = \lambda^{-N+\frac{1}{2}}\beta_N(x, \lambda) + \lambda^{-\frac{1}{2}} \times \int_a^b \lambda^{\frac{1}{2}} G(x, s; \lambda)[q(s, \lambda) - T(s, \lambda)]w(s, \lambda)\, ds, \tag{6.8}$$

where

$$\beta_N(x, \lambda) = \int_a^b \lambda^{\frac{1}{2}} G(x, s; \lambda)\alpha_N(s, \lambda)\, ds.$$

Any solution of (6.8) is a solution of (6.7).

To obtain a suitable extimate on the magnitude of $w(x, \lambda)$, $G(x, s; \lambda)$ must be chosen to be uniformly bounded for $x \in [a, b]$, $|\lambda| \geq |\lambda_0|$. One such function will be uniquely specified by the condition

$$\lim_{|x| \to \infty} G(x, s; \lambda) = 0. \tag{6.9}$$

For $|\arg \lambda| \leq \frac{1}{2}\pi - 2\Delta$, an appropriate Green's function satisfying (6.9) is given by

$$G(x, s; \lambda) = \begin{cases} \dfrac{Y_0(x, \lambda) Y_2(s, \lambda)}{W_x(Y_0, Y_2)}, & s \leq x, \\ \dfrac{Y_0(s, \lambda) Y_2(x, \lambda)}{W_x(Y_0, Y_2)}, & x \leq s. \end{cases}$$

Observing (6.4), (6.5), and using the asymptotic representation (5.3) for the parabolic cylinder function, we have

$$|Y_0(x, \lambda) e^{\lambda|x|x/2}(\lambda^{\frac{1}{2}}x)^{\frac{1}{2} - 2kx/|x|}| < M, \quad |Y_2(x, \lambda) e^{-\lambda|x|x/2}(\lambda^{\frac{1}{2}}x)^{\frac{1}{2} + 2kx/|x|}| < M, \tag{6.10}$$

for some constant M, and sufficiently large $|\lambda^{\frac{1}{2}}x|$. Since the parabolic cylinder function is entire, it is sufficient to consider the estimate (6.10) to determine an upper bound on $|G(x, s; \lambda)|$. By splitting the exponential term that arises into two parts [one being used to absorb factors of the form $(x/s)^{2k}$ which could be arbitrarily large], one can eventually obtain

$$|\lambda^{\frac{1}{2}}G(x, s; \lambda)| \leq M_1(\rho, \Delta) \exp [-\epsilon(\Delta)\cdot|\lambda|\cdot|x^2 - s^2|],$$

where $M_1(\rho, \Delta)$ is independent of x, λ and $\operatorname{Re} \lambda > 4\epsilon(\Delta)\,|\lambda|$ for some $\epsilon(\Delta) > 0$. Consequently,

$$\int_a^b |\lambda^{\frac{1}{2}}G(x, s; \lambda)|\, ds < (b - a)M_1(\rho, \Delta). \tag{6.11}$$

It is a known result from integral equation theory that if $r > 0$, $|g(x, \lambda)| \leq M_1$, $\int_a^b |K(x, s; \lambda)|\, ds < M_1$ where r, M_1 are constants and $x \in [a, b]$, $|\lambda| \geq |\lambda_0|$, then there exists a unique solution $w(x, \lambda)$ of the Fredholm integral equation

$$w(x, \lambda) = \lambda^{-n}g(x, \lambda) + \lambda^{-r}\int_a^b K(x, s; \lambda)w(s, \lambda)\, ds$$

and a constant M such that

$$|w(x, \lambda)| \leq M\,|\lambda|^{-n}.$$

Since q, T and α_N are bounded, we can apply this result to (6.8) and conclude that there exists a unique solution $w(x, \lambda) = O(\lambda^{-N+\frac{1}{2}})$ of (6.8). Consequently, there exists a solution of (6.2) such that

$$y(x, \lambda) = \tilde{y}_N(x, \lambda) + O(\lambda^{-N+\frac{1}{2}}).$$

Since

$$\tilde{y}_N(x, \lambda) - \tilde{y}_{N-1}(x, \lambda) = O(\lambda^{-N+1}),$$

we have

$$y(x, \lambda) = \tilde{y}_{N-1}(x, \lambda) + O(\lambda^{-(N-1)})$$

and since N is arbitrary,

$$y(x, \lambda) = \tilde{y}_N(x, \lambda) + O(\lambda^{-N})$$

uniformly in x for $|\lambda| \geq |\lambda_0|$, $x \in [a, b]$.

If $|\arg \lambda - \pi| \leq \frac{1}{2}\pi - 2\Delta$, the same analysis leads to the desired conclusion provided Y_1, Y_3 are used in place of Y_0, Y_2.

Proof of Theorem 3. The proof differs from that of Theorem 2 in only one place. To obtain the estimate (6.11) we cannot use part of the exponential term that arises to absorb the factors of the form $(x/s)^{2k}$ which occur (since here $\operatorname{Re} \lambda$ is allowed to be 0). However, for $|\arg \lambda| \leq \frac{1}{2}\pi$, estimate (6.11) can be obtained from (6.10) provided $\operatorname{Re} k \geq \delta$, $0 < \delta \ll 1$. Since k is arbitrarily close to $-\frac{1}{4}p(0)$ for $|\lambda| \geq |\lambda_0|$, (6.11) follows provided $\operatorname{Re} p(0) \geq +\delta$. If $|\arg \lambda - \pi| \leq \frac{1}{2}\pi$, (6.11) follows provided $\operatorname{Re} p(0) \leq -\delta$.

Actually, it is possible to prove Theorem 2 [and the special case $p(x) \equiv 0$ of Theorem 3] without using the results of McKelvey by constructing a Green's function (which is immediately expressible in terms of parabolic cylinder functions) for the operator

$$L^{**}w \equiv d^2w/dx^2 - [\lambda^2x^2 + \lambda p(0)]w.$$

The special case, $p(x) \equiv q(x, \lambda) \equiv 0$, has been considered by Sanders and Liepins.[6] In the notation of this paper, they have shown that a solution of

$$d^2y/dx^2 - \lambda^2x^2y = \lambda^{\frac{1}{2}}f_0(x) + O(\lambda) \tag{6.12}$$

can be represented asymptotically as

$$y(x, \lambda) = -\lambda f_0(0)T_{0,0}(z) - \lambda^{\frac{1}{2}}f_0'(0)T_{1,0}(z) - [f_0(x) - f_0(0) - xf_0'(0)]/x^2 + O(1) \tag{6.13}$$

for $x \in [a, b]$ and λ sufficiently large and positive. On the other hand, for this special case it can be shown by an inductive argument that all $a_n(x)$, $b_n(x)$ are constant and all $c_n(x)$, $d_n(x)$, k_n, $\hat{k}_n$ are identically zero in our expansion. Consequently, Theorem 3 not only yields (6.13), but asserts that this result remains valid for $|\arg \lambda| \leq \frac{1}{2}\pi$ and can be extended, by use of other special functions, to the case $|\arg \lambda - \pi| \leq \frac{1}{2}\pi$. Moreover, if we knew the differential equation (6.12) more precisely, it would be possible to calculate additional terms in the expansion.

ACKNOWLEDGMENTS

The results of this paper are contained in the author's Ph.D. thesis, written at Case Institute of Technology while he held a National Defense Education Act fellowship. The author wishes to thank Professor R. A. Clark for his guidance in the research.

A Short Simple Evaluation of Expressions of the Debye–Waller Form

N. DAVID MERMIN
Laboratory of Atomic and Solid State Physics, Cornell University, Ithaca, New York
(Received 15 November 1965)

Averages like those encountered in the theory of the Debye-Waller factor are evaluated in one sentence.

WHEN calculating absorption, emission, or scattering cross sections for crystalline matter in the harmonic approximation one needs the thermal equilibrium average of exponentials of operators linear in the atomic displacements and/or momenta:

$$\langle e^{\sum_i c_i a_i + d_i a_i^\dagger}\rangle = \mathrm{Tr}\, e^{-\beta H} e^{\sum_i c_i a_i + d_i a_i^\dagger}/\mathrm{Tr}\, e^{-\beta H},$$
$$H = \sum \omega_i (a_i^\dagger a_i + \tfrac{1}{2}),\quad \beta = 1/k_B T,\quad [a_i, a_i^\dagger] = \delta_{ij}. \tag{1}$$

This can be evaluated in a variety of ways,[1] some difficult, some direct, but all annoyingly cumbersome considering the simplicity of the final form. Here is a derivation as simple as the result:

As in most approaches, begin by using the well-known formula[1]

$$e^{A+B} = e^A e^B e^{-\frac{1}{2}[A,B]} \qquad ([A, B] \text{ a } c\text{-number}) \tag{2}$$

to reduce (1) to

$$\langle e^{\sum_i c_i a_i + d_i a_i^\dagger}\rangle = \langle e^{\sum_i c_i a_i} e^{\sum_i d_i a_i^\dagger}\rangle e^{-\frac{1}{2}\sum_i c_i d_i}, \tag{3}$$

but instead of proceeding with the clumsy direct evaluation of

$$g(c_i d_i) = \langle e^{\sum_i c_i a_i} e^{\sum_i d_i a_i^\dagger}\rangle, \tag{4}$$

note that (2) also entitles one to conclude

$$\langle e^{\sum_i c_i a_i + d_i a_i^\dagger}\rangle = \langle e^{\sum_i d_i a_i^\dagger} e^{\sum_i c_i a_i}\rangle e^{\frac{1}{2}\sum_i c_i d_i}, \tag{5}$$

which is consistent with (3) only if

$$\begin{aligned} g(c_i d_i) &= e^{\sum_i c_i d_i}\langle e^{\sum_i d_i a_i^\dagger} e^{\sum_i c_i a_i}\rangle \\ &= e^{\sum_i c_i d_i}\langle e^{\beta H} e^{\sum_i c_i a_i} e^{-\beta H} e^{\sum_i d_i a_i^\dagger}\rangle \\ &\quad \text{(cyclical permutation within trace)} \\ &= e^{\sum_i c_i d_i}\langle e^{\sum_i c_i e^{\beta H} a_i e^{-\beta H}} e^{\sum_i d_i a_i^\dagger}\rangle \\ &= e^{\sum_i c_i d_i}\langle e^{\sum_i c_i e^{-\beta\omega_i} a_i} e^{\sum_i d_i a_i^\dagger}\rangle \\ &\quad \text{(harmonic approximation)} \\ &= e^{\sum_i c_i d_i} g(c_i e^{-\beta\omega_i}, d_i), \end{aligned} \tag{6}$$

from which identity it follows at once (by iteration or induction on n) that

$$g(c_i d_i) = e^{\sum_i c_i d_i (1 + e^{-\beta\omega_i} + \cdots + e^{-n\beta\omega_i})} g(c_i e^{-(n+1)\beta\omega_i}, d_i), \tag{7}$$

and hence, taking the limit $n \to \infty$ (each ω_i is positive),

$$g(c_i d_i) = e^{\sum_i c_i d_i (1 - e^{-\beta\omega_i})^{-1}} g(0, d_i), \tag{8}$$

which, since it follows trivially from (4) that

$$g(0, d_i) = \langle e^{\sum_i d_i a_i^\dagger}\rangle = 1, \tag{9}$$

completes the derivation:

$$\begin{aligned} \langle e^{\sum_i c_i a_i + d_i a_i^\dagger}\rangle &= e^{-\frac{1}{2}\sum_i c_i d_i} g(c_i d_i) \\ &= e^{\frac{1}{2}\sum_i c_i d_i \coth \frac{1}{2}\beta\omega_i}. \end{aligned} \tag{10}$$

[1] See, for instance, A. A. Maradudin, E. W. Montroll, and G. H. Weiss, Solid State Phys. Suppl. 3, 239 (1963).

Derivation of the Generalized Boltzmann Equation in Quantum Statistical Mechanics*

TOHRU MORITA†
Department of Physics, The Catholic University of America, Washington, D. C.
(Received 23 August 1965)

The hierarchy of the equations of motion for the reduced density matrices in quantum statistical mechanics is solved and the (cumulant) reduced density matrices at a time t are expressed in terms of those at an earlier time t_0. Diagrams are introduced to express the results. With the aid of the technique of partial summations, the general term in the kinetic equation for the one-particle reduced density matrix or the generalized Boltzmann equation in quantum statistical mechanics is obtained. The equation is non-Markovian. A method of reducing the equation to Markovian is sketched.

I. INTRODUCTION

THE average value of one-particle quantity is calculated when we know the one-particle phase-space distribution function for a classical gas, which is classical in regard to the mechanics as well as to the statistics. The motion of the distribution function is described by the Boltzmann equation or its generalization. For a gas which obeys quantum mechanics, the one-particle reduced density matrix or the Wigner function plays the role of the distribution function. Recently Fujita[1] gave the general term in the generalized Boltzmann equation, which describes the motion of the Wigner function, for a gas which is governed by quantum mechanics but obeys classical statistics. In the derivation, the diagrammatical techniques which were developed by Prigogine and his collaborators[2] for their discussions of inhomogeneous classical gas, have been used.

For a quantum statistical gas which is quantum with respect to the mechanics as well as to the statistics, Fujita[3] applied the same arguments but could not reach the general result and alternatively gave the general term in the equation of motion for a one-particle two-time correlation function. It is the purpose of this paper to give the general term in the equation of motion for the one-particle reduced density matrix for the quantum statistical gas. This is done by integrating the hierarchy of the equations of motion for the reduced density matrices and expressing the result with the aid of the diagrams and reducing the result by taking partial summations.

Our result reduces to the one by Fujita[1] when we neglect the diagrams which contain the effect of the exchange of particles.

Konstantinov and Perel' and separately Resibois[4] have investigated the linear response of many fermion systems to an external field, where they obtained the general term in the kinetic equation for the one-particle reduced density matrix in the linear approximation. In the linear approximation, our generalized Boltzmann equation must be the same as Resibois',[4] but, because of the entirely different ways of introducing the diagrams, we have not been able to find the correspondence in the results. In both methods, eight diagrams are drawn in order to get the second-order terms in the kinetic equation, though the structures of the diagrams are entirely different.

The hierarchy of equations of motion for the reduced density matrices are set up in Sec. II and solved formally in Sec. III. The reduced density matrices at time t are expressed in terms of those at time t_0. The results are expressed in terms of the diagrams. In Sec. IV, the cumulant reduced density matrices are introduced, and the reduced density matrices and the cumulant reduced density matrices at t are expressed in terms of the cumulant reduced density matrices at t_0. In Sec. V, the partial summations are taken and the (cumulant) reduced density matrices at t are expressed in terms of the one-particle reduced density matrices at earlier times and the cumulant reduced density matrices at the initial time t_0. The result is used to obtain the kinetic equation for the one-particle reduced density matrix. A farther reduction of dia-

* This work was supported by the U. S. Air Force Office of Scientific Research, Grant No. AF-AFOSR-445-63.
† On leave of absence from Shizuoka University, Shizuoka, Japan.
[1] S. Fujita, J. Math. Phys. **6**, 1004 (1965).
[2] I. Prigogine and R. Balescu, Physica **25**, 281, 302 (1959). I. Prigogine, *Non-Equilibrium Statistical Mechanics* (Interscience Publishers, Inc., New York, 1962).
[3] S. Fujita, (preprint).
[4] P. Resibois, Phys. Rev. **138**, B281 (1965). O. V. Konstantinov and V. I. Perel', Zh. Eksperim. i Teor. Fiz. **39**, 197 (1960) [English transl.: Soviet Phys.—JETP **12**, 142 (1961)]. See references in these articles for other related papers on this problem.

grams is done in Sec. VI. The equation obtained is nonMarkovian. A method of reducing the resultant equation to the Markovian form is suggested in Sec. VII.

Following a suggestion by Fujita,[1] the plane-wave representation is not used and the representation which diagonalizes the one-particle time-independent part of the Hamiltonian is used; because then the result becomes applicable also, for instance, to a gas of charged particles under strong magnetic field. Since the case where the plane-wave representation is proper has, however, been of the primary importance, comments are given in footnotes for this special case.

II. EQUATIONS OF MOTION FOR THE REDUCED DENSITY MATRICES

We are interested in the change in time of the one-particle reduced density matrix given by[5]

$$\langle a_{k_1}^* a_{l_1} \rangle_t, \tag{1}$$

where

$$\langle A \rangle_t \equiv \mathrm{tr}\, A\rho(t). \tag{3}$$

$\rho(t)$ is the normalized density matrix of the system at time t and satisfies the equation of motion:

$$(\partial/\partial t)\rho(t) = -(i/\hbar)[H, \rho(t)], \tag{4}$$

where H is the Hamiltonian of the system. Taking a time derivative of Eq. (3) and using Eq. (4), one gets the equation of motion for $\langle A \rangle_t$ as follows:

$$\frac{\partial}{\partial t}\langle A \rangle_t = \mathrm{tr}\left\{A \frac{\partial \rho(t)}{\partial t}\right\} = -\frac{i}{\hbar}\mathrm{tr}\,\{A[H, \rho(t)]\} = -\frac{i}{\hbar}\langle [A, H] \rangle_t. \tag{5}$$

It is noted that this is valid even when H is time-dependent.

The Hamiltonian of the system is assumed to be[6]

$$H = \sum_k \epsilon(k) a_k^* a_k + \sum_k \sum_l \phi(l; k) a_k^* a_l + \tfrac{1}{2} \sum_{k_1} \sum_{k_2} \sum_{l_1} \sum_{l_2} v(l_1 l_2; k_1 k_2) a_{k_1}^* a_{k_2}^* a_{l_2} a_{l_1}, \tag{6}$$

where the labels k, l, $\cdots$ represent one-particle states which are eigenstates of the one-particle time-independent part of the Hamiltonian and $\epsilon(k)$, $\epsilon(l)$, $\cdots$ are the corresponding eigenvalues. The second term represents the effect of the external field which may depend on time, and the third represents the interaction of two particles.

The equation of motion for the one-particle reduced density matrix is obtained by substituting $a_{k_1}^* a_{l_1}$ for A, and the expression given by Eq. (6) for H in Eq. (5). As the result, one gets

$$\begin{aligned}(\partial/\partial t)\langle a_{k_1}^* a_{l_1} \rangle_t &= -(i/\hbar)[\epsilon(l_1) - \epsilon(k_1)]\langle a_{k_1}^* a_{l_1} \rangle_t \\ &- (i/\hbar) \sum_{l_1'} \phi(l_1'; l_1)\langle a_{k_1}^* a_{l_1'} \rangle_t \\ &+ (i/\hbar) \sum_{k_1'} \phi(k_1; k_1')\langle a_{k_1'}^* a_{l_1} \rangle_t \\ &- (i/\hbar) \sum_{k_2} \sum_{l_1'} \sum_{l_2'} v(l_1' l_2'; l_1 k_2)\langle a_{k_1}^* a_{k_2}^* a_{l_2'} a_{l_1'} \rangle_t \\ &+ (i/\hbar) \sum_{l_2} \sum_{k_1'} \sum_{k_2'} v(k_1 k_2; k_1' k_2')\langle a_{k_1'}^* a_{k_2}^* a_{l_2} a_{l_1} \rangle_t.\end{aligned} \tag{7}$$

This contains a two-particle reduced density matrix. The equation of motion for the n-particle reduced density matrix is generally given by

$$\begin{aligned}(\partial/\partial t)\langle a_{k_1}^* a_{k_2}^* \cdots a_{k_n}^* a_{l_n} \cdots a_{l_1} \rangle_t &= -(i/\hbar)\langle [a_{k_1}^* a_{k_2}^* \cdots a_{k_n}^* a_{l_n} \cdots a_{l_1}, H] \rangle_t \\ &= -(i/\hbar)\langle a_{k_1}^* a_{k_2}^* \cdots a_{k_n}^* [a_{l_n} \cdots a_{l_1}, H] \rangle_t - (i/\hbar)\langle [a_{k_1}^* a_{k_2}^* \cdots a_{k_n}^*, H] a_{l_n} \cdots a_{l_1} \rangle_t \\ &= -(i/\hbar) \sum_{i=1}^{n} [\epsilon(l_i) - \epsilon(k_i)]\langle a_{k_1}^* a_{k_2}^* \cdots a_{k_n}^* a_{l_n} \cdots a_{l_1} \rangle_t \\ &- (i/\hbar) \sum_{i=1}^{n} \sum_{l_i'} \langle a_{k_1}^* a_{k_2}^* \cdots a_{k_n}^* a_{l_n} \cdots a_{l_i'} \cdots a_{l_1} \rangle_t \phi(l_i'; l_i) \\ &+ (i/\hbar) \sum_{i=1}^{n} \sum_{k_i'} \phi(k_i; k_i')\langle a_{k_1}^* a_{k_2}^* \cdots a_{k_i'}^* \cdots a_{k_n}^* a_{l_n} \cdots a_{l_1} \rangle_t\end{aligned}$$

[5] The one-particle reduced density matrix $\langle a_{\mathbf{k}}^* a_{\mathbf{l}} \rangle$ in the plane wave representation is related with the Wigner function $\rho_w(\mathbf{k}, \mathbf{r})$, which is the quantum analogue of the one-particle phase-space distribution function, by

$$\rho_w(\mathbf{k}, \mathbf{r}) = \sum_{\mathbf{l}} \langle a_{\mathbf{k}+1/2}^* a_{\mathbf{k}-1/2} \rangle \exp\{-i\mathbf{l}\cdot\mathbf{r}\}. \tag{2}$$

[6] If the Hamiltonian is given by

$$H = -\frac{\hbar^2}{2m} \sum_i \frac{\partial^2}{\partial \mathbf{r}_i^2} + \sum_i \phi(\mathbf{r}_i) + \sum_{i>j} v(\mathbf{r}_i - \mathbf{r}_j)$$

as usual, Eq. (6) is written as

$$H = \sum_{\mathbf{k}} \epsilon(\mathbf{k}) a_{\mathbf{k}}^* a_{\mathbf{k}} + \sum_{\mathbf{k}} \sum_{\mathbf{q}} \phi(\mathbf{q}) a_{\mathbf{k}+\mathbf{q}}^* a_{\mathbf{k}} + \tfrac{1}{2} \sum_{\mathbf{k}_1} \sum_{\mathbf{k}_2} \sum_{\mathbf{q}} v(\mathbf{q}) a_{\mathbf{k}_1+\mathbf{q}}^* a_{\mathbf{k}_2-\mathbf{q}}^* a_{\mathbf{k}_2} a_{\mathbf{k}_1}. \tag{6'}$$

$$- (i/\hbar) \sum_{1\le i<j\le n} \sum_{l_i'} \sum_{l_j'} \langle a_{k_1}^* \cdots a_{k_n}^* a_{l_n} \cdots a_{l_j'} \cdots a_{l_i'} \cdots a_{l_1}\rangle_t v(l_i' l_j'; l_i l_j)$$

$$+ (i/\hbar) \sum_{1\le i<j\le n} \sum_{k_i'} \sum_{k_j'} v(k_i k_j; k_i' k_j') \langle a_{k_1}^* \cdots a_{k_i'}^* \cdots a_{k_j'}^* \cdots a_{k_n}^* a_{l_n} \cdots a_{l_1}\rangle_t$$

$$- (i/\hbar) \sum_{i=1}^{n} \sum_{k_{n+1}} \sum_{l_i'} \sum_{l_{n+1}'} \langle a_{k_1}^* \cdots a_{k_n}^* a_{k_{n+1}}^* a_{l_{n+1}} a_{l_n} \cdots a_{l_i'} \cdots a_{l_1}\rangle_t v(l_i' l_{n+1}'; l_i k_{n+1})$$

$$+ (i/\hbar) \sum_{i=1}^{n} \sum_{l_{n+1}} \sum_{k_i'} \sum_{k_{n+1}'} v(k_i l_{n+1}; k_i' k_{n+1}') \langle a_{k_1}^* \cdots a_{k_i'}^* \cdots a_{k_n}^* a_{k_{n+1}}^* a_{l_{n+1}} a_{l_n} \cdots a_{l_1}\rangle_t . \quad (8)$$

The first sum on the right-hand side represents free motions of particles. The second and third sums represent the effect of the external field. The fourth and fifth sums represent the interaction among the n particles of which motion are investigated. The last two sums represent their interaction with another particle in the system.

III. FORMAL SOLUTIONS IN TERMS OF THE INITIAL REDUCED DENSITY MATRICES

By taking the terms representing the effect of the external field and the interactions of particles as the inhomogeneous part, these equations are integrated formally as

$$\langle a_{k_1}^* a_{l_1}\rangle_t = \langle a_{k_1}^* a_{l_1}\rangle_{t_0} \exp\{-i[\epsilon(l_1) - \epsilon(k_1)](t - t_0)/\hbar\}$$

$$- \frac{i}{\hbar} \int_{t_0}^{t} dt_1 \exp\{-i[\epsilon(l_1) - \epsilon(k_1)](t - t_1)/\hbar\}$$

$$\times [\textstyle\sum_{l_1'} \langle a_{k_1}^* a_{l_1'}\rangle_{t_1} \phi(l_1'; l_1) - \sum_{k_1'} \phi(k_1; k_1') \langle a_{k_1'}^* a_{l_1}\rangle_{t_1}$$

$$+ \textstyle\sum_{k_2} \sum_{l_1'} \sum_{l_2'} \langle a_{k_1}^* a_{k_2}^* a_{l_2'} a_{l_1'}\rangle_{t_1} v(l_1' l_2'; l_1 k_2)$$

$$- \textstyle\sum_{l_2} \sum_{k_1'} \sum_{k_2'} v(k_1 l_2; k_1' k_2') \langle a_{k_1'}^* a_{k_2'}^* a_{l_2} a_{l_1}\rangle_{t_1}], \quad (9)$$

and,

$$\langle a_{k_1}^* a_{k_2}^* \cdots a_{k_n}^* a_{l_n} \cdots a_{l_1}\rangle_t$$

$$= \langle a_{k_1}^* a_{k_2}^* \cdots a_{k_n}^* a_{l_n} \cdots a_{l_1}\rangle_{t_0}$$

$$\times \exp\left\{-i \sum_{i=1}^{n} [\epsilon(l_i) - \epsilon(k_i)](t - t_0)/\hbar\right\}$$

$$- \frac{i}{\hbar} \int_{t_0}^{t} dt_1 \exp\left\{-i \sum_{i=1}^{n} [\epsilon(l_i) - \epsilon(k_i)](t - t_1)/\hbar\right\}$$

$$\times [\text{The expression which is obtained from the right-hand side of Eq. (8) by removing the first sum and replacing } t \text{ by } t_1]. \quad (10)$$

These are expressed diagrammatically as in Fig. 1.

Substituting the (diagrammatical) expression for a more-particle reduced density matrix for that of a fewer-particle one, successively, one gets the expression for the reduced density matrix at time t in terms of those at t_0. Fig. 2 gives the diagrammatical expression for $\langle a_{k_1}^* a_{l_1}\rangle_t$. In general (in terms of the diagrams where time increases as one goes up from below),

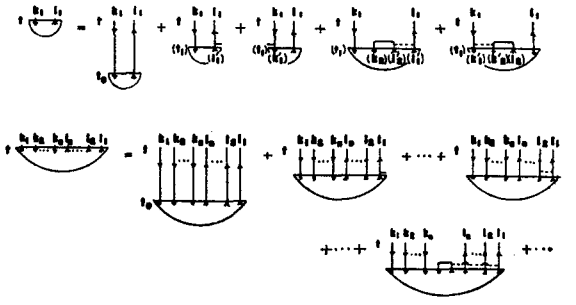


FIG. 1. Diagrammatical expressions for Eqs. (9) and (10).

$\langle a_{k_1}^* a_{k_2}^* \cdots a_{k_n}^* a_{l_n} \cdots a_{l_1}\rangle_t$ is equal to the sum of all the diagrams which are composed of a (transposed) D-type figure and a number, equal to or larger than n, of upward lines and the same number of downward lines, horizontal short bars attached to an upward or downward line, and horizontal dotted lines, connecting two upward or two downward lines. All the upward lines are drawn right to the downward lines. They are connected to the D-type figure at t_0 from above. The outermost n parts of downward and upward lines end at t, where they are labeled as k_1, k_2, $\cdots$, k_n, and l_n, $\cdots$, and l_1, respectively, from left. The other pairs of upward and downward lines must end earlier than the outer pairs and later than the inner pairs or t_0. At the end time, two lines in each of those pairs are connected with each other. Just before the end time, one line of the pair must be connected to another line by a dotted line. (11)

The rule to recover the original integral from the corresponding diagram is as follows:[7]

[7] For the system with the Hamiltonian (6′), rule 2 is rewritten as: 2′. To jth end point of a pair of upward and downward lines, a wave vector p_j is associated. To ith of short bars and dotted lines, a wave vector q_i is associated. A wave vector is associated to each interval of upward and downward lines between two successive interactions (connections to a short bar or a dotted line) in such a way that the total momentum is conserved at each vertex of lines. In doing this, it may be convenient to draw an arrow on the dotted lines; e.g., from left to right.

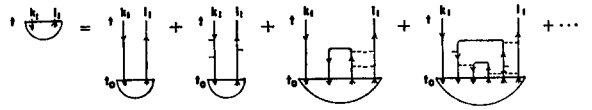


FIG. 2. Diagrammatical expression for Eq. (11) for $n = 1$.

1. To ith dotted line or a short bar, a time t_i is associated.

2. To jth interval of upward or downward lines between two connections with a short bar or a dotted line, one-particle state p_j is associated.

3a. The factor for a short bar attached to an upward or downward line, outgoing with state p and incoming with p', is

$$\mp(i/\hbar)\phi(p';p).$$

3b. The factor for a dotted line connecting two upward or downward lines, outgoing with states p_i and p_j and incoming with states p_i' and p_j', is

$$\mp(i/\hbar)v(p_i'p_j';p_ip_j).$$

The sign is negative or positive according as the line is upward or downward.

4. The factor for an upward or downward line with state p between t_1 and t_2, where $t_1 > t_2$, is

$$\exp\left[\mp i\epsilon(p)(t_1 - t_2)/\hbar\right].$$

The sign is negative or positive according as the line is upward or downward.

5a. The factor for the (transposed) D-type figure from which n pairs of upward and downward lines with states $k_1, k_2, \cdots, k_n, l_n, \cdots, l_1$ from left start at time t is

$$\langle a_{k_1}^* a_{k_2}^* \cdots a_{k_n}^* a_{l_n} \cdots a_{l_1}\rangle_t.$$

6. The summations over all states which are introduced in the rule 2 must be taken.

7. The integrals over all times which are introduced in the rule 1 must be taken as far as the topology of the diagram is not changed.

IV. SOLUTION IN TERMS OF INITIAL CUMULANT REDUCED DENSITY MATRICES

In the preceding section the reduced density matrices at time t have been solved in terms of those at time t_0. In this section, the corresponding relation for the cumulant reduced density matrices is established.

The cumulant reduced density matrices are introduced by the Ursell expansion from the reduced density matrices. The two-particle cumulant reduced density matrix is defined by

$$\langle a_{k_1}^* a_{k_2}^* a_{l_2} a_{l_1}\rangle_t = \langle a_{k_1}^* a_{l_1}\rangle_t \langle a_{k_2}^* a_{l_2}\rangle_t + \epsilon\langle a_{k_1}^* a_{l_2}\rangle_t \langle a_{k_2}^* a_{l_1}\rangle_t + \langle a_{k_1}^* a_{k_2}^* a_{l_2} a_{l_1}\rangle_t^c, \quad (12)$$

where ϵ is unity for Bose system and minus unity for Fermi system. Diagrammatically this is written as Fig. 3. Generally, the n-particle cumulant reduced density matrices are introduced, in terms of the diagram, as follows:

$\langle a_{k_1}^* a_{k_2}^* \cdots a_{k_n}^* a_{l_n} \cdots a_{l_1}\rangle_t$ = the sum of all the diagrams where n pairs of downward and upward lines, labeled as $k_1, k_2, \cdots, k_n, l_n, \cdots$, and l_1, are connected to D-type figures at time t, in such a way that each of the D-type figures is connected to the same number of downward and upward lines. The D-type figures connecting four or more lines are hatched. (13)

Now rule 5 is supplemented by rule 5b:

5b. The factor for a hatched diagram is the cumulant reduced density matrix. For a fermion system the sign of each term is positive or negative according as the number of permutations of lines, necessary to shift the D-type figures horizontally such that the upward or downward lines do not cross the D-type figures, is even or odd.

The n-particle cumulant reduced density matrix represents the correlation between n particles.

Substituting Eq. (13) into Eq. (11), one gets $\langle a_{k_1}^* a_{k_2}^* \cdots a_{k_n}^* a_{l_n} \cdots a_{l_1}\rangle_t$ = the sum of all the diagrams which are obtained from those in Eq. (11) by removing the D-type figure at t_0 and inserting one or more D-type figures in such a way that each D-type figure is connected with the same numbers of downward and upward lines. All the downward and upward lines must be connected to one of the D-type figures at t_0. The D-type figures connected to four or more lines are hatched. (14)

Taking account of Eq. (13) again, one gets

$\langle a_{k_1}^* a_{k_2}^* \cdots a_{k_n}^* a_{l_n} \cdots a_{l_1}\rangle_t^c$ = the sum of all the *connected* diagrams in the sum of Eq. (14). (15)

V. REDUCTION OF DIAGRAMS AND THE GENERALIZED BOLTZMANN EQUATION

We introduce the "s-pair of lines" for the purpose of the following reduction of the diagrams as follows:

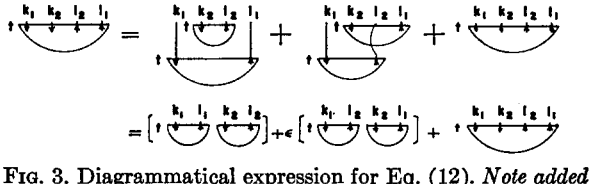


FIG. 3. Diagrammatical expression for Eq. (12). *Note added in proof.* The diagrams as drawn here should be corrected as follows: The D-type figures on the extreme right of both lines should be hatched.

An s-pair of lines is a pair of downward and upward lines such that, if one cuts those lines referred to the same time, the diagram is divided into two separate parts. In the diagrams for $\langle a_k^* a_l \rangle_t$, the final part of the pair of downward and upward lines, that is, the part of lines which take one-particle states k and l between the final time t and the time when one of the lines is connected with a short bar or a dotted line, will not be called the s-pair of lines for the convenience of the following reductions.

The diagrams in Eq. (14) or (15) are grouped by the basic diagrams which are obtained by replacing the past part—a part which is chronologically earlier—by a D-type figure at each s-pair of lines. All the diagrams in Eq. (14) or (15) are reproduced if one substitutes from one of the basic diagrams by replacing each of the D-type figures by one of the diagrams in the sum of Eq. (14) for $n = 1$ and for the corresponding time. Hence we have

$\langle a_{k_1}^* a_{k_2}^* \cdots a_{k_n}^* a_{l_n} \cdots a_{l_1} \rangle_t$ = the sum of all the diagrams which are obtained from those in Eq. (14) which have only such s-pairs of lines of which past parts are free from short bars and dotted lines, by replacing the past parts by D-type figures. (16)

$\langle a_{k_1}^* a_{k_2}^* \cdots a_{k_n}^* a_{l_n} \cdots a_{l_1} \rangle_t^c$ = the sum of all the connected diagrams in the sum of Eq. (16). (17)

A discussion has been given by Fujita[1,3] that one may neglect the terms which contain the cumulant reduced density matrices at the initial time t_0 in the right-hand side of Eq. (16) or (17), when $t - t_0$ is sufficiently large compared with the collision time t_0 characteristic to the system. If one neglects the contributions in which the cumulant reduced density matrices appear, Eq. (16) for $n = 1$ represents a closed integral equation for the one-particle reduced density matrix.[8]

If one takes a differentiation of Eq. (16) for $n = 1$ with respect to time, one gets an integro-differential equation. This is the same as

the equation which is obtained by substituting Eq. (16) for $n = 2$ into Eq. (7). (18)

The same equation is obtained by substituting Eq. (12) with Eq. (17) for $n = 2$ into Eq. (7); that is,

the equation which is obtained by substituting Eq. (17) for $n = 2$ into Eq. (20), (19)

where Eq. (20) is

$$\frac{\partial}{\partial t}\langle a_{k_1}^* a_{l_1} \rangle_t = -\frac{i}{\hbar}[\epsilon(l_1) - \epsilon(k_1)]\langle a_{k_1}^* a_{l_1} \rangle_t$$

$$- \frac{i}{\hbar}\sum_{l_1'} \phi(l_1'; l_1)\langle a_{k_1}^* a_{l_1'} \rangle_t$$

$$+ \frac{i}{\hbar}\sum_{k_1'} \phi(k_1; k_1')\langle a_{k_1'}^* a_{l_1} \rangle_t$$

$$- \frac{i}{\hbar}\sum_{k_2}\sum_{l_1'}\sum_{l_2'} v(l_1' l_2'; l_1 k_2)[\langle a_{k_1}^* a_{l_1'} \rangle_t \langle a_{k_2}^* a_{l_2'} \rangle_t$$

$$+ \epsilon\langle a_{k_1}^* a_{l_2'} \rangle_t \langle a_{k_2}^* a_{l_1'} \rangle_t + \langle a_{k_1}^* a_{k_2}^* a_{l_2'} a_{l_1'} \rangle_t^c]$$

$$+ \frac{i}{\hbar}\sum_{l_2}\sum_{k_1'}\sum_{k_2'} v(k_1 l_2; k_1' k_2')[\langle a_{k_1'}^* a_{l_1} \rangle_t \langle a_{k_2'}^* a_{l_2} \rangle_t$$

$$+ \epsilon\langle a_{k_1'}^* a_{l_2} \rangle_t \langle a_{k_2'}^* a_{l_1} \rangle_t + \langle a_{k_1'}^* a_{k_2'}^* a_{l_2} a_{l_1} \rangle_t^c]. \quad (20)$$

If one neglects the terms containing the cumulant reduced density matrices at t_0 on the right-hand side, Eq. (18) or (19) reduces to a closed integro-differential equation for the one-particle reduced density matrix. This equation is the kinetic equation for the one-particle reduced density matrix or the generalized Boltzmann equation for a quantum statistical mechanical system. This is nonlinear but also non-Markovian.

The generalized Boltzmann equation up to the first order in perturbation is obtained by neglecting the two-particle cumulant reduced density matrices at time t on the right-hand side of Eq. (20).

VI. A FURTHER REDUCTION AND THE EQUATION IN THE SECOND-ORDER PERTURBATION

There are 24 diagrams which contribute to the second-order perturbation in Eq. (18) or (19). Six of them are given in Fig. 4. It is noted that the

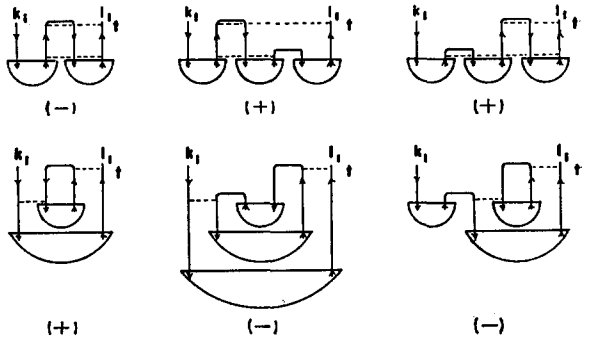


FIG. 4. Six of the diagrams which contribute to the second order in perturbation in Eq. (18) or (19). The sign (+) or (−) under the diagram shows whether the diagram is reduced from the standard form by an even or odd number of permutations of upward or downward lines.

[8] As is seen below, Eq. (16) for $n = 1$ is the integral equation for the one-particle reduced density matrix, which is equivalent to the generalized Boltzmann equation. This equation has been obtained by adopting the above definition of the s-pair lines, where the final pair of downward and upward lines in the diagrams for $\langle a_k^* a_l \rangle_t$ has been excluded. If this were not excluded, Eq. (16) for $n = 1$ would be an identity, $\langle a_k^* a_l \rangle_t = \langle a_k^* a_l \rangle_t$, and the integral equation would be obtained by integrating the generalized Boltzmann equation (19).

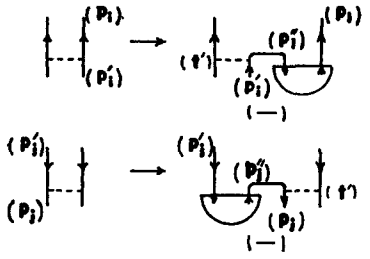


FIG. 5. Insertion of a reducible D-type figure. The labelings in parenthesis are introduced for the convenience of comparing with Eq. (21).

diagrams in the second and third column contain a D-type figure such that the pair of lines connected to it are not connected with a dotted line before one of them ends and are considered to be constructed from the far left diagram on each row by the replacement illustrated in Fig. 5. Let us call such a D-type figure "reducible." It is easy to show that, in the sum in Eq. (16) or (17), if we have a diagram which contains no reducible D-type figure and a pair of lines, both of which do not end immediately after being connected by a dotted line, then we have two diagrams which are obtained from the diagram by introducing a reducible D-type figure to one of these lines. Hence we can omit the diagrams which contain the reducible D-type figures, by adopting the following rule 3b′ instead of rule 3b:

3b′. The factor for a dotted line, at time t', connecting two upward or downward lines, outgoing with states p_i and p_j and incoming with states p_i' and p_j', is

$$-\frac{i}{\hbar}\sum_{\nu_i''}\sum_{\nu_j''} v(p_i'p_j';\, p_i''p_j'')[\delta_{\nu_i''\nu_i}\,\delta_{\nu_j''\nu_j} + \epsilon\langle a^*_{\nu_i''}a_{\nu_i}\rangle_{t'}\,\delta_{\nu_j''\nu_j} + \epsilon\langle a^*_{\nu_j''}a_{\nu_j}\rangle_{t'}\,\delta_{\nu_i''\nu_i}] \tag{21}$$

or

$$\frac{i}{\hbar}\sum_{\nu_i''}\sum_{\nu_j''} v(p_i''p_j'';\, p_ip_j)[\delta_{\nu_i''\nu_i}\,\delta_{\nu_j''\nu_j} + \epsilon\langle a^*_{\nu_i''}a_{\nu_i}\rangle_{t'}\,\delta_{\nu_j''\nu_j} + \epsilon\langle a^*_{\nu_j''}a_{\nu_j}\rangle_{t'}\,\delta_{\nu_i''\nu_i}].$$

when both of lines do not end immediately after the dotted line; otherwise rule 3b is used.

Then we have

$\langle a^*_{k_1k_2}a \cdots a^*_{k_n}a_{l_n} \cdots a_{l_1}\rangle_t^c$ = the sum of all the connected diagrams which are obtained from those in Eq. (14) which have only s-pairs of lines such that the past parts are free from short bars and dotted lines and at least one of two lines is connected to another line by dotted lines before one of them ends, by replacing the past parts by the D-type figures. The factor for the dotted line is given by rule 3b′. (22)

The equation of motion for the one-particle reduced density matrix is now given by

the equation obtained by substituting Eq. (22) for $n = 2$ into Eq. (20). (23)

All the diagrams which contribute to the second order are given in Fig. 6. For instance, the sum of contributions from the first two diagrams to the terms which contain the cumulant reduced density matrices in Eq. (20) is given by

$$\begin{aligned}&\left(\frac{i}{\hbar}\right)^2\sum_{k_2}\sum_{l_1'}\sum_{l_2'} v(l_1'l_2';\, l_1k_2)\\ &\times\int_{t_0}^{t} dt' \exp\{-i[\epsilon(l_1')+\epsilon(l_2')-\epsilon(k_1)-\epsilon(k_2)](t-t')/\hbar\}\\ &\times\sum_{l_1'''}\sum_{l_2'''}[v(l_1'''l_2''';\, l_1'l_2') + \epsilon v(l_1'''l_2''';\, l_1'l_2')]\\ &\times\sum_{l_1''}\sum_{l_2''}[\delta_{l_1''l_1'''}\,\delta_{l_2''l_2'''}\\ &+ \epsilon\langle a^*_{l_1''}a_{l_1'''}\rangle_{t'}\,\delta_{l_2''l_2'''}\\ &+ \epsilon\langle a^*_{l_2''}a_{l_2'''}\rangle_{t'}\,\delta_{l_1''l_1'''}]\langle a^*_{k_1}a_{l_1''}\rangle_{t'}\langle a^*_{k_2}a_{l_2''}\rangle_{t'}.\end{aligned} \tag{24}$$

If we use as $v(l_1l_2;\, k_1k_2)$ the symmetrized one, the two diagrams in each of this and the following pairs contribute the same value.

For a uniform system for which the Hamiltonian is given by

$$H = \sum_{\mathbf{k}}\epsilon(\mathbf{k})a^*_{\mathbf{k}}a_{\mathbf{k}} + \tfrac{1}{2}\sum_{\mathbf{k}_1}\sum_{\mathbf{k}_2}\sum_{\mathbf{q}} v(\mathbf{q})a^*_{\mathbf{k}_1+\mathbf{q}}a^*_{\mathbf{k}_2-\mathbf{q}}a_{\mathbf{k}_2}a_{\mathbf{k}_1}, \tag{25}$$

Eq. (24) is read as

$$\begin{aligned}&\left(\frac{i}{\hbar}\right)^2\sum_{\mathbf{l}}\sum_{\mathbf{q}} v(\mathbf{q})\int_{t_0}^{t} dt'\\ &\times\exp\{-i[\epsilon(\mathbf{k}+\mathbf{q})+\epsilon(\mathbf{l}-\mathbf{q})-\epsilon(\mathbf{k})-\epsilon(\mathbf{l})](t-t')/\hbar\}\\ &\times[v(\mathbf{q}) + \epsilon v(\mathbf{k}+\mathbf{q}-\mathbf{l})]\\ &\times[1+\epsilon f(\mathbf{k}+\mathbf{q},\, t') + \epsilon f(\mathbf{l}-\mathbf{q},\, t')]f(\mathbf{k},\, t')f(\mathbf{l},\, t'),\end{aligned} \tag{26}$$

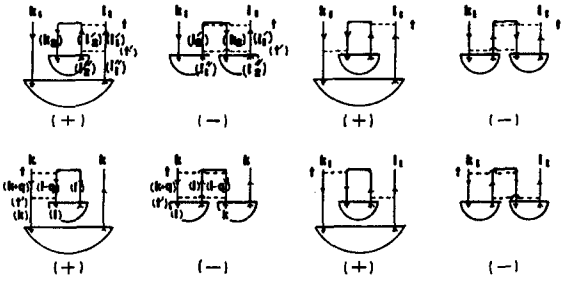


FIG. 6. All the diagrams which contribute to the second order in perturbation in Eq. (23). The meaning of the signs under the diagrams is the same as in Fig. 4. The labelings in the first pair of the diagrams are used to get the corresponding explicit expression given by Eq. (24). The labelings in the third pair of the diagrams are used to get the corresponding explicit expression for a uniform system of Eq. (28).

where

$$f(\mathbf{k}, t) \equiv \langle a_{\mathbf{k}}^* a_{\mathbf{k}} \rangle_t . \tag{27}$$

Assuming that $f(\mathbf{k}, t)$ is slowly varying and replacing $f(\mathbf{k}', t')$ on the right-hand side of Eq. (26) and the corresponding expressions for the other diagrams by $f(\mathbf{k}', t)$—or more precisely making use of Eq. (30), in Sec. VII, which makes the equation Markovian, for the $f(\mathbf{k}', t')$ and neglecting the higher-order terms in the perturbation—Eq. (20), up to the second order in perturbation, is written as follows:

$$\frac{\partial}{\partial t} f(\mathbf{k}, t) = \frac{2\pi}{\hbar} \sum_{\mathbf{l}} \sum_{\mathbf{q}} v(\mathbf{q})[v(\mathbf{q}) + \epsilon v(\mathbf{k} + \mathbf{q} - \mathbf{l})]$$
$$\times \delta[\epsilon(\mathbf{k} + \mathbf{q}) + \epsilon(\mathbf{l} - \mathbf{q}) - \epsilon(\mathbf{k}) - \epsilon(\mathbf{l})]$$
$$\times \{[1 + \epsilon f(\mathbf{k}, t) + \epsilon f(\mathbf{l}, t)] f(\mathbf{k} + \mathbf{q}, t) f(\mathbf{l} - \mathbf{q}, t)$$
$$- [1 + \epsilon f(\mathbf{k} + \mathbf{q}, t) + \epsilon f(\mathbf{l} - \mathbf{q}, t)] f(\mathbf{k}, t) f(\mathbf{l}, t)\}, \tag{28}$$

which is identical to the well-known Uhling–Uhlenbeck equation for the uniform dilute quantum gas. It is easily seen that the first-order terms cancel each other.

VII. CONCLUDING REMARKS

The generalized Boltzmann equation, Eq. (23) [(19) or (18)], which we have obtained in this paper is non-Markovian. The method of converting a non-Markovian equation to a Markovian one has already been prepared by the present author[9] when he discussed the perturbation formula for degenerate problems of many fermion systems. The same method is applicable to our equation. The result is not given here because we need modified diagrams, but the basic idea is sketched as follows. Our integral equation (16) for $n = 1$ may be written as

$$\langle a_{k_1}^* a_{l_1} \rangle_t = \langle a_{k_1}^* a_{l_1} \rangle_{t_0} \exp\{-i[\epsilon(l_1) - \epsilon(k_1)](t - t_0)/\hbar\}$$
$$- \frac{i}{\hbar} \int_{t_0}^{t} dt' \exp\{-i[\epsilon(l_1) - \epsilon(k_1)](t - t')/\hbar\}$$
$$\times I\{t'; \langle a^* a \rangle_{t''},\ t'' < t'; \langle a^* \cdots a^* a \cdots a \rangle_{t_0}^c\}$$
$$= \langle a_{k_1}^* a_{l_1} \rangle_{t_1} \exp\{-i[\epsilon(l_1) - \epsilon(k_1)](t - t_1)/\hbar\}$$
$$- \frac{i}{\hbar} \int_{t_1}^{t} dt' \exp\{-i[\epsilon(l_1) - \epsilon(k_1)](t - t')/\hbar\}$$
$$\times I\{t'; \langle a^* a \rangle_{t''},\ t'' < t'; \langle a^* \cdots a^* a \cdots a \rangle_{t_0}^c\}. \tag{29}$$

Solving for $\langle a_{k_1}^* a_{l_1} \rangle_{t_1}$, one gets

$$\langle a_{k_1}^* a_{l_1} \rangle_{t_1} = \langle a_{k_1}^* a_{l_1} \rangle_t \exp\{i[\epsilon(l_1) - \epsilon(k_1)](t - t_1)/\hbar\}$$
$$+ \frac{i}{\hbar} \int_{t_1}^{t} dt' \exp\{i[\epsilon(l_1) - \epsilon(k_1)](t' - t_1)/\hbar\}$$
$$\times I\{t'; \langle a^* a \rangle_{t''},\ t'' < t'; \langle a^* \cdots a^* a \cdots a \rangle_{t_0}^c\}. \tag{30}$$

By an iterative procedure, this equation is solved for $\langle a_{k_1}^* a_{l_1} \rangle_{t_1}$ for $t_1 < t$ in terms of $\langle a_k^* a_l \rangle_t$. This result is used to make our non-Markovian equation a Markovian one.

Our resultant generalized Boltzmann equation, given by Eq. (23), and its integral equation version, given by Eq. (16) for $n = 1$, are given in terms of diagrams. We have drawn the diagrams for the generalized Boltzmann equation up to the second-order terms. In order to get the higher-order terms we have to draw more and more complicated diagrams. However for various many-body problems the exclusive consideration of diagrams with some simple topology has been shown to be a practical way to get physically meaningful results and hence the use of the diagram should not be considered unpractical.

When we are interested in the linear irreversible processes, the transport coefficients are of the primary interest and they are calculated by solving the kinetic equation or the generalized Boltzmann equation, or more directly by calculating the suitable two-time correlation functions or two-time Green's functions.[10] The equations for them have the same or a similar structure to our equations for the reduced density matrices and our method of solving the hierarchy and then reducing it to an integral or integrodifferential equation is expected to be useful for this problem.

ACKNOWLEDGMENTS

The author is indebted to Professor S. Fujita and Professor T. Tanaka for their stimulating discussions.

[9] T. Morita, Progr. Theoret. Phys. (Kyoto) 29, 351 (1963).

[10] R. Kubo, J. Phys. Soc. Japan 12, 570 (1957); V. L. Bonch-Bruevich and S. V. Tyablikov, *The Green Function Method in Statistical Mechanics* (North-Holland Publishing Company, Amsterdam, 1962).

Theorem on the Clebsch–Gordan Series in $SU(n)$*

S. NUSSINOV
University of Washington, Seattle, Washington
(Received 23 July 1965)

A theorem on the Clebsch–Gordan series conjectured by B. Vitale is proved using simple Young diagram techniques.

I. INTRODUCTION

THE number and type of irreducible representations (I.R.'s) (P'')[1] appearing in a reduction of a direct product of two $SU(n)$ I.R.'s (P) and (P')[2],

$$(P) \otimes (P') = \oplus \sum K(P, P', P'')(P''), \qquad (1)$$

is, in general, a complicated function of the I.R.'s (P) and (P'). On the basis of an extensive analysis of the Clebsch–Gordan (C.G.) series in $SU(3)$, Vitale[3] conjectured that, in a special case [Eq. (3) below], the representation (P') determines completely the structure of the (C.G.) series. This conjecture is proved in the following, using the connection between representations of $SU(n)$ and the Young diagrams. Since very little is known at present about the *explicit* structure of the C.G. series of Lie groups, this theorem is quite interesting and may find useful applications.

II. THEOREM

Definition: The I.R. (P') is called "representation determining with respect to (P)" if, for any basis, vector in the I.R. (P') with weight $\mathbf{m}'$, there corresponds a representation (P'') in Eq. (1) such that[4]

$$\Lambda'' = \Lambda + \mathbf{m}', \qquad (2)$$

where Λ'' and Λ are the highest weights of the I.R.'s (P'') and (P), respectively.

Theorem: A necessary and sufficient condition for (P') to be representation determining with respect to (P) is

$$P_i \geq \sum_{k}^{n} P'_k = f'_1, \qquad i = 1 \cdots n-1. \qquad (3)$$

III. SOME PRELIMINARY DISCUSSION

In this section, we recall some facts about the connection between I.R.'s of $SU(n)$ and Young diagrams and tableaux.[5]

(1) To each "allowed" Young diagram (f'), there corresponds a unique I.R. (P').

(2) The basis vectors of the I.R. (P') can be associated with the various "allowed" Young tableaux $[f']$ generated from the Young diagram (f') as follows: "tensor indices" $1, \cdots, n$ are distributed in the boxes of (f'). This filling yields an "allowed" table if the tensor indices increase as we move down in any column, and do not decrease when we move to the right in any line.

The table can be specified by the numbers

$$n'^{i}_{i}, \qquad i \leq n, \qquad j \leq n-1,$$

where n'^{i}_{i} = number of times the tensor index i appears in the jth row of f'. Thus,

$$N'_i = \sum_{j=0}^{n-1} n'^{i}_{i} \qquad (4)$$

is the total number of indices i in the table $[f']$. Because no two identical indices appear in the same column:

$$N'_i \leq \text{total number of columns in } (f')$$
$$= \sum P'_i = f'_1. \qquad (4')$$

The weight vector $\mathbf{m}' = (m'_1 \cdots m'_i \cdots m'_{n-1})$ of the basis vector considered is related to

$$N' = (N'_1 \cdots N'_i \cdots N'_n) \text{ by } \mathbf{m}' = A\mathbf{N}' + \mathbf{C}', \qquad (5)$$

---

* This work was supported in part by the U. S. Atomic Energy Commission under contract A. T. (45-1)-1388, Program B.

[1] (P) is used as a shorthand for $(P_1 \ldots P_{n-1})$ and labels an Irreducible Representation (I. R.) of $SU(n)$. The corresponding Young diagram with rows

$$f_j = \textstyle\sum_{k=1}^{k=n-j} P_k \text{ is denoted by } (f).$$

For allowed Young diagrams, $f_j \geq f_{j+1}$ $j = 1 \ldots n$. $[f]$ will denote any of the Young tableaux corresponding to (f).

[2] Special cases of the series (1) were considered by H. Goldberg, Nuovo Cimento **27**, 532 (1963); A. Simoni and B. Vitale, *ibid.* **33**, 1199 (1964); B. Preziosi, A. Simoni, and B. Vitale, *ibid.* **34**, 1101 (1965).

[3] B. Vitale, "On the Structure of the Clebsch–Gordan Series for Semi-Simple Lie Groups," preprint from the University of Wisconsin. Vitale has given the conditions relative to A_2, C_2, G_2, and A_3 by using the "geometrical" weight diagram approach expounded by J. P. Antoine and D. Speiser, J. Math. Phys. **5**, 1226 and 1560 (1964), and besides has conjectured the validity of our condition (3) for A_l.

[4] In particular, if the weight m' is r-fold degenerate in the I.R. (P'), then the I.R. (P'') will occur r times in Eq. (1).

[5] See, for example, H. Weyl, *The Classical Groups* (Princeton University Press, Princeton, New Jersey, 1946); D. E. Littlewood, *The Theory of Group Characters* (Oxford University Press, Oxford, England, 1950).

where A is a constant matrix and C' a trivial "shift" vector which we omit in the following.[6]

In particular, to the table in which N_i is the length of the ith row in[7] $(f') = f'_i$ $(i = 1 \cdots n-1)$, corresponds the maximum weight of (P'):

$$\Lambda' = Af'. \tag{6}$$

(3) Using Young diagrams, the I.R.'s (P'') appearing in Eq. (1) may be obtained by the following process:

Process A: (1) f'_i labels a_i are put into the jth row of (f'), $j = 1 \cdots n-1$.

The labeled squares of (f') are then added, row after row, to the diagram (f). Among all such processes, the "allowed A processes" are singled out according to the following requirements:

(i) At each stage the figure consisting of the original diagram (f) and the squares added from (f') is a valid Young diagram.

(ii) No two identical labels appear in the same column.

(iii) In (f''), the diagram finally obtained, the added letters are "read" from right to left and the number of a_i's accumulated at an arbitrary stage and is denoted by M_i. We must always have $M_{i+1} \leq M_i$. To each of the diagrams (f'') obtained in an allowed process corresponds an I.R. (P'') in Eq. (1).

IV. THE PROOF OF THE THEOREM

To prove the theorem, we first show that if the condition of Eq. (3) is satisfied, then a one-to-one correspondence can be established between allowed tableaux $[f']$ and diagrams (f'') constructed in allowed A process.

The correspondence is achieved as follows:

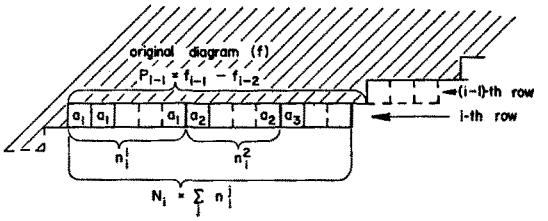


FIG. 1. The $(i-1)$th and ith rows of (f) after adjoining the labels a_i according to (a).

(a) Given any table $[f']$ with n_i^j tensor indices i in the jth row, we construct a diagram (f'') in the following A process: we add n_i^1 labels a_1 to the ith row of (f), $i = 1 \cdots n$, then n_i^2 labels a_2 to the ith row, and so on.

(b) Given any A process we construct a Young table $[f']$ in which n_i^j = number of a_j labels added to the ith row of (f).

We now show that:

(1) If $[f']$ is an allowed table, then the A process generated from it according to (a) satisfies conditions (i)–(iii) specified in the preceding section.

(2) If a process A satisfies the conditions (i)–(iii), then the corresponding table constructed according to (b) is an allowed table.

1. It follows from Eqs. (4'), (3), that the total number of labels added, in an A process corresponding to an allowed table, to the ith row of (f) is

$$N'_i \leq f'_1 = \sum^{n-1} P'_i \leq P_i, \qquad i = 1 \cdots n.$$

Thus, the added squares do not protrude beyond the original $(i-1)$th line of (f) (see Fig. 1) and conditions (i), (ii) are automatically satisfied.

The total number of labels a_j added to the first i rows of (f) is according to the construction (a) $\sum_{k=1}^{i} n_k^j$. We have

$$\sum_{k=1}^{i} n_k^j \leq \sum_{k=1}^{i-1} n_k^{j-1}. \tag{7}$$

Equation (7) can be verified by considering the $(j-1)$th and the jth rows in the table, $[f']$ (see Fig. 2). However, $\sum_{k=1}^{i-1} n_k^{j-1}$ is equal to the total number of a_{j-1} labels added to the rows $1 \cdots i-1$ of (f). Hence, (iii) is also satisfied.

2. Since construction (b) is the inverse of construction (a), it suffices to show that if we start from any unallowed table $[f']$, then the resulting

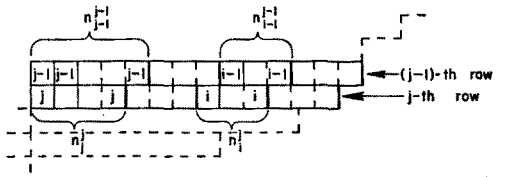


FIG. 2. The $(j-1)$th and jth rows of an allowed table $[f']$.

[6] To obtain A note that any vector $v \in (P)$ can be expressed by linear combinations of products $\Pi\, V_i^{N_i^{(1)}} \bar{V}_i^{N_i^{(2)}}$ where $V_i(\bar{V}_i)$, $i = 1 \ldots n$, are the n vectors in a fundamental ("quark") representation of $SU(n)$ and its conjugate ("antiquark") representation, respectively. The weight associated with such a product is $\mathbf{m} = \sum N_i \mathbf{m}_i$, where $N_i = N_i^{(1)} - N_i^{(2)}$ and $\mathbf{m}_i$ is the weight of V_i. The weight vectors $\mathbf{m}_i$ satisfy $\sum \mathbf{m}_i = 0$ so that if $N_i \to N_i + C$, m stays the same. N_i is the same as the N_i associated with a certain table f' [Eq. (4) above]. (The common practice of omitting the first C columns of length n in any Young table corresponds to $N_i \to N_i + C$, because a column of length n must contain all n indices at exactly one time.) The equation

$$\mathbf{m} = \sum N_i \mathbf{m}_i$$

is clearly equivalent to $\mathbf{m} = A\mathbf{N}$, $A = n-1 \times n$ matrix with elements $a_i{}^j = m_i{}^j = j$th component of $\mathbf{m}_i$. As $\mathbf{N}'$ runs over all tableaux $[f']$ the pattern of weights obtained will coincide in structure with the weight diagram of (P'), which is usually constructed by purely Lie-algebraic methods. This pattern will in general be "shifted" from the ordinary "centered" position of the weight diagram of (P'); hence the vector $\mathbf{C}$. This "centering" operation will always be assumed in the following and $\mathbf{C}$ will be omitted.

[7] This table is obtained by putting f'_j tensor indices in the jth row $j = 1 \ldots n-1$.

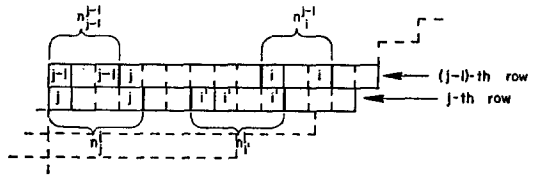


FIG. 3. Some pair of consecutive rows of an unallowed table $[f']$.

A process violates one of the conditions (i)–(iii). Since $[f']$ is unallowed it contains a pair of consecutive rows filled with tensor indices in the manner shown in Fig. 3, so that at least one tensor index, say i, appears above an index i' with $i' \leq i$. Thus, $\sum_{k=1}^{i-1} n_k^{i-1} \leq \sum_{k=1}^{i'} n_k^i$ with $i' \leq i$. On using (a), we have: total number of labels a_{i-1} added to the first $i-1$ rows of (f) $\leq$ total number of labels a_i added to the first i' rows of (f). Since $i' \leq i$, and since the labels a_{i-1} appear in any row of (f'') to the left of the levels a_i in the same row, it follows that condition (iii) will be violated at a certain stage while "reading" the labels of the i'th row, of (f').

This completes the proof that the correspondence described in (a) and (b) is one-to-one.

The length of the ith row in (f'') is

$$f_i'' = f_i + \sum_{i=0}^{n-1} n_i^i = f_i + N_i,$$

thus, using Eqs. (5) and (6), and remembering the remark (5), we obtain the required relation (2).

To complete the proof of the theorem, we have to show that condition (3) is necessary, and if it is violated the correspondence between basis vectors of (P') and I.R.'s (P'') in the C.G. series (1) is no longer one-to-one.

If Eq. (3) is not true, then at least for one number i_0, $1 \leq i_0 \leq n$:

$$P_{i_0} < \sum^{n} P_k' = f_1'. \tag{8}$$

We can in this case construct at least one allowable table $[f']$ for which (a) does not lead to a valid Young diagram (f''). Such a table is illustrated in Fig. 4. Since (f') has only $(n-1)$ rows, we can omit the tensor index $(i_0 - 1)$ in filling it. We put tensor indices (i_0) in all the squares of the $(i_0 - 1)$th row and in all the remaining allowable places in the rows above it, so that we have in all f_1' indices i_0.

If we apply now procedure (a) to this particular table, then f_1' squares will be added to the i_0th row of (f) and none to the $(i_0 - 1)$th row. Since the difference of length between these rows was originally P_{i_0}, Eq. (8) implies that in the final diagram (f'') the i_0th line will be longer than the $(i_0 - 1)$th line and (f'') will not be a valid Young diagram.

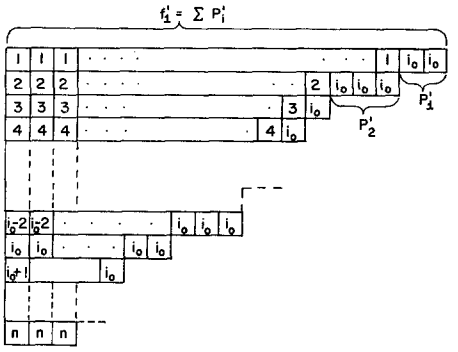


FIG. 4. A table $[f']$ which contains no tensor indices $(i_0 - 1)$ and f_1' tensor indices i_0.

On the other hand, the arguments in Sec. (2) above did not depend at all on Eq. (3) and the transition (b) always leads from an allowed A process to an allowed table $[f']$. Thus, when Eq. (3) is not valid, the set of I.R.'s (P'') in Eq. (1) will correspond to a subset of all tableaux $[f']$ [or basis vectors of (P')], a subset from which tableaux of the type shown in Fig. 4, and maybe other tableaux, have been omitted. The precise definition of this set, in the general case, is, according to our discussion, equivalent to determining the complete C.G. series [Eq. (1)].

It is quite conceivable that the correspondence suggested above may still yield useful information. For example, it was found in Ref. 2c that even when condition (3) was abandoned, the C.G. series (1) could still be represented by a set of points in the plane with a degeneracy structure analogous to that of an $SU(3)$ weight diagram. We think that the methods of this work can be extended in order to prove that this holds generally for $SU(n)$, although the proof may not be as straight forward as the one above.

Note added in proof. After submitting this paper we received a preprint by F. Zaccaria[8] in which the above theorem was proved by purely "Lie Algebraic" methods. Zaccaria's method, which is different from the one used here, allowed also for proving conditions analogous to Eq. (3) for all familiar semisimple Lie groups. I would like to thank Dr. Zaccaria for sending me his preprint.

ACKNOWLEDGMENTS

I would like to thank Professor B. Vitale for helpful correspondence and Professors M. Baker and I. J. Muzinich for encouragement.

[8] F. Zaccaria, Univ. of Naples preprint submitted to Journ. of Math. Phys.

On Alternate Commutation Relations*

W. F. PARKS
Department of Physics, Kansas State University, Manhattan, Kansas
(Received 10 August 1965)

An alternate quantization procedure for Bose fields is proposed. This procedure leads to a field algebra which is related to that of angular momentum rather than momentum. With this algebra is associated a maximum (n_0) to the number of particles allowed in a given momentum state. It is indicated by the investigation of models that physical observables converge for large n_0 to those obtained using the canonical commutation relations.

A FIELD quantization procedure is presented in this article, which leads to states that are symmetric under the permutations of identical particles. This procedure differs fundamentally from that used for Bose fields in that it has associated with it a maximum to the number of particles allowed in a given momentum state. Furthermore it is not among the parastatic theories recently investigated.[1]

In the first section, the commutation relations are assumed and an important property of the algebra satisfying these relations is obtained. In subsequent sections, models are investigated indicating the effect of changing the field quantization procedure. In particular, one model will demonstrate the connection between the static scalar field model[2] and the completely reduced Bardeen–Cooper–Schreiffer (BCS) model for superconductivity.[3]

1. THE COMMUTATION RELATIONS

It is assumed, first, that the system of interest is confined to a finite volume Ω thereby establishing a discrete single-particle momentum spectrum. The set of operators $\{a_\mathbf{k}, a_\mathbf{k}^\dagger, N_\mathbf{k}\}$ corresponding, respectively, to the annihilation, creation, and number operator for the free particle state of momentum $\mathbf{k}$ satisfy by their interpretation the relationships:

$$[a_\mathbf{k}, N_{\mathbf{k}'}] = \delta_{\mathbf{k}\mathbf{k}'} a_\mathbf{k}, \tag{1}$$

$$[N_\mathbf{k}, N_{\mathbf{k}'}] = 0, \qquad N_\mathbf{k}^\dagger = N_\mathbf{k}. \tag{2}$$

It is also expected with this interpretation, that there exists a vacuum state $|0\rangle$ such that

$$a_\mathbf{k} |0\rangle = N_\mathbf{k} |0\rangle = 0, \qquad \langle 0 \mid 0\rangle = 1, \tag{3}$$

from which the representation space of the operators can be generated by the application of the a's and $a^\dagger$'s. In order that the states generated from the vacuum be symmetric under the permutations of identical particles it is required that

$$[a_\mathbf{k}, a_{\mathbf{k}'}] = 0. \tag{4}$$

The form of the relationships between a and $a^\dagger$ appears to be free; therefore the following is assumed:

$$[a_\mathbf{k}^\dagger, a_{\mathbf{k}'}] = \delta_{\mathbf{k}\mathbf{k}'} C_\mathbf{k}, \tag{5}$$

where $C_\mathbf{k}$ is not necessarily a c-number. Equation (5) is inconsistent with the corresponding equation in the parastatic theories[1] unless $C_\mathbf{k} = -1$ [$C_\mathbf{k} = c$-number is sufficient] in which case Eqs. (1), (4), and (5) define a Bose system. In order to limit the choice of $C_\mathbf{k}$ consider what conditions it must satisfy. The first arises from the requirement that Eqs. (1) and (5) be consistent, i.e.,

$$[C_\mathbf{k}, N_{\mathbf{k}'}] = 0. \tag{6}$$

The second follows from the requirement that the single-particle states $a_\mathbf{k}^\dagger |0\rangle$ satisfy the orthonormality condition

$$\langle 0| a_\mathbf{k} a_{\mathbf{k}'}^\dagger |0\rangle = \delta_{\mathbf{k}\mathbf{k}'}, \tag{7}$$

i.e.,

$$\langle 0| C_\mathbf{k} |0\rangle = -1. \tag{8}$$

It is not clear what conditions are implied by the existence of the vacuum state, however the above relations are not inconsistent with its existence since they are satisfied by the Bose commutation relations. One choice of $C_\mathbf{k}$ satisfying Eqs. (6) and (8) is

$$C_\mathbf{k} = 2n_0^{-1}(N_\mathbf{k} - \tfrac{1}{2}n_0) \tag{9}$$

with n_0 a positive c-number. The existence of the vacuum along with the significance of n_0 is obtained

* This work was supported in part by the Air Force Office of Scientific Research.

[1] See, e.g., S. Kamefuchi and Y. Takahashi, Nucl. Phys. **36**, 177 (1962); S. Kamefuchi and J. Strathdee, Nucl. Phys. **42**, 166 (1963); M. Dresden, *Brandeis Summer Lectures in Theoretical Physics* (W. A. Benjamin Company, Inc., New York, 1963), Vol. 2, p. 377; A. M. L. Messiah and O. W. Greenberg, Phys. Rev. **136B**, 248 (1964).

[2] S. Schweber, *Relativistic Quantum Field Theory* (Row, Peterson, and Co., Evanston, Illinois, 1961), Chap. 12, p. 343.

[3] J. Bardeen, L. N. Cooper, and J. R. Schrieffer, Phys. Rev. **108**, 1175 (1957).

in the following analysis of the set of operators defined by Eqs. (1), (2), (4), (5), and (9).

Comparing the above commutation relations with those of angular momentum and noting that a discrete spectrum is isomorphic to the positive integers, one finds that

$$J_i^+ \leftrightarrow n_0^{\frac{1}{2}} a_{\mathbf{k}}^\dagger,\ J_i^- \leftrightarrow n_0^{\frac{1}{2}} a_{\mathbf{k}},\ \text{and}\ (J_z)_i \leftrightarrow N_{\mathbf{k}} - \tfrac{1}{2} n_0 \quad (10)$$

is an isomorphism if $N_{\mathbf{k}}$ is restricted to the values $0, 1, 2, \cdots, n_0$. Therefore the vacuum state exists if n_0 is identified as the maximum number of particles allowed in a given momentum state.

It is seen then that Eq. (9) leads to operators which are bounded in contrast to the unbounded operators obtained from the canonical commutation relations. This property is obtained, however, at the price of singling out the momentum-space representation as the only one in which the commutation relations assume a diagonal form. This is not necessarily objectionable since this representation has special significance due to the dynamics of free-particle theories.

Clearly Eqs. (6) and (8) allow choices for $C_{\mathbf{k}}$ other than that of Eq. (9). In fact, if $C_{\mathbf{k}}$ is chosen to be a polynomial (or series) in the number operator such that the leading term is -1, Eqs. (6) and (8) are satisfied. Such possibilities are ignored and for the purposes of this paper the relationship

$$[a_{\mathbf{k}}^\dagger, a_{\mathbf{k}'}] = 2n_0^{-1}\, \delta_{\mathbf{k}\mathbf{k}'}(N_{\mathbf{k}} - \tfrac{1}{2} n_0) \quad (11)$$

is assumed.

II. PARTICLE-CONSERVING THEORIES

If the usual forms are assumed for the nonrelativistic, particle-conserving Hamiltonians, the question arises as to what effect the change in quantization procedure has on the Schrödinger theory. It is easily shown that, for a single-particle system or a system of nonidentical particles (the fields of different particles commuting) the first quantized formulism is unchanged in the case of the assumed quantization. For a system of n identical particles a difference exists. This difference is indicated and examined in this section for some specific cases.

Consider first the case of $n_0 = 1$, i.e.,

$$[a_{\mathbf{k}}^\dagger, a_{\mathbf{k}'}] = 2\, \delta_{\mathbf{k}\mathbf{k}'}(N_{\mathbf{k}} - \tfrac{1}{2}). \quad (12)$$

The standard form for a particle-conserving, nonrelativistic, second quantized Hamiltonian is

$$H = \sum_{\mathbf{k}} \epsilon(\mathbf{k}) N_{\mathbf{k}} + (2\Omega)^{-1} \sum_{\mathbf{k}_1 \mathbf{k}_2 \mathbf{p}_1 \mathbf{p}_2} a_{\mathbf{k}_1}^\dagger a_{\mathbf{k}_2}^\dagger (\mathbf{k}_1\mathbf{k}_2|\ V\ |\mathbf{p}_2\mathbf{p}_1) a_{\mathbf{p}_2} a_{\mathbf{p}_1}, \quad (13)$$

where $\epsilon(\mathbf{k}) = \mathbf{k}^2/2M$ and $(\mathbf{k}_1\mathbf{k}_2|\ V_2\ |\mathbf{p}_2\mathbf{p}_1)$ contains a Kronecker-δ conservation of momentum. Let

$$|E, n\rangle = (n_!)^{-\frac{1}{2}} \sum_{\mathbf{k}_1, \cdots, \mathbf{k}_n} f(\mathbf{k}_1, \cdots, \mathbf{k}_n) a_{\mathbf{k}_1}^\dagger \cdots a_{\mathbf{k}_n}^\dagger |0\rangle, \quad (14)$$

with $f(\mathbf{k}_1, \cdots, \mathbf{k}_n)$ symmetric under the interchange of variables, satisfy the relation

$$H\,|E, n\rangle = E\,|E, n\rangle. \quad (15)$$

Then f must satisfy the Schrödinger equation

$$\begin{aligned}
&\Big(\textstyle\sum_{i=1}^n \epsilon(\mathbf{p}_i) - E\Big) \\
&\quad \times f(\mathbf{p}_1, \cdots, \mathbf{p}_n) \prod_{s>t}^n s, t(1 - \delta_{\mathbf{p}_s\mathbf{p}_t}) \\
&\quad + \prod_{s>t}^n s, t(1 - \delta_{\mathbf{p}_s\mathbf{p}_t}) \{(2\Omega)^{-1} \textstyle\sum_{\mathbf{k}\mathbf{k}'} (\mathbf{p}_1\mathbf{p}_2|\ V\ |\mathbf{k}'\mathbf{k}) \\
&\quad \times f(\mathbf{k}, \mathbf{k}', \mathbf{p}_3, \cdots, \mathbf{p}_n)(1 - \delta_{\mathbf{k}\mathbf{k}'}) \\
&\quad \times \prod_{i\neq 1,2}^n (1 - \delta_{\mathbf{k}\mathbf{p}_i})(1 - \delta_{\mathbf{k}'\mathbf{p}_i}) \\
&\quad + \text{permutations } (\mathbf{p}_i)\} = 0.
\end{aligned} \quad (16)$$

Or, in a more transparent form,

$$\begin{aligned}
&\Big(\textstyle\sum_{i=1}^n \epsilon(\mathbf{p}_i) - E\Big) f(\mathbf{p}_1, \cdots, \mathbf{p}_n) \\
&\quad + \{(2\Omega)^{-1} \textstyle\sum_{\mathbf{k}\mathbf{k}'} (\mathbf{p}_1\mathbf{p}_2|\ V\ |\mathbf{k}', \mathbf{k}) \\
&\quad \times f(\mathbf{k}, \mathbf{k}', \mathbf{p}_3, \cdots, \mathbf{p}_n) + \cdots\} \\
&\quad + \{\Omega^{-1} \textstyle\sum_{\mathbf{k}} \sum_{i\neq 1,2}^n (\mathbf{p}_1\mathbf{p}_2|\ V\ |\mathbf{p}_i\mathbf{k}) \\
&\quad \times f(\mathbf{k}, \mathbf{p}_i, \mathbf{p}_3, \cdots, \mathbf{p}_n) + \cdots\} \\
&\quad + \{(2\Omega)^{-1} \textstyle\sum_{i,j\neq 1,2} (\mathbf{p}_1\mathbf{p}_2|\ V\ |\mathbf{p}_i\mathbf{p}_j) \\
&\quad \times f(\mathbf{p}_j, \mathbf{p}_i, \mathbf{p}_3, \cdots, \mathbf{p}_n) + \cdots\} \\
&\quad + \{(2\Omega)^{-1} \textstyle\sum_{i\neq 1,2} \mathbf{p}_1\mathbf{p}_2|\ V\ |\mathbf{p}_i\mathbf{p}_i) \\
&\quad \times f(\mathbf{p}_i, \mathbf{p}_i, \mathbf{p}_3, \cdots, \mathbf{p}_n) + \cdots\} \\
&\quad + \{(2\Omega)^{-1} \textstyle\sum_{\mathbf{k}} (\mathbf{p}_1\mathbf{p}_2|\ V\ |\mathbf{k}\mathbf{k}) \\
&\quad \times f(\mathbf{k}, \mathbf{k}, \mathbf{p}_3, \cdots, \mathbf{p}_n) + \cdots\} = 0,
\end{aligned} \quad (17)$$

where the dots stand for terms involving permutations of $(\mathbf{p}_i)$ in the first term of the bracket and $\mathbf{p}_s \neq \mathbf{p}_t$ if $s \neq t$. The terms in the last four brackets are "corrections" to the theory. For a fixed n, i.e., not in the thermodynamic limit, these terms go as Ω^{-1} for large volume compared to the terms in the first bracket which are independent of Ω. It is clear, then, that if the number of particles is fixed and finite one obtains a close approximation to the Schrödinger theory for a system contained in a large volume. It is also clear, however, that in the thermodynamic limit the terms in the second and third brackets will not be negligible since the sums over the momenta of the particles will yield terms of the

same order in Ω as those terms involving sums over all momenta.

The fact that the Schrödinger theory is altered in the thermodynamic limit is easily seen in the case of a system of noninteracting particles. For a given Ω and $n > n_0$ one obtains a Fermi-sphere distribution of particles in the ground state of the system. For example, let a system of $6n_0 + p$, $p < n_0$, particles be contained in a box. Then the lowest energy state would be completely filled and the next-higher energy state would contain p particles. The "exclusion principle" is introduced through the inner product obtained from the second quantized theory:

$$(g, f) = \sum_{\mathbf{k}_1,\cdots,\mathbf{k}_n} g^*(\mathbf{k}_1, \cdots, \mathbf{k}_n) P_{n_0} f(\mathbf{k}_1, \cdots, \mathbf{k}_n), \tag{18}$$

where

$$P_{n_0} f(\mathbf{k}_1, \cdots, \mathbf{k}_n) = \begin{cases} 0, & \text{if more than } n_0 \text{ momenta are the same,} \\ f(\mathbf{k}_1, \cdots, \mathbf{k}_n), & \text{otherwise.} \end{cases} \tag{19}$$

It follows, then, that the norm of a state in which more than n_0 particles have the same momentum is zero. The number of single-particle momentum states in a given neighborhood of the zero-momentum state is proportional to Ω, therefore if n is held finite as $\Omega \to \infty$ the "Fermi momentum" of the particle distribution approaches zero. On the other hand, if n/Ω is fixed as $\Omega \to \infty$ the "Fermi momentum" is unchanged. Before leaving the subject of the exclusion principle, it should be noted that particles described by the $n_0 = 1$ theory satisfy Fermi–Dirac statistics in regard to their momentum-space distribution function. This is easily obtained from the grand canonical distribution function for arbitrary n_0:

$$n_{\mathbf{k}} = \{e^{\beta[\epsilon(\mathbf{k})-\mu]} - 1\}^{-1} - (n_0 + 1)\{e^{\beta[\epsilon(\mathbf{k})-\mu][n_0+1]} - 1\}^{-1}. \tag{20}$$

This theory cannot describe fermions however since there is no exclusion principle for other representations of single-particle states.

The above discussion of systems of interacting particles can be repeated for $n_0 > 1$ with a correspondingly more complicated expression for the Schrödinger equation. One cannot easily analyze the equation as to its behavior for large n_0 except in the two-particle theory in which the equation has the form

$$[\epsilon(\mathbf{k}_1) + \epsilon(\mathbf{k}_2) - E] f(\mathbf{k}_1, \mathbf{k}_2) + (2\Omega)^{-1} \sum_{\mathbf{k}\mathbf{k}'} (\mathbf{k}_1\mathbf{k}_2| V |\mathbf{k}'\mathbf{k}) f(\mathbf{k}, \mathbf{k}') = (2n_0\Omega)^{-1} \sum_{\mathbf{k}} (\mathbf{k}_1\mathbf{k}_2| V |\mathbf{k}, \mathbf{k}) f(\mathbf{k}, \mathbf{k}). \tag{21}$$

The correction term, which appears on the right-hand side of the equation, is proportional to $(n_0\Omega)^{-1}$. Therefore, one expects the results of the proposed theory to approach those of the Bose theory for large Ω and/or large n_0.

Finally, it is important that a theory satisfy the cluster decomposition property. Considering the conditions which are imposed to obtain the decomposition of the expectation values of products of creation and annihilation operators, i.e., two groups of particles separated by a large (approaching infinite) distance, one finds that these conditions are just those for which the proposed theory goes over into the Bose theory. This follows from the obvious fact that if the clusters are separated by a large distance they must be contained in a correspondingly large volume. This theory therefore satisfies the cluster decomposition property.

III. A PARTICLE-NONCONSERVING MODEL

To continue the investigation of the proposed commutation relations, the system described by the following model Hamiltonian is considered. Let

$$H = \sum_{\mathbf{k}} \omega(\mathbf{k}) N_{\mathbf{k}} + \sum_{\mathbf{k}} A(\mathbf{k})(a_{\mathbf{k}} + a_{\mathbf{k}}^\dagger) \tag{22}$$

with $\omega(k)$ the single-particle kinetic energy. If

$$\omega(\mathbf{k}) = (\mathbf{k}^2 + m_0^2)^{\frac{1}{2}} \tag{23}$$

and

$$A(\mathbf{k}) = \lambda f(\mathbf{k}^2)[2\Omega\omega(\mathbf{k})]^{-\frac{1}{2}},$$

this model corresponds to the static scalar field model[2] of the Bose theory. Due to the isomorphism with the angular momentum algebra, the Hamiltonian, written in the form

$$H = \tfrac{1}{2}n_0 \sum_{\mathbf{k}} \omega(\mathbf{k}) + \sum_{\mathbf{k}} W(\mathbf{k})[T(\mathbf{k})(N_{\mathbf{k}} - \tfrac{1}{2}n_0) + S(\mathbf{k})\tfrac{1}{2}n_0^{\frac{1}{2}}(a_{\mathbf{k}} + a_{\mathbf{k}}^\dagger)] \tag{24}$$

with

$$W(\mathbf{k}) = [\omega(\mathbf{k})^2 + 4A(\mathbf{k})^2 n_0^{-1}]^{\frac{1}{2}}, \tag{25a}$$

$$S(\mathbf{k}) = 2A(\mathbf{k})n_0^{-1}W(\mathbf{k})^{-1}, \tag{25b}$$

and

$$T(\mathbf{k}) = \omega(\mathbf{k})W(\mathbf{k})^{-1}, \tag{25c}$$

can be diagonalized by a "rotation about the y axis." This rotation is carried out by the "unitary" operator

$$U = \exp[\tfrac{1}{2}n_0^{\frac{1}{2}} \sum_{\mathbf{k}} \theta(\mathbf{k})(a_{\mathbf{k}} - a_{\mathbf{k}}^\dagger)] \tag{26a}$$

with

$$\theta(\mathbf{k}) = \sin^{-1} S(\mathbf{k}). \tag{26b}$$

If U exists, it is unitary. The conditions for its existence are found here in the case $n_0 = 1$.

Consider the operator

$$U_m \equiv \exp\left[\tfrac{1}{2}n_0^{\frac{1}{2}} \sum_{|\mathbf{k}|<m} \theta(\mathbf{k})(a_\mathbf{k} - a_\mathbf{k}^\dagger)\right]. \quad (27)$$

This operator exists and is unitary since it is a finite product of unitary operators. Clearly U can be defined by

$$U \equiv \lim_{m\to\infty} U_m \quad (28)$$

if this limit exists. What must be shown, therefore, is that the matrix elements of $U - U_m \to 0$ for large enough m. Let $|\psi\rangle$ and $|\psi'\rangle$ be states in the Hilbert space generated from $|0\rangle$. Consider

$$\langle\psi|\, U_m^\dagger(U - U_m)\, |\psi'\rangle = \langle\psi|\, V_m - 1\, |\psi'\rangle \quad (29a)$$

with

$$V_m = \exp\left[\tfrac{1}{2}n_0^{\frac{1}{2}} \sum_{|\mathbf{k}|\geq m} \theta(\mathbf{k})(a_\mathbf{k} - a_\mathbf{k}^\dagger)\right]. \quad (29b)$$

Using the fact that $n_0 = 1$ one obtains

$$V_m = \prod_{|\mathbf{k}|\geq m} [\cos\tfrac{1}{2}\theta(\mathbf{k}) + (a_\mathbf{k} - a_\mathbf{k}^\dagger)\sin\tfrac{1}{2}\theta(\mathbf{k})]. \quad (30)$$

If $A(\mathbf{k})\epsilon(\mathbf{k})^{-1} \to 0$ for $|\mathbf{k}| \to \infty$, $\sin\frac{1}{2}\theta(\mathbf{k})$ is a decreasing function of $\mathbf{k}$ for large $\mathbf{k}$. The only nondiagonal element which need be considered then is

$$\langle 0|\, a_\mathbf{k} V_m\, |0\rangle = \sin\tfrac{1}{2}\theta(\mathbf{k}), \qquad |\mathbf{k}| = m. \quad (31)$$

From the above condition this off-diagonal element can be made arbitrarily small by choosing m large enough. For the diagonal elements it suffices to consider

$$\langle 0|\, V_m - 1\, |0\rangle = \prod_{|\mathbf{k}|\geq m} \cos\tfrac{1}{2}\theta(\mathbf{k}) - 1. \quad (32)$$

Let

$$F_m \equiv \ln \prod_{|\mathbf{k}|\geq m} \cos\tfrac{1}{2}\theta(\mathbf{k}) = \sum_{|\mathbf{k}|\geq m} \ln\cos\tfrac{1}{2}\theta(\mathbf{k}) \quad (33)$$

From Eqs. (25) and (26b) one finds that

$$\cos\tfrac{1}{2}\theta(\mathbf{k}) = \tfrac{1}{2} + \tfrac{1}{2}\omega(\mathbf{k})W(\mathbf{k})^{-1} \quad (34a)$$

$$\to 1 - A(\mathbf{k})^2\omega(\mathbf{k})^{-2} \quad (34b)$$

for large $|\mathbf{k}|$. Therefore

$$F_m \to -\sum_{|\mathbf{k}|<m} A(\mathbf{k})^2\omega(\mathbf{k})^{-2} \quad (35a)$$

and

$$\exp(F_m) \to \exp\left[-\sum_{|\mathbf{k}|\geq m} A(\mathbf{k})^2\omega(\mathbf{k})^{-2}\right]. \quad (35b)$$

One sees, then, that $|\exp(F_m - 1)|$ can be made arbitrarily small for large enough m if and only if

$$\lim_{m\to\infty} \sum_{|\mathbf{k}|\geq m} A(\mathbf{k})^2\omega(\mathbf{k})^{-1} = 0. \quad (36)$$

This is equivalent to the requirement that

$$I_{k_0} \equiv \Omega \int_{|\mathbf{k}|\geq k_0} d^3k\, A(\mathbf{k})^2\omega(\mathbf{k})^{-2}, \quad (37)$$

with k_0 large enough to exclude any singularities of the integrand, exist. This requirement is essentially the same as that obtained for the static scalar field model.[2] One notes that, in the case of $\Omega \to \infty$, U exists and is not identically one only if $A \propto \Omega^{-\frac{1}{2}}$. From this discussion it is obvious that the standard perturbation theory would encounter the same problems in this quantization procedure as in the Bose quantization procedure. (It can be shown that there exists a nondenumerable number of inequivalent irreducible representations of the commutation relations.)

The ground-state energy of the system is given by

$$E_0 = \tfrac{1}{2}n_0 \sum_\mathbf{k} [\omega(\mathbf{k}) - W(\mathbf{k})] \quad (38)$$

with

$$N_0 = \tfrac{1}{2}n_0 \sum_\mathbf{k} [1 - \omega(\mathbf{k})W(\mathbf{k})^{-1}] \quad (39)$$

particles in the system. There are two cases of particular interest: $n_0 = 1$ and $n_0 \to \infty$. For $n_0 \to \infty$,

$$E_0 \to \sum_\mathbf{k} A(\mathbf{k})^2\omega(\mathbf{k})^{-1} \quad (40)$$

and

$$N_0 \to \sum_\mathbf{k} A(\mathbf{k})^2\omega(\mathbf{k})^{-2}, \quad (41)$$

which, if $\epsilon(\mathbf{k})$ and $A(\mathbf{k})$ are given by Eq. (23), are the same as the results of the static scalar field model.[2] In this same limit

$$W(\mathbf{k}) \to \omega(\mathbf{k}). \quad (42)$$

As expected, then, the theory converges to the Bose theory for large n_0. For $n_0 = 1$,

$$E_0 = \tfrac{1}{2} \sum_\mathbf{k} \{\omega(\mathbf{k}) - [\omega(\mathbf{k})^2 + 4A(\mathbf{k})^2]^{\frac{1}{2}}\} \quad (43)$$

and

$$N_0 = \tfrac{1}{2} \sum_\mathbf{k} \{1 - \omega(\mathbf{k})[\omega^2(\mathbf{k}) + 4A(\mathbf{k})^2]^{-\frac{1}{2}}\}. \quad (44)$$

If one sets

$$\omega(\mathbf{k}) = \mathbf{k}^2/m - 2\mu, \quad (45)$$

the form of Eqs. (43) and (44) is suggestive of the BCS model of superconductivity.[3] In fact this *is* the BCS model. To see this consider the commutation relations for the BCS pairs:

$$[b_{\mathbf{k}+}^\dagger b_{-\mathbf{k}-}^\dagger,\ b_{-\mathbf{k}'-}b_{\mathbf{k}'+}] = \delta_{\mathbf{k}\mathbf{k}'}(b_{\mathbf{k}+}^\dagger b_{\mathbf{k}+} + b_{-\mathbf{k}-}^\dagger b_{-\mathbf{k}-} - 1), \quad (46)$$

where the b's are fermi operators and $+$ $(-)$ signifies spin up (down). If one defines

$$N_{\mathbf{k}} \equiv \tfrac{1}{2}(b^{\dagger}_{\mathbf{k}+}b_{\mathbf{k}+} + b^{\dagger}_{-\mathbf{k}-}b_{-\mathbf{k}-}) \tag{47}$$

and

$$a_{\mathbf{k}} \equiv b_{-\mathbf{k}-}b_{\mathbf{k}+}, \tag{48}$$

then

$$[a^{\dagger}_{\mathbf{k}}, a_{\mathbf{k}'}] = 2\,\delta_{\mathbf{k}\mathbf{k}'}(N_{\mathbf{k}} - \tfrac{1}{2}), \tag{49}$$

$$[a_{\mathbf{k}}, N_{\mathbf{k}'}] = \delta_{\mathbf{k}\mathbf{k}'}a_{\mathbf{k}}, \tag{50}$$

and

$$[a_{\mathbf{k}}, a_{\mathbf{k}'}] = 0, \tag{51}$$

which are the proposed commutation relations for the case $n_0 = 1$. This last observation is noted as a curiosity rather than as a realization of the proposed quantization procedure since, the operators defined in Eq. (48) annihilate "particles" of zero total momentum rather than momentum $\mathbf{k}$.

CONCLUSION

Even though it appears that one can approximate results of the Bose theory with arbitrary accuracy, it is clear that, due to the added difficulties in the calculations, one would not elect to use this theory unless a saturation of the number of particles in a given momentum state is observed. Whether systems exhibiting this phenomena exist is unknown. There are two systems which one can presently consider. First is the photon field in masers or lasers. This system does not appear to show any saturation.[4] The second system is helium II. In this case the problem of a weakly coupled, finite, Bose gas with saturation needs to be carefully analyzed in order to obtain its properties.

Beyond the question of the realization of this theory is the deeper question already posed by the parastatics. Does there exist a fundamental principle which determines the "correct" quantization procedure for a given system?

ACKNOWLEDGMENT

I am indebted to Professor Max Dresden for his encouragement and suggestions during the course of this work.

[4] L. Mandel and E. Wolf, Rev. Mod. Phys. 37, 231 (1965).

An Axiomatic Approach to the Formalism of Quantum Mechanics. I.*

EDUARD PRUGOVEČKI
Palmer Physical Laboratory, Princeton University, Princeton, New Jersey
(Received 11 February 1965)

An axiomatic formulation of a quantum mechanical formalism is given. The formulation is not in terms of objects associated with the Hilbert space, but in terms of a different kind of objects for which the name "complex probability measures" has been chosen. It is shown that the conventional Hilbert-space formalism obeys the given axioms. A few consequences of the axioms are investigated, some of which are found useful in the second part of this work.

1. INTRODUCTION

THERE are today a few axiomatic approaches to the formalism of quantum mechanics which differ among themselves in almost every respect: in the content of the postulated material, in the mathematical apparatus employed, in the goals attempted, in the rationale motivating such an investigation, etc. Nevertheless, these approaches can be classified into two major groups.

The works assigned to the first group have in common a basically algebraic approach to the problem. A direct physical meaning is attached to various algebraic operations between the observables of the theory, and it is consequently required that the set of all observables has a definite algebraic structure. The origin of this kind of approach can be traced to the work of Jordan.[1] It has been elaborated by Jordan, von Neumann, and Wigner[2,3] in the hope that it would provide a generalization of quantum mechanics suitable to solve "...the (probably) fundamental difficulties resulting when one attempts to apply quantum mechanics to questions in relativistic and nuclear phenomena."[4] The same general approach can be found in the work of Segal[5] on a generalization of quantum mechanics.

The works belonging to the second group are closely related, in the nature of their approach, to mathematical logic and to lattice theory. The attention in these cases is not focused on the entire set of observables and the algebraic operations between them, as in the previous case, but on the set of a *special* kind of observables called questions[6] and on logical (lattice-type) operations within this set. This type of approach was initiated by Birkhoff and von Neumann,[7] and it can be traced to the work of Finkelstein, Jauch, and Speiser on quaternion quantum mechanics,[8] as well as in the axiomatic system proposed by Mackey[9] and further investigated by Pool.[10] It certainly culminates in the work of Piron.[11] This approach has also been used recently in an attempt to prove the impossibility of the existence of "hidden observables" in quantum mechanics.[12]

Our approach is essentially different in many respects from the above mentioned approaches. From the physical point of view, it is motivated by a consideration of the problem of simultaneous measurement of incompatible observables in quantum mechanics.[13,14] Mathematically, new tools can be found to generalize the apparatus of mathematical statistics (which is formulated only for the case of measurements on compatible observables) by introducing the concept of "complex probability measure."[14] This way, instead of starting with the set of observables or with the set of questions, the problem is attacked by first analyzing the concept of physical state and of observable, and then by giving a very general mathematical framework (Sec. 2.2) providing a general implicit definition of these con-

* This paper is based, in part, upon the author's doctoral dissertation submitted to the Physics Department of Princeton University.

[1] P. Jordan, Z. Phys. **80**, 285 (1933); Nachr. Akad. Wiss. Goettinger Math-Phys. Kl. **IIa**, 569 (1932); *ibid.*, 209 (1933).

[2] P. Jordan, J. von Neumann, and E. P. Wigner, Ann. Math. **35**, 29 (1934).

[3] J. von Neumann, Mat. Sborn. **1**, 415 (1936).

[4] Ref. 2, p. 29.

[5] I. E. Segal, Ann. Math. **48**, 930 (1948).

[6] A question is defined as an observable whose spectrum consists only of two points: 0 and 1. (See also Sec. 3.1.)

[7] G. Birkhoff and J. von Neumann, Ann. Math. **37**, 823 (1936).

[8] D. Finkelstein, J. M. Jauch, and D. Speiser, *Notes on Quaternion Quantum Mechanics*, CERN 59-7 (Part I), 59-9 (Part II), 59-17 (Part III), Geneve, 1959.

[9] G. W. Mackey, *Mathematical Foundations of Quantum Mechanics* (W. A. Benjamin Company, Inc., 1963).

[10] J. C. T. Pool, *Simultaneous Observability and the Logic of Quantum Mechanics*, SUI 63-17, Dept. of Physics and Astronomy, State University of Iowa, Iowa City, 1963.

[11] C. Piron. Helv. Phys. Acta **37**, 439 (1964).

[12] J. M. Jauch and C. Piron, Helv. Phys. Acta **36**, 827 (1963).

[13] H. Margenau, Phil. Sci. **25**, 23 (1958); R. N. Hill and H. Margenau, Progr. Theoret. Phys. (Kyoto), **26**, 722 (1961).

[14] E. Prugovečki, Ph.D. Thesis, Princeton University. We intend to give an analysis of this problem in future papers on the theory of measurement in quantum mechanics.

cepts. Thus we arrive at the axioms formulated in Sec. 2.2. In Sec. 2.3, we show that the conventional Hilbert-space formalism of quantum mechanics can be interpreted in such a way that it obeys the proposed axiomatic system which simultaneously demonstrates the logical compatibility of these axioms. In Sec. 3, we derive some consequences of the given axioms which sometimes enable us to compare them with other axiomatic approaches.

In Part II of this work, we investigate the conditions under which our axiomatic system is equivalent to the conventional Hilbert-space formalism of quantum mechanics. We show that one additional axiom is a necessary and sufficient condition for this equivalence (defined in a precise sense). It is interesting that we get this equivalence only by imposing a certain algebraic and topological structure on the basic objects (observables and physical states) of our language without having to resort to some *ad hoc* external condition which would explicitly require the equivalence of our language with the Hilbert-space formalism.[15]

From the physical point of view the axioms of Sec. 2.2, together with those in Part II, provide a language which, with a certain physical interpretation,[14] indicates a whole variety of generalizations of the conventional Hilbert space of quantum mechanics simply by dropping any of the axioms,[16] with the exception of Axiom I. From the mathematical point of view, these axioms and hypotheses represent an algebraic and topological analysis of the conventional Hilbert-space formalism from an angle which enables us to make use of the modern tools used in the measure theory and in the theory of Banach algebras.

2. AN AXIOMATIC SYSTEM FOR A GENERAL QUANTUM MECHANICAL FORMALISM

2.1. Preliminaries

Following the currently widely accepted point of view on the structure of a scientific theory,[17] we can analyze a physical theory into the following four major constituents:

(a) The *language* or *formalism* of the theory, consisting of a mathematical and logical framework formulated in the form of axioms and defining implicitly (definition by postulation or synthetic definition) the fundamental concepts of the theory (the primitive concepts or primitives).

(b) The *rules of deduction* telling us which are the allowed manipulations by means of which we can derive theorems from the axioms.

(c) The *dynamical law*, which, by imposing additional conditions on some primitives appearing in the theory, gives it predictive power.

(d) The *correspondence rules* relating some of the symbols appearing in the formalism to laboratory procedures, i.e., relating theory and experience.

The primitive concepts of our theory are called physical states and observables. The set of all observables is denoted by $\mathcal{O}$ and the set of all physical states by $\mathcal{S}$. We hope to analyze somewhere else in more detail the relation between observables and measurable quantities (i.e., quantities which are directly measured in the laboratory) as well as between physical states and the states which can be and are constructed through laboratory procedures. It suffices here to say that, in order to have a link between theory and experience, we have to relate explicitely with the help of correspondence rules at least some of the observables to measureable quantities, i.e., to laboratory procedures. Such observables to which correspondence rules are attached are called fundamental observables. Historically, in the course of the development of quantum mechanics, observables such as position coordinates and momentum or spin components have been considered as fundamental.

Through the axioms themselves we enlarge the concept of fundamental observable and obtain the general concept of observable, which does not have to stay in a direct relation to measurable quantities. However, the first axiom already uses this concept and deals with the set $\mathcal{O}$ of all observables, while it is only in Axioms IV and V that we begin to discuss fundamental observables. The reasons for presenting the axioms in the particular fashion of the next section is a matter of convenience in presentation and does not contradict the aforesaid. Namely, we have to remember that the axioms *have to be taken as a whole*, and that the division of this whole in different axioms is made to ease the exposition and for epistomological reasons; the individual axioms represent the natural links in the chains—links which will be easiest to change or replace in case of neces-

[15] Like Mackey's axiom VII (Ref. 9, p. 71) in which it is required that the introduced partially ordered set of all questions is isomorphic to the partially ordered set of all closed subspaces of a separable infinite dimensional Hilbert space.

[16] This certainly is not the case with the Hilbert space formalism which, under the current physical interpretation, forms a monolitic whole, unyeilding to any simple generalizations.

[17] A discussion of these matters, suitable for the needs of physicists, can be found in a paper by L. Tisza, Rev. Mod. Phys. 35, 151 (1963).

sity. It has to be remembered that each such change changes the logical structure of the axioms as a whole, and that we then deal with a *different* language. There are no measures to enable us to say that such changes are "small" or "great" in some rigorous sense.

In defining the concept of a physical state P, we use the concept of complex probability measure as fundamental. We do not try to justify here the introduction of this concept,[14] but we hope to do this elsewhere. We limit ourselves to simply define such a concept, while, in Sec. 2.3, we establish that such objects can really be found in the conventional Hilbert-space formalism of quantum mechanics.

For a given ordered set of n compatible or incompatible observables $\alpha_1, \cdots \alpha_n (1 \leq n < +\infty)$, a *complex probability measure* $P^{\alpha_1,\cdots\alpha_n}$ is defined as an ordered pair $[P_+^{\alpha_1,\cdots\alpha_n}(B), P_-^{\alpha_1,\cdots\alpha_n}(B)]$ of, in general, signed measures on the class $\mathfrak{B}^{(\alpha_1,\cdots\alpha_n)}$ of all Borel sets B of an n-dimensional Euclidean space $\mathbf{R}^{(\alpha_1,\cdots\alpha_n)}$, which satisfy the conditions[18]

$$P_+^{\alpha_1,\cdots\alpha_n}(\mathbf{R}^n) = 1, \qquad P_-^{\alpha_1,\cdots\alpha_n}(\mathbf{R}^n) = 0.$$

In condition (3) Axiom I, we will explicitly require that in case $\alpha_1, \cdots \alpha_n$ are compatible, $P_-^{\alpha_1,\cdots\alpha_n}(B)$ is identically equal to zero while $P_+^{\alpha_1,\cdots\alpha_n}(B)$ is a probability measure in the conventional sense, i.e., a positive-definite normalized measure, thus implicitly defining the concept of compatibility in theoretical terms.

2.2. The Axioms

Axiom I: A physical state P is given by a rule which assigns to each n-tuple $(\alpha_1, \cdots \alpha_n)$, $n = 1, 2, 3, \cdots$, of observables $\alpha_1, \cdots, \alpha_n$ a complex probability measure

$$[P_+^{\alpha_1,\cdots\alpha_n}(B), P_-^{\alpha_1,\cdots\alpha_n}(B)], \qquad B \in \mathfrak{B}^{(\alpha_1,\cdots\alpha_n)} \tag{2.1}$$

for which the shorter notation

$$P^{\alpha_1,\cdots\alpha_n}(B) = P_+^{\alpha_1,\cdots\alpha_n}(B) + iP_-^{\alpha_1,\cdots\alpha_n}(B) \tag{2.2}$$

can be introduced.

These complex measures have to fulfill the following conditions[19]:

(1) $$P^{\alpha_1;\cdots;\alpha_{j-1};\alpha_j;\alpha_{j+1};\cdots;\alpha_n} \times (B_1 \times \cdots \times B_{j-1} \times \mathbf{R}^1 \times B_{j+1} \times \cdots \times B_n)$$
$$= P^{\alpha_1;\cdots;\alpha_{j-1};\alpha_{j+1};\cdots;\alpha_n} \times (B_1 \times \cdots \times B_{j-1} \times B_{j+1} \times \cdots \times B_n), \tag{2.3}$$

where $B_1, \cdots B_{j-1}, B_{j+1}, \cdots B_n \in \mathfrak{B}^1$, $n = 1, 2, \cdots$, $j = 1, \cdots n$.

(2) If $\hat{\mathfrak{O}}$ denotes the set of n-tuples of observables, including the empty set, $n = 0, 1, 2, 3, \cdots$, then for any $\hat{\beta}_1', \hat{\beta}_2'' \in \hat{\mathfrak{O}}$ we must have[20]

$$P^{\hat{\beta}';\alpha_1;\cdots;\alpha_n;\hat{\beta}''}(B' \times B_1 \times \cdots \times B_n \times B'') = P^{\hat{\beta}';\alpha;\hat{\beta}''}[B' \times (B_1 \cap \cdots \cap B_n) \times B''],$$
$$B_1, B_2, \cdots B_n \in \mathfrak{B}^1,$$
$$n = 2, 3, 4, \cdots, B' \in \mathfrak{B}^{\hat{\beta}'}, \quad B'' \in \mathfrak{B}^{\hat{\beta}''}, \tag{2.4}$$

if $\alpha_1, \cdots, \alpha_n$ denote the same observable α, i.e., $\alpha = \alpha_1 = \cdots = \alpha_n$.

(3) A set of observables $\alpha_1, \cdots, \alpha_n$ is called compatible, in shortened notation $\{\alpha_1, \cdots \alpha_n\} = C$, if and only if the following conditions are satisfied:

(a) $$0 \leq P_+^{\alpha_1,\cdots\alpha_n}(B) \leq 1, \qquad B \in \mathfrak{B}^{(\alpha_1,\cdots\alpha_n)},$$

(b) $$P_-^{\alpha_1,\cdots\alpha_n}(B) = 0, \qquad B \in \mathfrak{B}^{(\alpha_1,\cdots\alpha_n)},$$

(c) $$P^{\hat{\beta}';\alpha_1;\cdots;\alpha_n;\hat{\beta}''}(B' \times B_1 \times \cdots \times B_n \times B'') = P^{\hat{\beta}';\alpha_{k_1};\cdots;\alpha_{k_n};\hat{\beta}''} \times (B' \times B_{k_1} \times \cdots \times B_{k_n} \times B''),$$
$$B' \in \mathfrak{B}^{\hat{\beta}'}, \quad B'' \in \mathfrak{B}^{\hat{\beta}''},$$
$$\hat{\beta}', \hat{\beta}'' \in \hat{\mathfrak{O}}, \quad B_1, \cdots B_n \in \mathfrak{B}^1,$$

for every permutation $(k_1, \cdots k_n)$ of the indices $1, \cdots n$.

(4) Every observable is compatible with itself, i.e.,

$$0 \leq P_+^{\alpha}(B) \leq 1, \quad P_-^{\alpha}(B) = 0, \quad B \in \mathfrak{B}^{(\alpha)}, \quad \alpha \in \mathfrak{O},$$

with $P_+^{\alpha}(\mathbf{R}^1) = 1$.

Remarks. As we see, the concept of compatibility is defined *within* our object-language itself in such a manner that, in case of compatible observables, our language reduces to the conventional language of statistics. Perhaps the most outstanding feature of the compatibility concept introduced by Axiom I(3) compared with the case of Hilbert-space formalism is that the compatibility of n observables is not guaranteed, so far, by the compatibility of each pair of these n observables. In the Hilbert-

[18] We also use the notation in which $\mathbf{R}^n$ denotes any n-dimensional Euclidian space $(1 \leq n < +\infty)$ and $\mathfrak{B}^n$ denotes the class of all Borel subsets of such a space.

[19] We introduce the obvious notation in which, when we have for arguments of $P^{\cdots\alpha\cdots}$ products of sets, the kth factor of this product is related to the kth *group* of observables between two *semicolons*.

[20] The convenient notation $\mathfrak{B}^{\hat{\alpha}}$ is introduced, meaning the same as $\mathfrak{B}^{(\alpha_1,\cdots,\alpha_n)}$, i.e. the family of all Borel sets in the space $\mathbf{R}^{(\alpha_1,\cdots,\alpha_n)}$ corresponding to the n-tuple $\hat{\alpha}$ of observables. Similarly, we will sometimes write $\mathbf{R}^{\hat{\alpha}}$ for $\mathbf{R}^{(\alpha_1,\cdots,\alpha_n)}$.

space formalism, where compatibility is translated in commutativity between the respective operators, this is true. It is only through the later introduction of additional hypotheses which guarantee the logical equivalence of our formalism to the Hilbert-space formalism that our language also aquires this feature. Otherwise it is disputable that this last feature is desirable in itself, e.g., it is not evident on empirical basis whether three quantities can be measured with arbitrary precision simultaneously if each two of them are compatible.

The last point of Axiom I introduces explicitly the assumption that each observable is considered compatible with itself in the sense that there is no limitation on the precision of its measurement. We do not consider that compatibility with itself is an intrinsic property of the concept of observable, but it is a feature which has proven convenient to assume until now, and which is not *necessarily* a feature of everlasting value; e.g., some of Heisenberg's ideas about a fundamental length would make the position coordinate observables incompatible with themselves.

It is convenient to state here an obvious proposition.

Proposition 1: For any $\alpha_1, \cdots, \alpha_n \in \mathfrak{O}$, $P \in \mathfrak{S}$,

$$P^{\alpha_1,\cdots\alpha_n}(\mathbf{R}^n) = 1.$$

Proof. To prove that $P^{\alpha_1,\cdots,\alpha_n}(\mathbf{R}^n) = 1$, we proceed by induction. Axiom I(4) tells us that this is true for $n = 1$. Assuming that the statement is true for any $n - 1$ observables, we get $P^{\alpha_1,\cdots,\alpha_n}(\mathbf{R}^n) = P^{\alpha_1,\cdots,\alpha_{n-1};\alpha_n}(\mathbf{R}^{n-1} \times \mathbf{R}^1) = P^{\alpha_1,\cdots,\alpha_{n-1}}(R^{n-1}) = 1$, where the last step follows from (2.3). Q.E.D.

Axiom II (*$\mathfrak{S}$–identity axiom*): Two physical states, P_1 and P_2, are considered identical if and only if $P_1^{\hat{\alpha}}(B) = P_2^{\hat{\alpha}}(B)$ for all $\hat{\alpha} \in \hat{\mathfrak{O}}$, $B \in \mathfrak{B}^{\hat{\alpha}}$.

Remarks. This axiom states the obvious. However, sometimes it is better to state explicitly the "obvious," because this then stresses a certain logical structure which does not have to be accepted on *a priori* grounds.

As a consequence of well-known measure-theoretical theorems[21] we can state Axiom II in the following equivalent form.

Proposition 2: Two physical states P_1, P_2 are identical if and only if

$$P_1^{\alpha_1;\cdots;\alpha_n}(B_1 \times \cdots \times B_n) = P_2^{\alpha_1;\cdots;\alpha_n}(B_1 \times \cdots \times B_n)$$

for any $\alpha_1, \cdots \alpha_n \in \mathfrak{O}$, $B_1, \cdots B_n \in \mathcal{G}^1$.

[21] P. R. Halmos, *Measure Theory* (D. van Nostrand Co., Princeton, N. J., 1961), Theorem *A*, p. 54 and §§ 33, 35.

Here $\mathcal{G}^1$ denotes the set of all nondegenerate (i.e., infinite or finite, closed, open, or half-open on any side, but consisting of more than one point) one-dimensional intervals.

Axiom III (*$\mathfrak{O}$–identity axiom*): Two elements α and β of the set of all observables are identical if and only if

$$P^{\hat{\varphi}';\alpha;\hat{\varphi}''}(B' \times B \times B'') = P^{\hat{\varphi}';\beta;\hat{\varphi}''}(B' \times B \times B'')$$

for all $\hat{\varphi}', \hat{\varphi}'' \in \hat{\mathfrak{O}}$, $B' \in \mathfrak{B}^{\hat{\varphi}'}$, $B'' \in \mathfrak{B}^{\hat{\varphi}''}$, $B \in \mathfrak{B}^1$, $P \in \mathfrak{S}$.

Remarks. Axiom III states the relation of identity between observables in terms of physical states. From the logical point of view, this axiom can be regarded as a synthetic *definition* of the concept of identity between two observables. From the epistomological point of view, in the process of building actual theoretical languages in terms of these axioms, it can be considered as a *condition* on the set $\mathfrak{S}$, which is required to be sufficiently rich to ensure that the concept of identity between observables introduced by Axiom III coincides with the concept of identity formulated in terms of equivalent operational definitions of measurable quantities represented by these observables.

It is convenient at this point to state, in the form of propositions, two equivalent formulations of Axiom III.

Proposition 3: If $\alpha, \beta \in \mathfrak{O}$, then $\alpha = \beta$ if and only if

$$P^{\varphi_1';\cdots;\varphi_k';\alpha;\varphi_1'';\cdots;\varphi_l''} \times (B_1' \times \cdots \times B_k' \times B \times B_1'' \times \cdots \times B_l'') = P^{\varphi_1';\cdots;\varphi_k';\beta;\varphi_1'';\cdots;\varphi_l''} \times (B_1' \times \cdots \times B_k' \times B \times B_1'' \times \cdots \times B_l'')$$

for any $\varphi_1', \cdots \varphi_k', \varphi_1, \cdots \varphi_l'' \in \mathfrak{O}$ and any $B, B_1', \cdots B_k', B_1'', \cdots B_l' \in \mathcal{G}^1$.

This proposition is a consequence of the definition of a complex signed measure as a σ-additive set function and of measure-theoretical theorems[22] which enable us to state that $\mathfrak{B}^{k+l+1}$ is the Boolean σ-algebra[23] generated by $\mathcal{G}^{k+l+1}$. Here $\mathcal{G}^n$ denotes the class of sets which are products of n sets belonging to $\mathcal{G}^1$, i.e., $\mathcal{G}^n = \mathcal{G}^1 \times \cdots \times \mathcal{G}^1$ (n times). Because of the statement above, it is easy to see that Proposition 2 is logically equivalent to the following proposition.

[22] Ibid., pp. 140, 143.
[23] Ibid., p. 21.

Proposition 4: If $\alpha, \beta \in \mathcal{O}$ then $\alpha = \beta$ if and only if

$$P^{\hat{\beta}';\alpha;\hat{\beta}''}(B) = P^{\hat{\beta}';\beta;\hat{\beta}''}(B), \qquad B \in \mathfrak{B}^{k+l+1}, \quad k, l = 0, 1, 2, \cdots$$

for any $\hat{\varphi}', \hat{\varphi}'' \in \hat{\mathcal{O}}$ ($\hat{\varphi}'$ is a k-tuple and $\hat{\varphi}''$ is an l-tuple).

As we have the logical chain Axiom II → Proposition 3 → Proposition 4 → Axiom II, we have established the logical equivalence of these three statements.

To be able to formulate the next axiom, we have to enlarge our notation and terminology. By $\mathfrak{F}_B^n$ we denote the set of all real-valued Borel-measurable functions on $\mathbf{R}^n$. We will say that an observable α is a *function of the compatible observables* $\alpha_1, \cdots \alpha_n$ if there is a $f(\lambda_1, \cdots \lambda_n) \in \mathfrak{F}_B^n$ such that

$$P^{\hat{\beta}';\alpha;\hat{\beta}''}(B' \times B \times B'') = P^{\hat{\beta}';\alpha_1,\cdots,\alpha_n;\hat{\beta}''}(B' \times f^{-1}(B) \times B''), \qquad (2.5)$$
$$B \in \mathfrak{B}^1, \quad B' \in \mathfrak{B}^{\hat{\beta}'}, \quad B'' \in \mathfrak{B}^{\hat{\beta}''},$$

for all $P \in \mathcal{S}$. In that case we write $\alpha = f(\alpha_1, \cdots \alpha_n)$. It is obvious that, because of Axiom III, we can associate with any $f \in \mathfrak{F}_B^n$ only one observable (if any!).

Axiom IV: For any finite number of compatible observables $\alpha_1, \cdots \alpha_n$ and, for any $f(\lambda_1, \cdots \lambda_n) \in \mathfrak{F}_B^n$, there exists an observable $\alpha = f(\alpha_1, \cdots \alpha_n)$.

Remarks. Axiom IV is introduced for mathematical convenience and its role will be better understood in the light of later mathematical developments. However, it has to be understood that, contrary to the "intuitive" view on this matter, we consider that (rigorously speaking) Axiom IV introduces a more general concept of observable than that which associates an observable through correspondence rules with measurable quantities. The intuitive conception of the measuring (or preparing) process is based on a simplified visualization of this process which compares it with a picking of a (finite) random sample out of a (in general, infinite) parent population. Such a conception, which is correct in case of observables with pure point spectrum[24] and very helpful in many cases of observables having a continuous spectrum, can be very misleading, because it does not take into account the existence of the errors of each individual measurement, which are integral parts of each measurement or state preparation. These matters have been discussed from a general point of view in Ref. 14. Therefore, we would only like to consider here the consequences of this point of view—namely, that the existence of correspondence rules for an observable α does not imply the existence of such rules for a function $\beta = f(\alpha)$ of α; or, loosely speaking, that a preparation or measurement of α does not imply in general a preparation or measurement of $\beta = f(\alpha)$.

Take the following example: α is an observable with continuous spectrum ranging over all $\mathbf{R}^1$ and $f(\lambda)$ is a step function:

$$f(\lambda) = \begin{cases} 0 & \text{for } \lambda < \lambda_0, \\ 1 & \text{for } \lambda \geq \lambda_0. \end{cases}$$

Now, it can happen (and it sometimes happens!) that we determine an empirical state localized in the interval $(\lambda_0 - \epsilon, \lambda_0 + \epsilon)$ around λ_0 with an apparatus having such characteristics (least count, etc.) that the errors of individual measurement are practically as big as ϵ. We are then dealing with a perfectly meaningful physical situation in which the empirical state[14] is represented by histograms which are ≈ 1 inside the interval $(\lambda_0 - \epsilon, \lambda_0 + \epsilon)$, and zero outside it (the detailed description depends on the actual situation). With the help of rules of correspondence we can assign to such an empirical state physical states determined by distributions $P^\alpha(B)$, $B \in \mathfrak{B}^1$, which are in concordance with that empirical state. However, in such a case we can say nothing about the empirical state in terms of $\beta = f(\alpha)$, and we can assign $P^\beta(B)$ only indirectly by the *definition* $P^\beta(B) = P^\alpha[f^{-1}(B)]$.

Conclusion: The apparatus which "measures" α does not necessarily have to measure a function $\beta = f(\alpha)$ of α; the question of whether there are any measurable quantities which can be directly associated with such a function $f(\alpha)$ is a purely empirical one and cannot be decided *a priori*. (Notice that the above considerations are valid for any given apparatus, because for *any given* apparatus we can choose an appropriate ϵ for which these considerations can be carried through in the same manner!)

Before introducing the next axiom, we have to enlarge again our terminology. We say that the elements of a given set of observables are *Borel-independent* if no observable in this set is a function of a finite number of compatible observables from the same set.

Axiom V: There is at least one fundamental set $\mathcal{O}'$ of observables, i.e., a subset $\mathcal{O}'$ of $\mathcal{O}$ such that the elements of $\mathcal{O}'$ are Borel-independent and each ele-

[24] A rigorous definition of the concept of the spectrum of an observable is given after Ax. VII.

ment of $\mathfrak{O}$ is a function of a finite number of compatible observables belonging to $\mathfrak{O}'$.

Remarks. Axiom V imposes a further condition on the set $\mathfrak{S}$ by requiring that there be at least one set of Borel independent observables. In many cases we can start with a set $\mathfrak{O}'$ of fundamental observables, which we consider to describe the system completely, and if we have defined on it a set of physical states such that the elements of $\mathfrak{O}'$ are Borel independent, then we can *define* the set $\mathfrak{O}$ of all observables as the set of all functions defined on finite compatible subsets of $\mathfrak{O}'$.

A concrete example of a fundamental set of observables describing a free electron are the three Cartesian coordinates x, y, z in three *given* orthogonal directions, the three momentum coordinates p_x, p_y, p_z along these given axes, and the projections of the spin in any direction (as well as the time t if we treat the problem relativistically). The coordinates x', y', z' along any other axes, the angular momentum M_x, M_y, M_z, the energy E, etc., will be functions of the fundamental observables belonging to the specified fundamental set [for example, $x' = a_{11}x + a_{12}y + a_{13}z$ where $||a_{ik}||$, $i, k = 1, 2, 3$ is the transformation matrix: $E = c(p_x^2 + p_y^2 + p_z^2)^{\frac{1}{2}}$, and $M_x = yp_z - zp_y$].

Axiom VI (*convexity of* $\mathfrak{S}$):

(1) If P_1 and P_2 are two physical states, then for any real number $0 \leq t \leq 1$ the complex signed measures

$$P^{\mathfrak{a}}(B) = tP_1^{\mathfrak{a}}(B) + (1 - t)P_2^{\mathfrak{a}}(B), \quad \mathfrak{a} \in \hat{\mathfrak{O}}, \quad B \in \mathfrak{B}^{\mathfrak{a}}$$

will define a physical state, denoted by $P = tP_1 + (1 - t)P_2$.

(2) If, for given P, $P_1 \in \mathfrak{S}$ and for some $0 < t < 1$, the signed probability measures

$$P_2^{\mathfrak{a}}(B) = \frac{1}{1 - t}(P^{\mathfrak{a}}(B) - tP_1^{\mathfrak{a}}(B)), \quad \mathfrak{a} \in \hat{\mathfrak{O}}, \quad B \in \mathfrak{B}^{\mathfrak{a}}$$

are real and positive definite when $\{\mathfrak{a}\} = C$, $\mathfrak{a} \in \hat{\mathfrak{O}}$, then the set $P_2^{\mathfrak{a}}(B)$, $\mathfrak{a} \in \hat{\mathfrak{O}}$ of complex probability measures defines a physical state $P_2 \in \mathfrak{S}$.

Remarks. From the fact that P_1 and P_2 satisfy the conditions of Axiom I, it is easy to establish that $tP_1 + (1 - t)P_2$ and $(1/1 - t)(P - tP_1)$ (with the imposed restrictions) must satisfy the same conditions; besides, Axioms II–V stay unaffected by the addition of Axiom VI. Therefore, from a logical point of view, Axiom VI is a permissible addition which gives to the set $\mathfrak{S}$ the desirable property of convexity and a certain kind of completeness in this respect.

It has to be noted at this point that our concept of physical state refers to pure states as well as mixtures. A physical state is called a *pure state* if the relation $P = tP_1 + (1 - t)P_2$ for some $0 < t < 1$ implies that $P = P_1 = P_2$. A physical state which is not a pure state is called a *mixture.*

Before formulating the next axiom we will have to introduce a "weak" topology in $\mathfrak{S}$. A *weak neighborhood* $W(P_0; \mathfrak{a}; B_1, \cdots B_k; \epsilon)$ of the point $P_0 \in \mathfrak{S}$ is the set of all $P \in \mathfrak{S}$ satisfying the relations

$$|P^{\mathfrak{a}}(B_1) - P_0^{\mathfrak{a}}(B_1)| < \epsilon, \cdots |P^{\mathfrak{a}}(B_k) - P_0^{\mathfrak{a}}(B_k)| < \epsilon. \tag{2.6}$$

It is very easy to check that the set of all neighborhoods of a physical state P_0 for all $\mathfrak{a} \in \hat{\mathfrak{O}}$, $B_1, \cdots B_k \in \hat{\mathfrak{O}}$, $k = 1, 2, \cdots$, and all $\epsilon > 0$ satisfy the conditions[25] for defining a neighborhood basis of the point P_0 in $\mathfrak{S}$. The weak topology is the topology which has as a basis the union of neighborhood bases of all points in $\mathfrak{S}$.

A sequence $P_1, P_2, P_3, \cdots$ of physical states is called a *Cauchy sequence in* $\mathfrak{S}$ if for any $\epsilon > 0$ and any $\mathfrak{a} \in \hat{\mathfrak{O}}$, $B \in \mathfrak{B}^{\mathfrak{a}}$, there is a positive integer $N_0(\mathfrak{a}, B)$ such that

$$|P_m^{\mathfrak{a}}(B) - P_n^{\mathfrak{a}}(B)) < \epsilon \quad \text{for} \quad m, n > N_0(\mathfrak{a}, B).$$

It is obvious that the concept of a Cauchy sequence so defined coincides with the concept of a Cauchy sequence in the weak topology of the space $\mathfrak{S}$.

Axiom VII (*completeness of* $\mathfrak{S}$ *in the weak topology*): The space $\mathfrak{S}$ is complete in the weak topology, i.e., each Cauchy sequence in $\mathfrak{S}$ has a limit which belongs to $\mathfrak{S}$.

Remarks. If $P_1, P_2, \cdots$ is a Cauchy sequence, it is obvious that $\lim_{n\to\infty} P_n^{\mathfrak{a}}(B)$ exists for each $\mathfrak{a} \in \hat{\mathfrak{O}}$, $B \in \mathfrak{B}^{\mathfrak{a}}$. If we denote these limiting values by $P^{\mathfrak{a}}(B)$, it is readily seen that each $P^{\mathfrak{a}}(B)$ defines a complex, additive, finite set function on $\mathfrak{B}^{\mathfrak{a}}$ for each $\mathfrak{a} \in \hat{\mathfrak{O}}$. These set functions satisfy the conditions of the Axiom I because the elements $P_1^{\mathfrak{a}}, P_2^{\mathfrak{a}}, \cdots$ satisfy these conditions. Due to the Vitali–Hahn–Sachs theorem, $P^{\mathfrak{a}}(B)$ is also a *countably* additive set function,[26] i.e., a complex-valued measure. Therefore Axiom VII does not impose any serious restrictions on the structure of $\mathfrak{S}$, but it rather represents a mathematical convenience.

---

[25] M. A. Naimark, *Normed Rings*, English transl. by L. F. Boron (Stechert-Hafner Science Agency, Inc., New York 1959) p. 21.

[26] N. Dunford and J. T. Schwartz *Linear Operators, Part I* (Interscience Publishers, Inc., New York, 1958), Theorem 2, Corollary 4, p. 160.

Introductions of new concepts are necessary for the formulation of the last two axioms.

A Borel set B in $\mathbf{R}^{\hat{\alpha}}$ is called a *zero set* if we have $P^{\hat{\alpha}}(B') = 0$ for all $P \in \mathcal{S}$, $B' \subset B$, $(B' \in \mathfrak{B}^{\hat{\alpha}})$; otherwise B is called a *nonzero set.* We note that, if $\{\hat{\alpha}\} = C$, it is necessary and sufficient for a set $B \in \mathfrak{B}^{\hat{\alpha}}$ to be a zero set so that $P^{\hat{\alpha}}(B) = 0$ for all $P \in \mathcal{S}$.

Two Borel sets, $B_1, B_2 \in \mathfrak{B}^{\hat{\alpha}}$, are called equal modulo $\mathcal{S}$, symbolically $B_1 = B_2 \pmod{\mathcal{S}}$, if their symmetric difference $B_1 \triangle B_2 = (B_1 - B_2) \cup (B_2 - B_1)$ is a zero set. According to such terminology, a zero set is a Borel set which is equal to the empty set modulo $\mathcal{S}$.

The *spectrum of an observable* α is the set of all points in $\mathbf{R}^{(\alpha)}$ having the property that each of their open neighborhoods are nonzero sets. In other words, a point $\lambda \in R^{(\alpha)}$ belongs to the spectrum of $\alpha \in \mathcal{O}$ if for each open interval I containing λ there is at least one $P \in \mathcal{S}$ such that $P^{\alpha}(I) \neq 0$. The spectrum of $\alpha \in \mathcal{O}$ will be denoted by $\mathbf{S}^{(\alpha)}$.

Proposition 5: The spectrum of each observable $\alpha \in \mathcal{O}$ is a closed set.

To establish the truth of this statement, it has only to be noticed that if $\lambda \in \mathbf{R}^{(\alpha)}$ is the accumulation point of points belonging to $\mathbf{S}^{(\alpha)}$, then each open neighborhood of λ is a nonzero set because it contains at least one point belonging to $\mathbf{S}^{(\alpha)}$.

We can generalize the concept of spectrum in a straightforward manner by saying that the *spectrum of an n-tuple $\hat{\alpha}$ of observables* consists of all points of $\mathbf{R}^{\hat{\alpha}}$ having the property that each open interval containing one of such points is a nonzero set.

We will now extend the terminology used in the theory of self-adjoint (hypermaximal) operators in the Hilbert space.[27] We say that a point $\lambda \in \mathbf{R}^{\hat{\alpha}}$ belongs to the *point spectrum* $\mathbf{S}_p^{\hat{\alpha}}$ of $\hat{\alpha}$ if $P^{\hat{\alpha}}(\{\lambda\}) \neq 0$ for at least one $P \in \mathcal{S}$. A point in $\mathbf{R}^{\hat{\alpha}}$ which belongs to the spectrum of $\hat{\alpha}$ but not to the point spectrum is said to belong to the *continuous spectrum* $\mathbf{S}_c^{\hat{\alpha}}$ *of* $\hat{\alpha}$.

An ordered n-tuple $(n > 1)\hat{\alpha}$ of observables is said to have a *pure point spectrum* or *discrete spectrum* if its continuous spectrum is empty. Similarly, its spectrum is said to be a *pure continuous spectrum* if no points of its spectrum belong to the point spectrum. If the spectrum of $\hat{\alpha}$ is neither a pure point spectrum nor a pure continuous spectrum, it is said to be a *mixed spectrum.*

Proposition 6: The spectrum $\mathbf{S}^{\hat{\alpha}}$ of a compatible set $\{\alpha_1, \cdots \alpha_n\}$ of observables is contained in the subset $\mathbf{S}^{(\alpha_1)} \times \cdots \times \mathbf{S}^{(\alpha_n)}$ of $\mathbf{R}^{\hat{\alpha}}$.

Proof: If $\hat{\lambda} = (\lambda_1, \cdots \lambda_n) \in \mathbf{R}^{\hat{\alpha}}$ does not belong to $\mathbf{S}^{(\alpha_1)} \times \cdots \times \mathbf{S}^{(\alpha_n)}$, then at least one $\lambda_j (1 \leq j \leq n)$ does not belong to $\mathbf{S}^{(\alpha_j)}$. Hence, there must be an open zero interval I_j in $\mathbf{R}^{(\alpha_j)}$ containing λ_j. Using the fact that $\{\alpha_1, \cdots \alpha_n\} = C$ as well as (2.3), we can write

$$0 \leq P^{\alpha_1, \cdots \alpha_{j-1}; \alpha_j; \alpha_{j+1}, \cdots \alpha_n}(B' \times I_j \times B'') \leq P^{\alpha_j}(I_j) = 0,$$

$$B' \in \mathfrak{B}^{(\alpha_1, \cdots \alpha_{j-1})}, \qquad B'' \in \mathfrak{B}^{(\alpha_{j+1}, \cdots \alpha_n)}.$$

Hence, we conclude that $\hat{\lambda}$ does not belong to $\mathbf{S}^{\hat{\alpha}}$. Q.E.D.

Proposition 7: The point spectrum $\mathbf{S}_p^{\hat{\alpha}}$ of a compatible set $\{\alpha_1, \cdots \alpha_n\}$ of observables is contained in $\mathbf{S}_p^{(\alpha_1)} \times \cdots \times \mathbf{S}_p^{(\alpha_n)}$.

The proof of this proposition proceeds along similar lines as the proof of Proposition 6.

The proofs of Propositions 6 and 7 are based on the validity of the following statement which is a direct consequence of Axiom I, (1):

Proposition 8: If $\{\alpha_1, \cdots \alpha_n\} = C$ and at least one of the sets $B_1 \in \mathfrak{B}^{(\alpha_1)}, \cdots B_n \in \mathfrak{B}^{(\alpha_n)}$ is a zero set, then $B_1 \times \cdots \times B_n \in \mathfrak{B}^{(\alpha_1, \cdots \alpha_n)}$ is a zero set too.

Axiom VIII: If $\{\alpha_1, \cdots \alpha_n\}$ is a set of incompatible observables and at least one of the Borel sets $B_1 \in \mathfrak{B}^{(\alpha_1)}, \cdots B_n \in \mathfrak{B}^{(\alpha_n)}$ is a zero set, then $B_1 \times \cdots \times B_n \in {}^{(\alpha_1, \cdots \alpha_n)}$ is also a zero set.

Remarks. It is very easy to see that this axiom implies the following proposition.

Proposition 9: The spectrum $\mathbf{S}^{\hat{\alpha}}$ of any n-tuple $\hat{\alpha} = (\alpha_1, \cdots \alpha_n)$ of observables is contained in the direct product $\mathbf{S}^{(\alpha_1)} \times \cdots \times \mathbf{S}^{(\alpha_n)}$ of the spectra of its component observables, while its point spectrum $\mathbf{S}_p^{\hat{\alpha}}$ is contained in the direct product $\mathbf{S}_p^{(\alpha_1)} \times \cdots \times \mathbf{S}_p^{(\alpha_n)}$ of the point spectra of its component observables.

Hence, we see that Axiom VIII enables us to extend Propositions 7 and 8 to any n-tuple of observables. The desirability of this feature is obvious if we realize that, in considering simultaneous measurements of observables $\alpha_1, \cdots \alpha_n$, we consider, *per definitionem*, that the possible numerical outcome of such a measurement (or state-preparation) should always be contained in $\mathbf{S}^{(\alpha_1)} \times \cdots \times \mathbf{S}^{(\alpha_n)}$. This feature is an important characteristic of our concept

[27] R. G. Cook, *Linear Operators* (Macmillan and Company Ltd., London, 1953), p. 182.

of an observable and of simultaneous measurements of observables. The space $\mathbf{R}^{(\alpha_1,\cdots\alpha_n)}$ is only a mathematical artifice when we are dealing with a closed theory, and it would seem, on the first sight, much simpler and more advisable to define each P^A as complex measures over $\mathbf{S}^{(\alpha_1)}\times\cdots\times\mathbf{S}^{(\alpha_n)}$ instead of over $\mathbf{R}^A$.

The situation is, however, quite different under dynamical conditions, i.e., when we are dealing with theories in the developing stage. In that case, we might have to deal not only with observables whose spectrum is completely determined on *a priori* grounds (such as the spectrum of the number of particles of a certain kind in field theory, which consists of all nonnegative integers), but also with observables for which a great number of possible alternatives are compatible with the available data. As an illustration, let us take the energy spectrum of the hydrogen atom.

At the beginning of this century Balmer's formula was established empirically. The discrete spectrum for the energy of the hydrogen atom introduced by Bohr's model was leading to Balmer's formula. However, a pure continuous spectrum consisting of bands which would be much "thinner" than the resolution of the spectroscopes available at that time and placed at the points of Bohr's energy spectrum would have "explained" the empirical facts as well. Of course, such a model would have come in contradiction with the empirical data when the experimental techniques became advanced enough to discover the fine and hyperfine structure of the spectra; but then so did Bohr's model. In such a case, it is desirable that the theory be formulated in such a way so that possible changes in the spectra of some observables do not affect its formal structure.

This example shows that, by formulating the axioms in terms of complex probability measures $P^{\alpha_1,\cdots\alpha_n}(B)$ over $\mathbf{R}^{(\alpha_1,\cdots\alpha_n)}$ instead of over $\mathbf{S}^{(\alpha_1)}\times\cdots\times\mathbf{S}^{(\alpha_n)}$, our language acquires a dynamism reflected in the possibility of correcting, should new experimental material necessitate it, the spectrum of some observables with minimum formal alterations.

Axiom IX: If B is a nonzero Borel set in the space $R^{(\alpha_1,\cdots\alpha_n)}$ assigned to the n compatible observables $\alpha_1,\cdots\alpha_n$, then there is at least one physical state P such that $P^{\alpha_1,\cdots\alpha_n}(B)=1$.

2.3. The compatibility of the axioms and their relation to the Hilbert-space formalism.

We will now show that the conventional Hilbert-space formalism of quantum mechanics is in agreement with the axioms introduced in the previous subsection. Thus, we will at the same time demonstrate that the introduced axioms are logically compatible with one another. To do this we have to translate the Hilbert-space language into the language of our axioms by establishing a correspondence between the primitive concepts of these two languages.

The set $\mathcal{O}$ introduced in the axioms will consist in the case of the Hilbert-space formalism of all the self-adjoint operators representing the fundamental observables and of all Borel functions of these operators. If the spectral decompositions of the operators $A_1,\cdots A_n$ are $A_1=\int_{-\infty}^{+\infty}\lambda_1 dE^{(1)}_{\lambda_1}$, $\cdots A_n=\int_{-\infty}^{+\infty}\lambda_n dE^{(n)}_{\lambda_n}$, then we write for a vector $\Psi\in\mathfrak{H}$ representing a pure state

$$P_\Psi^{A_1;\cdots;A_n}(B_1\times\cdots\times B_n)=\langle\Psi|\,E^{(1)}_{B_1}\cdots E^{(n)}_{B_n}\,|\Psi\rangle. \tag{2.7}$$

In general, a physical state is represented in the conventional formalism by a positive-definite bounded operator U, with $\mathrm{Tr}\,U=1$, called the statistical operator or density operator. In that case we have the generalization of (2.7):

$$P_U^{A_1;\cdots;A_n}(B_1\times\cdots\times B_n)=\mathrm{Tr}\,[UE^{(1)}_{B_1}\cdots E^{(n)}_{B_n}]. \tag{2.8}$$

It is obvious that the so defined set functions are finite and countably additive, and hence they uniquely determine[28] a complex probability measure. It is easy to establish that the conditions of Axiom I are satisfied; e.g.,

$$P_U^{A_1;A_2;\cdots;A_n}(\mathbf{R}^1\times B_2\times\cdots\times B_n)=P_U^{A_2;\cdots A_n}(B_2\times\cdots\times B_n)$$

because $E^{(1)}_{\mathbf{R}^1}=\mathbf{1}$. Similarly, Ax. I (2) is valid because the projections belong to the spectral decomposition of a self-adjoint operator.

To see that Axiom I (3) is satisfied, one has to take into consideration that compatibility between observables is expressed in the Hilbert-space formalism by commutativity between the corresponding operators. A necessary and sufficient condition for bounded operators to commute is that their spectral projections commute[29]; for unbounded operators, their commutativity can be defined only in terms of the commutativity of the corresponding spectral functions.

Axiom I(4) is obviously true because P^A_U is a

[28] Ref. 21, Theorem *A*, p. 22 and Thoerem *A*, p. 54.
[29] Ref. 25, p. 248.

positive probability measure for each self-adjoint operator A.

It is obvious that, if Ψ_1 and Ψ_2 differ only in a phase factor and therefore represent the same pure state, then for each $(A_1, \cdots A_n) \in \hat{\mathfrak{O}}$,

$$P_{\Psi_1}^{A_1;\cdots;A_n}(B_1 \times \cdots \times B_n) = P_{\Psi_2}^{A_1;\cdots;A_n}(B_1 \times \cdots \times B_n), \quad B_1, \cdots B_n \in \mathfrak{B}^1,$$

and hence $P_{\Psi_1} = P_{\Psi_2}$. However, it is not obvious that the above relations (valid for any $\hat{A} \in \hat{\mathfrak{O}}$) are sufficient to guarantee the identity of the two rays in $\mathfrak{H}$ satisfying them, or that the relations (2.8) will uniquely define a statistical operator U. This will be demonstrated in the course of showing that the language introduced by our axioms is equivalent, with the addition of some hypotheses, to the Hilbert-space formalism. At this stage, we only conclude that the identity of physical states defined in the Hilbert-space formalism fulfills the requirement of Axiom II.

Axiom III is certainly fulfilled, because the spectral decomposition of a self-adjoint operator is unique, and hence, if $A_1 = A_2$, then $E_B^{(1)} = E_B^{(2)}$ for any $B \in \mathfrak{B}^1$.[30] It can be readily seen that Axiom III also guarantees the identity of two self-adjoint operators satisfying it. Assume first that we deal with a case without superselection rules. Then, if we deal with a complete set $\mathfrak{O}^c = \{A, B, \cdots\}$ of commuting observables with discrete spectrum,

$$P_{\Psi}^{A;B;\cdots;A_1;A;B;\cdots} \times (\{a'\} \times \{b'\} \times \cdots \times B \times \{a\} \times \{b\} \times \cdots) = P_{\Psi}^{A;B;\cdots;A_2;A;B;\cdots} \times (\{a'\} \times \{b'\} \times \cdots \times B \times \{a\} \times \{b\} \times \cdots)$$

can be written as

$$\langle\Psi| E_{\{a'\}}^{A} E_{\{b'\}}^{B} \cdots E_B^{A_1} E_{\{a\}}^{A} E_{\{b\}}^{B} \cdots |\Psi\rangle = \langle\Psi| E_{\{a'\}}^{A} E_{\{b'\}}^{B} \cdots E_B^{A_2} E_{\{a\}}^{A} E_{\{b\}}^{B} \cdots |\Psi\rangle. \tag{2.9}$$

If we choose $|\Psi\rangle = 2^{-\frac{1}{2}}(|a, b, \cdots\rangle + |a', b' \cdots\rangle)$ in the case when $(a, b, \cdots) \neq (a', b', \cdots)$, then (2.9) becomes

$$\langle ab \cdots| E_B^{A_1} |a'b' \cdots\rangle = \langle ab \cdots| E_B^{A_2} |a'b' \cdots\rangle. \tag{2.10}$$

Such a choice of $|\Psi\rangle$ is permissible if we assume that each $|\Psi\rangle \in \mathfrak{H}$ represents a physical state, i.e., that we do not deal with superselection rules. From (2.9) it follows that (2.10) is true even when $(ab \cdots) = (a'b' \cdots)$, and hence we can conclude that $E_B^{A_1} = E_B^{A_2}$, $B \in \mathfrak{B}^1$, or $A_1 = A_2$.

In case of a theory with superselection rules, we can carry out this reasoning for each coherent subspace $\mathfrak{H}_i$ in which $\mathfrak{H}$ splits because of the superselection rules. As each self-adjoint operator A representing an observable has to leave each coherent subspace invariant, the self-adjoint operator A_i induced in $\mathfrak{H}_i$ by A is determined in this way.

Axiom IV can always be introduced in defining $\mathfrak{O}$ in terms of $\mathfrak{O}'$, because the Borel function of a self-adjoint operator is both a well-defined concept and a self-adjoint operator whose spectral function satisfies (2.5).

Axiom V is not peculiar of the Hilbert-space formalism. However, it is characteristic of all present quantum theories that $\mathfrak{O}$ is constructed out of a set $\mathfrak{O}_0'$ of fundamental observables from which it is very easy to pick a fundamental set $\mathfrak{O}'$, i.e., a subset of $\mathfrak{O}_0'$ containing only Borel-independent observables.[31] All the other observables are then defined as Borel functions of the observables in $\mathfrak{O}'$.

To show that Axiom VI is satisfied by the Hilbert-space formalism, we need a somewhat longer argument.

In the Hilbert-space formalism, the general form of a physical state (pure state or mixture) is the following: Given any complete orthonormal basis $|\Psi_i\rangle$, $i \in I$ in $\mathfrak{H}$ (I contains an enumerable number of elements because in the conventional formalism $\mathfrak{H}$ is supposed to be separable), any bounded operator of the form $\sum_{i\in I} |\Psi_i\rangle t_i \langle\Psi_i|$ with $\sum_{i\in I} t_i = 1$ represents a physical state, where $|\Psi_i\rangle\langle\Psi_i|$ denotes the projection operator on the one-dimensional space determined by $|\Psi_i\rangle$. It is easy to see that the statistical operator $\sum_{i\in I} |\Psi_i\rangle t_i \langle\Psi_i|$ is positive-definite and that its trace is equal to one. Hilbert has shown that any self-adjoint positive-definite operator U with a finite trace (and therefore bounded) has a discrete spectrum. The detailed proof[32] shows that if the operator is positive definite and its trace is equal to one, then its spectrum is of the form $t_1 > t_2 > t_3 > \cdots$, where $0 \leq t_k \leq 1$, $k \in K$.[33]

[30] In the case of unbounded operators, the identity of two self-adjoint operators representing observables should be defined in the rigorous mathematical sense and should imply equality of domains of definition as well as equality in the taken values.

[31] We can clarify the difference between $\mathfrak{O}_0'$ and an $\mathfrak{O}'$ on the example of a free electron: $\mathfrak{O}_0'$ can consist of *all* components of the momentum, all space coordinates, etc., while $\mathfrak{O}'$ contains only the components of the momentum along a certain set of axes, the position coordinates along *that* set of axes, etc.

[32] J. von Neumann, *Mathematical Foundations of Quantum Mechanics*, translation by R. T. Beyer (Princeton University Press, Princeton, 1955), p. 189.

[33] K contains a finite or denumerable number of elements.

Hence such an operator can be written in the form $U = \sum_{k\in K} t_k E_k$, where E_k is the projector corresponding to the eigenvalue t_k.[34]

In general, we are dealing with superselection rules which, in the present quantum mechanical theories, require a decomposition of the Hilbert space $\mathfrak{H}$ into a direct sum

$$\mathfrak{H} = \bigoplus_{j\in J} \mathfrak{H}_j, \tag{2.11}$$

where J contains, because of the separability of $\mathfrak{H}$, only a denumerable number of elements. Only the vectors belonging to a coherent subspace $\mathfrak{H}_j$ represent a physical state.

If we impose on the previously considered positive-definite operator U, with $\mathrm{Tr}\ U = 1$, the additional condition that it has to commute with each of the projectors $E(\mathfrak{H}_j)$, $j \in J$, on the coherent spaces $\mathfrak{H}_j$, then we can write U in the form

$$U = \sum_{j\in J} \sum_{k\in K} t_k E(\mathfrak{H}_j) E_k. \tag{2.12}$$

As $[U, E(\mathfrak{H}_j)]_- = 0$ we have also $[E_k, E(\mathfrak{H}_j)]_- = 0$ ($j \in J$, $k \in K$) and therefore each $E(\mathfrak{H}_j)E_k$ is a projection operator. In each of the subspaces $\mathfrak{H}_{jk}$ corresponding to $E(\mathfrak{H}_j)E_k$, we can choose a complete orthornormal basis $|\Psi_{jkl}\rangle$, $l \in L_{jk}$, each $|\Psi_{jkl}\rangle$ representing now a pure physical state. We can now write (2.12) in the form

$$U = \sum_{j\in J} \sum_{k\in K} \sum_{l\in L_{jk}} |\Psi_{jkl}\rangle\, t_{jkl}\, \langle\Psi_{jkl}|, \tag{2.13}$$

where $t_{jkl} = t_k$, $j \in J$, $l \in L_{jk}$. Besides

$$\sum_{j,k,l} t_{jkl} = \mathrm{Tr}\ U = 1.$$

Hence we have proved the following proposition.

Proposition: Each positive-definite operator of trace equal to 1 which leaves invariant each coherent space of the decomposition (2.11), i.e., commutes with each $E(\mathfrak{H}_j)$, $j \in J$, represents a physical state in the conventional Hilbert-space quantum mechanical theories, and therefore can be called a statistical operator.

This proposition shows that if U_1 and U_2 are statistical operators, then $tU_1 + (1 - t)U_2$, $0 \leq t \leq 1$ is a statistical operator, too. It can be easily checked, by using formula (2.8), that if U_1 and U_2 determine the complex probability measures $P_1^{\hat{\alpha}}$ and $P_2^{\hat{\alpha}}$, respectively, then $tU_1 + (1 - t)U_2$ determines the complex probability measure $tP_1^{\hat{\alpha}} + (1 - t)P_2^{\hat{\alpha}}$. Hence Axiom VI(1) is obeyed by the Hilbert-space formalism. In a similar manner, we can decide about the second part of Axiom VI.

In order to establish that Axiom VII is a true statement in the case of Hilbert-space formalism, consider first the case without superselection rules. Then the set $\mathcal{S}$ of all physical states can be identified with the set of all normed positive linear functionals on the C^*-algebra $\mathfrak{B}(\mathfrak{H})$ of all bounded operators on the given Hilbert space $\mathfrak{H}$.[35] This set is closed in the weak topology in $\mathcal{S}$ because the weak limit of any sequence of normalized linear positive functionals is again such a functional, and therefore, with the help of a theorem of Gleason,[35] we can easily conclude that it uniquely determines a density operator, i.e., a physical state. The more general case with superselection rules can be reduced to the previous case in each coherent subspace $\mathfrak{H}_j(\mathfrak{H} = \bigoplus_j \mathfrak{H}_j)$.

The fact that the Hilbert-space formalism is in accordance with Axioms VIII and IX is too obvious to be considered in any detail.

3. MISCELLANEOUS THEOREMS DIRECTLY DEDUCIBLE FROM THE AXIOMS.

3.1. The set of questions.

One of the basic concepts in some other axiomatic formulations[36] of quantum mechanics is the concept of question.

Definition: A *question* is an observable with a pure point spectrum concentrated in two points: 0 and 1.

The first obvious consequence of this definition is the following proposition.

Proposition 1: If q is a question, then

$$P^{\hat{\alpha};q;\hat{\beta}}(B' \times \{0\} \times B'') + P^{\hat{\alpha};q;\hat{\beta}}(B' \times \{1\} \times B'') = P^{\hat{\alpha};\hat{\beta}}(B' \times B'')$$

for any $\hat{\alpha}, \hat{\beta} \in \hat{\mathfrak{O}}$, $B' \in \mathfrak{B}^{\hat{\alpha}}$, $B'' \in \mathfrak{B}^{\hat{\beta}}$. In particular,

$$P^q(\{0\}) + P^q(\{1\}) = 1.$$

This proposition is an evident consequence of Axiom I(1).

Proposition 2: An observable α is a question if and only if $\alpha = \alpha^2$, i.e., the set of questions coincides with the set of indempotent elements of $\mathfrak{O}$.

---

[34] Note that $E_{k_1} \perp E_{k_2}$ for $k_1 \neq k_2$.

[35] We can establish this by using, for example, a theorem of A. M. Gleason, J. Math. Mech. 6, 885 (1953)—at least when we are dealing with a separable Hilbert-space $\mathfrak{H}$. To apply such a theorem, we have to take into consideration that each normed positive functional $f(A)$ on $\mathfrak{B}(\mathfrak{H})$ is bounded (Ref. 25, Sec. 10.4, Proposition I) and therefore continuous.

[36] See Refs. 7, 9, 10, 11.

Proof: If α is a question, then the spectrum of α^2 is obviously concentrated at 0 and 1, because, if $0, 1 \notin B$ for some $B \in \mathfrak{B}^1$, then $0, 1 \notin f^{-1}(B)$, where $f(\lambda) = \lambda^2$. Hence α^2 is a question too.

Assume that $\alpha = \alpha^2$. Then, by the definition of $f(\alpha) = \alpha^2$, for any open interval $I = (1, \lambda_0)$, $\lambda_0 > 1$, we have

$$P^{\alpha^2}(I) = P^{\alpha}(I' \cup I'') = P^{\alpha}(I') + P^{\alpha}(I''), \qquad (3.1)$$

where $I' = (1, \lambda^{\frac{1}{2}})$, $I'' = (-\lambda^{\frac{1}{2}}, -1)$. But, as $\alpha = \alpha^2$, we have, on the basis of Axiom II, that $P^{\alpha}(I) = P^{\alpha^2}(I)$ and therefore

$$P^{\alpha}((-\infty, 0)) = P^{\alpha^2}((-\infty, 0)) = 0$$

because α^2 has a positive spectrum. Hence, $P^{\alpha}(I'') = 0$ and (3.1) becomes

$$0 = P^{\alpha}(I) - P^{\alpha}(I') = P^{\alpha}(I - I').$$

This means that $P^{\alpha}((\lambda_0^{\frac{1}{2}}, \lambda_0)) = 0$ for any $\lambda_0 > 1$ and therefore $P^{\alpha}((1, +\infty)) = P^{\alpha}(\bigcup_{n=-\infty}^{+\infty} (2^{1/n}, 2^{1/n+1}]) = \sum_{n=-\infty}^{+\infty} P^{\alpha}((2^{1/n}, 2^{1/n+1}]) = 0$.

In a similar manner, we conclude that, for any $0 < \lambda < 1$, $P^{\alpha}((0, \lambda^{\frac{1}{2}})) = P^{\alpha}((-\lambda^{\frac{1}{2}}, 0)) + P^{\alpha}((0, +\lambda^{\frac{1}{2}})) = P^{\alpha^2}((0, \lambda)) = P^{\alpha}((0, \lambda))$, and hence $P^{\alpha}((\lambda, \lambda^{\frac{1}{2}})) = 0$. As this is true for any $0 < \lambda < 1$ we get $P^{\alpha}((0, 1)) = 0$ because $(0, 1)$ can be written in the form of a countable union of intervals $(\lambda, \lambda^{\frac{1}{2}})$. Thus, we have proved that $P^{\alpha}((-\infty, 0)) = P^{\alpha}((0, 1)) = P^{\alpha}((1, +\infty)) = 0$. Q.E.D.

Many axiomatic systems[37] are based on the possibility of introducing a partial ordering in the set $\mathfrak{Q}$ of all questions. This partial ordering in $\mathfrak{Q}$ is always achieved[38] by defining $q_1 \leq q_2$ if and only if $P^{q_1}(\{1\}) \leq P^{q_2}(\{1\})$ for all $P \in \mathfrak{S}$. However, in our case the axioms of Sec. 2.2 are not sufficient to guarantee that the equality of two questions introduced by such a partial ordering (i.e., $q_1 = q_2$ if $q_1 \leq q_2$ and $q_2 \leq q_1$) coincides with the definition of identity of two observables introduced by Axiom II. We can easily check this on an example.

We can build a very simple language obeying Axioms I–IX by starting with a fundamental set $\mathfrak{O}^f$ consisting of only two questions, α and β. A physical state P is determined by *any* given complex probability measure $P^{\alpha;\beta}$ so defined at the points $(0, 0)$, $(1, 0)$, $(0, 1)$, and $(1, 1)$ of the spectrum $\mathbf{S}^{(\alpha;\beta)}$ that the following conditions are fulfilled:

$$P^{\alpha;\beta}(\{1\} \times \{0\}) = P^{\alpha;\beta}(\{0\} \times \{1\}), \qquad (3.2)$$

$$P^{\alpha;\alpha}(\{0\} \times \{1\}) + 2P^{\alpha;\beta}(\{0\} \times \{1\}) + P^{\alpha;\beta}(\{1\} \times \{1\}) = 1.$$

[37] See, e. g., Refs. 9, 10.
[38] Ref. 9, p. 64.

Having determined $\mathfrak{S}$ this way, we can define $\mathfrak{O}$ by taking all functions of α and β. It is trivial to check that all Axioms I–IX are satisfied by this formalism. On the other hand, $P^{\alpha}(\{1\}) = P^{\beta}(\{1\})$ but obviously $\alpha \neq \beta$.

The conclusion is that, at this stage, the Axioms I–IX allow much more general structures than the axioms of the approaches which make use of quantum logics. Nevertheless, as mentioned in the introduction, simple physical interpretations can be assigned even to such general structures.

According to Axiom V, given a fundamental set $\mathfrak{O}^f$, we can write each question q as a function of a compatible finite number of observables belonging to $\mathfrak{O}^f$. Of course, though not possible in general, we can introduce in each set $\mathfrak{Q}(\hat{\alpha})$ of questions which can be written as functions of the same n-tuple $\hat{\alpha}$ of *compatible* observables from $\mathfrak{O}^f$ all the fundamental concepts employed in the quantum logics. Thus, we introduce in each $\mathfrak{Q}(\hat{\alpha})$ a partial ordering by writing $q_1 \leq q_2$ for some $q_1, q_2 \in \mathfrak{Q}(\hat{\alpha})$, if and only if $P^{q_1}(\{1\}) \leq P^{q_2}(\{1\})$ for all $P \in \mathfrak{S}$. It is straightforward to check that we are dealing with a partial ordering indeed.

We can define an operation of involution $q \to q'$ in each $\mathfrak{Q}(\hat{\alpha})$ by defining q' as that question of $\mathfrak{Q}(\hat{\alpha})$ for which $P^{q'}(\{1\}) = 1 - P^{q}(\{1\})$ for all $P \in \mathfrak{S}$. It will soon be clear that q' really always exists. Further, we will say that two questions $q_1, q_2 \in \mathfrak{Q}(\hat{\alpha})$ are orthogonal, writing $q_1 \perp q_2$, if and only if $P^{q_1}(\{1\}) + P^{q_2}(\{1\}) \leq 1$ for all $P \in \mathfrak{S}$.

We will now deduce a few easy results concerning questions which will be found useful in later developments. They are mostly a weaker counterpart of assumptions made in other axiomatic approaches.[9,10]

Given compatible observables $\alpha_1, \cdots \alpha_n$, it is clear that the characteristic function $\chi_B^{\alpha_1,\cdots\alpha_n}(\hat{\lambda})$, $\hat{\lambda} \in \mathbf{R}^{(\alpha_1,\cdots\alpha_n)}$, of a Borel set $B \in \mathfrak{B}^{(\alpha_1,\cdots\alpha_n)}$ defines a question $q = \chi_B^{\alpha_1,\cdots\alpha_n}(\hat{\alpha})$, $\hat{\alpha} = (\alpha_1, \cdots \alpha_n)$. Vice versa, according to Axiom *V*, each question q has to be a function $\chi(\hat{\alpha})$ of a finite number of compatible observables $\hat{\alpha}$ belonging to a fundamental set of observables $\mathfrak{O}^f$. As the spectrum of a question consists only of 0 and 1, this function $\chi(\hat{\lambda})$, $\hat{\lambda} \in \mathbf{R}^{\hat{\alpha}}$ can take on only the values 0 and 1 and therefore is a characteristic function of a set $B \subset \mathbf{R}^{\hat{\alpha}}$; as χ, by definition, is a Borel function, B has to be a Borel set. In general, if the spectrum of $\hat{\alpha}$ does not coincide with $\mathbf{R}^{\hat{\alpha}}$, there will be many Borel sets in $\mathbf{R}^{\hat{\alpha}}$ corresponding to q.

Proposition 3: To each question q and each fundamental set $\mathfrak{O}^f$ of observables corresponds at least one

n-tuple $\hat{\alpha}$ of compatible observables from $\mathfrak{O}'$, and one Borel set B in $\mathbf{R}^{\hat{\alpha}}$, which are such that $q = \chi_B^{\hat{\alpha}}(\hat{\alpha})$, where $\chi_B^{\hat{\alpha}}(\hat{\lambda})$, $\hat{\lambda} \in \mathbf{R}^{\hat{\alpha}}$ is the characteristic function of B. If $q = \chi^{\hat{\alpha}}(\hat{\alpha})$, $\hat{\alpha} \in \hat{\mathfrak{O}}$, there is a Borel set B_q in $\mathbf{R}^{\hat{\alpha}}$ which corresponds to q, i.e., $q = \chi_B^{\hat{\alpha}}(\hat{\alpha})$, and which is contained in any other Borel set from $\mathbf{R}^{\hat{\alpha}}$ which corresponds to q; B_q will be called the *minimal Borel set corresponding to* q.

The first part of the proposition has been elaborated earlier. We proceed with the proof of the existence of B_q.

Proof: Introduce a partial ordering in the family $\mathfrak{F}_q^{\hat{\alpha}}$ of Borel sets corresponding to q by the relation of set inclusion. We will establish that the conditions of Zorn's lemma are fulfilled by showing that for any linearly ordered sets $B_1 \supset B_2 \supset \cdots$ from $\mathfrak{F}_q^{\hat{\alpha}}$, the set $B = \bigcap_{k=1}^{\infty} B_k$ belongs again to $\mathfrak{F}_q^{\hat{\alpha}}$. To this purpose, we have to demonstrate that

$$P^{\hat{\beta}';\hat{\alpha};\hat{\beta}''}(B' \times B \times B'') = P^{\hat{\beta}';q;\hat{\beta}''}(B' \times \{1\} \times B'')$$

for any $P \in \mathcal{S}$, $\hat{\beta}'$, $\hat{\beta}'' \in \hat{\mathfrak{O}}$, and $B' \in \mathfrak{B}^{\hat{\beta}'}$, $B'' \in \mathfrak{B}^{\hat{\beta}''}$. But this is a consequence of continuity from above[39] of the complex probability measures $P^{\hat{\beta}';\hat{\alpha};\hat{\beta}''}$ because $P^{\hat{\beta}';q;\hat{\beta}''}(B' \times \{1\} \times B'') = P^{\hat{\beta}';\hat{\alpha};\hat{\beta}''}(B' \times B_1 \times B'') = P^{\hat{\beta}';\hat{\alpha};\hat{\beta}''}(B' \times B_2 \times B'') = \cdots$, and

$$B' \times B \times B'' = \lim_{k\to\infty} B' \times B_k \times B'',$$

where $B' \times B_k \times B''$ is a monotonously decreasing sequence.

The conditions of Zorn's lemma are therefore satisfied because B is the greatest lower bound of $B_1, B_2, \cdots \in \mathfrak{F}_q^{\hat{\alpha}}$. Hence, there exists a greatest lower bound B_q of the set $\mathfrak{F}_q^{\hat{\alpha}}$. We will show that this greatest lower bound is unique and therefore satisfies the conditions stated in the proposition.

To show that B_q is unique, assume that there is another greatest lower bound $B_q^1 \neq B_q$. In that case, obviously the relations $B_q^1 \subset B_q$ or $B_q \subset B_q^1$ are not possible. Then $B_q^1 - B_q$ and $B_q - B_q^1$ are not empty. From

$$P^{\hat{\alpha}}(B_q) = P^q(\{1\}) = P^{\hat{\alpha}}(B_q^1),$$

we obtain

$$P^{\hat{\alpha}}(B_q - B_q^1) = P^{\hat{\alpha}}(B_q^1 - B_q), \qquad P \in \mathcal{S}. \tag{3.3}$$

Therefore, either both sets $B_q - B_q^1$ and $B_q^1 - B_q$ are zero sets or both are nonzero sets. Because of Axiom IX, the second possibility is excluded. Namely, if $B_q - B_q^1$ is a nonzero set, then there is a $P_1 \in \mathcal{S}$ such that

$$P_1^{\hat{\alpha}}(B_q - B_q^1) = 1. \tag{3.4}$$

For this P_1, (3.3) is not true, because if it were, we could conclude from $(B_q - B_q^1) \cap (B_q^1 - B_q) = \varnothing$ and from (3.4) that

$$P_1^{\hat{\alpha}}((B_q - B_q) \cup (B_q^1 - B_q)) = 2P_1^{\hat{\alpha}}(B_q - B_q^1) = 2,$$

which is impossible.

Thus we showed that $B_q - B_q^1$ is a zero set. This means that $B_q - B_q^1$ is empty; namely Axiom VIII requires[40] $B' \times (B_q - B_q^1) \times B''$ to be a zero set too for any $B' \in \mathfrak{B}^{\hat{\beta}'}$, $B'' \in \mathfrak{B}^{\hat{\beta}''}$ and therefore

$$\begin{aligned} P^{\hat{\beta}';\hat{\alpha};\hat{\beta}''}(B' \times (B_q \cap B_q^1) \times B'') &= P^{\hat{\beta}';\hat{\alpha};\hat{\beta}''}(B' \times B_q \times B'') \\ &= P^{\hat{\beta}';q;\hat{\beta}''}(B' \times \{1\} \times B''). \end{aligned} \tag{3.5}$$

Relation (3.5) shows that $B_q \cap B_q^1$ belongs to the question q too; because B_q is the greatest lower bound in $\mathfrak{F}_q^{\hat{\alpha}}$, this is possible only if $B_q - B_q^1 = \varnothing$. In a similar manner we can show that $B_q^1 - B_q = \varnothing$ and therefore $B_q = B_q^1$. Q.E.D.

We have to remark, concerning the proof of Proposition 4, that the use of Zorn's lemma is unavoidable. If we would consider without further arguing that $\bigcap_{B\in\mathfrak{F}_q \exp\hat{\alpha}} B$ represents our B_q, there would be no other way to show that this is a Borel set at all and that it corresponds to q.

We also note that Axioms VIII and IX have played an essential role in proving this proposition.

Proposition 4: If $q_1 \perp q_2$, $q_1, q_2 \in Q(\hat{\alpha})$, where $\hat{\alpha} \in \hat{\mathfrak{O}}$ and $\{\hat{\alpha}\} = C$, then $B_{q_1} \cap B_{q_2} = \varnothing$.

Proof: Assume that $B_{q_1} \cap B_{q_2} \neq \varnothing$ and that it is not a zero set. Then, according to Axiom VIII, there is a $P_1 \in \mathcal{S}$ such that $P_1^{\hat{\alpha}}(B_{q_1} \cap B_{q_2}) = 1$ and, therefore,

$$\begin{aligned} P_1^{q_1}(\{1\}) + P_1^{q_2}(\{1\}) &= P_1^{\hat{\alpha}}(B_{q_1}) + P_1^{\hat{\alpha}}(B_{q_2}) \geq 2P_1^{\alpha}(B_{q_1} \cap B_{q_2}) = 2. \end{aligned}$$

This contradicts $q_1 \perp q_2$. Hence $B_{q_1} \cap B_{q_2}$ is a zero set. This implies that $B_{q_1} \cap B_{q_2}$ is empty; otherwise we could conclude, by the same reasonings as in (3.5), that $B_{q_1} - (B_{q_1} \cap B_{q_2})$ belongs to q_1 and $B_{q_2} - (B_{q_1} \cap B_{q_2})$ to q_2. If $B_{q_1} \cap B_{q_2} \neq \varnothing$ this would mean that B_{q_1} and B_{q_2} are not the minimal Borel sets belonging to their respective questions. Q.E.D.

[39] Ref. 21, p. 39.

[40] Strictly speaking, Ax. VIII requires that $B' \times (B_q - B_q{}^1) \times B''$ is of the form $B_1 \times \ldots \times B_n$ and similarly for $B_q - B_q{}^1$; but this is sufficient to guarantee the validity of our conclusions because $\mathfrak{B}^1 \times \ldots \times \mathfrak{B}^1$ (n-times) generates the Boolean σ-algebra $\mathfrak{B}^n$.

We can now easily prove a weaker form of Mackey's fifth axiom (Ref. 9, p. 64). To that purpose we define the orthogonal sum $q_1 \bigoplus q_2 \bigoplus \cdots$ of an at most denumerable number of mutually orthogonal questions $q_1, q_2, \cdots$ as that question q for which

$$P^q(\{1\}) = \sum_k P^{q_k}(\{1\}).$$

Proposition 5: If $q_1, q_2, q_3, \cdots$ is a set of mutually orthogonal questions which can be considered to be functions of the same set $\{\hat{\alpha}\}$, $\hat{\alpha} \in \hat{\mathfrak{O}}$ of compatible observables, then their orthogonal sum $q = q_1 \bigoplus q_2 \bigoplus \cdots$ exists.

Proof: Denote by q the question $\chi_B^{\hat{\mathfrak{o}}}(\hat{\alpha})$ belonging to $B = B_{q_1} \cup B_{q_2} \cup \cdots$. As on the basis of Proposition 4, $B_{q_i} \cap B_{q_j} = \varnothing$ for $i \neq j$, we have

$$P^{\hat{\beta}';q;\hat{\beta}''}(B' \times \{1\} \times B'') = P^{\hat{\beta}';\hat{\alpha};\hat{\beta}''}(B' \times B \times B'') = \sum_{k=1}^{\infty} P^{\hat{\beta}';\hat{\alpha};\hat{\beta}'''}(B' \times B_k \times B'' = \sum_{k=1}^{\infty} P^{\hat{\beta}';q_k;\hat{\beta}''}(B' \times \{1\} \times B'').$$

This means that $q = q_1 \bigoplus q_2 \bigoplus \cdots$. Q.E.D.

3.2. The algebras of compatible observables.

We say for a subset $\mathfrak{C}$ of $\mathfrak{O}$ that it is a *set of compatible observables* if any finite number of elements of $\mathfrak{C}$ are compatible. We will denote by $\mathfrak{F}_c$ the family of all sets of compatible observables.

Proposition 1: Each observable belongs to a maximal set of compatible observables; therefore $\mathfrak{O} = \bigcup_{\mathfrak{C} \in \mathfrak{M}_c} \mathfrak{C}$.

Proof: The relation of set inclusion introduces a partial ordering in the set $\mathfrak{F}_c^\alpha$ of all sets of compatible observables which contain the observable α. If $\mathfrak{C}_1 \subset \mathfrak{C}_2 \subset \cdots$ is a linearly ordered set, then $\mathfrak{C} = \bigcup_{k=1}^{\infty} \mathfrak{C}_k$ belongs to $\mathfrak{F}_c^\alpha$. Namely, if $\alpha_1, \cdots \alpha_n \in \mathfrak{C}$, then, for each $r = 1, \cdots n$, there must be a $\mathfrak{C}_{k_r}$ containing α_r. Then $\alpha_1, \cdots \alpha_r \in \mathfrak{C}_k$ if $k \geq \max\{k_1, \cdots k_n\}$ and, therefore, $\{\alpha_1, \cdots \alpha_n\} = C$. As this is true for any $\alpha_1, \cdots \alpha_n \in \mathfrak{C}$ and any $n = 1, 2, \cdots$, and as $\alpha \in \mathfrak{C}$, it follows that $\mathfrak{C} \in \mathfrak{F}_c^\alpha$. Hence, any linearly ordered family of sets in $\mathfrak{F}_c^\alpha$ has a least upper bound. Thus, according to Zorn's lemma, there is at least one maximal element $\mathfrak{C}^\alpha$ of $\mathfrak{F}_c^\alpha$. $\mathfrak{C}^\alpha$ is also a maximal element of $\mathfrak{F}_c$, because if $\mathfrak{C}_1 \supset \mathfrak{C}^\alpha$, $\mathfrak{C}_1 \in \mathfrak{F}_c$, then $\alpha \in \mathfrak{C}_1$ and therefore $\mathfrak{C}_1 \in \mathfrak{F}_c^\alpha$; thus $\mathfrak{C}_1 = \mathfrak{C}^\alpha$. Q.E.D.

Proposition 2: Each maximal set of compatible observables becomes a commutative real algebra if the operations of addition, multiplication, and scalar multiplication are defined by the respective Borel functions $\alpha + \beta$, $\alpha\beta$, $c\alpha$ (c is a real constant) of $\alpha, \beta \in \mathfrak{C}$.

Proof: The functions $f(\lambda, \mu) = \lambda + \mu$ and $g(\lambda, \mu) = \lambda\mu$, $\lambda, \mu \in \mathbf{R}^1$, are Borel functions on $\mathbf{R}^2$ while $h(\lambda) = c\lambda$, $\lambda \in \mathbf{R}^1$, is a Borel function on $\mathbf{R}^1$ (c is a real constant). Therefore, $f(\alpha, \beta) = \alpha + \beta$, $g(\alpha, \beta) = \alpha\beta$, and $h(\beta) = c\alpha$ are well-defined observables for any $\alpha, \beta \in \mathfrak{C}$ if $\mathfrak{C} \in \mathfrak{F}_c$. If $\mathfrak{C} \in \mathfrak{M}_c$, then $\alpha + \beta$, $\alpha\beta$, $c\alpha \in \mathfrak{C}$ for $\alpha, \beta \in \mathfrak{C}$, because $\{\alpha + \beta, \alpha_1, \cdots \alpha_n\} = C$ if $\{\alpha, \alpha_1, \cdots \alpha_n\} = C$ and $\{\beta, \alpha_1, \cdots \alpha_n\} = C$, and similar statements are true for $\alpha\beta$ and $c\alpha$. It is obvious that the operations of addition and multiplication are commutative, because we have for the scalar functions $\lambda + \mu$ and $\lambda\mu$ that $\lambda + \mu = \mu + \lambda$ and $\lambda\,\mu = \mu\lambda$. If we choose the two questions q^0 and q^1 defined by

$$P^{\hat{\beta}';q^0;\hat{\beta}''}(B' \times \{1\} \times B'') = 0,$$

$$P^{\hat{\beta}';q^1;\hat{\beta}''}(B' \times \{1\} \times B'') = P^{\hat{\beta}';\hat{\beta}''}(B' \times B''),$$

$$\hat{\beta}', \hat{\beta}'' \in \hat{\mathfrak{O}}, \qquad B' \in \mathfrak{B}^{\hat{\beta}'}, \qquad B'' \in \mathfrak{B}^{\hat{\beta}''},$$

to represent the zero and unity of this algebra,[41] it is easy to check that all the axioms of a commutative algebra are fulfilled. Q.E.D.

We will now consider the set $\mathfrak{O}_b$ of all bounded observables. An observable is called *bounded* if its spectrum is a bounded set in $\mathbf{R}^1$. The *bound* $||\alpha||$ of a bounded observable α is the least upper bound of the absolute values of the elements of $\mathbf{R}^1$ belonging to the spectrum of α, i.e., $||\alpha|| = \sup_{\lambda \in \mathfrak{s}^{(\alpha)}} |\lambda|$.

We will denote by $\mathfrak{F}_c^b$ the family of all sets of compatible bounded observables, and by $\mathfrak{M}_c^b$ the family of all maximal sets of compatible bounded observables; it should be understood that a maximal set of compatible bounded observables is maximal in $\mathfrak{F}_c^b$.

In a completely analogous manner as for observables in general, we can prove the replica of Proposition 1 for bounded observables.

Proposition 3: Each bounded observable belongs to a maximal set of compatible bounded observables; therefore $\mathfrak{O}_b = \bigcup_{\mathfrak{C} \in \mathfrak{M}_c^b} \mathfrak{C}$.

---

[41] Such questions certainly exist in $\mathfrak{O}$. Namely $q^0 = \chi_\phi{}^\alpha(\alpha)$ and $q^1 = \chi_{\mathbf{R}^1}{}^\alpha(\alpha)$ for any $\alpha \in \mathfrak{O}$, where $\chi_\phi{}^\alpha(\lambda) \equiv 0$, $\chi_{\mathbf{R}^1}{}^\alpha(\lambda) \equiv 1$, $\lambda \in \mathbf{R}^1$ are the characteristic functions of the empty set and of the real line.

Proposition 4: Each maximal set of compatible bounded observables $\mathfrak{C}$ becomes a *normed commutative* real algebra if the operations of addition, multiplication, and scalar multiplication are defined by the respective Borel functions $\alpha + \beta$, $\alpha\beta$, $c\alpha$ of $\alpha, \beta \in \mathfrak{C}$, and if we take for the norm of $\alpha \in \mathfrak{C}$ the bound $||\alpha||$.

Proof: Once we notice that the questions q^0 and q^1 defined by (3.1), are bounded observables (as any question is), the proof that $\mathfrak{C}$ is an algebra proceeds along the same line as the proof of Proposition 2. The only new feature is introduced by the norm $||\alpha||$ ($\alpha \in \mathfrak{C}$).

Obviously $||\alpha|| \geq 0$; if $||\alpha|| = 0$, then the spectrum of α is concentrated at 0 and therefore $\alpha = q^0$. It is also obvious that $||q^1|| = 1$.

The fact that $||c\alpha|| = |c|\,||\alpha||$ is easy to establish. The relations $||\alpha + \beta|| \leq ||\alpha|| + ||\beta||$ and $||\alpha\beta|| \leq ||\alpha||\,||\beta||$ are consequences of the fact that the relations $|\lambda + \mu| > ||\alpha|| + ||\beta||$ and $|\lambda\mu| > ||\alpha||\,||\beta||$ cannot be satisfied with values of $|\lambda| \leq ||\alpha||$ and $|\mu| \leq ||\beta||$. Namely, if $B \in \mathfrak{B}^{(\alpha+\beta)}$ does not contain any points ν, $|\nu| \leq ||\alpha|| + ||\beta||$, then the set $f^{-1}(B)$, $f(\lambda, \mu) = \lambda + \mu$, will not contain any points of $\mathbf{S}^\alpha \times \mathbf{S}^\beta$ and hence it will be a zero set; similarly for $\alpha\beta$. Q.E.D.

We have to stress that we do not claim that the sets $\mathfrak{C} \in \mathfrak{M}_c^b$ are Banach algebras. The question of completeness of $\mathfrak{O}_b$ in the given norm is of no interest to us for the time being.

3.3. The formulation of the axioms in terms of fundamental sets of observables.

In this subsection, we will reformulate some of our axioms in a logically equivalent way, with the help of the concept of a fundamental set of observables; in the following, $\mathfrak{O}^f$ will denote *any* fundamental set observables.

Proposition 1: Two physical states P_1, P_2 are identical, $P_1 = P_2$, if and only if

$$P_1^{\alpha_1,\cdots\alpha_n}(B) = P_2^{\alpha_1,\cdots\alpha_n}(B) \tag{3.6}$$

for any $\alpha_1, \cdots \alpha_n \in \mathfrak{O}^f$, $n = 1, 2, \cdots$, $B \in \mathfrak{B}^{(\alpha_1,\cdots\alpha_n)}$.

Proof: The necessity is obvious because of Axiom II and because $\mathfrak{O}^f \subset \mathfrak{O}$.

To establish the sufficiency, it has to be noticed that according to the definition of a fundamental set, we can write

$$\beta_1 = f_1(\alpha_1^{(1)}, \cdots \alpha_{k_1}^{(1)}), \cdots \beta_m = f_m(\alpha_1^{(m)}, \cdots \alpha_{k_m}^{(m)}),$$

$$\alpha_1^{(1)}, \cdots \alpha_{k_m}^{(m)} \in \mathfrak{O}^f$$

for any given $\beta_1, \cdots \beta_m \in \mathfrak{O}$. Therefore,

$$P_1^{\beta_1;\cdots;\beta_m}(B_1 \times \cdots \times B_m) = P_2^{\beta_1;\cdots;\beta_m}(B_1 \times \cdots \times B_m) \tag{3.7}$$

follows from (3.6), because (3.7) can be written in the form:

$$P_1^{\alpha_1(1),\cdots\alpha_{k_1}(1);\cdots;\alpha_1(m),\cdots\alpha_{k_m}(m)} \times (f_1^{-1}(B_1) \times \cdots \times f_m^{-1}(B_m)) = P_2^{\alpha_1(1),\cdots\alpha_{k_1}(1);\cdots;\alpha_1(m),\cdots\alpha_{k_m}(m)} \times (f_1^{-1}(B_1) \times \cdots \times f_m^{-1}(B_m)).$$

According to Proposition 2, Sec. 2.2, (3.7) implies $P_1 = P_2$. Q.E.D.

Using similar arguments we can prove easily the next two propositions.

Proposition 2: Two observables α, β are identical if and only if

$$P^{\wp_1';\cdots;\wp_k';\alpha;\wp_1'';\cdots;\wp_l''} \times (B_1' \times \cdots \times B_k' \times B \times B_1'' \times \cdots \times B_l') = P^{\wp_1'';\cdots;\wp_k';\beta;\wp_1'';\cdots;\wp_l''} \times (B_1' \times \cdots \times B_k' \times B \times B_1'' \times \cdots \times B_l')$$

for any $\wp_1', \cdots \wp_k', \wp_1'', \cdots \wp_l'' \in \mathfrak{O}_1^f$, $B, B_1', \cdots B_k', B_1'', \cdots B_l' \in \mathfrak{B}^1$, $k, l = 1, 2, 3, \cdots$.

Proposition 3: A sequence $P_1, P_2, \cdots$ of physical states is a Cauchy sequence if and only if

$$|P_m^\Delta(B) - P_n^\Delta(B)| < \epsilon \quad \text{for} \quad m, n > N_0(\epsilon, P)$$

and for any $\Delta \in \hat{\mathfrak{O}}^f$, $B \in \mathfrak{B}^\Delta$. (See Ref. 42.)

3.4. The relation of questions to other observables and to physical states.

For questions we can prove some statements which in forms resemble very much the propositions of the preceding subsection.

Proposition 1: Two physical states P_1, P_2 are identical if and only if

$$P_1^{q_1;\cdots;q_n}(\{1\} \times \cdots \times \{1\}) = P_2^{q_1;\cdots;q_n}(\{1\} \times \cdots \times \{1\}) \tag{3.8}$$

for any $q_1, \cdots q_n \in \mathfrak{Q}$, $n = 1, 2, \cdots$.

Proof: It is obvious that (3.8) is necessary for $P_1 = P_2$ to be true.

As we can write the relation

[42] $\hat{\mathfrak{O}}^f$ denotes the set of all $(\alpha_1, \cdots \alpha_n)$, $\alpha_1, \cdots \alpha_n \in \mathfrak{O}^f$, $n = 0, 1, 2, \cdots$.

$$P_1^{\alpha_1;\cdots;\alpha_n}(B_1 \times \cdots \times B_n) = P_2^{\alpha_1;\cdots;\alpha_n}(B_1 \times \cdots \times B_n)$$

in the form of (3.8) if we choose $q_1 = \chi_{B_1}^{\alpha_1}(\alpha_1)$, $\cdots$ $q_n = \chi_{B_n}^{\hat{a}_n}(\alpha_n)$. The sufficiency is proved on the base of Proposition 2 of Sec. 1.2. Q.E.D.

It is as easy to see that the next two propositions are a restatement of Axioms III and IX.

Proposition 2: If $\alpha, \beta \in \mathfrak{O}$, then $\alpha = \beta$ if and only if

$$P^{q_1';\cdots;q_m';\alpha;q_1'';\cdots;q_n''} \times (\{1\} \times \cdots \times \{1\} \times B \times \{1\} \times \cdots \times \{1\}) = P^{q_1';\cdots;q_m';\beta;q_1'';\cdots;q_n''} \times (\{1\} \times \cdots \times \{1\} \times B \times \{1\} \times \cdots \times \{1\})$$

for any $q_1', \cdots q_m', q_1'', \cdots q_n'' \in Q$, $B \in \mathfrak{B}^1$.

Proposition 3: If q is a question different from the null-question, i.e., if $P^q(\{1\}) \neq 0$ for some $P \in \mathfrak{s}$, then there is a $P_1 \in \mathfrak{s}$ such that $P_1^q(\{1\}) = 1$.

3.5. Properties of the spectrum of observables.

Proposition 1: A Borel set B in $\mathbf{R}^{\hat{a}}$ which does not contain any points of the spectrum, i.e., $B \cap \mathbf{S}^{(\alpha)} = \varnothing$, is a zero set.

Proof: For each $\hat{\lambda} \in B$, we can find a zero open interval $I(\hat{\lambda})$ containing $\hat{\lambda}$; otherwise, if such an interval does not exist, then $\hat{\lambda} \in \mathbf{S}^{(\alpha)}$, contrary to the assumption.

Denote by $\mathfrak{R}(B)$, the family of Borel sets consisting of a countable union of intervals $I(\hat{\lambda})$, $\hat{\lambda} \in B$. Each set of this family is a zero set because of the countable additivity of each complex probability measure $P^{\hat{a}}$, $P \in \mathfrak{s}$. The least upper bound of any linearly ordered family of sets in $\mathfrak{R}(B)$ (ordered by inclusion) obviously belongs to $\mathfrak{R}(B)$ on account of the continuity from below of each $P^{\hat{a}}$, $P \in \mathfrak{s}$. Therefore, according to Zorn's lemma, there is a maximal element B_0 in $\mathfrak{R}(B)$. B is a subset of B_0 because, if $\hat{\lambda} \in B$ and $\hat{\lambda} \notin B_0$, then $B_0 \cup I(\hat{\lambda}) \neq B_0$ and $B_0 \cup I(\hat{\lambda}) \in \mathfrak{R}(B)$; hence B_0 would not be maximal.

But B_0 is a zero set because it belongs to $\mathfrak{R}(B)$ and therefore each of its Borel subsets B', including B, and all the subsets of B, have the property $P^{\hat{a}}(B') = 0$, $P \in \mathfrak{s}$. Hence B is a zero set. Q.E.D.

It will be found useful to have a somewhat differently formulated version of this proposition which implicitly takes Axiom IX into consideration.

Proposition 2: Two Borel sets B_1, $B_2 \in \mathfrak{B}^{\hat{a}}$ are equal modulo $\mathfrak{s}$ if and only if their symmetric difference does not contain any points of the spectrum of $\hat{a}$, i.e., $(B_1 \Delta B_2) \cap \mathbf{S}^{\hat{a}} = \varnothing$.

Proposition 3: If $\{\hat{a}\} = C$, then each $P^{\hat{a}}$, $P \in \mathfrak{s}$, is different from zero on an at most countable number of points belonging to the point spectrum $\mathbf{S}_p^{\hat{a}}$.

Proof: Assume that $P^{\hat{a}}(\{\hat{\lambda}\}) > 0$ for an uncountable number of $\hat{\lambda} \in \mathbf{S}^{\hat{a}}$. Consider the number of such $\hat{\lambda}$'s for which $P^{\hat{a}}(\{\hat{\lambda}\})$ is in the intervals $(\frac{1}{2}, 1], \cdots (1/n + 1, 1/n], \cdots$. There has to be at least one such interval, say $(1/n_0 + 1, 1/n_0]$, such that there is an infinite number of $\hat{\lambda}$'s from $\mathbf{S}^{\hat{a}}$ for which $1/n_0 + 1 < P^{\alpha}(\{\hat{\lambda}\}) \leq 1/n_0$ (otherwise the number of $\hat{\lambda}$'s from $\mathbf{S}^{\hat{a}}$ for which $P^{\hat{a}}(\{\hat{\lambda}\}) > 0$ would be enumerable). But this means that $P^{\hat{a}}(\mathbf{R}^{\hat{a}}) \geq \sum_{\hat{\lambda} \in \mathbf{S}_p^{\alpha}} P^{\hat{a}}(\{\hat{\lambda}\}) = +\infty$, which is contrary to the request (Sec. 1.2, Proposition 2) that $P^{\hat{a}}(\mathbf{R}^{\hat{a}}) = 1$. Q.E.D.

Proposition 4: The spectrum $\mathbf{S}^{\hat{a}}$ of an ordered finite set $\hat{a}$ of observables contains the union of all carriers of all complex signed measures $P^{\hat{a}}$ for all $P \in \mathfrak{s}$, i.e., $\bigcup_{P \in \mathfrak{s}} \operatorname{carr} P^{\hat{a}} \subset \mathbf{S}^{\hat{a}}$, where carr $P^{\hat{a}}$ denotes the carrier of P^{α}. If $\bigcup_{P \in \mathfrak{s}} \operatorname{carr} P^{\hat{a}}$ is a closed set, then $\mathbf{S}^{\hat{a}} = \bigcup_{P \in \mathfrak{s}} \operatorname{carr} P^{\hat{a}}$.

Proof: According to the definition, the carrier of a measure $P^{\hat{a}}$ is the complement of the maximal open set which is zero with respect to $P^{\hat{a}}$.[43] Hence, if $\hat{\lambda} \in \operatorname{carr} P^{\hat{a}}$, then, each open interval containing $\hat{\lambda}$ cannot be a zero set with respect to $P^{\hat{a}}$, because, if such an open interval would exist, it would necessarily be a subset of the maximal open set which is zero with respect to $P^{\hat{a}}$ and, therefore, $\hat{\lambda}$ could not belong to carr $P^{\hat{a}}$. We conclude that $\hat{\lambda} \in \mathbf{S}^{\hat{a}}$. As our reasoning applied to any $P^{\hat{a}}$, $P \in \mathfrak{s}$ and any $\hat{\lambda} \in \operatorname{carr} P^{\hat{a}}$, we can write $\bigcup_{P \in \mathfrak{s}} \operatorname{carr} P^{\hat{a}} \subset \mathbf{S}^{\hat{a}}$.

To show that $\mathbf{S}^{\hat{a}} \subset \bigcup_{P \in \mathfrak{s}} \operatorname{carr} P^{\hat{a}}$ when $\bigcup_{P \in \mathfrak{s}} \operatorname{carr} P^{\hat{a}}$ is closed, we will prove that

$$(\bigcup_{P \in \mathfrak{s}} \operatorname{carr} P^{\hat{a}})' \subset (\mathbf{S}^{\hat{a}})'$$

(note that the fact that each carr $P^{\hat{a}}$ is a closed set does not imply that their union, in general infinite, is a closed set). If $\hat{\lambda} \notin \bigcup_{P \in \mathfrak{s}} \operatorname{carr} P^{\hat{a}}$ then, as $(\bigcup_{P \in \mathfrak{s}} \operatorname{carr} P^{\hat{a}})'$ is an open set, there is an open interval $I(\hat{\lambda})$ containing $\hat{\lambda}$ and having no common points with $\bigcup_{P \in \mathfrak{s}} \operatorname{carr} P^{\hat{a}}$. Therefore $I(\hat{\lambda})$ is a zero set with respect to all $P^{\hat{a}}$; this means that $I(\hat{\lambda})$ is a zero set in $\mathbf{R}^{\hat{a}}$. Hence $\hat{\lambda} \notin \mathbf{S}^{\hat{a}}$. Q.E.D.

Propositions 3 and 4 are more interesting because of what they tacitly suggest: Namely, that the pos-

[43] Ref. 25, p. 279.

sibility that the point spectrum of some observables contains an uncountable number of points and the possibility that $\mathbf{S}^{\mathfrak{a}} \neq \bigcup_{P \in \mathfrak{S}} \operatorname{carr} P^{\mathfrak{a}}$ are not excluded by the axioms.

It is easy to construct simple examples to show that.

Example 1 (showing that the point spectrum can contain an uncountable number of points). Choose a fundamental set $\mathcal{O}'$ consisting of only one observable α. Take for $\mathbf{S}^{(\alpha)}$ any uncountable subset of $\mathbf{R}^1$. Define $\mathfrak{S}$ as the convex set generated by the set of pure states. Each pure state P is defined by the measure $P^{\alpha}(\{\lambda\}) = 1$ for some $\lambda \in \mathbf{S}^{(\alpha)}$. It is easy to see that all axioms of Sec. 2.2. are obeyed by such a language, many of them in a trivial fashion. However, $\mathbf{S}^{(\alpha)}$ is a pure point spectrum with an uncountable number of points.

Example 2 (showing that it can happen that $\mathbf{S}^{(\alpha)} \neq \bigcup_{P \in \mathfrak{S}} \operatorname{carr} P^{\mathfrak{a}}$). Take, as in the preceding example, an $\mathcal{O}'$ consisting of only one observable. Choose an infinite subset S of $\mathbf{R}^1$ which has an accumulation point λ_0. Construct $\mathfrak{S}$ as the convex set generated by the set of pure states defined by

$$\mathfrak{S}_0 = \{P : P^{\alpha}(\{\lambda\}) = 1, \lambda \in S - \{\lambda_0\}\}.$$

As $\mathbf{S}^{(\alpha)}$ is a closed set (Sec. 2.2, Proposition 5) we get $\mathbf{S}^{(\alpha)} = S$, but $\bigcup_{P \in \mathfrak{S}} \operatorname{carr} P^{\mathfrak{a}} = S - \{\lambda_0\}$ and we have $\mathbf{S}^{(\alpha)} \neq \bigcup_{P \in \mathfrak{S}} \operatorname{carr} P^{\alpha}$.

ACKNOWLEDGMENTS

The author would like to thank Professor V. Bargmann for very interesting discussions and helpful comments. He would also like to thank O. Lanford for bringing to his attention some measure-theoretical theorems.

An Axiomatic Approach to the Formalism of Quantum Mechanics. II.*

EDUARD PRUGOVEČKI
Palmer Physical Laboratory, Princeton University, Princeton, New Jersey
(Received 30 March 1965)

Another axiom has to be introduced to make the formalism given in Part I physically equivalent to the conventional Hilbert-space formalism. Then it is shown that, given a certain fundamental set of observables, a B^*-algebra $\mathfrak{A}$ can be built into which the set $\mathcal{O}_b$ of all bounded observables can be mapped injectively. The closure of the algebra generated by the image of $\mathcal{O}_b$ into $\mathfrak{A}$ is $\mathfrak{A}$ itself. A Hilbert space $\mathfrak{H}$ exists into which the set $\mathcal{S}_0$ of all pure physical states can be mapped injectively. The closure of the subset of $\mathfrak{H}$ which is the image of this mapping is $\mathfrak{H}$ itself. The algebra $\mathfrak{A}$ can be mapped in such a fashion into the C^*-algebra $\mathfrak{B}(\mathfrak{H})$ of all bounded observables that these mappings provide, essentially, just a translation of the original formalism in a Hilbert-space formalism.

I. INTRODUCTION

IN this paper, we investigate the relation of the formalism given in I[1] to the conventional Hilbert-space formalism.

In the conventional Hilbert-space formalism, we deal with a Hilbert space $\mathfrak{H}$ in which, in the general case when superselection rules might be present, only *some* of the vectors represent pure physical states. At present, it is considered that separable Hilbert spaces $\mathfrak{H}$ with discrete superselection rules offer us sufficient generality. It is tacitly assumed that each vector in a coherent subspace of $\mathfrak{H}$ represents a physical state and that each self-adjoint bounded operator on $\mathfrak{H}$, leaving each coherent subspace invariant, represents an observable.

Both these assumptions are unnecessarily strong in so far as the physical content of a theory is concerned. First, in the laboratory, we always make measurements and state preparation with a finite precision. As a consequence, from the theoretical point of view, we never determine only one physical state but rather a weak environment of such a state.[2] Therefore, it is sufficient to require that the vectors of a coherent subspace (representing physical states) form a subset dense everywhere in that subspace of some topology which is not weaker than the weak topology.

Similar considerations apply to the set of all observables and lead us to the following definitions.

A mapping $\alpha \to \varphi(\alpha) = A$ of the set $\mathcal{O}_b$ of bounded observables into a Banach *-algebra is called an *embedding of* $\mathcal{O}_b$ *into* $\mathfrak{H}$ if it fulfills the following conditions.

(a) It is an injective mapping of $\mathcal{O}_b$ into $\mathfrak{A}$, i.e., $\alpha \neq \beta$ implies $\varphi(\alpha) \neq \varphi(\beta)$.

(b) It maps isometrically each normed real algebra $\mathcal{C} \in \mathfrak{M}_c^b$, determined by a maximal set of compatible bounded observables (see I, Sec. 3.2, Proposition 4), into a (not necessarily closed) *real* subalgebra of $\mathfrak{A}$.

(c) The closure, in some suitable topology which is not weaker than the weak topology, of any subalgebra of $\mathfrak{A}$ which contains the image of $\mathcal{O}_b$ in $\mathfrak{A}$ is $\mathfrak{A}$ itself.

We do not require that the set $\mathcal{O}_b$ is itself an algebra. Take, for example, two observables such as the position coordinate x and the corresponding momentum coordinate p_x along a certain axis represented (in a given theory) by the self-adjoint operators $X = \int_{-\infty}^{+\infty} \lambda dE_\lambda$ and $P_x = \int_{-\infty}^{+\infty} \mu dF_\mu$. According to Axiom IV in I, we consider that, for any λ and μ, E_λ and F_μ represent observables. It would be, however, completely *redundant* to enlarge our definitions and require that $E_\lambda + F_\mu$ is an observable. Namely, there are no experimental procedures attached to $E_\lambda + F_\mu$ intended to measure it *directly*. The only roundabout way one can imagine of "measuring" $E_\lambda + F_\mu$ would be to determine the physical state of the system by performing measurements on fundamental observables,[3] and then to compute from the theory the mean "value" of $E_\lambda + F_\mu$. However, the functional and predictive role of the theory is determined by the knowledge of the physical states and fundamental observables. Therefore, the definition of algebraic operations between the elements of $\mathcal{O}_b$, in general, is a matter of convenience in which we have complete freedom as long as these opera-

* This paper is based, in part, upon the author's doctoral dissertation submitted to the Physics Department of Princeton University, Princeton, New Jersey.

[1] E. Prugovečki, J. Math. Phys. 7, 1054 (1966), previous paper, referred to as I.

[2] These matters are discussed in detail in the first part of the author's Princeton Thesis (unpublished). These ideas are of the same type as the ones which lead to the concept of physical equivalence which plays a central role in the paper: R. Haag and D. Kastler, J. Math. Phys. 5, 848 (1964).

[3] See I, Sec. 1.

tions retain their old meaning in the cases where they were already defined. Besides, we have to be careful that our construction does not overburden the framework by introducing unnecessarily many redundant objects in the theory. All these requirements are met by the above-introduced concept of "embedding."

A Hilbert space $\mathfrak{H}$ (not necessarily separable) is said to provide a *representation of the quantum mechanical language* $\mathfrak{L}$ (given by $\mathfrak{O}$ and $\mathfrak{S}$ obeying Axioms I–IX in I), if the set $\mathfrak{O}_b$ of bounded observables can be embedded in a closed subalgebra $\mathfrak{A}$ of the algebra $\mathfrak{B}(\mathfrak{H})$ of bounded operators on $\mathfrak{H}$, and, if there exists a mapping of the set $\mathfrak{S}_0$ of pure physical states into $\mathfrak{H}$ such that the following conditions are fulfilled.

(a) The mapping of $\mathfrak{S}$ into $\mathfrak{H}$ is an injective mapping.

(b) The closure of the linear manifold spanned by the image of $\mathfrak{S}$ in $\mathfrak{H}$ is the entire Hilbert space $\mathfrak{H}$.

(c) If $A_1, \cdots A_n$ are the images in $\mathfrak{B}(\mathfrak{H})$ of any n ($n = 1, 2, \cdots$) bounded observables $\alpha_1, \cdots \alpha_n$, respectively, and Ψ is the image in $\mathfrak{H}$ of the pure physical state P, then we have

$$P^{\alpha_1;\cdots;\alpha_n}(B_1 \times \cdots \times B_n) = \langle \Psi | E^{A_1}_{B_1} \cdots E^{A_n}_{B_n} | \Psi \rangle \tag{1.1}$$

for any $B_1, \cdots B_n \in \mathfrak{B}^1$. Here $E^{A_1}_{B_1}, \cdots E^{A_n}_{B_n}$ are the projectors of the spectral decompositions

$$A_1 = \int_{-\infty}^{+\infty} \lambda_1 \, dE^{A_1}_{\lambda_1}, \cdots, A_n = \int_{-\infty}^{+\infty} \lambda_n \, dE^{A_n}_{\lambda_n}. \tag{1.2}$$

Our aim in the next two sections is to investigate under which conditions a Hilbert-space representation of the language $\mathfrak{L}$, formulated in terms of our axioms, exists. The resulting central theorems are Theorem 1 in Sec. 2.3, Theorem 6 in Sec. 2.5, and Theorem 3 in Sec. 3.3.

2. THE EMBEDDING OF THE SET OF BOUNDED OBSERVABLES IN A B^*-ALGEBRA[4]

2.1. The algebra $\mathfrak{A}_1(\mathfrak{O}')$

Choose a fundamental set $\mathfrak{O}'$ of observables and denote by $\mathfrak{Q}(\mathfrak{O}')$ the set of all questions of the form $\chi^\alpha_B(\alpha)(\alpha \in \mathfrak{O}', B \in \mathfrak{B}^{(\alpha)})$. It is later convenient to have a shorter notation $\mathfrak{Q}'$ for $\mathfrak{Q}(\mathfrak{O}')$, but it has to be remembered that such a $\mathfrak{Q}'$ is always defined only in relation to a certain $\mathfrak{O}'$ which is then not explicitly displayed. We then say that $\mathfrak{Q}'$ is the family of *simple questions relative to* $\mathfrak{O}'$, because the concept of a simple question is again defined only in relation to a fundamental set of observables.

We now consider the family $\hat{\mathfrak{Q}}'$ of n-tuples $(q_1, q_2, \cdots)$, $n = 1, 2, 3, \cdots$, on $\mathfrak{Q}'$. We define equivalence classes in $\hat{\mathfrak{Q}}'$ by the following rules.

Rule 1: If $(q'_1, \cdots q'_n), (q''_1, \cdots q''_n) \in \hat{\mathfrak{Q}}'$ are two n-tuples consisting of the same questions, then they are equivalent if the permutation

$$\begin{pmatrix} q'_1, & \cdots & q'_n \\ q''_1, & \cdots & q''_n \end{pmatrix}$$

can be written as a finite number of successive permissible inversions.

Any inversion of two *neighboring* questions

$$\begin{pmatrix} \cdots & q_i, & q_{i+1}, & \cdots \\ \cdots & q_{i+1}, & q_i, & \cdots \end{pmatrix}$$

is called a *permissible inversion* if $\{q_i, q_{i+1}\} = C$.

Rule 2: Let $(q'_1, \cdots q'_k), (q''_1, \cdots q''_l) \in \hat{\mathfrak{Q}}'$ be any two ordered sets of simple questions. These two sets are considered equivalent if, after eliminating from both of them all questions which are equal to one modulo $\mathfrak{S}$, and after retaining the old order in the so derived subsets $(q^1_1, \cdots q^1_{k_0})$ and $(q^2_1, \cdots q^2_{l_0})$ of the original sets, then the sets $(q^1_1, \cdots q^1_{k_0})$ and $(q^2_1, \cdots q^2_{l_0})$ are equivalent according to Rule 1; i.e., $k_0 = l_0$, they consist of the same observables, each kind of observable occuring the same number of times, and they can be derived from one another by permissible inversions.

It is easy to see that Rules 1 and 2 really define equivalence classes, i.e., that $\hat{q} \sim \hat{q}$, $\hat{q}_1 \sim \hat{q}_2$ imply $\hat{q}_2 \sim \hat{q}_1$, and $\hat{q}_1 \sim \hat{q}_2$, $\hat{q}_2 \sim \hat{q}_3$ imply $\hat{q}_1 \sim \hat{q}_3$ for any $\hat{q}, \hat{q}_1, \hat{q}_2, \hat{q}_3 \in \hat{\mathfrak{Q}}'$.

For reasons which become clear later, it is more convenient (and suggestive) to introduce a new symbol for the equivalence class of ordered sets containing the set $\hat{q} = (q_1, \cdots q_n)$, namely $q_1 \circ \cdots \circ q_n$. We call a symbol of the above form an *ordered product* of simple questions relative to $\mathfrak{O}'$, and the corresponding equivalence class represented by it is the *class of equal ordered products*. The family of all such equivalence classes is denoted by $\mathfrak{P}(\mathfrak{O}')$ or shortly $\mathfrak{P}'$.

Two equivalence classes are of special importance to us: the *zero equivalence class* and the *unit equivalence class* of ordered products. The zero equivalence class consists of all ordered products containing *at least* one zero question. The unit equivalence class consists of the ordered products consisting of *only*

[4] The terminology is the one used in C. E. Rickart, *General Theory of Banach Algebras* (D. van Nostrand, Inc., New York 1960). In this terminology, a Banach *-algebra, with the property $||x^*x|| = ||x||^2$ for each of its elements x, is called a B^*-algebra. A Banach algebra of bounded operators on a Hilbert space is called a C^*-algebra.

unit questions. The zero and unit questions (it is irrelevant for the purpose of this definition whether they are simple questions or not) are any q^0 and q^1, respectively, defined in I, Sec. 3.2 in the course of proving Proposition 2.

Consider now the family $\mathbf{C}^1 \tilde{\times} \mathcal{P}^f$ of finite sets $\{(a_r, p_r), r = 1, 2, \cdots \rho, \rho < +\infty\}$ of ordered pairs (a_r, p_r), where a_r are complex numbers and p_r are equivalence classes of ordered products, i.e., $a_r \in \mathbf{C}^1$, $p_r = q_1^{(r)} \circ \cdots \circ q_{n_r}^{(r)} \in \mathcal{P}^f$, $r = 1, \cdots \rho$. The complex number a_r in each ordered pair (a_r, p_r) is called the *coefficient* of that ordered pair.

We define equivalence classes in $\mathbf{C}^1 \tilde{\times} \mathcal{P}^f$ with the help of our next three rules.

Rule 3: All ordered pairs (a, p), $a \in \mathbf{C}^1$, $p \in \mathcal{P}^f$, which fulfill at least one of the following two conditions:

(1) $a = 0$;

(2) the equivalence class p of ordered products is the zero equivalence class of $\mathcal{P}^f$,

are classified together into a subset of $\mathbf{C}^1 \times \mathcal{P}^f$ called the *zero class of ordered pairs* from $\mathbf{C}^1 \times \mathcal{P}^f$.

To each element $\tilde{p} = \{(a_1, p_1), \cdots (a_\rho, p_\rho)\}$ of $\mathbf{C}^1 \tilde{\times} \mathcal{P}^f$ [5] we adjoin, with the help of the next rule, another element $\check{p} = \{(\check{a}_1, \check{p}_1), \cdots, (\check{a}_\sigma, \check{p}_\sigma)\}$ having the property that all its constituent ordered pairs from $\mathbf{C}^1 \times \mathcal{P}^f$ do not belong to the zero class in $\mathbf{C}^1 \times \mathcal{P}^f$ and that $\check{p}_1 \neq \cdots \neq \check{p}_\sigma$. To each $\tilde{p} \in \mathbf{C}^1 \tilde{\times} \mathcal{P}^f$, a certain $\check{p}$ can always be adjoint with the help of Rule 4, and $\check{p}$ is a set of ordered pairs from $\mathbf{C}^1 \times \mathcal{P}^f$ uniquely determined by $\tilde{p}$. Therefore such a $\check{p}$ is called the *standard form* of $\tilde{p}$.

Rule 4: To obtain the standard form $\check{p}$ of an element $\tilde{p}$ of $\mathbf{C}^1 \tilde{\times} \mathcal{P}^f$, replace each subset of $\tilde{p}$ of *all* ordered pairs (having equal equivalence classes of ordered products) with a single ordered pair [which has as coefficient the (algebraic) sum of the coefficients, and as element from $\mathcal{P}^f$ the common equivalence class of ordered products of the ordered pairs which it replaces]. To elaborate this first step, assume that $\tilde{p}$ is written in the form (to which, obviously, any $\tilde{p}$ can be reduced):

$$\{(a_1^{(1)}, p_1^{(1)}), \cdots (a_{\rho_1}^{(1)}, p_{\rho_1}^{(1)}); \cdots; (a_1^{(\sigma)}, p_1^{(\sigma)}), \cdots (a_{\rho_\sigma}^{(\sigma)}, p_{\rho_\sigma}^{(\sigma)})\}, \tag{2.1}$$

where $p_1^{(1)} = \cdots p_{\rho_1}^{(1)} = p^{(1)}, \cdots, p_1^{(\sigma)} = \cdots = p_{\rho_\sigma}^{(\sigma)} = p^{(\sigma)}$, and $p^{(1)}, \cdots p^{(\sigma)}$ are all different from one another, i.e., $p^{(\mu)} \neq p^{(\nu)}$ for $\mu \neq \nu$, $\mu, \nu = 1, 2, \cdots \sigma$. Then the first step of our procedure consists in replacing $\tilde{p}$ by

$$\{(a^{(1)}, p^{(1)}), \cdots (a^{(\sigma)}, p^{(\sigma)})\}, \tag{2.2}$$

where $a^{(1)} = a_1^{(1)} + \cdots + a_{\rho_1}^{(1)}, \cdots, a^{(\sigma)} = a_1^{(\sigma)} + \cdots + a_{\rho_\sigma}^{(\sigma)}$.

The second step consists in eliminating from (2.2) all the ordered pairs $(a^{(r)}, p^{(r)})$ which belong to the zero class of $\mathbf{C}^1 \times \mathcal{P}^f$. The resulting set of ordered pairs from $\mathbf{C}^1 \times \mathcal{P}^f$ is the element $\check{p}$ of $\mathbf{C}^1 \tilde{\times} \mathcal{P}^f$ called the *standard form* of $\tilde{p}$.

Rule 5: Two elements $\tilde{p}'$, $\tilde{p}''$ of $\mathbf{C}^1 \tilde{\times} \mathcal{P}^f$ are equivalent if the corresponding standard forms $\check{p}'$ and $\check{p}''$ are equivalent, i.e., if $\check{p}'$ and $\check{p}''$ are identical sets of elements from $\mathbf{C}^1 \tilde{\times} \mathcal{P}^f$.

It is easy to see that the "equivalence" relation introduced by Rule 5 can be used for defining equivalence classes in $\mathbf{C}^1 \tilde{\times} \mathcal{P}^f$ by grouping together in one class all the equivalent (according to Rule 5) elements of $\mathbf{C}^1 \tilde{\times} \mathcal{P}^f$. We can do this because, evidently, the "equivalence" relation introduced by Rule 5 satisfies the axioms of reflexiveness, symmetry and transitivity required for an equivalence relation.

We denote the set of these equivalence classes by $\mathcal{R}(\mathcal{P}^f)$, or shortened, by $\mathcal{R}^f$, while keeping in mind that the whole procedure depends on our initial choice of $\mathcal{P}^f$. An equivalence class of $\mathcal{R}^f$ is represented with the help of any of its elements $\{(a_1, p_1), \cdots (a_\rho, p_\rho)\}$. For such an element, representing an equivalence class, a more convenient notation is introduced by writing $\{(a_1, p_1), \cdots (a_\rho, p_\rho)\}$ in the form

$$a_1 p_1 \oplus a_2 p_2 \oplus \cdots \oplus a_\rho p_\rho \tag{2.3}$$

or, sometimes, in the shorter notation

$$\bigoplus_r a_r p_r, \qquad a_r \in \mathbf{C}^1, \qquad p_r \in \mathcal{P}^f. \tag{2.4}$$

Symbols of the form (2.3) and (2.4) are called *polynomial forms on* $\mathcal{P}^f$ or, when no ambiguities can arise, *polynomial forms.*

The earlier introduced standard forms in $\mathbf{C}^1 \tilde{\times} \mathcal{P}^f$ written in this new notation are called *standard polynomial forms on* $\mathcal{P}^f$ or simply *standard polynomial forms.*

We need, for future use, the following notation: If $p' = q_1' \circ \cdots \circ q_m'$ and $p'' = q_1'' \circ \cdots \circ q_n''$ represent two elements of $\mathcal{P}^f$, then the symbol $p' \circ p''$ means the same and represents the same element of $\mathcal{P}^f$ as the symbol $q_1' \circ \cdots \circ q_m' \circ q_1'' \circ \cdots \circ q_n''$. We express this notational agreement by writing

$$p' \circ p'' = q_1' \circ \cdots \circ q_m' \circ q_1'' \circ \cdots \circ q_n''. \tag{2.5}$$

[5] We consistently use the following notation: if S denotes any set, then the family of n-tuples ($n = 0, 1, 2, \ldots$) of elements from S is denoted by $\hat{S}$, while the class of finite subsets of S is denoted by $\tilde{S}$. Due to technical reasons in printing, for $S = \mathbf{C}^1 \times \mathcal{P}^f$ the tilde symbol $\sim$ appears only above $\times$, i.e., $\tilde{S} = \mathbf{C}^1 \tilde{\times} \mathcal{P}^f$.

Lemma 1: Operations of addition and multiplication can be defined between the elements of $\mathfrak{R}^f$, and a "multiplication with scalar" operation of elements of $\mathfrak{R}^f$ with complex numbers can be defined in such a way that, with appropriate definitions of the zero and unity, all the axioms of an (in general noncommutative) algebra are satisfied.

If a denotes any complex number and $\check{p}$, $\check{p}'$, $\check{p}''$ are standard polynomial forms over $\mathfrak{O}^f$ representing, respectively, the elements A, A', A'' of $\mathfrak{R}^f$, where $\check{p} = a_1 p_1 \oplus \cdots \oplus a_\sigma p_\sigma$, $\check{p}' = a'_1 p'_1 \oplus \cdots \oplus a'_{\sigma_1} p'_{\sigma_1}$ and $\check{p}'' = a''_1 p''_1 \oplus \cdots \oplus a''_{\sigma_2} p''_{\sigma_2}$, then the mentioned operations are defined in the following way:

Addition. The sum $A' + A''$ is the element of $\mathfrak{R}^f$ represented by the polynomial form (which is not necessarily standard) $a'_1 p'_1 \oplus \cdots \oplus a'_{\sigma_1} p'_{\sigma_1} \oplus a''_1 p''_1 \oplus \cdots \oplus a''_{\sigma_2} p''_{\sigma_2}$, which is also written in the shorter form $\check{p}' \oplus {}''\check{p}$. Symbolically

$$\begin{aligned} &A' + A'' \\ &= (a'_1 p'_1 \oplus \cdots \oplus a'_{\sigma_1} p'_{\sigma_1}) + (a''_1 p''_1 \oplus \cdots \oplus a''_{\sigma_2} p''_{\sigma_2}) \\ &= a'_1 p'_1 \oplus \cdots \oplus a'_{\sigma_1} p'_{\sigma_1} \oplus a''_1 p''_1 \oplus \cdots \oplus a''_{\sigma_2} p''_{\sigma_2}, \end{aligned} \tag{2.6}$$

where it has to be clear, at least for the time being, that the polynomials in the parentheses of the left-hand side of (2.6) are standard.

Multiplication. The product $A'A''$ is the element of $\mathfrak{R}^f$ represented by the polynomial form (in general not standard) $\bigoplus_{r=1}^{\sigma_1} \bigoplus_{s=1}^{\sigma_2} (a'_r a''_s) p'_r \circ p''_s$, which is also written in the shorter form $\check{p}' \circ \check{p}''$. Symbolically:

$$\begin{aligned} A'A'' &= (a'_1 p'_1 \oplus \cdots \oplus a'_{\sigma_1} p'_{\sigma_1})(a''_1 p''_1 \oplus \cdots \oplus a''_{\sigma_2} p''_{\sigma_2}) \\ &= a'_1 a''_1 p'_1 \circ p''_1 \oplus \cdots \oplus a'_1 a''_{\sigma_1} p'_1 \circ p''_{\sigma_1} \\ &\quad \oplus \cdots \oplus a'_{\sigma_1} a''_1 p'_{\sigma_1} \circ p''_1 \\ &\quad \oplus \cdots \oplus a'_{\sigma_1} a''_{\sigma_2} p'_{\sigma_1} \circ p''_{\sigma_2}, \end{aligned} \tag{2.7}$$

where it is again self-understood that the expressions in the parentheses of the left-hand side of (2.7) are standard polynomials.

Multiplication with scalar. The product aA is the element of $\mathfrak{R}^f$ represented by $\bigoplus_{r=1}^{\sigma} (a\, a_r) p_r = (a\, a_1) p_1 \oplus \cdots \oplus (a\, a_\sigma) p_\sigma$. Symbolically:

$$\begin{aligned} aA &= a(a_1 p_1 \oplus \cdots \oplus a_\sigma p_\sigma) \\ &= (aa_1) p_1 \oplus \cdots \oplus (a_1 a_\sigma) p_\sigma. \end{aligned} \tag{2.8}$$

The zero element **0** and the unit element **1** of algebra $\mathfrak{A}_1(\mathfrak{O}^f)$, whose elements are the elements of $\mathfrak{R}(\mathfrak{O}^f)$ and in which the operations are defined by (2.6), (2.7), and (2.8), are the equivalence classes in $\mathbf{C}^1 \tilde{\times} \mathfrak{O}^f$, i.e., elements of $\mathfrak{R}(\mathfrak{O}^f)$, defined in the following way:

The *zero* element **0** of $\mathfrak{A}_1^f$[6] is the equivalence class of $\mathbf{C}^1 \tilde{\times} \mathfrak{O}^f$ which contains sets consisting only of ordered pairs belonging to the zero class of $\mathbf{C}^1 \times \mathfrak{O}^f$.

The *unit* element **1** of $\mathfrak{A}_1^f$ is the equivalence class of $\mathbf{C}^1 \tilde{\times} \mathfrak{O}^f$ which has as standard polynomial form over $\mathfrak{O}^f$ the set $\mathfrak{p}^1 = \{(1, p^1)\}$ consisting of only one ordered pair, where p^1 is the unit equivalence class of ordered products.

Proof: It is clear from the way we have defined the equivalence classes in $\mathbf{C}^1 \tilde{\times} \mathfrak{O}^f$, i.e., the elements of $\mathfrak{R}^f$, that to each element of $\mathfrak{R}^f$ belongs one and only one standard polynomial form. Hence, the definitions (2.6), (2.7), and (2.8) determine uniquely the sum and product of two elements of $\mathfrak{R}^f$ (or, respectively, the product of an element of $\mathfrak{R}^f$ with a scalar) as an element of $\mathfrak{R}^f$. The only detail which has to be mentioned to make the validity of these statements obvious is the evident fact that formula (2.5) defines the same element of $\mathfrak{O}^f$ no matter what representative ordered products p', p'' we choose to stand for the component elements of $\mathfrak{O}^f$. This remark makes the uniqueness of the product (2.7) quite obvious.

Some of the axioms for an algebra can be now checked very easily in a direct way. From now on, we denote by A, B, C, $\cdots$ the elements of $\mathfrak{A}_1^f$, i.e., the equivalence classes of $\mathbf{C}^1 \tilde{\times} \mathfrak{O}^f$, for which the operations (2.6), (2.7), and (2.8) are defined.

To check that

$$A + B = B + A, \qquad A, B \in \mathfrak{A}_1^f, \tag{2.9}$$

we pick up the standard polynomial forms $a_1 p'_1 \oplus \cdots \oplus a_{\sigma_1} p'_{\sigma_1}$ and $b_1 p''_1 \oplus \cdots \oplus b_{\sigma_2} p''_{\sigma_2}$ representing A and B, respectively, and remembering that, according to the definition of the elements of $\mathbf{C}^1 \tilde{\times} \mathfrak{O}^f$, the two ordered polynomial forms $a_1 p'_1 \oplus \cdots \oplus a_{\sigma_1} p'_{\sigma_1} \oplus b_1 p''_1 \oplus \cdots \oplus b_{\sigma_2} p''_{\sigma_2}$ and $b_1 p''_1 \oplus \cdots \oplus b_{\sigma_2} p''_{\sigma_2} \oplus a_1 p'_1 \oplus \cdots \oplus a'_{\sigma_1} p'_{\sigma_1}$ represent the same element of $\mathfrak{R}^f$, i.e., of $\mathfrak{A}_1^f$. After taking into consideration formula (2.6), (2.9) is obviously established.

A similar reasoning helps us to prove that

$$a(bA) = (ab)A, \quad a, b \in \mathbf{C}^1, \quad A \in \mathfrak{A}_1^f. \tag{2.10}$$

The only additional detail to be taken into consideration in proving (2.10) is that, if $a \neq 0$ and $b \neq 0$, then the fact that one of the polynomial forms $(aba_1)p_1 \oplus \cdots \oplus (aba_\sigma)p_\sigma$ and $(ba_1)p_1 \oplus \cdots \oplus (ba_\sigma)p_\sigma$ is standard implies that the other one is a

[6] In the same spirit as in similar previous cases, we introduce $\mathfrak{A}_1{}^f$, as a shorter notation for $\mathfrak{A}_1(\mathfrak{O}^f)$.

standard polynomial form too. The case that either $a = 0$, $b = 0$, or $a = b = 0$ is too obvious to be commented upon.

However, already the relation

$$a(A + B) = aA + aB, \qquad a \in \mathbf{C}^1, \qquad A, B \in \mathfrak{A}_1'$$

cannot be proven in such a straightforward manner, because, if $\check{p}'$ and $\check{p}''$ are the standard polynomial forms representing the equivalence classes A and B, respectively, then indeed $a\check{p}'$ and $a\check{p}''$ are also standard polynomials (for $a \neq 0$), but $\check{p}' \oplus \check{p}''$ is not in general a standard polynomial form and we do not know yet immediately whether $a(\check{p}' \oplus \check{p}'')$ represents $a(A + B)$. (Remember that the expression aA is defined by (2.8) only when we take for the polynomial in the left-hand side, representing A, the standard polynomial form!)

Our problems are easily solvable after we have proved the following lemma.

Lemma 2: If $\bar{p}, \bar{p}', \bar{p}'' \in \mathfrak{R}(\mathcal{O}')$ and $a \in \mathbf{C}^1$, then the polynomial forms on $\mathcal{O}'$, $\bar{p}' \oplus \bar{p}''$, $\bar{p}' \circ \bar{p}''$, and $a\bar{p}$, defined from the polynomial forms $\bar{p}'$, $\bar{p}''$, and $\bar{p}$ by the formulas (2.12), (2.13), and (2.20), respectively, represent the same equivalence class in $\mathfrak{R}'$ as the polynomial forms $\check{p}' \oplus \check{p}''$, $\check{p}' \circ \check{p}''$, and $a\check{p}$, defined by (2.6), (2.7), and (2.8) from the standard polynomial forms $\check{p}$, $\check{p}'$, $\check{p}''$ corresponding, respectively, to $\bar{p}$, $\bar{p}'$, $\bar{p}''$.

Proof: Take $\bar{p}'$ and $\bar{p}''$ written in the form

$$\begin{aligned} \bar{p}' &= \{(a_1', p_1'), \cdots (a_{\rho_1}', p_{\rho_1}')\} \\ &= a_1'p_1' \oplus \cdots \oplus a_{\rho_1}'p_{\rho_1}', \\ \bar{p}'' &= \{(a_1'', p_1''), \cdots (a_{\rho_2}'', p_{\rho_2}'')\} \\ &= a_1''p_1'' + \cdots + a_{\rho_2}''p_{\rho_2}''. \end{aligned} \tag{2.11}$$

We denote by the symbol $\bar{p}' \oplus \bar{p}''$ the following polynomial form on $\mathcal{O}'$ [7]

$$\begin{aligned} &\bar{p}' \oplus \bar{p}'' \\ &= a_1'p_1' \oplus \cdots \oplus a_{\rho_1}'p_{\rho_1}' \oplus a_1''p_1'' \oplus \cdots \oplus a_{\rho_2}''p_{\rho_2}'' \\ &= \{\{(a_1'p_1'), \cdots (a_{\rho_1}', p_{\rho_1}'), (a_1''p_1''), \cdots (a_{\rho_2}'', p_{\rho_2}'')\}\}. \end{aligned} \tag{2.12}$$

A careful study of Rule 3 should reveal that the standard polynomial form corresponding to (2.12) is the same as the standard polynomial form which corresponds to $\check{p}' \oplus \check{p}''$. We will not dwell any longer on this case, but turn to the case of the product $\bar{p}' \circ \bar{p}''$, which possesses (besides others) all the features of the case of the sum $\bar{p}' \oplus \bar{p}''$, only in a more complicated form.

The symbol $\bar{p}' \circ \bar{p}''$, in which $\bar{p}'$ and $\bar{p}''$ are the polynomials forms of general type (2.12), is defined to stand for

$$\begin{aligned} \bigoplus_{r=1}^{\rho_1} \bigoplus_{s=1}^{\rho_2} a_r'a_s''p_r' \circ p_s'' = \{\{&a_1'a_1''p_1' \circ p_1'', \cdots \\ &a_1'a_{\rho_2}''p_1' \circ p_{\rho_2}'', \cdots a_{\rho_1}'a_{\rho_2}''p_{\rho_1}' \circ p_{\rho_2}''\}\}. \end{aligned} \tag{2.13}$$

We have to show that (2.13) has the same standard polynomial form as the expression $\check{p}' \circ \check{p}''$. With that purpose in mind, we write $\bar{p}'$ and $\bar{p}''$ in the form

$$\begin{aligned} \bar{p}' &= \{\{(a_1^{(1)}, p_1^{(1)}), \cdots (a_{\mu_1}^{(1)}, p_{\mu_1}^{(1)}), \\ &\qquad \cdots (a_1^{(\sigma)}, p_1^{(\sigma)}) \cdots (a_{\mu_\sigma}^{(\sigma)}, p_{\mu_\sigma}^{(\sigma)})\}\}, \\ \bar{p}'' &= \{\{(b_1^{(1)}, r_1^{(1)}), \cdots (b_{\nu_1}^{(1)}, r_{\nu_1}^{(1)}), \\ &\qquad \cdots (b_1^{(\tau)}, r_1^{(\tau)}), \cdots (b_{\nu_\tau}^{(\tau)}, r_{\nu_\tau}^{(\tau)})\}\}, \end{aligned} \tag{2.14}$$

where

$$\begin{aligned} &p_1^{(1)} = \cdots = p_{\mu_1}^{(1)} = p^{(1)}, \cdots, p_1^{(\sigma)} = \cdots \\ &\quad = p_{\mu_\sigma}^{(\sigma)} = p^{(\sigma)}, r_1^{(1)} = \cdots = r_{\nu_1}^{(1)} = r^{(1)}, \cdots \\ &\quad r_1^{(\tau)} = \cdots = r_{\nu_\tau}^{(\tau)} = r^{(\tau)}, \quad \text{and} \quad p^{(i)} \neq p^{(j)} \end{aligned}$$

for $i \neq j$, $r^{(i)} \neq r^{(j)}$ for $i \neq j$, $(p_1^{(1)}, \cdots p_{\tau\mu}^{(\sigma)})$, $(r_1^{(1)}, \cdots, r_\tau^{(\tau)}) \in \mathcal{O}'$). This can always be done because it involves only the rearrangement of a finite set.

The corresponding standard polynomial forms $\check{p}'$ and $\check{p}''$ can now be written as

$$\begin{aligned} \check{p}' &= \{\{(a_1^{(1)} + \cdots + a_{\mu_1}^{(1)}, p^{(1)}), \\ &\qquad \cdots (a_1^{(\sigma)} + \cdots + a_{\mu_\sigma}^{(\sigma)}, p^{(\sigma)})\}\}, \\ \check{p}'' &= \{\{(b_1^{(1)} + \cdots + b_{\nu_1}^{(1)}, r^{(1)}), \\ &\qquad \cdots (b_1^{(\tau)} + \cdots + b_{\nu_\tau}^{(\tau)}, r^{(\tau)})\}\}. \end{aligned} \tag{2.15}$$

To prove that $\bar{p}' \circ \bar{p}'' = \check{p}' \circ \check{p}''$, i.e., that $\bar{p}' \circ \bar{p}''$ and $\check{p}' \circ \check{p}''$ have the same standard polynomial form, we show the following: Take

$$\begin{aligned} \bar{p} &= \{\{\cdots ; (a_1^{(\lambda)}, p^{(\lambda)}), \\ &\qquad \cdots (a_k^{(\lambda)}, p^{(\lambda)}), (a^{(\lambda)}, p^{(\lambda)}); \cdots\}\}, \\ \bar{r} &= \{\{\cdots ; (a_1^{(\lambda)}, p^{(\lambda)}), \\ &\qquad \cdots (a_k^{(\lambda)}, p^{(\lambda)}), (b^{(\lambda)}, p^{(\lambda)}), (c^{(\lambda)}, p^{(\lambda)}); \cdots\}\}, \end{aligned} \tag{2.16}$$

to be two polynomial forms on $\mathcal{O}'$, written in the same form as (2.14); in (2.16), all the ordered pairs, except the ones explicitly written, are the same in $\bar{p}$ and $\bar{r}$, and such that their corresponding equivalence classes of ordered products $p^{(1)}, \cdots p^{(\lambda-1)}$, $p^{(\lambda+1)}, \cdots p^{(\sigma)}$ are different from $p^{(\lambda)}$ (as well as different between themselves). If further we have $a^{(\lambda)} = b^{(\lambda)} + c^{(\lambda)}$, then $\bar{p} \circ \bar{p}'' = \bar{r} \circ \bar{p}''$, where $\bar{p}''$

[7] The two curly brackets on the right-hand side of (2.12) indicate that we are dealing with an equivalence class containing as an element the set within the inner curly brackets.

is the polynomial form given explicitly by the second expressions of (2.14).

To see this, we consider in more detail the method of constructing the standard polynomial forms for $\tilde{p} \circ \tilde{p}''$ and $\tilde{r} \circ \tilde{p}''$. This is done by grouping together ordered pairs having equal equivalence classes of ordered products:

$$p^{(1)} \circ r^{(\kappa_1{}^1)} = \cdots = p^{(1)} \circ r^{(\kappa_{n^1})} = \cdots p^{(\lambda)} \circ r^{(\kappa_1{}^\lambda)} = \cdots = p^{(\lambda)} \circ r^{(\kappa_{n^\lambda})} = \cdots = p^{(\sigma)} \circ r^{(\kappa_1{}^\sigma)} = \cdots = p^{(\sigma)} \circ r^{(\kappa_{n^\sigma})}. \tag{2.17}$$

The coefficient, in the standard polynomial form corresponding to $\tilde{p} \circ \tilde{p}''$, of an ordered pair from $\mathbf{C}^1 \times \mathcal{P}'$ belonging to an equivalence class of ordered products equal to (2.17), can be written, after some rearrangements, in the convenient form (in which the summands which do not correspond to λ are not explicitly written, because they are obviously the same in $\tilde{p} \circ \tilde{p}''$ and $\tilde{r} \circ \tilde{p}''$):

$$\cdots + (a_1^{(\lambda)} + \cdots + a_k^{(\lambda)} + a^{(\lambda)})[(b_1^{(\rho)} + \cdots + b_{\nu_\rho}^{(\rho)}) + \cdots + (b_1^{(\omega)} + \cdots + b_{\nu_\omega}^{(\omega)})] + \cdots ; \quad \rho = \kappa_1^\lambda, \cdots \omega = \kappa_{n_\sigma}^\sigma. \tag{2.18}$$

The same object in the standard polynomial form corresponding to $\tilde{r} \circ \tilde{p}''$ is (after suppressing, for the same reasons as before, the same summands):

$$\cdots + (a_1^{(\lambda)} + \cdots + a_k^{(\lambda)} + b^{(\lambda)} + c^{(\lambda)}) \times [(b_1^{(\rho)} + \cdots + b_{\nu_\rho}^{(\rho)}) + \cdots + (b_1^{(\omega)} + \cdots + b_{\nu_\omega}^{(\omega)})] + \cdots ; \quad \rho = \kappa_1^\lambda, \cdots \omega = \kappa_{n_\sigma}^\sigma. \tag{2.19}$$

As we have assumed that $a^{(\lambda)} = b^{(\lambda)} + c^{(\lambda)}$, (2.18) and (2.19) are obviously equal. Therefore, either both these coefficients are equal to zero and therefore the corresponding ordered pairs from $\mathbf{C}^1 \times \mathcal{P}'$ do not belong to the standard polynomial forms for $\tilde{p} \circ \tilde{p}''$ and $\tilde{r} \circ \tilde{p}''$ which we denote with $(\tilde{p} \circ \tilde{p}'')\check{}$ and $(\tilde{r} \circ \tilde{p}'')\check{}$, or they are different from zero but equal. In any case, we are led to the conclusion that $(\tilde{p} \circ \tilde{p}'')\check{}$ and $(\tilde{r} \circ \tilde{p}'')\check{}$ are the same, and therefore $\tilde{p} \circ \tilde{p}'' = \tilde{r} \circ \tilde{p}''$. Of course, the equality $a^{(\lambda)} = b^{(\lambda)} + c^{(\lambda)}$ also guarantees that $\tilde{p} = \tilde{r}$.

Comparing (2.14) and (2.15), we easily see that we can construct (in many ways) a finite sequence of polynomial forms $\tilde{p}_1, \tilde{p}_2, \cdots \tilde{p}_n$ such that $\tilde{p}_1 = \tilde{p}'$, $\tilde{p}_n = \check{p}'$, and in which $\tilde{p}_{k+1}$ stays in the same relation to $\tilde{p}_k$ as $\tilde{p}$ stays to $\tilde{r}$ in (2.16). Then, we can write $\tilde{p}_1 \circ \tilde{p}'' = \tilde{p}_2 \circ \tilde{p}'' = \cdots = \tilde{p}_n \circ \tilde{p}''$, i.e., $\tilde{p}' \circ \tilde{p}'' = \check{p}' \circ \tilde{p}''$. In a similar manner, we can show that $\check{p}' \circ \tilde{p}'' = \check{p}' \circ \check{p}''$ and therefore $\tilde{p}' \circ \tilde{p}'' = \check{p}' \circ \tilde{p}'' = \check{p}' \circ \check{p}''$.

Finally, in the last part of Lemma 2, the product $a\tilde{p}$, where

$$\tilde{p} = \{\{(a_1, p_1), \cdots (a_\rho, p_\rho)\}\} = a_1 p_1 \oplus \cdots \oplus a_\rho p_\rho,$$

is defined as

$$a\tilde{p} = \{\{(aa_1, p_1), \cdots (aa_\rho, p_\rho)\}\} = aa_1 p_1 \oplus \cdots \oplus aa_\rho p_\rho. \tag{2.20}$$

The fact that $(a\tilde{p})\check{} = a\check{p}$ is too transparent to be considered in more detail. Q.E.D.

With Lemma 2 proved, it is easy to finish the proof of Lemma 1. Let us show that

$$(A + B) + C = A + (B + C), \qquad A, B, C \in \mathfrak{A}_1'. \tag{2.21}$$

Assume that A, B, C can be represented by the following polynomial forms, in respective order:

$$\tilde{p}_A = a_1 p_1^A \oplus \cdots \oplus a_\rho p_\rho^A,$$
$$\tilde{p}_B = b_1 p_1^B \oplus \cdots \oplus b_\sigma p_\sigma^B,$$
$$\tilde{p}_C = c_1 p_1^C \oplus \cdots \oplus c_\tau p_\tau^C.$$

Then, according to Lemma 2, $(A + B) + C$ will be represented by

$$\begin{aligned}(\tilde{p}_A \oplus \tilde{p}_B) + \tilde{p}_C &= (a_1 p_1^A \oplus \cdots \oplus a_\rho p_\rho^A + b_1 p_1^B \oplus \cdots \oplus b_\sigma p_\sigma^B) \oplus (c_1 p_1^C \oplus \cdots \oplus c_\tau p_\tau^C) \\ &= a_1 p_1^A \oplus \cdots \oplus a_\rho^A p_\rho^A \oplus b_1 p_1^B \oplus \cdots \oplus b_\sigma p_\sigma^B \oplus c_1 p_1^C \oplus \cdots \oplus c_\tau p_\tau^C,\end{aligned}$$

and, by the same manipulations, this last expression can be written in the form $\tilde{p}_A \oplus (\tilde{p}_B \oplus \tilde{p}_C)$ which represents $A + (B + C)$; hence (2.21) is proved.

By similar formal manipulations, which are allowable on grounds of Lemma 2, we can show that, for any $a, b \in \mathbf{C}^1$, $A, B, C \in \mathfrak{A}_1'$, we have:

$$(a + b)A = aA + bA, \tag{2.22}$$

$$a(A + B) = aA + bB, \tag{2.23}$$

$$a(AB) = (aA)B = A(aB), \tag{2.24}$$

$$(A + B)C = AC + BC, \tag{2.25}$$

$$C(A + B) = CA + CB, \tag{2.26}$$

$$(AB)C = A(BC), \tag{2.27}$$

$$A + \mathbf{O} = \mathbf{O} + A = A, \tag{2.28}$$

$$OA = AO = O, \tag{2.29}$$

$$1A = A1 = A. \tag{2.30}$$

The relations (2.9), (2.10), and (2.21)–(2.30) show that $\mathfrak{A}(\mathcal{O}^f)$ is an algebra. Q.E.D.

2.2. The *-algebra $\mathfrak{A}_2(\mathcal{O}^f)$

An operation of involution, $A \to A^*$, can be introduced in $\mathfrak{A}_1(\mathcal{O}^f)$ so that we get an algebra with involution (*-algebra) $\mathfrak{A}_2(\mathcal{O}^f)$ out of $\mathfrak{A}_1(\mathcal{O}^f)$. As is well known, an involution in a complex algebra $\mathfrak{A}$ is defined as an operation which satisfies the following axioms for any $a, b \in \mathbf{C}^1$ and any $A, B \in \mathfrak{A}$:

$$(aA + bB)^* = \bar{a}A^* + \bar{b}B^*, \tag{2.31}$$

$$(A^*)^* = A[\text{we denote } (A^*)^* \text{ by } A^{**}], \tag{2.32}$$

$$(AB)^* = B^*A^*. \tag{2.33}$$

Lemma 3: Introduce the following operation in $\hat{\mathcal{Q}}^f$:

$$(q_1, q_2, \cdots, q_{n-1}, q_n)^* = (q_n, q_{n-1}, \cdots, q_2, q_1). \tag{2.34}$$

If $q_1^{(1)} \circ q_2^{(1)} \circ \cdots \circ q_m^{(1)}$ and $q_1^{(2)} \circ q_2^{(2)} \circ \cdots \circ q_n^{(2)}$ represent the same equivalence class p of $\mathcal{O}^f$, then $q_m^{(1)} \circ \cdots \circ q_2^{(1)} \circ q_1^{(1)}$ and $q_n^{(2)} \circ \cdots \circ q_2^{(2)} \circ q_1^{(2)}$ also represent the same equivalence class of $\mathcal{O}^f$ which will be denoted by p^*.

Proof: If $q_1^{(1)} \circ \cdots \circ q_n^{(1)} = q_1^{(2)} \circ \cdots \circ q_n^{(2)} = p (p \in \mathcal{O}^f)$ and if none of the simple questions $q_1^{(1)}, \cdots q_n^{(1)}, q_1^{(2)}, \cdots q_n^{(2)}$ is equal to one modulo $\mathfrak{s}$, then according to Rules 1 and 2 (Sec. 2.1), there exists a sequence $i_1, \cdots i_r$ of permissible inversions

$$i_s = \begin{pmatrix} \cdots & q_{k_s}, & q_{l_s} & \cdots \\ \cdots & q_{l_s}, & q_{k_s} & \cdots \end{pmatrix}, \tag{2.35}$$

which, when applied in the written order on $(q_1^{(1)}, \cdots q_n^{(1)})$ takes it over into $(q_1^{(2)}, \cdots q_n^{(2)})$.

The inversions

$$i'_s = \begin{pmatrix} \cdots & q_{l_s}, & q_{k_s}, & \cdots \\ \cdots & q_{k_s}, & q_{l_s}, & \cdots \end{pmatrix}, \qquad s = 1, \cdots r, \tag{2.36}$$

are also permissible. The sequence of permissible inversions $i'_r, i'_{r-1} \cdots i'_2, i'_1$ applied in the written order on $(q_n^{(1)}, q_{n-1}^{(1)}, \cdots q_2^{(1)}, q_1^{(1)})$ obviously takes it over into $(q_n^{(2)}, q_{n-1}^{(2)}, \cdots q_2^{(2)}, q_1^{(2)})$.

The general case of ordered sets $(q_1^{(1)}, \cdots q_m^{(1)})$ and $(q_1^{(2)}, \cdots q_n^{(2)})$ belonging to the same equivalence class p can be first, according to Rule 2, Sec. 2.1, reduced to the previous case by eliminating all the questions equal to one modulo $\mathfrak{s}$ and obtaining $(q_1^{(1)\prime}, \cdots q_k^{(1)\prime})(q_1^{(2)\prime}, \cdots q_k^{(2)\prime}) \in p$. Then, according to the above considerations, $(q_k^{(1)\prime}, \cdots q_1^{(1)\prime})$, $(q_k^{(2)\prime}, \cdots q_1^{(2)\prime}) \in p^*$. As we can insert in an ordered product an arbitrary finite number of questions equal to one modulo $\mathfrak{s}$, at arbitrary places, we conclude that $(q_m^{(1)}, \cdots q_1^{(1)})$, $(q_n^{(2)}, \cdots q_1^{(2)}) \in p^*$. Q.E.D.

Lemma 4: We define the following operation in $\mathfrak{R}(\mathcal{O}^f)$: If $\tilde{p} = a_1p_1 \oplus \cdots \oplus a_np_n$, $a_1, \cdots a_n \in \mathbf{C}^1$, $p_1, \cdots p_n \in \mathcal{O}^f$, then $p^* = \bar{a}_1p_1^* + \cdots + \bar{a}_np_n^*$, where p_k^*, $k = 1, \cdots n$ is defined in Lemma 3.

If $\tilde{p}' = \tilde{p}'' = A[A \in \mathfrak{A}_1(\mathcal{O}^f)]$, then $\tilde{p}'^*$ and $\tilde{p}''^*$ represent the same equivalence class which will be denoted by A^*, i.e., $\tilde{p}'^* = \tilde{p}''^* = A^*$. The so defined *-operation in $\mathfrak{A}_1(\mathcal{O}^f)$ satisfies the relations (2.31), (2.32), (2.33) and therefore can be used to define out of $\mathfrak{A}_1(\mathcal{O}^f)$ an algebra with involution (*-algebra) which will be denoted by $\mathfrak{A}_2(\mathcal{O}^f)$.

Proof: If $\tilde{p}'$ and $\tilde{p}''$ are written in the form (2.14), then, on the basis of Lemma 3, we can conclude from

$$p_1^{(1)} = \cdots = p_{\mu_1}^{(1)} = p^{(1)}, \cdots r_1^{(\tau)} = \cdots = r_{\nu_\tau}^{(\tau)} = r^{(\tau)}$$

that

$$p_1^{(1)*} = \cdots = p_{\mu_1}^{(1)*} = p^{(1)*}, \cdots r_1^{(\tau)*} = \cdots r_{\nu_\tau}^{(\tau)} = r^{(\tau)}.$$

We also have that $p^{(i)*} \neq p^{(j)*}$ for $i \neq j$, because if $p^{(i)*} = p^{(j)*}$ for some $i \neq j$ then, by applying Lemma 3 again, we get $(p^{(i)*})^* = (p^{(j)*})^*$; from relation (2.34) it is obvious that $(q^*)^* = q$ for any $q \in \hat{\mathcal{Q}}^f$ and therefore $(p^*)^* = p$ for any $p \in \mathcal{O}^f$; hence $p^{(i)*} = p^{(j)*}$, $i \neq j$ implies $p^{(i)} = p^{(j)}$, $i \neq j$ which contradicts the definitions of the $p^{(i)}$'s. A similar conclusion is valid for the r's. This enables us to deduce that $\tilde{p}'^*$ and $\tilde{p}''^*$ have the same standard polynomial form.

Very simple algebraic manipulations enable us to verify (2.31), (2.32), and (2.33); e.g., to verify (2.31), choose $\tilde{p}_A, \tilde{p}_B \in \mathfrak{R}(\mathcal{O}^f)$ such that $\tilde{p}_A = A$, $\tilde{p}_B = B$. Then if

$$\tilde{p}_A = a_1p_1^A \oplus \cdots \oplus a_mp_m^A,$$

$$\tilde{p}_B = b_1p_1^B \oplus \cdots \oplus b_np_n^B,$$

we have

$$\begin{aligned}
&(a\tilde{p}_A + b\tilde{p}_B)^* \\
&= [aa_1p_1^A \oplus \cdots \oplus aa_mp_m^A \oplus bb_1p_1^B \oplus \cdots \oplus bb_np_n^B]^* \\
&= \bar{a}\bar{a}_1(p_1^A)^* \oplus \cdots \oplus \bar{a}\bar{a}_m(p_m^A)^* \\
&\quad \oplus \bar{b}\bar{b}_1(p_1^B)^* \oplus \cdots \oplus \bar{b}\bar{b}_n(p_n^B)^* \\
&= \bar{a}(\bar{a}_1(p_1^A)^* \oplus \cdots \oplus \bar{a}_m(p_m^A)^*) \\
&\quad \oplus \bar{b}(\bar{b}_1(p_1^B)^* \oplus \cdots \oplus \bar{b}_n(p_n^B)^* \\
&= \bar{a}\tilde{p}_A^* + \bar{b}\tilde{p}_B^*.
\end{aligned}$$

The last expression represents $\bar{a}A^* + \bar{b}B^*$. Q.E.D.

2.3. The normed *-algebra $\mathfrak{A}_3(\mathcal{O}^f)$

To each polynomial form on $\mathcal{P}^f$,

$$\tilde{p} = a_1 p_1 \oplus \cdots \oplus a_n p_n,$$
$$a_1, \cdots a_n \in \mathbf{C}^1, \qquad p_1, \cdots p_n \in \mathcal{P}^f, \tag{2.37}$$

corresponds a whole class of *polynomial forms on* $\mathfrak{Q}(\mathcal{O}^f)$, or *polynomial forms of simple questions relative to* $\mathcal{O}^f$:

$$p = a_1 q_1' + \cdots + a_n q_n',$$
$$a_1, \cdots a_n \in \mathbf{C}^1, \qquad q_1', \cdots q_n' \in \mathcal{P}^f, \tag{2.38}$$

obtained from (2.37) by replacing $p_1, \cdots p_n$ with $q_1' = q_1^{(1)} \circ \cdots \circ q_{k_1}^{(1)}, \cdots q_n' = q_1^{(n)} \circ \cdots \circ q_{k_n}^{(n)}$, where $q_1', \cdots q_n'$ are any ordered products representing the equivalence classes $p_1, \cdots p_n$, respectively.

Each of the questions $q_1^{(1)}, \cdots q_n^{(k_n)}$ is simple and can be considered to be functions of $\alpha_1^{(1)}, \cdots \alpha_n^{(k_n)} \in \mathcal{O}^f$ respectively, with corresponding minimal Borel sets $B_1^{(1)}, \cdots B_n^{(k_n)}$. We denote with $\chi_j^{(i)}(\lambda)$ $(\lambda \in \mathbf{R}^{(\alpha_i^{(j)})}$, $i = 1, \cdots n$, $j = 1, \cdots k_i)$ the characteristic function of the set $B_j^{(i)}$. To the polynomial form (2.38) we can now adjoin a polynomial $p(\hat{\lambda}_1, \cdots \hat{\lambda}_n)$ in the real variables $\hat{\lambda}_1 = (\lambda_1^{(1)}, \cdots \lambda_{k_1}^{(1)})$, $\cdots \hat{\lambda}_n = (\lambda_1^{(n)}, \cdots \lambda_{k_n}^{(n)})$ having the form:

$$p(\hat{\lambda}_1, \cdots \hat{\lambda}_n) = a_1 \chi_1^{(1)}(\lambda_1^{(1)}) \cdots \chi_{k_1}^{(1)}(\lambda_{k_1}^{(1)}) + \cdots + a_n \chi_1^{(n)}(\lambda_1^{(n)}) \cdots \chi_{k_n}^{(n)}(\lambda_{k_n}^{(n)}). \tag{2.39}$$

We are now able to adjoin to each physical state $P \in \mathcal{S}$ a functional F_P on polynomial forms of simple questions by the formula

$$F_P[p(q_1', \cdots q_n')] = \int p(\hat{\lambda}_1, \cdots \hat{\lambda}_n)\, dP^{\hat{\alpha}_1; \cdots; \hat{\alpha}_n}(\hat{\lambda}_1, \cdots \hat{\lambda}_n); \tag{2.40}$$

here $\hat{\alpha}_i = (\alpha_1^{(i)}, \cdots \alpha_{k_i}^{(i)})$ while $p(\hat{\lambda}_1, \cdots \hat{\lambda}_n)$ is defined by (2.39) and the rest of the symbols are the ones already introduced at the beginning of this subsection.

We have first to establish that (2.40) adjoins a unique value to each polynomial form, $p(q_1', \cdots q_n')$, on $\mathfrak{Q}^f$. Namely, the step from (2.38) to (2.39) might not be unique because it might happen that, for a question $q_i^{(i)}$, there exist two different $\alpha, \beta \in \mathcal{O}^f, \alpha \neq \beta$, such that $q_i^{(i)} = \chi_{B'}^{\alpha}(\alpha) = \chi_{B''}^{\beta}(\beta)$. Do we have in such a case (the uninteresting terms in the polynomials, which are common to both of them are not explictly written)

$$F_P[\cdots + a_i \chi_1^{(i)}(\alpha_1^{(i)}) \cdots \chi_{B'}^{\alpha}(\alpha_i^{(i)}) \cdots \chi_{k_i}^{(i)}(\alpha_{k_i}^{(i)}) + \cdots] = F_P[\cdots + a_i \chi_1^{(i)}(\alpha_1^{(i)}) \cdots \chi_{B''}^{\beta}(\beta_i^{(i)}) \cdots \chi_{k_i}^{(i)}(\alpha_{k_i}^{(i)}) \times \cdots]? \tag{2.41}$$

The answer is—yes! To see that, insert (2.39) in (2.40). We obtain

$$\int p(\hat{\lambda}_1, \cdots \hat{\lambda}_n)\, dP^{\hat{\alpha}_1; \cdots; \hat{\alpha}_n}(\hat{\lambda}_1, \cdots \hat{\lambda}_n) = a_1 P^{\alpha_1^{(1)}; \cdots; \alpha_{k_1}^{(1)}}(B_1^{(1)} \times \cdots \times B_{k_1}^{(1)}) + \cdots + a_n \times P^{\alpha_1^{(n)}; \cdots; \alpha_{k_n}^{(n)}}(B_1^{(n)} \times \cdots \times B_{k_n}^{(n)}). \tag{2.42}$$

Written with the help of (2.42), relation (2.41) becomes, after ignoring the unessential common terms (it is assumed that $a_i \neq 0$)

$$P^{\alpha_1^{(i)}; \cdots; \alpha; \cdots; \alpha_{k_i}^{(i)}}(B_1^{(i)} \times \cdots \times B' \times \cdots \times B_{k_i}^{(i)}) = P^{\alpha_1^{(i)}; \cdots; \beta; \cdots; \alpha_{k_i}^{(i)}}(B_1^{(i)} \times \cdots \times B'' \times \cdots \times B_{k_i}^{(i)}). \tag{2.43}$$

But according to I, Sec. 2.2, Proposition 3, relation (2.43) is necessary for the validity of the equation $\chi_{B'}^{\alpha}(\alpha) = \chi_{B''}^{\beta}(\beta)$. Hence, we can indeed consider F_P to be a functional of polynomial forms on $\mathfrak{Q}^f$.

Lemma 5: The functional F_P defined by (2.40) on polynomial forms on $\mathfrak{Q}^f$ has the same value on all polynomial forms of simple questions which belong to the same element of $\mathfrak{K}^f$. Hence, a functional

$$\langle A \rangle_P, \qquad A \in \mathfrak{A}_2(\mathcal{O}^f), \qquad P \in \mathcal{S} \tag{2.44}$$

can be defined by the formula

$$\langle A \rangle_P = F_P[\tilde{p}(q_1', \cdots q_n')], \quad q_1', \cdots q_n' \in \mathcal{P}^f, \tag{2.45}$$

where $\tilde{p}(q_1', \cdots q_n')$ is any polynomial form on $\mathcal{P}^f$ which is equal to A. This functional is linear and normalized.

Proof: The first step is to show that

$$F_P[a q_1 \circ \cdots \circ q_n]$$

has the same value for all ordered products $q_1 \circ \cdots \circ q_n$ representing the same equivalence class $p \in \mathcal{P}^f$. According to (2.42), if $q_1 = \mathcal{P}_{B_1}^{\alpha_1}(\alpha_1), \cdots q_n = \chi_{B_n}^{\alpha_n}(\alpha_n)$, then $(a \in \mathbf{C}^1)$:

$$F_P[a q_1 \circ \cdots \circ q_n] = a P^{\alpha_1; \cdots; \alpha_n}(B_1 \times \cdots \times B_n). \tag{2.46}$$

Axiom I(3) in I requires that if $\{\alpha, \beta\} = C$, then

$$P^{\cdots; \alpha; \beta; \cdots}(\cdots \times B' \times B'' \times \cdots) = P^{\cdots; \beta; \alpha; \cdots}(\cdots \times B'' \times B' \times \cdots). \tag{2.47}$$

Relation (2.47) shows that (2.46) has the same value on all ordered products equivalent according to Rule 1, Sec. 2.1; the relation (2.3) in I, Sec. 2.2 shows that (2.46) has the same value on all ordered products equivalent according to Rule 2, Sec. 2.1. Hence, the first step has been proved.

Next, we have to show that $F_P = 0$ on all elements of the zero class of ordered pairs form $\mathbf{C}^1 \times \mathcal{O}'$ (Sec. 2.1, Rule 3). Again, from (2.42), we get $F_P[aq_1 \circ \cdots \circ q_n] = 0$ for $a = 0$, because $P^{\alpha_1 \cdots \alpha_n}$ is always finite. If $q_1 \circ \cdots \circ q_n$ represents the zero equivalence class of $\mathcal{O}'$, then, according to definition, there must be at least one zero simple question q_j in $q_1 \circ \cdots \circ q_n$. But then, according to the definition of a zero question in I, Sec. 3.2, q_j is a zero question if and only if B_j is a zero set. Therefore,

$$P^{\alpha_1;\cdots;\alpha_j;\cdots;\alpha_n}(B_1 \times \cdots \times B_j \times \cdots \times B_n) = 0.$$

By showing that F_P is zero for the zero class of ordered pairs from $\mathbf{C}^1 \times \mathcal{O}'$, we have simultaneously established that the value of F_P on polynomial forms on $\mathcal{O}'$ (which differ from their standard polynomial form on $\mathcal{O}'$ only by a finite number of ordered pairs from the zero class of $\mathbf{C}^1 \times \mathcal{O}'$) is equal to its value on this standard polynomial.

To see that the value of F_P on a polynomial form $\tilde{p}$ on $\mathcal{O}'$ is equal to the value of F_P on the standard polynomial form $\check{p}$, we insert $\tilde{p}$ of (2.1) into (2.40) to get

$$\begin{aligned} F_P(\tilde{p}) &= \sum_{\nu=1}^{\sigma} (a_1^{(\nu)} F_P[1p_1^{(\nu)}] + \cdots + a_{\rho_\nu}^{(\nu)} F_P[1p_{\rho_\nu}^{(\nu)}]) \\ &= \sum_{\nu=1}^{\sigma} (a_1^{(\nu)} + \cdots + a_{\rho_\nu}^{(\nu)}) F_P[1p^{(\nu)}] \\ &= \sum_{\nu=1}^{\sigma} a_\nu F_P[1p^{(\nu)}]. \end{aligned} \tag{2.48}$$

In the same fashion, we can show that

$$\sum_{\nu=1}^{\sigma} a^{(\nu)} F_P[1p^{(\nu)}]$$

is the value of F_P on the polynomial form (2.2), which differs from $\check{p}$ only by zero class ordered pairs from $\mathbf{C}^1 \times \mathcal{O}'$. Hence $F_P[\tilde{p}] = F_P[\check{p}]$, and the first part of Lemma 4 is proved.

To show that $\langle A \rangle_P$ is linear, i.e., that

$$\langle aA + bB \rangle_P = a\langle A \rangle_P + b\langle B \rangle_P, \qquad a, b \in \mathbf{C}^1, \quad A, B \in \mathfrak{A}_2', \tag{2.49}$$

we pick up any polynomial forms $\tilde{p}' = A$, $\tilde{p}'' = B$. According to lemmas 1 and 2, $a\tilde{p}' \oplus b\tilde{p}'' = aA + bB$. Therefore, we have:

$$\begin{aligned} \langle aA + bB \rangle_P &= F_P[a\tilde{p}' \oplus b\tilde{p}''] = aF_P[\tilde{p}'] \\ &\quad + bF_P[\tilde{p}''] = a\langle A \rangle_P + b\langle B \rangle_P, \end{aligned}$$

where the second step is easily derivable from (2.40).

Finally, to show that $\langle A \rangle_P$ is normalized, i.e., that $\langle \mathbf{1} \rangle_P = 1$, take any function $\chi^\alpha(\lambda) \equiv 1 \pmod{\mathcal{S}}$ for some $\alpha \in \mathcal{O}'$. By definition, $1\chi^\alpha(\alpha) \in 1$. From (2.40) and (2.45) we get

$$\langle \mathbf{1} \rangle_P = \int \chi^\alpha(\lambda)\, dP^\alpha(\lambda) = 1. \qquad \text{Q.E.D.}$$

As has been pointed out in Sec. 1, our ultimate interest is to get a representation of the language $\mathfrak{L}$ introduced in I, Sec. 2.2. This requires the embedding of $\mathfrak{L}$ in a C^*-algebra $\mathfrak{B}(\mathfrak{H})$ of bounded operators on a Hilbert space $\mathfrak{H}$, and therefore $\mathfrak{A}_2'$ has to be isomorphic to a subalgebra $\mathfrak{B}'$ of $\mathfrak{B}(\mathfrak{H})$. This implies that each physical state will be represented by a positive bounded operator U on $\mathfrak{H}$ with $TrU = 1$ (I, Sec. 2.3): Each such U determines a positive functional[8] on $\mathfrak{B}'$:

$$F_U(A) = Tr(UA), \; A \in \mathfrak{B}'. \tag{2.50}$$

After taking in consideration relation (2.8) in I, it is clear that, for an element A of $\mathfrak{B}'$ which corresponds to an element of $\mathfrak{A}_2'$, the functionals (2.40) and (2.50) should yield the same values when U represents the physical state P. This indicates that we have to require that the linear functionals $\langle A \rangle_P (A \in \mathfrak{A}_2', P \in \mathcal{S})$ on $\mathfrak{A}_2'$, introduced in lemma 5, should be positive.

Lemma 6: A necessary and sufficient condition for the linear functional $\langle A \rangle_P$, $A \in \mathfrak{A}_2'$, to be positive, i.e.,

$$\langle A^* A \rangle_P \geq 0, \qquad A \in \mathfrak{A}_2(\mathcal{O}'), \tag{2.51}$$

is that the determinant of any $n \times n$ matrix $(n = 1, 2, 3, \cdots)$ having for its (i, j) element

$$\begin{aligned} &F_P[q_{k_i}^{(i)} \circ \cdots \circ q_1^{(i)} \circ q_1^{(j)} \circ \cdots \circ q_{k_j}^{(j)}], \\ &q_1^{(1)}, \cdots q_{k_i}^{(1)}, \cdots q_1, \cdots q_{k_j} \in Q(\mathcal{O}'), \end{aligned} \tag{2.52}$$

is non-negative, and the matrix itself Hermitian.

Proof: Any polynomial form on $\mathfrak{Q}'$, having the general form (2.38), belongs to an element A of $\mathfrak{A}_2'$, and according to Lemma 4

$$\begin{aligned} &p(q_1', \cdots q_n')^* \\ &= \bar{a}_1 q_{k_1}^{(1)} \circ \cdots \circ q_1^{(1)} \oplus \cdots \oplus \bar{a}_n q_{k_n}^{(n)} \circ \cdots \circ q_1^{(n)} \end{aligned} \tag{2.53}$$

belongs to A^*. According to the definition (2.45) in Lemma 5, we have

$$\begin{aligned} \langle A^* A \rangle_P &= F_P\Big[\bigoplus_{i,j=1}^{n} \bar{a}_i a_j \\ &\quad \times q_{k_i}^{(i)} \circ \cdots \circ q_1^{(i)} \circ q_1^{(j)} \circ q_2^{(j)} \circ \cdots \circ q_{k_j}^{(j)}\Big] \\ &= \sum_{i,j} \bar{a}_i a_j F_P[q_{k_i}^{(i)} \circ \cdots \circ q_1^{(i)} \circ q_1^{(j)} \circ \cdots \circ q_{k_j}^{(j)}]. \end{aligned} \tag{2.54}$$

[8] J. von Neumann, *Mathematical Foundations of Quantum Mechanics*, translated by R. T. Beyer (Princeton University Press, Princeton, New Jersey, 1955), p. 318.

We have on the right-hand side of (2.54) a bilinear form in $a_1, \cdots a_n \in \mathbf{C}^1$, which has to be positive definite or semi-definite if $\langle A\rangle_P$ is a positive functional. Such a bilinear form is positive definite or semi-definite if and only if the corresponding matrix, i.e., in our case the matrix whose (i, j) element is given by (2.52), is *Hermitian* positive definite or positive semi-definite. As is well known from linear algebra, the sufficient and necessary condition for Hermitian matrix to be semi-definite is that all its principal minors are nonnegative[9] which is exactly the condition expressed by Lemma 6, Q.E.D.

Lemma 6 can be expressed in a different way which has the advantage of having the form of a restriction imposed on, *and only on*, the set $\mathcal{S}$ of physical states (and, implicitly, on the concept of a fundamental set $\mathcal{O}^f$ of observables). This condition, which is sufficient and necessary for the linear functionals $\langle A\rangle_P$, $P \in \mathcal{S}$ to be positive, is given by the following hypothesis:

Hypothesis 1: The complex probability measures $P^{\hat{\alpha}}$, $\hat{\alpha} \in \hat{\mathcal{O}}^f$, for a given fundamental set $\mathcal{O}^f$ of observables are such that the following three conditions are fulfilled for any $\hat{\alpha} = (\alpha_1, \cdots \alpha_k) \in \hat{\mathcal{O}}^f$, $B_i = B_i^{(1)} \times \cdots \times B_i^{(k)}$ $(B_i^{(1)}, \cdots B_i^{(k)} \in \mathcal{B}^1, i = 1, \cdots n)$:

a) $$P^{\hat{\alpha}}(B_i) = \overline{P^{\hat{\alpha}^*}(B_i^*)}, \qquad i = 1, \cdots n. \tag{2.55a}$$

We used the notation $\hat{\alpha}^* = (\alpha_k, \cdots \alpha_1)$ and $B^* = B_k \times \cdots \times B_1$ if

$$\hat{\alpha} = (\alpha_1, \cdots \alpha_k), B = B_1 \times \cdots \times B_k.$$

b)

$$\begin{vmatrix} P^{\hat{\alpha}^*;\hat{\alpha}}(B_1^*\times B_1) & P^{\hat{\alpha}^*;\hat{\alpha}}(B_1^*\times B_2)\cdots P^{\hat{\alpha}^*;\hat{\alpha}}(B_1^*\times B_n) \\ P^{\hat{\alpha}^*;\hat{\alpha}}(B_2^*\times B_1) & P^{\hat{\alpha}^*;\hat{\alpha}}(B_2^*\times B_2)\cdots P^{\hat{\alpha}^*;\hat{\alpha}}(B_2^*\times B_n) \\ \cdots & \cdots \\ P^{\hat{\alpha}^*;\hat{\alpha}}(B_n^*\times B_1) & P^{\hat{\alpha}^*;\hat{\alpha}}(B_n^*\times B_2)\cdots P^{\hat{\alpha}^*;\hat{\alpha}}(B_n^*\times B_n) \end{vmatrix} \geq 0 \tag{2.55b}$$

c) If the ordered products

$$\chi_{B_i^{(1)}}^{\alpha_1}(\alpha_1) \circ \cdots \circ \chi_{B_i^{(n)}}^{\alpha_n}(\alpha_n)$$

over $\mathcal{P}(\mathcal{O}^f)$ are all different from one another elements of $\mathcal{P}(\mathcal{O}^f)$, which do not belong to the zero class of $\mathcal{P}^f$, then there is at least one physical state $P \in \mathcal{S}$ for which the determinant (2.55b) is greater than zero.

From here on, Hypothesis 1 will be assumed throughout. The reason why it has been stated at this place, and not as one of the axioms, is that, on an intuitive level, this hypothesis seems to be more ad hoc and arbitrary than the rest of the axioms. It has to be noted that the axioms of I, Sec. 2.2 already define a sufficiently rich language, in terms of which dynamical laws of the type of a Schrödinger equation, and hence a complete theory, can be formulated. The introduction of Hypothesis 1, and later of Hypothesis 2, serves the only purpose of reducing such a language to the conventional Hilbert space language.

In introducing a norm in $\mathfrak{A}_2^f$, we will again follow the analogy with the Hilbert space formalism. If we deal with a theory without super-selection rules, the norm $||A||$ of any bounded operator A representing an observable is given by

$$\begin{aligned} ||A|| &= \sup_{\substack{\Psi\in\mathcal{H} \\ ||\Psi||=1}} (\langle A\Psi \mid A\Psi\rangle)^{\frac{1}{2}} \\ &= \sup_{\substack{\Psi\in\mathcal{H} \\ ||\Psi||=1}} (\langle\Psi|\ A^*A\ |\Psi\rangle)^{\frac{1}{2}}. \end{aligned} \tag{2.56}$$

In case superselection rules are present, not every Hilbert space vector represents a physical state; the set $\mathcal{S}_0$ of vectors representing pure states is (using the notation of I, Sec. 2.3)

$$\mathcal{S}_0 = \{\Psi\ ; \Psi \in \mathcal{H}_j, j \in J\}, \tag{2.57}$$

where $\mathcal{H}_j$ can be any of the coherent subspaces [I, (2.11)]. The norm of a bounded operator is again defined by (2.56). Any normalized vector $\Psi \in \mathcal{H}$ can be written in the form

$$\begin{aligned} &\Psi = a_1\Psi_{j_1} + a_2\Psi_{j_2} + \cdots, \\ &\Psi_{j_1} \in \mathcal{H}_{j_1}, \quad \Psi_{j_2} \in \mathcal{H}_{j_2}, \cdots, \\ &\qquad ||\Psi_{j_1}|| = ||\Psi_{j_2}|| = \cdots = 1, \\ &\qquad |a_{j_1}|^2 + |a_{j_2}|^2 + \cdots = 1. \end{aligned} \tag{2.58}$$

Hence, in case that a bounded operator A represents an observable, it leaves each coherent space $\mathcal{H}_j$, $j \in J$ invariant, and we have

$$\begin{aligned} \langle A\Psi \mid A\Psi\rangle &= |a_{j_1}|^2 \langle A\Psi_{j_1} \mid A\Psi_{j_1}\rangle \\ &\quad + |a_{j_2}|^2 \langle A\Psi_{j_2} \mid A\Psi_{j_2}\rangle + \cdots \\ &\leq (|a_{j_1}|^2 + |a_{j_2}|^2 + \cdots) \max_{k=1,2,\cdots} \langle A\Psi_{j_k} \mid A\Psi_{j_k}\rangle \\ &\leq \max_{k=1,2,\cdots} \langle A\Psi_{j_k} \mid A\Psi_{j_k}\rangle. \end{aligned} \tag{2.59}$$

Inequality (2.59) enables us to assert that the norm of a bounded operator A representing an observable is in general given by $(||\Psi|| = 1)$:

[9] L. Mirsky, *An Introduction to Linear Algebra* (Clarendon Press, Oxford, England, 1955), Theorem 13.3.5., p. 405.

$$||A|| = \sup_{j\in J}\sup_{\Psi\in\mathcal{K}_j}(\langle A\Psi \mid A\Psi\rangle)^{\frac{1}{2}} = \sup_{\Psi\in\mathcal{S}_0}(\langle A\Psi \mid A\Psi\rangle)^{\frac{1}{2}}. \tag{2.60}$$

After taking into consideration the fact that, in our language, $\langle A\Psi|A\Psi\rangle = \langle\Psi|A^*A|\Psi\rangle = \langle A^*A\rangle_P$, where Ψ represents the pure physical state $P \in \mathcal{S}_0$, we will attempt to introduce

$$||A|| = \sup_{P\in\mathcal{S}_0}(\langle A^*A\rangle_P)^{\frac{1}{2}} \tag{2.61}$$

as a norm in $\mathfrak{A}_2(\mathcal{O}^f)$.

Lemma 7: The functional

$$||A|| = \sup_{P\in\mathcal{S}}(\langle A^*A\rangle_P)^{\frac{1}{2}}, \qquad A \in \mathfrak{A}_2(\mathcal{O}^f) \tag{2.62}$$

is convex, symmetric, and normalized:

$$||A|| \geq 0; \quad ||A|| = 0 \text{ if and only if } A = \mathbf{0}, \tag{2.63}$$

$$||aA|| = |a|\,||A||, \qquad a \in \mathbf{C}^1, \tag{2.64}$$

$$||A + B|| \leq ||A|| + ||B||, \tag{2.65}$$

$$||\mathbf{1}|| = 1. \tag{2.66}$$

Proof: The first part of relation (2.63) is a direct consequence of the definition (2.62) and of the fact that the linear functionals $\langle A\rangle_P$ are positive. The fact that $||A|| = 0$ if and only if $A = \mathbf{0}$ is an obvious consequence of Hypotheses 1c. In the case of our positive functionals $\langle A\rangle_P$, $P \in \mathcal{S}$, this inequality has the form

$$|\langle A^*B\rangle_P|^2 \leq \langle A^*A\rangle_P\,\langle B^*B\rangle_P \tag{2.67}$$

for any $A, B \in \mathfrak{A}_2(\mathcal{O}^f)$.

By using (2.67) and (2.31), we get

$$\begin{aligned}\langle (A+B)^*(A+B)\rangle_P &= \langle A^*A + A^*B + B^*A + B^*B\rangle_P \\ &\leq \langle A^*A\rangle_P + |\langle A^*B\rangle_P| + |\langle B^*A\rangle_P| + \langle B^*B\rangle_P \\ &\leq \langle A^*A\rangle_P + (\langle A^*A\rangle_P\,\langle B^*B\rangle_P)^{\frac{1}{2}} \\ &\quad + (\langle B^*B\rangle_P\,\langle A^*A\rangle_P)^{\frac{1}{2}} + \langle B^*B\rangle_P.\end{aligned}$$

This can be written in the form

$$\langle (A+B)^*(A+B)\rangle_P \leq ((\langle A^*A\rangle_P)^{\frac{1}{2}} + (\langle B^*B\rangle_P)^{\frac{1}{2}})^2$$

which immediately yields (2.65).

The last relation, (2.66), is a direct consequence of the fact that all functionals $\langle A\rangle_P$ are normalized. Q.E.D.

At this stage, we note that definition (2.62) is different from (2.61), because, in (2.62), the supremum is taken over the entire $\mathcal{S}$ and not only over the set $\mathcal{S}_0$ of pure physical states. In Sec. 3, it will be shown that these two definitions are equivalent.

Lemma 8: If the functional (2.62) satisfies the inequality

$$||AB|| \leq ||A||\,||B||, \qquad A, B \in \mathfrak{A}_2(\mathcal{O}^f) \tag{2.68}$$

for any $A, B \in \mathfrak{A}_2(\mathcal{O}^f)$, then it has also the following two properties:

$$||A^*|| = ||A||, \qquad A \in \mathfrak{A}_2(\mathcal{O}^f), \tag{2.69}$$

$$||A^*A|| = ||A||^2, \qquad A \in \mathfrak{A}_2(\mathcal{O}^f). \tag{2.70}$$

Proof: To prove (2.69), we first derive from (2.68):

$$||A^*A|| \leq ||A^*||\,||A||. \tag{2.71}$$

Using the Cauchy-Schwartz inequality (2.67) we get:

$$|\langle A^*A\rangle_P|^2 = |\langle \mathbf{1}(A^*A)\rangle_P|^2 \leq \langle \mathbf{1}^*\mathbf{1}\rangle_P\,\langle (A^*A)^*A^*A\rangle_P. \tag{2.72}$$

However, very easy algebraic manipulations show that $\mathbf{1}^* = \mathbf{1}$ and therefore $\mathbf{1}^*\mathbf{1} = \mathbf{1}^2 = \mathbf{1}$. Hence from (2.72) we get

$$|\langle A^*A\rangle_P|^2 \leq \langle (A^*A)^*\,(A^*A)\rangle_P,$$

and therefore

$$\begin{aligned}||A||^2 &= \sup_{P\in\mathcal{S}}\langle A^*A\rangle_P \leq \sup_{P\in\mathcal{S}}(\langle (A^*A)^*\,(A^*A)\rangle)^{\frac{1}{2}} \\ &= ||A^*A||.\end{aligned} \tag{2.73}$$

Combining (2.71) and (2.73), we obtain

$$||A||^2 \leq ||A||\,||A^*||; \tag{2.74}$$

interchanging the roles of A and A^*, we also get

$$||A^*||^2 \leq ||A^*||\,||A||. \tag{2.75}$$

Hence, if $||A|| = 0$ then $||A^*|| = 0$, and vice versa. If both $||A||$ and $||A^*||$ are different from zero, we conclude $||A^*|| \leq ||A||$, $||A|| \leq ||A^*||$, and therefore (2.69) is true in general.

We can now write (2.71) in the form

$$||A^*A|| \leq ||A||^2. \tag{2.76}$$

By combining (2.73) and (2.76) we obtain (2.70). Q.E.D.

From Lemma 8, we see that, if we would like to construct out of $\mathfrak{A}_2^f$ a normed *-algebra with the norm (2.62), then we have to assume that the structure of the set $\mathcal{S}$ of all physical states is such that the relation

$$||AB||^2 \leq ||A||^2\,||B||^2 \tag{2.77}$$

is valid for any $A, B \in \mathfrak{A}_2^f$.

In our later consideration of the next section, we

will need a hypothesis from which relation (2.77) can be derived. We immediately arrive at this hypothesis if we consider the way in which (2.77) is derived in the Hilbert-space formalism, i.e., when A and B are bounded operators representing observables on a Hilbert space $\mathfrak{H}$. Namely, if $\Psi \in \mathfrak{H}$, then we can write

$$\begin{aligned}\langle\Psi|\,(AB)^*\,(AB)\,|\Psi\rangle &= \langle B\Psi|\,A^*A\,|B\Psi\rangle \\ &\leq ||A||^2\,\langle\Psi|\,B^*B\,|\Psi\rangle.\end{aligned} \tag{2.78}$$

It is easy to see that, if we know that relation (2.77) is satisfied for any Hilbert vector Ψ representing a pure physical state, then we can immediately derive (2.77) from the very definition of the norm of an operator. But relation (2.78) is a direct consequence of the essential characteristic of the Hilbert-space formalism, which is true even in the case when superselection rules are present: if $\Psi \in \mathfrak{H}$ represents a physical state and A is a bounded operator of the form

$$\textstyle\sum_{i=1}^{n} a_i A_1^{(i)} \cdots A_{k_i}^{(i)},$$

where $A_1^{(i)}, \cdots, A_{k_i}^{(i)}$ represent observables, and such that $\langle A\Psi|A\Psi\rangle \neq 0$, then the normed vector $(\langle A\Psi|A\Psi\rangle)^{-\frac{1}{2}}A\Psi$ also represents a physical state. This induces us to introduce the following assumption:

Hypothesis 2A: If the element $A \in \mathfrak{A}_2(\mathfrak{O}')$ is such that $\langle A^*A\rangle_P \neq 0$, then the positive normed functional

$$F(B) = (\langle A^*A\rangle_P)^{-1}\,\langle A^*BA\rangle_P \tag{2.79}$$

corresponds to a physical state P_1, i.e.,

$$F(B) = \langle B\rangle_{P_1}. \tag{2.80}$$

To see that Hypothesis 2A implies (2.78) and hence (2.77), we have to write (2.79) and (2.80) in the form

$$\langle A(B^*B)A\rangle_P = \langle A^*A\rangle_P\,\langle B^*B\rangle_{P_1}. \tag{2.81}$$

It is naturally possible to express this hypothesis in an equivalent form, employing only algebraic relations between the complex probability measures defining physical states.

Hypothesis 2B: Any fundamental set $\mathfrak{O}'$ of observables is such that, if, for some complex numbers $a_1, \cdots a_n$, for some $\hat{\beta} \in \hat{\mathfrak{O}}'$, $B_1, \cdots B_n \in \mathfrak{B}^{\hat{\beta}}$, and for a given $P \in \mathfrak{S}$ the expression

$$Z = \sum_{i,j=1}^{n} \bar{a}_i a_j P^{\hat{\beta}^*;\hat{\beta}}(B_i^* \times B_j) \neq 0, \tag{2.82}$$

then the complex probability measures

$$P_1^{\hat{\alpha}}(B) = Z^{-1} \sum_{i,j\in 1}^{n} \bar{a}_i a_j P^{\hat{\beta}^*;\hat{\alpha};\hat{\beta}}(B_i^* \times B \times B_j), \qquad \hat{\alpha} \in \hat{\mathfrak{O}}', \quad B \in \mathfrak{B}^{\hat{\alpha}}, \tag{2.83}$$

correspond to a physical state $P_1 \in \mathfrak{S}$.

It is, indeed, very easy to check that the expressions (2.83) define complex probability measures satisfying all the requirements of Ax. I, except for the points 3*a* and 3*b*. To verify that these two points are satisfied too, we have to make use of Hypothesis 1, and then the verification is carried out easily. Naturally, Hypothesis 2B does not affect the rest of the axioms in I, Sec. 2.2.

It is also obvious from the construction of the algebra $\mathfrak{A}_2'$ that Hypothesis 2B is equivalent to Hypothesis 2A.

Both introduced hypotheses can be replaced by the following single axiom, which is evidently equivalent to these two hypotheses taken together.

Axiom X

a) For any $\hat{\alpha} = (\alpha_1, \cdots \alpha_n) \in \hat{\mathfrak{O}}'$ and any $B \in \mathfrak{B}^{(\alpha_1)} \times \cdots \times \mathfrak{B}^{(\alpha_n)}$, we have for any $P \in \mathfrak{S}$[10]:

$$P^{\hat{\alpha}}(B) = \overline{P^{\hat{\alpha}^*}(B^*)}. \tag{2.84}$$

b) Given any $\hat{\alpha}, \hat{\beta} \in \hat{\mathfrak{O}}'$, $\hat{\beta} = (\beta_1, \cdots \beta_n)$, any $B \in \mathfrak{B}^{\hat{\alpha}}$, $B_1, \cdots B_k \in \mathfrak{B}^{(\beta_1)} \times \cdots \times \mathfrak{B}^{(\beta_n)} (k = 1, 2, \cdots)$, the inequality

$$\begin{vmatrix} P^{\hat{\beta}^*;\hat{\alpha}^*\hat{\beta}}(B_1^* \times B \times B_1) & \cdots & P^{\hat{\beta}^*;\hat{\alpha};\hat{\beta}}(B_1^* \times B \times B_k) \\ \cdots\cdots & \cdots & \cdots\cdots \\ P^{\hat{\beta}^*;\hat{\alpha};\hat{\beta}}(B_k^* \times B \times B_1) & \cdots & P^{\hat{\beta}^*;\hat{\alpha};\hat{\beta}}(B_k^* \times B \times B_k) \end{vmatrix} \geq 0 \tag{2.85}$$

is satisfied for any $P \in \mathfrak{S}$ in the case $\{\hat{\alpha}\} = C$. The sets of complex probability measures (2.83) defined when (2.82) is different from zero, correspond to an element $P_1 \in \mathfrak{S}$.

c) If the above introduced sets $B_1, \cdots B_k$ are all nonzero sets and different (modulo $\mathfrak{S}$) from one another, then the determinant in (2.85) is positive for at least one $P \in \mathfrak{S}$ when $\hat{\alpha} = \varnothing$.

We can summarize the most important conclusions of this subsection in the following theorem.

Theorem 1: Axiom X represents a necessary condition for the equivalence of the language $\mathfrak{L}$, introduced in I, Sec. 2.2, to the Hilbert-space formalism. It is also a sufficient condition for the introduction of the norm (2.62) by means of which the algebra $\mathfrak{A}_2(\mathfrak{O}')$ becomes a normed *-algebra $\mathfrak{A}_3(\mathfrak{O}')$.

2.4. The normed *-algebra $\mathfrak{A}_4(\mathfrak{O}')$

In the next subsection, we will turn to the task of embedding $\mathfrak{O}_b$ into a Banach algebra. This algebra

[10] See the notation introduced in, e. g., (2.55a).

is not the completion in norm of $\mathfrak{A}_3(\mathcal{O}')$ itself, but of another normed *-algebra, $\mathfrak{A}_4(\mathcal{O}')$, whose elements will be denoted by the letters $x, y, z, \cdots$, and which will be obtained from $\mathfrak{A}_3'$ with the help of the following rules:

Restrictive Rule. Consider sequences $A_1, A_2, \cdots$ of elements of $\mathfrak{A}_3(\mathcal{O}')$ having the following form:

$$A_k = \sum_{i=1}^{M} a_i A_{\hat{\alpha}}(B_{ik}^{(1)}) \cdots A_{\hat{\alpha}}(B_{ik}^{(n_i)}), \quad k = 1, 2, \cdots. \tag{2.86}$$

Here, each sequence $B_{i1}^{(m)}, B_{i2}^{(m)}, \cdots$ is a monotone sequence (decreasing or increasing) of sets which are unions of finite members of disjoint sets belonging to $B^{(\alpha_1)} \times \cdots \times B^{(\alpha_n)}$, where $\hat{\alpha} = (\alpha_1, \cdots \alpha_n) \in \hat{\mathcal{O}}'$ and $m = 1, \cdots n_i$, $i = 1, \cdots M$. If B is such a union of mutually disjoint sets

$$B_1, \cdots B_r \in \mathfrak{B}^{(\alpha_1)} \times \cdots \times \mathfrak{B}^{(\alpha_n)},$$

then $A_{\hat{\alpha}}(B)$ is defined as:

$$A_{\hat{\alpha}}(B) = A_{\hat{\alpha}}(B_1) + \cdots + A_{\hat{\alpha}}(B_r).$$

For a set $B \in \mathfrak{B}^{(\alpha_1)} \times \cdots \times \mathfrak{B}^{(\alpha_n)}$, $A_{\hat{\alpha}}(B)$ is defined as the element of $\mathfrak{A}_3(\mathcal{O}')$ adjoint to the question $\chi_B^{\hat{\alpha}}(\hat{\alpha})$ by the rules 1, 2, and 3 in Sec. 2.1.

We will say that such a sequence $A_1, A_2, \cdots$ of elements of $\mathfrak{A}_3(\mathcal{O}')$ obeys the Restrictive Rule if the following conditions are fulfilled:

1. Both sequences $A_1, A_2, \cdots$ and $A_1^*, A_2^* \cdots$[11] are Cauchy sequences in the strong topology of $\mathfrak{A}_3(\mathcal{O}')$, in which the neighborhood basis of an element $A_0 \in \mathfrak{A}_3'$ consists of all elements of $\mathfrak{A}_3'$ for which

$$(\langle (A - A_0)^* (A - A_0)\rangle_{P_1})^{\frac{1}{2}} < \epsilon,$$
$$\cdots (\langle (A - A_0)^* (A - A_0)\rangle_{P_n})^{\frac{1}{2}} < \epsilon, \tag{2.87}$$

for some $\epsilon > 0$ and $P_1, \cdots P_n \in \mathcal{S}$, $n = 1, 2, \cdots$. The necessary and sufficient condition for that is, obviously, that

$$(\langle (A_m - A_n)^* (A_m - A_n)\rangle_P)^{\frac{1}{2}} < \epsilon,$$
$$(\langle (A_m - A_n)(A_m - A_n)^*\rangle_P)^{\frac{1}{2}} < \epsilon$$

for some $n > N(\epsilon, P)$ and for any $P \in \mathcal{S}$, $m > n$.

2. If we denote by X the sequence $A_1, A_2, \cdots$ and by $\langle X^*X\rangle_P$ the limit of $\langle A_1^*A_1\rangle_P, \langle A_2^*A_2\rangle_P, \cdots$:

$$\langle X^*X\rangle_P = \lim_{n\to\infty} \langle A_n^*A_n\rangle_P, \tag{2.88}$$

(we will prove in our next lemma that these limits exist) then

$$||X|| = \sup_{P\in\mathcal{S}} (\langle X^*X\rangle_P)^{\frac{1}{2}} \tag{2.89}$$

has to be finite. The same condition is imposed on $\sup_{P\in\mathcal{S}}(\langle X^*X\rangle_P)^{\frac{1}{2}}$.

3. The set

$$\{||A_1||, ||A_2||, \cdots\}$$

of norms of elements in the sequence is a bounded set of real numbers.

Lemma 9: If the sequence $A_1, A_2, \cdots$ of elements of $\mathfrak{A}_3'$ is convergent in the above introduced strong topology of $\mathfrak{A}_3'$, then

$$(\langle A_1^*A_1\rangle_P)^{\frac{1}{2}}, \quad (\langle A_2^*A_2\rangle_P)^{\frac{1}{2}}, \cdots, \tag{2.90}$$

and

$$\langle A_1\rangle_P, \langle A_2\rangle_P, \langle A_3\rangle_P, \cdots \tag{2.91}$$

are convergent sequences of numbers for any $P \in \mathcal{S}$.

Proof: First we prove that

$$(\langle (A + B)^* (A + B)\rangle_P)^{\frac{1}{2}} \leq (\langle A^*A\rangle_P)^{\frac{1}{2}} + (\langle B^*B\rangle_P)^{\frac{1}{2}}, \quad A, B \in \mathfrak{A}_3' \tag{2.92}$$

by the same method as in lemma 7 (Sec. 2.3).

From (2.92) we deduce that

$$(\langle A_m^*A_m\rangle_P)^{\frac{1}{2}} - (\langle A_n^*A_n\rangle_P)^{\frac{1}{2}} \leq (\langle (A_m - A_n)^* (A_m - A_n)\rangle_P)^{\frac{1}{2}}$$

and

$$(\langle A_m^*A_m\rangle_P)^{\frac{1}{2}} - (\langle A_n^*A_n\rangle_P)^{\frac{1}{2}} \geq -(\langle (A_n - A_m)^* (A_n - A_m)\rangle_P)^{\frac{1}{2}} = -(\langle (A_m - A_n)^* (A_m - A_n)\rangle_P)^{\frac{1}{2}},$$

and therefore

$$|(\langle A_m^*A_m\rangle_P)^{\frac{1}{2}} - (\langle A_n^*A_n\rangle_P)^{\frac{1}{2}}| \leq (\langle (A_m - A_n)^* (A_m - A_n)\rangle_P)^{\frac{1}{2}}. \tag{2.93}$$

The convergence of the sequence (2.89) is now evident.

The convergence of the sequence (2.90) follows immediately from the relation

$$|\langle A_m - A_n\rangle_P|^2 \leq \langle (A_m - A_n)^* (A_m - A_n)\rangle_P, \tag{2.94}$$

which can be derived directly from the Cauchy-Schwartz inequality. Q.E.D.

Theorem 2: Consider classes of all sequences

$$x = \{(A_1, A_2, \cdots), (A_1', A_2', \cdots), \cdots\},$$
$$A_1, A_2, \cdots, A_1', A_2', \cdots \in \mathfrak{A}_3' \tag{2.95}$$

[11] It should be noticed that, if a sequence $A_1, A_2, \ldots$ of bounded operators in a Hilbert space is a Cauchy sequence, the sequence $A_1^*, A_2^*, \ldots$ of their Hermitian adjoint operators is not necessarily a Cauchy sequence too [Ref. 12, p. 442]. Therefore, we have to require explicitly that, in our case, the sequence $A_1^*, A_2^*, \ldots$ is also a Cauchy sequence.

fulfilling the conditions of the Restrictive Rule, and such that $A_n - A'_n \to 0$ in the strong topoligy (2.87) in case that $(A_1, A_2, \cdots)$ and $(A'_1, A'_2, \cdots)$ are any two sequences belonging to the same class. The set $\mathfrak{A}_4(\mathfrak{O}')$ of all such classes becomes a normed *-algebra with the norm defined by (2.89), if the following operations of addition, multiplication, scalar multiplication, and involution are introduced

$$x + y = \{(A_1 + B_1, A_2 + B_2, \cdots), \cdots\}, \quad (2.96)$$

$$xy = \{(A_1A_1, A_2B_2, \cdots), \cdots\}, \quad (2.97)$$

$$ax = \{(aA_1, aA_2, \cdots), \cdots\}, \quad a \in \mathbf{C}^1, \quad (2.98)$$

$$x^* = \{(A_1^*, A_2^*, \cdots), \cdots\}, \quad (2.99)$$

where $y = \{(B_1, B_2, \cdots), \cdots\}$, $B_1, B_2, \cdots \in \mathfrak{A}'_3$. The norm (2.88) is completely regular,[12] i.e.,

$$||x^*x|| = ||x||^2. \quad (2.100)$$

Proof: We have first to prove that the operations (2.96)–(2.99) do not lead outside the introduced set $\mathfrak{A}_4(\mathfrak{O}')$.

First, it is evident that the sequence in the expressions (2.96), $\cdots$, (2.99) are really of the form (2.86) when the original sequences belonging to x and y are of that form.

The fact that the strong convergence of the sequence $(A_1 + B_1, A_2 + B_2 \cdots)$ in (2.96) follows from the strong convergence of the sequences $(A_1, A_2, \cdots)$ and $(B_1, B_2, \cdots)$ is a direct consequence of relation (2.92). It follows also from (2.92) that $||x + y||$ exists and is finite because

$$||x + y|| \leq ||x|| + ||y||. \quad (2.101)$$

On the basis of (2.92) it should also be clear that all sequences $(A_1 + B_1, \cdots)$, $(A'_1 + B'_1, \cdots)$ belong to the same class, i.e., to the same element of $\mathfrak{A}'_4$, when $(A_1, A_2, \cdots)$ and $(A'_1, A'_2, \cdots)$ together, as well as $(B_1, B_2. \cdots)$ and $(B'_1, B'_2, \cdots)$ together, belong to the same class.

Finally, the fact that the set

$$\{||A_1 + B_1||, ||A_2 + B_2||, \cdots\}$$

is bounded follows from the relation (2.65) when consideration is given to the fact that the sets $\{||A_1||, ||A_2||, \cdots\}$ and $\{||B_1||, ||B_2||, \cdots\}$ are bounded.

To verify the strong convergence of the sequence $(A_1B_1, A_2B_2, \cdots)$, we write the expression

$$\langle (B_m^*A_m^* - B_n^*A_n^*)(A_mB_m - A_nB_n)\rangle_P \quad (2.102')$$

in the form

$$\begin{aligned}&\langle B_n^*(A_m^* - A_n^*)(A_m - A_n)B_n\rangle_P \\ &\quad + \langle (B_m^* - B_n^*)A_m^*(A_m - A_n)B_n\rangle_P \\ &\quad + \langle B_n^*(A_m^* - A_n^*)A_m(B_m - B_n)\rangle_P \\ &\quad + \langle (B_m^* - B_n^*)A_m^*A_m(B_m - B_n)\rangle_P. \quad (2.102)\end{aligned}$$

We will demonstrate on the first term of (2.102) that it can be made arbitrarily small when m and n are sufficiently great. Using Cauchy-Schwartz inequality, the definition (2.62) of the norm in $\mathfrak{A}'_3$, and (2.77), we get

$$\begin{aligned}&|\langle B_n^*(A_m - A_n)^*(A_m - A_n)B_n\rangle_P|^2 \\ &\leq \langle B_n^*B_n\rangle_P\langle (A_m - A_n)^*(A_m - A_n) \\ &\quad \times B_nB_n^*(A_m - A_n)^*(A_m - A_n)\rangle_P \\ &\leq \langle (A_m - A_n)^*(A_m - A_n)\rangle_P\ ||B_n||^2 \\ &\quad \times ||(A_m - A_n)B_nB_n^*(A_m - A_n)^*(A_m - A_n)||^2 \\ &\leq \langle (A_m - A_n)^*(A_m - A_n)\rangle_P\ ||B_n||^4\ ||A_m - A_n||^6. \\ &\quad (2.103)\end{aligned}$$

It is evident that the right-hand side of inequality (2.103) can be made arbitrarily small because its first factor becomes arbitrarily small when m and n are sufficiently great, while the rest of the factors stay bounded when m and n increase. It can be shown in a similar manner that the rest of the terms in (2.102), and therefore (2.102′) also, can be made arbitrarily small for m and n sufficiently great.

In a completely analogous way, it can be established that $A_nB_n - A'_nB'_n \to \mathbf{0}$ when $n \to \infty$, in case that $(A_1, A_2, \cdots)$ and $(A'_1, A'_2, \cdots)$ as well as $(B_1, B_2, \cdots)$ and $(B'_1, B'_2, \cdots)$ belong to the same class.

We will now show that $||xy||$ is finite and that

$$||xy|| \leq ||x||\ ||y||. \quad (2.104)$$

For that purpose, we have to employ relation (2.81) when writing

$$\langle (A_nB_n)^*(A_nB_n)\rangle_P = \langle A_n^*A_n\rangle_{P_1}\langle B_n^*B_n\rangle_P, \quad (2.105)$$

where $P_1 \in \mathfrak{S}$ exists according to Hypothesis 2A. From (2.105) we immediately get

$$\langle (A_nB_n)^*(A_nB_n)\rangle_P \leq ||x||^2\ ||y||^2 \quad (2.106)$$

and hence follows (2.104).

Finally, the fact that the set

$$\{||A_1B_1||, ||A_2B_2||, \cdots\}$$

is bounded can be derived straightforwardly from the relation (2.77).

The case of the relation (2.98) is too trivial to be considered in any detail.

[12] M. A. Naimark, *Normed Rings*, transl. by L. F. Boron (P. Noordhoff N. V., Groningen, The Netherlands, 1959), p. 228.

In the case of the operation of involution (2.99), the strong convergence of the sequence $A_1^*, A_2^*, \cdots$ and the existence of $||x^*||$ have been explicitly required. In fact, because of the symmetry of the problem in this respect, it is evident that this requirement is preserved by the above considered operations (2.96), (2.97), and (2.98). Therefore we did not consider it explicitly.

It obviously follows from (2.69) that the set $\{||A_1^*||, ||A_2^*||, \cdots\}$ is bounded whenever the set $\{||A_1||, ||A_2||, \cdots\}$ is.

It is very straightforward to establish that $\mathfrak{A}_4'$ is a *-algebra by checking that the operations of addition, multiplication, scalar multiplication, and involution satisfy relations (2.21) – (2.30) and (2.31) – (2.33). In this context, the unity 1 and the zero **0** of $\mathfrak{A}_4'$ are defined as classes of sequences of elements of $\mathfrak{A}_3'$, containing the sequences $(\mathbf{1}, \mathbf{1}, \cdots)$, $\mathbf{1} \in \mathfrak{A}_3'$, $(\mathbf{0}, \mathbf{0}, \cdots)$, and $\mathbf{0} \in \mathfrak{A}_3'$, respectively.

To prove that we are dealing with a normed *-algebra and that the norm is completely regular, we still have to demonstrate that relations (2.63)–(2.66) and (2.68)–(2.70) are obeyed. Now, (2.65) and (2.68) have already been proved and written as relations (2.101) and (2.104). Relations (2.63), (2.64), and (2.65) are very easy to establish. To prove (2.69) and (2.70), we will first notice that we can define uniquely linear positive functionals $\langle x\rangle_P$, $x \in \mathfrak{A}_4(\mathcal{O}')$ on $\mathfrak{A}_4'$ by writing:

$$\langle x\rangle_P = \lim_{n\to\infty} \langle A_n\rangle_P, \qquad P \in \mathcal{S}, \tag{2.107}$$

where $A_1, A_2, \cdots$ is any sequence of elements of $\mathfrak{A}_3'$ which belongs to $x \in \mathfrak{A}_4'$. It is easy to see that this definition determines $\langle x\rangle_P$ uniquely. Namely, if $(A_1, A_2, \cdots)$ and $(A', A', \cdots)$ both belong to x, then we can derive, as in the case of (2.94):

$$|\langle A_n\rangle_P - \langle A_n'\rangle_P|^2 = |\langle A_n - A_n'\rangle_P|^2$$

$$\leq \langle (A_n - A_n')^*(A_n - A_n')\rangle_P \to 0 \quad \text{for} \quad n \to \infty.$$

The Cauchy-Schwartz inequality for the functionals $\langle x\rangle_P$ on the algebra $\mathfrak{A}_4'$:

$$|\langle x^*y\rangle_P|^2 \leq \langle x^*x\rangle_P \langle y^*y\rangle_P, \tag{2.108}$$

can easily be derived from the same relations valid for the functionals $\langle A\rangle_P$ on the algebra $\mathfrak{A}_3'$ by going to the limit. Once this is established, we can apply Lemma 8 to obtain

$$||x^*|| = ||x||, \qquad x \in \mathfrak{A}_4', \tag{2.109}$$

$$||x^*x|| = ||x^*||\,||x||, \qquad x \in \mathfrak{A}_4'. \tag{2.110}$$

2.5 The B^*-algebra $\mathfrak{A}(\mathcal{O}')$ and its relation to the set $\mathcal{O}_b$ of bounded observables

We can now, finally, begin with the task of embedding $\mathcal{O}_b$ in a B^*-algebra. This will be done in a few steps by introducing rules relating elements of $\mathcal{O}_b$ to elements of some of the already constructed algebras.

Rule 1: Assume that the question q can be written in the form $q = \chi_{B_1\times\cdots\times B_n}^{\alpha_1,\cdots,\alpha_n}(\alpha_1, \cdots \alpha_n)(\alpha_1, \cdots \alpha_n) \in \mathcal{O}'$, $B_n \in \mathfrak{B}^1$ with $\{\alpha_1, \cdots \alpha_n\} = C$. We will denote the set of all such questions by $Q_0(\mathcal{O}')$. To each such question we put, in correspondence in $\mathfrak{A}_3(\mathcal{O}')$, the element A_q containing the polynomial form over $\mathfrak{Q}(\mathcal{O}')$:

$$1\chi_{B_1}^{\alpha_1}(\alpha_1) \circ \cdots \circ \chi_{B_n}^{\alpha_n}(\alpha_n), \tag{2.111}$$

and in $\mathfrak{A}_4(\mathcal{O}')$ the element x_q which is assigned to A_q when constructing $\mathfrak{A}_4'$ out of $\mathfrak{A}_3'$.

According to the definition of a fundamental set and of a bounded observable, each such observable α can be written as a function $\alpha = f(\alpha_1, \cdots \alpha_n)$ of n compatible observables $\alpha_1, \cdots \alpha_n$ belonging to a fundamental set $\mathcal{O}'$. The corresponding real-valued function $f(\lambda_1, \cdots \lambda_n)$ on $\mathbf{R}^{(\alpha_1,\cdots,\alpha_n)}$ is, according to the definition in I, Sec. 2.2, a Borel-measurable function, i.e., $f(\lambda_1, \cdots \lambda_n) \in \mathfrak{F}_B^n$. As α is bounded, the range of the function $f(\lambda_1, \cdots \lambda_n)$ is a bounded set in $\mathbf{R}^n$ and therefore $f(\lambda_1, \cdots \lambda_n)$ is a Lebesgue-integrable function on $\mathbf{R}^n$. Therefore, according to the definition of an integrable function,[13] there is at least one mean fundamental sequence $\{f_m\}$ of integrable simple functions

$$f_m(\lambda_1, \cdots \lambda_n) = \sum_{i=1}^{k_m} a_m^{(i)} \chi_{B_m^{(i)}}(\lambda_1, \cdots \lambda_n), \tag{2.112}$$

$$B_m^{(i)} \cap B_m^{(j)} = \varnothing \quad \text{for} \quad i \neq j, \qquad i, j = 1, \cdots k_m,$$

converging in measure to $f(\lambda_1, \cdots \lambda_n)$.

These considerations indicate that, as the next step, we have to assign elements of $\mathfrak{A}_4'$ to questions of the form

$$\chi_B^{\hat{\alpha}}(\hat{\alpha}), \qquad \hat{\alpha} \in \hat{\mathcal{O}}', \qquad B \in \mathfrak{B}^{\hat{\alpha}}, \tag{2.113}$$

where B can be *any* element of $\mathfrak{B}^{\hat{\alpha}}$, and afterwards to

[13] At this point it should be clear why we have taken the closure of $\mathfrak{A}_3(\mathcal{O}')$ in the strong topology instead of the uniform topology determined by the norm of $\mathfrak{A}_3'$. Namely, we can have a case when $B_1 \supset B_2 \supset \ldots$ with $\cap_{k=1}^{\infty} B_k = \varnothing$, without the sequence $A_1, A_2, \ldots$ being a Cauchy sequence in the uniform topology. For example, take an observable α with a pure point spectrum $\mathbf{S}^\alpha = \{1/n, n = 1, 2, \ldots\}$. If we choose $B_k = (0, 1/k]$, $k = 1, 2, \ldots$, for the above Borel sets, then the imposed conditions are satisfied, but we have

$$||A_n - A_{n-1}|| = \sup_{P\in\mathcal{S}} P^\alpha(B_n - B_{n-1})$$

$$= \sup_{P\in\mathcal{S}} P^\alpha([1/n, 1/(n-1)]) = 1$$

according to Ax. IX.

observables of the form

$$\sum_{i=1}^{k} a^{(i)}\chi^{\mathfrak{A}}_{B^{(i)}}(\mathfrak{A}), \qquad B^{(i)} \cap B^{(j)} = \varnothing \quad \text{for } i \neq j. \tag{2.114}$$

Rule 2: If the elements $A_i \in \mathfrak{A}_3^f$, $i = 1, \cdots k$, are assigned by Rule 1 to the questions $\chi^{\mathfrak{A}}_{B_1}, \cdots \chi^{\mathfrak{A}}_{B_k} \in Q_0(\mathfrak{O}^f)$, $B_i \cap B_j = \varnothing$ for $i \neq j$, then the element $A = A_1 + \cdots + A_k \in \mathfrak{A}_3^f$ is assigned to the question $\chi^{\mathfrak{A}}_B$, $B = B_1 \cup \cdots \cup B_k$. We denote by $Q_1(\mathfrak{O}^f)$ the set of all questions of this form.

Lemma 10: For given compatible observables $\alpha_1, \cdots \alpha_n \in \mathfrak{O}^f$ the class of all sets, which are unions of finite numbers of disjoint sets belonging to $\mathfrak{B}^{(\alpha_1)} \times \cdots \times \mathfrak{B}^{(\alpha_n)}$, is a Boolean algebra, which we will conveniently denote with $\overset{\circ}{\mathfrak{B}}{}^{(\alpha_1 \cdots \alpha_n)}$ or $\overset{\circ}{\mathfrak{B}}{}^{\mathfrak{A}}(\mathfrak{A} = (\alpha_1, \cdots \alpha_n))$. The questions $\chi^{\mathfrak{A}}_B$ belonging to sets $B \in \overset{\circ}{\mathfrak{B}}{}^{(\alpha_1 \cdots \alpha_n)}$ are questions from the family $Q_1(\mathfrak{O}^f)$ of questions. Rule 2 assigns *uniquely* to such questions elements of the algebra $\mathfrak{A}_3^f$, i.e., if we have $B_1' \cup \cdots \cup B_{k_1}' = B_1'' \cup \cdots \cup B_{k_2}'' \in \overset{\circ}{\mathfrak{B}}{}^{\mathfrak{A}}$, where $B_1', \cdots B_{k_1}', B_1'', \cdots B_{k_2}'' \in \mathfrak{B}^{(\alpha_1)} \times \cdots \times \mathfrak{B}^{(\alpha_n)}$, and $A_1' \in \mathfrak{A}_3^f$ is assigned to $\chi^{\mathfrak{A}}_{B_1'}$, A_1'' to $\chi^{\mathfrak{A}}_{B_1''}$, etc., then $A_1' + \cdots + A_{k_1}' = A_1'' + \cdots + A_{k_2}''$; in general, if $\chi^{\mathfrak{A}_1}_{B_1} = \chi^{\mathfrak{A}_2}_{B_2}$ and $\chi^{\mathfrak{A}_1}_{B_1}, \chi^{\mathfrak{A}_2}_{B_2} \in Q_1(\mathfrak{O}^f)$, then Rule 2 assigns the same element of $\mathfrak{A}_3(\mathfrak{O}^f)$ to both these questions.

Proof: The fact that $\overset{\circ}{\mathfrak{B}}{}^{\mathfrak{A}}$ is a Boolean algebra is a direct consequence of Theorem E in Ref. 12, Sec. 33 when that theorem is reformulated for the case of Boolean algebras instead of Boolean rings.

To show that $A_1' + \cdots + A_{k_1}' = A_1'' + \cdots + A_{k_2}''$, we consider the expression

$$\langle (A_1' + A_2' + \cdots + A_{k_1}' - A_1'' - \cdots - A_{k_2}'')^2 \rangle_P, \tag{2.115}$$

which is equal to

$$F_P[(\chi^{\mathfrak{A}}_{B_1'} \oplus \cdots \oplus \chi^{\mathfrak{A}}_{B_{k_1}'} - \chi^{\mathfrak{A}}_{B_1''} - \cdots - \chi^{\mathfrak{A}}_{B_{k_2}''})^2], \tag{2.116}$$

where the functional F_P, defined on elements of $\mathbf{C}^1 \tilde{\times} \mathfrak{O}^f$, is given by formula (2.40). As we have $\chi^{\mathfrak{A}}_{B_1'}(\hat{\lambda}) + \cdots + \chi^{\mathfrak{A}}_{B_{k_1}'}(\hat{\lambda}) = \chi^{\mathfrak{A}}_{B_1''}(\hat{\lambda}) + \cdots + \chi^{\mathfrak{A}}_{B_{k_2}''}(\hat{\lambda})$, $\hat{\lambda} \in \mathbf{R}^{\mathfrak{A}}$, we conclude from (2.40) that expression (2.115) equals zero. In general, if $\chi^{\mathfrak{A}_1}_{B_1}(\mathfrak{A}_1) = \chi^{\mathfrak{A}_2}_{B_2}(\mathfrak{A}_2)$ ($B_1 \in \overset{\circ}{\mathfrak{B}}{}^{\mathfrak{A}_1}$, $B_2 \in \overset{\circ}{\mathfrak{B}}{}^{\mathfrak{A}_2}$), we have, according to Ax. II in I, that $P^{\alpha_1}(\{1\}) = P^{\alpha_2}(\{1\})$ for all $P \in \mathfrak{S}(q_1 = \chi^{\mathfrak{A}_1}_{B_1}(\mathfrak{A}_1)$, $q_2 = \chi^{\mathfrak{A}_2}_{B_2}(\mathfrak{A}_2))$. From (2.40), we can derive easily that expression (2.115) must equal

$$P^{\alpha_1;\alpha_1}(\{1\} \times \{1\}) - P^{\alpha_1;\alpha_2}(\{1\} \times \{1\}) - P^{\alpha_2;\alpha_1}(\{1\} \times \{1\}) + P^{\alpha_2;\alpha_2}(\{1\} \times \{1\}). \tag{2.117}$$

Expression (2.117) is equal to zero when $q_1 = q_2$, because then

$$P^{\alpha_1;\alpha_2}(\{1\} \times \{1\}) = P^{\alpha_1}(\{1\}) = P^{\alpha_2;\alpha_1}(\{1\} \times \{1\})$$

and

$$P^{\alpha_1;\alpha_1}(\{1\} \times \{1\}) = P^{\alpha_1}(\{1\}) = P^{\alpha_2;\alpha_2}(\{1\} \times \{1\}).$$

Therefore, in general, we get

$$\begin{aligned} &||A_1' + \cdots + A_{k_1}' - (A_1'' + \cdots + A_{k_2}'')||^2 \\ &= \sup_{P \in \mathfrak{S}} \langle (A_1' + \cdots + A_{k_1}' - A_1'' - \cdots - A_{k_2}'')^2 \rangle_P \\ &= 0, \end{aligned} \tag{12.18}$$

and hence

$$A_1' + \cdots + A_{k_1}' = A_1'' + \cdots + A_{k_2}''. \qquad \text{Q.E.D.}$$

Rule 3: If $\{\mathfrak{A}\} = C$ and $B_1, B_2, \cdots \in \overset{\circ}{\mathfrak{B}}{}^{\mathfrak{A}}$ is any monotone sequence of sets from $\overset{\circ}{\mathfrak{B}}{}^{\mathfrak{A}}$ for which the corresponding elements $A_1, A_2, \cdots \in \mathfrak{A}_3^f$, assigned by Rule 2 to $\chi^{\mathfrak{A}}_{B_1}(\mathfrak{A})$, $\chi^{\mathfrak{A}}_{B_2}(\mathfrak{A})$, $\cdots$, form a Cauchy sequence in the strong topology of $\mathfrak{A}_3^f$, then we assign to the question $\chi^{\mathfrak{A}}_B(\mathfrak{A})$, $B = \lim_{k\to\infty} B_k$, the element $x \in \mathfrak{A}_n^f$ determined by this Cauchy sequence.

Lemma 11: The monotone class generated by the family of sets $\overset{\circ}{\mathfrak{B}}{}^{\mathfrak{A}}$, $\mathfrak{A} \in \hat{\mathfrak{O}}^f$, $\{\mathfrak{A}\} = C$, is a Boolean σ-algebra which coincides with $\mathfrak{B}^{\mathfrak{A}}$. Any sequence of elements of $\mathfrak{A}_3^f$, which is assigned by Rule 3 to a monotone sequence $B_1, B_2, \cdots \in \mathfrak{B}^{\mathfrak{A}}$, is a Cauchy sequence in the strong topology of $\mathfrak{A}_3^f$ and therefore defines an element of $\mathfrak{A}_4^f$. Rule 3 maps injectively the set $Q_3(\mathfrak{O}^f)$ of all questions of the form $\chi^{\mathfrak{A}}_B(\mathfrak{A})$, $\mathfrak{A} \in \hat{\mathfrak{O}}^f$, $B \in \mathfrak{B}^{\mathfrak{A}}$, into the algebra $\mathfrak{A}_4(\mathfrak{O})$.

Proof: The fact that $\mathfrak{B}^{\mathfrak{A}}$ is equal to the monotone class generated by $\overset{\circ}{\mathfrak{B}}{}^{\mathfrak{A}}$ is an obvious consequence of Theorem B in Ref. 14, Sec. 6; it has only to be remembered that $\overset{\circ}{\mathfrak{B}}{}^{\mathfrak{A}} \subset \mathfrak{B}^{\mathfrak{A}}$ because $\overset{\circ}{\mathfrak{B}}{}^{\mathfrak{A}}$ is a Boolean algebra generated by $\mathfrak{B}^{(\alpha_1)} \times \cdots \times \mathfrak{B}^{(\alpha_n)}$, which is contained in $\mathfrak{B}^{\mathfrak{A}}$.

Let us say that we are now dealing with a decreasing sequence $B_1, B_2, B_3, \cdots$ of elements from $\overset{\circ}{\mathfrak{B}}{}^{\mathfrak{A}}$ to which the sequence $A_1, A_2, \cdots$ of elements of $\mathfrak{A}_3^f$ is assigned by Rule 2. It is easy to check with formula (2.40) and Ax. I, 2 that for $m > n$

$$\begin{aligned} \langle (A_m - A_n)^*(A_m - A_n) \rangle_P &= \langle (A_m - A_n)^2 \rangle_P \\ &= P^{\mathfrak{A};\mathfrak{A}}(B_m \times B_m) - P^{\mathfrak{A};\mathfrak{A}}(B_m \times B_n) \\ &\quad - P^{\mathfrak{A};\mathfrak{A}}(B_n \times B_m) + P^{\mathfrak{A};\mathfrak{A}}(B_n \times B_n) \\ &= P^{\mathfrak{A}}(B_m - B_n). \end{aligned} \tag{2.119}$$

The sequence $P^{\mathfrak{A}}(B_1), P^{\mathfrak{A}}(B_2), \cdots$ is a monotone nonincreasing sequence of nonnegative numbers, which is therefore convergent. But we can write

[14] P. R. Halmos, *Measure Theory* (D. van Nostrand Company, Princeton, New Jersey, 1961).

$$P^{\mathfrak{a}}(B_m - B_n) = P^{\mathfrak{a}}(B_m) - P^{\mathfrak{a}}(B_n). \qquad (2.120)$$

Hence, the sequence $A_1, A_2, \cdots$ of self-adjoint elements of $\mathfrak{A}_3'$ satisfies the first condition[13] of the Restrictive Rule in Sec. 2.5. It is also obvious that

$$\sup_{P \in \mathcal{S}} [\lim_{n\to\infty} (\langle A_n^* A_n \rangle_P)^{\frac{1}{2}}] \leq ||A_1^* A_1|| < +\infty$$

and that $\{||A_1||, ||A_2||, \cdots\}$ is a bounded set. Therefore $(A_1, A_2, \cdots)$ belongs to an element x of $\mathfrak{A}_4(\mathfrak{O}')$.

In the case of an increasing sequence $B_1 \subset B_2 \subset \cdots$, the corresponding sequence $A_1, A_2, \cdots$ would again be a Cauchy sequence because

$$P^{\mathfrak{a}}(B_1 \cup B_2 \cup \cdots) \leq 1,$$

and therefore the sequence $P^{\mathfrak{a}}(B_1), P^{\mathfrak{a}}(B_2), \cdots$ is a monotone nondecreasing convergent sequence of numbers. Because of (2.120) we conclude that the first condition of the Restrictive Rule in Sec. 2.4 is satisfied. As we also have that

$$\langle A_n^* A_n \rangle_P = \langle A_n \rangle_P = P(B_n) \leq 1,$$

we conclude that the other two conditions are satisfied too. Hence, the sequence $(A_1, A_2, \cdots)$ belongs again to an element of $\mathfrak{A}_4(\mathfrak{O}')$.

The next step is to establish the fact that we are really dealing with a mapping of $Q_3(\mathfrak{O}')$ into $\mathfrak{A}_4(\mathfrak{O}')$, i.e., that only one element of $\mathfrak{A}_4(\mathfrak{O}')$ is assigned by our procedure to each element

$$\chi_B^{\mathfrak{a}}(\hat{\alpha})(\hat{\alpha} \in \hat{\mathfrak{O}}', \{\hat{\alpha}\} = C) \quad \text{of} \quad Q_3(\mathfrak{O}').$$

Assume that $B_1' \subset B_2' \subset \cdots$ and $B_1'' \subset B_2'' \subset \cdots$ are two increasing sequences of elements of $\mathring{\mathfrak{B}}^{\mathfrak{a}'}$ and $\mathring{\mathfrak{B}}^{\mathfrak{a}''}$ respectively, such that $\chi_B^{\mathfrak{a}'}(\hat{\alpha}') = \chi_B^{\mathfrak{a}''}(\hat{\alpha}'')$, $B' = B_1' \cup B_2' \cup \cdots$, $B'' = B_1'' \cup B_2'' \cup \cdots$. We have to prove that the two sequences $(A_1', A_2', \cdots)$ and $(A_1'', A_2'', \cdots)$ of elements of $\mathfrak{A}_3(\mathfrak{O}')$, corresponding to $B_1', B_2', \cdots$ and $B_1'', B_2'', \cdots$ respectively, belong to the same element of $\mathfrak{A}_4(\mathfrak{O}')$, i.e., that

$$\lim_{n\to\infty} \langle (A_n' - A_n'')^*(A_n' - A_n'') \rangle_P = 0, \qquad P \in \mathcal{S}. \qquad (2.121)$$

Again, we can easily establish that

$$\begin{aligned}
&\langle (A_n' - A_n'')^*(A_n' - A_n'') \rangle_P \\
&= \langle A_n'^* A_n' \rangle_P - \langle A_n''^* A_n' \rangle_P - \langle A_n'^* A_n'' \rangle_P + \langle A_n''^* A_n'' \rangle_P \\
&= P^{\mathfrak{a}';\mathfrak{a}'}(B_n' \times B_n') - P^{\mathfrak{a}'';\mathfrak{a}'}(B_n'' \times B_n') \\
&\quad - P^{\mathfrak{a}';\mathfrak{a}''}(B_n' \times B_n'') + P^{\mathfrak{a}'';\mathfrak{a}''}(B_n'' \times B_n''). \qquad (2.122)
\end{aligned}$$

As the sequences $(B_1' \times B_1', B_2' \times B_2', \cdots)$, $(B_1'' \times B_1', B_2'' \times B_2', \cdots)$, etc. are increasing sequences of Borel sets in, respectively, $\mathbf{R}^{\mathfrak{a}',\mathfrak{a}'}$, $\mathbf{R}^{\mathfrak{a}'',\mathfrak{a}'}$, etc., it follows from the continuity from below of measures that when $n \to \infty$, the right-hand side of (2.122) has the limit (when $n \to \infty$),

$$\begin{aligned}
&P^{\mathfrak{a}';\mathfrak{a}'}(B' \times B') - P^{\mathfrak{a}'';\mathfrak{a}'}(B'' \times B') \\
&- P^{\mathfrak{a}';\mathfrak{a}''}(B' \times B'') + P^{\mathfrak{a}'';\mathfrak{a}''}(B'' \times B''). \qquad (2.123)
\end{aligned}$$

Because of the equality $\chi_{B'}^{\mathfrak{a}'}(\hat{\alpha}') = \chi_{B''}^{\mathfrak{a}''}(\hat{\alpha}'')$, all terms in (2.123) are equal to one another and therefore that expression is equal to zero.

Consider now the case when we are dealing with the following two sequences: $B_1' \subset B_2' \subset \cdots$ and $B_1'' \supset B_2'' \supset \cdots$, $B_1', B_2', \cdots, B_1'', B_2'', \cdots \in \mathring{\mathfrak{B}}^{\mathfrak{a}}$, $\bigcup_{n=1}^{\infty} B_n' = \bigcap_{n=1}^{\infty} B_n'' = B$. We will prove that the corresponding two sequences $(A_1', A_2', \cdots)$ and $(A_1'', A_2'', \cdots)$ of elements of $\mathfrak{A}_3(\mathfrak{O}')$ determine the same element of $\mathfrak{A}_4(\mathfrak{O}')$, by showing that (2.121) is true again. Now we can write

$$\begin{aligned}
&\langle (A_n'' - A_n')^*(A_n'' - A_n') \rangle_P \\
&\quad = P^{\mathfrak{a};\mathfrak{a}}((B_n'' - B_n') \times (B_n'' - B_n')). \qquad (2.124)
\end{aligned}$$

Since $(B_1'' - B_1') \supset (B_2'' - B_1') \supset \cdots$ is a decreasing sequence of Borel sets, having the empty set, as a limit, we can again infer from the continuity from above of the complex measure $P^{\mathfrak{a};\mathfrak{a}}(B)(B \in \mathfrak{B}^{\mathfrak{a};\mathfrak{a}})$ that the expressions (2.124) tend to zero when $n \to \infty$.

The more general case, dealing with an increasing sequence, $B_1' \subset B_2' \subset \cdots$, and a decreasing sequence, $B_1'' \supset B_2'' \supset \cdots$, such that $B_1', B_2', \cdots \in \mathring{\mathfrak{B}}^{\mathfrak{a}'}$, $B_1'', B_2'', \cdots \in \mathring{\mathfrak{B}}^{\mathfrak{a}''}$, and $\chi_{B'}^{\mathfrak{a}'}(\hat{\alpha}') = \chi_{B''}^{\mathfrak{a}''}(\hat{\alpha}'')$, $\hat{\alpha}' \neq \hat{\alpha}''$, $B' = B_1' \cup B_2' \cup \cdots$, $B'' = B_1'' \cap B_2'' \cap \cdots$, can be easily reduced to the above two cases in the following way: take an increasing sequence of sets $B_1''' \subset B_2''' \subset \cdots \subset B''$ from $\mathring{\mathfrak{B}}^{\mathfrak{a}''}$ having B'' as a limit[15]. If $(A_1''', A_2''', \cdots)$ is the corresponding sequence of elements of $\mathfrak{A}_3(\mathfrak{O}')$, then according to the first case, $(A_1', A_2', \cdots)$ and $(A_1''', A_2''', \cdots)$ represent the same element of $\mathfrak{A}_4'$, while according to the second case, $(A_1'', A_2'', \cdots)$ and $(A_1''', A_2''', \cdots)$ represent the same element of $\mathfrak{A}_4'$ and therefore $(A_1', A_2', \cdots)$ and $(A_1'', A_2'', \cdots)$ belong to the same element of $\mathfrak{A}_4'$ too.[16]

Now that it is shown that we are dealing with a mapping, we are left to demonstrate that this mapping is injective. To this purpose assume that two sequences $(A_1', A_2', \cdots)$ and $(A_1'', A_2'', \cdots)$ of elements of $\mathfrak{A}_3(\mathfrak{O}')$ correspond to the increasing sequences $B_1' \subset B_2' \subset \cdots$ and $B_1'' \subset B_2'' \subset \cdots$

[15] This can always be done—as is shown in more detail in the course of proving lemma 13.

[16] It should be clear that the elements of $\mathfrak{A}_4(\mathfrak{O}')$ are equivalence classes of sequences $(A_1, A_2, \ldots)$ of elements of $\mathfrak{A}_3(\mathfrak{O}')$.

respectively, where $B_1', B_2', \cdots \in \mathring{\mathfrak{B}}^{\hat{\alpha}'}$, $B_1'', B_1'' \cdots \in \mathring{\mathfrak{B}}^{\hat{\alpha}''}$. We will show that if condition (2.121) is fulfilled, i.e., if these two sequences represent the same element of $\mathfrak{A}_4'$, then we must have

$$\chi_{B'}^{\hat{\alpha}'}(\hat{\alpha}') = \chi_{B''}^{\hat{\alpha}''}(\hat{\alpha}''), \quad B' = \bigcup_{n=1}^{\infty} B_n', \quad B'' = \bigcup_{n=1}^{\infty} B_n''. \tag{2.125}$$

Take any $\hat{\alpha}, \hat{\beta} \in \hat{\mathfrak{O}}^f$, $\hat{\alpha} = (\alpha_1, \cdots \alpha_k)$, $\hat{\beta} = (\beta_1, \cdots \beta_l)$, and $B_1 = B_1^1 \times \cdots \times B_1^k \in \mathfrak{B}^{(\alpha_1)} \times \cdots \times \mathfrak{B}^{(\alpha_k)}$, $B_2 = B_2^1 \times \cdots \times B_2^l \in \mathfrak{B}^{(\beta_1)} \times \cdots \times B^{(\beta_l)}$. Denote by A_1^1 the element of $\mathfrak{A}_3(\mathfrak{O}^f)$ assigned by Rule 1 to $\chi_{B_1^1}^{\alpha_1}(\alpha_1)$, by A_1^2 the one assigned to $\chi_{B_1^2}^{\alpha_2}(\alpha_2)$ etc., by A_2^1 the one assigned to $\chi_{B_2^1}^{\alpha_1}(\alpha_2)$ etc., etc.; then denote by A_{q1} and A_{q2} the respective elements $A_1^1 \cdots A_1^k$ and $A_2^1 \cdots A_2^l$ of $\mathfrak{A}_3(\mathfrak{O}^f)$, where $q_1 = \chi_{B_1}^{\hat{\alpha}}(\hat{\alpha})$, $q_2 = \chi_{B_2}^{\hat{\beta}}(\hat{\alpha}_2)$. We can easily establish with the help of (2.40) that

$$\begin{aligned} \langle A_{q_1} A_n' A_{q_2} \rangle_P &= P^{\hat{\alpha};\hat{\alpha}';\hat{\beta}}(B_1 \times B_n' \times B_2), \\ \langle A_{q_1} A_n'' A_{q_2} \rangle_P &= P^{\hat{\alpha};\hat{\alpha}'';\hat{\beta}}(B_1 \times B_n'' \times B_2). \end{aligned} \tag{2.126}$$

Therefore, by using (2.67) we get from (2.126) that

$$\begin{aligned} &|P^{\hat{\alpha};\hat{\alpha}'';\hat{\beta}}(B_1 \times B_n' \times B_2) - P^{\hat{\alpha};\hat{\alpha}'';\hat{\beta}}(B_1 \times B_n'' \times B_2)|^4 \\ &= |\langle A_{q_1}(A_n' - A_n'')A_{q_2}\rangle_P|^4 \\ &\leq |\langle (A_n' - A_n'')A_{q_1}^* A_{q_1}(A_n' - A_n'')\rangle_P|^2 \ |\langle A_{q_2}^* A_{q_2}\rangle_P|^2 \\ &\leq \langle (A_n' - A_n'')^2\rangle_P \ ||A_n' - A_n''||^4 \ ||A_{q_1}||^4 \ ||A_{q_2}||^4. \end{aligned} \tag{2.127}$$

As the factor of $\langle (A_n' - A_n'')^2\rangle_P$ in the right-hand side of inequality (2.127) stays bounded when $n \to \infty$, we conclude that (2.121) has, as a consequence,

$$\begin{aligned} &\lim_{n\to\infty} P^{\hat{\alpha};\hat{\alpha}';\hat{\beta}}(B_1 \times B_n' \times B_2) \\ &= \lim_{n\to\infty} P^{\hat{\alpha};\hat{\alpha}'';\hat{\beta}}(B_1 \times B_n'' \times B_2), \quad P \in \mathfrak{S}. \end{aligned} \tag{2.128}$$

On the other hand, $B_1 \times B_1' \times B_2 \subset B_1 \times B_2' \times B_2 \subset \cdots$ is an increasing sequence of Borel sets in $\mathbf{R}^{\hat{\alpha},\hat{\alpha}',\hat{\beta}}$, and hence $P^{\hat{\alpha};\hat{\alpha}';\hat{\beta}}(B_1 \times B' \times B_2) = \lim_{n\to\infty} P^{\hat{\alpha};\hat{\alpha}';\hat{\beta}}(B_1 \times B_n' \times B_2)$. Similarly in case of $P^{\hat{\alpha};\hat{\alpha}'';\hat{\beta}}(B_1 \times B_n'' \times B_2)$; therefore (2.128) implies

$$\begin{aligned} &P^{\hat{\alpha};\hat{\alpha}';\hat{\beta}}(B_1 \times B' \times B_2) \\ &= P^{\hat{\alpha};\hat{\alpha}'';\hat{\beta}}(B_1 \times B'' \times B_2), \quad P \in \mathfrak{S}. \end{aligned} \tag{2.129}$$

According to Proposition 3 in I, Sec. 3.3, we can conclude that (2.125) is true.

The case when the sequence $B_1'', B_2'', \cdots$ of Borel sets is decreasing instead of increasing can be treated in exactly the same fashion. These being the only two logically distinct cases, we have proved that our mapping is really injective. Q.E.D.

Rule 4: Denote by $\mathfrak{O}_b(\mathfrak{O}^f)$ the set of all bounded observables β of the form

$$\begin{aligned} &\beta = f(\hat{\alpha}), \quad \hat{\alpha} \in \hat{\mathfrak{O}}^f, \quad f(\lambda) = \sum_{i=1}^{k} a^{(i)} \chi_{B_i}^{\hat{\alpha}}(\lambda), \quad \lambda \in \mathbf{R}^{\hat{\alpha}}, \\ &a^{(1)}, \cdots a^{(k)} \in \mathbf{C}^1, \qquad B_1, \cdots B_k \in \mathfrak{B}^{\hat{\alpha}}, \\ &\qquad B_i \cap B_j = \varnothing \quad \text{for} \quad i \neq j. \end{aligned} \tag{2.130}$$

If the elements $x^{(1)}, \cdots x^{(k)} \in \mathfrak{A}_4(\mathfrak{O}^f)$ are assigned by Rule 3 to the questions $\chi_{B_1}^{\hat{\alpha}}(\hat{\alpha}), \cdots \chi_{B_k}^{\hat{\alpha}}(\hat{\alpha})$ respectively, then the element

$$x_\beta = \sum_{i=1}^{k} a^{(i)} x^{(i)}, \qquad x_\beta \in \mathfrak{A}_4(\mathfrak{O}^f), \tag{2.131}$$

is to be assigned to the observable β.

Lemma 12: a) If the observable $\beta \in \mathfrak{O}_b(\mathfrak{O}^f)$ is given by formula (2.130), then

$$\langle X_\beta \rangle_P = \sum_{i=1}^{k} a^{(i)} P^{\hat{\alpha}}(B_i) \tag{2.132}$$

and therefore

$$||x_\beta|| \leq (a^{(1)^2} + \cdots + a^{(k)^2})^{\frac{1}{2}} \tag{2.133}$$

if all questions $\chi_{B_1}^{\hat{\alpha}}, \cdots \chi_{B_n}^{\hat{\alpha}}$ are different from the null-question. Hence:

b) Rule 4 assigns uniquely to each observable $\beta \in \mathfrak{O}_b(\mathfrak{O}^f)$ an element of $\mathfrak{A}_4(\mathfrak{O}^f)$. The ensuing mapping is injective.

Proof: a) Relation (2.132) follows straightforwardly from (2.40). To arrive at (2.133), we have to remember that, according to Ax. I(2), $P^{\alpha;\alpha}(B_i \times B_j) = P^{\hat{\alpha}}(B_i \cap B_j) = \delta_{ij} P^{\hat{\alpha}}(B_i)$ and according to Ax. I(3), $P^{\hat{\alpha}}(B_i) \leq 1$ ($\{\hat{\alpha}\} = C$) if B_i is any Borel set. Hence,

$$\begin{aligned} ||x_\beta||^2 &= \sup_{P\in\mathfrak{S}} \sum_{i,j=1}^{k} a^{(i)} a^{(j)} \langle x^{(i)} x^{(j)} \rangle_P \\ &= \sup_{P\in\mathfrak{S}} \sum_{i,j=1}^{k} a^{(i)} a^{(j)} P^{\hat{\alpha};\hat{\alpha}}(B_i \times B_j) \leq \sum_{i=1}^{n} a^{(i)^2}. \end{aligned}$$

b) Take any two observables $\beta = f_1(\hat{\alpha}_1)$, $\wp = f_2(\hat{\alpha}_2)$ belonging to $\mathfrak{O}_b(\mathfrak{O}^f)$. It is easy to establish in the manner used so many times in this section that

$$\begin{aligned} &\langle (x_\beta - x_\wp)^2 \rangle_P \\ &= \int f_1(\lambda_1) f_2(\lambda_2) \, dP^{\hat{\alpha}_1;\hat{\alpha}_1}(\lambda_1, \lambda_2) \\ &\quad - \int f_1(\lambda_1) f_2(\lambda_2) \, dP^{\hat{\alpha}_1;\hat{\alpha}_2}(\lambda_1, \lambda_2) \\ &\quad - \int f_1(\lambda_1) f_2(\lambda_2) \, dP^{\hat{\alpha}_2;\hat{\alpha}_1}(\lambda_2, \lambda_1) \\ &\quad + \int f_2(\lambda_1) f_2(\lambda_2) \, dP^{\hat{\alpha}_2;\hat{\alpha}_1}(\lambda_1, \lambda_2) \end{aligned}$$

$$= \int \lambda_1\lambda_2\, dP^{\beta;\beta}(\lambda_1, \lambda_2) - \int \lambda_1\lambda_2\, dP^{\beta;\varphi}(\lambda_1, \lambda_2)$$
$$- \int \lambda_1\lambda_2\, dP^{\varphi;\beta}(\lambda_1, \lambda_2) + \int \lambda_1\lambda_2\, dP^{\varphi;\varphi}(\lambda_1, \lambda_2). \tag{2.134}$$

It is now obvious that, if $\beta = \varphi$, then we have $||x_\beta - x_\varphi||^2 = 0$ and therefore $x_\beta = x_\varphi$. Hence, Rule 4 determines indeed a mapping of $\mathfrak{O}_b(\mathfrak{O}')$ into $\mathfrak{A}_4(\mathfrak{O}')$.

To show that this mapping is injective, assume that $x_\beta = x_\varphi$ for some β, $\varphi \in \mathfrak{O}_b(\mathfrak{O}')$.

First we have to notice that, from the very definition of the observables β and φ as elements of $\mathfrak{O}_b(\mathfrak{O}')$, both these observables must have a pure point spectrum containing a finite number of elements. Assume that the sets $\{a^{(1)}, \cdots a^{(m)}\}$ and $\{b^{(1)}, \cdots b^{(n)}\}$ are the spectra of β and φ respectively. Then we can write

$$\begin{aligned}
&\beta = a^{(1)}\chi^{\mathfrak{A}'}_{B_1'}(\hat{\alpha}') + \cdots + a^{(m)}\chi^{\mathfrak{A}'}_{B_m'}(\hat{\alpha}'),\\
&B_i' \cap B_j' = \varnothing \quad \text{for} \quad i \neq j (=1, \cdots m),\\
&\varphi = b^{(1)}\chi^{\mathfrak{A}''}_{B_1''}(\hat{\alpha}'') + \cdots + b^{(n)}\chi^{\mathfrak{A}''}_{B_n''}(\hat{\alpha}''),\\
&B_i'' \cap B_j'' = \varnothing \quad \text{for} \quad i \neq j (=1, \cdots n),
\end{aligned} \tag{2.135}$$

where $B_1', \cdots B_m' \in \mathfrak{B}^{\mathfrak{A}'}$ and $B_1'', \cdots B_n'' \in \mathfrak{B}^{\mathfrak{A}''}$ are nonzero Borel sets. If the elements

$$x^{(1)}, \cdots x^{(m)}, \qquad y^{(1)}, \cdots y^{(n)}$$

of $\mathfrak{A}_4(\mathfrak{O}')$ are assigned to the questions

$$\chi^{\mathfrak{A}'}_{B_1'}, \cdots \chi^{\mathfrak{A}'}_{B_m'}, \qquad \chi^{\mathfrak{A}''}_{B_1''}, \cdots \chi^{\mathfrak{A}''}_{B_n''}$$

respectively by Rule 3, then we can write

$$\begin{aligned}
&x_\beta = a^{(1)}x^{(1)} + \cdots + a^{(m)}x^{(m)},\\
&\quad x^{(i)}x^{(j)} = \delta_{ij}x^{(i)}, \qquad i, j = 1, \cdots m,\\
&x_\varphi = b^{(1)}y^{(1)} + \cdots + b^{(n)}y^{(n)},\\
&\quad y^{(i)}y^{(j)} = \delta_{ij}y^{(i)}, \qquad i, j = 1, \cdots n.
\end{aligned} \tag{2.136}$$

On the other hand, $\mathfrak{A}_4(\mathfrak{O}')$ is a *-algebra with a regular norm. On the basis of Theorem 5, Sec. 24.2 in Ref. 11, we can assert that $\mathfrak{A}_4(\mathfrak{O}')$ is completely isomorphic to some algebra of operators on a Hilbert space. But then formulas (2.135), written down in terms of bounded operators assigned by the mentioned isomorphism

$$A(x_\beta) = \sum_{i=1}^{m} a^{(i)}A(x^{(i)}), \qquad A(x_\varphi) = \sum_{j=1}^{n} b^{(j)}A(y^{(j)}), \tag{2.137}$$

represent the spectral decompositions of the self-adjoint bounded operators $A(x_\beta) = A(x_\varphi)$. According to the spectral theorem[17] for self-adjoint operators, the two spectra $\mathbf{S}^\beta$ and $\mathbf{S}^\varphi$ have to be identical, i.e., $m = n$ and (after a possible rearrangement of indices) $a^{(i)} = b^{(i)}$, $i = 1, \cdots n$. Besides, we must have $A(x^{(i)}) = A(y^{(i)})$ and therefore $x^{(i)} = y^{(i)}$. According to lemma 11, the equality $x^{(i)} = y^{(i)}$ implies $\chi^{\mathfrak{A}'}_{B_i'}(\hat{\alpha}') = \chi^{\mathfrak{A}''}_{B_i''}(\hat{\alpha}'')(i = 1, \cdots n)$. Hence we can conclude that $\beta = \varphi$. Q.E.D.

We will denote now by $\mathfrak{A}(\mathfrak{O}')$ the completion of $\mathfrak{A}_4(\mathfrak{O}')$ in the uniform topology defined by the norm. The elements of the B^*-algebra $\mathfrak{A}(\mathfrak{O}')$ will be denoted by the letters $x, y, z, \cdots$ too. The algebra $\mathfrak{A}_4(\mathfrak{O}')$ will now be completely isomorphic to a subalgebra $\mathfrak{A}_4^{(0)}(\mathfrak{O}')$ of $\mathfrak{A}(\mathfrak{O}')$.

Rule 5: Assume that the bounded observable β can be written as $\beta = f(\hat{\alpha})$, $\{\hat{\alpha}\} = C$. Take a sequence $\pi_1, \pi_2, \cdots$ of finer and finer partitions of the interval $[-||\beta||, +||\beta||]$, i.e.,

$$\pi_k = \{I_1^{(k)}, \cdots I_{n_k}^{(k)}\}, \qquad I_i \cap I_j = \varnothing \quad \text{if} \quad i \neq j,$$

which are such that, if ϵ_k is the length of the greatest interval in the partition π_k, then $\lim_{k\to\infty} \epsilon_k = 0$. If $B_i^{(k)} = f^{-1}(I_i^{(k)})$, then write

$$\beta_k = f_k(\hat{\alpha}) = \sum_{i=1}^{n_k} a_i^{(k)}\chi^{\mathfrak{A}}_{B_i^{(k)}}(\hat{\alpha}), \tag{2.138}$$

where $a_i^{(k)} \in I_i^{(k)}$ are some chosen values. Introduce the general notation in which $x^{\mathfrak{A}}_B(\hat{\alpha} \in \hat{\mathfrak{O}}', B \in \mathfrak{B}^{\mathfrak{A}})$ denotes the element of $\mathfrak{A}_4^{(0)}(\mathfrak{O}')$ assigned by the rules 1 to 4 to the question $\chi^{\mathfrak{A}}_B(\hat{\alpha})$. Then, if the sequence $x_{\beta_1} \cdots, x_{\beta_2}, \cdots,$

$$x_{\beta_k} = \sum_{i=1}^{n_k} a_i^{(k)}x^{\mathfrak{A}}_{B_i^{(k)}}, \qquad k = 1, 2, 3, \cdots \tag{2.139}$$

converges in the norm to an element x_β of $\mathfrak{A}(\mathfrak{O}')$, assign to the bounded observable β this element x_β.

Theorem 3: Rule 5 defines an injective mapping of the set $\mathfrak{O}_b$ of all observables into the B^*-algebra $\mathfrak{A}(\mathfrak{O}')$.

Proof: To show that we are dealing with a mapping, we have to demonstrate that, if $x_{\beta_1}, x_{\beta_2}, \cdots$ and $x_{\varphi_1}, x_{\varphi_2}, \cdots$ are any two sequences of the type mentioned in Rule 5, which correspond to sequences $\beta_1, \beta_2, \cdots$ and $\varphi_1, \varphi_2, \cdots$ of observables converging in the mentioned sense to the same observable β, then

$$\lim_{m\to\infty} x_{\beta_m} = \lim_{n\to\infty} x_{\varphi_n} = x_\beta. \tag{2.140}$$

Assume therefore that

$$\beta_n = f_n(\hat{\alpha}), \ \varphi_n = g_n(\hat{\beta}) (n = 1, 2, \cdots),$$

where the mean sequence $f_1(\hat{\lambda}), f_2(\hat{\lambda}), \cdots (\hat{\lambda} \in \mathbf{R}^{\mathfrak{A}})$

[17] Reference 12, Proposition IV, p. 248.

of simple step-functions of the mentioned type tends in measure towards $f(\hat{\lambda})$, and $g_1(\hat{\mu})$, $g_2(\hat{\mu})$, $\cdots$ ($\hat{\mu} \in \mathbf{R}^{\hat{\beta}}$) tend toward $g(\hat{\mu})$: the assumption is that $f(\hat{\alpha}) = g(\hat{\beta}) = \beta$. We can write

$$\begin{aligned}\langle (x_{\beta_n} - x_{\varphi_n})^2 \rangle_P = &\int f_n(\hat{\lambda}) f_n(\hat{\lambda}')\, dP^{\hat{\alpha};\hat{\alpha}}(\hat{\lambda}, \hat{\lambda}') \\ &- \int f_n(\lambda) g_n(\hat{\mu})\, dP^{\hat{\alpha};\hat{\beta}}(\hat{\lambda}, \hat{\mu}) \\ &- \int g_n(\hat{\mu}) f_n(\hat{\lambda})\, dP^{\hat{\beta};\hat{\alpha}}(\hat{\mu}, \hat{\lambda}) \\ &+ \int g_n(\hat{\mu}) g_n(\hat{\mu}')\, dP^{\hat{\beta};\hat{\beta}}(\hat{\mu}, \hat{\mu}'). \qquad (2.141)\end{aligned}$$

If we assume that

$$\beta_m = \sum_{j=1}^{r_m} b_j^{(m)} \chi_{B_j^{(m)}}^{\hat{\alpha}}(\hat{\alpha}), \qquad \varphi_n = \sum_{k=1}^{s_n} c_k^{(n)} \chi_{B_k^{(n)}}^{\hat{\beta}}(\hat{\beta}), \qquad (2.142)$$

then we can write the right-hand side of (2.141), after making use of Ax. I(2), in the form

$$\begin{aligned}&\sum_{j=1}^{r_n} \sum_{j_1=1}^{r_n} b_j^{(n)} b_{j_1}^{(n)} P^{\beta}(I_j^{(n)} \cap I_{j_1}^{(n)}) \\ &\quad - 2 \sum_{j=1}^{r_n} \sum_{k=1}^{s_n} b_j^{(n)} c_k^{(n)} P^{\beta}(I_j^{(n)} \cap I_k'^{(n)}) \\ &\quad + \sum_{k=1}^{s_n} \sum_{k_1=1}^{s_n} c_k^{(n)} c_{k_1}^{(n)} P^{\beta}(I_k'^{(n)} \cap I_{k_1}'^{(n)}) \\ &= \sum_{j=1}^{r_m} \sum_{k=1}^{s_n} (b_j^{(n)2} - 2 b_j^{(n)} c_k^{(n)} + c_k^{(n)2}) P^{\beta}(I_j^{(n)} \cap I_k'^{(n)}) \\ &\leq 2\, \epsilon_n'^2 \sum_{j=1}^{r_n} \sum_{k=1}^{s_n} P^{\beta}(I_j^{(m)} \cap I_k'^{(n)}) = 2\, \epsilon_n'^2. \qquad (2.141')\end{aligned}$$

Here, $I_j^{(n)} = f(B_j^{(n)})$, $I_k'^{(n)} = g(B_k^{(n)})$, and ϵ_n' is the length of the greatest of all intervals $I_j^{(n)}$, $I_k'^{(n)}$, $j = 1, \cdots r_n$, $k = 1, \cdots s_n$. From (2.141) and (2.141'), we can deduce that $||x_{\beta_n} - x_{\varphi_n}||^2 = \sup_{P \in \mathfrak{s}} \langle (x_{\beta_n} - x_{\varphi_n})^2 \rangle_P \leq 2\, \epsilon_n'^2 \to 0$ for $n \to \infty$.

Thus we have proved that Rule 5 defines a mapping. As each element β of $\mathfrak{O}_b$ can be written as $\beta = f(\hat{\alpha})$, $\hat{\alpha} \in \hat{\mathfrak{O}}'$, we have to show that each sequence of type (2.139) converges to an element of $\mathfrak{A}'$ in order to prove that we deal with a mapping of $\mathfrak{O}_b$ into $\mathfrak{A}'$.

We can arrive at

$$\begin{aligned}&\langle (x_{\beta_m} - x_{\beta_n})^2 \rangle_P \\ &\quad = \sum_{j,k=1}^{r_n} (b_j^{(m)2} - 2 b_j^{(m)} b_k^{(n)} + b_k^{(n)2}) P^{\beta}(I_j^{(m)} \cap I_k^{(n)}) \\ &\quad \leq 2\, \epsilon_n^2. \qquad (2.143)\end{aligned}$$

where $m > n$, in the same way as we arrived from (2.141) to (2.141'). Hence,

$$\langle (x_{\beta_m} - x_{\beta_n})^2 \rangle_P \leq 2\, \epsilon_n^2 \to 0 \text{ for } n \to \infty, \qquad (2.144)$$

and therefore $x_{\beta_1}, x_{\beta_2}, \cdots$ is a Cauchy sequence in $\mathfrak{A}'$. As $\mathfrak{A}'$ is a B^*-algebra and thus complete, each Cauchy sequence must have a limit in $\mathfrak{A}'$.

To prove that the obtained mapping is injective, we will attack the problem along the same lines as we did approach a similar problem in proving Lemma 12b. Assume therefore that $x_\alpha = x_\beta$. Write $q^\alpha(\lambda) = \chi_{B_\lambda}^{\alpha}(\alpha)$ and $q^\beta(\lambda) = \chi_{B_\lambda}^{\beta}(\beta)$ where $B_\lambda = (-\infty, \lambda)$. Denote by $\xi^\alpha(\lambda)$ and $\xi^\beta(\mu)$ the elements of $\mathfrak{A}'$ adjoined by rules 1–5 to $q^\alpha(\lambda)$ and $q^\beta(\mu)$ respectively.

The B^*-algebra $\mathfrak{A}(\mathfrak{O}')$ is completely isomorphic[18] to some algebra of bounded operators on a Hilbert space. The bounded operators $A(\xi^\alpha(\lambda))$ and $A(\xi^\beta(\mu))$ corresponding to $\xi^\alpha(\lambda)$ and $\xi^\beta(\mu)$ respectively have to be, because of the isomorphism, indempotent elements, i.e., they are projection operators. Obviously, because of the definitions of x_α and x_β, and because of the complete isomorphism, the projector-valued functions $A(\xi^\alpha(\lambda))$ and $A(\xi^\beta(\mu))$ represent the spectral decompositions of $A(x_\alpha)$ and $A(x_\beta)$ respectively. As the Spectral Theorem says that a spectral decomposition is unique, it follows that $A(\xi^\alpha(\lambda)) = A(\xi^\beta(\lambda))$ and therefore $\xi^\alpha(\lambda) = \xi^\beta(\lambda)$, $\lambda \in \mathbf{R}^1$. According to Lemma 11, which asserts that Rule 3 defines an injective mapping of the set $Q_3(\mathfrak{O}')$[19] into $\mathfrak{A}_4'$ and therefore into $\mathfrak{A}'$, we can conclude that $q^\alpha(\lambda) = q^\beta(\lambda)$, $\lambda \in \mathbf{R}^1$. But then we can write

$$P^{\hat{\varphi}';\alpha;\hat{\varphi}''}(B' \times B \times B'') = P^{\hat{\varphi}';\beta;\hat{\varphi}''}(B' \times B \times B'') \qquad (2.145)$$

for all $\hat{\varphi}', \hat{\varphi}'' \in \hat{\mathfrak{O}}$, $B' \in \mathfrak{B}^{\hat{\varphi}'}$, $B'' \in \mathfrak{B}^{\hat{\varphi}''}$, and $B \in \mathfrak{I}^1$. According to Proposition 3 in I, Sec. 2.2, we can conclude that $\alpha = \beta$. Q.E.D.

Theorem 4: The mapping of $\mathfrak{O}_b$ into $\mathfrak{A}(\mathfrak{O}')$ (defined by Rule 5) determines a complete isomorphism of each normed commutative real algebra $\mathfrak{C} \in \mathfrak{M}_c^b$ of compatible bounded observables into a subalgebra $\mathfrak{A}(\mathfrak{C})$ of $\mathfrak{A}$.

Proof: It should be obvious that the algebra $\mathfrak{A}(\mathfrak{O}')$ has been so constructed, and the rules 1 to 5 so formulated, that the mapping of each $\mathfrak{C} \in \mathfrak{M}_c^b$ into $\mathfrak{A}(\mathfrak{O}')$ is a homomorphism which maps each element of $\mathfrak{C}$ into a Hermitian element of $\mathfrak{A}(\mathfrak{O}')$. Theorem 5

[18] Reference 12, p. 314.

[19] Note that we can write $\alpha = f(\hat{\varphi})$, $\hat{\varphi} \in \hat{\mathfrak{O}}'$, $f(\lambda) \in \mathfrak{F}_B^{\hat{\alpha}}$, and therefore $q^\alpha(\lambda) = \chi_{B_\lambda}^{\hat{\alpha}}(\hat{\varphi})$, $B_\lambda' = f^{-1}(B_\lambda)$. Therefore we have $q^\alpha(\lambda) \in Q_3(\mathfrak{O}')$ (and similarly $q^\beta(\lambda) = Q_3(\mathfrak{O}')$) for all $\lambda \in \mathbf{R}^1$.

tells us that this mapping is injective, i.e., we deal with an isomorphism. The proof that this is a complete isomorphism will be completed if we show that the norm of each element α of $\mathcal{C}$, which coincides with the bound $||\alpha||$ of the observable, coincides with $||x_\alpha||$. We have

$$\langle x_\alpha^2\rangle_P = \int_{-\infty}^{+\infty} \lambda^2\, dP^\alpha(\lambda) = \int_{-||\alpha||}^{+||\alpha||} \lambda^2\, dP^\alpha(\lambda)$$

$$\leq ||\alpha||^2 \int_{-\infty}^{+\infty} dP^\alpha(\lambda) = ||\alpha||^2. \qquad (2.146)$$

Therefore

$$||x_\alpha|| = \sup_{P\in\mathcal{S}} (\langle x_\alpha^2\rangle_P)^{\frac{1}{2}} \leq ||\alpha||. \qquad (2.147)$$

On the other hand, according to the definition of the bound of an observable, each interval $[||\alpha|| - \epsilon, ||\alpha||]$ in $\mathbf{R}^\alpha$ has to be a nonzero interval for each $0 < \epsilon < ||\alpha||$. Hence, on the basis of Ax. IX, there should exist for each such ϵ at least one such $P_\epsilon \in \mathcal{S}$ that $P_\epsilon^\alpha((||\alpha|| - \epsilon, ||\alpha||)) = 1$. Therefore

$$||x_\alpha||^2 = \int_{-||\alpha||}^{+||\alpha||} \lambda^2\, dP_\epsilon^\alpha(\lambda) \geq (||\alpha|| - \epsilon)^2. \qquad (2.148)$$

As ϵ is arbitrarily small, we get

$$||x_\alpha||^2 \geq \sup_{0<\epsilon<||\alpha||} \langle x_\alpha^2\rangle_{P_\epsilon}$$

$$= \sup_{0<\epsilon\leq||\alpha||} (||\alpha|| - \epsilon)^2 = ||\alpha||^2. \qquad (2.149)$$

From (2.147) and (2.149) we deduce that

$$||x_\alpha|| = ||\alpha|| \qquad (2.150)$$

and our assertion is proved. Q.E.D.

Lemma 13: To each ordered pair $(\hat{\alpha}, B)$, $\hat{\alpha} \in \hat{\mathcal{O}}'$, $B \in \mathfrak{B}^{\hat{\alpha}}$,[20] can be assigned an element $x_B^{\hat{\alpha}}$ of $\mathfrak{A}(\mathcal{O}')$ in such a fashion that $x_B^{\hat{\alpha}}$ is the element of $\mathfrak{A}_4^{(0)}(\mathcal{O}')$ assigned by the rules 1–5 to the question $\chi_B^{\hat{\alpha}}(\hat{\alpha})$ in case that $\{\hat{\alpha}\} = C$, and the linear manifold in $\mathfrak{A}(\mathcal{O}')$ spanned by all elements $x_B^{\hat{\alpha}}$, $\hat{\alpha} \in \hat{\mathcal{O}}'$, $B \in \mathfrak{B}^{\hat{\alpha}}$, is identical with $\mathfrak{A}_4^{(0)}(\mathcal{O}')$, i.e., the general element of $\mathfrak{A}_4^{(0)}(\mathcal{O}')$ has the form

$$x = a_1 x_{B_1}^{\hat{\alpha}_1} + \cdots + a_n x_{B_n}^{\hat{\alpha}_n}, \qquad \hat{\alpha}_1, \cdots \hat{\alpha}_n \in \hat{\mathcal{O}}',$$

$$B_1 \in \mathfrak{B}^{\hat{\alpha}_1}, \cdots B_n \in \mathfrak{B}^{\hat{\alpha}_n}. \qquad (2.151)$$

For such an x we have

$$\langle x\rangle_P = a_1 P^{\hat{\alpha}_1}(B_1) + \cdots + a_n P^{\hat{\alpha}_n}(B_n). \qquad (2.152)$$

Proof: Pick up a fundamental set $\mathcal{O}'$ and any n questions $\chi_{B_1}^{\alpha_1}(\alpha_1), \cdots \chi_{B_n}^{\alpha_n}(\alpha_n)$, $\alpha_k \in \mathcal{O}'$, $B_k \in \mathfrak{B}^{(\alpha_k)}$, $k = 1, \cdots n$. We will denote with $x_B^{\hat{\alpha}}$, $B = B_1 \times \cdots \times B_n$ the element of $\mathfrak{A}_4^{(0)}(\mathcal{O}')$ assigned to the element of $\mathfrak{A}_2(\mathcal{O}')$ containing the ordered product $\chi_{B_1}^{\alpha_1}(\alpha_1) \circ \cdots \circ \chi_{B_n}^{\alpha_n}(\alpha_n)$. To any set B of the form $B = B_1 \cup \cdots \cup B_k$, $B_1, \cdots, B_k \in \mathfrak{B}^{(\alpha_1)} \times \cdots \times B^{(\alpha_n)}$, $B_i \cap B_j = \varnothing$ for $i \neq j$, the element

$$x_B^{\hat{\alpha}} = x_{B_1}^{\hat{\alpha}} + \cdots + x_{B_n}^{\hat{\alpha}}$$

is adjoined. It is obvious, from the definition (2.40), that

$$\langle x_B^{\hat{\alpha}}\rangle_P = P^{\alpha_1,\cdots\alpha_n}(B).$$

Denote the class of all sets of the form $B_1 \cup \cdots \cup B_k, B_1, \cdots B_k \in \mathfrak{B}^{(\alpha_1)} \times \cdots \times \mathfrak{B}^{(\alpha_n)}, k = 1, 2, \cdots$, by $\overset{\circ}{\mathfrak{B}}{}^{(\alpha_1,\cdots\alpha_n)}$. According to Ref. 12, Theorem E, Sec. 33, $\overset{\circ}{\mathfrak{B}}{}^{\hat{\alpha}}$ is a Boolean-algebra.

On the other hand, Theorem B, Sec. 6 in Ref. 12, tells us that the monotone class of sets generated by $\overset{\circ}{\mathfrak{B}}{}^{\hat{\alpha}}$ coincides with the Boolean σ-algebra generated by $\overset{\circ}{\mathfrak{B}}{}^{\hat{\alpha}}$, i.e., with $\mathfrak{B}^{\hat{\alpha}}$. If $B_1, B_2, \cdots \in \overset{\circ}{\mathfrak{B}}{}^{\hat{\alpha}}$ is a monotone sequence and $\lim_{k\to\infty} B_k = B$, then we can show that the sequences $x_{B_1}^{\hat{\alpha}}, x_{B_2}^{\hat{\alpha}}, \cdots$ and $(x_{B_1}^{\hat{\alpha}})^*, (x_{B_2}^{\hat{\alpha}})^*, \cdots$ are Cauchy sequences in the strong topology of $\mathfrak{A}_4'$ which satisfy the conditions of the Restrictive Rule in Sec. 2.4. Namely, we get as in similar previous cases in this section that

$$\langle (x_{B_m}^{\hat{\alpha}} - x_{B_n}^{\hat{\alpha}})^*(x_{B_m}^{\hat{\alpha}} - x_{B_n}^{\hat{\alpha}})\rangle_P$$

$$= P^{\hat{\alpha}^*;\hat{\alpha}}(B_m^* \times B_m) - P^{\hat{\alpha}^*;\hat{\alpha}}(B_m^* \times B_n)$$

$$- P^{\hat{\alpha}^*;\hat{\alpha}}(B_n^* \times B_m) + P^{\hat{\alpha}^*;\hat{\alpha}}(B_n^* \times B_n). \qquad (2.153)$$

Each of the terms on the right-hand side of (2.153) differs arbitrarily little from $P^{\hat{\alpha}^*;\hat{\alpha}}(B^* \times B)$ (because of the continuity from above and below of the measure $P^{\hat{\alpha}^*;\hat{\alpha}}$) for m and n sufficiently great, and therefore $x_{B_1}^{\hat{\alpha}}, x_{B_2}^{\hat{\alpha}}, \cdots$ is a Cauchy sequence in the strong topology; a similar procedure proves the same for $(x_{B_1}^{\hat{\alpha}})^*, (x_{B_2}^{\hat{\alpha}})^*, \cdots$. As

$$||x_{B_k}^{\hat{\alpha}}||^2 = \sup_{P\in\mathcal{S}} P^{\hat{\alpha}^*;\hat{\alpha}}(B_k^* \times B_k)$$

and $P^{\hat{\alpha}^*;\hat{\alpha}}(B_k^* \times B_k) \to P^{\hat{\alpha}^*;\hat{\alpha}}(B^* \times B)$ when

[20] Strictly speaking, we have been able to carry through each step of the proof for the case $\{\hat{\alpha}\} \neq C$ only when $B \in \overset{\circ}{\mathfrak{B}}{}^{\hat{\alpha}}$ (see footnote 21). However, usually all arguments in this paper using Lemma 13 as is stated can be replaced by somewhat longer arguments using the weaker form of this lemma in which $\overset{\circ}{\mathfrak{B}}{}^{\hat{\alpha}}$ replaces $\mathfrak{B}^{\hat{\alpha}}$ in case that $\{\hat{\alpha}\} \neq C$. All these modified arguments would make use of the fact that $\mathfrak{B}^{\hat{\alpha}}$ is the monotone class generated by $\overset{\circ}{\mathfrak{B}}{}^{\hat{\alpha}}$ (Lemma 11). The only exception might be the proof of Proposition 1 is Sec. 3.2, We believe, however, that this lemma is true as is stated and that it is not necessary to introduce the additional axioms that $\sup_{P\in\mathcal{S}} P^{\hat{\alpha}^*;\hat{\alpha}}(B^* \times B) < +\infty$ for any $B \in \mathfrak{B}^{\hat{\alpha}}$.

$k \to \infty$, while $\sup_{P\in\mathcal{S}} P^{\hat{a}^*;\hat{a}}(B^* \times B) < +\infty$,[21] it is obvious that all the conditions of the Restrictive Rule are satisfied and $x_B^{\hat{a}}$ exists. Because of Lemma 9 in Sec. 2.4, we have

$$\langle x_B^{\hat{a}}\rangle_P = \lim_{k\to\infty} \langle x_{B_k}^{\hat{a}}\rangle_P = P^{\hat{a}}(B). \qquad (2.154)$$

Thus, we have shown that, for any $\hat{a} \in \hat{\mathfrak{O}}$ and any $B \in \mathfrak{B}^{\hat{a}}$, there is an element $x_B^{\hat{a}}$ of $\mathfrak{A}_4^f$, and therefore an element of $\mathfrak{A}^f$ which will be denoted also by $x_B^{\hat{a}}$, such that

$$\langle x_B^{\hat{a}}\rangle_P = P^{\hat{a}}(B). \qquad (2.155)$$

Considering the form of the sequence of elements of $\mathfrak{A}_3^f$, given by (2.83), which are submitted to the Restrictive Rule, it is obvious that we have proved our lemma entirely if we show that any sequence of elements of $\mathfrak{A}_4^f$, having the form

$$x_{B_{1k}}^{\hat{a}} \cdots x_{B_{nk}}^{\hat{a}}, \quad \hat{a} \in \hat{\mathfrak{O}}^f, \quad k = 1, 2, \cdots \qquad (2.156)$$

is convergent in the strong topology when all sequences $B_{j1}, B_{j2}, \cdots$ $(j = 1, \cdots n)$ are monotone sequences of sets from $\overset{\circ}{\mathfrak{B}}{}^{\hat{a}}$. To simplify the notation, consider the case when $n = 2$ and when the sequence $B_{11}, B_{12}, \cdots \in \overset{\circ}{\mathfrak{B}}{}^{\hat{a}}$ is increasing while $B_{21}, B_{22}, \cdots \in \overset{\circ}{\mathfrak{B}}{}^{\hat{a}}$ is decreasing; this case possesses all the significant features of the most general case.

Write $B_k' = \bigcup_{i=1}^k B_{1i}$, $B_k'' = \bigcap_{i=1}^k B_{2i}$. The sequence $B_1'', B_2'', \cdots$ is a decreasing sequence of sets from $\overset{\circ}{\mathfrak{B}}{}^{\hat{a}}$. If we write $B_k'' = B_k^{(1)} \cup \cdots \cup B_k^{(s_k)}$ $(B_k^{(i)} \bigcap B_k^{(j)} = \varnothing$ if $i \neq j)$, where $B_k^{(1)}, \cdots B_k^{(s_k)} \in \mathfrak{B}^{(\alpha_1)} \times \cdots \times \mathfrak{B}^{(\alpha_r)}$, $\hat{a} = (\alpha_1, \cdots \alpha_r)$, then, obviously, B_{k+1}'' can also be written as a union of disjoint sets, belonging to $B^{(\alpha_1)} \times \cdots \times B^{(\alpha_r)}$, which are of the following form:

$$B_{k+1}'' = \bigcup_{i=1}^{s_k} (B_{k1}^{(i)} \cup \cdots \cup B_{km_i}^{(i)}),$$
$$B_{kp}^{(i)} \subset B_k^{(i)}, \qquad p = 1, \cdots m.$$

This indicates that we can write $B_2 = \bigcap_{k=1}^\infty B_k''$ in the form of a union of an, at most, enumerable number of sets belonging to $\mathfrak{B}^{(\alpha_1)} \times \cdots \times \mathfrak{B}^{(\alpha_r)}$. Therefore, we will have $B_2 = \bigcup_{k=1}^\infty B_k'''$ where $B_1''' \subset B_2''' \subset \cdots$ is an increasing sequence of sets belonging to $\overset{\circ}{\mathfrak{B}}{}^{\hat{a}_3}$. Thus, according to our earlier considerations, $x_{B_k'}^{\hat{a}_1} x_{B_k''}^{\hat{a}_2}$ will converge towards $x_B^{(\hat{a}_1,\hat{a}_2)}$ in the strong topology when $k \to \infty$.

On the other hand, the sequence $(B_k' \times B_k'') - (B_k' \times B_k''') = B_k' \times (B_k'' - B_k''')$ is a decreasing sequence of sets in $\overset{\circ}{\mathfrak{B}}{}^{(\hat{a},\hat{a})}$ tending to the empty set. This implies that $x_{B_k'}^{\hat{a}}(x_{B_k''}^{\hat{a}} - x_{B_k'''}^{\hat{a}})$ tends to $\mathbf{0}$ in the strong topology when $k \to \infty$. Thus, we conclude that $x_B^{(\hat{a},\hat{a})}$ is the strong limit of the sequence $x_{B_k'}^{\hat{a}} x_{B_k''}^{\hat{a}}$, $k = 1, 2, \cdots$. Q.E.D.

Theorem 5: Denote by $\mathfrak{A}_s(\mathfrak{O}^f)$ the closure of $\mathfrak{A}(\mathfrak{O}^f)$ in the strong topology. All the algebras $\mathfrak{A}_s(\mathfrak{O}^f)$ determined by different fundamental sets of observables are indistinguishable from an algebraic point of view, i.e., they are *-isomorphic.

Proof: Take any two fundamental sets $\mathfrak{O}_1^f$ and $\mathfrak{O}_2^f$ and consider the images $\mathfrak{g}(\mathfrak{O}_1^f)$ and $\mathfrak{g}(\mathfrak{O}_2^f)$ of the set $\mathfrak{O}_b$ in the B^*-algebras $\mathfrak{A}_s(\mathfrak{O}_1^f)$ and $\mathfrak{A}_s(\mathfrak{O}_2^f)$ respectively. Already from (2.150) we can conclude that

$$||x_\beta|| = ||y_\beta||, \quad x_\beta \in \mathfrak{g}(\mathfrak{O}_1^f), \quad y_\beta \in \mathfrak{g}(\mathfrak{O}_2^f) \qquad (2.157)$$

is true for any $\beta \in \mathfrak{O}_b$. From Lemma 12 we can easily derive that

$$\langle x_\beta\rangle_P = \int_{-\infty}^{+\infty} \lambda\, dP^\beta(\lambda) = \langle y_\beta\rangle_P,$$
$$x_\beta \in \mathfrak{g}(\mathfrak{O}_1^f), \qquad y_\beta \in \mathfrak{g}(\mathfrak{O}_2^f) \qquad (2.158)$$

again for any $\beta \in \mathfrak{O}_b$. The mapping $x_\beta \to y_\beta = \varphi(x^\beta)$ is obviously an injective mapping (having therefore an inverse) of $\mathfrak{g}(\mathfrak{O}_1^f)$ *onto* $\mathfrak{g}(\mathfrak{O}_2^f)$, which preserves the norm and the value of any of the functionals $\langle x\rangle_P = \langle\varphi(x)\rangle_P$, $x \in \mathfrak{g}(\mathfrak{O}_1^f)$.

Pick up a monotone sequence $B_1, B_2, \cdots$ of sets from $\overset{\circ}{\mathfrak{B}}{}^{\hat{a}}$, $\hat{a} \in \mathfrak{O}_1^f$, having as a limit a set B in $\mathfrak{B}^{\hat{a}}$. According to Lemma 13, the corresponding sequence $x_{B_1}^{\hat{a}}, x_{B_2}^{\hat{a}}, \cdots$ of elements of $\mathfrak{A}_4^{(0)}(\mathfrak{O}_1^f)$ converges strongly to an element $x_B^{\hat{a}} \in \mathfrak{A}_4^{(0)}(\mathfrak{O}_1^f)$. As we can write each $x_{B_k}^{\hat{a}}$, $k = 1, 2, \cdots$, in the form of a finite sum of products of elements of $\mathfrak{A}(\mathfrak{O}_2^f)$ assigned to questions, we can denote by $y_{B_k}^{\hat{a}}$ the elements of $\mathfrak{A}(\mathfrak{O}_2^f)$ defined as the same sums of products of elements of $\mathfrak{g}(\mathfrak{O}_2^f)$ adjoined by the mapping φ to the corresponding elements of $\mathfrak{g}(\mathfrak{O}_2^f)$. (Note that these $y_{B_k}^{\hat{a}}$ will not belong

[21] We can easily prove that $\sup_{P\in\mathcal{S}} P^{\hat{a}^*;\hat{a}}(B^* \times B) < +\infty$ in case that $B \in \overset{\circ}{\mathfrak{B}}{}^{\hat{a}}$. Take first $B \in \mathfrak{B}^{(\alpha_1)} \times \ldots \times \mathfrak{B}^{(\alpha_n)}$, $\hat{a} = (\alpha_1, \ldots \alpha_n)$. In case that $n = 1$, we have $P^{\alpha;\alpha}(B \times B) = P^\alpha(B) \leq 1$. Using (2.83) (Hypothesis 2), we get that, for any $\beta \in \mathfrak{O}^f$, $B_1 \in \mathfrak{B}^{(\beta)}$, $P^{\hat{a}^*;\beta;\beta;\hat{a}}(B^* \times B_1 \times B_1 \times B) \leq P_1^{\beta;\beta}(B_1 \times B_1) \times P^{\hat{a}^*;\hat{a}}(B^* \times B) \leq P^{\hat{a}^*;\hat{a}}(B^* \times B)$. Therefore, by induction, we prove that $P^{\hat{a}^*;\hat{a}}(B^* \times B) \leq 1$ for any $P \in \mathcal{S}$, $\alpha_1, \ldots \alpha_n \in \mathfrak{O}^f$ and $B \in \mathfrak{B}^{(\alpha_1)} \times \ldots \times \mathfrak{B}^{(\alpha_n)}$. If we have $B = B_1 \cup B_2$, where $B_1 \cap B_2 = \varnothing$ and $B_1, B_2 \in \mathfrak{B}^{(\alpha_1)} \times \ldots \times \mathfrak{B}^{(\alpha_n)}$, then, by using the above result, we get easily that

$$P^{\hat{a}^*;\hat{a}}(B^* \times B) < 2(1 + |\, P^{\hat{a}^*;\hat{a}}(B_1^* \times B_2)\,|\,).$$

By applying Cauchy–Schwartz inequality on $\langle A_{\hat{a}}{}^*(B_1{}^*) A_{\hat{a}}(B_2)\rangle_P$, where $A_{\hat{a}}(B)$ is the element of $\mathfrak{A}_3{}^f$ corresponding to the question $\chi_B{}^{\hat{a}}$ (see Sec. 2.4), we get

$$|\, P^{\hat{a}^*;\hat{a}}(B_1^* \times B_2)\,| \leq P^{\hat{a}^*;\hat{a}}(B_1^* \times B_1)\, P^{\hat{a}^*;\hat{a}}(B_2^* \times B_2).$$

This and the previous formula enable us to conclude that $\sup_{P\in\mathcal{S}} P^{\hat{a}^*;\hat{a}}(B^* \times B) < +\infty$ for the above introduced B. By induction, we can prove the same result for any $B \in \overset{\circ}{\mathfrak{B}}{}^{\hat{a}}$. On the other hand, we do not know whether the same is true for any element of $\mathfrak{B}^{\hat{a}}$, but, as mentioned in the previous footnote, this does not essentially influence the later results.

in general to $\mathfrak{A}_4^{(0)}(\mathcal{O}_2')$, because, in general, $\hat{a} \notin \mathcal{O}_2'$ when $\hat{a} \in \mathcal{O}_1'$.) To show that $y_{B_1}^{\hat{a}}, y_{B_2}^{\hat{a}}, \cdots$ converges in the strong topology, we have to prove that

$$\langle y_{B_k}^{\hat{a}} \rangle_P = \langle x_{B_k}^{\hat{a}} \rangle_P,$$
$$\langle (y_{B_m}^{\hat{a}} - y_{B_n}^{\hat{a}})^* (y_{B_m}^{\hat{a}} - y_{B_n}^{\hat{a}}) \rangle_P = \langle (x_{B_m}^{\hat{a}} - x_{B_n}^{\hat{a}})^* (x_{B_m}^{\hat{a}} - x_{B_n}^{\hat{a}}) \rangle_P. \tag{2.159}$$

Relations (2.159) are established if we show that, in general,

$$\langle y_{B_1}^{\alpha_1} \cdots y_{B_n}^{\alpha_n} \rangle_P = P^{\alpha_1;\cdots;\alpha_n}(B_1 \times \cdots \times B_n),$$
$$\alpha_1, \cdots \alpha_n \in \mathcal{O}, \qquad B_1 \in \mathfrak{B}^{(\alpha_1)}, \cdots B_n \in \mathfrak{B}^{(\alpha_n)}. \tag{2.161}$$

If $\alpha_1 = f_1(\hat{\beta}_1), \cdots \alpha_n = f_n(\hat{\beta}_n), \hat{\beta}_1, \cdots \hat{\beta}_n \in \mathcal{O}_2'$ then (2.161) is proved by noticing that $y_{B_1}^{\alpha_1} \cdots y_{B_n}^{\alpha_n} = y_{f_1^{-1}(B_1)}^{\hat{\beta}_1} \cdots y_{f_n^{-1}(B_n)}^{\hat{\beta}_n}$ and that we have, according to Lemma 13,

$$\langle y_{f_1^{-1}(B_1)}^{\hat{\beta}_1} \cdots y_{f_n^{-1}(B_n)}^{\hat{\beta}_n} \rangle_P = P^{\hat{\beta}_1;\cdots;\hat{\beta}_n}(f_1^{-1}(B_1) \times \cdots \times f_n^{-1}(B_n)). \tag{2.162}$$

Thus we conclude that $y_{B_1}^{\hat{a}}, y_{B_2}^{\hat{a}}, \cdots$ converges in the strong topology to an element of $\mathfrak{A}_s(\mathcal{O}_2')$ [but not necessarily to an element of $\mathfrak{A}(\mathcal{O}_2')$], which will be denoted by $y_B^{\hat{a}}$. As the linear manifold spanned in $\mathfrak{A}_s(\mathcal{O}_1')$ by all $x_B^{\hat{a}}$, $\hat{a} \in \mathcal{O}_1'$, $B \in \mathfrak{B}^{\hat{a}}$, is, according to Lemma 13, identical with $\mathfrak{A}_4^{(0)}(\mathcal{O}_1')$ and the algebra $\mathfrak{A}_4^{(0)}(\mathcal{O}_1')$ is everywhere dense (in the strong topology) in $\mathfrak{A}_s(\mathcal{O}_1')$, we can say that there exists a *-isomorphism between $\mathfrak{A}_s(\mathcal{O}_1')$ and a subalgebra of $\mathfrak{A}_s(\mathcal{O}_2')$. By reversing the roles of $\mathcal{O}_1'$ and $\mathcal{O}_2'$, we may conclude that this *-isomorphism is between $\mathfrak{A}_s(\mathcal{O}_1')$ and $\mathfrak{A}_s(\mathcal{O}_2')$. If we extend the domain of definition of $\langle x \rangle_P$ and $\langle y \rangle_P$, $x \in \mathfrak{A}(\mathcal{O}_1')$, $y \in \mathfrak{A}(\mathcal{O}_2')$, in a natural manner to the entire $\mathfrak{A}_s(\mathcal{O}_1')$ and $\mathfrak{A}_s(\mathcal{O}_2')$ respectively, then it is easy to see that

$$\langle x \rangle_P = \langle \varphi(x) \rangle_P, \qquad x \in \mathfrak{A}_s(\mathcal{O}_1'), \qquad \varphi(x) \in \mathfrak{A}_s(\mathcal{O}_2'), \tag{2.163}$$

and that the functionals $\langle x \rangle_P$, $P \in \mathcal{S}$, will be positive and defined everywhere in $\mathfrak{A}_s(\mathcal{O}_1')$. If we define

$$||x|| = \sup_{P \in \mathcal{S}} (\langle x^* x \rangle_P)^{\frac{1}{2}}, \qquad x \in \mathfrak{A}_s(\mathcal{O}_1'),$$
$$||y|| = \sup_{P \in \mathcal{S}} (\langle y^* y \rangle_P)^{\frac{1}{2}}, \qquad y \in \mathfrak{A}_s(\mathcal{O}_2'), \tag{2.164}$$

this functional will not be necessarily finite everywhere, i.e., it will not be necessarily a norm. Q.E.D.

Finally, we can summarize the most important results of this section in the following theorem:

Theorem 6: The language $\mathfrak{L}$ given by the Axioms I to IX in I, Sec. 2.2., can be embedded in a B^*-algebra $\mathfrak{A}(\mathcal{O}')$ if the set of all physical states is such that Axiom X is also fulfilled. This embedding is such that there is an algebra $\mathfrak{A}_4(\mathcal{O}')$ which is everywhere dense in $\mathfrak{A}(\mathcal{O}')$ in the uniform topology defined by the norm of $\mathfrak{A}(\mathcal{O}')$. The algebra $\mathfrak{A}_4(\mathcal{O}')$ itself is the closure in a strong topology (Sec. 2.4) of the algebra generated by the image in the algebra $\mathfrak{A}(\mathcal{O}')$ of the set $\mathfrak{Q}(\mathcal{O}')$ of all the questions on the fundamental set $\mathcal{O}'$. The algebra $\mathfrak{A}(\mathcal{O}')$ will, in general, depend (in a physically unessential way) on the fundamental set $\mathcal{O}'$ of observables which has been taken as the starting point of the construction of $\mathfrak{A}(\mathcal{O}')$.

3. THE HILBERT-SPACE REPRESENTATION OF THE QUANTUM-MECHANICAL LANGUAGE $\mathfrak{L}$

3.1. The physical states as positive functionals on the B^*-algebra $\mathfrak{A}(\mathcal{O}')$

In the last section, we have proved that the set of bounded observables $\mathcal{O}_b$ can be embedded in the B^*-algebra $\mathfrak{A}(\mathcal{O}')$. A well-known theorem of the theory of Banach algebras[22] tells us that any B^*-algebra is completely isomorphic[23] to some subalgebra of the algebra $\mathfrak{B}(\mathfrak{H})$ of all bounded operators on a Hilbert space $\mathfrak{H}$ (which is therefore a C^*-algebra). On the first sight, it would be sufficient to take such a Hilbert space $\mathfrak{H}$ to provide a representation for our language $\mathfrak{L}$. However, a careful study of the proof of the above mentioned theorem (see especially the proof of theorem 1, p. 260 in Ref. 12) will soon convince us that the Hilbert space $\mathfrak{H}$, introduced in the course of this proof, does not provide at all an injective mapping of the positive indecomposable[24] functionals on a B^*-algebra into the set $\mathfrak{H}$ of Hilbert-space vectors. Therefore we will proceed in our own way.

Lemma 1: The mapping $P \to \langle x \rangle_P$, $P \in \mathcal{S}$, $x \in \mathfrak{A}_4(\mathcal{O}')$, is an injective mapping of the set $\mathcal{S}$ of all physical states into the set $\mathfrak{K}(\mathfrak{A}_4')$ of all positive linear functionals on $\mathfrak{A}_4'$, i.e., if $P_1 \neq P_2$ then $\langle x \rangle_{P_1} \neq \langle x \rangle_{P_2}$, $x \in \mathfrak{A}_4'$.

Proof: It is self-evident that $\langle x \rangle_{P_1} \neq \langle x \rangle_{P_2}$ implies $P_1 \neq P_2$ because it is evident from the construction of $\langle x \rangle_P$ [see (2.40)] that each P determines $\langle x \rangle_P$ uniquely. Hence, we are indeed dealing with a mapping.

Now, it follows directly from Proposition 1, Sec.

[22] Reference 12, Theorem 5, p. 314.
[23] Reference 12, p. 189.
[24] Reference 12, p. 265.

2.3 in I, that if $P_1 \neq P_2$ then there exists an n-tuple $(\alpha_1, \cdots \alpha_n) \in \hat{\mathfrak{O}}^f$ and Borel sets $B_1, \cdots B_n \in \mathfrak{B}^1$ such that

$$P_1^{\alpha_1,\cdots\alpha_n}(B_1 \times \cdots \times B_n) \neq P_2^{\alpha_1,\cdots\alpha_n}(B_1 \times \cdots \times B_n). \quad (3.1)$$

If x_0 is an element of $\mathfrak{A}_4'$ corresponding to an element $A_0 \in \mathfrak{A}_2(\mathfrak{O}^f)$ which contains the polynomial form on $\mathfrak{Q}^f$:

$$1q_1 \circ \cdots \circ q_n, \qquad q_1 = \chi_{B_1}^{\alpha_1}(\alpha_1), \cdots q_n = \chi_{B_n}^{\alpha_n}(\alpha_n)$$

then, according to (2.40) and (2.42),

$$\langle x_0\rangle_{P_i} = \langle A_0\rangle_{P_i} = P_i^{\alpha_1;\cdots;\alpha_n}(B_1 \times \cdots \times B_n), \quad i = 1, 2, \quad (3.2)$$

and therefore $\langle x_0\rangle_{P_1} \neq \langle x_0\rangle_{P_2}$. Q.E.D.

Lemma 2: The functional $\langle x\rangle_P$, assigned to $P \in \mathfrak{S}$ and defined on $\mathfrak{A}_4^{(0)}(\mathfrak{O}^f)$ by (3.2), is uniquely extendable on the completion $\mathfrak{A}(\mathfrak{O}^f)$ of the normed *-algebra $\mathfrak{A}_4(\mathfrak{O}^f)$.

Proof: According to Lemma 1, $\langle x\rangle_P$ is a positive linear functional, and therefore Cauchy-Schwartz inequality is applicable. Hence, if by 1 we denote the unity of $\mathfrak{A}(\mathfrak{O}^f)$, we can write,

$$|\langle x\rangle_P| = |\langle 1x\rangle_P| \leq (\langle 1^*1\rangle_P\langle x^*x\rangle_P)^{\frac{1}{2}} = (\langle x^*x\rangle_P)^{\frac{1}{2}} \leq ||x||, \quad x \in \mathfrak{A}_4^{(0)}(\mathfrak{O}^f). \quad (3.3)$$

Inequality (3.3) shows that $\langle x\rangle_P$ is a continuous functional on $\mathfrak{A}_4^{(0)}(\mathfrak{O}^f)$. Hence, it can be uniquely extended to a functional on the completion $\mathfrak{A}^f$ of $\mathfrak{A}_4^{(0)}(\mathfrak{O}^f)$, which can be shown, by standard methods, to be linear and positive. Q.E.D.

Lemmas 1 and 2 yield the following theorem:

Theorem 1: The functionals $\langle x\rangle_P (x \in \mathfrak{A}(\mathfrak{O}^f), P \in \mathfrak{S})$ defined as extensions of the positive functionals $\langle x\rangle_P$, $x \in \mathfrak{A}_4^{(0)}(\mathfrak{O}^f)$, on the everywhere in $\mathfrak{A}^f$ dense algebra $\mathfrak{A}_4^{(0)}(\mathfrak{O}^f)$, which is completely isomorphic to $\mathfrak{A}_4(\mathfrak{O}^f)$, are positive continuous linear functionals on $\mathfrak{A}(\mathfrak{O}^f)$. The set of these functionals will be denoted by $\mathfrak{K}^f$.

3.2. The set $\mathfrak{K}^f$ of positive functionals on $\mathfrak{A}(\mathfrak{O}^f)$ representing physical states

Following a widely adopted notation, we will denote by $\mathfrak{A}'(\mathfrak{O}^f)$ the space conjugated to $\mathfrak{A}(\mathfrak{O}^f)$, i.e., the space of all continuous linear functionals on the normed space $\mathfrak{A}(\mathfrak{O}^f)$.

Proposition 1: The set $\mathfrak{K}^f$ of all positive linear functionals representing physical states is convex and closed in the weak topology of $\mathfrak{A}'(\mathfrak{O}^f)$.

Proof: Let $\langle x\rangle_{P_1}, \langle x\rangle_{P_2}, \cdots (x \in \mathfrak{A}(\mathfrak{O}^f))$ be a Cauchy sequence in the weak topology of $\mathfrak{A}'(\mathfrak{O}^f)$ of elements of $\mathfrak{K}^f$. Then, for any $x \in \mathfrak{A}(\mathfrak{O}^f)$,

$$|\langle x\rangle_{P_m} - \langle x\rangle_{P_n}| < \epsilon \quad \text{for} \quad n, m > N(\epsilon, x), \quad (3.4)$$

and therefore $\lim_{n\to\infty}\langle x\rangle_{P_n}$ exists and defines a functional $f(x)$ on $\mathfrak{A}^f$. It is easy to check that $f(x)$ is a positive linear functional on $\mathfrak{A}^f$, and therefore a continuous one. We will show that it belongs to $\mathfrak{K}^f$.

According to Lemma 13 in Sec. 2.5, there exist for each $\hat{\alpha} \in \hat{\mathfrak{O}}^f$ and each $B \in \mathfrak{B}^{\hat{\alpha}}$ such elements $x_B^{\hat{\alpha}}$ of $\mathfrak{A}_4^{(0)}(\mathfrak{O}^f)$ that $\langle x_B^{\hat{\alpha}}\rangle_P = P^{\hat{\alpha}}(B)$. Therefore, on the basis of Proposition 4 in I, Sec. 3.3, we can conclude that $P_1, P_2, \cdots$ is a Cauchy sequence in the weak topology of $\mathfrak{S}$, because, for any $\hat{\alpha} \in \hat{\mathfrak{O}}^f$, $B \in \mathfrak{B}^{\hat{\alpha}}$, we have

$$|P_m^{\hat{\alpha}}(B) - P_n^{\alpha}(B)| = |\langle x_B^{\hat{\alpha}}\rangle_{P_m} - \langle x_B^{\hat{\alpha}}\rangle_{P_n}| < \epsilon \quad \text{for} \quad m, n > N(\epsilon, x_B^{\hat{\alpha}}). \quad (3.5)$$

According to Axiom VII, there exists a weak limit $P \in \mathfrak{S}$ of this sequence. Hence we will have

$$\langle x_B^{\hat{\alpha}}\rangle_P = f(x_B^{\hat{\alpha}}), \qquad \hat{\alpha} \in \mathfrak{O}^f, \qquad B \in \mathfrak{B}^{\hat{\alpha}}. \quad (3.6)$$

As Lemma 13 in Sec. 2.5 tells us that the linear manifold spanned by all $x_B^{\hat{\alpha}}$, $\hat{\alpha} \in \hat{\mathfrak{O}}^f$, $B \in \mathfrak{B}^{\hat{\alpha}}$, is identical with $\mathfrak{A}_4^{(0)}(\mathfrak{O}^f)$, we can write

$$\langle x\rangle_P = f(x), \qquad x \in \mathfrak{A}_4^{(0)}(\mathfrak{O}^f). \quad (3.7)$$

Due to the fact that both functional $\langle x\rangle_P$ and $f(x)$ are continuous linear functionals, we can deduce from (3.7) that $f(x) = \langle x\rangle_P$ everywhere on $\mathfrak{A}(\mathfrak{O}^f)$. Hence $\mathfrak{K}^f$ is indeed closed in the weak topology of $\mathfrak{A}'(\mathfrak{O}^f)$.

To prove the convexity of $\mathfrak{K}^f$ we have only to note that

$$\langle x\rangle_{tP_1+(1-t)P_2} = t\langle x\rangle_{P_1} + (1-t)\langle x\rangle_{P_2}, \quad 0 \leq t \leq 1 \quad (3.8)$$

which is easy to derive from (2.155). Relation (3.8) and Axiom VI show immediately that $\mathfrak{K}^f$ is convex. Q.E.D.

Proposition 2: The set $\mathfrak{K}^f$ is bicompact in the weak topology of $\mathfrak{A}'(\mathfrak{O}^f)$.

Proof: The set $\mathfrak{N}$ of all the functionals $f(x)$, $x \in \mathfrak{A}(\mathfrak{O}^f)$ in $\mathfrak{A}'(\mathfrak{O}^f)$ which satisfy the relation

$$|f(x)| \leq ||x|| \quad (3.9)$$

is, according to Proposition III, p. 56 of Ref. 12, a bicompact set in $\mathfrak{A}'(\mathfrak{O}^f)$. It can be immediately inferred from (3.7) that all the functionals from $\mathfrak{K}^f$ satisfy (3.9). Therefore $\mathfrak{K}^f$ is a subset of $\mathfrak{N}$, which

is closed (Proposition 1). Hence $\mathcal{K}^f$ is a biocompact set in the weak topology of $\mathfrak{A}'(\mathcal{O}^f)$. Q.E.D.

Once the convexity and bicompactness of $\mathcal{K}^f$ in $\mathfrak{A}'(\mathcal{O}^f)$ has been established, we can apply Krein-Milman theorem[25] and obtain the following proposition:

Proposition 3: The set $\mathcal{K}^f$ contains extremal points and is the smallest convex closed set which contains all the extremal points of $\mathcal{K}^f$.

Proposition 4: A continuous functional $f(x) \in \mathfrak{A}'(\mathcal{O}^f)$ is an extremal point of $\mathcal{K}^f$ if and only if it belongs to $\mathcal{K}^f$, i.e., $f(x) = \langle x \rangle_P$, $P \in \mathcal{S}$, and if the corresponding physical state P is a pure state, $P \in \mathcal{S}_0$.

Proof: If $P \in \mathcal{S}_0$ then, naturally, $f(x) = \langle x \rangle_P \in \mathcal{K}^f$ by definition. Assume that there are such $f_1, f_2 \in \mathcal{K}^f$, $0 < t < 1$ so that

$$f(x) = tf_1(x) + (1 - t)f_2(x). \tag{3.10}$$

Then $f_1(x) = \langle x \rangle_{P_1}$, $f_2(x) = \langle x \rangle_{P_2}$. By inserting in (3.10) an element $x_B^{\hat{a}}$ of $\mathfrak{A}_4^{(0)}(\mathcal{O}^f)$, where $\hat{a} \in \hat{\mathcal{O}}^f$, $B \in \mathfrak{B}^{\hat{a}}$, we obtain

$$P^{\hat{a}}(B) = tP_1^{\hat{a}}(B) + (1 - t)P_2^{\hat{a}}(B). \tag{3.11}$$

As $\hat{a}$ is any element of $\hat{\mathcal{O}}^f$ and B is any element of $\mathfrak{B}^{\hat{a}}$, (3.11) cannot be true if P is a pure state. Hence $f(x)$ is an extremal point of $\mathcal{K}^f$.

Take now any extremal point $f(x)$ of $\mathcal{K}^f$. According to Proposition 3, $f(x)$ belongs to $\mathcal{K}^f$. Therefore there is such a $P \in \mathcal{S}$ that $f(x) = \langle x \rangle_P$. Assume that P is not a pure state, i.e., there are such $P_1, P_2 \in \mathcal{S}$ that for any $\hat{a} \in \hat{\mathcal{O}}^f$, $B \in \mathfrak{B}^{\hat{a}}$, (3.11) is valid for some $0 < t < 1$. Then we have

$$\langle x_B^{\hat{a}} \rangle_P = f(x_B^{\hat{a}}) = t\langle x_B^{\hat{a}} \rangle_{P_1} + (1 - t)\langle x_B^{\hat{a}} \rangle_{P_2}, \qquad 0 < t < 1, \tag{3.12}$$

for any $x_B^{\hat{a}} \in \mathfrak{A}_4^{(0)}(\mathcal{O}^f)$ corresponding to some $\hat{a} \in \hat{\mathcal{O}}^f$, $B \in \mathfrak{B}^{\hat{a}}$. As the linear manifold spanned by all $x_B^{\hat{a}}$, $\hat{a} \in \hat{\mathcal{O}}^f$, $B \in \mathfrak{B}^{\hat{a}}$ is identical with $\mathfrak{A}_4^{(0)}(\mathcal{O}^f)$ (Lemma 13, Sec. 2.5), and therefore is everywhere dense in $\mathfrak{A}^f$, we conclude that (3.12) is true for any $x \in \mathfrak{A}^f$. But this would mean that $f(x)$ is not an extremal point of $\mathcal{K}^f$. Q.E.D.

We will call a positive functional $f(x) \in \mathcal{K}^f$ *indecomposable on* $\mathcal{K}^f$ if every (positive) functional $f_1(x) \in \mathcal{K}^f$ dominated by $f(x)$ is a multiple of $f(x)$, i.e., if

$$f_1(x^*x) \leq \lambda f(x^*x) \tag{3.13}$$

for some $\lambda > 0$ implies that $f_1(x) = \lambda_1 f(x)$. Note that this definition differs somewhat from the definition of an indecomposable positive functional,[26] because, in our definition, not all positive functionals from $\mathfrak{A}'(\mathcal{O}^f)$ are considered, but only the ones belonging to $\mathcal{K}^f$ (which does not necessarily consist all positive functionals).

Proposition 5: Each positive functional $f(x) \in \mathcal{K}^f$ indecomposable on $\mathcal{K}^f$ is an extremal point of $\mathcal{K}^f$. Conversely, each extremal point of $\mathcal{K}^f$ is an indecomposable functional on $\mathcal{K}^f$.

Proof: Assume that $f(x) \in \mathcal{K}^f$ is an indecomposable on $\mathcal{K}^f$ positive functional. Assume further that there exist functionals $f_1(x)$, $f_2(x) \in \mathcal{K}^f$, and a real number $0 < t < 1$ in terms of which $f(x)$ can be written in the form (3.10). As $f_2(x) \in \mathcal{K}^f$ and $t < 1$, it follows that $(1 - t)f_2(x)$ is a positive functional. Therefore, we have

$$f_1(x^*x) < \frac{1}{t} f(x^*x), \qquad f_1(x) \in \mathcal{K}^f. \tag{3.14}$$

But (3.14) contradicts the assumption that $f(x)$ is indecomposable on $\mathcal{K}^f$. Hence $f(x)$ must be an extremal point of $\mathcal{K}^f$.

Assume now that $f(x)$ is an extremal point of $\mathcal{K}^f$. Then, according to Proposition 3, $f(x) \in \mathcal{K}^f$. Let us assume that there is a functional $f_1(x) \in \mathcal{K}^f$ dominated by $f(x)$, i.e., staying in relation (3.13) to $f(x)$. Then, necessarily, $\lambda \geq 1$, because $f_1(\mathbf{1}^*\mathbf{1}) = f_1(\mathbf{1}) = 1 = f(\mathbf{1}) = f(\mathbf{1}^*\mathbf{1})$ ($\mathbf{1}$ denotes the unity of $\mathfrak{A}^f$). We can assume that $\lambda > 1$, because, if (3.13) is true for $\lambda = 1$, then it will be certainly true for $\lambda > 1$.

From (3.13) we derive that

$$f_2(x) = \frac{f(x) - 1/\lambda f_1(x)}{1 - 1/\lambda} \tag{3.15}$$

is a positive functional. If x_q is the element of $\mathfrak{A}^f$ corresponding to the question $q = \chi_B^{\hat{a}}(\hat{a})$ ($\hat{a} \in \hat{\mathcal{O}}$, $\{\hat{a}\} = C$, $B \in \mathfrak{B}^{\hat{a}}$), then, by inserting it in (3.15), we get

$$f_2(x_q) = \frac{1}{1 - 1/\lambda}\left(P^{\hat{a}}(B) - \frac{1}{\lambda} P_1^{\hat{a}}(B)\right), \quad 0 < \frac{1}{\lambda} < 1. \tag{3.16}$$

It can be easily verified that, for any $q \in \mathcal{Q}$, we have

$$x_q^* x_q = x_q^2 = x_q. \tag{3.17}$$

Therefore,

$$\frac{1}{1 - 1/\lambda}\left(P^{\hat{a}}(B) - \frac{1}{\lambda} P_1^{\hat{a}}(B)\right) \geq 0$$

for $\{\hat{a}\} = C$, and the conditions of Axiom VI(2) are

[25] Reference 12, p. 62.

[26] Reference 12, p. 265.

fulfilled. Hence, there is a physical state $P_2 \in \mathcal{S}$ such that

$$f_2(x_a) = P_2^{\hat{a}}(B) = \langle x_a \rangle_{P_2}, \quad \hat{a} \in \hat{\mathcal{O}}, \quad B \in \mathfrak{B}^{\hat{a}}. \tag{3.18}$$

We will also have, for any $\hat{a} \in \hat{\mathcal{O}}^f$, $B \in \mathfrak{B}^{\hat{a}}$, that

$$f_2(x_B^{\hat{a}}) = P_2^{\hat{a}}(B) = \langle x_B^{\hat{a}} \rangle_{P_2}.$$

As the linear manifold spanned by $x_B^{\hat{a}}$, $\hat{a} \in \hat{\mathcal{O}}^f$, $B \in \mathfrak{B}^{\hat{a}}$ is everywhere dense in $\mathfrak{A}^f$, we conclude that $f_2(x) = \langle x \rangle_{P_2}$, and therefore $f_2(x) \in \mathcal{K}^f$. But then (3.15) can be written in the form

$$f(x) = \frac{1}{\lambda} f_1(x) + \left(1 - \frac{1}{\lambda}\right) f_2(x), \tag{3.19}$$

where $f_1(x)$, $f_2(x) \in \mathcal{K}^f$ and $0 < 1/\lambda < 1$. As (3.19) is impossible if $f(x)$ is an extremal point of $\mathcal{K}^f$, we conclude that $f(x)$ is indecomposable on $\mathcal{K}^f$. Q.E.D.

The most important results of this section can be summarized in the following theorem:

Theorem 2: The set $\mathcal{K}_0^f$ of positive functionals on $\mathfrak{A}(\mathcal{O}^f)$ corresponding to pure physical states coincides with the set of all extremal points of the set $\mathcal{K}^f$ of all positive functionals $\langle x \rangle_P$ corresponding to physical states $P \in \mathcal{S}$. This set also coincides with the set of positive functionals which are indecomposable on $\mathcal{K}^f$.

3.3 The representation Hilbert space $\mathfrak{H}$ of the language $\mathcal{L}$

By formulating Theorem 1, Proposition I, and Theorem 2 of Ref. 12 to fit our purposes, we get the following Proposition:

Proposition 1: To every functional $\langle x \rangle_P \in \mathcal{K}^f$ corresponds a continuous cyclic representation $x \to A_x^P$ of the Banach algebra $\mathfrak{A}(\mathcal{O}^f)$ in a representation Hilbert space $\mathfrak{H}^P$. The cyclic vector Ψ^P is such that

$$\langle x \rangle_P = (\Psi^P, A_x^P \Psi^P). \tag{3.20}$$

If, for a given $P \in \mathcal{S}$, there are two representations of $\mathfrak{A}^f$, $x \to A_x^P$ and $x \to B_x^P$, with representation Hilbert spaces $\mathfrak{H}_1^P$, $\mathfrak{H}_2^P$, and cyclic vectors Ψ_1^P, Ψ_2^P respectively, for which

$$(\Psi_1^P, A_x^P \Psi_1^P) = \langle x \rangle_P = (\Psi_2^P, B_x^P \Psi_2^P), \tag{3.21}$$

then these two representations are equivalent, i.e., there exists an isometric operator V, with domain of definition $\mathfrak{H}_1$, and range $\mathfrak{H}_2$ such that

$$\mathfrak{H}_2^P = V\mathfrak{H}_1^P, \qquad B_x V = V A_x. \tag{3.22}$$

Theorem 3: There is such a Hilbert space $\mathfrak{H}$ for which the algebra $\mathfrak{A}(\mathcal{O}^f)$ is completely isomorphic to a C^*-algebra $\mathfrak{B}(\mathcal{O}^f)$ [which, in general, is a C^*-subalgebra of the algebra $\mathfrak{B}(\mathfrak{H})$ of all bounded operators on $\mathfrak{H}$], and there exists an injective mapping $P \to \Psi^P$ of the set $\mathcal{S}_0$ of all pure states into the Hilbert space $\mathfrak{H}$. This mapping is such that, for each

$$\langle x \rangle_P = (\Psi^P, A_x \Psi^P), \quad x \in \mathfrak{A}(\mathcal{O}^f), \quad A_x \in \mathfrak{B}(\mathcal{O}^f), \tag{3.23}$$

where A_x is the image of x in $\mathfrak{B}(\mathcal{O}^f)$. Besides, the linear manifold spanned by all vectors Ψ^P, $P \in \mathcal{S}_0$, is everywhere dense in $\mathfrak{H}$. Obviously $\mathfrak{H}$ provides a representation space for the language $\mathcal{L}$ because it follows from (3.23) that requirement (1.1) is satisfied and it will become evident from the construction of $\mathfrak{H}$ that (1.2) is also fulfilled.

Proof: Consider the set H of all Hilbert spaces $\mathfrak{H}^P$, $P \in \mathcal{S}_0$, which are defined on the basis of Proposition 1. Introduce a partial ordering in H by defining that $\mathfrak{H}^{P_1} \leq \mathfrak{H}^{P_2}$ if $\mathfrak{H}^{P_1}$ is isometrically isomorphic to a subspace $\mathfrak{H}'^{P_1}$ of $\mathfrak{H}^{P_2}$ in such a manner that there exists in relation to $\mathfrak{H}'^{P_1}$ a cyclic vector Ψ'^{P_1} of the image $\mathfrak{B}'^{P_2}(\mathcal{O}^f)$ in $\mathfrak{B}(\mathfrak{H}'^{P_1})$ of the C^*-algebra $\mathfrak{B}^{P_2}(\mathcal{O}^f)$[27] for which

$$\langle x \rangle_{P_1} = (\Psi'^{P_1}, A_x'^{P_2} \Psi'^{P_1}), \qquad x \in \mathfrak{A}(\mathcal{O}^f). \tag{3.24}$$

Here, $A_x'^{P_2}$ denotes the image $VA_x^{P_2}V^{-1}$ of $A_x^{P_2} \in \mathfrak{B}^{P_2}(\mathcal{O}^f)$ under the isometric mapping V from $\mathfrak{H}^{P_2}$ onto $\mathfrak{H}'^{P_1}$. Obviously, all the axioms of partial ordering are satisfied, in particular, $\mathfrak{H}^{P_1} \leq \mathfrak{H}^{P_2}$, $\mathfrak{H}^{P_2} \leq \mathfrak{H}^{P_1}$ imply that $\mathfrak{H}^{P_1} \cong \mathfrak{H}^{P_2}$ (meaning that these two spaces are isometrically isomorphic) and $\mathfrak{B}^{P_1}(\mathcal{O}^f) \cong \mathfrak{B}^{P_2}(\mathcal{O}^f)$ (meaning that these two C^*-algebras are equivalent).

Denote by H_0 the set of all finite direct sums $\mathfrak{H}^{P_i} \oplus \cdots \oplus \mathfrak{H}^{P_n}$ of elements of H, which are such that $\mathfrak{H}^{P_i} \leq \mathfrak{H}^{P_j}$ can be true only for $i = j$, and that there does not exist for $i \neq j$ such an $\mathfrak{H}^P \in H$ that $\mathfrak{H}^P \leq \mathfrak{H}^{P_i}$ and $\mathfrak{H}^P \leq \mathfrak{H}^{P_j}$. In case that, for n Hilbert spaces $\mathfrak{H}^{P_1}, \mathfrak{H}^{P_2}, \cdots \mathfrak{H}^{P_n} \in H$, the first condition is fulfilled but not the second, build a corresponding element of H_0 in the following way: denote by $\mathfrak{H}'$ the Hilbert space which is isometrically isomorphic to the subspace of $\mathfrak{H}^{P_1}$ spanned by all the vectors which belong to at least one $\mathfrak{H}'^P \subset \mathfrak{H}^{P_1}$ for which $\mathfrak{H}^P \leq \mathfrak{H}^{P_1}$ (in the partial ordering of H). Then build the space $\mathfrak{H}' \oplus (\mathfrak{H}^{P_1} \ominus \mathfrak{H}') \oplus (\mathfrak{H}^{P_2} \ominus \mathfrak{H}'')$, where $\mathfrak{H}''$ is the corresponding subspace of $\mathfrak{H}^{P_2}$ which is isometrically isomorphic to $\mathfrak{H}'$. Continue in this fashion, consecutively, with $\mathfrak{H}^{P_3}, \cdots \mathfrak{H}^{P_n}$. It should be noticed, incidentally, that the construction does not depend on the order in which $\mathfrak{H}^{P_1}, \cdots \mathfrak{H}^{P_n}$ are taken.

To each $\mathfrak{H}_0 \in H_0$ will correspond a C^*-algebra

[27] $\mathfrak{B}^P(\mathcal{O}^f)$ denotes the C^*-subalgebra of $\mathfrak{B}(\mathfrak{H}^P)$ which is *-homomorphic to $\mathfrak{A}(\mathcal{O}^f)$ (Proposition 1, Sec. 3.3).

$\mathfrak{B}(\mathcal{O}', \mathcal{H}_0)$ determining a representation of $\mathfrak{A}(\mathcal{O}')$; e.g., in the case when $\mathcal{H}_0$ is built from $\mathcal{H}^{P_1}$ and $\mathcal{H}^{P_2}$, $\mathfrak{B}^{P_1}(\mathcal{O}')$ leaves invariant $\mathcal{H}'$ as well as $\mathcal{H}'^{P_1} \ominus \mathcal{H}'$, and $\mathfrak{B}^{P_2}(\mathcal{O}')$ leaves invariant $\mathcal{H}''$ and $\mathcal{H}^{P_2} \ominus \mathcal{H}''$. Hence, we get the element $A_x \in \mathfrak{B}(\mathcal{O}', \mathcal{H}_0)$ by taking the corresponding direct sums $A_x = A_x' \oplus A_x'' \oplus A_x'''$, where $A_x' \in \mathfrak{B}(\mathcal{H}')$, $A_x'' \in \mathfrak{B}(\mathcal{H}^{P_1} \ominus \mathcal{H}')$, $A_x''' \in \mathfrak{B}(\mathcal{H}^{P_2} \ominus \mathcal{H}'')$.

We introduce a partial ordering in H_0 by writing $\mathcal{H}_1 \leq \mathcal{H}_2 (\mathcal{H}_1, \mathcal{H}_2 \in H_0)$ if and only if $\mathcal{H}_1$ is isometrically isomorphic to a subspace of $\mathcal{H}_2$, and if, for any $\mathcal{H}^{P_k}$ which entered in the building of $\mathcal{H}_1$, there is such an $\mathcal{H}^{P_l}$ which entered in the building of $\mathcal{H}_2$, then $\mathcal{H}^{P_k} \leq \mathcal{H}^{P_l}$ (in the partial ordering of H).

Denote by H_1 the set of all Hilbert spaces which are either elements of H_0 or which are the limits of monotone increasing sequences of elements of H_0. We will introduce a partial ordering in H_1 defined in the following way: $\mathcal{H}_1 \leq \mathcal{H}_2 (\mathcal{H}_1, \mathcal{H}_2 \in H_1)$ if and only if, for any given $\mathcal{H}_1^{(n)} \in H_0$ in the sequence $\mathcal{H}_1^{(1)}, \mathcal{H}_1^{(2)}, \cdots$ determining $\mathcal{H}_1$, there exists an $\mathcal{H}_2^{(m)} \in H_0$ in the sequence $\mathcal{H}_2^{(1)}, \mathcal{H}_2^{(2)}, \cdots$ determining $\mathcal{H}_2$, which is such that $\mathcal{H}_1^{(n)} \leq \mathcal{H}_2^{(m)}$.

It is easy to see that H_1 satisfies the conditions of the Zorn's lemma. Namely, take a linearly ordered sequence $\mathcal{H}_1 \leq \mathcal{H}_2 \leq \cdots$ of elements of H_1 which is such that $\mathcal{H}_n \neq \mathcal{H}_{n+1}$ $(n = 1, 2, 3, \cdots)$. (This condition obviously does not affect the generality of the argument). Denote by $\mathcal{H}_0$ the Hilbert space which is the limit of this sequence, i.e.,

$$\mathcal{H}_0 = \mathcal{H}_1 \oplus (\mathcal{H}_2 \ominus \mathcal{H}_1) \oplus (\mathcal{H}_3 \ominus \mathcal{H}_2) \oplus \cdots . \quad (3.25)$$

To show that $\mathcal{H}_0 \in H_1$, choose from the sequence determining $\mathcal{H}_n$ an $\mathcal{H}_n^{(k_n)} \in H_0$ such that $\mathcal{H}_{n-1}^{(k_{n-1})} \leq \mathcal{H}_n^{(k_n)}$. Such an $\mathcal{H}_n^{(k_n)}$ obviously exists if

$$\mathcal{H}_{n-1} \leq \mathcal{H}_n (\mathcal{H}_{n-1}, \mathcal{H}_n \in H_1).$$

We evidently have $\mathcal{H}_0$ as the limit of the sequence $\mathcal{H}_1^{(k_1)} \leq \mathcal{H}_2^{(k_2)} \leq \cdots$ of elements of H_0.

Denote by $\mathcal{H}$ a maximal element (in the introduced partial ordering) of H_1, and denote by $\mathfrak{B}(\mathcal{O}')$ the corresponding C^*-subalgebra of $\mathfrak{B}(\mathcal{H})$, which is completely isomorphic to $\mathfrak{A}(\mathcal{O}')$.[28] We will show that $\mathcal{H}$ is a representation Hilbert space for the language $\mathfrak{L}$.

[28] It will be clear from the immediately following argument that we are dealing with an isomorphism. The only additional detail that has to be remembered is that if $x \neq \mathbb{0}$ then $\| x \| > 0$, i. e., there exists a $P \in \mathcal{S}_0$ for which $\langle x^* x \rangle_P > 0$.

$\mathcal{H}$ can be written as a limit of $\mathcal{H}_1, \mathcal{H}_2, \cdots$, where $\mathcal{H}_n \in H_0 (n = 1, 2, \cdots)$. If, for some $P \in \mathcal{S}_0$, $\mathcal{H}^P \leq \mathcal{H}_n$ (in the partial ordering of H_0), then we adjoin, naturally, to P the vector $\Psi'^P \in \mathcal{H}_n$ (it has been explained earlier what is denoted by Ψ'^P). We obviously must have, for each $P \in \mathcal{S}_0$, at least one such $\mathcal{H}_n$. Namely, if this were not so in the case of some $P_0 \in \mathcal{S}_0$, then we would have $\mathcal{H} \oplus \mathcal{H}^{P_0} \in H_1$ and $\mathcal{H} \leq \mathcal{H} \oplus \mathcal{H}^{P_0}$; this would mean that $\mathcal{H}$ is not maximal because $\mathcal{H} \neq \mathcal{H} \oplus \mathcal{H}^{P_0}$.

The elements of H_0 have been so built that, to each $P \in \mathcal{S}_0$, there can correspond only one Ψ^P in an $\mathcal{H}_n \in H_0$. Therefore we are dealing with a mapping. Evidently, this mapping is injective, because, if $\Psi^{P_1} = \Psi^{P_2}$, then $\langle x \rangle_{P_1} = \langle x \rangle_{P_2}$ for all $x \in \mathfrak{A}(\mathcal{O}')$, and therefore $P_1 = P_2$.

Finally, the linear manifold spanned by all Ψ^P, $P \in \mathcal{S}_0$, has to be everywhere dense in $\mathcal{H}$. If that were not so, we could denote by $\mathcal{H}'$ the closure of this manifold. We have, because of Hypothesis 2A in Sec. 2.3, that, in case $\| A_x \Psi^P \| > 0$, $A_x \in \mathfrak{B}^P(\mathcal{O}')$, $x \in \mathfrak{A}_3(\mathcal{O}')$, then $\| A_x \Psi^P \|^{-1} A_x \Psi^P$ corresponds to a $P \in \mathcal{S}$. The linear manifold spanned by $\{A_x \Psi^P : x \in \mathfrak{A}_3(\mathcal{O}')\}$ is everywhere dense in $\mathcal{H}^P$ because Ψ^P is cyclic and because of the way $\mathfrak{A}_4(\mathcal{O}')$ and $\mathfrak{A}(\mathcal{O}')$ are constructed out of $\mathfrak{A}_3(\mathcal{O}')$. This would imply that $\mathcal{H}'$ contains all $\mathcal{H}_n$, $n = 1, 2, \cdots$, which is impossible if $\mathcal{H}' \neq \mathcal{H}$. Q.E.D.

We are also able to say something about the relation of the pure physical states, represented by vectors in $\mathcal{H}$, to mixtures, if we reformulate Proposition 1, Sec. 40.3 in Ref. 12 to suit our needs:

Proposition 2: The set Ψ' of all positive linear functionals $\langle x \rangle_P$ on $\mathfrak{A}(\mathcal{O}')$ corresponding to physical states $P \in \mathcal{S}$, is identical with the set of all functionals

$$F(x) = \int_{f \in \mathfrak{X}_0'} f(x)\, d\rho_F(f) = \int_{P \in \mathcal{S}_0} \langle x \rangle_P \, d\rho_F'(P)$$

corresponding to all nonnegative measures on the weakly closed and bounded in norm subset $\mathcal{K}_0'$ of $\mathfrak{A}'(\mathcal{O}')$, which are such that $\rho_F(\mathcal{K}_0') = \rho_F'(\mathcal{S}_0) = 1$.

ACKNOWLEDGMENTS

The author would like to thank Professor V. Bargmann for very interesting discussions and helpful comments.

The Dirac Bra and Ket Formalism*

J. E. ROBERTS
Emmanuel College, Cambridge, England
(Received 5 August 1965)

The place of the Dirac formalism in quantum theory is investigated using rigged Hilbert spaces. Emphasis is laid on the representation of observables by continuous linear operators and on the existence of sufficient eigenkets. Using the concept of labeled observables, a canonical procedure is given for constructing a rigged Hilbert space and the bra and ket spaces are constructed for nonrelativistic quantum systems of n interacting particles. Spectral theory is investigated in this framework and the results are compared with the Dirac formalism.

1. INTRODUCTION

IN the preface to his book on the foundations of quantum mechanics, von Neumann[1] says of Dirac's own formulation of quantum theory,[2] that it is "scarcely to be surpassed in brevity and elegance," but that it "in no way satisfies the requirements of mathematical rigour." These requirements were admirably satisfied by von Neumann's own work, for he showed that the techniques of Hilbert space provide an adequate mathematical framework for quantum theory. Since this work was published, little has changed to affect the validity of these remarks. The improper functions of Dirac, the δ function and its derivatives have stimulated the growth of a new branch of mathematics, the theory of distributions thus winning for themselves a firm place in the realm of rigorous mathematics. Yet the Dirac formalism remains far from rigorous, and the formulation in terms of Hilbert space is still the only adequate framework for quantum theory. The very elegance and success of the Dirac formalism have ensured its survival. Most of the current generation of books on quantum theory prefer to take it as their guide, rather than give more than a passing reference to the niceties of Hilbert space. The most unsatisfactory feature of the present situation is that the gulf between the Dirac formalism and Hilbert space is quite substantial, so that a lot of rethinking is necessary before grasping the "correct" way of expressing things in Hilbert space.

Under these circumstances it is surprising how little work has been done on the Dirac formalism itself. This paper reports on the possibilities of realizing this formalism within the present framework of quantum theory. As we are dealing with a formalism rather than a precise mathematical theory, we must not interpret too rigidly what constitutes the Dirac formalism. However, two basic features are taken as characteristic of the Dirac formalism as opposed to the Hilbert space theory. These are that algebraic operations on the basic physical observables should be permitted, without having to worry about the domains of definition as in Hilbert space, and that sufficient eigenvectors should exist so that even the elements of the continuous spectrum of the observables should have corresponding eigenvectors. From a purely formal point of view, it is here that the Dirac formalism enjoys its greatest advantage. We have only to compare the simplicity of the Heisenberg commutation relations between position and momentum with the corresponding Weyl form, which is preferred by rigorous theory as it involves only bounded operators. Formal eigenvectors corresponding to elements of the continuous spectrum can often be constructed explicitly—a plane wave for the momentum operator or a δ function for the position operator—and their use may then be justified in the particular context, but to find a mathematical model for the Dirac formalism we must be able to introduce such operators for a sufficiently large class of observables and extend Hilbert space to a larger space containing all such eigenvectors.

2. LABELED HILBERT SPACES

The method of approach adopted in this paper is a cautious one, which is suited to investigations of such a successful theory as quantum mechanics. We simply take the Hilbert space formulation of quantum theory as exemplified by its application

* The research reported in this document has been sponsored in part by the Air Force Office of Scientific Research under Grant AF EOAR 63-79, through the European Office of Aerospace Research (OAR), U. S. Air Force.

[1] J. Von Neumann, *Mathematical Foundations of Quantum Mechanics* (Princeton University Press, Princeton, New Jersey, 1955).

[2] P. A. M. Dirac, *The Principles of Quantum Mechanics* (Oxford University Press, New York, 1958), 4th ed.

to a specific class of physical systems and look for a realization of the Dirac formalism within the common structure of such systems. We are thus not so much engaged in modifying quantum theory as in interpreting the place of the Dirac formalism in quantum theory. Now to preserve a degree of abstraction and to avoid having to deal with a specified class of quantum systems from the start, it is necessary to make explicit a tacitly assumed feature of the mathematical structure of quantum theory.

Suppose we consider physical systems without superselection rules; under these circumstances it is often held that the (pure) states are in $1-1$ correspondence with the set of all one dimensional subspaces of an infinite-dimensional separable Hilbert space, and that the observables are in $1-1$ correspondence with the set of all self-adjoint operators on Hilbert space. While the validity or at least the physical significance of such an identification is questionable the point here is that, if this view is accepted, then the mathematical structures assigned to these quantum systems are mutually isomorphic. Yet we can clearly distinguish between the different physical systems and their realizations in terms of Hilbert space, so there is some feature of the mathematical structure, which we have not yet made explicit. This is related to a remark by Dirac[3] to the effect that we need a deeper classification of representations of the Lorentz group than that to with a unitary equivalence, because some unitary transformations, for example the S matrix, have physical significance and cannot thus be regarded as physical equivalences.

Now it seems that for quantum theory an adequate way of making this additional structure explicit is to suppose that certain of the observables have physical labels such as "energy," "momentum," or "spin." (Compare Segal[4].) To deal with this we introduce the concept of a labeled Hilbert space.

Definition. A *labeled Hilbert space* consists of an Hilbert space $\mathcal{H}$, an index set I, and a mapping from I into the set of all operators on $\mathcal{H}$. Two labeled Hilbert spaces $(\mathcal{H}_1, \{A_i\}_{i\in I_1})$ and $(\mathcal{H}_2, \{B_i\}_{i\in I_2})$ are said to be isomorphic if $I_1 = I_2$ and if there exists a unitary transformation $U\colon \mathcal{H}_1 \to \mathcal{H}_2$ such that

$$UA_iU^{-1} = B_i, \quad \text{for all} \quad i \in I_1 = I_2.$$

We now suppose that each quantum system has an associated labeled Hilbert space $(\mathcal{H}, \{A_i\}_{i\in I})$ and that each A_i is a self-adjoint operator corresponding to an observable. The set of those observables corresponding to $\{A_i\}_{i\in I}$ will be called the *labeled observables* and be denoted by $\mathcal{O}_I$. A fully developed theory would associate an equivalence class of labeled Hilbert spaces with a given quantum system to allow for ambiguities in the choice and labeling of the labeled observables of the system. However, this is best done after specification of the set of quantum systems to be considered. In this paper we are only dealing with quantum systems at a given time, so we do not need to specify whether we are using the Heisenberg or Schrödinger pictures. However, we are assuming that there exist sufficient observables defined at an instant of time for the description of a quantum system at an instant in time to have a meaning. In other words, we are taking the existence of the Schrödinger picture for granted.

The labeled observables must be sufficient to describe the quantum system and in particular all the observables of the system $\mathcal{O}$. We simply suppose that $\mathcal{O}' = \mathcal{O}_I'$, where $\mathcal{O}'$ denotes the set of all bounded operators commuting with $\mathcal{O}$. Setting up a representation[2] can now be understood as giving a description of the Hilbert space in terms of certain of the labeled observables. If such a description is to have any meaning, it must be invariant under isomorphisms of labeled Hilbert spaces. As an example consider a particle of mass m, with no internal structure, which moves under the action of a potential $V(x_1, x_2, x_3)$. We can take the labeled observables as the components of position Q_1, Q_2, Q_3, and of momentum P_1, P_2, P_3 of the particle, and the total energy of the system, given by the Hamiltonian H. Now the configuration space representation is that representation which diagonalizes the complete set of commuting observables $\{Q_1, Q_2, Q_3\}$. In this representation the Hilbert space is described as $L^2(\mathbf{R}^3)$, the space of square-integrable functions over $\mathbf{R}^3$, and the states by normalized wavefunctions $\psi(x_1, x_2, x_3)$. This representation is of course only defined up to a phase factor depending on x_1, x_2 and x_3. If we further insist that P_1, P_2 and P_3 should be represented by $-i\hbar(\partial/\partial x_1)$, $-i\hbar(\partial/\partial x_2)$, and $-i\hbar(\partial/\partial x_3)$, respectively, then the representation is defined up to a constant phase factor, so that the representation of states is uniquely defined. This is the Schrödinger representation, which we now see has been completely described in terms of the labeled observables. All physically meaningful concepts in a quantum system must be described in terms of

[3] P. A. M. Dirac, Rev. Mod. Phys. **34**, 592 (1962).
[4] I. E. Segal, *Mathematical Problems of Relativistic Physics* (American Mathematical Society, 1963).

the labeled observables, and must thus be invariant under labeled Hilbert space isomorphisms.

Setting up a representation must be distinguished from canonical quantization.[5] From the point of view of labeled Hilbert spaces, canonical quantization may be described as an invariant procedure for specifying a labeled Hilbert space to within an equivalence. Thus considering the free particle of mass m, with no internal structure, we could describe its canonical quantization by saying that $\{P_1, P_2, P_3; Q_1, Q_2, Q_3\}$ form an irreducible set of self-adjoint operators, satisfying the Weyl form of the commutation relations, and that

$$H = (1/2m)(P_1^2 + P_2^2 + P_3^2).$$

In axiomatic field theory we could take the field operators, together with the projection on the (unique) vacuum state to be the labeled operators for an Hermitian scalar field. Prescribing a set of Wightman functions, satisfying the usual conditions, is then by the reconstruction theorem equivalent to the canonical quantization of the system.[6] It is now within a presupposed structure of a labeled Hilbert space that we look for a natural method of introducing the Dirac formalism.

3. RIGGED HILBERT SPACES

At this stage it is convenient to introduce a notation which can be used either to describe the Dirac formalism, or Hilbert space, as well as to describe the structure, we finally adopt. Let Φ be a complex vector space; we denote by $\bar{\Phi}$ the complex conjugate space, that is the space whose elements are in 1–1 correspondence with those of Φ, $\phi \to \bar{\phi}$, but where the vector space operations are defined by the following equation:

$$\lambda\bar{\phi} + \mu\bar{\psi} = \overline{\bar{\lambda}\phi + \bar{\mu}\psi}.$$

Φ^* denotes the algebraic dual of Φ, that is the space of linear functionals on Φ.

We suppose:

(1) Φ is a complex vector space,

(2) Ψ is a vector subspace of $\bar{\Phi}^*$,

(3) χ is a 1–1 linear mapping $\chi: \Phi \to \Psi$ such that $(\phi_1, \phi_2) = \langle \phi_1, \chi\phi_2 \rangle$ defines a positive-definite scalar product on Φ.

We remark that $\langle \Phi, \Psi \rangle$ is a dual system,[7] and $\dim \Phi \leq \dim \Psi \leq \dim \bar{\Phi}^*$. To obtain the Dirac formalism from this, we call Ψ ket space and $\bar{\Phi}$ bra space. Our choice of $\bar{\Phi}$ rather than Φ has no deeper significance than that Dirac happened to define the bras as linear rather than antilinear functionals on the kets. If we take Dirac literally at this point, we would be forced to add that $\bar{\Phi} \cong \Psi^*$ and that χ is onto. This makes Φ isomorphic to its algebraic dual so that Φ and Ψ become finite-dimensional vector spaces. If Hilbert space is taken to represent the Dirac formalism, then χ is still taken to be onto and Φ is taken to be an (infinite-dimensional, separable) Hilbert space under the scalar product introduced above. In this case Ψ is naturally isomorphic to the space of continuous antilinear functionals on Φ, $\Phi^\times$.

We suppose a given labeled Hilbert space $\{\mathfrak{H}, \mathcal{O}\}$ and we want to give a meaning to algebraic operations on $\mathcal{O}$ and algebraic relations between elements of $\mathcal{O}$. Such operations have an unambiguous definition on any domain D contained in the domain of definition of all labeled observables and invariant under these observables.

$$AD \subset D, \quad \text{for all} \quad A \in \mathcal{O}.$$

The restriction of $\mathcal{O}$ to act only on D will serve no useful purpose in determining the properties of $\mathcal{O}$ unless D contains sufficiently many vectors. We thus impose the condition that the labeled observables have a common dense invariant domain. As we want a domain determined uniquely by the labeled Hilbert space $\mathfrak{H}$, we take D to be the maximal invariant domain for $\mathcal{O}$. D is a vector space and we denote the restriction of $\mathcal{O}$ to D by $\mathcal{O}^\circ$, and let the algebra with unit generated by $\mathcal{O}^\circ$ be $\mathfrak{A}^\circ$. If we take D as our vector space Φ, and work with $\mathcal{O}^\circ$ instead of the labeled observables, then the algebraic pathologies have been eliminated. Now to enable us to employ analytic methods as well as algebraic ones, it is necessary to give Φ a topology making it a topological vector space and making $\mathcal{O}^\circ$ a set of continuous linear operators on Φ. There is a natural way of doing this and we have the following results.

Theorem 1:[8] If $\mathcal{O}$ has a common dense invariant domain then a topology on Φ is uniquely specified by the following conditions:

(1) The topology of Φ should be finer than that induced by the Hilbert space $\mathfrak{H}$ (a finer topology is one with more open sets).

[5] G. W. Mackey, *The Mathematical Foundations of Quantum Mechanics* (W. A. Benjamin Company Inc., New York, 1963).

[6] R. F. Streater and A. S. Wightman, *PCT, Spin and Statistics, and All That* (W. A. Benjamin Company Inc., New York, 1964).

[7] G. Köthe, *Topologische lineare Räume* (Springer-Verlag, Berlin, 1960). N. Bourbaki, *Espaces vectoriels topologiques* (Hermann & Cie, Paris, 1964).

[8] J. E. Roberts, thesis, Cambridge University.

(2) Φ should have the coarsest such topology making each operator A° in $\mathcal{O}^\circ$ a continuous mapping

$$A^\circ: \quad \Phi \to \Phi$$

Further, this topology has the following properties:

(a) It is locally convex being given by the set of seminorms[7]

$$\phi \to ||A^\circ \phi||, \qquad (A^\circ \in \mathcal{A}^\circ).$$

(b) If $\mathcal{O}$ is countable then Φ is a reflexive Fréchet space.[7]

(c) In any case Φ is complete and semi-reflexive.[7]

The proof of this result is too technical to be given here and we content ourselves with the remark that the topology in question is the projective topology (*Kerntopologie*)[7] with respect to the mappings

$$A^\circ: \quad \Phi \to \mathcal{H} \qquad (A^\circ \in \mathcal{A}^\circ).$$

Let T denote the embedding mapping $T: \Phi \to \mathcal{H}$ that identifies Φ with a dense subspace of $\mathcal{H}$. If we denote the adjoint of T by $T^\times$ and identify $\mathcal{H}$ with $\mathcal{H}^\times$ under the natural isomorphism, then $T^\times : \mathcal{H} \to \Phi^\times$. $T^\times$ is a 1–1 mapping of $\mathcal{H}$ onto a dense subspace of $\Phi^\times$, if we give this latter space the strong topology[7] with respect to the dual system $\langle \Phi, \Phi^\times \rangle$. Φ and $\mathcal{H}$ may thus both be identified with dense subspaces of $\Phi^\times$, and we have derived a canonical triplet of spaces $\Phi \subset \mathcal{H} \subset \Phi^\times$ from the labeled Hilbert space $\{\mathcal{H}, \mathcal{O}\}$. From now on we identify $\Phi^\times$ with the ketspace previously denoted by Ψ. Hence $\chi = T^\times T : \Phi \to \Phi^\times$ is no longer required to be a mapping onto the whole of $\Phi^\times$, so that the bras are only in 1–1 correspondence with a subset of the kets.

We have now solved the problem of the meaning of algebraic relations between labeled observables in a satisfactory way because if $A \in \mathcal{O}$, then $A^\circ : \Phi \to \Phi$. Hence the adjoint of A°, $A^\times$ is a continuous linear mapping $A^\times : \Phi^\times \to \Phi^\times$. Since A was self-adjoint on $\mathcal{H}$, this means that $A^\circ \subset A \subset A^\times$, reflecting the inclusions $\Phi \subset \mathcal{H} \subset \Phi^\times$. Any algebraic relations between the labeled observables which we originally interpreted as being valid on Φ, may equally well be interpreted as being valid on the much larger space $\Phi^\times$.

Turning to the other main feature of the Dirac formalism (the existence of sufficient eigenvectors), we must now distinguish between eigenkets and eigenbras. There will be many more eigenkets than eigenbras, and while it is clear that we cannot hope to find enough eigenbras since such vectors must automatically lie in $\mathcal{H}$, we might hope to find sufficient eigenkets. If Φ is what is known as a nuclear space,[9] then a number of accounts have been given in the mathematical literature showing that there are sufficient eigenkets to describe the spectral resolution of a continuous linear operator on Φ with a self-adjoint extension to $\mathcal{H}$.[10] Indeed, Foiaş[11] has already proposed using such a scheme to account for the Dirac formalism, although his paper contains no attempt to identify the space for any given class of quantum systems. To ensure the existence of sufficient eigenkets, it is imperative to know when the space Φ we have constructed is a nuclear space. The following result provides an answer.

Theorem 2:[8] Let $\Phi \subset \mathcal{H} \subset \Phi^\times$ be a canonical triplet constructed from the labeled Hilbert space $\{\mathcal{H}, \mathcal{O}\}$, then Φ is nuclear if and only if there exists an operator $A^\circ \in \mathcal{A}^\circ$ with a self-adjoint extension A to $\mathcal{H}$ whose inverse A^{-1} is a nuclear mapping, that is a mapping of trace class.[12] As A^{-1} is to be of trace class, A must have a purely discrete spectrum of finite multiplicity. Further, if the eigenvalues of A are λ_i, $i = 1, 2, \cdots$, allowing for multiplicity, then $\sum_i (1 + |\lambda_i|)^{-1} < \infty$. If Φ is nuclear then we call $\Phi \subset \mathcal{H} \subset \Phi^\times$ a *rigged Hilbert space.*[13]

4. APPLICATIONS TO NONRELATIVISTIC QUANTUM SYSTEMS

We consider the nonrelativistic quantum system of n interacting distinguishable particles and allow each particle to have a finite number of internal states, m_i, $i = 1, 2, \cdots n$. We assume further that the labeled observables include Q_{ij} and P_{ij} $i = 1, 2, \cdots n$ and $j = 1, 2, 3$, the position and momentum observables of the particles. We can now take our Hilbert space to be

$$\mathcal{H} = \bigotimes_{i=1}^{n} \left(\bigoplus_{k=1}^{m_i} \mathcal{H}_{ik} \right),$$

where each $\mathcal{H}_{ik}$ is a copy of $L^2(\mathbf{R}^3)$ the space of square-integrable functions over $\mathbf{R}^3$ with respect to Lebesgue measure. We use the Schrödinger representation so that Q_{ij} is represented by multiplication by x_{ij} and P_{ij} by the operator $-i\hbar(\partial/\partial x_{ij})$.

[9] A. Grothendieck, Mem. Am. Math. Soc. No. 16 (1955).
[10] I. M. Gel'fand and N. Ya. Vilenkin, *Generalized Functions 4* (Moscow 1961).
[11] C. Foiaş, Compt. Rend. **248**, 1105 (1959).
[12] R. Schatten, "Norm Ideals of Completely Continuous Operaters," Ergeb. Math 27 (1960).
[13] It is not assumed that Φ is necessarily a countably Hilbert space in the sense of I. M. Gel'fand and N. Ya. Vilenkin.[10]

These operators have their usual domains as self-adjoint operators on $\mathfrak{H}$. Let

$$A = \sum_{i=1}^{n} \sum_{j=1}^{3} (P_{ij}^2 + Q_{ij}^2),$$

then A is essentially a harmonic oscillator Hamiltonian and has discrete eigenvalues of finite multiplicity. Further A^{-2} must be of trace class because $\sum \frac{1}{2}^n < \infty$. The restriction of A^2 to Φ must lie in the algebra $\mathfrak{A}^\circ$ so that we may deduce from Theorem 2 that the labeled Hilbert space of such a quantum mechanical system will always generate a rigged Hilbert space provided that only the labeled observables have a common invariant dense domain.

If we assume for the moment that $\Phi \subset \mathfrak{H} \subset \Phi^\times$ is in fact determined just by the position and momentum operators, then we can actually identify Φ. For convenience, we write P_i, $i = 1, 2, \cdots, 3n$ instead of P_{ij} and suppress the internal degrees of freedom of the particles. The maximal invariant domain D will consist of infinitely often differentiable functions, i.e., C^∞ functions and of just those functions for which $x^k\phi^{(m)}(x) \in L^2(\mathbf{R}^{3n})$ for all $k = (k_1, \cdots k_{3n})$, $m = (m_1, \cdots, m_{3n})$. Here we have written x^k for $x_1^{k_1}x_2^{k_2} \cdots x_{3n}^{k_{3n}}$ and $\phi^{(m)}$ for

$$\left(\frac{\partial}{\partial x_1}\right)^{m_1}\left(\frac{\partial}{\partial x_2}\right)^{m_2} \cdots \left(\frac{\partial}{\partial x_{3n}}\right)^{m_{3n}}\phi.$$

Φ may thus be identified with the space of C^∞ functions, for which $(\phi, \phi)_{k,m} = \int x^{2k}\overline{\phi^{(m)}(x)}\,\phi^{(m)}(x)\,dx < \infty$ for all k, m, and with the topology given by the norms $\phi \to [(\phi, \phi)_{k,m}]^{\frac{1}{2}}$. Now the Schwartz space of test functions[14] may be characterized as the space of C^∞ functions for which

$$(\phi, \phi)_p = \int (1 + |x|)^p \sum_{|q| \le p} \overline{\phi^{(q)}(x)}\phi^{(q)}(x)\,dx < \infty$$

for $p = 0, 1, 2, \cdots$, and its topology is given by the norms $\phi \to [(\phi, \phi)_p]^{\frac{1}{2}}$, where $|x|^2 = \sum_{i=1}^{3n} |x_i|^2$ and where $|q| = q_1 + q_2 \cdots + q_{3n}$.[10] The isomorphism of the spaces Φ and $\mathfrak{S}^{3n}$ follows from the inequalities

$$1 + |x| \le 2 + x_1^2 + \cdots + x_{3n}^2,$$

$$x^{2k} \le (1 + |x|)^{2|k|}.$$

Of the other observables (which have a clear right to be called observables with a direct physical interpretation, and thus to be included in the labeled observables), only the Hamiltonian σ affects the existence of a common invariant dense domain D. The observables associated with the internal states will of themselves give no trouble, because they will be bounded and it will suffice to consider the subset of states with a given internal configuration. We again ignore these internal degrees of freedom. Specifying the Hamiltonian H means specifying a self-adjoint operator on $L^2(\mathbf{R}^{3n})$, but for our purposes it suffices to assume that, for those vectors which are both in $D(H)$ and in $\mathfrak{S}(\mathbf{R}^{3n}) \subset L^2(\mathbf{R}^{3n})$, the Hamiltonian has the form

$$H = \sum_{i=1}^{3n} \frac{P_i^2}{2m_i} + V(x),$$

where $V(x) = V(x_1, x_2, \cdots x_{3n})$ is the operator of multiplication by $V(x)$. We have already shown that the maximal invariant domain D satisfies $D \subset \mathfrak{S}(\mathbf{R}^{3n})$. D thus consists of C^∞ functions, so it is clear that, for D to be dense, $V(x)$ must in some sense be C^∞. If $V(x)$ is not C^∞ at $x = x_0$, then $f(x_0) = 0$ for all $f \in D$. This leads to the following criterion:

Theorem 3: A necessary and sufficient condition for the labeled Hilbert space of a quantum system to generate a rigged Hilbert space is that the potential energy function should be C^∞ on some open set Ω whose complement has zero Lebesgue measure.

Proof: Under the hypothesis of the theorem the space of C^∞ functions with compact support in Ω, $\mathfrak{D}(\Omega)$, is an invariant domain for the labeled observables. If the complement of Ω has zero Lebesgue measure, then $\mathfrak{D}(\Omega)$ will be dense in $L^2(\mathbf{R}^{3n})$. Conversely, let S be the set of points where $V(x)$ is not C^∞. Then if $f \in D$, $f(x) = 0$ on $\bar{S}$, the closure of S. $V(x)$ is C^∞ on the open set $\mathbf{R}^{3n} \sim \bar{S}$. If $\bar{S}$ has no zero Lebesgue measure we can find a C^∞ function with support in $\bar{S}$, which represents a nonzero vector for $\mathfrak{H}$, but which is orthogonal to D. Hence D cannot be dense in $\mathfrak{H}$.

The condition of Theorem 3 is not one that is normally imposed on the potential in quantum theory. Nevertheless, the potentials normally used, such as those derived from the Coulomb interaction, do satisfy this condition. There seems no reason to suppose, that it would lead to any contradiction with experiment to assume that this condition is always satisfied, within the range of validity of non-relativistic quantum mechanics.

For particles with internal degrees of freedom, the potential energy becomes a matrix $V_{jk}(x_1, x_2, \cdots x_{3n})$, where j, k label the internal degrees of freedom. For a rigged Hilbert space to exist it is now necessary and sufficient that each component should be C^∞

[14] L. Schwartz, *Théorie des distributions* (Hermann & Cie, Paris, 1957).

on an open set Ω, whose complement has zero Lebesgue measure. If some of the particles are indistinguishable, the potential energy is symmetric in the corresponding variables, and the wavefunctions are antisymmetric or symmetric depending on whether we have fermions or bosons. The effect on the Hilbert space $\mathcal{H} = \otimes_{i=1}^{n}(\oplus_{k=1}^{m_i} \mathcal{H}_{ik})$ is to replace certain of the tensor products $\otimes$ by the corresponding antisymmetric tensor products Ⓐ or symmetric tensor products Ⓢ. Replacing $\mathcal{D}(\mathbf{R}^{3n})$ by the correspondingly symmetrized space of test functions the analysis would go through as before leaving the conclusions unmodified.

5. SPECTRAL THEORY

Although, as we have remarked, it is $\bar{\Phi}$ which really corresponds to the space of bra vectors introduced by Dirac, it is technically more convenient to avoid introducing this space and from now on we refer to Φ as the space of bra vectors. Let us call an observable A *continuous* if its restriction to Φ, A°, is a continuous mapping $A^\circ : \Phi \to \Phi$, i.e., if $A^\circ \in \mathcal{L}(\Phi)$. By construction the labeled observables are continuous. A continuous observable has three manifestations; a self-adjoint operator A on $\mathcal{H}$, a continuous operator A° on Φ, and the adjoint of A°, $A^\times$, a continuous operator on $\Phi^\times$. We have $A^\circ \subset A \subset A^\times$.

An *eigenvector* of the observable is an eigenvector of A, an *eigenbra* of the observable is an eigenvector of A° and an *eigenket* of the observable is an eigenvector of $A^\times$. The concept, which is best suited to describe the spectral theory in a rigged Hilbert space is that of an eigenoperator.

Definition. A nonzero continuous linear mapping γ, $\gamma : \Phi \to \Phi^\times$, is said to be an eigenoperator of the observable if $\langle \phi, \gamma\phi \rangle \geq 0$ for all $\phi \in \Phi$ and if $A^\times\gamma = \gamma A^\circ = \lambda\gamma$. λ is said to be the corresponding operator eigenvalue.

To some extent the eigenoperators play the part of the eigenprojections in Hilbert space. Thus if γ is an eigenoperator and if $\gamma\phi \neq 0$, $\phi \in \Phi$, $\gamma\phi$ is an eigenket of the observable.

The operator eigenvalues of an observable must be real, although the same does not apply to the ket eigenvalues. This is a familiar situation in wave mechanics, where formally at least we have eigenfunctions of observables, where the corresponding eigenvalue is not real. For example, we could consider $e^{i\mathbf{k}\cdot\mathbf{x}}$ for the momentum operators, even when $\mathbf{k}$ is not real. On the other hand, for real ket eigenvalues, an eigenoperator is easily constructed. For let $f \in \Phi^\times$ be a corresponding eigenket, then $\gamma = \bar{f} \otimes f$ defined by $\gamma\phi = \overline{\langle \phi, f \rangle}\, f$ is an eigenoperator corresponding to the same eigenvalue.

Yet even after introducing the concept of an eigenoperator of an observable, we cannot expect all of the operator eigenvalues to belong to the Hilbert space spectrum. The following example should serve as a warning.

Example. Let $\mathcal{D}$ denote the usual Schwartz space of test functions,[14] then $\mathcal{D}$ is a nuclear space[9] and we may make it into a rigged Hilbert space by embedding it in $L^2(-\infty, \infty)$.

$$\mathcal{D} \subset L^2(-\infty, \infty) \subset \mathcal{D}^\times$$

$-\left(\frac{d}{dx}\right)^2 + x^2$ is continuous on $\mathcal{D}$ and is essentially self-adjoint on $\mathcal{D} \subset L^2(-\infty, \infty)$. Consider $e^{\frac{1}{2}x^2} \in \mathcal{D}^\times$ defined by

$$\langle \phi, e^{\frac{1}{2}x^2} \rangle = \int \overline{\phi(x)} e^{\frac{1}{2}x^2}\, dx.$$

Then $[-(d/dx)^2 + x^2]^\times e^{\frac{1}{2}x^2} = -e^{\frac{1}{2}x^2}$, so that -1 is an operator eignevalue of $-(d/dx)^2 + x^2$, but it is not in the spectrum of $-\overline{(d/dx)^2 + x^2}$ on $L^2(-\infty, \infty)$.

Definition. An integral A° eigendecomposition of Φ is a triplet $\{\gamma(z), Z, \mu\}$, where

(1) $\gamma(z) \in \mathcal{L}(\Phi, \Phi^\times)$ is an eigenoperator of A, if nonzero.

(2) μ is a positive regular Borel measure on a locally compact space Z.

(3) $z \to \langle \phi, \gamma(z)\psi \rangle$ is μ-measurable for all ϕ, $\psi \in \Phi$.

(4) $(\phi, \psi) = \int_Z \langle \phi, \gamma(z)\psi \rangle\, d\mu(z)$ for all $\phi, \psi \in \Phi$.

The eigendecomposition is said to be real if it is of the form $\{\gamma(\lambda), \mathbf{R}, \mu\}$, where $A^\times\gamma(\lambda) = \gamma(\lambda)A^\circ = \lambda\gamma(\lambda)$.

Let $\mathcal{H} = \int \mathcal{H}(\lambda)\, d\mu(\lambda)$ be the direct integral decomposition of $\mathcal{H}$ which corresponds to the spectral resolution of the continuous observable A.[10] So that if

$$\phi = \int \phi(\lambda)\, d\mu(\lambda), \qquad \phi \in \mathcal{H},$$

then

$$A\phi = \int \lambda\phi(\lambda)\, d\mu(\lambda), \qquad \phi \in D(A).$$

Theorem 4.[8,15] There is a real integral A° eigendecomposition such that

$$\langle \phi, \gamma(\lambda)\psi \rangle = (\phi(\lambda), \psi(\lambda))_\lambda$$

[15] C. Foiaş, Rev. Math. Pures Appl. 7, 241 and 571 (1962).

for all $\phi, \psi \in \Phi$, except on a set of μ-measure zero.

This theorem shows that there are sufficient eigenoperators, and hence *a priori* sufficient eigenkets of a continuous observable to describe the spectral resolution of the observable. There will be a unique integral A° eigendecomposition if and only if A° is essentially self-adjoint, and it is only in this case that A° itself is sufficient to represent the observable. Given a set of continuous observables which commute in the usual Hilbert space sense of having commuting spectral resolutions, then we may find a corresponding simultaneous integral eigendecomposition. If there are no superselection rules and we take a complete set of continuous commuting observables, so that the corresponding direct integral decomposition of the Hilbert space has multiplicity 1, then the simultaneous eigenoperators in the corresponding integral eigendecomposition of Φ are of rank 1.

Let us consider this case in more detail, supposing that $\{A_1, A_2, \cdots A_n\}$ form a complete set of continuous commuting observables, then we can find a direct integral decomposition of $\mathcal{H}$

$$\mathcal{H} = \int \mathcal{H}(\lambda_1, \cdots \lambda_n)\, d\mu(\lambda_1, \cdots, \lambda_n),$$

where A_i is represented by multiplication by λ_i and where $\mathcal{H}$ is the space of functions on $\mathbf{R}^n$, square-integrable with respect to $\mu(\lambda_1, \cdots, \lambda_n)$.[10,16] Constructing the corresponding simultaneous integral $\{A_1^\circ, A_2^\circ, \cdots, A_n^\circ\}$ eigendecomposition gives

$$(\phi, \psi) = \int \langle \phi, f(\lambda_1, \cdots, \lambda_n)\rangle \times \overline{\langle \psi, f(\lambda_1, \cdots \lambda_n)\rangle}\, d\mu\,(\lambda_1, \cdots, \lambda_n),$$

$$A_i f(\lambda_1, \cdots, \lambda_n) = \lambda_i f(\lambda_1, \cdots, \lambda_n),\ i = 1, 2 \cdots n.$$

Writing the ket corresponding to $f(\lambda_1, \cdots \lambda_n)$ as $|\lambda_1, \cdots, \lambda_n\rangle$ we have

$$(\phi, \psi) = \int \langle \phi \mid \lambda_1, \cdots, \lambda_n\rangle \times \overline{\langle \psi \mid \lambda_1, \cdots, \lambda_n\rangle}\, d\mu\,(\lambda_1, \cdots, \lambda_n)$$

or

$$(\phi, \psi) = \int \overline{\phi(\lambda_1, \cdots, \lambda_n)} \times \psi(\lambda_1, \cdots, \lambda_n)\, d\mu\,(\lambda_1, \cdots, \lambda_n).$$

We have indeed achieved just the Dirac type of representation as a function space over the simultaneous eigenvalues of the complete set of commuting observables $\{A_1, A_2, \cdots, A_n\}$. The equation

$$\phi(\lambda_1, \cdots, \lambda_n) = \overline{\langle \phi, f(\lambda_1, \cdots, \lambda_n)\rangle}$$

shows that $\bar{f}(\lambda_1, \cdots, \lambda_n)$ may be regarded as a kind of Dirac δ function. We do not have the so-called δ-function normalization, because we do not have a weight function unity.[2] The measure

$$d\mu\,(\lambda_1, \cdots, \lambda_n)$$

is not in general Lebesgue measure, and conceals a sort of generalized weight function. Although we can decompose the measure into a discrete part and a continuous part, we can only achieve a true δ-function normalization of the continuous part if it is absolutely continuous with respect to Lebesgue measure. It may well be that this is the only case which arises for physically interesting operators.

6. CONCLUSIONS

The Dirac formalism may be regarded as being valid for a wide range of quantum systems provided we make a number of modifications, the most important of which are:

(1) The bras are in 1–1 correspondence with a subset of the kets and not with all the kets. This is already implicit in Dirac's work, because he relaxes the requirement that the complete bracket expression should always be defined.

(2) The observables used in representation theory should be continuous.

(3) The term "commuting observable" is to be understood in the usual Hilbert space sense of commuting spectral resolutions.

(4) A δ-function normalization of a continuous spectrum is only possible when the corresponding spectral measure is absolutely continuous with respect to Lebesgue measure.

(5) An eigenket of an observable is only of direct physical significance if it forms part of the integral eigendecomposition, associated with the spectral resolution of the corresponding self-adjoint operator on $\mathcal{H}$.

Although we have only treated the case of a non-relativistic system of n-interacting particles, the methods used are of wider application, and the difficulties seem to lie more in posing a suitable mathematical problem than in solving it. Basically the difficulty lies in deciding whether the method of labeled observables can be usefully employed, and what these labeled observables are to be in any particular case. Thus we can treat relativistic free particle equations using Newton and Wigner

[16] J. M. Jauch and B. Misra, Helv. Phys. Acta **38**, 30 (1965).

position operators.[17] Axiomatic field theory can be made to fit reasonably well into the scheme as has essentially been shown by Borchers and Maurin.[18] Here Φ is the space of sequences of test functions (ϕ_n), ϕ_n being a test function in $4n$ dimensions. The only difference is that the scalar product on Φ is now degenerate so that we get a rigged Hilbert space

$$\Phi/N \subset \mathfrak{H} \subset (\Phi/N)^\times,$$

where $N = \{\phi \in \Phi \mid (\phi, \phi) = 0\}$. This still allows the use of eigenkets of continuous operators on Φ. All eigenkets which play a part in the integral eigendecomposition must be elements of $(\Phi/N)^\times$, which is just the set of continuous antilinear functionals, which vanish on N. The spectral conditions on the Wightman functions[6] imply that simultaneous eigenkets of energy momentum can only lie in $(\Phi/N)^\times$ if the corresponding eigenvalue lies in the closed forward light cone. Hence, as is well known, the spectral conditions on the Wightman functions imply that the energy-momentum spectrum is contained in the closed forward light cone.

Some of the methods used in this work have been previously applied in a quantum mechanical context. Thus Grossman[19] has given examples of eigenkets in quantum theory using nested Hilbert spaces, a concept closely related to our rigged Hilbert spaces. Kristensen *et al.*[20] have used the method of making operators continuous in their investigations of representations of the canonical commutation relations in locally convex spaces, both for the finite and the infinite dimensional case. Their maximal spaces are then just spaces which are obtained by using the projective topology.

The model we have derived for the Dirac formalism exhibits some of the features both satisfactory and unsatisfactory of the wave mechanics, which preceded the introduction of the Hilbert space formulation. We can find all the eigenkets we want, but also many we do not want, and the problem of what boundary conditions to use has its abstract reflection in the failure of certain operators to be essentially self-adjoint. From a mathematical point of view, it is interesting that this model uses spaces, which may be regarded as an abstract generalization of spaces of distributions, both as regards their method of construction and as regards their topological vector space properties. Where this analogy breaks down is in the absence of a distinguished representation, in which we have well-behaved support properties.

ACKNOWLEDGMENTS

The author wishes to acknowledge the help and encouragement given him by his supervisors, Dr. R. J. Eden and Professor R. Jost, during the course of this work. He is particularly grateful to Professor Jost for allowing him to enjoy the hospitality of the Institut für Theoretische Physik, ETH, Zurich, for three semesters. The author is also duly grateful for the financial support of a D.S.I.R. grant and of a recent Research Fellowship at Emmanuel College, Cambridge.

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

[18] H. J. Borchers, Nuovo Cimento 24, 214 (1962). K. Maurin, Bull. Acad. Pol. Sci. Ser. Math. 11, 115 and 121 (1963).

[19] A. Grossmann, J. Math. Phys. 5, 1025 (1964), 6, 54 (1965).

[20] P. Kristensen, L. Mejelbo, and E. T. Poulsen, Commun. Math. Phys. 1, 175 (1965), and Math. Scand. 14, 129 (1964).

Polynominal Bases and Isoscalar Factors for $SU(3)$*

R. T. SHARP AND HANS VON BAEYER
Department of Physics, McGill University, Montreal, Canada
(Received 9 November 1965)

The generators of $SU(3)$ are represented by differential operators which act on states represented by functions of six independent variables. Orthonormal bases for irreducible representations are presented in the form of polynomials. These are used to construct bases for product representations, to derive the Clebsch–Gordan series, and to find explicit expressions for certain special isoscalar factors. The general isoscalar factors are then expressed as sums over products of these special ones. Also, two types of recursion relation are derived for the general isoscalar factors; we illustrate our method of deriving these by deriving recursion relations for $SU(2)$ which have been obtained recently by other means.

I. INTRODUCTION

IT is the purpose of this paper to present a set of polynomial bases for the irreducible representations (IR's) of $SU(3)$ and to construct the Clebsch–Gordan (CG) coefficients for an arbitrary product explicitly. Our method is a generalization of the treatment of $SU(2)$ by one of us[1] and requires very little of the machinery of classical representation theory.

As noted by Moshinsky[2] there are two related methods for treating the unitary groups. One, based on the classic work of Weyl,[3] utilizes the connection with the symmetric group to construct tensors which carry the representation. This method is discussed in detail by Baird and Biedenharn.[4] Considerable simplification is gained by the technique introduced by Schwinger,[5] who noted that the vectors which are used to construct tensors can be represented by boson creation operators operating on a suitably defined vacuum state. A variation of this method utilizes the properties of tensors directly to derive many results of the classical theory. A review of this approach and clarification of its relation to the canonical method was given by Mukunda and Pandit.[6]

The other method for treating the unitary groups is a generalization of Wigner's work on the rotation group. The generators are represented by differential operators and the states by solutions to partial differential equations. For $SU(2)$ this program has been carried out in two ways: In spherical coordinates it leads to the spherical harmonics and in Cartesian coordinates to the functions derived in I. For $SU(3)$ the problem was studied in spherical coordinates by Beg and Ruegg.[7] This paper is concerned with the Cartesian case. [*Note added in proof.* A number of our results have been derived by A. Ponzano, Nuovo Cimento **41**, 142 (1966), in a paper which appeared after the present work was submitted.]

Moshinsky, in studying the relation between the two methods, was led to differential equations similar to ours. Since he was guided by the tensor calculus he found bases for $SU(3)$ which are polynomials in nine variables, the components of three vectors in three dimensions.[8] The generalization to $SU(n)$ is then straightforward in principle.[2] Our method, on the other hand, leads to differential equations in six independent variables corresponding to the states of the defining representation and its conjugate. It is thus simpler and more natural for $SU(3)$. In the language of particle physics, we can say roughly that Moshinsky builds particles from three quarks while we use a quark and an antiquark.

Several interesting and well-known results can be found very simply by our method. In particular, we derive the eigenvalues of the quadratic and cubic Casimir operators, the unitary content of the Clebsch–Gordan series, and the Biedenharn expression for the matrix elements of the generators. The isotopic content of an irreducible representation and the dimension formula are found by inspection of the basis states.

The main result, the CG coefficients, are not so simply derived, but we have carried the program to a point beyond which the manipulations are strictly

* Work supported by the National Research Council of Canada.

[1] R. T. Sharp, Am. J. Phys. **28**, 116 (1960). Equation (n) from this paper will be referred to as (I, n).

[2] M. Moshinsky, J. Math. Phys. **4**, 1128 (1963).

[3] H. Weyl, *The Classical Groups* (Princeton University Press, Princeton, New Jersey, 1946).

[4] L. C. Biedenharn, J. Math. Phys. **4**, 436 (1963); G. E. Baird and L. C. Biedenharn, *ibid.* **4**, 1449 (1963); **5**, 1723 (1964); **5**, 1730 (1964).

[5] J. Schwinger, *On Angular Momentum* reprinted in *The Quantum Theory of Angular Momentum*, L. C. Biedenharn and H. Van Dam, Eds. (Academic Press Inc., New York, 1965).

[6] N. Mukunda and L. K. Pandit, J. Math. Phys. **6**, 746 (1965).

[7] M. A. B. Beg and H. Ruegg, J. Math. Phys. **6**, 677 (1965).

[8] M. Moshinsky, Rev. Mod. Phys. **34**, 813 (1962).

arithmetic but tedious. Certain special CG coefficients are given in the form of finite sums; the general CG coefficients are then expressed as finite sums over products of these special ones. We also derive two types of recursion relations analogous to those discussed recently for $SU(2)$ by Shelepin.[9] The CG coefficients also contain stretched X coefficients for $SU(2)$. Certain techniques for evaluating the latter are believed to be new. A short review of the existing literature on the CG coefficients for $SU(3)$, with a careful discussion of the phase conventions, is contained in Ref. 10.

The plan of the work is as follows: In Sec. II the generators of $SU(3)$ are represented by differential operators and the unique heaviest state of an IR is found. The Casimir operators are found to have simple differential forms. In Sec. III the basis functions are given explicitly. Biedenharn's formulas are derived, and at the same time the basis functions are normalized. The scalar product, which is Schwinger's work follows from the commutation relations of boson operators, is justified in our framework. In Sec. IV the Clebsch–Gordan series is read off from the explicit form of the heaviest state of an arbitrary product representation. In Sec. V we derive certain special isoscalar factors. Using these, and a judicious recoupling procedure by means of X coefficients, we find in Sec. VI the isoscalar factors for the case $p = p_1 + p_2$, $q = q_1 + q_2$ which we call the stretched case. We derive a simple form for the required X coefficients at this point. In order to find more general isoscalar products we derive two types of recursion relations in Secs. VII and VIII. The method is illustrated by an independent derivation of $SU(2)$ recursion relations. Expressions for the general isoscalar factors in terms of the "special" and stretched ones are also derived in Secs. VII and VIII.

II. OPERATORS AND HEAVIEST STATES

We denote the eight generators of $SU(3)$ by $R_\pm$, $S_\pm$, $T_\pm$, T_3, and Y. They are related to the generators of de Swart[11] by

$$R_\pm = -K_\mp, \qquad S_\pm = L_\pm,$$

$$T_\pm = I_\pm, \qquad T_3 = I_3, \qquad Y = Y.$$

In de Swart's triangular eigenvalue diagram[11] with T_3 plotted horizontally and $\frac{1}{2}\sqrt{3}Y$ plotted vertically,

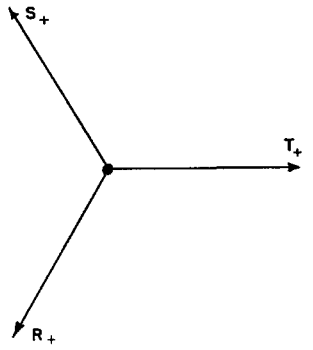


FIG. 1. The ladder operators R_+, S_+, T_+ move states in a weight diagram in the directions indicated.

the ladder operators move eigenvalues through unit distances in the directions indicated in Fig. 1.

The Lie algebra of $SU(3)$ is defined by the commutation relations of the generators. In order to write these in a symmetric fashion we define

$$R_3 \equiv -\tfrac{3}{4}Y - \tfrac{1}{2}T_3, \qquad S_3 \equiv \tfrac{3}{4}Y - \tfrac{1}{2}T_3. \tag{2.1}$$

The nonvanishing commutators are then given by (2.2) and those obtained from (2.2) by cyclic permutations of (RST).

$$\begin{aligned} [T_+, T_-] &= 2T_3, \qquad [T_3, T_\pm] = \pm T_\pm, \\ [T_3, S_\pm] &= \mp\tfrac{1}{2}S_\pm, \\ [R_+, S_+] &= T_-, \qquad [R_-, S_-] = -T_+. \end{aligned} \tag{2.2}$$

We now introduce six independent variables ξ, ξ^*, η, η^*, ζ, ζ^*. The asterisks will be an aid in distinguishing between representations and their conjugates and do not represent complex conjugation. Note that we do not impose a subsidiary condition of the form $\xi\xi^* + \eta\eta^* + \zeta\zeta^* = \text{const}$. The commutation relations are faithfully reproduced if we take

$$T_+ = -\eta\frac{\partial}{\partial\xi} + \xi^*\frac{\partial}{\partial\eta^*}, \qquad T_- = -\xi\frac{\partial}{\partial\eta} + \eta^*\frac{\partial}{\partial\xi^*},$$

$$T_3 = \frac{1}{2}\left(\eta\frac{\partial}{\partial\eta} - \xi\frac{\partial}{\partial\xi} - \eta^*\frac{\partial}{\partial\eta^*} + \xi^*\frac{\partial}{\partial\xi^*}\right). \tag{2.3}$$

$R_\pm$, R_3 are obtained from $T_\pm$, T_3 and $S_\pm$, S_3 from $R_\pm, R_3$ by the cyclic permutation $(\xi\eta\zeta)$. The operands and hence the basis vectors of the IR's are functions of the six variables. The heaviest state $|H\rangle$ of a finite dimensional representation is defined by

$$T_+|H\rangle = R_-|H\rangle = S_-|H\rangle = 0. \tag{2.4}$$

The most general solution of these equations is

$$|H\rangle = f(\xi\xi^* + \eta\eta^* + \zeta\zeta^*, \xi^*, \eta).$$

Now from the form, $R = \xi\xi^* + \eta\eta^* + \zeta\zeta^*$ is invariant with respect to the generators of $SU(3)$. In order to have unique basis vectors we must there-

[9] L. A. Shelepin, Zh. Eksperim. i Teor. Fiz. 46, 1033 (1964) [English transl.: Soviet Phys.—JETP 19, 702 (1964)].
[10] J. G. Kuriyan, D. Lurie, and A. J. Macfarlane, J. Math. Phys. 6, 722 (1965).
[11] J. J. de Swart, Rev. Mod. Phys. 35, 916 (1963).

fore discard from the space of functions all those which depend on R. (This requirement stems from the unimodularity condition.) We find then

$$|H\rangle = f(\xi^*, \eta). \tag{2.5}$$

In a finite representation, a finite number of applications of T_- annihilates $|H\rangle$; it follows that $|H\rangle$ is a polynomial in ξ^* and η. Since the generators leave invariant the degrees in unstarred and starred variables (denoted by p and q), we can label an IR by (p, q). We call a function of degree p in unstarred and q in starred variables a function of degree p, q. The heaviest state of an IR labeled by (p, q) is given uniquely by

$$\left|\begin{matrix} p, & q \\ H & \end{matrix}\right\rangle = \frac{\xi^{*q}\eta^p}{(q!\,p!)^{\frac{1}{2}}}, \tag{2.6}$$

where phase and normalization are arbitrary and p and q are nonnegative integers. The other states are obtained by applying polynomials of the generators to $|H\rangle$. In the next section we present an orthonormal basis and a scalar product for the resulting vector space. The basis functions will be labeled by p, q, n, T, and T_3, where $n \equiv Y - \frac{1}{3}(p - q)$ is an integer denoting the number of lines above (or below) the line containing $|H\rangle$ in the weight diagram.

In the notation

$$\left|\begin{matrix} p, & q \\ n, & T, \; T_3 \end{matrix}\right\rangle,$$

we find

$$|H\rangle \equiv \left|\begin{matrix} p, & q \\ H & \end{matrix}\right\rangle = \left|\begin{matrix} p, & q \\ 0, & \frac{1}{2}(p+q), \; \frac{1}{2}(p+q) \end{matrix}\right\rangle. \tag{2.7}$$

The Casimir operators can be written in terms of the A matrix as defined by de Swart.[11] In our notation it is given by

$$A = N - \bar{N} - \tfrac{1}{3}P + \tfrac{1}{3}Q, \tag{2.8}$$

where

$$N = \begin{bmatrix} \eta\frac{\partial}{\partial\eta} & -\eta\frac{\partial}{\partial\xi} & \eta\frac{\partial}{\partial\zeta} \\ -\xi\frac{\partial}{\partial\eta} & \xi\frac{\partial}{\partial\xi} & -\xi\frac{\partial}{\partial\zeta} \\ \zeta\frac{\partial}{\partial\eta} & -\zeta\frac{\partial}{\partial\xi} & \zeta\frac{\partial}{\partial\zeta} \end{bmatrix} \tag{2.9}$$

and $\bar{N}$ is obtained from N by starring all variables and transposing. The Casimir operators $F^2 = \mathrm{Tr}\, A^2/2$ and $G^3 = \mathrm{Tr}\, A^3$ can be simplified by observing that $N^2 = N(P+2)$, $\bar{N}^2 = Q + (Q-1)\bar{N}$, $\mathrm{Tr}\, N = P$, $\mathrm{Tr}\, \bar{N} = Q$ and that P and Q applied to any vector of an IR are p and q, respectively. Using these identies we find

$$F^2 = \tfrac{1}{3}[(p+q)(p+q+3) - pq]. \tag{2.10}$$

Writing $G^3(p, q)$ for G^3, we find for the part of G^3 symmetric in p, q

$$\tfrac{1}{2}[G^3(p, q) + G^3(q, p)] = 3F^2, \tag{2.11}$$

and for the antisymmetric part

$$\tfrac{1}{2}[G^3(p, q) - G^3(q, p)] = (p-q)[\tfrac{2}{9}(p+q)(p+q+\tfrac{3}{2}) + \tfrac{1}{9}pq + 1]. \tag{2.12}$$

III. THE BASIS FUNCTIONS

The basis functions are given explicitly by the formula

$$\begin{aligned}
&\left|\begin{matrix} p, & q & \\ n, & T, & T_3 \end{matrix}\right\rangle \\
&= \left[\frac{\{\frac{1}{2}(p+q-n)+T+1\}!\,\{\frac{1}{2}(p+q+n)+T+1\}!\,\{\frac{1}{2}(p+q-n)-T\}!\,\{\frac{1}{2}(p+q+n)-T\}!}{(p+q+1)!}\right]^{\frac{1}{2}} \\
&\times \sum_{n'} [\{\tfrac{1}{2}(p+q-n)+T+n'+1\}!\,\{\tfrac{1}{2}(p+q-n)-T+n'\}!\,(n-n')!\,(-n')!]^{-\frac{1}{2}} \\
&\times \sum_{T_3'} \left|\begin{matrix} p, & 0 & \\ n', & \frac{1}{2}(p+n'), & T_3' \end{matrix}\right\rangle \left|\begin{matrix} 0, & q & \\ n-n', & \frac{1}{2}(q-n+n'), & T_3 - T_3' \end{matrix}\right\rangle \left\langle\begin{matrix} \frac{1}{2}(p+n'), & \frac{1}{2}(q-n+n') \\ T_3', & T_3 - T_3' \end{matrix}\right| \left.\begin{matrix} T \\ T_3 \end{matrix}\right\rangle.
\end{aligned} \tag{3.1}$$

The simple functions

$$\left|\begin{matrix} p, & 0 & \\ n, & \frac{1}{2}(p+n), & T_3 \end{matrix}\right\rangle \text{ and } \left|\begin{matrix} 0, & q & \\ n, & \frac{1}{2}(q-n), & T_3 \end{matrix}\right\rangle,$$

to which

$$\left|\begin{matrix} p, & q & \\ n, & T, & T_3 \end{matrix}\right\rangle$$

reduce when $q = 0$ or $p = 0$, respectively, are given by

$$\left|\begin{matrix} p, & 0 & \\ n, & \frac{1}{2}(p+n), & T_3 \end{matrix}\right\rangle = \frac{\zeta^{-n}(-\xi)^{\frac{1}{2}(p+n)-T_3}\eta^{\frac{1}{2}(p+n)+T_3}}{[(-n)!\,\{\frac{1}{2}(p+n)-T_3\}!\,\{\frac{1}{2}(p+n)+T_3\}!]^{\frac{1}{2}}}, \tag{3.2}$$

$$\left|\begin{matrix}0, & q \\ n, & \frac{1}{2}(q-n), \quad T_3\end{matrix}\right\rangle$$

$$= \frac{(-\zeta^*)^n \eta^{*\frac{1}{2}(q-n)-T_3} \xi^{*\frac{1}{2}(q-n)+T_3}}{[n!\,\{\frac{1}{2}(q-n)-T_3\}!\,\{\frac{1}{2}(q-n)+T_3\}!]^{\frac{1}{2}}}. \quad (3.3)$$

The last factor on the right side of (3.1) is an $SU(2)$ CG coefficient. The parameters n, n' in (3.1) take on all integer values for which the arguments of factorials in the denominator are nonnegative. Thus according to (3.2) and (3.3) we have, respectively, $-p \leq n' \leq 0$ and $0 \leq n - n' \leq q$. It follows that

$$-p \leq n \leq q. \quad (3.4)$$

For n fixed, n' varies from a minimum of $-p$ or $(-q+n)$, whichever is greater, to a maximum of 0 or n, whichever is smaller. Thus T, as the vector sum of isospins $\frac{1}{2}(p+n')$ and $\frac{1}{2}(q-n+n')$, satisfies the inequality

$$\tfrac{1}{2}|p-q+n| \leq T \leq \tfrac{1}{2}(p+q-|n|). \quad (3.5)$$

For n and T fixed, n' thus varies from a minimum of $T - \frac{1}{2}(p+q-n)$ to a maximum of 0 or n, whichever is lesser. The number of terms in the n' sum is $\frac{1}{2}(p+q-|n|) - T + 1$; for states on the perimeter of the weight diagram we find $T=\frac{1}{2}(p+q-|n|)$ and the sum reduces to a single term.

To count states we notice by (3.5) that for fixed $n \geq 0$ there are $(p+1)(q-n+1)$ states while for fixed $n \leq 0$ there are $(p+n+1)(q+1)$; then according to (3.4) the total number of states is

$$(p+1)\sum_{n=0}^{q}(q-n+1) + (q+1)\sum_{n=-p}^{-1}(p+n+1) = \tfrac{1}{2}(p+1)(q+1)(p+q+2).$$

The n, T content of an IR is given by (3.4) and (3.5), which are equivalent to Eqs. (II.4) and (II.5) of Ref. 10. The dimension formula is given, e.g., by de Swart.[11]

In Figs. 2, 3, 4 we show the weight diagrams and basis functions for the representations (1, 0), (0, 1), and (1, 1).

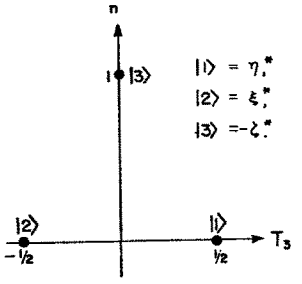


FIG. 2. The weight diagram and basis vectors for the representation (1, 0).

The proof of (3.1) is by induction. In the case of the heaviest state, for which $n = 0$, $T = \frac{1}{2}(p+q)$, it agrees with

$$\left|\begin{matrix}p, & q \\ H & \end{matrix}\right\rangle$$

as given by (2.6); if we apply $T_\pm$ to (3.1) we get a multiple of the neighboring state $|n, T, T_3 \pm 1\rangle$ (we suppress p, q in the state symbol when no confusion arises); and if we apply $R_\pm$ we get a linear combination of the neighboring states

$$|n \mp 1, T + \tfrac{1}{2}, T_3 \mp \tfrac{1}{2}\rangle, \; |n \mp 1, T - \tfrac{1}{2}, T_3 \mp \tfrac{1}{2}\rangle.$$

All the states (3.1) and no others are generated by repeated application of $T_\pm$, $R_\pm$. Also (3.1) is seen to be an eigenstate of Y, T^2, T_3 with the eigenvalues indicated. Therefore (3.1) are the required eigenstates but the correctness of their normalization has not yet been demonstrated.

We could derive the normalization factors in (3.1) by means of the cranking procedure described in the preceding paragraph, if we assumed the well-known values for the matrix elements of $T_\pm$ and Biedenharn's values[4] for those of $R_\pm$; but it is possible to derive the normalization and the matrix elements of $T_\pm$, $R_\pm$ *ab initio*. In (3.1) multiply the right-hand side by a normalizing factor C_{nTT_3} to be determined (it will turn out to be unity). Then, operating with R_- and R_+ we find

$$R_-|n, T, T_3\rangle = -\frac{C_{nTT_3}}{C_{n+1,T+\frac{1}{2},T_3+\frac{1}{2}}} a_+ |n+1, T+\tfrac{1}{2}, T_3+\tfrac{1}{2}\rangle - \frac{C_{nTT_3}}{C_{n+1,T-\frac{1}{2},T_3+\frac{1}{2}}} a_- |n+1, T-\tfrac{1}{2}, T_3+\tfrac{1}{2}\rangle, \quad (3.6)$$

$$R_+|n+1, T+\tfrac{1}{2}, T_3+\tfrac{1}{2}\rangle = -\frac{C_{n+1,T+\frac{1}{2},T_3+\frac{1}{2}}}{C_{nTT_3}} a_+ |n, T, T_3\rangle - \frac{C_{n+1,T+\frac{1}{2},T_3+\frac{1}{2}}}{C_{n,T+1,T_3}} b_- |n, T+1, T_3\rangle, \quad (3.7)$$

where

$$a_+ = \left[\frac{(T+T_3+1)\{\frac{1}{2}(p-q+n)+T+1\}\{\frac{1}{2}(p+q+n)+T+2\}\{\frac{1}{2}(p+q-n)-T\}}{2(T+1)(2T+1)}\right]^{\frac{1}{2}}, \quad (3.8a)$$

$$a_- = \left[\frac{(T - T_3)\{\frac{1}{2}(q - p - n) + T\}\{\frac{1}{2}(p + q + n) - T + 1\}\{\frac{1}{2}(p + q - n) + T + 1\}}{2T(2T + 1)}\right]^{\frac{1}{2}}. \quad (3.8b)$$

b_-, which we do not need, is obtained from a_+ by the replacement $T \to T + 1$.

Take the scalar product of (3.6) with

$$|n + 1, T + \tfrac{1}{2}, T_3 + \tfrac{1}{2}\rangle$$

and of (3.7) with $|n, T, T_3\rangle$. We can equate

$$\langle n + 1, T + \tfrac{1}{2}, T_3 + \tfrac{1}{2}\,|R_-|\, n, T, T_3\rangle$$

to

$$\langle n, T, T_3\,|R_+|\, n + 1, T + \tfrac{1}{2}, T_3 + \tfrac{1}{2}\rangle;$$

they are complex conjugates because $R_\pm$ are Hermitian conjugate operators, as demanded by their role as generators; and they are negative by Biedenharn's phase convention, which we adopt. We find immediately

$$C_{n,T,T_3} = C_{n+1,T+\frac{1}{2},T_3+\frac{1}{2}} \quad (3.9)$$

and

$$\langle n + 1, T + \tfrac{1}{2}, T_3 + \tfrac{1}{2}\,|R_-|\, n, T, T_3\rangle = -a_+. \quad (3.10)$$

Similarly we can deduce

$$C_{n,T,T_3} = C_{n+1,T-\frac{1}{2},T_3+\frac{1}{2}},\ C_{n,T,T_3} = C_{n,T,T_3+1}, \quad (3.11)$$

and

$$\langle n + 1, T - \tfrac{1}{2}, T_3 + \tfrac{1}{2}\,|R_-|\, n, T, T_3\rangle = -a_-. \quad (3.12)$$

Equations (3.10) and (3.12) are Biedenharn's values for the matrix elements of R_-. By repeated application of (3.9) and (3.11) we see that the C_{n,T,T_3} are all equal, and therefore equal to unity, the value arbitrarily chosen for the heaviest state.

We have defined a scalar product in the space spanned by the functions (3.1). For later application we wish to extend our definition of scalar product to include all polynomials of degree p, q. Specifically we want to justify the scalar product formula

$$\left\langle \frac{\xi^{*a}\xi^{*b}\eta^{c}\eta^{*d}\zeta^{e}\zeta^{*f}}{(a!\,b!\,c!\,d!\,e!\,f!)^{\frac{1}{2}}} \,\middle|\, \frac{\xi^{*a'}\xi^{*b'}\eta^{c'}\eta^{*d'}\zeta^{e'}\zeta^{*f'}}{(a'!\,b'!\,c'!\,d'!\,e'!\,f'!)^{\frac{1}{2}}} \right\rangle = \delta_{aa'}\delta_{bb'}\delta_{cc'}\delta_{dd'}\delta_{ee'}\delta_{ff'}. \quad (3.13)$$

It is not hard to show that the scalar product (3.13) gives the generators their correct Hermiticity properties, i.e., $T_\pm$, $R_\pm$, $S_\pm$, are pairs of Hermitian conjugates, T_3, Y are Hermitian, and hence the Casimir operators F^2, G^3 are Hermitian.

To span the space of the polynomials of degree p, q we augment the functions (3.1) by including

$$N_G(\xi\xi^* + \eta\eta^* + \zeta\zeta^*)^G \left|\begin{matrix} p - G, & q - G \\ n, & T, \quad T_3 \end{matrix}\right\rangle, \quad (3.14)$$

where N_G is a normalizing factor independent of n, T, T_3 and G ranges from 0 to the smaller of p, q. The polynomials (3.14) are orthogonal since for different G they belong to different representations and hence to different eigenvalues of the Casimir operators, and their number,

$$\sum_{G=0}^{p}(p - G + 1)(q - G + 1)(p + q - 2G + 2) = \tfrac{1}{4}(p + 1)(p + 2)(q + 1)(q + 2),$$

is equal to the number of independent monomials of degree p, q; hence they form a complete orthonormal set. The N_G are arbitrary so we may choose them to normalize (3.14) on the basis of the scalar product (3.13); the polynomials (3.14) and the monomials of degree p, q are then connected by a unitary transformation and the scalar product (3.13) follows from the orthonormality and completeness of the polynomials (3.14). Incidentally, the value of the normalizing factor is

$$N_G = \left[\frac{(p + q + 2 - 2G)!}{G!\,(p + q + 2 - G)!}\right]^{\frac{1}{2}},$$

as may be verified by normalizing the heaviest state of (3.14), for which $n = 0$, $T = \frac{1}{2}(p + q) - G$. The polynomials (3.14) form a basis for the group $U(3)$.

IV. THE CLEBSCH–GORDAN SERIES

The product of two representations can be reduced;

$$(p_1, q_1) \otimes (p_2, q_2) = \Sigma\,(p, q). \quad (4.1)$$

We refer to terms in the sum on the right as product representations and to the factors on the left as the factor representations. We find all the representations (p, q) by solving the equations

$$T_+\psi = 0, \qquad R_-\psi = 0, \qquad S_-\psi = 0, \quad (4.2)$$

satisfied by the heaviest state of any representation. The general solution of (4.2) is

$$\psi_a = f(\xi_1\xi_2^* + \eta_1\eta_2^* + \zeta_1\zeta_2^*, \xi_1^*\xi_2 + \eta_1^*\eta_2 + \zeta_1^*\zeta_2, \eta_1\zeta_2 - \zeta_1\eta_2, \eta_1, \eta_2, \xi_1^*, \xi_2^*) \quad (4.3a)$$

or

$$\psi_b = f(\xi_1\xi_2^* + \eta_1\eta_2^* + \zeta_1\zeta_2^*, \xi_1^*\xi_2 + \eta_1^*\eta_2 + \zeta_1^*\zeta_2, \zeta_1^*\xi_2^* - \xi_1^*\zeta_2^*, \eta_1, \eta_2, \xi_1^*, \xi_2^*). \quad (4.3b)$$

We refer to these as case (a) and case (b), respectively. The number of independent arguments of ψ was reduced from 12 to 7 by means of the three conditions (4.2) and by ignoring a possible dependence on $\xi_1\xi_1^* + \eta_1\eta_1^* + \zeta_1\zeta_1^*$, $\xi_2\xi_2^* + \eta_1\eta_2^* + \zeta_2\zeta_2^*$. (Note that ψ_a and ψ_b are not necessarily orthogonal to all functions containing

$$\xi_1\xi_1^* + \eta_1\eta_1^* + \zeta_1\zeta_1^* \quad \text{or} \quad \xi_2\xi_2^* + \eta_2\eta_2^* + \zeta_2\zeta_2^*$$

as a factor, a fact which causes complications later.)

Expand (4.3a), (4.3b) in a power series (polynomial):

$$\psi_a = (-1)^v NP(\xi_1\xi_2^* + \eta_1\eta_2^* + \zeta_1\zeta_2^*)^u \times (\xi_1^*\xi_2 + \eta_1^*\eta_2 + \zeta_1^*\zeta_2)^v(\eta_1\zeta_2 - \zeta_1\eta_2)^s\eta_1^{p_1-u-s} \times \eta_2^{p_2-v-s}\xi_1^{*q_1-v}\xi_2^{*q_2-u}, \quad (4.4a)$$

$$\psi_b = (-1)^v NP(\xi_1\xi_2^* + \eta_1\eta_2^* + \zeta_1\zeta_2^*)^u \times (\xi_1^*\xi_2 + \eta_1^*\eta_2 + \zeta_1^*\zeta_2)^v(\zeta_1^*\xi_2^* - \xi_1^*\zeta_2^*)^s \times \eta_1^{p_1-u}\eta_2^{p_2-v}\xi_1^{*q_1-v-s}\xi_2^{*q_2-u-s}. \quad (4.4b)$$

The "polynomial" can be chosen as a monomial since the conditions that it have definite degrees p_1, q_1 and p_2, q_2 and be an eigenfunction of Y, T^2, T_3 do not introduce relations between the coefficients of different terms. The exponents of η_1, η_2, ξ_1^*, ξ_2^* are fixed by the degrees p_1, q_1 and p_2, q_2. P is an operator which projects out of ψ the polynomials (3.14) of spaces 1 and 2 for which $G \neq 0$; except for the $u = v = 0$ case these polynomials will otherwise be present. N is a normalization factor.

The p, q values of the representation of which (4.4a) or (4.4b) is the heaviest state are found by determining the eigenvalues of T $(= T_3)$ and Y using (2.7). They are, respectively,

$$\begin{aligned} p &= p_1 + p_2 - u - v - 2s, \\ q &= q_1 + q_2 - u - v + s, \end{aligned} \quad (4.5a)$$

and

$$\begin{aligned} p &= p_1 + p_2 - u - v + s, \\ q &= q_1 + q_2 - u - v - 2s. \end{aligned} \quad (4.5b)$$

The possible values for u, v are

$$0 \leq u \leq \text{smaller of } p_1, q_2;$$

$$0 \leq v \leq \text{smaller of } p_2, q_1, \quad (4.6)$$

and for s

$$0 \leq s \leq \text{smaller of } p_1 - u, p_2 - v \quad (4.7a)$$

or

$$0 \leq s \leq \text{smaller of } q_1 - v, q_2 - u \quad (4.7b)$$

for the respective cases (a), (b).

By giving u, v, s all integer values consistent with these restrictions we obtain the complete Clebsch–Gordan series (4.1). If two representations have the same values of $u + v$ and s [and are both type (a) or both type (b)] they have the same values of p, q and are said to be degenerate. We distinguish these by specifying u.

In choosing the phase of ψ_a, ψ_b in [(4.4a), (4.4b)] we have followed the convention suggested by de Swart.[11] Of those states involving the heaviest state of representation 1, the one with the largest T_2 should have a positive coefficient. de Swart's conjecture that this will determine the phase in general is borne out by examination of ψ_a or ψ_b. The heaviest state of representation 1 is $\eta_1^{p_1}\xi_1^{*q_1}$ and its coefficient is the vector sum of states of representation 2 with isospins $\frac{1}{2}(p_2 - s)$ and $\frac{1}{2}q_2$ from unstarred and starred variables, respectively; the state of largest T_2 has the stretched value $\frac{1}{2}(p_2+q_2-s)$ and appears with positive coefficient.

V. SPECIAL ISOSCALAR FACTORS

The values of p, q, u are sufficient to specify a product representation. We denote by

$$\left|\begin{matrix} p, q, u \\ n, T, T_3 \end{matrix}\right\rangle$$

the general state of the product representation, properly normalized; the values of p_1, q_1, p_2, q_2 are suppressed and must be inferred from the context; n is not additive like Y but satisfies

$$n = n_1 + n_2 + \tfrac{1}{3}(p_1 + p_2 - p - q_1 - q_2 + q). \quad (5.1)$$

Since the product state has definite T, T_3 it is a superposition of vector sums of isospins T_1 and T_2 formed with $SU(2)$ Clebsch–Gordan coefficients:

$$\left|\begin{matrix} p, q, n \\ n, T, T_3 \end{matrix}\right\rangle = \sum_{\substack{n_1, n_2, T_1, T_2 \\ T_{13}}} \left|\begin{matrix} p_1, q_1 \\ n_1, T_1, T_{13} \end{matrix}\right\rangle \left|\begin{matrix} p_2, q_2 \\ n_2, T_2, T_3 - T_{13} \end{matrix}\right\rangle \left\langle\begin{matrix} T_1, T_2 \\ T_{13}, T_3 - T_{13} \end{matrix}\right| \left.\begin{matrix} T \\ T_3 \end{matrix}\right\rangle \left\langle\begin{matrix} p_1, q_1, p_2, q_2 \\ n_1, T_1, n_2, T_2 \end{matrix}\right| \left.\begin{matrix} p, q, u \\ n, T \end{matrix}\right\rangle, \quad (5.2)$$

i.e., the $SU(3)$ Clebsch-Gordan coefficient factors into an $SU(2)$ Clebsch-Grodan coefficient

$$\left\langle \begin{matrix} T_1, & T_2 \\ T_{13}, & T_3 - T_{13} \end{matrix} \middle| \begin{matrix} T \\ T_3 \end{matrix} \right\rangle$$

and an isoscalar factor

$$\left\langle \begin{matrix} p_1, q_1, p_2, q_2 \\ n_1, T_1, n_2, T_2 \end{matrix} \middle| \begin{matrix} p, q, u \\ n, T \end{matrix} \right\rangle ;$$

here n_2 is not independent but is given by (5.1). Since the $SU(2)$ Clebsch-Gordan coefficients are well known we have only to determine the isoscalar factors.

In this section we derive some simple special isoscalar factors in terms of which the general ones are expressed later.

By cranking we can find the eigenstates of the representation $(p_1 + p_2, 0)$ formed from $(p_1, 0)$ and $(p_2, 0)$. According to (3.1) the eigenstates of $(p_1, 0)$ with $T_3 = T$ are

$$\left| \begin{matrix} p_1, & 0 \\ n_1, & \frac{1}{2}(p_1 + n_1), & \frac{1}{2}(p_1 + n_1) \end{matrix} \right\rangle = \frac{\zeta_1^{-n_1} \eta_1^{p_1+n_1}}{[(-n_1)!\,(p_1 + n_1)!]^{\frac{1}{2}}} \tag{5.3}$$

[n_1 is negative for these states, and $T = \frac{1}{2}(p_1 + n_1)$]. The heaviest state of the product representation according to (4.4) is, normalized (we suppress $u = 0$)

$$\left| \begin{matrix} p_1 + p_2, 0 \\ H \end{matrix} \right\rangle = \frac{\eta_1^{p_1} \eta_2^{p_2}}{(p_1!\, p_2!)^{\frac{1}{2}}}$$

and the general state (but with $T = T_3$) is

$$\left| \begin{matrix} p_1 + p_2, & 0 \\ n, & \frac{1}{2}(p_1 + p_2 - n), & \frac{1}{2}(p_1 + p_2 - n) \end{matrix} \right\rangle = \left[\frac{(p_1 + p_2 + n)!}{(p_1 + p_2)!\,(-n)!} \right]^{\frac{1}{2}} (-R_+)^{-n} \frac{\eta_1^{p_1} \eta_2^{p_2}}{(p_1!\, p_2!)^{\frac{1}{2}}}$$

$$= \sum_{n_1} \left[\frac{(p_1 + p_2 + n)!\,(-n)!\,p_1!\,p_2!}{(p_1 + p_2)!} \right]^{\frac{1}{2}} \frac{\zeta_1^{-n_1} \eta_1^{p_1+n_1} \zeta_2^{n_1-n} \eta_2^{p_2+n-n_1}}{(-n_1)!\,(n_1 - n)!\,(p_1 + n_1)!\,(p_2 + n - n_1)!}.$$

By (I.12) and (5.3) we recognize this as

$$\sum_{n_1} \left[\frac{(p_1 + p_2 + n)!\,(-n)!\,p_1!\,p_2!}{(p_1 + p_2)!\,(-n_1)!\,(n_1 - n)!\,(p_1 + n_1)!\,(p_2 + n - n_1)!} \right]^{\frac{1}{2}}$$

$$\times \left| \begin{matrix} p_1, & 0 \\ n_1, & \frac{1}{2}(p_1 + n_1), & \frac{1}{2}(p_1 + n_1) \end{matrix} \right\rangle \left| \begin{matrix} p_2, & 0 \\ n - n_1, & \frac{1}{2}(p_2 + n - n_1), & \frac{1}{2}(p_2 + n - n_1) \end{matrix} \right\rangle,$$

and hence, comparing with (5.2), we find

$$\left\langle \begin{matrix} p_1, & 0, & p_2, & 0 \\ n_1, & \frac{1}{2}(p_1 + n_1), & n - n_1, & \frac{1}{2}(p_2 + n - n_1) \end{matrix} \middle| \begin{matrix} p_1 + p_2, & 0 \\ n, & \frac{1}{2}(p_1 + p_2 + n) \end{matrix} \right\rangle$$

$$= \left[\frac{(p_1 + p_2 + n)!\,(-n)!\,p_1!\,p_2!}{(p_1 + p_2)!\,(-n_1)!\,(n_1 - n)!\,(p_1 + n_1)!\,(p_2 + n - n_1)!} \right]^{\frac{1}{2}}. \tag{5.4}$$

Similarly we find

$$\left\langle \begin{matrix} 0, & q_1, & 0, & q_2 \\ n_1, & \frac{1}{2}(q_1 - n_1), & n - n_1, & \frac{1}{2}(q_2 - n + n_1) \end{matrix} \middle| \begin{matrix} 0, & q_1 + q_2 \\ n, & \frac{1}{2}(q_1 + q_2 - n) \end{matrix} \right\rangle$$

$$= \left[\frac{(q_1 + q_2 - n)!\,n!\,q_1!\,q_2!}{(q_1 + q_2)!\,n_1!\,(q_1 - n_1)!\,(n - n_1)!\,(q_2 - n + n_1)!} \right]^{\frac{1}{2}}. \tag{5.5}$$

Next we wish to combine $(p, 0)$ with $(0, q)$ to form (p, q); but we have already solved this problem in constructing the basis functions for (p, q) as given in Eq. (3.1). The isoscalar factors can be

read off:

$$\left\langle \begin{matrix} p, & 0, & 0, & q \\ n', & \frac{1}{2}(p+n'), & n-n', & \frac{1}{2}(q-n+n') \end{matrix} \middle| \begin{matrix} p, & q \\ n, & T \end{matrix} \right\rangle$$
$$= \left[\frac{\{\frac{1}{2}(p+q-n)+T+1\}!\,\{\frac{1}{2}(p+q+n)+T+1\}!\,\{\frac{1}{2}(p+q-n)-T\}!\,\{\frac{1}{2}(p+q+n)-T\}!}{(p+q+1)!\,\{\frac{1}{2}(p+q-n)+T+n'+1\}!\,\{\frac{1}{2}(p+q-n)-T+n'\}!\,(-n')!\,(n-n')!} \right]^{\frac{1}{2}}. \tag{5.6}$$

The three quantities $\eta_1\zeta_1 - \zeta_1\eta_2$, $\zeta_1\xi_2 - \xi_1\zeta_2$, $\xi_1\eta_2 - \eta_1\xi_2$ transform like ξ^*, η^*, ζ^* as can be seen by applying $T_\pm$, $R_\pm$ to them. So from the representations $(s, 0)_1$ and $(s, 0)_2$ we can construct a representation $(0, s)$ whose general state (but with $T = T_3$) is

$$\left| \begin{matrix} 0, & s \\ n, & \frac{1}{2}(s-n), \quad \frac{1}{2}(s-n) \end{matrix} \right\rangle = \frac{(-\xi_1\eta_2 + \eta_1\xi_2)^n (\eta_1\zeta_2 - \zeta_1\eta_2)^{s-n}}{[(s+1)!\,n!\,(s-n)!]^{\frac{1}{2}}}. \tag{5.7}$$

The normalization can be checked easily for the heaviest state. When the second factor in the numerator is expanded, (5.7) becomes

$$\left[\frac{(s-n)!}{(s+1)!} \right]^{\frac{1}{2}} \sum_{n_1} \frac{(-\xi_1\eta_2 + \eta_1\xi_2)^n (\eta_1\zeta_2)^{s-n+n_1} (-\zeta_1\eta_2)^{-n_1}}{(n!)^{\frac{1}{2}} (s-n+n_1)!\,(-n_1)!};$$

and using (I, 9) and (5.3) we get

$$\left| \begin{matrix} 0, & s \\ n, & \frac{1}{2}(s-n), \quad \frac{1}{2}(s-n) \end{matrix} \right\rangle = \frac{1}{(s-n+1)^{\frac{1}{2}}} \sum_{n_1, T_{13}} (-1)^{n-n_1} \left| \begin{matrix} s, & 0 \\ n_1, & \frac{1}{2}(s+n_1), \quad T_{13} \end{matrix} \right\rangle_1$$
$$\times \left| \begin{matrix} s, & 0 \\ n-s-n_1, & \frac{1}{2}(n-n_1), \quad \frac{1}{2}(s-n)-T_{13} \end{matrix} \right\rangle \left\langle \begin{matrix} \frac{1}{2}(s+n_1), & \frac{1}{2}(n-n_1) \\ T_{13}, & \frac{1}{2}(s-n)-T_{13} \end{matrix} \middle| \begin{matrix} \frac{1}{2}(s-n) \\ \frac{1}{2}(s-n) \end{matrix} \right\rangle, \tag{5.8}$$

from which we can read off the isoscalar factors

$$\left\langle \begin{matrix} s, & 0, & s, & 0 \\ n_1, & \frac{1}{2}(s+n_1), & n-s-n_1, & \frac{1}{2}(n-n_1) \end{matrix} \middle| \begin{matrix} 0, & s \\ n, & \frac{1}{2}(s-n) \end{matrix} \right\rangle = \frac{(-1)^{n-n_1}}{(s-n+1)^{\frac{1}{2}}}. \tag{5.9}$$

Similarly

$$\left\langle \begin{matrix} 0, & s, & 0, & s \\ n_1, & \frac{1}{2}(s-n_1), & n+s-n_1, & \frac{1}{2}(n_1-n) \end{matrix} \middle| \begin{matrix} s, & 0 \\ n, & \frac{1}{2}(s+n) \end{matrix} \right\rangle = \frac{(-1)^{n-n_1}}{(s+n+1)^{\frac{1}{2}}}. \tag{5.10}$$

It may be verified that the expression

$$\sum_{n, T, T_3} \left| \begin{matrix} u, & v \\ n, & T, \quad T_3 \end{matrix} \right\rangle_1 \left| \begin{matrix} v, & u \\ -n, & T, \quad -T_3 \end{matrix} \right\rangle_2 \frac{(-1)^{\frac{1}{2}(p-q-n)-T_3}}{[\frac{1}{2}(p+1)(q+1)(p+q+2)]^{\frac{1}{2}}} = \left| \begin{matrix} 0, & 0 \\ 0, & 0, \quad 0 \end{matrix} \right\rangle$$

is an $SU(3)$ scalar by noticing that it vanishes on applying $T_\pm$ or $R_\pm$. Also it is normalized to unity. Hence we may read off the isoscalar factor

$$\left\langle \begin{matrix} u, & v, & v, & u \\ n, & T, & -n, & T \end{matrix} \middle| \begin{matrix} 0, & 0 \\ 0, & 0 \end{matrix} \right\rangle = (-1)^{\frac{1}{2}(p-q-n)-T} \left[\frac{2T+1}{\frac{1}{2}(p+1)(q+1)(p+q+2)} \right]^{\frac{1}{2}}. \tag{5.11}$$

Equations (5.4)–(5.6) and (5.9)–(5.11) are special isoscalar factors which we use to generate the general ones.

VI. STRETCHED ISOSCALAR FACTORS

By the "stretched" case we mean the product representation with $p = p_1 + p_2$, $q = q_1 + q_2$. From (4.5) we find $u = v = s = 0$ and $n = n_1 + n_2$. According to (4.4) the heaviest state is, with proper normalization,

$$\left| \begin{matrix} p_1 + p_2, & q_1 + q_2 \\ H \end{matrix} \right\rangle = \frac{\eta_1^{p_1} \xi_1^{*q_1} \eta_2^{p_2} \xi_2^{*q_2}}{(p_1!\,q_1!\,p_2!\,q_2!)^{\frac{1}{2}}}, \tag{6.1}$$

and the general state is

$$\left|\begin{matrix} p_1+p_2, & q_1+q_2 \\ n, & T, \quad T_3 \end{matrix}\right\rangle = \sum_{\substack{T_{1s}', T_{1s}'', T_{2s}' \\ n_s', n_1', n_1''}} \left|\begin{matrix} p_1, & 0 \\ n_1', & \tfrac{1}{2}(p_1+n_1'), \quad T_{13}' \end{matrix}\right\rangle_1 \left|\begin{matrix} p_2, & 0 \\ n_3'-n_1', & \tfrac{1}{2}(p_2+n_3'-n_1'), \quad T_{33}'-T_{13}' \end{matrix}\right\rangle_2$$
$$\times \left|\begin{matrix} 0, & q_1 \\ n_1'', & \tfrac{1}{2}(q_1-n_1''), \quad T_{13}'' \end{matrix}\right\rangle_1 \left|\begin{matrix} 0, & q_2 \\ n-n_3'-n_1'', & \tfrac{1}{2}(q_2-n+n_3'+n_1''), \quad T_3-T_{33}'-T_{13}'' \end{matrix}\right\rangle_2$$
$$\times \left\langle \begin{matrix} \tfrac{1}{2}(p_1+n_1'), & \tfrac{1}{2}(p_2+n_3'-n_1') \\ T_{13}', & T_{33}'-T_{13}' \end{matrix} \middle| \begin{matrix} \tfrac{1}{2}(p_1+p_2+n_3') \\ T_{33}' \end{matrix} \right\rangle \left\langle \begin{matrix} \tfrac{1}{2}(q_1-n_1''), & \tfrac{1}{2}(q_2-n+n_3'+n_1'') \\ T_{13}'', & T_3-T_{33}'-T_{13}'' \end{matrix} \middle| \begin{matrix} \tfrac{1}{2}(q_1+q_2-n+n_3') \\ T_3-T_{33}' \end{matrix} \right\rangle$$
$$\times \left\langle \begin{matrix} \tfrac{1}{2}(p_1+p_2+n_3'), & \tfrac{1}{2}(q_1+q_2-n+n_3') \\ T_{33}', & T_3-T_{33}' \end{matrix} \middle| \begin{matrix} T \\ T_3 \end{matrix} \right\rangle$$
$$\times \left\langle \begin{matrix} p_1, & 0, & p_2, & 0 \\ n_1', & \tfrac{1}{2}(p_1+n_1'), & n_3'-n_1', & \tfrac{1}{2}(p_2+n_3'-n_1') \end{matrix} \middle| \begin{matrix} p_1+p_2, & 0 \\ n_3', & \tfrac{1}{2}(p_1+p_2+n_3') \end{matrix} \right\rangle$$
$$\times \left\langle \begin{matrix} 0, & q_1, & 0, & q_2 \\ n_1'', & \tfrac{1}{2}(q_1-n_1''), & n-n_3'-n_1'', & \tfrac{1}{2}(q_2-n+n_3'+n_1'') \end{matrix} \middle| \begin{matrix} 0, & q_1+q_2 \\ n-n_3', & \tfrac{1}{2}(q_1+q_2-n+n_3') \end{matrix} \right\rangle$$
$$\times \left\langle \begin{matrix} p_1+p_2, & 0, & 0, & q_1+q_2 \\ n_3', & \tfrac{1}{2}(p_1+p_2+n_3'), & n-n_3', & \tfrac{1}{2}(q_1+q_2-n+n_3') \end{matrix} \middle| \begin{matrix} p_1+p_2, & q_1+q_2 \\ n, & T \end{matrix} \right\rangle. \tag{6.2}$$

(6.2) was constructed by coupling the unstarred variables which transform like $(p_1, 0)$ and $(p_2, 0)$ to form a representation $(p_1 + p_2, 0)$, coupling the starred variables which transform like $(0, q_1)$ and $(0, q_2)$ to form a representation $(0, q_1 + q_2)$, and finally coupling $(p_1 + p_2, 0)$ and $(0, q_1 + q_2)$ to form the required stretched representation $(p_1 + p_2, q_1 + q_2)$. The isoscalar factors needed to effect these couplings are the special ones given by (5.4)–(5.6). The correctness of the normalization may be verified by noting that (6.2) reduces to (6.1) for $n = 0$, $T_3 = T = \frac{1}{2}(p_1 + p_2 + q_1 + q_2)$.

To identify the isoscalar factors in (6.2) we must recouple the isospins by means of the well-known X coefficient[12]; also write $n_1'' = n_1 - n_1'$, $n_3' = n_1' + n_2'$:

$$\left|\begin{matrix} p_1+p_2, & q_1+q_2 \\ n, & T, \quad T_3 \end{matrix}\right\rangle = \sum_{\substack{T_{1s}', T_{2s}', T_{1s}, \\ T_1, T_2, \\ n_1', n_2', n_1}} \left|\begin{matrix} p_1, & 0 \\ n_1', & \tfrac{1}{2}(p+n_1'), \quad T_{13}' \end{matrix}\right\rangle_1 \left|\begin{matrix} 0, & q_1 \\ n_1-n_1', & \tfrac{1}{2}(q_1-n_1+n_1'), \quad T_{13}-T_{13}' \end{matrix}\right\rangle_1$$
$$\times \left|\begin{matrix} p_2, & 0 \\ n_2', & \tfrac{1}{2}(p_2+n_2'), \quad T_{23}' \end{matrix}\right\rangle_2 \left|\begin{matrix} 0, & q_2 \\ n-n_1-n_2', & \tfrac{1}{2}(q_2-n+n_1+n_2'), \quad T_3-T_{13}-T_{23}' \end{matrix}\right\rangle_2$$
$$\times \left\langle \begin{matrix} \tfrac{1}{2}(p_1+n_1'), & \tfrac{1}{2}(q_1-n_1+n_1') \\ T_{13}', & T_{13}-T_{13}' \end{matrix} \middle| \begin{matrix} T_1 \\ T_{13} \end{matrix} \right\rangle \left\langle \begin{matrix} \tfrac{1}{2}(p+n_2'), & \tfrac{1}{2}(q_2-n+n_1+n_2') \\ T_{23}', & T_3-T_{13}-T_{23}' \end{matrix} \middle| \begin{matrix} T_2 \\ T_3-T_{13} \end{matrix} \right\rangle$$
$$\times \left\langle \begin{matrix} T_1, & T_2 \\ T_{13}, & T_3-T_{13} \end{matrix} \middle| \begin{matrix} T \\ T_3 \end{matrix} \right\rangle [(2T_1+1)(2T_2+1)(p_1+p_2+n_1'+n_2'+1)(q_1+q_2-n+n_1'+n_2'+1)]^{\frac{1}{2}}$$
$$\times X \begin{bmatrix} \tfrac{1}{2}(p_1+n_1') & \tfrac{1}{2}(q_1-n_1+n_1') & T_1 \\ \tfrac{1}{2}(p_2+n_2') & \tfrac{1}{2}(q_2-n+n_1+n_2') & T_2 \\ \tfrac{1}{2}(p_1+p_2+n_1'+n_2') & \tfrac{1}{2}(q_1+q_2-n+n_1'+n_2') & T \end{bmatrix}$$
$$\times \left\langle \begin{matrix} p_1, & 0, & p_2, & 0 \\ n_1', & \tfrac{1}{2}(p_1+n_1'), & n_2', & \tfrac{1}{2}(p_2+n_2') \end{matrix} \middle| \begin{matrix} p_1+p_2, & 0 \\ n_1'+n_2', & \tfrac{1}{2}(p_1+p_2+n_1'+n_2') \end{matrix} \right\rangle$$

[12] M. E. Rose, *Elementary Theory of Angular Momentum* (John Wiley & Sons, Inc., New York, 1957), p. 191.

$$\times \left\langle \begin{matrix} 0, & q_1, & 0, & q_2 \\ n_1 - n_1', & \tfrac{1}{2}(q_1 - n_1 + n_1'), & n - n_1 - n_2', & \tfrac{1}{2}(q_2 - n + n_1 + n_2') \end{matrix} \middle| \begin{matrix} 0, & q_1 + q_2 \\ n - n_1' - n_2', & \tfrac{1}{2}(q_1 + q_2 - n + n_1' + n_2') \end{matrix} \right\rangle$$
$$\times \left\langle \begin{matrix} p_1 + p_2, & 0, & 0, & q_1 + q_2 \\ n_1' + n_2', & \tfrac{1}{2}(p_1 + p_2 + n_1' + n_2'), & n - n_1' - n_2', & \tfrac{1}{2}(q_1 + q_2 - n + n_1' + n_2') \end{matrix} \middle| \begin{matrix} p_1 + p_2, & q_1 + q_2 \\ n, & T \end{matrix} \right\rangle. \tag{6.3}$$

Now write out the expansion (5.2) for

$$\left| \begin{matrix} p_1 + p_2, & q_1 + q_2 \\ n, & T, \quad T_3 \end{matrix} \right\rangle,$$

which defines the stretched isoscalar factor and substitute for the factor states using (5.6):

$$\left| \begin{matrix} p_1 + p_2, & q_1 + q_2 \\ n, & T, \quad T_3 \end{matrix} \right\rangle = \sum_{\substack{T_{13}, T_{13}', T_{23}' \\ T_1, T_2 \\ n_1', n_2', n_1}} \left| \begin{matrix} p_1, & 0 \\ n_1', & \tfrac{1}{2}(p_1 + n_1'), \quad T_{13}' \end{matrix} \right\rangle_1 \left| \begin{matrix} 0, & q_1 \\ n_1 - n_1', & \tfrac{1}{2}(q - n_1 + n_1'), \quad T_{13} - T_{13}' \end{matrix} \right\rangle_1$$
$$\times \left| \begin{matrix} p_2, & 0 \\ n_2', & \tfrac{1}{2}(p_2 + n_2'), \quad T_{23}' \end{matrix} \right\rangle_2 \left| \begin{matrix} 0, & q_2 \\ n - n_1 - n_2', & \tfrac{1}{2}(q_2 - n + n_1 + n_2'), \quad T_3 - T_{13} - T_{23}' \end{matrix} \right\rangle_2$$
$$\times \left\langle \begin{matrix} \tfrac{1}{2}(p_1 + n_1'), & \tfrac{1}{2}(q_1 - n_1 + n_1') \\ T_{13}', & T_{13} - T_{13}' \end{matrix} \middle| \begin{matrix} T_1 \\ T_{13} \end{matrix} \right\rangle \left\langle \begin{matrix} \tfrac{1}{2}(p_2 + n_2), & \tfrac{1}{2}(q_2 - n + n_1 + n_2') \\ T_{23}', & T_3 - T_{13} - T_{23}' \end{matrix} \middle| \begin{matrix} T_2 \\ T_3 - T_{13} \end{matrix} \right\rangle$$
$$\times \left\langle \begin{matrix} T_1, & T_2 \\ T_{13}, & T_3 - T_{13} \end{matrix} \middle| \begin{matrix} T \\ T_3 \end{matrix} \right\rangle \left\langle \begin{matrix} p_2, & 0, & 0, & q_2 \\ n_2', & \tfrac{1}{2}(p_2 + n_2'), & n - n_1 - n_2', & \tfrac{1}{2}(q_2 - n + n_1 + n_2') \end{matrix} \middle| \begin{matrix} p_2, & q_2 \\ n - n_1, & T_2 \end{matrix} \right\rangle$$
$$\times \left\langle \begin{matrix} p_1, & 0, & 0, & q_1 \\ n_1', & \tfrac{1}{2}(p_1 + n_1'), & n_1 - n_1', & \tfrac{1}{2}(q_1 - n_1 + n_1') \end{matrix} \middle| \begin{matrix} p_1, & q_1 \\ n_1, & T_1 \end{matrix} \right\rangle$$
$$\times \left\langle \begin{matrix} p_1, & q_1, & p_2, & q_2 \\ n_1, & T_1, & n_2, & T_2 \end{matrix} \middle| \begin{matrix} p_1 + p_2, & q_1 + q_2 \\ n, & T \end{matrix} \right\rangle. \tag{6.4}$$

Comparing (6.3) with (6.4) we see that

$$\left\langle \begin{matrix} p_1, & q_1, & p_2, & q_2 \\ n_1, & T_1, & n_2, & T_2 \end{matrix} \middle| \begin{matrix} p_1 + p_2, & q_1 + q_2 \\ n, & T \end{matrix} \right\rangle$$
$$= [(2T_1 + 1)(2T_2 + 1)(p_1 + p_2 + n_1' + n_2' + 1)(q_1 + q_2 - n + n_1' + n_2' + 1)]^{\frac{1}{2}}$$
$$\times \left\langle \begin{matrix} p_1, & 0, & p_2, & 0 \\ n_1', & \tfrac{1}{2}(p_1 + n_1'), & n_2', & \tfrac{1}{2}(p_2 + n_2') \end{matrix} \middle| \begin{matrix} p_1 + p_2, & 0 \\ n_1' + n_2', & \tfrac{1}{2}(p_1 + p_2 + n_1' + n_2') \end{matrix} \right\rangle$$
$$\times \left\langle \begin{matrix} 0, & q_1, & 0, & q_2 \\ n_1 - n_1', & \tfrac{1}{2}(q_1 - n_1 + n_1'), & n - n_1 - n_2', & \tfrac{1}{2}(q_2 - n + n_1 + n_2') \end{matrix} \middle| \begin{matrix} 0, & q_1 + q_2 \\ n - n_1' - n_2', & \tfrac{1}{2}(q_1 + q_2 - n + n_1' + n_2') \end{matrix} \right\rangle$$
$$\times \left\langle \begin{matrix} p_1 + p_2, & 0, & 0, & q_1 + q_2 \\ n_3', & \tfrac{1}{2}(p_1 + p_2 + n_1' + n_2'), & n_1 - n_1' - n_2', & \tfrac{1}{2}(q_1 + q_2 - n + n_1' + n_2') \end{matrix} \middle| \begin{matrix} p_1 + p_2, & q_1 + q_2 \\ n, & T \end{matrix} \right\rangle$$
$$\times X \begin{bmatrix} \tfrac{1}{2}(p_1 + n_1') & \tfrac{1}{2}(q_1 - n_1 + n_1') & T_1 \\ \tfrac{1}{2}(p_2 + n_2') & \tfrac{1}{2}(q_2 - n + n_2' + n_1) & T_2 \\ \tfrac{1}{2}(p_1 + p_2 + n_1' + n_2') & \tfrac{1}{2}(q_1 + q_2 - n + n_1' + n_2') & T \end{bmatrix}$$
$$\times \frac{1}{\left\langle \begin{matrix} p_1, & 0, & 0, & q_1 \\ n_1', & \tfrac{1}{2}(p_1 + n_1'), & n_1 - n_1', & \tfrac{1}{2}(q_1 - n_1 + n_1') \end{matrix} \middle| \begin{matrix} p_1, & q_1 \\ n_1, & T_1 \end{matrix} \right\rangle}$$
$$\times \frac{1}{\left\langle \begin{matrix} p_2, & 0, & 0, & q_2 \\ n_2', & \tfrac{1}{2}(p_2 + n_2'), & n - n_1 - n_2', & \tfrac{1}{2}(q_2 - n + n_1 + n_2') \end{matrix} \middle| \begin{matrix} p_2, & q_2 \\ n - n_1, & T_2 \end{matrix} \right\rangle}. \tag{6.5}$$

The stretched isoscalar factors do not depend on n_1', n_2'; they must cancel out of (6.5). That this is the case is shown by actual evaluation of the X coefficient; it is doubly stretched and takes a rather simple form. We have not found its evaluation in the literature, so we give it here.

The X coefficient is defined by

$$\sum_{\alpha\gamma\theta} \left|\begin{matrix} a \\ \alpha \end{matrix}\right\rangle \left|\begin{matrix} b \\ \theta-\alpha \end{matrix}\right\rangle \left\langle\begin{matrix} a, & b \\ \alpha, & \theta-\alpha \end{matrix}\right| \left|\begin{matrix} j \\ \theta \end{matrix}\right\rangle \left|\begin{matrix} c \\ \gamma \end{matrix}\right\rangle \left|\begin{matrix} d \\ e-\theta-\gamma \end{matrix}\right\rangle \left\langle\begin{matrix} c, & d \\ \gamma, & e-\theta-\gamma \end{matrix}\right| \left|\begin{matrix} k \\ e-\theta \end{matrix}\right\rangle \left\langle\begin{matrix} j, & k \\ \theta, & e-\theta \end{matrix}\right| \left|\begin{matrix} e \\ e \end{matrix}\right\rangle$$

$$= \sum_{\substack{\alpha\gamma\theta \\ f_\sigma}} \left|\begin{matrix} a \\ \alpha \end{matrix}\right\rangle \left|\begin{matrix} c \\ \gamma \end{matrix}\right\rangle \left\langle\begin{matrix} a, & c \\ \alpha, & \gamma \end{matrix}\right| \left|\begin{matrix} f \\ \alpha+\gamma \end{matrix}\right\rangle \left|\begin{matrix} b \\ \theta-\alpha \end{matrix}\right\rangle \left|\begin{matrix} d \\ e-\theta-\gamma \end{matrix}\right\rangle$$

$$\times \left\langle\begin{matrix} b, & d \\ \theta-\alpha, & e-\theta-\gamma \end{matrix}\right| \left|\begin{matrix} g \\ e-\alpha-\gamma \end{matrix}\right\rangle \left\langle\begin{matrix} f, & g \\ \alpha+\gamma, & e-\alpha-\gamma \end{matrix}\right| \left|\begin{matrix} e \\ e \end{matrix}\right\rangle$$

$$\times X\begin{bmatrix} a & b & j \\ c & d & k \\ f & g & e \end{bmatrix} [(2j+1)(2k+1)(2f+1)(2g+1)]^{\frac{1}{2}}. \tag{6.6}$$

Take the scalar product of (6.6) with

$$\left|\begin{matrix} a \\ a \end{matrix}\right\rangle \left|\begin{matrix} c \\ c \end{matrix}\right\rangle.$$

The left side is

$$\sum_{\theta} \left|\begin{matrix} b \\ \theta-a \end{matrix}\right\rangle \left\langle\begin{matrix} a, & b \\ a, & \theta-a \end{matrix}\right| \left|\begin{matrix} j \\ \theta \end{matrix}\right\rangle \left|\begin{matrix} d \\ e-\theta-c \end{matrix}\right\rangle \left\langle\begin{matrix} c, & d \\ c, & e-\theta-c \end{matrix}\right| \left|\begin{matrix} k \\ e-\theta \end{matrix}\right\rangle \left\langle\begin{matrix} j, & k \\ \theta, & e-\theta \end{matrix}\right| \left|\begin{matrix} e \\ e \end{matrix}\right\rangle.$$

On the right the only value of f that contributes is $f = a + c$ so the right side becomes

$$\sum_{\theta\sigma} \left|\begin{matrix} b \\ \theta-a \end{matrix}\right\rangle \left|\begin{matrix} d \\ e-\theta-c \end{matrix}\right\rangle \left\langle\begin{matrix} b, & d \\ \theta-a, & e-\theta-c \end{matrix}\right| \left|\begin{matrix} g \\ e-a-c \end{matrix}\right\rangle \left\langle\begin{matrix} a+c, & g \\ a+c, & e-a-c \end{matrix}\right| \left|\begin{matrix} e \\ e \end{matrix}\right\rangle$$

$$\times X\begin{bmatrix} a & b & j \\ c & d & k \\ a+c & g & e \end{bmatrix} [(2j+1)(2k+1)(2g+1)(2a+2c+1)]^{\frac{1}{2}};$$

now take the scalar product of both sides with

$$\sum_{\theta} \left|\begin{matrix} b \\ \theta-a \end{matrix}\right\rangle \left|\begin{matrix} d \\ e-\theta-c \end{matrix}\right\rangle \left\langle\begin{matrix} b, & d \\ \theta-a, & e-\theta-c \end{matrix}\right| \left|\begin{matrix} g \\ e-a-c \end{matrix}\right\rangle.$$

The left side becomes

$$\sum_{\theta} \left\langle\begin{matrix} b, & d \\ \theta-a, & e-\theta-c \end{matrix}\right| \left|\begin{matrix} g \\ e-a-c \end{matrix}\right\rangle \left\langle\begin{matrix} a, & b \\ a, & \theta-a \end{matrix}\right| \left|\begin{matrix} j \\ \theta \end{matrix}\right\rangle \left\langle\begin{matrix} c, & d \\ c, & e-\theta-c \end{matrix}\right| \left|\begin{matrix} k \\ e-\theta \end{matrix}\right\rangle \left\langle\begin{matrix} j, & k \\ \theta, & e-\theta \end{matrix}\right| \left|\begin{matrix} e \\ e \end{matrix}\right\rangle,$$

while the right side is

$$[(2a+2c+1)(2g+1)(2j+1)(2k+1)]^{\frac{1}{2}} \left\langle\begin{matrix} a+c, & g \\ a+c, & e-a-c \end{matrix}\right| \left|\begin{matrix} e \\ e \end{matrix}\right\rangle X\begin{bmatrix} a & b & j \\ c & d & k \\ a+c & g & e \end{bmatrix}.$$

So we find

$$X\begin{bmatrix} a & b & j \\ c & d & k \\ a+c & g & e \end{bmatrix} = \frac{\sum_\theta \left\langle \begin{matrix} b, & d \\ \theta-a, & e-\theta-c \end{matrix} \middle| \begin{matrix} g \\ e-a-c \end{matrix} \right\rangle \left\langle \begin{matrix} a, & b \\ a, & \theta-a \end{matrix} \middle| \begin{matrix} j \\ \theta \end{matrix} \right\rangle \left\langle \begin{matrix} c, & d \\ c, & e-\theta-c \end{matrix} \middle| \begin{matrix} k \\ e-\theta \end{matrix} \right\rangle \left\langle \begin{matrix} c, & d \\ c, & e-\theta-c \end{matrix} \middle| \begin{matrix} k \\ e-\theta \end{matrix} \right\rangle}{\left\langle \begin{matrix} a+c, & g \\ a+c, & e-a-c \end{matrix} \middle| \begin{matrix} e \\ e \end{matrix} \right\rangle [(2a+2c+1)(2j+1)(2k+1)(2g+1)]^{\frac{1}{2}}} \tag{6.7}$$

If $g = b + d$, X is doubly stretched and none of the Clebsch–Gordan coefficients involves sums. Inserting the values of the Clebsch–Gordan coefficients then gives

$$\begin{aligned} X\begin{bmatrix} a & b & j \\ c & d & k \\ a+c & b+d & e \end{bmatrix} &= \left[\frac{(b+d+e-a-c)!\,(b+d-e+a+c)!\,(2b)!\,(2d)!\,(2a)!\,(j-a+b)!\,(2c)!}{(2b+2d+1)!\,(a+b-j)!\,(j+a-b)!\,(a+b+j+1)!\,(c+d-k)!\,(k+c-d)!}\right]^{\frac{1}{2}} \\ &\times \left[\frac{(j+k-e)!\,(a+b+c+d+e+1)!\,(e+a+c-b-d)!\,(k-c+d)!}{(c+d+k+1)!\,(j+k+e+1)!\,(j-k+e)!\,(e-j+k)!\,(2a+2c+1)!}\right]^{\frac{1}{2}} \\ &\times \sum_\theta (-1)^{j-\theta} \frac{(j+\theta)!\,(k+e-\theta)!}{(b+\theta-a)!\,(d+e-\theta-c)!\,(k-e+\theta)!\,(j-\theta)!}. \end{aligned} \tag{6.8}$$

[*Note added in proof.* It has come to our attention (A. Ponzano, referred to in note added above) that (6.8) has been derived by A. Bandzaitis, A. Karosiene, and A. Jucys, Liet. Fiz. Rin. **4**, 457 (1964).]

The sum over θ is the same as that which occurs in Wigner's formula[13] for the Clebsch–Gordan coefficient; like it, it can be transformed to the more symmetric Racah form.

When the expressions for X coefficient and the special isoscalar factors are substituted in (6.5) the terms involving n_1', n_2' cancel out and there results

$$\begin{aligned} \left\langle \begin{matrix} p_1, & q_1, & p_2, & q_2 \\ n_1, & T_1, & n-n_1, & T_2 \end{matrix} \middle| \begin{matrix} p_1+p_2, & q_1+q_2 \\ n, & T \end{matrix} \right\rangle &= \left[\frac{(2T_1+1)(2T_2+1)p_2!\,p_1!\,q_1!\,q_2!\,(p_1+q_1+1)!}{(p_1+p_2)!\,(q_1+q_2)!\,(p_1+p_2+q_1+q_2+1)!}\right]^{\frac{1}{2}} \\ &\times \left[\frac{(p_2+q_2+1)!\,(T_1+T_2-T)!\,\{\frac{1}{2}(p_1+p_2+q_1+q_2-n)+T+1\}!\,\{\frac{1}{2}(p_1+p_2+q_1+q_2+n)+T+1\}!}{(T_1+T_2+T+1)!\,(T+T_1-T_2)!\,(T-T_1+T_2)!\,\{\frac{1}{2}(p_1+q_1-n_1)+T_1+1\}!\,\{\frac{1}{2}(p_1+q_1+n_1)+T_1+1\}!}\right]^{\frac{1}{2}} \\ &\times \left[\frac{\{\frac{1}{2}(p_1+q_1+p_2+q_2-n)-T\}!\,\{\frac{1}{2}(p_1+p_2+q_1+q_2+n)-T\}!\,\{\frac{1}{2}(q_1+q_2-p_1-p_2-n)+T\}!}{\{\frac{1}{2}(p_1+q_1-n_1)-T_1\}!\,\{\frac{1}{2}(p_1+q_1+n_1)-T_1\}!\,\{\frac{1}{2}(p_2+q_2-n+n_1)+T_2+1\}!\,\{\frac{1}{2}(p_2+q_2-n+n_1)-T_2\}!}\right]^{\frac{1}{2}} \\ &\times \left[\frac{\{\frac{1}{2}(q_1-p_1-n_1)+T_1\}!\,\{\frac{1}{2}(q_2-p_2-n+n_1)+T_2\}!\,\{\frac{1}{2}(p_1+p_2-q_1-q_2+n)+T\}!}{\{\frac{1}{2}(p_2+q_2+n-n_1)+T_2+1\}!\,\{\frac{1}{2}(p_2+q_2+n-n_1)-T_2\}!\,\{\frac{1}{2}(p_1-q_1+n_1)+T_1\}!\,\{\frac{1}{2}(p_2-q_2+n-n_1)+T_2\}!}\right]^{\frac{1}{2}} \\ &\times \sum_\theta (-1)^{T_1-\theta} \frac{(T_1+\theta)!\,(T_2+T-\theta)!}{\{\frac{1}{2}(q_1-p_1-n_1)+\theta\}!\,\{\frac{1}{2}(q_2-p_2-n+n_1)+T-\theta\}!\,(T_2-T+\theta)!\,(T_1-\theta)!}. \end{aligned} \tag{6.9}$$

When one of p_1, q_1, p_2, q_2, vanishes the sum over θ either reduces to a single term or can be executed, and a simpler formula results.

---

[13] Reference 12, p. 39.

VII. RECURSION FORMULAS IN s AND ISOSCALAR FACTORS FOR $u = v = 0$

From Eq. (4.4a) the heaviest state ψ_s of the product representation $(p_1 + p_2 - 2s - u - v, q_1 + q_2 + s - u - v, u)$ is a product of the heaviest states of seven representations (0, 0), (0, 0), (0, s), $(p_1-s-u, 0)$, $(p_2-v-s, 0)$, $(0, q_1-v)$, $(0, q_2-u)$. Since they appear in a fully stretched configuration, we may combine them in any convenient order, with the appropriate Clebsch–Gordan coefficients, to get the general state of the product representation. In the preceding section, where the first three were absent, we combined the fourth with the fifth and the sixth with the seventh. With all seven present, there is considerably more freedom, and complication.

More generally, any of the seven factor representations can be factored further, giving rise to relationships of recursion type between Clebsch–Gordan coefficients. A few of these are derived in this and the following section. First let us see by way of illustration as well as for its own inherent interest, how the same device may be used to find relations among $SU(2)$ Clebsch–Gordan coefficients. Equations (I, 12) with replacement $j_1 \to j_1 + A$, $j_2 \to j_2 + A$ and a factor $(\xi_1\eta_2 - \xi_2\eta_1)^{2A}$ split off give

$$\psi^{j_1+A,\,j_2+A}_{j,m} = \left[\frac{(j_1 + j_2 - j)!\,(j + j_1 + j_2 + 1)!}{(j_1 + j_2 + 2A - j)!\,(j + j_1 + j_2 + 2A + 1)!}\right]^{\frac{1}{2}} (\xi_1\eta_2 - \xi_2\eta_1)^{2A}\psi^{j_1 j_2}_{jm}. \tag{7.1}$$

When $(\xi_1\eta_2 - \xi_2\eta_1)^{2A}$ and $\psi^{j_1 j_2}_{jm}$ are expanded in terms of the factor eigenstates we get a relationship between the two sets of Clebsch–Gordan coefficients

$$\left\langle \begin{matrix} j_1 + A, & j_2 + A \\ m_1, & m_2 \end{matrix} \middle| \begin{matrix} j \\ m \end{matrix} \right\rangle \quad \text{and} \quad \left\langle \begin{matrix} j_1, & j_2 \\ m_1, & m_2 \end{matrix} \middle| \begin{matrix} j \\ m \end{matrix} \right\rangle.$$

In fact

$$\begin{aligned} \left\langle \begin{matrix} j_1 + A, & j_2 + A \\ m_1, & m_2 \end{matrix} \middle| \begin{matrix} j \\ m \end{matrix} \right\rangle = {} & \left[\frac{(j_1 + j_2 - j)!\,(j + j_1 + j_2 + 1)!}{(j_1 + j_2 + 2A - j)!\,(j + j_1 + j_2 + 2A + 1)!}\right]^{\frac{1}{2}} \\ & \times \sum_{m_1'} \left\langle \begin{matrix} j_1, & j_2 \\ m_1', & m_2' \end{matrix} \middle| \begin{matrix} j \\ m \end{matrix} \right\rangle \begin{bmatrix} 2A \\ A - m_1' + m_1 \end{bmatrix} (-1)^{A+m_2-m_2'} \\ & \times \left[\frac{(j_1 + A + m_1)!\,(j_1 + A - m_1)!\,(j_2 + A + m_1)!\,(j_2 + A - m_2)!}{(j_1 + m_1')!\,(j_1 - m_1')!\,(j_2 + m_2')!\,(j_2 - m_2')!}\right]^{\frac{1}{2}}; \end{aligned} \tag{7.2}$$

as one extreme we may take $j_1 + j_2 = j$; then (7.2) is the usual Racah formula for Clebsch–Gordan coefficients. As the other extreme we may take $A = \frac{1}{2}$; then (7.3) gives a recursion formula for Clebsch–Gordan coefficients:

$$\begin{aligned} \left\langle \begin{matrix} j_1 + \frac{1}{2}, & j_2 + \frac{1}{2} \\ m_1, & m_2 \end{matrix} \middle| \begin{matrix} j \\ m \end{matrix} \right\rangle = {} & \Big\{[(j_1 + m_1 + \tfrac{1}{2})(j_2 - m_2 + \tfrac{1}{2})]^{\frac{1}{2}} \left\langle \begin{matrix} j_1, & j_2 \\ m_1 - \frac{1}{2}, & m_2 + \frac{1}{2} \end{matrix} \middle| \begin{matrix} j \\ m \end{matrix} \right\rangle \\ & - [(j_1 - m_1 + \tfrac{1}{2})(j_2 + m_2 + \tfrac{1}{2})]^{\frac{1}{2}} \left\langle \begin{matrix} j_1, & j_2 \\ m_1 + \frac{1}{2}, & m_2 - \frac{1}{2} \end{matrix} \middle| \begin{matrix} j \\ m \end{matrix} \right\rangle \Big\} [(j_1 + j_2 - j + 1)(j_1 + j_2 + j + 2)]^{-\frac{1}{2}}. \end{aligned} \tag{7.3}$$

This is equivalent to the recursion formula recently derived by Shelepin[9] by other methods.

In the same spirit we may split off a factor $(\eta_1\zeta_2 - \zeta_1\eta_2)^{\sigma}$ from ψ_s in (4.4a). Let us denote by (s) the representation formed from (p_1q_1) and (p_2q_2) whose heaviest state is ψ_s, and by $(s - \sigma)$ the similar representation formed from $(p_1 - \sigma, q_1)$ and $(p_2 - \sigma, q_2)$ whose heaviest state is ψ_s with s replaced by $s - \sigma$. Call the corresponding normalization factors N_s and $N_{s-\sigma}$, respectively. We have not been able to get a simple expression for N in the general case because of the complication occasioned by projecting out $G \neq 0$ states. [See, however, the discussion of orthonormalization following Eq. (8.1).] But if $u = v = 0$, $G \neq 0$ states do not enter (this is easy to see in the case of the heaviest state ψ_s

which contains only perimeter states of the factor representations). Then the normalization is straightforward:

$$N_s=\left[\frac{(p_1+p_2-2s+1)!}{s!\,q_1!\,q_2!\,(p_1-s)!\,(p_2-s)!\,(p_1+p_2-s+1)!}\right]^{\frac{1}{2}},$$
$$u=v=0, \qquad (7.4)$$

and

$$N_{s-\sigma}=\left[\frac{s!\,(p_1+p_2-s+1)!}{(s-\sigma)!\,(p_1+p_2-s-\sigma+1)!}\right]^{\frac{1}{2}}N_s,$$
$$u=v=0. \qquad (7.5)$$

A word here is required about the method of projecting out $G\neq 0$ states from the factor spaces. Two representations of the same space may be combined to give a product representation in much the same way as two representations from different spaces. However a glance at (4.4a) shows that only the stretched combination arises. For if $s\neq 0$ the product representation vanishes and if $u\neq 0$ or $v\neq 0$, the product states have $G\neq 0$.

We can multiply

$$P\left|\begin{matrix}p_1,\ q_1\\ n_1,\ T_1,\ T_{13}\end{matrix}\right\rangle_1\left|\begin{matrix}p_2,\ q_2\\ n_2,\ T_2,\ T_{23}\end{matrix}\right\rangle_1$$

by the identity operator in the form

$$1=\sum_{p,q,n,T,T_3}\left|\begin{matrix}p,\ q\\ n,\ T,\ T_3\end{matrix}\right\rangle_1{}_1\left\langle\begin{matrix}p,\ q\\ n,\ T,\ T_3\end{matrix}\right|$$

and get, according to the Wigner–Eckart theorem,

$$P\left|\begin{matrix}p_1,\ q_1\\ n_1,\ T_1,\ T_{13}\end{matrix}\right\rangle_1\left|\begin{matrix}p_2,\ q_2\\ n_2,\ T_2,\ T_{23}\end{matrix}\right\rangle_1=\left[\frac{(p_1+p_2)!\,(q_1+q_2)!}{p_1!\,p_2!\,q_1!\,q_2!}\right]^{\frac{1}{2}}\sum_T\left|\begin{matrix}p_1+p_2,\ q_1+q_2\\ n_1+n_2,\ T,\ T_{13}+T_{23}\end{matrix}\right\rangle$$
$$\times\left\langle\begin{matrix}T_1,\ T_2\\ T_{13},\ T_{23}\end{matrix}\middle|\begin{matrix}T\\ T_{13}+T_{23}\end{matrix}\right\rangle\left\langle\begin{matrix}p_1,\ q_1,\ p_2,\ q_2\\ n_1,\ T_1,\ n_2,\ T_2\end{matrix}\middle|\begin{matrix}p_1+p_2,\ q_1+q_2\\ n_1+n_2,\ T\end{matrix}\right\rangle. \qquad (7.6)$$

The normalization factor in (7.6) which must be independent of the quantum numbers n, T, T_3 is evaluated easily by examining the heaviest state.

Since the projection operator P is just an instruction: "retain only stretched combinations of representations in the same space," such combinations may be made in any convenient order, i.e., P inserted wherever convenient in a product. The normalization factor in (7.6) is compatible with this observation.

The general state of (s) may be written

$$\left|\begin{matrix}s\\ n,\ T,\ T_3\end{matrix}\right\rangle=\frac{N_s}{N_{s-\sigma}}\,[s!(s+1)!]^{\frac{1}{2}}\,P\sum_{n',T',T_3'}\left|\begin{matrix}s-\sigma\\ n',\ T',\ T_3'\end{matrix}\right\rangle\left|\begin{matrix}0,\ \sigma\\ n-n',\ \frac{1}{2}(\sigma-n+n'),\ T_3-T_3'\end{matrix}\right\rangle$$
$$\times\left\langle\begin{matrix}T',\ \frac{1}{2}(\sigma-n+n')\\ T_3',\ T_3-T_3'\end{matrix}\middle|\begin{matrix}T\\ T_3\end{matrix}\right\rangle\left\langle\begin{matrix}s-\sigma,\ 0,\ \sigma\\ n',\ T',\ n-n',\ \frac{1}{2}(\sigma-n+n')\end{matrix}\middle|\begin{matrix}s\\ n,\ T\end{matrix}\right\rangle. \qquad (7.7)$$

The stretched isoscalar factor on the right is simple (no sums) because its "p_2" is zero. Now expand in terms of the factor representations and combine them using (7.6). The T_3 sums can then be done to give

$$\left|\begin{matrix}s\\ n,\ T,\ T_3\end{matrix}\right\rangle$$

directly in terms of the factor eigenstates and the desired isoscalar factors can be read off:

$$\left\langle\begin{matrix}p_1,\ q_1,\ p_2,\ q_2\\ n_1,\ T_1,\ n-s-n_1,\ T_2\end{matrix}\middle|\begin{matrix}s\\ n,\ T\end{matrix}\right\rangle$$
$$=\frac{N_s}{N_{s-\sigma}}\sum_{\substack{n',n_1',\\ T',T_1',T_3'}}[(2T'+1)(\sigma-n+n'+1)(2T_1+1)(2T_2+1)]^{\frac{1}{2}}X\begin{bmatrix}T_1' & \frac{1}{2}(\sigma+n_1') & T_1\\ T_2' & \frac{1}{2}(n-n'-n_1') & T_2\\ T' & \frac{1}{2}(\sigma-n+n') & T\end{bmatrix}$$
$$\times\left[\frac{[s!(s+1)!]^{\frac{1}{2}}\,p_1!\,p_2!}{(p_1-\sigma)!\,(p_2-\sigma)!\,(\sigma!)^2}\right]^{\frac{1}{2}}\left\langle\begin{matrix}p_1-\sigma,\ q_1,\ p_2-\sigma,\ q_2\\ n_1-n_1',\ T_1',\ n'-s+\sigma-n_1+n_1',\ T_2'\end{matrix}\middle|\begin{matrix}s-\sigma\\ n'T'\end{matrix}\right\rangle$$

$$\times \left\langle \begin{matrix} \sigma, & 0, & \sigma, & 0 \\ n_1', & \tfrac{1}{2}(\sigma + n_1'), & n - n' - \sigma - n_1', & \tfrac{1}{2}(n - n' - n_1') \end{matrix} \middle| \begin{matrix} 0, & \sigma \\ n - n', & \tfrac{1}{2}(\sigma - n + n') \end{matrix} \right\rangle$$

$$\times \left\langle \begin{matrix} s - \sigma, & 0, & \sigma \\ n', & T', & n - n', & \tfrac{1}{2}(\sigma - n + n_1') \end{matrix} \middle| \begin{matrix} s \\ n, T \end{matrix} \right\rangle \left\langle \begin{matrix} p_1 - \sigma, & q_1, & \sigma, & 0 \\ n_1 - n_1', & T_1', & n_1', & \tfrac{1}{2}(\sigma + n_1') \end{matrix} \middle| \begin{matrix} p_1, & q_1 \\ n_1, & T_1 \end{matrix} \right\rangle$$

$$\times \left\langle \begin{matrix} p_2 - \sigma, & q_2, & \sigma, & 0 \\ n' - s + \sigma - n_1 + n_1', & T_2', & n - n' - \sigma - n_1', & \tfrac{1}{2}(n - n' - n_1') \end{matrix} \middle| \begin{matrix} p_2, & q_2 \\ n - s - n_1, & T_2 \end{matrix} \right\rangle. \tag{7.8}$$

Equation (7.8) expresses

$$\left\langle \begin{matrix} p_1, & q_1, & p_2, & q_2 \\ n_1, & T_1, & n - s - n_1, & T_2 \end{matrix} \middle| \begin{matrix} s \\ n, T \end{matrix} \right\rangle$$

in terms of

$$\left\langle \begin{matrix} p_1 - \sigma, & q_1, & p_2 - \sigma, & q_2 \\ n_1, & T_1, & n - s + \sigma - n_1, & T_2 \end{matrix} \middle| \begin{matrix} s - \sigma \\ n, T \end{matrix} \right\rangle$$

and other isoscalar factors which are known; three of these are stretched with "p_2" = 0 or "q_2" = 0 and the fourth is given by (5.9).

The quintuple sums and the X coefficient make Eq. (7.8) rather formidable in general. But in the limiting case $\sigma = 1$ it becomes a rather simple recursion formula in s. There are three possible sets of values of n_1', n', namely, $n_1' = 0$, $n' = n$; $n_1' = 0$, $n' = n - 1$; $n_1' = -1$, $n' = n$. The X coefficient is given for the cases by, respectively,

$$X \begin{bmatrix} T_1' & \tfrac{1}{2} & T_1 \\ T_2' & 0 & T_2 \\ T' & \tfrac{1}{2} & T \end{bmatrix}, \quad X \begin{bmatrix} T_1' & \tfrac{1}{2} & T_1 \\ T_2' & \tfrac{1}{2} & T_2 \\ T' & 0 & T \end{bmatrix}, \quad X \begin{bmatrix} T_1' & 0 & T_1 \\ T_2' & \tfrac{1}{2} & T_2 \\ T' & \tfrac{1}{2} & T \end{bmatrix}.$$

There are four possible values of T_1', T_2', T' in each case, so the quintuple sum in (7.8) consists of twelve terms. Because of the presence of 0 and $\frac{1}{2}$ the X coefficients are simple; for example, the first one, with $T_1' = T_1 - \frac{1}{2}$, $T' = T - \frac{1}{2}$ is

$$\left[\frac{(T_1 + T_2 + T + 1)(T + T_1 - T_2)}{4T_1(2T_1 + 1)2T(2T - 1)(2T_2 + 1)}\right]^{\frac{1}{2}}.$$

The other limiting case of (7.8), $\sigma = s$, gives the isoscalar factors for finite s in terms of those with $s = 0$, i.e., in the case $u = v = 0$, in terms of stretched isoscalar factors. For the case $u = v = 0$, however, it is possible to derive a simpler formula. As a first step we find the isoscalar factors for the case $u = v = q_1 = q_2 = 0$. The general state of the product representation is then

$$\left| \begin{matrix} p_1 + p_2 - 2s, & s \\ n, & T, \ T_3 \end{matrix} \right\rangle$$

$$= \left[\frac{(p_1 + p_2 - 2s + 1)!\,(s + 1)!}{(p_1 + p_2 - s + 1)!}\right]^{\frac{1}{2}} \sum_{n', T_3'} \left| \begin{matrix} p_1 + p_2 - 2s, & 0 \\ n', & \tfrac{1}{2}(p_1 + p_2 - 2s + n'), \ T_3' \end{matrix} \right\rangle$$

$$\times \left| \begin{matrix} 0, & s \\ n - n', & \tfrac{1}{2}(s - n + n'), \ T_3 - T_3' \end{matrix} \right\rangle \left\langle \begin{matrix} \tfrac{1}{2}(p_1 + p_2 - 2s + n'), & \tfrac{1}{2}(s - n + n') \\ T_3', & T_3 - T_3' \end{matrix} \middle| \begin{matrix} T \\ T_3 \end{matrix} \right\rangle$$

$$\times \left\langle \begin{matrix} p_1 + p_2 - 2s, & 0, & 0, & s \\ n', & \tfrac{1}{2}(p_1 + p_2 - 2s + n'), & n - n', & \tfrac{1}{2}(s - n + n') \end{matrix} \middle| \begin{matrix} p_1 + p_2 - 2s, & s \\ n, & T \end{matrix} \right\rangle.$$

Now expand in terms of the factor representations and recouple the isospins. The desired isoscalar factors can then be read off:

$$\left\langle \begin{matrix} p_1, & 0, & p_2, & 0 \\ n_1, & \tfrac{1}{2}(p_1 + n_1), & n - s - n_1, & \tfrac{1}{2}(p_2 + n - s - n_1) \end{matrix} \middle| \begin{matrix} p_1 + p_2 - 2s, & s \\ n, & T \end{matrix} \right\rangle$$

$$= \sum_{n_1', n_2'} \left[\frac{(p_1 + p_2 - 2s + 1)!\,(s + 1)!\,(-n_1)!\,(p_1 + n_1)!\,(s - n + n_1)!\,(p_2 - s + n - n_1)!}{(p_1 + p_2 - s + 1)!\,(-n_1')!\,(-n_1 + n_1')!\,(p_1 - s + n_1')!\,(s + n_1 - n_1')!\,(-n_2')!\,(s - n + n_2' + n_1)!}\right]^{\frac{1}{2}}$$

$$\times \frac{1}{[(p_2 - s + n_2')!\,(n - n_2' - n_1)!]^{\frac{1}{2}}} X \begin{bmatrix} \tfrac{1}{2}(p_1 - s + n_1') & \tfrac{1}{2}(s + n_1 - n_1') & \tfrac{1}{2}(p_1 + n_1) \\ \tfrac{1}{2}(p_2 - s + n_2') & \tfrac{1}{2}(n - n_2' - n_1) & \tfrac{1}{2}(p_2 - s + n - n_1) \\ \tfrac{1}{2}(p_1 + p_2 - 2s + n_1' + n_2') & \tfrac{1}{2}(s - n + n_2' + n_1') & T \end{bmatrix}$$

$$\times [(p_1+n_1+1)(p_2-s+n-n_1+1)(p_1+p_2-2s+n_1'+n_2'+1)(s-n+n_2'+n_1')]^{\frac{1}{2}}$$

$$\times \left\langle \begin{matrix} p_1-s, & 0, & p_2-s, & 0 \\ n_1', & \frac{1}{2}(p_1-s+n_1'), & n_2', & \frac{1}{2}(p_2-s+n_2) \end{matrix} \middle| \begin{matrix} p_1+p_2-2s, & 0 \\ n_1'+n_2', & \frac{1}{2}(p_1+p_2-2s+n_1'+n_2') \end{matrix} \right\rangle$$

$$\times \left\langle \begin{matrix} s, & 0, & s, & 0 \\ n_1-n_1', & \frac{1}{2}(s+n_1-n_1'), & -s+n-n_2'-n_1, & \frac{1}{2}(n-n_2'-n_1) \end{matrix} \middle| \begin{matrix} 0, & s \\ n-n_2'-n_1', & \frac{1}{2}(s-n+n_2'+n_1') \end{matrix} \right\rangle$$

$$\times \left\langle \begin{matrix} p_1+p_2-2s, & 0, & 0, & s \\ n_2'+n_1', & \frac{1}{2}(p_1+p_2-2s+n_2'+n_1'), & n-n_2'-n_1', & \frac{1}{2}(s-n+n_2'+n_1') \end{matrix} \middle| \begin{matrix} p_1+p_2-2s, & s \\ n, & T \end{matrix} \right\rangle. \tag{7.9}$$

The isoscalar factors on the right are given by (5.4), (5.9), and (5.6), respectively. The X coefficient is doubly stretched and is given by (6.8).

The $u = v = 0$ isoscalar factors can now be derived. The general state of the product representation is

$$\left| \begin{matrix} p_1+p_2-2s, & q_1+q_2+s \\ n, & T, \quad T_3 \end{matrix} \right\rangle$$

$$= \sum_{n',T',T_3'} \left| \begin{matrix} p_1+p_2-2s, & s \\ n', & T', \quad T_3' \end{matrix} \right\rangle \left| \begin{matrix} 0, & q_1+q_2 \\ n-n', & \frac{1}{2}(q_1+q_2-n+n'), \quad T_3-T_3' \end{matrix} \right\rangle \left\langle \begin{matrix} T', & \frac{1}{2}(q_1+q_2-n+n') \\ T_3', & T_3-T_3' \end{matrix} \middle| \begin{matrix} T \\ T_3 \end{matrix} \right\rangle$$

$$\times \left\langle \begin{matrix} p_1+p_2-2s, & s, & 0, & q_1+q_2 \\ n', & T', & n-n', & \frac{1}{2}(q_1+q_2-n+n') \end{matrix} \middle| \begin{matrix} p_1+p_2-2s, & q_1+q_2+s \\ n, & T \end{matrix} \right\rangle.$$

Expand in terms of the factor states and recouple isospins; then compare with the same state expanded directly in one-particle states in the manner of (6.4). One finds for the $u = v = 0$ isoscalar factor

$$\left\langle \begin{matrix} p_1, & q_1, & p_2, & q_2 \\ n_1, & T_1, & n-s-n_1, & T_2 \end{matrix} \middle| \begin{matrix} p_1+p_2-2s, & q_1+q_2+s \\ n, & T \end{matrix} \right\rangle$$

$$= \sum_{T_3'} \left\langle \begin{matrix} p_1, & 0, & p_2, & 0 \\ n_1', & \frac{1}{2}(p_1+n_1'), & -s+n_2', & \frac{1}{2}(p_2-s+n_2') \end{matrix} \middle| \begin{matrix} p_1+p_2-2s, & s \\ n_1'+n_2', & T_3' \end{matrix} \right\rangle$$

$$\times \left\langle \begin{matrix} 0, & q_1, & 0, & q_2 \\ n_1-n_1', & \frac{1}{2}(q_1-n_1+n_1'), & n-n_1-n_2', & \frac{1}{2}(q_2-n+n_1+n_2') \end{matrix} \middle| \begin{matrix} 0, & q_1+q_2 \\ n-n_1'-n_2', & \frac{1}{2}(q_1+q_2-n+n_1'+n_2') \end{matrix} \right\rangle$$

$$\times \left\langle \begin{matrix} p_1+p_2-2s, & s, & 0, & q_1+q_2 \\ n_1'+n_2', & T_3', & n-n_1'-n_2', & \frac{1}{2}(q_1+q_2-n+n_1'+n_2') \end{matrix} \middle| \begin{matrix} p_1+p_2-2s, & q_1+q_2+s \\ n, & T \end{matrix} \right\rangle$$

$$\times [(2T_3'+1)(2T_1+1)(2T_2+1)(q_1+q_2-n+n_1'+n_2'+1]^{\frac{1}{2}} X \begin{bmatrix} \frac{1}{2}(p_1+n_1') & \frac{1}{2}(q_1-n_1+n_1') & T_1 \\ \frac{1}{2}(p_2-s+n_2') & \frac{1}{2}(q_2-n+n_2'-n_1) & T_2 \\ T_3' & \frac{1}{2}(q_1+q_2-n+n_1'+n_2') & T \end{bmatrix}$$

$$\times \frac{1}{\left\langle \begin{matrix} p_1, & 0, & 0, & q_1 \\ n_1', & \frac{1}{2}(p_1+n_1'), & n-n_1', & \frac{1}{2}(q_1-n_1+n_1') \end{matrix} \middle| \begin{matrix} p_1, & q_1 \\ n_1, & T_1 \end{matrix} \right\rangle}$$

$$\times \frac{1}{\left\langle \begin{matrix} p_2, & 0, & 0, & q_2 \\ n_2'-s, & \frac{1}{2}(p_2+n_2'-s), & n-n_1-n_2', & \frac{1}{2}(q_2-n+n_1+n_2') \end{matrix} \middle| \begin{matrix} p_2, & q_2 \\ -s+n-n_1, & T_2 \end{matrix} \right\rangle}. \tag{7.10}$$

The five isoscalar factors on the right are given by (7.9), (5.5), (6.9) with "p_2" = 0, (5.6), (5.6). The X coefficient is singly stretched and given by (6.7). The right side of course cannot depend on n_1', n_2'. Factors containing them cancel when the expressions for the isoscalar factors and X coefficient are inserted.

We have worked with case (a) in this section, i.e., representations whose heaviest state is given by (4.4a). Case (b), corresponding to (4.4b) can, of course, be dealt with by the same methods.

VIII. RECURSION FORMULAS IN u, v AND GENERAL ISOSCALAR FACTORS

We have not found a simple formula for the normalizing factor in the general case; nor have we solved the related problem of explicitly orthogonalizing the degenerate representations with the same $u + v$. In order to define a set of orthogonal basis vectors we can regard u as an operator. This will not be Hermitian as it stands, but we can find its Hermitian part. The eigenvectors of this operator constitute a complete orthonormal set labeled by u.[14]

To construct the general isoscalar factor we first project the $G \neq 0$ states out of the product

$$(\xi_1\xi_2^* + \eta_1\eta_2^* + \zeta_1\zeta_2^*)^u(\xi_1^*\xi_2 + \eta_1^*\eta_2 + \zeta_1^*\zeta_2)^v$$

which appears in (4.4a). It is apparent that the state we want is

$$C \sum_{n,T,T_3} \left|\begin{matrix} u, & & v \\ n, & T, & T_3 \end{matrix}\right\rangle_1 \left|\begin{matrix} v, & u & \\ -n, & T, & -T_3 \end{matrix}\right\rangle_2 \left\langle\begin{matrix} T, & T \\ T_3, & -T_3 \end{matrix}\right| \left.\begin{matrix} 0 \\ 0 \end{matrix}\right\rangle \left\langle\begin{matrix} u, & v, & v, & u \\ n, & T, & -n, & T \end{matrix}\right| \left.\begin{matrix} 0, & 0 \\ 0, & 0 \end{matrix}\right\rangle.$$

The isoscalar factor is given by (5.11). By comparing coefficients of $(\eta_1\eta_2^*)^u(\xi_1^*\xi_2)^v$ the constant is found to be $C = u!v![\frac{1}{2}(u+1)(v+1)(u+v+2)]^{\frac{1}{2}}$.

We may write the general state of the product representation [for case (a)]

$$\left|\begin{matrix} p_1+p_2-u-v-2s, & & q_1+q_2-u-v+s \\ n, & T, & T_3 \end{matrix}\right\rangle = NP \sum_{n_1',T_1',T_2',T_{13}'} \left|\begin{matrix} u, & & v \\ n', & T', & T_3' \end{matrix}\right\rangle_1 \left|\begin{matrix} v, & & u \\ -n', & T', & -T_3' \end{matrix}\right\rangle_2$$

$$\times \left\langle\begin{matrix} T', & T' \\ T_3', & -T_3' \end{matrix}\right| \left.\begin{matrix} 0 \\ 0 \end{matrix}\right\rangle \left\langle\begin{matrix} u, & v, & v, & u \\ n, & T, & -n, & T \end{matrix}\right| \left.\begin{matrix} 0, & 0 \\ 0, & 0 \end{matrix}\right\rangle \left|\begin{matrix} p_1-u, & & q_1-v \\ n_1', & T_1', & T_{13}' \end{matrix}\right\rangle_1 \left|\begin{matrix} p_2-v, & & q_2-u \\ n-n_1'-s, & T_2', & T_3-T_{13}' \end{matrix}\right\rangle_2$$

$$\times \left\langle\begin{matrix} T_1', & T_2' \\ T_{13}', & T_3-T_{13}' \end{matrix}\right| \left.\begin{matrix} T \\ T_3 \end{matrix}\right\rangle \left\langle\begin{matrix} p_1-u, & q_1-v, & p_2-v, & q_2-u \\ n_1', & T_1', & n-n_1'-s, & T_2' \end{matrix}\right| \left.\begin{matrix} p_1+p_2-u-v-2s, & q_1+q_2-u-v+s \\ n, & T \end{matrix}\right\rangle.$$

Now the factor states can be combined using (7.6), the T_3 sums performed and the general isoscalar factor read off:

$$\left\langle\begin{matrix} p_1, & q_1, & p_2, & q_2 \\ n_1, & T_1, & n-s-n_1, & T_2 \end{matrix}\right| \left.\begin{matrix} p_1+p_2-u-v-2s, & q_1+q_2-u-v+s \\ n, & T \end{matrix}\right\rangle$$

$$= N \frac{1}{u!\,v!} \left[\frac{p_1!\,q_1!\,p_2!\,q_2!\,(2T_1+1)(2T_2+1)(2T+1)}{(p_1-u)!\,(q_1-v)!\,(p_2-v)!\,(q_2-u)!}\right]^{\frac{1}{2}} \sum_{\substack{n_1',T' \\ T_1',T_2'}} \left\langle\begin{matrix} u, & v, & p_1-u, & q_1-v \\ n_1-n_1', & T', & n_1', & T_1' \end{matrix}\right| \left.\begin{matrix} p_1, & q_1 \\ n_1, & T_1 \end{matrix}\right\rangle$$

$$\times \left\langle\begin{matrix} v, & u, & p_2-v, & q_2-u \\ -n_1+n_1', & T', & n-n_1'-s, & T_2' \end{matrix}\right| \left.\begin{matrix} p_2, & q_2 \\ n-s-n_1, & T_2 \end{matrix}\right\rangle \left\langle\begin{matrix} u, & v, & v, & u \\ n_1-n_1', & T', & -n_1+n_1', & T' \end{matrix}\right| \left.\begin{matrix} 0, & 0 \\ 0, & 0 \end{matrix}\right\rangle$$

$$\times \left\langle\begin{matrix} p_1-u, & q_1-v, & p_2-v, & q_2-u \\ n_1', & T_1', & n-n_1'-s, & T_2' \end{matrix}\right| \left.\begin{matrix} p_1+p_2-u-v-2s, & q_1+q_2-u-v+s \\ n, & T \end{matrix}\right\rangle X \begin{bmatrix} T' & T' & 0 \\ T_1' & T_2' & T \\ T_1 & T_2 & T \end{bmatrix}. \tag{8.1}$$

The four isoscalar factors on the right are given, respectively, by (6.9), (6.9), (5.11), (7.10). The X coefficient, having one vanishing element, can be expressed in terms of a Racah coefficient.[15] For the purpose of orthonor-

[14] This procedure is similar to a more general method for resolving degeneracies developed by L. O'Raifeartaigh and A. J. Macfarlane (private communication).

[15] Reference 12, p. 192, Eq. (11.12) and p. 227, Table I. 3.

malizing the degenerate representations it is necessary to consider only the heaviest states, for which

$$n = 0,\ T = \tfrac{1}{2}(p_1 + p_2 + q_1 + q_2 - 2u - 2v - s).$$

For those states $T_1' = \frac{1}{2}(p_1 + q_1 - u - v + n_1')$, $T_2' = \frac{1}{2}(p_2 + q_2 - u - v + n - n_1' - s)$ so that in (8.1) the quadruple sum reduces to a double sum and the X coefficient simplifies further by virtue of its middle row being stretched.

We can generalize (8.1) by splitting off only

$$(\xi_1\xi_2^* + \eta_1\eta_2^* + \zeta_1\zeta_2^*)^\mu(\xi_1^*\xi_2 + \eta_1^*\eta_2 + \zeta_1^*\zeta_2)^\nu,$$

where $\mu \leq u$, $\nu \leq v$. The same procedure followed in deriving (8.1) then gives

$$\left\langle \begin{matrix} p_1, & q_1, & p_2, & q_2 \\ n_1, & T_1, & n-s-n_1, & T_2 \end{matrix} \middle| \begin{matrix} p_1+p_2-u-v-2s, & q_1+q_2-u-v+s \\ n, & T \end{matrix} \right\rangle$$

$$= \frac{N_{u,v}}{N_{u-\mu,v-\nu}} \frac{1}{\mu!\,\nu!} \left[\frac{p_1!\,q_1!\,p_2!\,q_2!\,(2T_1+1)(2T_2+1)(2T+1)}{(p_1-\mu)!\,(q_1-\nu)!\,(p_2-\nu)!\,(q_2-\mu)!}\right]^{\frac{1}{2}} \sum_{\substack{n_1',T_1' \\ n-n_1',T_2'}} \left\langle \begin{matrix} \mu, & \nu, & p_1-\mu, & q_1-\nu \\ n_1-n_1', & T', & n_1', & T_1' \end{matrix} \middle| \begin{matrix} p_1, & q_1 \\ n_1, & T_1 \end{matrix} \right\rangle$$

$$\times \left\langle \begin{matrix} \nu, & \mu, & p_2-\nu, & q_2-\mu \\ -n_1+n_1', & T', & n-n_1'-s, & T_2' \end{matrix} \middle| \begin{matrix} p_2, & q_2 \\ n-s-n_1, & T_2 \end{matrix} \right\rangle \left\langle \begin{matrix} \mu, & \nu, & \nu, & \mu \\ n_1-n_1', & T', & -n_1+n_1', & T' \end{matrix} \middle| \begin{matrix} 0, & 0 \\ 0, & 0 \end{matrix} \right\rangle$$

$$\times \left\langle \begin{matrix} p_1-\mu, & q_1-\nu, & p_2-\nu, & q_2-\mu \\ n_1', & T_1', & n-n_1'-s, & T_2' \end{matrix} \middle| \begin{matrix} p_1+p_2-u-v-2s, & q_1+q_2-u-v+s \\ n, & T \end{matrix} \right\rangle X \begin{bmatrix} T' & T' & 0 \\ T_1' & T_2' & T \\ T_1 & T_2 & T \end{bmatrix}. \tag{8.2}$$

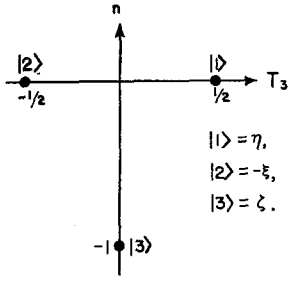


FIG. 3. The weight diagram and basis vectors for the representation (0, 1).

Equation (8.2) expresses the general isoscalar factor in terms of isoscalar factors with p_1, q_2, u reduced by μ and p_2, q_1, v, reduced by ν. $N_{u,v}$ in (8.2) is just N of Eq. (8.1) while $N_{u-\mu,v-\nu}$ is $N_{u,v}$ with p_1, q_2, u reduced by μ and p_2, q_1, v reduced by ν.

As one extreme we may take $\mu = u$, $\nu = v$. Then (8.2) reduces to (8.1). At the other extreme we may take $\mu = 1$, $\nu = 0$ or $\mu = 0$, $\nu = 1$ and obtain a recurrence formula for u or v, respectively. For $\mu = 1$, $\nu = 0$ we have $T' = \frac{1}{2}(1 + n_1 - n_1')$ with $n_1 - n_1' = 0$ or -1. Thus the X coefficient in (8.2) becomes

$$X \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ T_1' & T_2' & T \\ T_1 & T_2 & T \end{bmatrix}$$

for $n_1 - n_1' = 0$ and

$$X \begin{bmatrix} 0 & 0 & 0 \\ T_1 & T_2 & T \\ T_1 & T_2 & T \end{bmatrix} = 1$$

for $n_1 - n_1' = -1$. The quadruple sum in (8.2) reduces to five terms with rather simple coefficients.

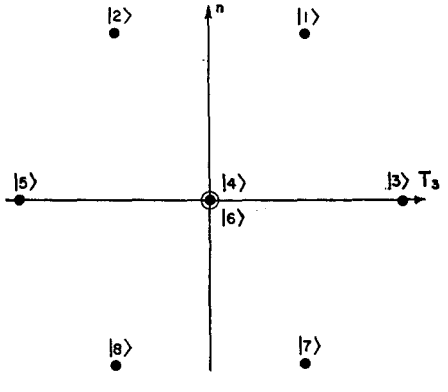


FIG. 4. The weight diagram and basis vectors for the representation (1, 1). The diagrams in Figs. 2–4 are identical to those of de Swart with the definition $n = Y - \frac{1}{3}(p - q)$.

Ising Model Spin Correlations on the Triangular Lattice. II. Fourth-Order Correlations

JOHN STEPHENSON
Mathematics Department, University of Adelaide, Adelaide, South Australia
(Received 5 October 1965)

A class of fourth-order spin correlations whose sum is closely related to the specific heat is calculated exactly by Pfaffian perturbation theory. In the particular case when the four spins lie on a lattice axis, it is shown that the reduced fourth-order correlation function has the asymptotic form

$$\langle \sigma_{0,0}\sigma_{1,1}\sigma_{k,k}\sigma_{k+1,k+1}\rangle - \langle \sigma_{0,0}\sigma_{1,1}\rangle\langle \sigma_{k,k}\sigma_{k+1,k+1}\rangle \doteq A'(e^{-2\beta k}/k^2),$$

where β^{-1} is the (positive) mean range of order. The decay of correlations with spin separation k is symmetric about the critical point $\beta = 0$. The amplitude A' is $1/\pi^2$ at the critical point and is $1/(2\pi)$ at all other temperatures. Correlations on ferro- and antiferromagnetic triangular and quadratic lattices are discussed in some detail.

1. INTRODUCTION

INTEREST in fourth-order spin correlations for the Ising model[1] stems from their close relation to the specific heat. Thus the spin correlation

$$S_4 = \langle \sigma_0\sigma_\delta\sigma_r\sigma_{r+\gamma}\rangle \tag{1.1}$$

and the corresponding reduced correlation function

$$\omega_4(\mathbf{r}) \equiv \omega(0, \boldsymbol{\delta}; \mathbf{r}, \mathbf{r}+\boldsymbol{\gamma}) = \langle \sigma_0\sigma_\delta\sigma_r\sigma_{r+\gamma}\rangle - \langle \sigma_0\sigma_\delta\rangle\langle \sigma_r\sigma_{r+\gamma}\rangle, \tag{1.2}$$

where $\boldsymbol{\delta}$ and $\boldsymbol{\gamma}$ are lattice vectors, are of particular interest in this respect. The calculation of higher (even)-order spin correlations is advantageously approached via the Pfaffian method, pictorially interpreted in terms of the associated dimer problem. The spin correlation may be expressed as a Pfaffian, the elements of which may be written down immediately on inspection of the corresponding dimer lattice problem.[2,3] The details of the method are fully described in the paper by Montroll, Potts, and Ward (Ref. 2), and its extension to the calculation of higher-order spin correlations on the triangular lattice is described in a previous paper (Ref. 3, hereafter referred to as I). The spin correlation S_4, above, may be calculated explicitly as a 4×4 Pfaffian for arbitrary spin separation $\mathbf{r}$. The simplicity of the calculation for this particular class of fourth-order correlation contrasts markedly with that for the pair correlation function

$$\omega_2(\mathbf{r}) = \langle \sigma_0\sigma_r\rangle, \tag{1.3}$$

the latter being expressed as a Toeplitz determinant whose order *increases* with increasing spin separation. For the reduced fourth-order correlation function $\omega_4(\mathbf{r})$ it is a relatively simple matter to determine its dependence on spin separation, which is of the form

$$\omega_4(\mathbf{r}) \doteq A'(e^{-2\beta r}/r^2) \tag{1.4}$$

for large spin separation r.[4] Here β^{-1} is the mean range of order. This form of the decay has only been verified for particular relative orientations of the spin pairs σ_0, σ_δ and σ_r, $\sigma_{r+\gamma}$. The specific heat C is related to the sum over all possible correlations $\omega_4(\mathbf{r})$. If one assumes that the correlations have a radial decay of the form in Eq. (1.4) in all directions, where A' *may* depend on direction, one may replace this sum by an integral and thus obtain

$$C \doteq A'' \int_1^\infty e^{-2\beta r}\frac{dr}{r} \propto \beta^0, \tag{1.5}$$

where β^0 is to be interpreted as a logarithmic dependence of the specific heat on $\beta \propto |T - T_c|$, at temperatures close to the critical point T_c. On the other hand, an attempt to evaluate C at the critical point itself would lead one to the integral

$$C \doteq B' \int_1^d \frac{dr}{r} \propto \log d, \tag{1.6}$$

where d is related to the linear dimensions of the lattice. Such a relation for the two-dimensional

[1] (a) C. Domb, Advan. Phys. 9, Nos. 34, 35 (1960). (b) M. E. Fisher, J. Math. Phys. 4, 278 (1963). (c) H. S. Green and C. A. Hurst, *Order Disorder Phenomena*, Vol. 5 of *Monographs in Statistical Physics and Thermodynamics* (Interscience Publishers, New York, 1964). (d) C. A. Hurst and H. S. Green, J. Chem. Phys. **33**, 1059 (1960). (e) P. W. Kasteleyn, J. Math. Phys. **4**, 287 (1963).

[2] E. W. Montroll, R. B. Potts, and J. C. Ward, J. Math. Phys. **4**, 308 (1963).

[3] J. Stephenson, J. Math. Phys. **5**, 1009 (1964).

[4] M. E. Fisher (private communication) has suggested that this form of the correlation decay would hold near the critical point.

lattices was conjectured by Kramers and Wannier[5] on the basis of numerical calculations of C and was confirmed rigorously by Onsager for the quadratic lattice.[6] The decay of the fourth-order correlations is therefore, within the tentative nature of the above calculation, consistent with this behavior. The decay of the correlations is symmetric in $\beta \propto |T - T_c|$ about the critical point. The amplitude A may be calculated by first fixing the temperature, and then allowing the spin separation to increase. A subsequent interchange of the limiting process, as attempted in Eq. (1.5), is not permissible. The value of A' at the critical point *differs* from its value at temperatures above or below T_c. This is indicative that calculations of the decay amplitudes for pair correlations at the critical point are also beset by this kind of complication.

2. FOURTH-ORDER CORRELATION FUNCTIONS

It is convenient to calculate the correlations for a triangular lattice, which reduces to a quadratic lattice on setting one of the interaction energies equal to zero. Consider a triangular lattice of N spins, each of which interacts with its six nearest neighbors (nn). Let $-J_1$, $-J_2$, $-J_3$ be the interaction energies between horizontal, vertical, and diagonal neighboring pairs of spins, respectively, and define

$$K_i = J_i/kT, \qquad z_i = \exp(-2K_i), \tag{2.1}$$

$$v_i = \tanh K_i = \frac{1 - z_i}{1 + z_i} = \exp(-2K_i^*), \tag{2.2}$$

where v_i is essentially the "dual" transformation variable.[7] Then the square of the partition function may be written[2,3]

$$Z_N^2 = (2 \cosh K_1 \cosh K_2 \cosh K_3)^{2N} |\mathbf{A}|, \tag{2.3}$$

where $\mathbf{A}$ is a $6N \times 6N$ matrix which has a doubly cyclic structure (in the limit of large N, when edge effects may be neglected), and its nonzero elements are 6×6 matrices.[3] The matrix $\mathbf{A}$ can be block diagonalized by a Fourier-type unitary transformation, and the value of $|\mathbf{A}|$ can be obtained from the determinant:

$$\mathbf{A}(\phi_1, \phi_2) = \begin{array}{c|cccccc} & R & S & U & L & T & D \\ R & 0 & 1 & 1 & 1 - v_1 e^{i\phi_1} & 1 & 1 \\ S & -1 & 0 & 1 & 1 & 1 - v_3 e^{i(\phi_1+\phi_2)} & 1 \\ U & -1 & -1 & 0 & 1 & 1 & 1 - v_2 e^{i\phi_2} \\ L & -1 + v_1 e^{-i\phi_1} & -1 & -1 & 0 & 1 & 1 \\ T & -1 & -1 + v_3 e^{-i(\phi_1+\phi_2)} & -1 & -1 & 0 & 1 \\ D & -1 & -1 & -1 + v_2 e^{-i\phi_2} & -1 & -1 & 0 \end{array}. \tag{2.4}$$

The lattice axes are labeled *RSULTD* in counterclockwise order from the horizontal axis, and the lattice orientation has been modified slightly from that used in I. It is most convenient for correlation calculations to have the elements signed in a regular manner as in (2.4). Then for large N,

$$\ln |\mathbf{A}| \simeq \frac{N}{(2\pi)^2} \iint_{-\pi}^{\pi} \ln \Delta(\phi_1, \phi_2)\, d\phi_1\, d\phi_2, \tag{2.5}$$

where $\Delta(\phi_1, \phi_2) = |\mathbf{A}(\phi_1, \phi_2)|$,[3] and the partition function for the triangular Ising lattice is given by

$$\frac{\ln Z_N}{N} = \ln 2 + \frac{1}{2(2\pi)^2} \iint_{-\pi}^{\pi} d\phi_1\, d\phi_2 \times \ln [C_1C_2C_3 + S_1S_2S_3 - S_1 \cos\phi_1 - S_2 \cos\phi_2 - S_3 \cos(\phi_1 + \phi_2)], \tag{2.6}$$

where

$$C_i = \cosh 2K_i \quad \text{and} \quad S_i = \sinh 2K_i.$$

Further, the inverse of $\mathbf{A}$ is required

$$\mathbf{A}^{-1}(p, q; p', q') = \frac{1}{(2\pi)^2} \iint_{-\pi}^{\pi} \exp i[(p - p')\phi_1 + (q - q')\phi_2] \times \mathbf{A}^{-1}(\phi_1, \phi_2)\, d\phi_1\, d\phi_2, \tag{2.7}$$

where $\mathbf{A}^{-1}(\phi_1, \phi_2)$ is the inverse matrix of $\mathbf{A}(\phi_1, \phi_2)$. The elements of $\mathbf{A}^{-1}$ depend only on the differences $p - p'$, $q - q'$, so the abbreviated notation

$$[p' - p, q' - q]_{IJ} = \mathbf{A}^{-1}(p, q; p', q')_{IJ} \tag{2.8}$$

is used. $\mathbf{A}^{-1}$ is antisymmetric:

$$[p' - p, q' - q]_{IJ} = -[p - p'; q - q']_{JI}. \tag{2.9}$$

[5] H. A. Kramers and G. H. Wannier, Phys. Rev. 60, 252, 263 (1941).
[6] L. Onsager, Phys. Rev. 65, 117 (1944).
[7] The notation is the same as that in I (Ref. 3).

Further, it can be shown that

$$[k, 0]_{RR} = [k, 0]_{LL} = [k, k]_{SS} = [k, k]_{TT} = [0, k]_{UU} = [0, k]_{DD} = 0. \quad (2.10)$$

The elements of $\mathbf{A}^{-1}(\phi_1, \phi_2)$ depend on the cofactors of $\mathbf{A}(\phi_1, \phi_2)$. These are denoted by C_{IJ}. They are tabulated in I, but the formulas given there must be amended to agree with (2.4) by judicious use of the factor -1.

The Pfaffian perturbation theory presented in Refs. 2 and 3 can be applied to evaluate the fourth-order spin correlation

$$S_4 = \langle \sigma_0 \sigma_\delta \sigma_\tau \sigma_{\tau+\gamma} \rangle$$

between the two neighboring spins σ_0, σ_δ with interaction energy J_δ (say) and the two neighboring spins σ_τ, $\sigma_{\tau+\gamma}$ with interaction energy J_γ, where δ and γ are lattice vectors. In every case two "bonds" of the dimer lattice must be "perturbed," so the correlations are 4×4 Pfaffians. Three specific cases are discussed below.

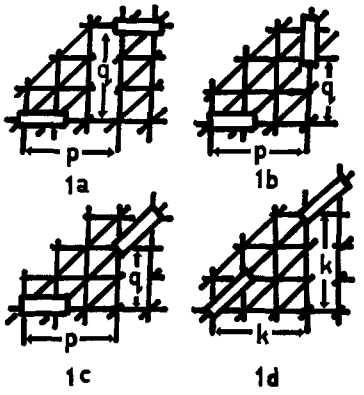


FIG. 1. Dimer configurations corresponding to fourth-order correlations. The complete hexagon decorating each lattice site has been omitted (for clarity).

(i) $\langle \sigma_{0,0}\sigma_{1,0}\sigma_{p,q}\sigma_{p+1,q} \rangle$

Referring to Fig. 1(a), one sees that this correlation concerns two horizontal bonds. The Pfaffian for the correlation may be written down immediately,

$$\langle \sigma_{0,0}\sigma_{1,0}\sigma_{p,q}\sigma_{p+1,q} \rangle = \begin{vmatrix} v_1 + (1-v_1^2)[1,0]_{RL} & (1-v_1^2)[p,q]_{RR} & (1-v_1^2)[p+1,q]_{RL} \\ & (1-v_1^2)[p-1,q]_{LR} & (1-v_1^2)[p,q]_{LL} \\ & & v_1 + (1-v_1^2)[1,0]_{RL} \end{vmatrix}. \quad (2.11)$$

Now the first term in this expression is just the nearest neighbor pair correlation

$$\omega_2(1,0) = \langle \sigma_{0,0}\sigma_{1,0} \rangle = v_1 + (1-v_1^2)[1,0]_{RL}. \quad (2.12)$$

So the reduced correlation function, defined in Eq. (1.2), may be written

$$\omega_4(0,0;1,0;p,q;p+1,q) = (1-v_1^2)^2 \times \{[p+1,q]_{RL}[p-1,q]_{LR} - [p,q]_{RR}[p,q]_{LL}\}. \quad (2.13)$$

The corresponding correlation $\langle \sigma_{0,0}\sigma_{0,1}\sigma_{p,q}\sigma_{p,q+1} \rangle$ which involves two vertical bonds can be obtained by interchanging v_1 with v_2, and RL with UD; and similarly for the diagonal correlation

$$\langle \sigma_{0,0}\sigma_{1,1}\sigma_{p,q}\sigma_{p+1,q+1} \rangle.$$

(ii) $\langle \sigma_{0,0}\sigma_{1,0}\sigma_{p,q}\sigma_{p,q+1} \rangle$

Referring to Fig. 1(b), one sees that this correlation concerns one horizontal bond and one vertical bond. Proceeding as in (i) above, one gets

$$\langle \sigma_{0,0}\sigma_{1,0}\sigma_{p,q}\sigma_{p,q+1} \rangle = \{v_1 + (1-v_1^2)[1,0]_{RL}\}\{v_2 + (1-v_2^2)[0,1]_{UD}\} + (1-v_1^2)(1-v_2^2)\{[p-1,q]_{LU}[p,q+1]_{RD} - [p,q]_{RU}[p-1,q+1]_{LD}\}. \quad (2.14)$$

Again the first term in this expression contains nn pair correlations $\langle \sigma_{0,0}\sigma_{1,0} \rangle$ and $\langle \sigma_{0,0}\sigma_{0,1} \rangle$. The reduced correlation function is

$$\omega_4(0,0;1,0;p,q;p,q+1) = (1-v_1^2)(1-v_2^2)\{[p-1,q]_{LU}[p,q+1]_{RD} - [p,q]_{RU}[p-1,q+1]_{LD}\}. \quad (2.15)$$

(iii) $\langle \sigma_{0,0}\sigma_{1,0}\sigma_{p,q}\sigma_{p+1,q+1} \rangle$

Referring to Fig. 1(c), one sees that this correlation concerns one horizontal and one diagonal bond. It may be evaluated along the same lines as above.

These three cases cover the possible combinations of two nn pairs of spins on a triangular lattice—apart from geometrically equivalent configurations, which can be obtained by appropriate interchange of $v_1 v_2$, v_3 and R, L; U, D; S, T.

3. SPECIFIC HEAT

The perturbation method employed to calculate the correlations is capable of handling any fourth or higher even order correlation. However, the particular correlations calculated in the previous section arise in a natural way in an evaluation of the specific heat of an Ising lattice. In the following, the inter-

action energies are set equal to shorten the various expressions, but there is no difficulty in principle in carrying through the analysis for the general case. It is easy to show that the internal energy of an Ising lattice is

$$U_N = -J \sum_{k,l} \langle \sigma_k \sigma_l \rangle, \tag{3.1}$$

and that the specific heat is

$$C_N = \frac{J^2}{kT^2} \{ \sum_{k,l} \sum_{m,n} \langle \sigma_k \sigma_l \sigma_m \sigma_n \rangle - (\sum_{k,l} \langle \sigma_k \sigma_l \rangle)^2 \}, \tag{3.2}$$

where k, l and m, n are neighboring pairs of spins. For an isotropic asymptotically large lattice of coordination number q, the sum over nn spins k, l reduces to multiplication by $\frac{1}{2}qN$, and so (3.1) reduces to

$$U_N = -\tfrac{1}{2} qNJ \langle \sigma_0 \sigma_1 \rangle, \tag{3.3}$$

and (3.2) reduces to

$$C_N = \frac{J^2}{kT^2} \{ \tfrac{1}{2} qN \sum_{m,n} \langle \sigma_0 \sigma_1 \sigma_m \sigma_n \rangle - (\tfrac{1}{2} qN \langle \sigma_0 \sigma_1 \rangle)^2 \}. \tag{3.4}$$

The remaining sum is over all fourth-order correlations between two sets of neighboring pairs of spins σ_0, σ_1 and σ_m, σ_n, and in addition contains those pair correlations for which the spins σ_m or σ_n overlap σ_0 or σ_1, and a single $+1$ arising from the complete overlap: $\langle \sigma_0^2 \sigma_1^2 \rangle = 1$. Now introducing the reduced correlation functions ω_4, one sees that the pair correlation terms in (3.4), which are of order N^2, cancel out, leaving

$$C_N = \frac{\frac{1}{2} qNJ^2}{kT^2} \sum_{\mathbf{r}} \omega_4(0, 0; 1, 0; \mathbf{r}; \mathbf{r} + \boldsymbol{\gamma}). \tag{3.5}$$

This expression must be modified for more general interaction energies to

$$C_N = \frac{N}{kT^2} \sum_{\mathbf{r}} \sum_{\boldsymbol{\delta}} J_\delta J_\gamma \omega_4(0, \boldsymbol{\delta}; \mathbf{r}, \mathbf{r} + \boldsymbol{\gamma}), \tag{3.6}$$

where the sum over $\boldsymbol{\delta}$ incorporates the points (1,0), (1,1), and (0,1) in turn.

The calculations in the previous section provide all the reduced fourth-order correlation functions, and also those pair correlations contributing to the sum, excluding, however, the term when the four spins overlap completely, which must be added on separately: it is just unity. The explicit summation of the correlation functions is not attempted here.

4. DIAGONAL CORRELATIONS

The correlation $\langle \sigma_{0,0} \sigma_{1,1} \sigma_{k,k} \sigma_{k+1,k+1} \rangle$ concerns two diagonal bonds: Fig. 1(d). It is equivalent to a special case of the correlations derived in Sec. 2(i), with $q = 0$ in Eq. (2.11), and the lattice axes renamed. So one obtains

$$\begin{aligned} \langle \sigma_{0,0} \sigma_{1,1} \sigma_{k,k} \sigma_{k+1,k+1} \rangle &= \{ v_3 + (1 - v_3^2)[1, 1]_{ST} \}^2 \\ &+ (1 - v_3^2)^2 \{ [k + 1, k + 1]_{ST} [k - 1, k - 1]_{TS} \\ &- [k, k]_{SS} [k, k]_{TT} \}, \end{aligned} \tag{4.1}$$

where ST replaces LR, v_3 replaces v_1, and J_3 replaces J_1 in (2.11) for the diagonal direction. Now the first term is just the square of the pair correlation $\langle \sigma_{0,0} \sigma_{1,1} \rangle$, and the third term vanishes by virtue of the relations in Eq. (2.10). The second term can be rewritten by use of the antisymmetry of the matrix elements, Eq. (2.9), which gives

$$[k - 1, k - 1]_{TS} = -[-k + 1, -k + 1]_{ST}.$$

Thus one obtains

$$\begin{aligned} \langle \sigma_{0,0} \sigma_{1,1} \sigma_{k,k} \sigma_{k+1,k+1} \rangle &= \langle \sigma_{0,0} \sigma_{1,1} \rangle^2 - (1 - v_3^2)^2 \\ &\times [k + 1, k + 1]_{ST} [-k + 1, -k + 1]_{ST}. \end{aligned} \tag{4.2}$$

The remaining matrix elements are the same as those required for the calculation of the general diagonal *pair* correlation $\langle \sigma_{0,0} \sigma_{k,k} \rangle$, and employing the definition

$$a_n = v_3 \, \delta_{0n} + (1 - v_3^2)[n + 1, n + 1]_{ST}, \tag{4.3}$$

we may quote the result from I [Ref. 3, Eq. (6.12)]

$$a_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-in\omega} \frac{(a - be^{+i\omega} - ce^{-i\omega})}{|a - be^{-i\omega} - ce^{+i\omega}|} \, d\omega, \tag{4.4}$$

where

$$\begin{aligned} a &= 2v_3(1 + v_1^2)(1 + v_2^2) + 4v_1 v_2 (1 + v_3^2), \\ b &= v_3^2 c = v_3^2 (1 - v_1^2)(1 - v_2^2). \end{aligned} \tag{4.5}$$

In terms of the integrals a_n, the correlation becomes

$$\langle \sigma_{0,0} \sigma_{1,1} \sigma_{k,k} \sigma_{k+1,k+1} \rangle = a_0^2 - a_k a_{-k}, \quad k \neq 0, \tag{4.6}$$

where the case $k = 0$, corresponding to complete overlap of the four spins, is excluded. The a_n may be written in the equivalent form, on substituting explicit values for a, b, and c:

$$\begin{aligned} a_n &= \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-in\omega} \\ &\times \frac{(C_1 C_2 S_3 + S_1 S_2 C_3 - C_3 \cos\omega + i \sin\omega)}{|C_1 C_2 S_3 + S_1 S_2 C_3 - C_3 \cos\omega + i \sin\omega|} \, d\omega, \end{aligned} \tag{4.7}$$

and again in the form given by Green and Hurst[1(o)]:

$$a_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-in\omega} \frac{(1 - Ae^{-i\omega})(1 - Be^{+i\omega})}{|(1 - Ae^{+i\omega})(1 - Be^{-i\omega})|} \, d\omega, \tag{4.8}$$

where, choosing $|A| > |B|$,

$$(1 + AB)/(A - B) = C_1 C_2 S_3 + S_1 S_2 C_3 \tag{4.9}$$

and $(A + B)/(A - B) = C_3$. This last form for the a_n is convenient for discussing the quadratic lattice and ferromagnetic triangular lattice, in which cases A and B are real. For the antiferromagnetic triangular lattice A and B are complex, and it is simpler to retain the v_i variables as in Eq. (4.4).

It is important to elucidate the temperature dependence of the parameters A and B. This is most easily done in terms of the *dual* and *inversion* transformation variables.[1,8] Suppose a triangular lattice $\triangle$ with interaction parameter K_3 in one direction is at a low temperature (low). Then it is related to an identical triangular lattice $\triangledown$ with interaction parameter K_3^+ at a high temperature (high) by the *inversion* transformation:

$$e^{-4K_3^+} = \frac{(v_1 + v_2v_3)(v_2 + v_3v_1)}{(v_3 + v_1v_2)(1 + v_1v_2v_3)}, + \text{cyclic}, \quad (4.10)$$

where $v_i = \tanh K_i$, $i = 1, 2, 3$, as before. The triangular lattice $\triangle$ (low) is also related to its dual honeycomb lattice (high) with interaction parameter K_3^* by the *dual* transformation

$$e^{-2K_3^*} = v_3 = \tanh K_3, + \text{cyclic}, \quad (4.11)$$

and similarly the triangular lattice $\triangledown$ (high) is related to its dual honeycomb lattice (low) by the *dual* transformation

$$e^{-2K_3^{+*}} = \tanh K_3^+, + \text{cyclic}. \quad (4.12)$$

The critical point for the triangular lattice occurs when $K_3^+ = K_3$ or $K_3^{+*} = K_3^*$. Now it may easily be confirmed that

$$\cosh 2K_3^{+*} = \coth 2K_3^+ = \frac{C_1C_2S_3 + S_1S_2C_3}{S_3}, \quad (4.13)$$

whence from the definition of the parameters A and B in Eq. (4.9) one obtains for the triangular lattice

$$A = e^{2(K_3^* - K_3^{+*})}, \qquad B = e^{-2(K_3^* + K_3^{+*})}. \quad (4.14)$$

Therefore the temperature dependence of A and B for the ferromagnetic lattice is, in summary,

$$\begin{aligned} &T < T_c : B < A < 1, \\ &T = T_c : A = 1, \qquad B = (\tanh K_3)^2, \\ &T > T_c : B < A^{-1} < 1. \end{aligned} \quad (4.15)$$

Thus the a_n are seen to be coefficients of $e^{in\omega}$ in the expansion of the generating function

$$A(\omega) = \left[\left(\frac{1 - Ae^{-i\omega}}{1 - Ae^{+i\omega}}\right)\left(\frac{1 - Be^{+i\omega}}{1 - Be^{-i\omega}}\right)\right]^{\frac{1}{2}} \quad (4.16)$$

below the critical point, and correspondingly coefficients of $e^{in\omega}$ in

$$A'(\omega) = -e^{-i\omega}\left[\left(\frac{1 - A^{-1}e^{+i\omega}}{1 - A^{-1}e^{-i\omega}}\right)\left(\frac{1 - Be^{+i\omega}}{1 - Be^{-i\omega}}\right)\right]^{\frac{1}{2}} \quad (4.17)$$

above the critical point. At the critical point itself, $A = 1$ and $B = (\tanh K_3)^2$, and the a_n are given by

$$a_n = \frac{1}{\pi}\int_0^\pi \frac{\sin(n + \frac{1}{2})\omega - B\sin(n - \frac{1}{2})\omega}{(1 - 2B\cos\omega + B^2)^{\frac{1}{2}}}\, d\omega. \quad (4.18)$$

5. ASYMPTOTIC BEHAVIOR

In order to discuss the behavior of the reduced fourth-order correlations [Eq. (4.6)] for large spin separation, we require a knowledge of the asymptotic behavior of the integrals a_n for large integer n. At the critical point, the integral for a_n in Eq. (4.18) has the asymptotic expansion

$$a_n \simeq \frac{n - \frac{1}{2}\cosh 2K_3}{\pi(n^2 - \frac{1}{4})} + O\left(\frac{1}{n^3}\right), \quad (5.1)$$

so that the leading term is $(\pi n)^{-1}$ for all (ferromagnetic) lattices. At temperatures well removed from the critical point, B may be neglected in comparison with A, or A^{-1} as the case may be, in the expansion of the generating functions. To this (good) approximation one may just pick out the coefficient of $e^{in\omega}$ *as if* B were zero. Further, to the same approximation only one term in A, or A^{-1}, need be retained in this coefficient, since higher powers of A, or A^{-1}, will make a negligible contribution. Note that in these asymptotic expansions the temperature variable A is held fixed, while the spin separation is increased.

Below the critical point, $AB \ll B \ll A \ll 1$, the coefficient of $e^{in\omega}$ in $A(\omega)$ is asymptotically

$$a_n \simeq \frac{\Gamma(n + \frac{1}{2})}{\Gamma(\frac{1}{2})\Gamma(n + 1)} A^n \simeq \frac{A^n}{\pi^{\frac{1}{2}}n^{\frac{1}{2}}}, \qquad n > 0, \quad (5.2)$$

and the coefficient of $e^{-in\omega}$ is

$$a_{-n} \simeq \frac{-\Gamma(n - \frac{1}{2})}{2\Gamma(\frac{1}{2})\Gamma(n + 1)} A^n \simeq \frac{-A^n}{2\pi^{\frac{1}{2}}n^{\frac{1}{2}}}, \quad n > 0. \quad (5.3)$$

It is convenient to retain $n > 0$, a positive integer, in all the asymptotic formulas, and to distinguish a_n and a_{-n}.

Above the critical point, $BA^{-1} \ll B \ll A^{-1} \ll 1$, the coefficients of $e^{in\omega}$ in $A'(\omega)$ is asymptotically

$$a'_n \simeq \frac{\Gamma(n + \frac{1}{2})}{2\Gamma(\frac{1}{2})\Gamma(n + 2)} A^{-n-1} \simeq \frac{A^{-n-1}}{2\pi^{\frac{1}{2}}n^{\frac{1}{2}}}, \quad n > 0, \quad (5.4)$$

and the coefficient of $e^{-in\omega}$ is

$$a'_{-n} \simeq \frac{-\Gamma(n - \frac{1}{2})}{\Gamma(\frac{1}{2})\Gamma(n)} A^{-n+1} \simeq \frac{-A^{-n+1}}{\pi^{\frac{1}{2}}n^{\frac{1}{2}}}, \quad n > 0. \quad (5.5)$$

[8] I. Syozi and S. Naya, Progr. Theor. Phys. 24, 829 (1960).

Primes are used to distinguish coefficients in high-temperature generating functions.

We are now in a position to discuss the asymptotic behavior of the diagonal reduced fourth-order correlation function, Eq. (4.6):

$$\omega_4(0, 0; 1, 1; k, k; k+1, k+1) = -a_k a_{-k} \equiv \omega_4(k, k), \quad \text{say.} \tag{5.6}$$

At the critical point, the asymptotic form of the a_k is given by Eq. (5.1), so that

$$\omega_4(k, k) \simeq 1/(\pi^2 k^2) \tag{5.7}$$

as the spin separation tends to infinity. The amplitude of the decay is $1/\pi^2$. At low temperatures, the asymptotic forms of a_k and a_{-k} are given by Eqs. (5.2) and (5.3), so that

$$\omega_4(k, k) \simeq A^{2k}/2\pi k^2, \tag{5.8}$$

and at high temperatures, from Eqs. (5.4) and (5.5):

$$\omega_4(k, k) \simeq A^{-2k}/2\pi k^2, \tag{5.9}$$

thus the results for temperatures well removed from the critical point may be combined into the single formula:

$$\omega_4(k, k) \simeq e^{-4\,|K_3^{*}-K_3^{+*}|\,k}/2\pi k^2. \tag{5.10}$$

These formulas have been derived for the ferromagnetic triangular lattice. The reduction to the quadratic lattice is considered in the next section. Here we note that the corresponding decay of the pair correlation $\omega_2(k, k) = \langle \sigma_{0,0}\sigma_{k,k} \rangle$ at *high* temperatures is

$$\omega_2(k, k) \simeq e^{-2(K_3^{*}-K_3^{+*})}/\pi^{\frac{1}{2}}k^{\frac{1}{2}}, \tag{5.11}$$

in which, by comparison with the Ornstein–Zernike theory, one may introduce the range of order β^{-1} so that

$$\omega_2(k, k) \propto e^{-\beta r}/r^{\frac{1}{2}}, \tag{5.12}$$

where r is a radial distance ($r = k$ for the triangular lattice). Thence the fourth-order correlation may be written

$$\omega_4(k, k) \simeq e^{-2\beta r}/2\pi r^2, \tag{5.13}$$

where $\beta = 2\,|K_3^{*} - K_3^{+*}|$. Although $\omega_4(k, k)$ decays by an "inverse square law" both at the critical point and at temperatures well removed from the critical point, the decay amplitude is different. This arises from the way in which the asymptotic formulas were obtained, the temperature parameter A being held fixed while the spin separation $r(=k)$ was increased. It is not permissible to let $T \to T_c$, or $\beta \to 0$ subsequently. A similar difficulty arises in the case of the pair correlations $\omega_2(k, k)$, which decay as $E/r^{\frac{1}{2}}$ at the critical point, so that the actual form of the correlation decay is different in this case.

Although we have so far only considered one particular relative orientation of bonds (spin pairs), and then only along a lattice axis, it is not unreasonable to assume that the correlations decay with a similar radial dependence in all directions (at least close to the critical point). The decay amplitude may now be direction dependent. Then, as indicated in the Introduction, the specific heat may be approximated as a sum over all reduced fourth-order correlations (neglecting the few pair correlations involved), and this sum replaced by an integral. Thus near the critical point

$$C_N \simeq \int_1^{\infty} e^{-2\beta r}\,\frac{dr}{r}, \tag{5.14}$$

which may be interpreted as representing a logarithmic divergence of C_N as $\log|T - T_c|$ at the critical point. The approach to the critical point by this means is not justifiable, because it involves reversing the limiting process used to obtain the asymptotic behavior of the correlations; and in any case the approximation $A \ll 1$ (or $A^{-1} \ll 1$) is invalid there.

At the critical point itself the integral for the specific heat is

$$C_N \simeq \frac{2}{\pi}\int_1^{d} \frac{dr}{r} \propto \log d, \tag{5.15}$$

where the upper limit has been restrained to a value d, representing the linear dimensions of the lattice, $d \simeq N^{\frac{1}{2}}$, which is really infinite in two dimensions. Such a logarithmic dependence of the "critical" value of the specific heat upon the size of the lattice was conjectured by Kramers and Wannier[5] on the basis of numerical calculations of C, which is finite for a finite lattice, but has a sharp maximum for a large lattice. It is this maximum value of C which diverges as $\log d$, suggesting that the specific heat is actually infinite there for an infinite lattice. This was confirmed rigorously by Onsager.[6]

6. QUADRATIC LATTICE

In this section the fourth-order correlations on the quadratic lattice are discussed. If in the triangular lattice calculations J_3 is set equal to zero, the formulas of the previous section now apply to diagonal correlations, between pairs of second nearest neighbor spins at lattice sites (0, 0), (1, 1), and (k, k), $(k+1, k+1)$. Also the corresponding reduced fourth-order correlation $\omega_4(k, k)$ does not now contribute to the specific heat, since the cor-

responding interaction energy J_3 is zero. Alternatively, one may set J_1 equal to zero, and then formally replace the subscript 3 by 1, thus converting the lattice into a quadratic one with parameters K_1 and K_2. The fourth-order correlation formulas of the previous sections now correspond to row 1 of the quadratic lattice and contribute to the specific heat.

Row correlations are considered first. Setting $J_1 = 0$, and replacing J_3 by (a new) J_1, one finds that

$$A = z_2/v_1 = e^{2(K_1^*-K_2)} \tag{6.1}$$

and

$$B = v_1 z_2 = e^{-2(K_1^*+K_2)},$$

where the asterisk denotes the dual variable:

$$v_i = \tanh K_i = e^{-2K_i^*}, \qquad i = 1, 2. \tag{6.2}$$

The new values of A and B satisfy Eq. (4.15) and so all the formulas of the previous section hold. The qualitative behavior of the ferromagnetic triangular and quadratic lattices is the same for row correlations. But the reciprocal range is of order $\beta = 2\,|K_1^* - K_2|$ in the row 1 direction. At temperatures well-removed from the critical point

$$\omega_4(k, 0) \doteq e^{-4\,|K_1^*-K_2|\,k}/2\pi k^2, \tag{6.3}$$

where the radial distance $r = k$ for row correlations.

To obtain the diagonal second nearest neighbor correlations, one sets $J_3 = 0$ in the formulas for A and B. This yields

$$A = (\sinh 2K_1 \sinh 2K_2)^{-1} \quad \text{and} \quad B = 0 \tag{6.4}$$

for the quadratic lattice with parameters K_1 and K_2. A and B clearly satisfy Eq. (4.15), for at the critical point $\sinh 2K_1 \sinh 2K_2 = 1$. The simplification provided by the absence of B means that a more sophisticated discussion of the integrals a_n, which are now denoted b_n to avoid confusion, is possible. It is evident that the formulas of the previous sections hold with the new values of A and B.

The integral representation of the b_n is

$$b_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-in\omega}\,\frac{(1 - Ae^{-i\omega})}{|1 - Ae^{-i\omega}|}\,d\omega. \tag{6.5}$$

The integral b_0 is now the second nearest neighbor diagonal correlation $\langle\sigma_{0,0}\sigma_{1,1}\rangle$, which is positive at all temperatures. There is a certain symmetry between the values of the b_n at low temperatures and the b_n' at high temperatures:

$$b_n' = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-in\omega}(-e^{-i\omega})\,\frac{(1 - A^{-1}e^{+i\omega})}{|1 - A^{-1}e^{+i\omega}|}\,d\omega. \tag{6.6}$$

Clearly, having calculated b_n at a *low* temperature given by $A < 1$, one may formally replace A by $A'(\equiv A^{-1})$ in b_n, thus obtaining the value of $-b'_{n-1}$ at a *high* temperature given by $A'(=A^{-1} > 1)$. The primes denote high-temperature functions as before. The integer n is not restricted to positive values here. The converse relation holds between b_n' and b_{-n-1}. In particular $b_0 = -b'_{-1}$ and $b_0' = -b_{-1}$. Thus we need consider b_n for only positive values of n, since on replacing A by $A'(=A^{-1})$ we obtain

$$b_n \to -b'_{n-1} \quad \text{and} \quad b_n' \to -b_{-n-1}. \tag{6.7}$$

This symmetry is related to the dual transformation which may be written in the form

$$\sinh 2K_i \sinh 2K_i^* = 1, \qquad i = 1, 2.$$

The integrals b_n may be expressed in terms of Legendre functions of the second kind of half-integer order:

$$b_n = \pi^{-1}[A^{-\frac{1}{2}}Q_{n-\frac{1}{2}}(x) - A^{\frac{1}{2}}Q_{n+\frac{1}{2}}(x)], \tag{6.8}$$

where $x = \frac{1}{2}(A + A^{-1})$, whence the symmetry in A and A^{-1} follows immediately. For the present purpose an alternative representation of the b_n as hypergeometric functions is preferable[9] (also see Appendix).

At low temperatures, $A < 1$,

$$b_n = \frac{1}{\pi^{\frac{1}{2}}}\frac{\Gamma(n+\frac{1}{2})}{\Gamma(n+1)}A^n\Big\{F(\tfrac{1}{2}, n+\tfrac{1}{2}; n+1; A^2) - \frac{(n+\frac{1}{2})}{(n+1)}A^2F(\tfrac{1}{2}, n+\tfrac{3}{2}; n+2; A^2)\Big\}, \tag{6.9}$$

and at high temperatures, $A > 1$,

$$b_n' = \frac{1}{\pi^{\frac{1}{2}}}\frac{\Gamma(n+\frac{1}{2})}{\Gamma(n+1)}\frac{1}{A^{n+1}}\Big\{F(\tfrac{1}{2}, n+\tfrac{1}{2}; n+1; 1/A^2) - \frac{(n+\frac{1}{2})}{(n+1)}F(\tfrac{1}{2}, n+\tfrac{3}{2}; n+2; 1/A^2)\Big\}. \tag{6.10}$$

At the critical point $A = 1$, the explicit value of the b_n may be obtained directly from Eq. (4.18); it is

$$b_n = 1/\pi(n + \tfrac{1}{2}) = -b_{-n-1}. \tag{6.11}$$

Thus the exact value of the fourth-order correlation,

$$\langle\sigma_{0,0}\sigma_{1,1}\sigma_{k,k}\sigma_{k+1,k+1}\rangle = \langle\sigma_{0,0}\sigma_{1,1}\rangle^2 - b_k b_{-k}, \tag{6.12}$$

may be obtained on substitution of the appropriate value for b_k above. At the critical point the reduced fourth-order correlation is

[9] Bateman Manuscript Project, *Higher Transcendental Functions*, A. Erdelyi, Ed. (McGraw-Hill Book Company, Inc., New York, 1953), Vol. II, p. 137, Eq. (45).

$$\omega_4(k, k) = 1/\pi^2(k^2 - \tfrac{1}{4}), \qquad (6.13)$$

which is asymptotically similar to Eq. (5.7). For temperatures well removed from the critical point the asymptotic behavior of b_n, $n > 0$, may be deduced from the representations (6.9) and (6.10). Keeping A fixed, one obtains the asymptotic form of b_n from the series expansion of the hypergeometric functions. The leading terms are of course similar to those of the a_n in Eqs. (5.2) and (5.4). The coefficients of $1/n^{\frac{1}{2}}$ and $1/n^{\frac{3}{2}}$, as appropriate, may be extracted from the hypergeometric functions. This process yields

$$b_n \doteq (A^n/\pi^{\frac{1}{2}}n^{\frac{1}{2}})(1 - A^2)^{\frac{1}{4}}, \qquad (6.14)$$

at low temperatures, $A < 1$, and

$$b'_n \doteq (A^{-n-1}/2\pi^{\frac{1}{2}}n^{\frac{1}{2}})(1 - A^{-2})^{-\frac{1}{4}}, \qquad (6.15)$$

at high temperatures, $A > 1$, where n is a positive integer. An alternative derivation of these formulas is given in the Appendix. Now by using the symmetry property of the b_n in Eq. (6.7), one may replace A by $A' = A^{-1}$, and hence obtain the behavior of the b_{-n}:

$$b_{-n} \doteq (-A^n/2\pi^{\frac{1}{2}}n^{\frac{1}{2}})(1 - A^2)^{-\frac{1}{4}}, \qquad (6.16)$$

at low temperatures, $A < 1$, and

$$b'_{-n} \doteq (-A^{-n+1}/\pi^{\frac{1}{2}}n^{\frac{1}{2}})(1 - A^{-2})^{\frac{1}{4}}, \qquad (6.17)$$

at high temperatures, $A > 1$. It is now apparent how the asymptotic expansions break down as $A \to 1^-$ from below T_c or $A \to 1^+$ from above T_c: which limiting processes are forbidden by the method of constructing the asymptotic expansions in the first place. n is positive in all these formulas. But when the reduced fourth-order correlation $\omega_4(k, k)$ is formed one obtains

$$\omega_4(k, k) \doteq A^{2k}/2\pi k^2, \qquad (6.18)$$

below the critical point and

$$\omega_4(k, k) \doteq A^{-2k}/2\pi k^2, \qquad (6.19)$$

above it, the factors $(1 - A^{\pm 2})^{\pm\frac{1}{4}}$ cancelling. A is now given by Eq. (6.4). Thus the final result is deceptive in that it appears one could take the limits $A^{\pm 1} \to 1$ satisfactorily, but this is not the case, since the asymptotic forms of the b_n (or a_n) being combined is radically different for temperatures below, at, or above the critical point.

For the diagonal direction on the quadratic lattice $r = \sqrt{2}k$, the corresponding high-temperature form of the pair correlation $\omega_2(k, k)$ is

$$\omega_2(k, k) \doteq A^k/\pi^{\frac{1}{2}}k^{\frac{1}{2}} \propto e^{-\beta r}/r^{\frac{1}{2}}, \qquad (6.20)$$

where the reciprocal range of order β is now

$$\beta = (1/\sqrt{2})\,|\ln(\sinh 2K_1 \sinh 2K_2)|. \qquad (6.21)$$

In terms of β and the radial distance $r = \sqrt{2}k$,

$$\omega_4(k, k) \doteq e^{-2\beta r}/\pi r^2. \qquad (6.22)$$

Thus the fourth-order correlations on the quadratic lattice are not radially symmetric, and just as for the high-temperature pair correlations there is a preference for order to be propagated along the lattice axes [see Ref. 1(a), p. 202]. However, the diagonal correlations discussed here do not contribute to the specific heat.

7. ANTIFERROMAGNETIC LATTICES

The quadratic lattice can be divided into two interpenetrating sublattices. At low temperatures an ordered structure is possible, the spins on the two sublattices pointing in opposite directions, and the ground state is twofold degenerate in zero magnetic field. The critical point is the same as for the ferromagnetic lattice, and the specific heat is logarithmically infinite there. Therefore one would expect the fourth-order correlations to be unaltered on replacing J_i by $-J_i$, $i = 1, 2$. This is evident on inspection of the formula defining A and B for this lattice. For row correlations A is replaced by $1/(v_1z_2)$ and B by v_1/z_2, since $z_i = e^{-2K_i} > 1$. A and B are now negative since $v_i = \tanh K_i$ is. At the critical point $A = -1$ and $B = -(\tanh K_1)^2$ for row 1 correlations. A and B are unaltered for diagonal correlations. Thus the formulas above for the fourth-order correlations are still valid for the completely antiferromagnetic quadratic lattice. However, the behavior of the pair correlations is affected. The row pair correlations oscillate in sign: $(-)^k$; but the diagonal pair correlations remain positive. A redefinition of β appropriate to the antiferromagnetic case is easily made.

More interesting is the antiferromagnetic triangular lattice, with three equal interactions. There is no ordering at the antiferromagnetic zero point, $T = -0$, and the zero point entropy is finite. Also there is no critical point in the usual sense, though $T = -0$ is certainly critical mathematically. At the zero point it is easy to evaluate the integrals a_n explicitly from Eq. (4.4) with $v_i = v = -1$ [see also Eq. (7.8) below]. Then one obtains

$$a_0 = -\tfrac{1}{3}, \qquad a_n = a_{-n} = -\frac{2\sin(2\pi n/3)}{\pi n} \qquad (7.1)$$

or a_n equals

$$-\tfrac{1}{3}, \text{ when } n = 0; \quad 0, \text{ when } n = 3l;$$
$$\sqrt{3}/\pi n, \text{ when } n = 3l - 1;$$
$$-\sqrt{3}/\pi n, \text{ when } n = 3l - 2; \qquad (7.2)$$

where l is a positive integer. This periodicity is connected with the three sublattices into which the triangular lattice may be divided.

a_0 is related to the internal energy which is

$$U_N(T = -0) = -N|J| = 3Na_0. \qquad (7.3)$$

The specific heat exhibits no anomaly and is a smooth function of temperature, being zero at infinite and zero temperatures. The reduced fourth-order correlations are exactly

$$\omega_4(k, k) = -a_k a_{-k} = -\left[\frac{2 \sin (2\pi k/3)}{\pi k}\right]^2 \qquad (7.4)$$

at the antiferromagnetic zero point. They vanish when the spins $\sigma_{0,0}$ and $\sigma_{k,k}$ are on the same sublattice, and are negative and equal to $-3/(\pi k)^2$ on the other two sublattices. The characteristic inverse square law decay is retained, with a superposed oscillation related to the periodic arrangement of the sublattices. It is interesting to compare the corresponding asymptotic behavior of the pair correlations at the zero point:

$$\omega_2(k, k) \doteq \epsilon_0 \cos (2\pi k/3)/k^{\frac{1}{2}}, \qquad (7.5)$$

which oscillate in magnitude and sign.[3]

Above the zero point the generating function for the integrals a_n may be manipulated into the form

$$A''(\omega) = -e^{-i\omega}\left(\frac{1 - 2v \cos\theta\, e^{+i\omega} + v^2 e^{+2i\omega}}{1 - 2v \cos\theta\, e^{-i\omega} + v^2 e^{-2i\omega}}\right)^{\frac{1}{2}}, \qquad (7.6)$$

where θ is a real angle defined by

$$\cos\theta = \tfrac{1}{2}(1 + e^{2K}), \qquad (7.7)$$

K being negative for an antiferromagnet. $A''(\omega)$ has zeros at $e^{+i\omega} = e^{\pm i\theta}/v$ and poles at $e^{-i\omega} = e^{\pm i\theta}/v$, which occur at complex values of ω except when $v = -1$ ($T = -0$, $\cos\theta = \frac{1}{2}$). Then these points coalesce, and $A''(\omega)$ must be rewritten:

$$A''(\omega) = -\frac{(1 + 2\cos\omega)}{|1 + 2\cos\omega|}, \qquad (7.8)$$

from which the zero point values of a_n in Eq. (7.1) may be derived. At very high temperatures v is small and negative, so that terms of order v^{r+1} may be neglected compared with v^r. To the same approximation $\cos\theta = 1$, and the generating function becomes

$$A''(\omega) \doteq -e^{-i\omega}\left(\frac{1 - 2ve^{+i\omega}}{1 - 2ve^{-i\omega}}\right)^{\frac{1}{2}}, \qquad (7.9)$$

which is similar to the high-temperature generating function $A'(\omega)$ in Eq. (4.17), with $B = 0$, for the ferromagnetic case. Therefore the asymptotic form of the a_n given by Eqs. (5.4) and (5.5) may be taken over to yield

$$a_n'' \doteq \frac{(2v)^{n+1}}{2\pi^{\frac{1}{2}} n^{\frac{3}{2}}} \text{ and } a_{-n}'' \doteq -\frac{(2v)^{n-1}}{\pi^{\frac{1}{2}} n^{\frac{3}{2}}}, \quad n > 0, \qquad (7.10)$$

where v is negative. Correspondingly, the reduced fourth-order correlation has the asymptotic form

$$\omega_4(k, k) \doteq (2v)^{2k}/2\pi k^2. \qquad (7.11)$$

So the antiferromagnetic triangular lattice is qualitatively similar to an ordinary antiferromagnetic lattice at very high temperatures, with A^{-1} replaced by $2v$. The necessary restriction $|v| < \frac{1}{2}$ is roughly equivalent to $T > \frac{1}{2}T_c$, where T_c is the ferromagnetic critical temperature.

CONCLUDING REMARKS

This paper contains a fairly complete discussion of fourth-order correlations for one particular relative orientation and one particular direction of spin separation (when the spins are collinear). This leaves a lot of other directions and orientations. An analysis of the general fourth-order correlation which contributes to the specific heat should provide information about the angular dependence of the decay amplitude, both at and well away from the critical point. Thus one would expect to obtain the value of the constant B' in Eq. (1.6), relating to the divergence of the specific heat at the critical point (it is already known for the quadratic lattice[6]).

The calculation of the particular fourth-order correlations discussed in this paper has involved the perturbation of two lattice bonds, corresponding to the selection of two pairs of nearest-neighbor spins. Such a calculation closely parallels that required for an analysis of the effects of two defect bonds. Much of the working here is preparatory for this problem. To avoid congestion here, details have been reserved for another communication.

Finally, it should be remembered that a knowledge of *all* possible second- and fourth-order correlations would enable one in principle to obtain the coefficients of $(H/kT)^2$ and $(H/kT)^4$ in an expansion of the partition function of an Ising lattice in a magnetic field H, H/kT being the "high-temperature" expansion variable. However, the exact derivation of these coefficients still appears to be an intractable problem, as approached via calculations of spin correlations.

ACKNOWLEDGMENTS

I am grateful to Professor M. E. Fisher for initially interesting me in this problem by drawing my attention to the importance of these fourth-order correlations because of their relation to the specific heat. I would also like to thank Professor H. S. Green and Professor C. A. Hurst for stimulating discussions concerning this and related problems in the Ising model.

APPENDIX

The Appendix contains an alternative representation for the b_n which leads to a rederivation of Eqs. (6.14)–(6.17), and also a brief account of the behavior of the b_n and the correlations near the critical point.

An alternative representation for the b_n, defined in Eq. (6.8), for the positive n when $T < T_c$ is[10]

$$b_n = \frac{1}{\pi^{\frac{1}{2}}} \frac{\Gamma(n+\frac{1}{2})}{\Gamma(n+1)} \frac{A^n}{(1-A^2)^{\frac{1}{4}}} \left\{ F(\tfrac{1}{2}, \tfrac{1}{2}; n+1; \zeta) - A^2 \left(\frac{n+\frac{1}{2}}{n+1} \right) F(\tfrac{1}{2}, \tfrac{1}{2}; n+2; \zeta) \right\}, \tag{A1}$$

where the argument ζ of the hypergeometric functions is

$$\zeta = -A^2/(1-A^2), \tag{A2}$$

$|\zeta|$ being less than unity provided $A^2 < \frac{1}{2}$. This representation is therefore valid for low temperatures such that $0 < A^2 < \frac{1}{2}$, but could be analytically continued into the region $\frac{1}{2} < A < 1$, since the b_n are well-behaved functions of A. Equation (6.9) would be such an analytic continuation. However, regarded as a series in inverse powers of n, which would be of an asymptotic nature for $\frac{1}{2} < A^2 < 1$, the above representation for b_n has the leading term

$$b_n \simeq \frac{1}{\pi^{\frac{1}{2}}} \frac{\Gamma(n+\frac{1}{2})}{\Gamma(n+1)} A^n (1-A^2)^{\frac{1}{4}}, \tag{A3}$$

in which the parameter A is held fixed, while n is permitted to increase. This provides an alternative, and illuminating, derivation of Eq. (6.14), while confirming the asymptotic nature of this expansion. Above the critical point, the representation corresponding to Eq. (A1) for b_n' is

$$b_n' = \frac{1}{\pi^{\frac{1}{2}}} \frac{\Gamma(n+\frac{1}{2})}{\Gamma(n+1)} \frac{A^{-n-1}}{(1-A^{-2})^{\frac{1}{4}}} \left\{ F(\tfrac{1}{2}, \tfrac{1}{2}; n+1; \zeta) - \left(\frac{n+\frac{1}{2}}{n+1} \right) F(\tfrac{1}{2}, \tfrac{1}{2}; n+2; \zeta) \right\}, \tag{A4}$$

where $A > 1$ and ζ is defined by

$$\zeta = -A^{-2}/(1-A^{-2}), \tag{A5}$$

$|\zeta|$ being less than unity provided $A^{-2} < \frac{1}{2}$. The leading term in an asymptotic expansion in inverse powers of n is

$$b_n' \simeq \frac{1}{2n\pi^{\frac{1}{2}}} \frac{\Gamma(n+\frac{1}{2})}{\Gamma(n+1)} \frac{A^{-n-1}}{(1-A^{-2})^{\frac{1}{4}}}, \tag{A6}$$

confirming Eq. (6.15). The corresponding formulas for the b_{-n} are obtained by use of the relations in Eq. (6.7).

The behavior of the b_n in the neighborhood of the critical point $A = 1$ may be investigated with the aid of the appropriate analytic continuation of the hypergeometric functions in Eqs. (6.9) and (6.10),[11] or again directly from the Legendre function representation in Eq. (6.8). Taken individually, the hypergeometric functions (and the Legendre functions) have a logarithmic singularity at $A = 1$, the arguments of these functions approaching $A = 1$ from above or below as appropriate. The b_n, however, depend on the difference between two of these divergent functions and remain finite for all temperatures. Thus for fixed n, as $A \to 1^-$,

$$b_n = \frac{1}{\pi(n+\frac{1}{2})} - \tfrac{1}{2}(1-A^2)\log(1-A^2) + O(1-A^2), \tag{A7}$$

and as $A \to 1^+$

$$b_n' = \frac{1}{\pi(n+\frac{1}{2})} + \tfrac{1}{2}(1-A^{-2})\log(1-A^{-2}) + O(1-A^{-2}). \tag{A8}$$

n is unrestricted in these formulas. Therefore the leading terms in the reduced fourth-order correlation $\omega_4(k, k)$ for fixed spin separation k, as $T \to T_c^-$, $A \to 1^-$ are

$$\omega_4(k,k) \simeq \frac{1}{\pi^2(k^2-\frac{1}{4})} \times [1 - \tfrac{1}{2}\pi(1-A^2)\log(1-A^2)], \tag{A9}$$

and as $T \to T_c^+$, $A \to 1^+$ are

$$\omega_4(k,k) \simeq \frac{1}{\pi^2(k^2-\frac{1}{4})} \times [1 + \tfrac{1}{2}\pi(1-A^{-2})\log(1-A^{-2})]. \tag{A10}$$

The symmetry of the correlation about the critical point is clearly displayed.

[10] Reference 9, Vol. II, p. 137, Eq. (44).

[11] Reference 9, Vol. II, p. 110, Eq. (12), with $m = 0$.

S-Matrix Formalism, Charge Renormalization, and Definition of the Hamiltonian in a Simple Field-Theoretical Model*

F. J. YNDURÁIN
University of Zaragoza, Zaragoza, Spain
(Received 29 November 1965)

A field-theoretical model (the relativistic version of the Lee model) is exhibited for which one can rigorously prove that the essential features of the S-matrix formalism (in particular analyticity) hold, even for point coupling, and that all charge renormalizations are finite. It is also shown that, although the interaction-picture description fails, the total Hamiltonian may still be defined as a symmetric operator and it governs the dynamics of the interaction. Finally, we point out that the usual pathologies which appear in the current (static) version of the model are not an intrinsic feature of the theory, but rather the result of too crude oversimplifications.

1. INTRODUCTION

THE difficulties inherent to the ordinary interaction picture approach to field theory have led many physicists to the belief that one should deal with the more directly significant S matrix,[1,2] then assuming that it is possible to describe all physical processes in terms of its properties. Thus, there have been conjectured suitable analytic, unitary, etc. properties of S, and assumed that they will hold and suffice to describe observables even if the Lagrangian–Hamiltonian formalism were meaningless.[2] One has, in this context, many positive results: Everything may be proved for potential scattering; and analytical properties are known to hold within certain regions (and, in general, for the first terms of a Feynman graph expansion). As a consequence, (single) dispersion relations have been proved starting from the fundamental axioms.[3]

However, one is still far from having totally clarified the problems involved; thus, e.g., much of the work done lacks mathematical rigor and, in addition, one knows that too much analyticity and unitarity are incompatible. There also exist, in another context, more difficulties: Although some people have put forward the hypothesis that the whole dynamics should be described through disperson relations,[4] it is still a rigorous consequence of the axioms that a unitary time-evolution operator (and thus a symmetric total Hamiltonian) must exist,[5] irrespectively of the failure of the interaction description.

The aim of the present article is to present a field-theoretical model for which the preceding analysis turns out to be essentially correct. To be precise, we want to show that, for the relativistic generalization of the Lee[6,7] model, and for the N–θ scattering, one has:

(1) The interaction picture fails for point coupling (no cutoff); the Møller operators cease to exist and thus so do the usual methods of the Hamiltonian formalism. However, we may still define a symmetric (Hermitian) total Hamiltonian and it governs the dynamics of the process.

(2) The S matrix exists and presents all the required (analytic, invariant, and unitary) properties, within the limitations of the model [see No. (4) below]. In particular, we may write a dispersion relation for it. It is also possible to introduce saturation[8] in a natural way; i.e., there exists a maximal value of the coupling constant compatible with unitarity.

(3) Taking into account the relativistic version of the model is by no means trivial (contrary to what it has sometimes been assumed[7,9]). Thus, we may show not only that the scattering does not vanish, but also that all charge renormalizations

* This work has been partially supported by the Fondo para Ayuda a la Investigación en la Universidad and Jen, Madrid.

[1] W. Heisenberg, Z. Phys. 65, 4 (1930); 120, 513, 673 (1943).

[2] G. F. Chew, *S-Matrix Theory of Strong Interactions* (W. A. Benjamin Company, Inc., New York, 1962).

[3] N. Bogolyoubov and D. V. Chirkov, *Théorie Quantique des Champs* (Dunod Cie., Paris, 1960). R. Omnes, *Relations de Dispersion*, (Dunod Cie., Paris, 1960).

[4] M. Gell-Mann, G. F. Chew, *et al.*; see, for instance, Ref. 2.

[5] J. M. Jauch, in *Dispersion Relations* (Oliver and Boyd, Edinburgh, 1961). F. J. Yndurain, Rev. Acad. Cienc. (Zaragoza) 20, 1 (1965).

[6] T. D. Lee, Phys. Rev. 95, 1329 (1954).

[7] S. Schweber, *Relativistic Quantum Field Theory* (Row, Peterson, and Company, Evanston, Illinois, 1961).

[8] G. F. Chew and S. C. Frautschi, Phys. Rev. 123, 1478 (1961).

[9] L. Van Hove, Physica 18, 145 (1952); *ibid.* 22, 343 (1956).

are finite. As a byproduct, the usual abnormalities,[10] in particular ghosts or indefinite metrics, need not appear,[11] this simply seems to be due to the oversimplification of neglecting the recoil of the V and N particles. All the results stated here and in No. (2) are to be understood as valid even for point coupling.

(4) The model, of course, presents some drawbacks. As is well known, the Lee model describes a nonlocal interaction and does not allow for antiparticles, so that, in particular, the S matrix lacks crossing symmetry. This is, however, a lucky circumstance (in a sense), since it shows that, at least in general, crossing symmetry is not a consequence of (in the sense of, e.g., Mandelstam[10]), but rather independent of the remaining properties of a theory (in particular, of the existence or nonexistence of ghosts).[10]

Finally, we note that all our results are mathematically rigorous.

2. FORM OF THE S-MATRIX

The model we consider is described by the Hamiltonians

$$H_0 = \int d^3p\, E_V(p) V_p^* V_p + \int d^3p\, E_N(p) N_p^* N_p \int d^3k\, \omega_k a_k^* a_k, \tag{2.1a}$$

$$H = H_0 + \int d^3p\, \delta E_V(p)\, V_p^* V_p + \int d^3p\, d^3k\, [g(p,k) V_p^* N_{p-k} a_k + \text{h.c.}]\, \lambda, \tag{2.1b}$$

where λ is proportional to the charge (we borrow the notation from Ref. 7),

$$\delta E_V(p) = |\lambda|^2 \times \int d^3k\, |g(p,k)|^2/[E_N(p-k) + \omega_k - E_V(p)] \tag{2.2}$$

is the energy renormalization counterterm, and where

$$g(p,k) = f(p,k)[8E_V(p)E_N(p-k)\omega_k]^{-\frac{1}{2}}. \tag{2.3}$$

$f(p,k)$ is a weight (cutoff) function; for point coupling, $f = 1$.[7,9] The relativistic expressions

$$E_V(p) = (M^2 + p^2)^{\frac{1}{2}}, \qquad E_N(p) = (m^2 + p^2)^{\frac{1}{2}}, \qquad \omega_k = (\mu^2 + k^2)^{\frac{1}{2}} \tag{2.4}$$

are used for the energies of the V, N, and θ particles. In the usually considered version, (2.4) is approximated by[6,10]

$$E_V = M, \qquad E_N = m, \qquad \omega_k = (\mu^2 + k^2)^{\frac{1}{2}} \tag{2.5}$$

(fixed V and N particles) and the interaction is further restricted by the condition $f(p,k) = f(k) =$ square-integrable, lack of the last condition implying the appearance of pathologies. Our chief task here is, precisely, the removal in a consistent way, of the last condition. Surprisingly enough, use of the exact expressions (2.4) [rather than (2.5)] suffices to avoid the mentioned abnormalities as well as to ensure the correct properties for S. To prove this, we compute the R matrix. This may be done in the standard fashion (see, e.g., Ref. 7); the result is

$$\langle N_{q'} a_{k'} |\, R\, | N_q a_k \rangle = \delta(q + k - q' - k')\, \delta[E_N(q) + \omega_k - E_N(q') - \omega_{k'}] \frac{|\lambda|^2\, \overline{g(q+k,k)} g(q'+k',k')}{E_N(q) + \omega_k - E_V(q+k)} \times \left\{1 + \int \frac{d^3p\, |\lambda|^2\, |g(q+k,p)|^2}{[E_N(q) + \omega_k - E_N(q+k-p) - \omega_p + i\epsilon][E_V(q+k) - E_N(q+k-p) - \omega_p]}\right\}^{-1}. \tag{2.6}$$

This shows at once why, using (2.5), infinite renormalization is needed: for point coupling, the integrand in (2.6) behaves as $|p|^2\, (\omega_p\, |p|^2)^{-1} \sim 1/|p|$ so that the integral is divergent and $R \to 0$. However, if we take the exact expression (2.4) for the energies, the integrand is $\sim [E(p)\omega_p]^{-1} \sim 1/|p|^2$, which makes the integral finite and $R \neq 0$: there is no necessity for (infinite) charge renormalization.

From (2.6) we may write at once the scattering amplitude. In the c.m. system and for point coupling it is

$$A(s) = \frac{\pi\, |\lambda|^2}{2M(s^{\frac{1}{2}} - M)} \left\{1 + \int \frac{|\lambda|^2\, \pi\, |p|^2\, d\, |p|}{2ME_N(p)\omega_p[s^{\frac{1}{2}} - E_N(p) - \omega_p + i\epsilon][E_N(p) + \omega_p - M]}\right\}^{-1} \tag{2.7}$$

[10] W. Heisenberg, *Theorie der Elementärteilchen* (Munich, 1961). N. Bogoljubov, in *Physikalische Abhandlungen aus der Sowjetunion* (Akademie-Verlag, Berlin, 1958). K. W. Ford, Phys. Rev. **105**, 320 (1957). G. Källén and W. Pauli, Kgl. Dansk. Videnskab. Selskab Mat. Fys. Medd. **30** (1955). S. Mandelstam, Phys. Rev. **112**, 1344 (1958) (especially the Introduction).

[11] We refer, of course, to the $N = \theta$ scattering case. We are not be concerned with other (multiparticle) problems; see, for instance, F. J. Yndurain, Ref. 5.

with $s = [\omega_k + E_N(k)]^2$. It may be noted that A does not depend on the momentum transfer; this reflects the fact that only S waves interact [cf. (2.1b), (2.6)]. In particular, crossing is impossible, which is to be expected from (2.1b).

A close look at (2.7) convinces us that $A(s)$ posesses all required properties (within the limitations of the model). Actually, (i) it is unitary [check for $S = 1 + 2\pi i k s^{-\frac{1}{2}} A(s)$: $SS^* = S^*S = 1$]; (ii) it corresponds to point coupling; (iii) it is relativistically invariant, and (iv) it is analytic in the entire s-plane with the exception of a (kinematical) cut from $s = 0$ to $-\infty$ (due to the square root of s), a cut from $(m + \mu)^2$ to ∞, due to the opening of the scattering channel, and a pole at $s = M^2$ due to the (virtual) V-particle intermediate state. In particular, we may write a dispersion relation for A to read

$$\begin{aligned}\text{Re } A(s) = {} & \frac{|\lambda|^2}{s - M^2} \\ & + (1/\pi) \int_{(m+\mu)^2}^{\infty} ds' \text{ Im } A(s')/(s' - s) \\ & + (1/\pi) \int_0^{\infty} ds' \text{ Im } A(s')/(s - s'), \end{aligned} \tag{2.8}$$

where the difference with the usual dispersion relation is again due to the absence of crossing, and where

$$|\lambda_d|^2 = \frac{|\lambda|^2}{8M} \left\{ 1 + \frac{\pi |\lambda|^2}{2M} \times \int_0^{\infty} \frac{p^2\, dp}{E_N(p)\omega_p [E_N(p) - \omega_p - M]^2} \right\}^{-1} \tag{2.9}$$

is the residue of A at $s = M^2$.

3. RENORMALIZATION AND SATURATION PROPERTIES

We show later (Sec. 4) that, for point coupling, the interaction picture fails, and thus the "bare" charge λ is meaningless. However, we have just proved that the S-matrix formalism still works; and this permits us to define an "effective" charge [through (2.8)], viz. λ_d (Ref. 12) which is finitely related to λ (2.9). We henceforth consider λ_d as the physically relevant quantity.

The Born renormalization may also be performed. We compute it according to the prescription of, say, Refs. 6 and 7, obtaining

$$|\lambda_{\text{Born}}|^2 = |\lambda|^2 \left\{ 1 + \int_0^{\infty} \frac{|\lambda|^2 \pi p^2\, dp}{2M\omega_p E_N(p)(m + \mu - E_N(p) - \omega_p)[E_N(p) - \omega_p - M]} \right\}^{-1} \tag{3.1}$$

which is also finite.

It is worth noting that (2.9) gives a maximal value $\lambda_{\max}$ for $|\lambda_d|$ if we want to maintain unitarity; we let the bare λ go to infinity:

$$|\lambda_{\max}|^2 = \lim_{\lambda \to \infty} |\lambda_d|^2 = \frac{2M}{\pi} \left\{ \int_0^{\infty} \frac{p^2\, dp}{E_N(p)\omega_p [E_N(p) - \omega_p - M]^2} \right\}^{-1}. \tag{3.2}$$

It is clearly finite, and the corresponding scattering amplitude

$$A_{\text{sat}}(s) = \frac{1}{s^{\frac{1}{2}} - M} \left\{ \int_0^{\infty} \frac{p^2\, dp}{E_N(p)\omega_p [s^{\frac{1}{2}} - E_N(p) - \omega_p + i\epsilon][E_N(p) + \omega_p - M]} \right\}^{-1} \tag{3.3}$$

then satisfies all properties (i) to (iv) and is saturated.[8] We also remark that in this limit the Born renormalization is still finite and meaningful.

4. HAMILTONIAN FORMALISM AND SOME COMMENTS

It only remains to prove assertion (1), Sec. 1. Precisely, we want to prove that, although the interaction picture formalism becomes meaningless (for point coupling), one may define, in an appropriate way, a total Hamiltonian. For the first part, we compute the Møller wave operators:

$$\Omega^{\pm} = \int d^3q\, d^3k\, |\psi^{\pm}_{q,k}\rangle\langle 0|\, (a_k N_q + V_{q+k}), \tag{4.1}$$

where the incoming (outgoing) waves are given by

$$|\psi^{\pm}_{q,k}\rangle = \left\{ N^*_q a^*_k + \lambda f_{\pm}(q, k) \left[V^*_{q+k} + \int \frac{d^3p\, \lambda g(q + k, p)}{E_N(q) + \omega_k - E_N(q + k - p) - \omega_p \pm i\epsilon} N^*_{q+k-p} a^*_p \right] \right\} |0\rangle, \tag{4.2}$$

[12] See, e.g., S. Gasirowicz, Fortschr. Physik 8, 665 (1960).

with

$$f_\pm(q,k) = \frac{g(q+k,k)}{E_N(q)+\omega_k - E_N(q+k)} \times \left\{1 + \int \frac{d^3p\ |\lambda|^2\ |g(q+k,p)|^2}{[E_N(q+k-p)+\omega_p - E_V(q+k)][E_N(q+k-p)+\omega_p - E_N(q)-\omega_k \mp i\epsilon]}\right\}^{-1}. \quad (4.3)$$

We may then calculate, e.g., $\langle\Phi|\ \Omega\ |\Phi\rangle$, $|\Phi\rangle = \int d^3p\Phi(p)V_p^*\ |0\rangle$:

$$\langle\Phi|\ \Omega\ |\Phi\rangle = \int d^3p\ |\Phi(p)|^2 \int d^3k\ \lambda f(p-k,k). \quad (4.4)$$

Now, the term $\{\cdots\}^{-1}$ in (4.3) is always finite and tends to 1 as $k \to \infty$; on the other hand, for large k, $g(p,k) \sim 1/|k|^2$, so that $f(p-k,k) \sim 1/|k|^2$ and the second integral in (4.4) diverges: the Møller operators become meaningless and, since $H\Omega^\pm = \Omega^\pm H_0$, the desired result follows.

To prove the second part, consider the set of vectors $|\chi\rangle$ of the form

$$|\chi\rangle = \int d^3p\ \Phi(p)V_p^*\ |0\rangle + \int d^3q\ d^3k \left[\Psi(q,k) - \frac{\lambda g(q+k,k)\Phi(q+k)}{E_N(q)+\omega_k - E_V(q+k)}\right] N_q^* a_k^*\ |0\rangle, \quad (4.5)$$

where $\Phi(p) \gtrsim |p|^{-3}$, $\Psi(q,k) \gtrsim |q|^{-3}\ |k|^{-3}$ for large p, q, k; it is clear that this is a linear manifold dense in the Hilbert space of superpositions of N, θ, and V states. Now, for such vectors, one may define H as follows: first compute $H\ |\chi\rangle$ for finite cutoff; observe that the δE_V term cancels. Thus, we may remove the cutoff to get

$$H\ |\chi\rangle = \int d^3p\ [E_V(p)\Phi(p) + \lambda\ d^3k\ g(p,k)\Psi(p-k,k)]V_p^*\ |0\rangle + \int d^3q\ d^3k \left\{[E_N(q)+\omega_k]\Psi(q,k) - \frac{\lambda E_V(q+k)g(q+k,k)}{E_N(q)+\omega_k - E_V(q+k)}\Phi(q+k)\right\} N_q^* a_k^*\ |0\rangle$$

which is still a normalizable vector of the Hilbert space (check), Q.E.D.

To conclude, we want to comment a little on the main features of our analysis. The situation is remarkable: we started with an interaction formalism, which was only valid for finite cutoff, and then we obtained the S matrix and the Hamiltonian. After this, we extended the results to point coupling by formally removing the cutoff in the previously derived formulas. This very heuristic role played by the interaction description mimics the situation in quantum electrodynamics (with which the Lee model has so many analogies[7,10]) where it has had the well-known spectacular successes—it differs from our model in the fact that here we can rigorously prove (and not only formally calculate and guess) everything. Perhaps the methods could be developed also in quantum electrodynamics to achieve rigorous results; however, nobody yet knows how this can be done.

ACKNOWLEDGMENT

I want to express my gratitude to Professor A. Galindo for valuable discussions and suggestions.

Topology of Some Spheroidal Metrics*

DAVID M. ZIPOY
University of Maryland, Department of Physics and Astronomy, College Park, Maryland
(Received 16 June 1964)

The solutions of Einstein's vacuum field equations, $R_{\mu\nu} = 0$, are found when quasi-oblate and prolate spheroidal coordinates are used. The solutions for the "Newtonian" potential can be written as a linear combination of Legendre polynomides of integral order l. For oblate coordinates the solutions for each l have a ring singularity and have a double sheeted topology; one can get from one sheet to the other by going through the ring. When the $l = 0$ and $l = 1$ solutions are combined an infinite-sheeted topology results from the nonlinear character of the field equations. In general the geometry is asymptotically flat on only one sheet; on the others it is highly distorted. In some cases the region near the "ring singularity" opens out into a multisheeted infinite space. For the prolate coordinates the solutions contain a line singularity of finite length. In general, the prolate coordinate solutions are much less rich in varied topologies than are the oblate solutions.

1. INTRODUCTION

POSSIBLY the most well-known solution of Einstein's vacuum field equations is the "Schwarzschild solution".[1] It is the only solution that has spherical symmetry and it can be interpreted as giving the gravitational field around a spherically symmetric mass distribution. In 1917 Weyl[2] presented the general form of the time-independent metric with axial symmetry. He also showed that the field equations took on a simple form and in practice could be solved to find the field around any axially symmetric mass distribution. Although this metric is well known, it has seldom been used to find simple explicit solutions. This seems to be due to the fact that the coordinates used by Weyl were the analog of cylindrical coordinates in flat space and therefore do not lend themselves to simple solutions for finite mass distributions. The solutions would normally be written as integrals over given mass distributions. It might be expected that a coordinate system that was tailored to fit a given mass distribution would give simple solutions. In the present paper spheroidal coordinates are used; these have ellipsoids and hyperbloids of revolution as coordinate surfaces.

The idea of fitting the coordinate system to the problem is used frequently in physics but normally only cartesian, cylindrical or spherical systems are employed; the other possibilities being relatively unknown and sometimes difficult to picture. The above idea was first mentioned to the author in 1961 by Vinti, who has used spheroidal coordinates to good effect in computing the exact Newtonian gravitational field around the earth. In 1960 Misra[3] used the idea in general relativity and found solutions in oblate spheroidal coordinates and toroidal coordinates. Erez and Rosen have studied the Weyl metric in ellipsoidal coordinates and obtained some solutions similar to those obtained here.[4] Darmois[5] has previously obtained one of the solutions presented here [Eqs. (43)–(45) with $\gamma = 0$] and Weyl also found a similar (but not identical) solution.[2] At large distance these last two solutions correspond to the field around two mass points.

In the present paper some solutions in oblate and prolate spheroidal coordinates are calculated. The main purpose of the paper is to study the topology of the solutions. To this end we do *not* put in a mass distribution but study the vacuum solutions alone; mass will manifest itself as "wormholes" or singularities in the Riemann tensor. This approach has been used to study the Schwarzschild metric.[6] All of the author's mentioned above either put in a mass distribution explicitly or treated their solutions as if they were the external fields of an actual distribution.

II. FORM OF THE METRIC

The starting point of this investigation is the familiar and completely general static cylindrically symmetric line element.[2]

* This research was supported by AFOSR Grant 409-63.
[1] K. Schwarzschild, Berlin. Ber., p. 189 (1916).
[2] H. Weyl, Ann. Physik 54, 117 (1917).
[3] M. Misra, Proc. Natl. Inst. Sci. India A26, 673 (1960); A27, 373 (1961).
[4] G. Erez and N. Rosen, Bull. Res. Council Israel 8F, 47, (1959).
[5] G. Darmois, *Les Equations de la Gravitation Einsteinienne* (Gauthier-Villars, Paris, 1927), p. 36.
[6] J. A. Wheeler, *Geometrodynamics* (Academic Press Inc., New York, 1962), p. 45. M. Kruskal, Phys. Rev. 119, 1743 (1960).

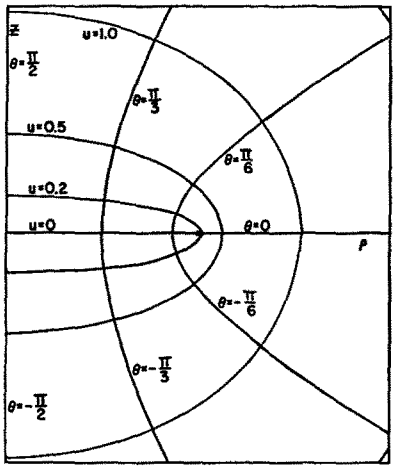


FIG. 1. A graph showing the relation between cylindrical and oblate spheroidal coordinates.

$$ds^2 = -e^{2(\lambda-\sigma)}(d\rho^2 + dz^2) - \rho^2 e^{-2\sigma}\, d\varphi^2 + e^{2\sigma}\, dt^2, \qquad (1)$$

where

$$\lambda = \lambda(\rho, z) \quad \text{and} \quad \sigma = \sigma(\rho, z) \qquad (2)$$

are two functions to be found by solving Einstein's vacuum field equations. The coordinate system that we consider first is "oblate spheroidal coordinates".[7] These are illustrated in Fig. 1; they have ellipsoids and hyperboloids of revolution making up the orthogonal system. The relation between ρ and z and the new coordinates u and θ are

$$\rho = a \cosh u \cos\theta, \qquad z = a \sinh u \sin\theta, \qquad (3)$$

where a is a constant. The new line element is

$$ds^2 = -a^2 e^{2(\lambda-\sigma)}(\sinh^2 u + \sin^2\theta)(du^2 + d\theta^2) - a^2 e^{-2\sigma} \cosh^2 u \cos^2\theta\, d\varphi^2 + e^{2\sigma}\, dt^2. \qquad (4)$$

With this metric the vacuum field equations reduce to the following four equations.[8]

$$\sigma_{11} + \sigma_{22} + \sigma_1 \tanh u - \sigma_2 \tan\theta = 0, \qquad (5)$$

$$\sigma_1^2 - \sigma_2^2 - \lambda_1 \tanh u - \lambda_2 \tan\theta = 0, \qquad (6)$$

$$2\sigma_1\sigma_2 + \lambda_1 \tan\theta - \lambda_2 \tanh u = 0, \qquad (7)$$

$$\lambda_{11} + \lambda_{22} + \sigma_1^2 + \sigma_2^2 = 0. \qquad (8)$$

Equation (5) is just Laplace's equation in these coordinates. Another change of variables converts it into a more suitable form. Let

$$x = \sinh u, \qquad y = \sin\theta. \qquad (9)$$

Then (5) becomes

$$\frac{\partial}{\partial x}\left[(x^2+1)\frac{\partial\sigma}{\partial x}\right] + \frac{\partial}{\partial y}\left[(1-y^2)\frac{\partial\sigma}{\partial y}\right] = 0. \qquad (10)$$

Equation (10) can be solved by separation of variables. Let $\sigma(x, y) = R(x)S(y)$. Then

$$\frac{\partial}{\partial x}\left[(1-y^2)\frac{\partial S}{\partial y}\right] + l(l+1)S = 0 \qquad (11)$$

and

$$\frac{\partial}{\partial x}\left[(x^2+1)\frac{\partial R}{\partial x}\right] - l(l+1)R = 0. \qquad (12)$$

If we require that σ be finite along $y = \pm 1$ ($\theta = \frac{1}{2}\pi$), then the parameter l is a positive integer or zero and the solution for $S(y)$ is given as a Legendre polynomial of the first kind. The solution for $R(x)$ will then be a sum of Legendre polynomials of the first and second kind with imaginary argument. Once σ is found we can obtain λ from (6) and (7). [Equation (8) is then an identity.]

$$\lambda = \int^x \frac{(1-y^2)[(x^2+1)\sigma_x^2 - (1-y^2)\sigma_y^2] - 2y\sigma_x\sigma_y}{x^2+y^2}\,dx + \int^y \frac{(x^2+1)[(x^2+1)\sigma_x^2 - (1-y^2)\sigma_y^2] - 2y\sigma_x\sigma_y}{x^2+y^2}\,dy. \qquad (13)$$

In order to make the solutions more definite we impose the boundary conditions that σ and λ approach zero as x approaches infinity; we require the space to be asymptotically flat. It may be well to state explicitly at this point that we interpret x as a radial coordinate and y as an angular coordinate at least when x is large. This is in analogy to r and $\cos\theta$ in spherical coordinates and we will see that this is consistant with the solutions that we obtain. Of course, when x is small it is obvious from Fig. 1 that we must be more careful. Since λ goes to zero as x goes to infinity, we can write (13) as

$$\lambda = \int_{+\infty}^{x} \frac{(1-y^2)[(x^2+1)\sigma_x^2 - (1-y^2)\sigma_y^2] - 2y\sigma_x\sigma_y}{x^2+y^2}\,dx. \qquad (14)$$

In (14) y is held fixed during the integration.

III. PARTICULAR SOLUTIONS

The solution of (11) and (12) for $l = 0$ is

$$\sigma = -\beta(\tfrac{1}{2}\pi - \arctan x) = -\beta \arctan(1/x), \qquad \beta = m/a. \qquad (15)$$

[7] P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill Book Company, Inc., New York, 1953), pp. 661–2.

[8] J. L. Synge, *Relativity, The General Theory* (North-Holland Publishing Company, Amsterdam, 1965), p. 312.

One of the integration constants has been chosen such that σ approaches zero as x approaches infinity; the other constant is m. Integration of (14) yields

$$\lambda = \tfrac{1}{2}\beta^2 \ln [(x^2 + y^2)/(x^2 + 1)]. \tag{16}$$

By redefining our variables we can write the line element as

$$ds^2 = -e^{2(\nu-\sigma)}[dr^2 + (r^2 + a^2)\, d\theta^2] - e^{-2\sigma}(r^2 + a^2)\cos^2\theta\, d\varphi^2 + e^{2\sigma}\, dt^2, \tag{17}$$

where

$$r = ax, \qquad \sin\theta = y, \qquad \beta = m/a, \tag{18}$$

$$\sigma = -\beta \arctan(a/r), \quad \text{with} \quad 0 \leq \arctan(a/r) \leq \pi, \tag{19}$$

and

$$e^{2\nu} = \left(\frac{r^2 + a^2\sin^2\theta}{r^2 + a^2}\right)^{\beta^2+1}. \tag{20}$$

The properties of the metric are now investigated. As r approaches infinity, (17) becomes the isotropic Schwarzschild line element with m as the mass of the field source (in suitable units), thus, r, θ, φ asymptotically become the usual spherical coordinates. For later convenience θ has been measured from the equator rather than the pole. We see then that at very large values of r the above field approaches that around a spherical mass. More interesting things happen for small r. (When $r \ll a$, $u = r/a$, and so Fig. 1 is useful for interpreting the following results.) The function σ is a bounded function for all r; however, $e^{2\nu}$ is zero at $r = \theta = 0$. We see later that the invariants of the Riemann tensor are infinite on this line. We can compute the circumference C of this ring singularity at $r = \theta = 0$.

$$C = \int_0^{2\pi} (-g_{33})^{\frac{1}{2}}\, d\varphi = 2\pi a e^{\frac{1}{2}\pi\beta}. \tag{21}$$

We can define its diameter as the proper length of a spacial geodesic joining the point $\varphi = 0$ to $\varphi = \pi$ and from symmetry we would expect this to be a line along $r = 0$ ($u = 0$ in Fig. 1); however, this line is *not* a geodesic. The properties of the space inside the ring can be resolved if we attempt to drop a test mass down the z axis (Fig. 1). The force acting on a stationary unit mass held on the z axis just above $r = 0$ ($r = 0^+$, $\theta = \frac{1}{2}\pi$) is

$$f^1 = -\Gamma^1_{00}(dt/d\tau)^2 = -(\beta/a)e^{-\pi\beta} \neq 0 \tag{22}$$

whereas the force just below $r = 0$ ($r = 0^+$, $\theta = -\frac{1}{2}\pi$) is the negative of (22). So it would appear that for no particular reason the force suffers a finite discontinuity as the test particle goes through $r = 0$. A further consideration resolves the difficulty because we note that nothing prevents the coordinate r from becoming negative. If the coordinate r is allowed to go negative and θ is fixed at $+\frac{1}{2}\pi$ the trajectory of the particle passes smoothly through $r = 0$. A more realistic picture of the topology is given in Fig. 2 where it is illustrated that any plane through the z axis really consists of two sheets which interpenetrate inside of the ring. On the second sheet the sign of r and θ are opposite from Fig. 1. A study of the metric (17) shows that there are only two sheets because if the singularity at $r = \theta = 0$ is "walked around" twice, one returns to the starting point [see Fig. 3 and Eq. (19)]; that is, the metric returns to its original value.

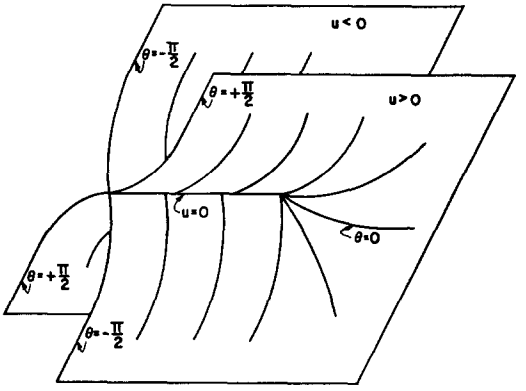


FIG. 2. This drawing illustrates the two-sheeted character of the two-space (α = const, t = const). The two sheets have been pulled apart to show the interconnections more clearly.

In Eqs. (17)–(20) we see that as r approaches minus infinity, ν approaches zero and σ approaches $-\pi\beta + m/r$; therefore the space is asymptotically Schwarzschildian but the scales of r and t are different on the two sheets and the corresponding Schwarzschild mass on the second sheet is $-me^{\pi\beta}$. In general the force on a stationary unit mass is

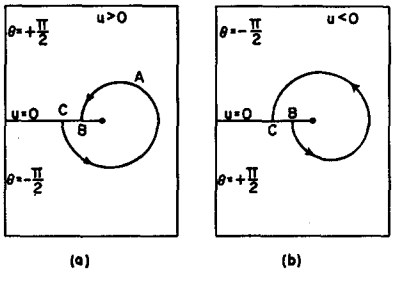


FIG. 3. The path of a person who walks around the ring singularity of the metric of Eqs. (18)–(20). He starts at A and walks to B, he then is automatically on the lower sheet. If he proceeds around the singularity he passes through C and comes out on the upper sheet where he returns to A.

$$f^\mu = -\Gamma^\mu_{00}\left(\frac{dt}{ds}\right)^2 = -\frac{m}{r^2+a^2}\,e^{2(\sigma-\upsilon)}\,\delta^\mu_1 \leq 0. \qquad (23)$$

Therefore for large r the force on the upper sheet ($r > 0$) is toward the ring whereas on the lower sheet ($r < 0$) it is *away* from the ring, a repulsive force.

It is interesting to conjecture about this solution. There are two independent parameters, the mass m and the "radius" a, which can be picked for convenience, say m equals 5 oz. and a equals 2 in. Now *if* we could find an "interior" solution around $r = \theta = 0$ without destroying the two sheeted topology, then we would have a toroidal object of reasonable size and mass that would have very strange properties. For instance if the toroid were viewed from the side as an object fell through the hole, the object would not be seen coming out the other side, whereas if viewed from above the ring it could be seen dropping through. We make no attempt to find such an interior solution but it should be noted that Weyl[2] has considered a similar case. He made the observation that since σ obeys the same equation as the Newtonian potential, then the solution for σ can be written down in terms of an integral over the mass distribution. In particular he calculated the field around a ring with rectangular cross section and, of course, found no unusual behavior. Apparently Weyl's solution consists of a linear combination of our solutions (with all l values present) which combine in such a way as to eliminate the double-sheeted topology. The details on how this happens would be very interesting but we do not pursue them further in this paper. We should clarify one other point. The term "viewed" was used above and now the question arises as to what a manifestation of the solution would actually look like. The geodesic equations for null rays are exceedingly hard to integrate analytically but a semiquantitative investigation of them indicates that the manifestation would indeed look like a ring. Although there is distortion near the ring singularity, the light rays do not seem to get bent violently out of line.

Finally a calculation of the three nonzero invariants of the Riemann tensor gives

$$\lambda_3 = -e^{2(\sigma-\upsilon)}[m(r-m)/(r^2+a^2)^2] \qquad (24)$$

and

$$\lambda_\pm = \tfrac{1}{2}e^{2(\sigma-\upsilon)}\,\frac{m}{(r^2+a^2)^2} \times\left[(r-m) \pm \left\{\frac{[-4m(r-m)^3 - a^2(r-m)^2 + 4(mr+a^2)(r^2+a^2)]\cos^2\theta + 9(r-m)^2(r^2+a^2)}{r^2+a^2\sin^2\theta}\right\}^{\frac{1}{2}}\right]. \qquad (25)$$

All three are infinite at $r = \theta = 0$. In the Petrov classification the metric is of type I.[9]

Equations (11), (12), and (14) can be solved with $l = 1$ and the result put into the form of Eq. (17) with

$$\sigma = \gamma\left(1 - \frac{r}{a}\arctan\frac{a}{r}\right)\sin\theta, \qquad 0 \leq \arctan\frac{a}{r} \leq \pi; \qquad (26)$$

$$\upsilon = \tfrac{1}{2}(1-\gamma^2)\ln\left(\frac{r^2+a^2\sin^2\theta}{r^2+a^2}\right) - \tfrac{1}{2}\gamma^2\cos^2\theta\left[\left(\arctan\frac{a}{r}\right)^2 + \left(1 - \frac{r}{a}\arctan\frac{a}{r}\right)^2\right], \qquad (27)$$

where $\gamma = 3p/a^2$, and p = const.

As r approaches plus infinity the "Newtonian potential" σ becomes that of a dipole with p being the dipole moment. As in the case of the "monopole" solution ($l = 0$), the "dipole" solution ($l = 1$) has a two-sheeted topology, however, as r approaches minus infinity the metric term $e^{2\upsilon}$ approaches zero. An investigation of the Riemann invariants shows them becoming infinite in this limit. A detailed discussion of this solution is postponed until we discuss the combined "monopole–dipole" solution, but first two other features of the "dipole" solution are mentioned. In (19) and (26) we chose the values of the arctangent function at a point to be between 0 and π. For the "monopole" solution this is no restriction because choosing a different sheet of the arctangent merely changes the scale of r and t; the geometry is unchanged. Such is *not* the case for the $l = 1$ solution; the metric tensor is no longer flat at $r = +\infty$ for any sheet other than the one chosen in (26). This aspect of the solution has not been investigated thoroughly, but it would appear that the properties of this solution on the other sheets is not radically different from those discussed in the next solution. The second point to be men-

[9] F. A. E. Pirani, Phys. Rev. 105, 1089 (1957).

tioned is that the factor $1 - \gamma^2$ in (27) can become negative and thus change the nature of singularity at $r = \theta = 0$; this too is discussed shortly.

Because of the nonlinearity of the field equations and in particular of (14), the combined solution for $l = 0$ and $l = 1$ is not just the superposition of the individual solutions but it has an "interference" term in addition. For the combined "monopole–dipole" solution the line element is in the form of (17) again with

$$\sigma = -\beta \arctan \frac{a}{r} + \gamma\left(1 - \frac{r}{a} \arctan \frac{a}{r}\right) \sin \theta, \quad (28)$$

and

$$\upsilon = \tfrac{1}{2}\alpha \ln \left(\frac{r^2 + a^2 \sin^2 \theta}{r^2 + a^2}\right) - 2\beta\gamma\left(\sin \theta \arctan \frac{a}{r} - \arctan \frac{a \sin \theta}{r}\right) - \tfrac{1}{2}\gamma^2 \cos^2 \theta\left[\left(\arctan \frac{a}{r}\right)^2 + \left(1 - \frac{r}{a} \arctan \frac{a}{r}\right)^2\right], \quad (29)$$

where $\beta = m/a$, $\gamma = 3p/a^2$, $\alpha = 1 + \beta^2 - \gamma^2$, $0 \leq \arctan a/r \leq \pi$.

Both β and γ can be taken as positive without loss in generality. The interference term appearing in the metric contains not only arctan (a/r) but also arctan $(a \sin \theta/r)$; this latter function completely alters the topology of the space. If we make a plot of arctan $(a \sin \theta/r)$ as a function of r and $\sin \theta$, a helical surface is formed which looks similar to the surface traced out by a propeller blade as it advances through the air. In particular it is *not* periodic. This means that if we walk around the point $r = \theta = 0$ in a clockwise direction, arctan $(a \sin \theta/r)$ increases monotonically and therefore the metric continuously changes and never returns to its original value. The only way to return to the starting point is to turn around and retrace our path. This is in contrast to the two-sheeted behavior of the separate "monopole" and "dipole" solutions where, after two circuits of $r = \theta = 0$, we returned to the starting point. The "interference" term, however small, changes the space from a two-sheeted to an infinitely sheeted topology. Of course, in the previous solutions we could have used more than two sheets also, but it was unnecessary. In the present case it is necessary.

Before undertaking a study of the singularity at $r = \theta = 0$, the singularity at $r = -\infty$ is investigated. It turns out that the line $\theta = \frac{1}{2}\pi$, $r < 0$ [Fig. 3(b)] is the spacial projection of any type of geodesic (spacelike, null, and timelike) and provides a convenient set of geodesics for studying $r = -\infty$. At $\theta = \frac{1}{2}\pi$, $e^{2\upsilon}$ equals unity on the first sheet of arctan $(a \sin \theta/r)$. Its value changes by a constant factor $e^{8\pi\beta\gamma}$ between adjacent sheets. For the present we confine our attention to the first sheet. The proper distance l from $r = 0$ to $r = -\infty$ is given by

$$l = -\int_0^{-\infty} e^{-\sigma}\, dr = -\int_0^{-\infty} \exp\left[\beta \arctan \frac{a}{r} - \gamma\left(1 - \frac{r}{a} \arctan \frac{a}{r}\right)\right] dr. \quad (30)$$

It is finite and of the order of $(a/\pi\gamma) \exp(\beta\pi - \gamma)$ for small γ. For a timelike geodesic the geodesic equations give at $\theta = \frac{1}{2}\pi$,

$$\left(\frac{dr}{d\tau}\right)^2 = k^2 - e^{2\sigma}, \qquad \frac{dt}{d\tau} = ke^{-2\sigma};$$

$$\tau = \text{proper time}, \quad k = \text{const.} \quad (31)$$

Equation (28) shows that $e^{2\sigma}$ blows up as r approaches minus infinity and therefore at some point $(dr/d\tau)$ equals zero. This means that a mass thrown down the z axis through the ring is reflected eventually. Similarly we find that a light ray reaches the singularity in a finite coordinate time. Finally, we can calculate the "circumference" C of the singularity at $r = -\infty$ $(\theta > 0)$.

$$C = \lim_{r \to -\infty} \int_0^{2\pi} e^{-\sigma}(r^2 + a^2)^{\frac{1}{2}} \cos \theta\, d\varphi = 0.$$

The space below the ring is highly distorted.

The z axis constitutes a singularity of sorts on all the sheets but the first one. Since $e^{2\upsilon}$ is not equal to unity on the z axis in general, the surfaces with constant r and t are not locally flat there. For example on the second sheet $(r < 0)$ in the neighborhood of $\theta = -\frac{1}{2}\pi$, the value of υ is $4\pi\beta\gamma$. This means that the ratio of circumference to radius of a small circle around the z axis is $2\pi \exp(-4\pi\beta\gamma)$. For $\beta\gamma$ positive this corresponds to the geometry near the vertex of a cone.

The qualitative behavior of particles in this space can be obtained by studying the force f^μ acting on a stationary unit test mass:

$$f^\mu = -\Gamma^\mu_{00}(dt/d\tau)^2 = -\Gamma^\mu_{00}e^{-2\sigma}, \quad (33)$$

$$f^1 = -(\partial\sigma/\partial r)e^{2(\sigma-\upsilon)}, \quad (34)$$

$$f^2 = -(\partial\sigma/\partial\theta)[1/(r^2 + a^2)]e^{2(\sigma-\upsilon)}. \quad (35)$$

A study of f^2 shows that it is always negative. The force f^1 is negative for negative θ and all r. For positive θ it is negative for large positive r but

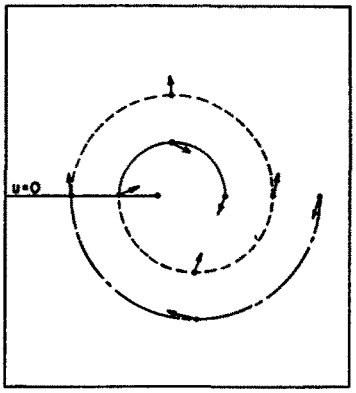


FIG. 4. A sketch showing the qualitative behavior of the force on a stationary mass in the space of a monopole–dipole. The arrows on the solid curve correspond to the directions of the forces on the first Riemann sheet, those on the dashed curve the second sheet, and those on the dot-dashed curve the third sheet. There are an infinite number of sheets.

positive for large negative r. A qualitative sketch of the directions of f^a for three of the sheets of the space is shown in Fig. 4.

We now turn to a study of the behavior of the metric near the "ring singularity" at $r = \theta = 0$. A calculation of the *leading* term in the Riemann invariants yields

$$\lambda_3 \approx \frac{1}{a^2}(\beta^2 + \gamma^2)e^{2(\sigma-v)} \propto (x^2 + \theta^2)^{-\alpha}, \tag{36}$$

$$\lambda_\pm \approx \pm\frac{1}{a^2}\left[\frac{(\beta^2 + \gamma^2)(\alpha^2 + 4\beta^2\gamma^2)}{x^2 + \theta^2}\right]^{\frac{1}{2}} e^{2(\sigma-v)} \propto (x^2 + \theta^2)^{-\alpha-\frac{1}{2}}, \tag{37}$$

where $\alpha = 1 + \beta^2 - \gamma^2$ and $x = r/a$. It should be noted that θ is no longer conveniently interpretable as an angle. We see from Fig. 1 that r and θ are rather similar when they are small. For $\alpha > -\frac{1}{2}$ the invariants are singular; for $\alpha < -\frac{1}{2}$ the invariants vanish and the space is asymptotically flat. For $\alpha = -\frac{1}{2}$, $\lambda_3 \to 0$ but $\lambda_\pm$ are finite and nonzero (but multivalued as is shown later).

It is convenient to introduce new coordinates for which the metric components neither vanish nor blow up at $r = \theta = 0$. When $\alpha \neq -1$ let

$$\rho = a(x^2 + \theta^2)^{\frac{1}{2}(\alpha+1)}, \qquad \psi = \arctan(\theta/x). \tag{38a}$$

Then

$$r = a(\rho/a)^{1/(\alpha+1)} \cos\psi \qquad \theta = (\rho/a)^{1/(\alpha+1)} \sin\psi. \tag{38b}$$

When $\alpha = -1$ let

$$\rho = -\tfrac{1}{2}a \ln(x^2 + \theta^2), \qquad \psi = \arctan(\theta/x). \tag{39a}$$

Then

$$r = a\cos\psi e^{-\rho/a}, \qquad \theta = \sin\psi e^{-\rho/a}. \tag{39b}$$

The form of the line element near $r = \theta = 0$ is

$$ds^2 \approx -(x^2 + \theta^2)^{-\alpha}e^{2(v-\sigma)}\left[\frac{d\rho^2}{(1+\alpha)^2} + \rho^2\, d\psi^2\right] - a^2e^{-2\sigma}\, d\varphi^2 + e^{2\sigma}\, dt^2, \qquad \alpha \neq -1; \tag{40}$$

$$ds^2 \approx -(x^2 + \theta^2)e^{2(v-\sigma)}(d\rho^2 + a^2\, d\psi^2) - a^2e^{-2\sigma}\, d\varphi^2 + e^{2\sigma}\, dt^2, \qquad \alpha = -1, \tag{41}$$

where

$$\sigma \approx -\tfrac{1}{2}\pi\beta \quad \text{and} \quad e^{2v} \approx (x^2 + \theta^2)^{\alpha}e^{-\gamma^2(\frac{1}{2}\pi+1)}e^{4\beta\gamma\psi}. \tag{42}$$

If we substitute (42) into (36) and (37) we see that $\psi = -\infty$ is a singularity. Equations (38) and (39) show that the point $r = \theta = 0$ corresponds to $\rho = 0$ for $\alpha > -1$ and $\rho = \infty$ for $\alpha \leq -1$. Since the metric is finite and nonzero (for ψ finite) this means that for $\alpha \leq -1$ the region near the "ring" is really an infinitely large space. That is, when the parameter α gets negative enough the space looses some of its singular character and grows a helical "wormhole."[6]

IV. ANOTHER CLASS OF SOLUTIONS

A glance at Fig. 1 shows that if we take the ρ axis as the axis of symmetry a qualitatively different class of solutions will arise, that is, solutions with prolate rather than oblate coordinate surfaces (see Fig. 5). Strangely enough we can obtain these new solutions from the old by assuming that a is purely imaginary. If we make the change of parameters and variables $a \to ia$; $x \to -ix$; $\beta \to -i\beta$; $\gamma \to -\gamma$ in the previous solutions (28) and (29) we get

$$ds^2 = -e^{2(v-\sigma)}[dr^2 + (r^2 - a^2)\, d\theta^2] - e^{-2\sigma}(r^2 - a^2)\cos^2\theta\, d\varphi^2 + e^{2\sigma}\, dt^2, \tag{43}$$

where

$$\sigma = -\tfrac{1}{2}\beta \ln\frac{x+1}{x-1} - \gamma\left(1 + \tfrac{1}{2}x \ln\frac{x+1}{x-1}\right)\sin\theta, \qquad r = ax, \tag{44}$$

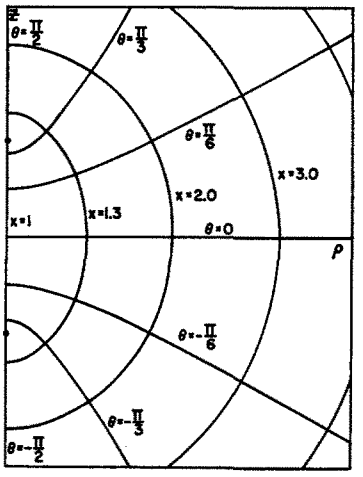


FIG. 5. A graph showing the relation between cylindrical and prolate spheroidal coordinates.

$$v = \tfrac{1}{2}(1 - \beta^2 - \gamma^2) \ln\left(\frac{x^2 - \sin^2\theta}{x^2 - 1}\right) - \beta\gamma \ln\frac{x + \sin\theta}{x - \sin\theta} + \beta\gamma \sin\theta \ln\frac{x+1}{x-1} - \tfrac{1}{2}\gamma^2 \cos^2\theta\left[-\frac{1}{4}\left(\ln\frac{x+1}{x-1}\right)^2 + \left(1 - \tfrac{1}{2}x \ln\frac{x+1}{x-1}\right)^2\right]. \quad (45)$$

Also

$$\lambda_3 = -\frac{m}{a^3}(x - \beta)(x+1)^{-(1+\beta+\beta^2)}(x-1)^{-(1-\beta+\beta^2)} \times (x^2 - \sin^2\theta)^{\beta^2-1} \quad (\text{for } \gamma = 0 \text{ only}). \quad (46)$$

For the case $\gamma = 0$ we see from (46) that $x = \pm 1$ are singularities for all β except $\beta = \pm 1$. Therefore, except for these two cases, x ranges from $+1$ to $+\infty$ or -1 to $-\infty$. This is in contrast to the oblate solutions in which it ranged from $-\infty$ to $+\infty$. The case for $\beta = 1$ is interesting because on substitution of $x' = x+1$ we obtain none other than the Schwarzschild spherically symmetric solution and the solution extends down to $x = -1$.[6]

For the case $\gamma = 0$, the line $\theta = 0$ is a geodesic. A computation of the proper distance along this line from any point r_0 to $r = a$ shows this distance to be finite for all values of β. If we drop an object along $\theta = 0$, we find that it takes finite coordinate and proper times for the object to reach $r = a$. This object arrives at $r = a$ and *stops* if $1 > \beta > \frac{1}{2}(1 - 5^{\frac{1}{2}})$, otherwise it arrives with infinite velocity [except that $dr/d\tau$ is finite at $r = a$ for $\beta = 1$ and $\frac{1}{2}(1 - 5^{\frac{1}{2}})$]. The "circumference" C around $x = 1$, $\theta = 0$, $t = \text{const}$, is interesting.

$$C = \lim_{x\to 1}\int_0^{2\pi} e^{-\sigma}(r^2 - a^2)^{\frac{1}{2}}\, d\varphi = \lim_{x\to 1} 2\pi a(x+1)^{\frac{1}{2}(\beta+1)}(x-1)^{\frac{1}{2}(1-\beta)}. \quad (47)$$

This is infinite for $\beta > 1$, zero for $\beta < 1$ and $4\pi a$ for $\beta = 1$. Therefore one has, for $\beta > 1$, an infinite "circumference" whereas the proper distance along $\theta = 0$ to any point on this line is finite.

For $\gamma \neq 0$ and any values of x, θ,

$$C = 2\pi a(x^2 - 1)^{\frac{1}{2}} \cos\theta\left(\frac{x+1}{x-1}\right)^{\frac{1}{2}\beta} \times \exp\left[\gamma\left(1 - \tfrac{1}{2}x \ln\frac{x+1}{x-1}\right)\right] \sin\theta \quad (48)$$

and the properties of the space near $x = 1$ are similar to above but lose their symmetry around $\theta = 0$.

The interpretation of the metric (43)–(45) for $|x| < 1$ has not been carried very far; however, one point has been noted. Take the case $\gamma = 0$. Equation (44) shows that in general $e^{2\sigma}$ and e^{2v} are complex for $|x| < 1$; but if β is a rational fraction with odd denominator the metric is real, although its signature may change.

V. CONCLUSION

A thorough study of the above class of solutions could prove to be a fruitful source of different types of topologies, but more important, a source of qualitative ideas about how the nonlinear character of the field equations affects an attempt to "superimpose" two different solutions. Undoubtedly even more spectacular changes can occur than those that happened above.

In the "oblate" solutions studied in this paper the properties of the metric near $r = \theta = 0$ for $\alpha = -\frac{1}{2}$ are still somewhat mysterious. For this case λ_3 is zero [Eq. (36)] and $\lambda_\pm$ are finite, but their values depend on the direction of approach to the ring. Further, a study of the geodesic equations near the ring seems to indicate that a test particle can reach there in a finite proper time, finite coordinate time, and arrive with finite velocity. This would indicate that it could continue on, although where it goes is not clear. Possibly this solution connects to one with negative β and γ at the ring; indeed this may happen for all $-1 < \alpha \leq -\frac{1}{2}$ because the λ's are not infinite for any α in this range.

The "prolate" solutions seem somewhat simpler than the oblate ones although the details of the geometry near $x = 1$ are still unclear. As mentioned above, the solutions for $|x| < 1$ have not been studied in detail but may prove interesting.

ACKNOWLEDGMENTS

I am indebted to C. Misner for pointing out the singularity along the z axis in the monopole–dipole solution, and to C. Misner and K. Thorne for pointing out previous work with similar solutions.

Note on Product Integrals

G. E. BACKUS AND F. GILBERT
Institute of Geophysics and Planetary Physics, University of California, San Diego, California
(Received 20 July 1965)

Attention is called to Volterra's product integral solution of ordinary linear differential equations with variable coefficients.

A RECENT paper by J. Kane and E. R. Surynarayan[1] discusses what amounts to Volterra's[2,3] theory of the product integral solution of ordinary linear differential equations. Since Volterra's work seems not to be as widely known as it should be, we would like to call attention to the fact that he has already solved, for linear ordinary differential equations of arbitrary order, the problem discussed by Kane and Surynarayan for second-order equations. The solution of

$$dV/dt = P(t)V(t) \tag{1}$$

when V is an $n \times 1$ column vector and P is an $n \times n$ matrix, is

$$V(t) = [\pi(I + P(t)\,dt)]V(0),$$

where I is the $n \times n$ unit matrix and $\pi[I + P(t)dt]$ denotes Volterra's product integral, sometimes called the matricant of (1). Product integrals are defined as the limits of Riemann partial products, just as ordinary integrals are limits of Riemann sums. The theories of convergence and approximation are the same for Riemann sums and products, and most of the properties of ordinary Riemann integrals have their analogs for product integrals.

The fact that Volterra has proved the convergence of the partial product approximations to a product integral implies that (1) can be solved with any degree of accuracy by subdividing the interval of integration with sufficient fineness and approximating $P(t)$ by a constant matrix within each interval of the subdivision.

[1] J. Kane and E. R. Surynarayan, J. Math. Phys. **6**, 966 (1965).
[2] V. Volterra, Mem. Soc. Ital. Sci. (3), Vol. 6, pp. 1–104 (1887); vol. 12, pp. 3–68 (1902).
[3] F. R. Gantmacher, *The Theory of Matrices*, (Chelsea Publishing Company, New York, 1959), Vol. 2, pp. 135–141.