JOURNAL OF
MATHEMATICAL
PHYSICS

VOLUME 7, NUMBER 3 MARCH 1966

Error Bounds for Approximations to Expectation Values of Unbounded Operators

NORMAN W. BAZLEY*
Institut Battelle, Geneva, Switzerland
AND
DAVID W. FOX*†
Applied Physics Laboratory, The Johns Hopkins University, Silver Spring, Maryland
(Received 26 July 1965)

A bound for $|(B\Psi, \Psi) - (B\varphi, \varphi)|$, the error between the expectation value of a self-adjoint operator B on an eigenfunction Ψ of a self-adjoint Hamiltonian H and the value given by an approximating vector φ, is obtained when H is a radial Schödinger operator and B is multiplicative and unbounded. The analysis makes use of point and asymptotic estimates for Ψ.

I. INTRODUCTION

IN this article we consider the problem of estimating the error in the approximation of expection values of an unbounded self-adjoint operator B on eigenfunctions of a self-adjoint Hamiltonian H. The operator H defined in a separable Hilbert space $\mathfrak{H}$ having the inner product (φ, ψ) is supposed to be bounded below and to have an initial point spectrum of ordered eigenvalues,

$$E_1 \leq E_2 \leq \cdots .$$

The corresponding eigenfunctions, $\Psi_1, \Psi_2, \cdots$, are presumed to be orthonormal; the square of their values is a probability distribution. The operator B corresponds to an observable, and the quantity $(B\Psi, \Psi)$ is interpreted as the expected value of the observable for the bound state[1] Ψ. We consider approximations to quantities of the form $(B\Psi_\mu, \Psi_\nu)$ when E_μ and E_ν are simple eigenvalues of H. Thus, we wish to obtain upper bounds for the quantities

$$|(B\Psi_\mu, \Psi_\nu) - (B\varphi_\mu, \varphi_\nu)|,$$

where φ_μ and φ_ν are approximations to Ψ_μ and Ψ_ν, respectively.

In Sec. II we recall[2] the kinds of results obtainable by function–analytic methods when B is bounded in norm or bounded relative to a norm generated by H. In Sec. III we develop an error bound for a class of expectation values associated with multiplicative operators B on eigenfunctions of radial Hamiltonians. Sec. IV discusses an elementary application of the result.

II. EIGENVALUES, EIGENVECTORS, AND BOUNDED EXPECTATION OPERATORS

A. Bounds to Eigenvalues

When we take B to be the Hamiltonian H itself, we have

$$(B\Psi, \Psi) = (H\Psi, \Psi) = E,$$

* Supported in part by the NATO Research Grants Programme under the number SA 5-2-05 (170)/278 (64) HH.
† Supported in part by a NSF travel grant, and in part by the Aerospace Research Laboratories, Office of Aerospace Research United States Air Force, Contract No. As-5-298.

[1] When there is no possibility of confusion, we often omit the subscript on eigenfunctions and eigenvalues.

[2] See N. W. Bazley and D. W. Fox, Rev. Mod. Phys. **35**, 712 (1963); and P. O. Löwdin, Ann. Rev. Phys. Chem. **11**, 107 (1960).

where Ψ is the eigenfunction corresponding to any simple eigenvalue E; and the estimation of the expectation value is the same as that of finding an upper bound E^u and a lower bound E^l for E. Upper bounds can be readily obtained by the Rayleigh–Ritz procedure and lower bounds can often be found by recently developed methods.[3]

B. Norm Estimates to Eigenvectors

When B is the orthogonal projection on a normalized vector φ that approximates Ψ, the expectation values take the form

$$(B\Psi, \Psi) = |(\Psi, \varphi)|^2.$$

Since these expectation values are always less than one, the problem reduces to finding a lower bound for $|(\Psi, \varphi)|^2$. This problem is in turn equivalent to that of estimating upper bounds for $||\varphi - \Psi||$. Bounds of this sort have been given for appropriate classes[4] of approximating vectors φ.

C. Bounded Expectation Operators

When B is bounded relative to the norm of $\mathfrak{H}$, we have $|(Bv, v)| \leq \alpha ||v||^2$ for some positive constant α and any vector v in $\mathfrak{H}$. Then it is easy to show that

$$|(B\Psi, \Psi) - (B\varphi, \varphi)| \leq 2\alpha ||\Psi - \varphi||.$$

More generally, when B is bounded relative to a norm generated by H, we have

$$|(Bv, v)| \leq \beta[(Hv, v) + (1 - E_1^l)(v, v)]$$

for some positive β and all v in the domain of H. Then the result

$$|(B\Psi, \Psi) - (B\varphi, \varphi)| \leq 2\beta\{1 + E^u - E_1^l\}^{\frac{1}{2}} \times \{E^u - E^l + 2(1 + E^u - E_1^l)\ ||\Psi - \varphi||^2\}^{\frac{1}{2}}$$

can be obtained for Rayleigh–Ritz vectors φ. Here E^u is given by $(H\varphi, \varphi)$, and E^l is a lower bound to the eigenvalue E.

III. ERROR ESTIMATES FOR EXPECTATION VALUES OF AN UNBOUNDED OPERATOR

The main result of this section is an error bound for approximations to expectation values of a class of unbounded multiplicative operators acting on the eigenfunctions of radial Hamiltonians. Our result makes use of asymptotic and point-wise estimates for the eigenfunctions.

We assume that the eigenfunctions Ψ and their approximations φ are in the domain of B. By Schwarz' inequality we have

$$\text{and} \qquad \begin{aligned} |(B[\Psi - \varphi], \Psi)| &\leq ||\Psi - \varphi||\ ||B\Psi||, \\ |(B\varphi, \Psi - \varphi)| &\leq ||\Psi - \varphi||\ ||B\varphi||, \end{aligned} \tag{1}$$

and hence, we obtain

$$|(B\Psi, \Psi) - (B\varphi, \varphi)| \leq ||\Psi - \varphi||\ \{||B\varphi|| + ||B\Psi||\}, \tag{2}$$

which gives an error bound[5] in terms of the unknown quantity $||B\Psi||$.

The remainder of this section is devoted to the estimation of upper bounds for $||B\Psi||$. We assume that $\mathfrak{H}$ is the real Hilbert space $\mathfrak{L}^2(0, \infty)$ and that the self-adjoint operator H has the expression

$$Hv = -v'' + V(r)v \tag{3}$$

on a suitably defined domain $\mathfrak{D}_H$. The potential V may have singularities at zero and infinity, but it is assumed to be continuous elsewhere; further, we assume that for r sufficiently large, $V(r)$ is monotonically increasing with r. The operator B is assumed to the multiplicative and given by

$$Bv(r) = b(r)v(r), \tag{4}$$

in which $b(r)$ is bounded on each finite interval but becomes unbounded as r approaches infinity.

The quantity $||B\Psi||^2$ can be expressed as

$$||B\Psi||^2 = \int_0^{r^*} b^2(r)\Psi^2(r)\ dr + \int_{r^*}^{\infty} b^2(r)\Psi^2(r)\ dr, \tag{5}$$

where r^* is defined by

$$r^* = \max\{r \mid V(r) \leq E^u\}. \tag{6}$$

The first integral of (5) is easily dominated[6] by the quantity b_0^2 defined by

$$b_0^2 = \sup_{0 \leq r \leq r^*} b^2(r),$$

and the second can be bounded by using estimates for the eigenfunction Ψ.

[3] A. Weinstein, Mém. Sci. Math. Fasc. **88** (Gauthier–Villars, Paris, 1937); N. Aronszajn, Proceedings of the Oklahoma Symposium on Spectral Theory and Differential Problems, Stillwater, Okla., 1950, (Dept. of Math. Oklahoma A and M, Stillwater, Okla, 1955), 179; H. F. Weinberger, Institute for Fluid Dynamics and Applied Mathematics, University of Maryland, College Park, Maryland Tech. Note BN-183 (1959); N. Bazley and D. W. Fox, Phys. Rev. **124**, 483 (1961); J. Math. Phys. **4**, 1147 (1963); J. G. Gay, Phys. Rev. **135**, A 1220 (1964).

[4] See T. Kato, J. Phys. Soc. Japan **4**, 334 (1949); H. F. Weinberger, J. Res. Natl. Bur. Std. **64B**, 217 (1960).

[5] A parallel argument shows that $|(B\Psi_\mu, \Psi_\nu) - (B\varphi_\mu, \varphi_\nu)|$ is bounded by $||\Psi_\mu - \varphi_\mu||\ ||B\Psi_\nu|| + ||\Psi_\nu - \varphi_\nu||\ ||B\varphi_\mu||$ and by $||\Psi_\mu - \varphi_\mu||\ ||B\varphi_\nu|| + ||\Psi_\nu - \varphi_\nu||\ ||B\Psi_\mu||$. Further, if B is positive and $||\Psi - \varphi||\ ||B\varphi|| \leq (B\varphi, \varphi)$, the inequality $|(B\Psi, \varphi)|^2 \leq (B\varphi, \varphi)(B\Psi, \Psi)$ can be combined with the second of (1) to yield the lower bound $(B\Psi, \Psi) \geq [(B\varphi, \varphi) - ||B\varphi||\ ||\Psi - \varphi||^2/(B\varphi, \varphi)$.

[6] Another estimate will be indicated later.

A minor modification of arguments of Titschmarsh[7] yields the estimate

$$0 < |\Psi(r)| \leq |\Psi(r^*)| \exp\left\{-\int_{r^*}^{r} [V(\rho) - E^u]^{\frac{1}{2}}\, d\rho\right\} \quad (8)$$

for all r greater than r^* under the assumption[8] that $V(r)$ is monotonically increasing with r beyond r^*. Thus the second integral of (5) is dominated by

$$\Psi^2(r^*) \int_{r^*}^{\infty} b^2(r) \exp\left\{-2\int_{r^*}^{r} [V(\rho) - E^u]^{\frac{1}{2}}\, d\rho\right\} dr. \quad (9)$$

Here we assume that $b(r)$ grows sufficiently slowly at infinity for the integral in (9) to converge. This is ensured if b grows no more rapidly than a power of r.

The error bound is completed by an estimate[9] for $\Psi^2(r^*)$. If $V(r)$ is bounded below by $V_{\min}$ and $\Psi(0) = 0$, then the following elementary estimates can be found without difficulty:

$$\Psi^2(r^*) \leq \tfrac{2}{3}(E - V_{\min})(r^*)^{\frac{3}{2}}, \quad \Psi^2(r^*) \leq (E - V_{\min})r^*,$$
$$\Psi^2(r^*) \leq 2(E - V_{\min})^{\frac{1}{2}}. \quad (10)$$

More generally, estimates can be found with the help of the resolvent of some related operator H^0. We suppose H^0 to be selfadjoint, expressed by

$$H^0 v = -v'' + V^0(r)v \quad (11)$$

on a suitably defined domain $\mathfrak{D}_{H^0}$ that contains Ψ and φ, and to have a known resolvent kernel $\mathcal{K}_0(r, \rho, E^0)$; that is,

$$(H^0 - E^0)^{-1}v(r) = \int_0^{\infty} \mathcal{K}_0(r, \rho, E^0)v(\rho)\, d\rho \quad (12)$$

for all v in $\mathfrak{L}^2(0, \infty)$ and all E^0 not in the spectrum of H^0. In terms of the approximating function $\varphi^*(r)$ defined by

$$\varphi^*(r) = \int_0^{\infty} \mathcal{K}_0(r, \rho, E^0) \times [V^0(\rho) - V(\rho) + E - E^0]\varphi(\rho)\, d\rho, \quad (13)$$

the value of $|\Psi(r^*)|$ can be estimated by

$$|\Psi(r^*)| \leq |\varphi^*(r^*)| + \Delta(r^*), \quad (14)$$

where $\Delta(r^*)$ is given by[10]

$$\Delta^2(r^*) = \int_0^{\infty} \mathcal{K}_0^2(r^*, \rho, E^0) \times [V^0(\rho) - V(\rho) + E - E^0]^2\, d\rho\, ||\varphi - \Psi||^2. \quad (15)$$

In terms of φ alone we can obtain the estimate

$$|\Psi(r^*)| \leq |\varphi(r^*)| + \Delta(r^*) + \Theta(r^*) \quad (16)$$

in which $\Theta(r^*)$ is given by

$$\Theta^2(r^*) = \int_0^{\infty} \mathcal{K}_0^2(r^*, \rho, E^0)\, d\rho\, ||(H - E)\varphi||^2. \quad (17)$$

The terms in which E appears can be estimated easily by use of bounds for E.

IV. EXAMPLE

In our example we choose H an explicitly resolvable operator, and we compare our bound for the error with the exact error itself.

We take H to be the Hamiltonian,

$$Hv = -v'' - (2/r)v, \qquad v(0) = 0,$$

and B the multiplicative operator $Bv(r) = rv(r)$. The smallest eigenvalue of H is given by $E = -1$, with corresponding eigenfunction Ψ by $\Psi = 2re^{-r}$; thus,

$$(B\Psi, \Psi) = 4\int_0^{\infty} r^3 e^{-2r}\, dr = \tfrac{3}{2},$$

and is the expected electron radius (atomic units) for a hydrogen atom in its ground state.

We consider approximating vectors of the form

$$\varphi(r) = 2\alpha^{\frac{3}{2}} re^{-\alpha r},$$

where α is a real parameter, and we apply our method to estimate $|(B\Psi, \Psi) - (B\varphi, \varphi)|$. In our calculations we choose for simplicity $E = -1$. The exact error is, of course, known and is given by

$$|(B\Psi, \Psi) - (B\varphi, \varphi)| = \tfrac{3}{2}\,|1 - 1/\alpha|.$$

The approximation (14) is used to estimate $\Psi(r^*)$, and the operator H^0 is chosen[11] as

$$H^0 v = -v'', \qquad v(0) = 0;$$

here $\mathcal{K}_0(r, \rho, -1)$ is given by

$$\mathcal{K}_0(r, \rho, -1) = \begin{cases} e^{-r}\sinh \rho, & 0 < \rho \leq r < \infty, \\ e^{-\rho}\sinh r, & 0 < r \leq \rho < \infty. \end{cases}$$

[7] E. Titschmarsh, *Eigenfunction Expansions*, Part I (Oxford University Press, Oxford, 1962), 2nd ed., pp. 107–110, 165.

[8] The restrictions on the monotonicity of V can be lightened and similar estimates obtained.

[9] Some of the estimates given here might be used to improve the bound given earlier for the first term on the right in (5). Demonstrations of these estimates are given in the Appendix.

[10] The convergence of this integral requires that the singularities of V^0 at zero and infinity do not differ too greatly from those of V.

[11] The domains of H and H^0 are characterized as those functions v, $v(0) = 0$, which are, together with their first derivatives, absolutely continuous and such that Hv belongs to $\mathfrak{H}$ and H^0v belongs to $\mathfrak{H}$, respectively. [See, for example, the discussion of boundary conditions given in T. Kato, *Quadratic Forms in Hilbert Spaces and Asymptotic Perturbation Series* (University of California, Berkeley, California, 1955), pp. 137–140].

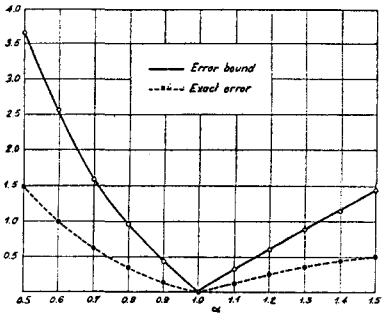

FIG. 1. Bound for $|(B\Psi, \Psi) - (B\varphi, \varphi)|$.

Explicit evaluation of the terms in our error bound gives

$$b_0^2 = 4, \qquad (B\varphi, B\varphi) = 3/\alpha^2,$$

$$||\Psi - \varphi||^2 = 2\,\frac{(1+\alpha)^3 - 8\alpha^{\frac{3}{2}}}{(1+\alpha)^3},$$

$$\Delta^2(2) = \left\{4e^{-4}\int_0^4 \frac{\sinh\rho}{\rho}\,d\rho - 8[\sinh 2]^2\int_4^\infty \frac{e^{-\rho}}{\rho}\,d\rho\right\}||\Psi - \varphi||^2,$$

$$|\varphi^*(2)| = 4\alpha^{\frac{3}{2}}\,\frac{e^{-2\alpha} - e^{-2}}{1-\alpha^2},$$

$$\int_{r^*}^\infty b^2(r)\exp\left\{-2\int_{r^*}^r [V(\rho)+1]^{\frac{1}{2}}\,d\rho\right\}dr = 4\int_2^\infty \frac{\exp[-2r(1-2/r)^{\frac{1}{2}}]}{[(1-2/r)^{\frac{1}{2}} - 1]^4}\,dr.$$

The last integral can be easily bounded from above.

We have plotted our error bound, together with the exact error, as a function of α in Fig. 1. A graph of the lower bound to $(B\Psi, \Psi)$ discussed in Ref. 5 is given in Fig. 2.

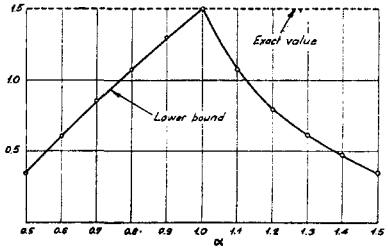

FIG. 2. Lower bound to $(B\Psi, \Psi)$.

APPENDIX

In this appendix we obtain the point estimates for Ψ used in Sec. III.

The estimates given in (10) suppose that Ψ vanishes at the origin. The first is that obtained by Titschmarsh,[7] p. 167, under slightly different hypothesis on V. The second is demonstrated for all positive r by

$$\Psi^2(r) = \left(\int_0^r \Psi'\,d\rho\right)^2 \le r\int_0^\infty (\Psi')^2\,d\rho \le r(E - V_{\min}),$$

and the third by

$$\Psi^2(r) = 2\int_0^r \Psi\Psi'\,d\rho \le 2\left\{\int_0^\infty (\Psi')^2\,d\rho\right\}^{\frac{1}{2}} \le 2(E - V_{\min})^{\frac{1}{2}}.$$

The demonstrations of the inequalities (14) and (16) start from the representation

$$\varphi(r) = \int_0^\infty \mathcal{K}_0(r, \rho, E^0)(H^0 - E^0)\varphi(\rho)\,d\rho$$

for each φ in $\mathfrak{D}_{H^0}$ as follows from (12). Further, for φ in $\mathfrak{D}_H$ as well, we may write

$$\varphi(r) = \int_0^\infty \mathcal{K}_0(r, \rho, E^0)(H - E)\varphi(\rho)\,d\rho + \varphi^*(r),$$

$$\varphi^*(r) = \int_0^\infty \mathcal{K}_0(r, \rho, E^0) \times [V^0(\rho) - V(\rho) + E - E^0]\varphi(\rho)\,d\rho;$$

and with Ψ in place of φ we have

$$\Psi(r) = \int_0^\infty \mathcal{K}_0(r, \rho, E^0) \times [V^0(\rho) - V(\rho) + E - E^0]\Psi(\rho)\,d\rho.$$

By applying Schwarz' inequality to the integral in the relationship

$$\Psi(r) = \varphi^*(r) + \int_0^\infty \mathcal{K}_0(r, \rho, E^0) \times [V^0(\rho) - V(\rho) + E - E^0][\Psi(\rho) - \varphi(\rho)]\,d\rho,$$

we obtain (14). Similarly, using Schwarz' inequality for the integrals in the relationship

$$\Psi(r) = \varphi(r) - \int_0^\infty \mathcal{K}_0(r, \rho, E^0)(H - E)\varphi(\rho)\,d\rho + \int_0^\infty \mathcal{K}_0(r, \rho, E^0)[V^0(\rho) - V(\rho) + E - E^0] \times [\Psi(\rho) - \varphi(\rho)]\,d\rho,$$

the inequality (16) is obtained.

On the Complex Structure of the Universe

EDMUND H. BROWN
Environment Science Services Administration,
Institute for Telecommunication Sciences and Aeronomy, Boulder, Colorado
(Received 16 June 1965)

A complex space–time is constructed from physical axioms. Both gravitational and electromagnetic fields are approximations to parts of a geometric object defined on this complex space, while other parts may represent strong and weak interactions. The intersections of the singularities of the three independent geometric invariants are identified as elementary particles. This identification leads to geometric definitions of mass and momentum and suggests the geometric significance of internal quantum numbers.

1. INTRODUCTION

THEORETICAL physics is encompassed by the problem of constructing an homeomorphic and isometric subspace of the universe which contains one image of every equivalence class of measureable physical experiments. This first paper considers only the simplest aspects of a subspace representable by a complex space–time manifold. However, the results lead to an understanding of the nature of the electromagnetic field and its relation to the gravitational field (that is, to a "unified field theory") and of the geometric structure of an "elementary particle." A sequel will consider the topology of this complex space–time and its singularities in much greater depth and may be useful in the interpretation of strong and weak interactions and the geometric significance of "internal" quantum numbers. Each critical axiom will be stated in such a form that it has well-defined converse (i.e., the assertion of the converse proposition), and the experimental observation which makes the converse inadmissible in a physical theory will be briefly noted.

2. TOPOLOGY

If Z is the subspace homeomorphic and isometric to the universe with elements $z \in Z$ and Z' is a second subspace, with T a transformation such that $Z' = TZ$, then Z' is homeomorphic to the universe if and only if T is bicontinuous (Ref. 1, p. 78).

If $w(Q)$ is a number-valued function and $w(Q) = w(Q')$, where $Q \subset Z$, $Q' \subset Z'$, $Q' = TQ$, and T is bicontinuous and isometric, then $w(Q)$ will be called a *physical* property; all other functions will be called *unphysical*. A transformation which does not *change* some physical property will be called unphysical; thus, if T is bicontinuous and isometric, it is unphysical.

Axiom 1 (Separability.): Z is a normal, separable space of finite dimension n.

By a theorem of Urysohn, that any regular separable space is metrizable, Kuratowski, p. 136,[1] Z admits a positive metric. Since physics is based on *measurement*, the existence of a metric on Z is clearly necessary.

By theorems of Hausdroff and Lavrientieff (Kuratowski,[1] p. 316), the space Z is homeomorphic and isometric to a subset $Z_{(i)}$ of a complete number-valued space of dimension n. This correspondence, or "imbedding", may impose a different metric on $Z_{(i)}$ from the Cartesian metric since the most general space of dimension n can be imbedded only in a Euclidean space of maximum dimension $2n + 1$ (Hurewicz and Wallman,[2] p. 60). The subset $Z_{(i)}$ will be called a *representation* of the space Z, and the *points* $z^i \in Z_{(i)}$, representations of the elements $z \in Z$.

If $Z_{(r)}$ is a similar representation of a second homeomorphic and isometric space Z' or, equivalently, if $Z_{(r)}$ is a second representation of Z, then any unphysical transformation $Z_{(i)} \to Z_{(r)}$, that is, any coordinate transformation $z^r = z^r(z^i)$, must be bicontinuous and isometric on $Z_{(i)}$ and $Z_{(r)}$. The following index conventions will be used throughout: (a) indices from the beginning of an alphabet designate one coordinate system; indices from the end, a second coordinate system; (b) the summation convention holds for diagonally repeated indices; (c) vector indices, Brown[3] designate higher-

[1] K. Kuratowski, *Topologie I* (Monografje Matematyczne, Warsaw, 1933).
[2] W. Hurewicz and H. Wallman, *Dimension Theory* (Princeton University Press, Princeton, New Jersey, 1941).
[3] E. H. Brown, Absolute Tensors of Arbitrary Order, J. Res. Natl. Bur. of Stds., **64** *B*, 99 (1960).

order objects, e.g., $w^{\mathbf{i}}(z^k)$ where $\mathbf{i}=\mathbf{i}_m=\{i_1, i_2, \cdots, i_m\}$, $i_1, i_2, \cdots, i_m = 1, \cdots, n$; (d) diagonally-repeated vector indices imply summation of terms for all states of the index vectors; full vertical bars indicate the covariant derivative, e.g., $w^{\mathbf{i}} \mid_k$.

Axiom 2 (Complexity): The points $z^{\mathbf{i}} \in Z_{(\mathbf{i})}$ of the representation space are n-sets of *complex numbers*.

If the points were always real, then reflection of the "space" coordinates alone on all of $Z_{(\mathbf{i})}$ to all of $Z_{(\mathbf{r})}$ (parity) would be bicontinuous and isometric and, thus, unphysical; then, "weak" interactions would be invariant under such a transformation in contradiction to observation.

Axiom 3 (Differentiability): Every representation space $Z_{(\mathbf{i})}$ admits a differential structure.

Since the set of all differentials $dz^{\mathbf{i}}$ at a fixed element z forms a linear space, the coordinate transformation $z^{\mathbf{r}} = z^{\mathbf{r}}(z^{\mathbf{i}})$ imposes a *linear* transformation on the differentials, $dz^{\mathbf{r}} = z^{\mathbf{r}}_{\mathbf{i}}(z)dz^{\mathbf{i}}$. (More generally, a locally Euclidean group is isomorphic to a Lie group, Montgomery and Zippin,[4] p.184.) The coefficients $z^{\mathbf{r}}_{\mathbf{i}}$ define partial derivatives, $\partial z^{\mathbf{r}}/\partial z^{\mathbf{i}} = z^{\mathbf{r}}_{\mathbf{i}}$, and thus, by the generalization of Cauchy's formula, Behnke and Thullen,[5] p. 40, imply partial derivatives of all orders $z^{\mathbf{r}}_{\mathbf{i}} = z^{\mathbf{r}}_{i_1 i_2} \cdots$. Accordingly, all coordinate transformations are represented by *analytic* functions. Since bicontinuity implies that these functions can have neither multiple values nor essential singularities, all coordinate transformations $z^{\mathbf{r}} = z^{\mathbf{r}}(z^{\mathbf{i}})$ must be meromorphic.

Axiom 4 (Symmetry): If $Z_{(\mathbf{i})}$ is a representation of Z, then the complex conjugate space $Z_{(\mathbf{r})} = \bar{Z}_{(\mathbf{i})}$ is also a representation of Z.

The comment on Axiom 3 suggests that the imaginary part of $Z_{(\mathbf{i})}$ is associated with particle–antiparticle character. The present axiom reinforces this association by a requirement that particle–antiparticle character be determined by a *relation* between points. Equivalently, whether an object is a particle or an antiparticle depends on an arbitrary assignment of some other object (this axiom could be called a principle of discrete relativity).

Theorem 1 (Product Space): A representation space $Z_{(\mathbf{i})}$ of the universe is a Cartesian product of two conjugate complex spaces $Z_{(\alpha)}$ and $Z_{(\dot{\alpha})}$ of equal dimension ν, and $n = 2\nu$.

The proof follows from Axioms 3 and 4: the transformations $dz^{\mathbf{i}} \to dz^{\mathbf{r}} = d\bar{z}^{\mathbf{i}}$ must be analytic; since by Axiom 2, $Z_{(\mathbf{i})}$ is not everywhere real, this can be true if and only if $dz^{\mathbf{i}} = \{d\zeta^{\alpha}, d\zeta^{\dot{\alpha}}\}$, or $z^{\mathbf{i}} = \{\zeta^{\alpha}, \zeta^{\dot{\alpha}}\}$, and $\zeta^{\dot{\alpha}} = \overline{\zeta^{\alpha}}$, $\alpha = 1, \cdots, \nu$, which implies that $dz^{\mathbf{i}} \to d\bar{z}^{\mathbf{i}}$ merely interchanges dz^{α}, $dz^{\dot{\alpha}}$.

The transformation $z^{\mathbf{r}}_{\mathbf{i}}(z^k)$ then separates into the analytic transformations $z^{\mathbf{r}}_{\mathbf{i}} = \{z^{\beta}_{\alpha}(z^{\beta}), z^{\dot{\beta}}_{\dot{\alpha}}(z^{\dot{\beta}})\}$. The two ν-dimensional spaces $Z_{(\alpha)}$ and $Z_{(\dot{\alpha})}$ are complex *analytic manifolds* (Steenrod,[6] p. 209), and each has an n-dimensional quasi-complex real representation in $X_{(\alpha)} \times Y_{(\alpha)}$. Also, the complex dimension and real dimension of $Z_{(\mathbf{i})}$ are equal, and $Z_{(\mathbf{i})}$ has a real representation in $X_{(\alpha)} \times Y_{(\alpha)}$.

Defining $z_{(\mathbf{r})} = \det z^{\mathbf{r}}_{\mathbf{i}}$, then $z_{(\mathbf{r})} = \zeta_{(\rho)}\bar{\zeta}_{(\dot{\rho})} \geq 0$, with equality occurring only at the isolated poles of $z^{\mathbf{r}}(z^{\mathbf{i}})$; thus, $Z_{(\mathbf{i})}$ is an *orientable* manifold and can be visualized as a two-sided surface, one side corresponding to the subspace $Z_{(\alpha)}$ and the other to the conjugate subspace $Z_{(\dot{\alpha})}$. Also, it is clear that the connection between a particle and its corresponding antiparticle will be an *order relation* of the type $\{t^{\alpha}, t^{\dot{\alpha}}\} \to \{t^{\dot{\alpha}}, t^{\alpha}\}$.

The relations $Z_{(\alpha)} = P_{\alpha}Z_{(\mathbf{i})}$, $Z_{(\dot{\alpha})} = P_{\dot{\alpha}}Z_{(\mathbf{i})}$, $X_{(\alpha)} = \operatorname{Re} Z_{(\mathbf{i})}$, and $Y_{(\alpha)} = \operatorname{Im} Z_{(\mathbf{i})}$ define four projections of $Z_{(\mathbf{i})}$. Since $P_{\alpha} \times P_{\dot{\alpha}} = I$, $P_{\dot{\alpha}}$ is the *adjoint* of P_{α}; also, $Z_{(\mathbf{i})}$ can be called a self-adjoint manifold. The operator P_{α} corresponds to *interpreting* $z^{\mathbf{i}}$ as a point, and $P_{\dot{\alpha}}$ to interpreting $z^{\mathbf{i}}$ as an "antipoint." Since P_{α} and $P_{\dot{\alpha}}$ are continuous operators (Kuratowski,[1] p. 227), if $z^k = cz^{\mathbf{i}}$, $\zeta^{\alpha} = P_{\alpha}z^{\mathbf{i}}$, and $\zeta^{\gamma} = P_{\alpha}z^k$, where c is an arbitrary constant, then $\zeta^{\gamma} = c\zeta^{\alpha}$ and $P_{\alpha}cz^{\mathbf{i}} = cP_{\alpha}z^{\mathbf{i}}$. Then $\zeta^{\dot{\gamma}} = P_{\dot{\alpha}}z^k = P_{\dot{\alpha}}cz^{\mathbf{i}} = cP_{\dot{\alpha}}z^{\mathbf{i}} = c\zeta^{\dot{\alpha}}$ and $\zeta^{\dot{\gamma}} = \overline{\zeta^{\gamma}} = \overline{c\zeta^{\alpha}} = \bar{c}\zeta^{\dot{\alpha}}$. Therefore, all multiplicative constants defined on $Z_{(\mathbf{i})}$ must be real.

Let $z^a = \delta^a_{\mathbf{i}}z^{\mathbf{i}} + \delta^a_{\mathbf{i}}c^{\mathbf{i}}$ be a coordinate translation with an arbitrary set of constants $c^{\mathbf{i}}$; and let $z^{\mathbf{r}} = A^{\mathbf{r}}_{ab}z^a z^b$ be a symmetric transformation in which, by the previous result, the coefficients $A^{\mathbf{r}}_{ab}$ must be real. Since $z^{\mathbf{r}} = A^{\mathbf{r}}_{\mathbf{i}\mathbf{j}}z^{\mathbf{i}}z^{\mathbf{j}} + 2A^{\mathbf{r}}_{\mathbf{i}\mathbf{j}}c^{\mathbf{i}}z^{\mathbf{j}} + A^{\mathbf{r}}_{\mathbf{i}\mathbf{j}}c^{\mathbf{i}}c^{\mathbf{j}}$ is a coordinate transformation, the coefficients $A^{\mathbf{r}}_{\mathbf{i}\mathbf{j}}c^{\mathbf{i}}$ are real, which implies the additive constants $c^{\mathbf{i}}$ are real. Combining the two results, *all* physical constants must be real. Also, the set of all unphysical transformations is a proper subset of the class of all analytic transformations: the coefficients of a Taylor or Laurent expansion of an unphysical trans-

[4] D. Montgomery and L. Zippin, *Topological Transformation Groups* (Interscience Publishers, Inc., New York, 1955).

[5] H. Behnke and P. Thullen, Theorie der Funktionen Mehrer Komplexer Veranderlichen, *Ergebnisse der Mathematik III*, (Julius Springer–Verlag, Berlin, 1934).

[6] N. Steenrod, *Topology of Fiber Bundles* (Princeton University Press, Princeton, New Jersey, 1951).

formation must be real. The following theorem is an immediate consequence.

Theorem 2 (Imaginary Origin): The subset $X \subset Z$ whose representation is $X_{(\alpha)}$ is invariant under unphysical transformations of $Z_{(i)}$ and, thus, X is both a proper subspace and a retract of Z.

Theorem 3 (Reality): All differentiable scalar functions $w(z^i)$ defined on $Z_{(i)}$ must be *real.*

Proof. by Axiom 4, if $u(x^\alpha, y^\alpha) = \text{Re } w(z^i)$, then $dw = \frac{1}{2}\partial_i w dz^i = \frac{1}{2}(\partial_\alpha w dz^\alpha + \partial_{\bar\alpha} w dz^{\bar\alpha}) = (dx^\alpha \partial_{x\alpha} + dy^\alpha \partial_{y\alpha})u = du$; thus, $w(z^i)$ can differ from a real function only by an additive constant. But, from the previous paragraph, this constant also must be real and, therefore, $w(z^i)$ is real.

Since the imaginary part of an analytic scalar is determined to within a constant of integration by its real part, the relations $w(z^i) = \text{Re } \omega(\zeta^\alpha) = \text{Re } \bar\omega(\zeta^{\bar\alpha})$ define two conjugate functions $\omega(\zeta^\alpha)$ and $\bar\omega(\zeta^{\bar\alpha})$ analytic on $Z_{(\alpha)}$ and $Z_{(\bar\alpha)}$, respectively; thus, any scalar $w(z^i)$ can be written in the form $w(z^i) = \frac{1}{2}[\omega(\zeta^\alpha) + \bar\omega(\zeta^{\bar\alpha})]$. Analyticity requires that $\omega(\zeta^\alpha)$ and $\bar\omega(\zeta^{\bar\alpha})$ satisfy the Cauchy–Riemann equations $\partial_{\bar\alpha}\omega = \partial_\alpha \bar\omega = 0$ and, thus, $w(z^i)$ and $\text{Im } \omega(\zeta^\alpha) = -\text{Im } \bar\omega(\zeta^{\bar\alpha})$ are *harmonic* on $X_{(\alpha)} \times Y_{(\alpha)}$. In particular, $w(x^\alpha, y^\alpha)$ must be an *even* function of y_α.

In a generalized sense, $w(z^i)$ is an analytic function of z^i since all of its derivatives $\partial_k w(z^i)$ exist and it always satisfies $\partial_{\bar\alpha}\partial_\beta w(z^i) = 0$; and, thus it is the property of analyticity that restricts w to the form $w = \frac{1}{2}(\omega + \bar\omega)$ rather than the exponentially corresponding function $w' = (\omega'\bar\omega')^{\frac{1}{2}}$ which is real but not analytic. Also, a study of scalars on Z cannot be reduced to a mere study of harmonic functions on $X_{(\alpha)} \times Y_{(\alpha)}$ since the only physical scalars will be topological or geometric invariants. In particular, the nonconstant geometric invariants, such as result from contraction of high-order tensors, will introduce a more complicated structure than that of harmonic functions defined *a priori* on a real space.

Since Urysohn's theorem affirms the constructability of a positive metric on Z, such a metric can be used to define *bounded* subsets of Z. For a bounded subset $Q \subset Z$, if a frontier of Q exists, this frontier must also be physical (that is, definable by an analytic relation on $Z_{(i)}$). But the frontier of any set of dimension n must have dimension $n - 1$ (Hurewicz and Wallman,[2] p.48), and thus cannot be complex analytic. Thus, no bounded, physical, infinite subset $Q \subset Z$ possesses a connected frontier. However, in any locally Euclidean space (a space for which a neighborhood of a point can be mapped isometrically on a neighborhood of the origin of R^n), sufficiently small infinite subsets exist which are closed and possess connected frontiers. The property of being not locally Euclidean may arise because of an indefinite metric or because the space is not simply connected; but, for Z the first reason cannot apply. Since X is an invariant subspace of Z, the next theorem follows immediately.

Theorem 4 (Compactness): Neither Z nor its real invariant subspace X are locally Euclidean, and Z is not simply connected.

Since none of the subsets $Y \subset Z$ defined by $x^i = \text{constant}$ can be locally Euclidean, then $X_{(\alpha)}$ or $Y_{(\alpha)}$, or both, must be "closed" (i.e., compact, or closed in the same sense as for surfaces). In the next section, a study of the Riemannian curvature will prove that if X is closed then Y is open, and conversely. But, if Y is noncompact, a fixed Y-origin cannot be introduced by the topology, that is, X cannot be an invariant subspace of Z for the Brouwer fixed-point theorem does not hold on a noncompact space; conversely, possession of an isometry-invariant point is characteristic of compact Riemannian spaces (Helgason,[7] p. 241 ff. or Dunford and Schwartz,[8] p. 459). Then, X must be open and, since it cannot be locally Euclidean, must have an indefinite metric.

Theorem 5 (Metric): The subspace $X \subset Z$ must be open (i.e., noncompact) and the subsets Y must be closed (i.e., compact); the *physical* metric on X induced by the positive-nondefinite physical metric of Z must be *indefinite.*

Although X must be open, a bounded subset $Q \subset Z$ can be chosen arbitrarily to be compact in $X_{(\alpha)}$ and, thus, in both $X_{(\alpha)}$ and $Y_{(\alpha)}$. Then, Q is representable by a topological product of quasi-complex spheres in $X_{(\alpha)} \times Y_{(\alpha)}$, that is, spheres representable by complex spheres in $Z_{(\alpha)}$ or $Z_{(\alpha)}$, of which at least two must differ in dimension since the X-metric is indefinite (Steenrod,[6] p. 207). Thus a theorem of A. Borel and Serre, see Steenrod,[6] p. 219, (that the only spheres which are quasi-complex have complex dimension 1 and 3) implies the inequalities $\nu \geq 4$ and $n \geq 8$. For the remainder of this study the *equalities* $\nu = 4$ and $n = 8$ will

[7] S. Helgason, *Differential Geometry and Symmetric Spaces,* (Academic Press Inc., New York, 1962).

[8] N. Dunford, and J. T. Schwartz, *Linear Operators* (Interscience Publishers, Inc., New York, 1958).

be introduced on empirical grounds. An important comment may be made that, since the restriction of the group of isometries on Z to a subset Y is a finite Lie group, the isometries on Y will lead to discrete "internal quantum numbers."

3. GEOMETRY

On any Riemannian space the metric will have the form of an integral along a geodesic of a line element defined by a symmetric quadratic form. Thus, the line element dx on $Z_{(i)}$ may be written in terms of the quadratic form

$$ds^2 = \tfrac{1}{2}g_{ij}(z^k)\, dz^i\, dz^j, \qquad (1)$$

in which the *metric tensor* g_{ij} separates into four blocks, $g_{ij} = \{g_{\alpha\beta}, g_{\dot\alpha\dot\beta}, g_{\alpha\dot\beta}, g_{\dot\alpha\beta}\}$. Since ds^2 is real and dz^i, self-adjoint, i.e., $\overline{d\zeta^\alpha} = d\zeta^{\dot\alpha}$, g_{ij} must be also self-adjoint. Then both of the partial quadratic forms $g_{\dot\alpha\dot\beta}d\zeta^{\dot\alpha}d\zeta^{\dot\beta} + g_{\alpha\dot\beta}d\zeta^\alpha d\zeta^{\dot\beta}$ and $g_{\alpha\dot\beta}d\zeta^\alpha d\zeta^{\dot\beta} + g_{\dot\alpha\dot\beta}d\zeta^{\dot\alpha}d\zeta^{\dot\beta}$ are also real. Mathematical simplicity led Kaehler to choose $g_{ij} = \{g_{\dot\alpha\beta}, g_{\alpha\dot\beta}\}$, with $g_{\alpha\beta} = g_{\dot\alpha\dot\beta} = 0$, as a metric tensor, and a number of mathematicians have conducted extensive research on manifolds with Kaehler metrics. But, on a complex manifold, $g_{\dot\alpha\beta}$ and $g_{\alpha\dot\beta}$ satisfy the equations

$$g_{\dot\alpha\beta} = \partial_{\dot\alpha}\partial_\beta w(\zeta^k), \qquad g_{\alpha\dot\beta} = \partial_\alpha\partial_{\dot\beta} w(\zeta^k),$$

(Yano and Bochner,[9] p. 123). Thus, by Theorem 3. on the reality of physical scalars, $g_{\dot\alpha\beta} = g_{\alpha\dot\beta} = 0$ on the physical space $Z_{(i)}$. The following theorem is an immediate consequence.

Theorem 6 (Line Element): The line element on $Z_{(i)}$ is defined by the quadratic form

$$ds^2 = \tfrac{1}{2}\{g_{\alpha\beta}(\zeta^\gamma)\, d\zeta^\alpha\, d\zeta^\beta + g_{\dot\alpha\dot\beta}(\zeta^{\dot\gamma})\, d\zeta^{\dot\alpha}\, d\zeta^{\dot\beta}\}, \qquad (2)$$

where $g_{\alpha\beta}(\zeta^\gamma)$ is a symmetric tensor analytic on $Z_{(\alpha)}$ and $g_{\dot\alpha\dot\beta}(\zeta^{\dot\gamma})$, a symmetric tensor analytic on $Z_{(\dot\alpha)}$. Equivalently, the line element on $X_{(\alpha)} \times Y_{(\alpha)}$ is defined by

$$ds^2 = g^{\mathrm{Re}}_{\alpha\beta}(x^\gamma, y^\gamma)[dx^\alpha\, dx^\beta - dy^\alpha\, dy^\beta] - 2g^{\mathrm{Im}}_{\alpha\beta}(x^\gamma, y^\gamma)\, dx^\alpha\, dy^\beta. \qquad (2a)$$

On Z, the *metric* (or distance function) for two arbitrary elements z_1, z_2, will be given by $d(z_1, z_2) = \int_{z_1 C}^{z_2} ds$, where ds is the line element of Theorem 6 and C is a geodesic common to z_1 and z_2. Since the metric on X is indefinite of signature 3, the "tensor" $g^{\mathrm{Re}}_{\alpha\dot\beta}$ also will be indefinite of signature 3. But, the imaginary "tensor" $g^{\mathrm{Im}}_{\alpha\dot\beta}$ can be obtained from $g^{\mathrm{Re}}_{\alpha\dot\beta}$ by an integration of the Cauchy–Riemann equations, which leaves the signature unchanged. Therefore, $g_{\alpha\beta}$, $g_{\dot\alpha\dot\beta}$, and g_{ij} must be all indefinite of signature 3. Then, the fact that the metric $d(z_1, z_2)$ on Z must be positive-nondefinite leads to the following theorem.

Theorem 7 (Geodesic): If two elements x_1, x_2 in the subspace X are separated by a spacelike coordinate interval, a geodesic C between them cannot lie entirely in X; if two elements y_1, y_2 in a subset Y are separated by a timelike coordinate interval, a geodesic between them cannot be entirely in Y; thus, all geodesics entirely in X must be timelike, and all geodesics entirely in Y must be spacelike.

Since the indefinite metric on X cannot be a function of y^i, that is, $d(x_1, x_2 + \delta x) - d(x_1, x_2)$ must depend only on δx, a geodesic between two elements x_1, x_2 in X separated by a spacelike coordinate interval must travel completely around a compact subset Y. Thus, the y-component of the path will contribute only a constant to the metric, and this constant will disappear in the comparison of two spacelike intervals or in the definition of distance by reference to a standard length. For two elements y_1, y_2 in Y separated by a timelike coordinate interval, the fact that X is open implies that a physical metric cannot be defined on Y alone since the distance must always depend on the x-component of the path.

For a positive-nondefinite metric, the axiom of *definiteness*, $d(z_1, z_2) = 0 \leftrightarrows z_1 = z_2$, Dunford and Schwartz,[8] p. 18, must be dropped. Frequently, the implication $z_1 = z_2 \rightarrow d(z_1, z_2) = 0$ is retained alone but, on Z, *both* this and the inverse implication must be discarded. For any element $z \in Z$ is connected with itself by a completely spacelike geodesic over a subset Y and, thus, $d(z, z) = kL$, where L is a constant and k an integer. Clearly, the remaining metric axioms remain unchanged.

A separable metric manifold will possess an affine structure, Helgason,[7] p. 84, and a unique connection which is torsion-free will always exist (Kobayashi,[10] p. 158). Thus, at an arbitrary point in $Z_{(i)}$, a coordinate system z^a can be chosen such that, for any arbitrary tensor $t^a(z^b)$ the covariant differential is equal to the ordinary differential $Dt^a = dt^a$. Then, writing $z^a_i = z^{a_1}_{i_1} z^{a_2}_{i_2} \cdots z^{a_m}_{i_m}$, in an arbitrary system z^i at the same point,

$$Dt^i = dt^i + t^j z^a_{jk} z^i_a\, dz^k,$$

[9] K. Yano, and S. Bochner, *Curvature and Betti Numbers*, (Princeton University Press, Princeton, New Jersey 1953).

[10] S. Kobayashi, and K. Nomizu, *Differential Geometry* (Interscience Publishers, Inc., New York, 1963).

or, defining the covariant derivative $t^i\ |_k = Dt^i/dz^k$ and the affine connection $\Gamma^i_{.jk} = z^a_{jk}z^i_a$ at the point,

$$t^i\ |_k = \partial_k t^i + t^j\Gamma^i_{.jk}. \tag{3}$$

Applying (3) to the metric condition $g_{ij}|_k = 0$, where $g_{ij} = g_{i_mj_m} = g_{i_1j_1}g_{i_2j_2}\cdots g_{i_mj_m}$, leads immediate to

$$\partial_k g_{ij} = \Gamma_{ijk} + \Gamma_{jik}, \tag{4}$$

where Γ_{ijk} is defined by $g_{ih}\Gamma^h_{.jk}$.

Since Γ_{ijk} is determined by $z^i_r = \{z^\alpha_\rho, z^{\bar\alpha}_{\bar\rho}\}$ and $g_{ij} = \{g_{\alpha\beta}, g_{\bar\alpha\bar\beta}\}$, the affine connection $\Gamma^i_{.jk}$, the Riemann–Christoffel curvature tensor $R^i_{.jkh}$, and the Ricci tensor $R_{ik} = R^h_{.ihk}$ must also separate into analytic adjoint blocks,

$$\Gamma^i_{.jk}(z^k) = \{\Gamma^\alpha_{.\beta\gamma}(z^\delta), \Gamma^{\bar\alpha}_{.\bar\beta\bar\gamma}(z^{\bar\delta})\}$$
$$R^i_{.jkh}(z^l) = \{R^\alpha_{.\beta\gamma\delta}(z^\epsilon), R^{\bar\alpha}_{.\bar\beta\bar\gamma\bar\delta}(z^{\bar\epsilon})\}$$
$$R_{ik}(z^h) = \{R_{\alpha\beta}(z^\gamma), R_{\bar\alpha\bar\beta}(z^{\bar\gamma})\}.$$

Also, since $z^\rho(z^\alpha)$ is analytic, $z^\rho_{\alpha\beta} = z^\rho_{\beta\alpha}$, which implies the affine connection for $m = 1$, is symmetric, $\Gamma^i_{.jk} = \Gamma^i_{.kj}$.

In General Relativity, Einstein postulated the indistinguishability of accelerations due to gravitational fields and accelerations due to curvature of space–time. Here, the "equivalence principle" will be taken to its limit. Later, the manner in which the concept of the trajectory of a particle moving in a field is abstracted from a more general geometric structure will become clear.

Axiom 5 (Equivalence): No forces exist in nature. Then, the world lines of all particles must be geodesics. Apparent forces or force fields are merely interpretations of the components of the complex curvature of space–time. If the unit vector $w^i(z^k) = dz^i/ds = \{w^\alpha(\zeta^\gamma), w^{\bar\alpha}(\zeta^{\bar\gamma})\}$ is the complex 4-velocity, the equations of a geodesic (the equations of motion of a particle) are then

$$dw^i/ds + \Gamma^i_{.jk}w^jw^k = 0, \tag{5}$$

or

$$dw^\alpha/ds + \Gamma^\alpha_{.\beta\gamma}w^\beta w^\gamma = 0, \tag{6}$$
$$dw^{\bar\alpha}/ds + \Gamma^{\bar\alpha}_{.\bar\beta\bar\gamma}w^{\bar\beta}w^{\bar\gamma} = 0.$$

Equivalently, in $X_{(\alpha)} \times Y_{(\alpha)}$ these equations may be written in terms of real and imaginary parts as

$$du^\alpha/ds + \Gamma^{\alpha\mathrm{Re}}_{.\beta\gamma}(u^\beta u^\gamma - v^\beta v^\gamma) - 2\Gamma^{\alpha\mathrm{Im}}_{.\beta\gamma}u^\beta v^\gamma = 0, \tag{7a}$$
$$dv^\alpha/ds + \Gamma^{\alpha\mathrm{Im}}_{.\beta\gamma}(u^\beta u^\gamma - v^\beta v^\gamma) + 2\Gamma^{\alpha\mathrm{Re}}_{.\beta\gamma}u^\beta v^\gamma = 0, \tag{7b}$$

where $\Gamma^\alpha_{.\beta\gamma} = \Gamma^{\alpha\mathrm{Re}}_{.\beta\gamma} + i\Gamma^{\alpha\mathrm{Im}}_{.\beta\gamma}$ and $w^\alpha = u^\alpha + iv^\alpha$. Since $u_\alpha = g^{\mathrm{Re}}_{\alpha\beta}u^\beta - g^{\mathrm{Im}}_{\alpha\beta}v^\beta$ and $v_\alpha = g^{\mathrm{Re}}_{\alpha\beta}v^\beta + g^{\mathrm{Im}}_{\alpha\beta}u^\beta$, the subsidiary condition that w^i is a unit vector, $\frac{1}{2}w_iw^i = 1$, can be written in the simple form $u_\alpha u^\alpha - v_\alpha v^\alpha = 1$.

In a fixed system, the relativistic equations of motion may be considered equivalent to the classical equations plus an apparent unit force, $-\Gamma^\alpha_{\beta\gamma}u^\beta u^\gamma$, but the classical *form* of this force (that derived from the Newtonian potential) can be only an *approximation*. On $Z_{(i)}$, however, even equivalence depends on approximation, since exact solutions $u^\alpha(s)$, $v^\alpha(s)$ are coupled by the pair of equations (7). Also, separation into real and imaginary parts is a covariant procedure only for scalars.

If powers of v^α can be neglected, (7a) reduces to

$$du^\alpha/ds + \Gamma^{\alpha\mathrm{Re}}_{.\beta\gamma}u^\beta u^\gamma - 2\Gamma^{\alpha\mathrm{Im}}_{.\beta\gamma}u^\beta v^\gamma = 0. \tag{8}$$

The similarity to the general relativistic equations of motion with electromagnetic field (Bergmann,[11] p. 193),

$$du^\alpha/ds + \Gamma^\alpha_{.\beta\gamma}u^\beta u^\gamma - e/mc^3\, F^\alpha_{.\beta}u^\beta = 0, \tag{9}$$

suggests that $F^\alpha_{.\beta}e/mc^3$ is a classical approximation to $2\Gamma^{\alpha\mathrm{Im}}_{.\beta\gamma}v^\gamma$, and that v^α (or, possibly, only its time-like component) is related to charge.

For any physical object $w^\alpha = u^\alpha + iv^\alpha$, y-symmetry implies $[\partial_{y\gamma}u^\alpha]_{y=0} = 0$, thus $[\partial_{x\gamma}v^\alpha]_{y=0} = 0$ for all x^γ, or $v^\alpha =$ constant on $X_{(\alpha)}$. But the only constants are scalars, linear combinations of δ^α_β, and zero tensors. Thus, on $X_{(\alpha)}$: $y^\alpha = 0$, $\partial_{x\alpha}y^\rho = 0$, $y^\rho_{\alpha\beta} = 0$, $g^{\mathrm{Im}}_{\alpha\beta} = 0$, $\Gamma^{\mathrm{Im}}_{\alpha\beta\gamma} = 0$, $R^{\mathrm{Im}}_{\alpha\beta} = 0$; also, for the 4-velocity, $v^\alpha =$ constant, and $dv^\alpha/ds = 0$. Then, for $y^\alpha \neq 0$, but sufficiently small, v^α is approximately constant. The usual affine transformation law now separates into

$$\Gamma^{\rho\mathrm{Re}}_{.\sigma\tau} \approx \Gamma^{\alpha\mathrm{Re}}_{.\beta\gamma}x^\rho_\alpha x^\beta_\sigma x^\gamma_\tau + x^\alpha_{\sigma\tau}x^\rho_\alpha, \tag{10}$$
$$\Gamma^{\rho\mathrm{Im}}_{.\sigma\tau} \approx \Gamma^{\alpha\mathrm{Im}}_{.\beta\gamma}x^\rho_\alpha x^\beta_\sigma x^\gamma_\tau,$$

that is, $\Gamma^{\alpha\mathrm{Im}}_{.\beta\gamma} \approx g^{\alpha\delta}_{\mathrm{Re}}\Gamma^{\mathrm{Im}}_{\delta\beta\gamma}$ transforms *approximately* as a tensor.

Writing

$$\Gamma^{\mathrm{Im}}_{\alpha\beta\gamma} = \Gamma^{\mathrm{Im}}_{(\alpha\beta)\gamma} + \Gamma^{\mathrm{Im}}_{[\alpha\beta]\gamma}$$
$$= \tfrac{1}{2}\partial_{x\gamma}g^{\mathrm{Im}}_{\alpha\beta} - \tfrac{1}{2}(\partial_{x\alpha}g^{\mathrm{Im}}_{\beta\gamma} - \partial_{x\beta}g^{\mathrm{Im}}_{\alpha\gamma}),$$

then in $\Gamma^{\mathrm{Im}}_{\alpha\beta\gamma}v^\gamma$ the term $\frac{1}{2}\partial_{x\gamma}g^{\mathrm{Im}}_{\alpha\beta}v^\gamma$ is equivalent to

$$-\frac{1}{2}\partial_{v\gamma}g^{\mathrm{Re}}_{\alpha\beta}v^\gamma = -\frac{1}{2}\frac{d}{ds}g^{\mathrm{Re}}_{\alpha\beta} + \frac{1}{2}u^\gamma\partial_{x\gamma}g^{\mathrm{Re}}_{\alpha\beta} \approx 0.$$

Thus, $\Gamma^{\mathrm{Im}}_{\alpha\beta\gamma}v^\gamma = \Gamma^{\mathrm{Im}}_{[\alpha\beta]\gamma}v^\gamma$, or $\Gamma^{\mathrm{Im}}_{\alpha\beta\gamma}v^\gamma$ is *approximately* skew-symmetric.

[11] P. G. Bergmann, *Introduction to the Theory of Relativity* (Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1946).

Using the approximate tensor character and approximate skew-symmetry of

$$\Gamma^{\alpha\,\mathrm{Im}}_{\cdot\beta\gamma}v^\gamma \approx g^{\alpha\delta}_{\mathrm{Re}}\Gamma^{\mathrm{Im}}_{\delta\beta\gamma}v^\gamma$$

leads immediately to the equation

$$R^{\mathrm{Im}}_{\alpha\beta}v^\beta \approx (\Gamma^{\gamma\,\mathrm{Im}}_{\cdot\alpha\beta}v^\beta\,|_{*\gamma}, \tag{11}$$

which appears similar to the Maxwell equation with current, (Barut,[12] p. 94)

$$j_\alpha = cF^\gamma_{\cdot\alpha}\,|_{*\gamma}. \tag{12}$$

The approximation $j_\alpha \approx R^{\mathrm{Im}}_{\alpha\beta}v^\beta$ can be valid only if the equation for the dual, $F^{\gamma^*}{}_\alpha|_{*\gamma} = 0$ (which is not an identity in a nonflat space), or equivalently the continuity equation $j^\alpha|_{*\alpha} = 0$, is an additional result. Writing the contracted Bianchi identity $G^\gamma_{\cdot\alpha}|_\gamma = 0$ in the form $(G^\gamma_{\cdot\alpha}w^\alpha|_{*\gamma} = 0$, then the three conditions $R^{\mathrm{Im}} \approx g^{\alpha\beta}_{\mathrm{Re}}R^{\mathrm{Im}}_{\alpha\beta} \approx 0$, $(G^{\gamma\,\mathrm{Re}}_{\cdot\alpha}u^\alpha|_{*\gamma} \approx 0$, and approximate skew-symmetry, imply

$$(G^\gamma_{\cdot\alpha}w^\alpha\,|_{*\gamma} - (G^{\gamma\,\mathrm{Re}}_{\cdot\alpha}u^\alpha\,|_{*\gamma} \approx -(R^{\gamma\,\mathrm{Im}}_{\cdot\alpha}v^\alpha\,|_{*\gamma} \approx 0, \tag{13}$$

which confirms that $j^\gamma = R^{\gamma\,\mathrm{Im}}_{\cdot\alpha}v^\alpha$ appears as a divergenceless 4-current density in the limiting approximation. Thus, the "electromagnetic field" is an approximation to the imaginary part of the affine connection on $Z_{(i)}$, and the two Maxwell's equations are approximations to "field equations" satisfied by this affine connection.

To clarify the nature of the approximation involved, an additional axiom may be included:

Axiom 6 (Asymptotic): At large real spacelike coordinate intervals between particles $g_{\alpha\beta}(z^\gamma) \to T\mathcal{G}_{\alpha\beta}$, where T is an orthogonal transformation and $\mathcal{G}_{\alpha\beta}$ is the Lorentz matrix

$$\mathcal{G}_{\alpha\beta} = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix}.$$

Then, since $\lim \Gamma^\alpha_{\cdot\beta\gamma}$ can be taken as zero, $g^{\mathrm{Im}}_{\alpha\beta}$, $\Gamma^{\alpha\,\mathrm{Im}}_{\cdot\beta\gamma}$, etc., are zero for all y^α at large x. Thus, the accuracy of the approximation for fixed y^α will increase with the *distance*.

A second type of approximation can be obtained by assuming $u_\alpha u^\alpha \to 1$ in the equation $u_\alpha u^\alpha - v_\alpha v^\alpha = 1$. However, $u_\alpha u^\alpha$ is an invariant only on $X_{(\alpha)}$ and, thus, $u_\alpha u^\alpha \to 1$ results in a physical approximate theory only if y^α is small. Therefore, the results are the same, that is, the classical theory is a limit for classical real 4-*velocities*. A different limit cannot be obtained by assuming $v_\alpha v^\alpha \to$ constant (since the only such limit which can be covariant is $v_\alpha v^\alpha \to 0$, which may occur when the "internal" coordinates $y^\alpha \to 0$).

Equation (7a) for the real part u^α of the velocity contains three apparent forces, $-\Gamma^{\alpha\,\mathrm{Re}}_{\cdot\beta\gamma}u^\beta u^\gamma$, $\Gamma^{\alpha\,\mathrm{Re}}_{\cdot\beta\gamma}v^\beta v^\gamma$, and $2\Gamma^{\alpha\,\mathrm{Im}}_{\cdot\beta\gamma}u^\beta v^\gamma$. Two of the "forces" depend on the "internal states" v^α of the particle moving in the "field" $\Gamma^\alpha_{\cdot\beta\gamma}$, but all three interactions conserve "parity," since all contain an even number of imaginary factors. However, the motion of a particle depends indirectly, through v^α and Eq. (7b), on three other interactions, $-\Gamma^{\alpha\,\mathrm{Im}}_{\cdot\beta\gamma}u^\beta u^\gamma$, $\Gamma^{\alpha\,\mathrm{Im}}_{\cdot\beta\gamma}v^\beta v^\gamma$, and $-2\Gamma^{\alpha\,\mathrm{Re}}_{\cdot\beta\gamma}u^\beta v^\gamma$, which cannot conserve parity, since all contain an odd number of imaginary factors.

If an *indicator* ϵ (Synge and Schild,[13] p. 97) is introduced in the form (1) to distinguish timelike from spacelike vectors, the magnitude of an arbitrary vector w^i becomes

$$|w|^2 = \tfrac{1}{2}\epsilon(w)g_{ij}w^iw^j,$$

where $\epsilon = \pm 1$ when sgn $(g_{ij}w^iw^j) = \pm 1$. In $Z_{(i)}$, four extreme cases must be distinguished: (a) w^i is real-timelike, $\epsilon = +1$; (b) w^i is real-spacelike, $\epsilon = -1$; (c) w^i is imaginary-timelike, $\epsilon = -1$; (d) w^i is imaginary-spacelike, $\epsilon = +1$. Thus, the indicator of an imaginary vector is the negative of the indicator of a real vector having the same timelike or spacelike character.

The Riemannian curvature $K_{(i)}$ of a section of $Z_{(i)}$ formed on two vectors w_1 and w_2 is given by

$$K_{(i)}(z; w_1, w_2) = \frac{R_{ijkh}w_1^iw_2^jw_1^kw_2^h}{(g_{ac}g_{bd} - g_{ad}g_{bc})w_1^aw_2^bw_1^cw_2^d}.$$

For two real vectors u_1 and u_2, $K_{(i)}$ becomes

$$K_{(i)} = \frac{R^{\mathrm{Re}}_{\alpha\beta\gamma\delta}u_1^\alpha u_2^\beta u_1^\gamma u_2^\delta}{(g^{\mathrm{Re}}_{\epsilon\lambda}g^{\mathrm{Re}}_{\eta\mu} - g^{\mathrm{Im}}_{\epsilon\lambda}g^{\mathrm{Im}}_{\eta\mu} - g^{\mathrm{Re}}_{\epsilon\mu}g^{\mathrm{Re}}_{\lambda\eta} + g^{\mathrm{Im}}_{\epsilon\mu}g^{\mathrm{Im}}_{\lambda\eta})u_1^\epsilon u_2^\eta u_1^\lambda u_2^\mu},$$

where $R^{\mathrm{Re}}_{\alpha\beta\gamma\delta}$ is the real part of $R_{\alpha\beta\gamma\delta}$; and, for two imaginary vectors v_1' and v_2', $K'_{(i)}$ is

$$K'_{(i)} = \frac{\mathrm{Re}^{\mathrm{Re}}_{\alpha\beta\gamma\delta}v_1'^\alpha v_2'^\beta v_1'^\gamma v_2'^\delta}{(g^{\mathrm{Re}}_{\epsilon\lambda}g^{\mathrm{Re}}_{\eta\mu} - g^{\mathrm{Im}}_{\epsilon\lambda}g^{\mathrm{Im}}_{\eta\mu} - g^{\mathrm{Re}}_{\epsilon\mu}g^{\mathrm{Re}}_{\lambda\eta} + g^{\mathrm{Im}}_{\epsilon\mu}g^{\mathrm{Im}}_{\lambda\eta})v_1'^\epsilon v_2'^\eta v_1'^\lambda v_2'^\mu}.$$

Thus, if the components of a pair of real vectors are numerically equal to the respective components of a pair of imaginary vectors, the real and imaginary sectional curvatures are equal, $K_{(i)} = K'_{(i)}$.

For a space with a positive-definite metric form, if $K \leq 0$ everywhere the space is noncompact and, if $K > 0$ everywhere the space is compact (Helgason,[7] p. 205). With an indefinite metric form, a simple connection with the topology is obtained only if

[12] A. O. Barut, *Electrodynamics and Classical Theory of Fields and Particles*, (The Macmillan Company, New York, 1964).

[13] J. L. Synge and A. Schild, *Tensor Calculus*, (University of Toronto Press, Toronto, 1949).

the two vectors in $K(w_1, w_2)$ are close enough to have the same indicator ϵ. Then, the above inequalities must be replaced by $\epsilon K \leq 0$ everywhere for a noncompact space, and $\epsilon K > 0$ everywhere for a compact space, Synge and Schild,[13] p. 97. But, the three conditions $K_{(i)}(u_1, u_2) = 0$, $K_{(i)}(v_1, v_2) = 0$, and the Bianchi identity may be used to *define* the points at real, spacelike, infinite coordinate intervals of Axiom 5 and thus, $K_{(i)}(u_1, u_2)$ and $K_{(i)}(v_1, v_2)$ can be taken as nonzero at all finite intervals. Then, either X is noncompact and Y compact, or conversely; but, as shown in the last section, the existence of an imaginary origin implies only the former is valid in Z, i.e., the subspace X is noncompact and the subsets Y are compact at all finite intervals.

4. SINGULARITIES AND PARTICLES

The above analysis permits identification of the geometric structure of an "elementary particle." By the Cayley–Hamilton theorem, the Einstein tensor G_{ij} has, at most, four principal invariants; but, the reduced Bianchi identity implies that, at most, three of these are independent.

Defining $\delta^{i_1 \cdots i_p}_{k_1 \cdots k_p} = \delta^{i_1}_{k_1} \delta^{i_2}_{k_2} \cdots \delta^{i_p}_{k_p}$, $G^{i_1 \cdots i_p}_{k_1 \cdots k_p} = G^{i_1}_{k_1} G^{i_2}_{k_2} \cdots G^{i_p}_{k_p}$, and $e^{i_1 \cdots i_p}_{k_1 \cdots k_p} = (1/p!)\delta^{[i_1 \cdots i_p]}_{k_1 \cdots k_p}$, where [ ] indicates complete skew-symmetrization of i_p, then the invariants of G_{ik} are given by $G_{(p)} = \frac{1}{2} e^{i_1 \cdots i_p}_{k_1 \cdots k_p} G^{k_1 \cdots k_p}_{i_1 \cdots i_p}$. None of the invariants $G_{(p)}$ can be constants for then the asymptotic Axiom 6 would imply that they were everywhere zero and Z would be empty of physical content. Then, the Liouville theorem requires that each of the invariants have at least one singularity in Z.

Each singularity forms an analytic subspace of three complex, or six real dimensions. However, the *intersection* of the singularities of three independent invariants is of one complex, or two real dimensions. By a theorem of Hartogs (Behnke, and Thullen,[5] p. 50) the singular subspaces must be unbounded and, since they are independent and analytic, their mutual intersection must also be unbounded. Since the space itself is closed in one direction and open in the other, the same must be true of the singular intersection; that is, the intersection must form an infinite *tube*.

The eigenvalues of G_{ik} are the extrema of $\frac{1}{2} G_{ik} w^i w^k$ under the subsidiary condition

$$\tfrac{1}{2}\lambda g_{\alpha\beta} w^\alpha w^\beta + \tfrac{1}{2}\bar{\lambda} g_{\bar{\alpha}\bar{\beta}} w^{\bar{\alpha}} w^{\bar{\beta}} = 1.$$

Setting $\lambda = \mu + i\nu$, the determinantal condition for a solution can be separated into the real equations

$$f(\mu, \nu) = \mu^4 - 4a_1\mu^3 + 3a_2(\nu)\mu^2 - 2a_3(\nu)\mu + a_4(\nu) = 0, \tag{14a}$$

$$g(\mu, \nu) = -a_2'(\nu)\mu^3 + a_3'(\nu)\mu^2 - a_4'(\nu)\mu + a_5(\nu) = 0, \tag{14b}$$

where primes indicate ν-derivatives and

$$a_1 = \tfrac{1}{4} G^{\mathrm{Re}}_{(1)},$$

$$a_2 = \tfrac{1}{3}\{G^{\mathrm{Re}}_{(2)} + 3G^{\mathrm{Im}}_{(1)}\nu - 6\nu^2\},$$

$$a_3 = \tfrac{1}{2}\{G^{\mathrm{Re}}_{(3)} + 2G^{\mathrm{Im}}_{(2)}\nu - 3G^{\mathrm{Re}}_{(1)}\nu^2\},$$

$$a_4 = G^{\mathrm{Re}}_{(4)} + G^{\mathrm{Im}}_{(3)}\nu - G^{\mathrm{Re}}_{(2)}\nu^2 - G^{\mathrm{Im}}_{(1)}\nu^3 + \nu^4,$$

$$a_5 = G^{\mathrm{Im}}_{(4)} - G^{\mathrm{Re}}_{(3)}\nu - G^{\mathrm{Im}}_{(2)}\nu^2 + G^{\mathrm{Re}}_{(1)}\nu^3.$$

Both of Eqs. (14) can be made of the fourth degree in μ and ν by a rotation in the η, ν plane. A criterion of Bôcher,[14] p. 206, shows that $f(\mu, \nu)$ and $g(\mu, \nu)$ are relatively prime, which implies that there are only a finite number of discrete solutions (μ_i, ν_i) of (14), determined by the intersections of the "root curves" of $f = 0$ and $g = 0$. The Cauchy–Riemann equations, $f_\mu = g_\nu$, $f_\nu = g_\mu$, require the curves $f = 0$ and $g = 0$ to be mutually orthogonal at every such intersection. If (14b) has *three* real root curves, then there are a minimum of six intersections; thus, the condition that G_{ik} has exactly four eigenvalues requires that (14b) has exactly *one* and (14a) exactly *four* real solutions. The invariants $G^{\mathrm{Re,Im}}_{(i)}$ must satisfy two real inequalities involving the discriminants of (14a) and (14b) for this number of real solutions. These inequalities do not have an immediately clear physical significance.

Denoting the solutions of (14) by $G^\alpha = \mu_{(\alpha)} + i\nu_{(\alpha)}$, the contracted Bianchi identity becomes $\partial_\alpha G^\alpha = 0$ in a Riemann normal coordinate system oriented in the "Ricci directions" determined by the solutions $w^\gamma_{(\alpha)}$ of the extremum problem. Since the indicator ϵ of a vector is an invariant, three spacelike vectors can be permuted into each other by an unphysical transformation; but timelike and spacelike vectors cannot be interchanged. This implies that three of the eigenvalues G^α can be permuted with each other and can be termed "spacelike," while the fourth, G^0, must be "timelike." Continuity requires the identity $\partial_\alpha G^\alpha = 0$ to hold at singularities if one or more spacelike eigenvalues G^α become singular; thus G^0 must also be singular and the time z^0 must be one of the coordinates in any singular subspace or intersection of singular subspaces.

Since the time z^0 cannot be determined from the

[14] M. Bôcher, *Higher Algebra*, The Macmillan Company, New York, 1907).

equations of a singularity $1/G^0 = 1/G^\alpha = 0$, a determinantal condition must be satisfied for the existence of solutions z^α (Graves[15], p. 138); equivalently, $1/G^\alpha$ must be *regular* in z^α and $1/G^0$, irregular in z^0. [Throughout this development, indices $\alpha, \beta \cdots$ will be reserved for spacelike components and the time index will be used explicitly; also parentheses, e.g. (α), will be used to prevent summation.] Therefore, a spacelike coordinate transformation $z^\gamma \to \zeta^\gamma$, $z^0 = \zeta^0$ will reduce the general Laurent expansion $G^\alpha = \sum_{k=-\infty}^{\infty} C^\alpha_{\gamma k} z^{\gamma k}$ of G^α in the neighborhood of its singularity into a Laurent expansion in one variable ζ^α alone (for *coordinates* only, $z^{\gamma - k_i} = z^{-\gamma k_i} = 1/z^{\gamma k_i}$ and $\sum k_i = k$).

Writing $1/G^\alpha = 1_{(\alpha)k}\zeta^{\alpha k}$, the asymptotic condition 6 implies, $1_{(\alpha)k} = 0$ for $k \leq 0$ and the regularity condition implies $\partial_{(\alpha)}(1/G^\alpha) \neq 0$, $1_{(\alpha)1} \neq 0$; also, for small ζ^α, higher terms can be neglected. All of the spacelike eigenvalues can then be expressed in the form $G^\alpha = 1^{(\alpha)}\zeta_\alpha$ in the neighborhood of their singularities. Integrating the reduced Bianchi identity then requires G^0 to have the form $G^0 = \zeta^0 1^\alpha \zeta_\alpha^2 + f(\zeta_\gamma)$. The function of integration $f(\zeta_\gamma)$ also can be taken as linear in each ζ_γ, that is, $f(\zeta_\gamma) = C^\alpha \zeta_\alpha$ γ $n 1^\beta 1^\gamma \zeta_\beta \zeta_\gamma + m\zeta_1\zeta_2\zeta_3$, $\beta \neq \gamma$. The condition that $G_{(1)}(\zeta^\gamma)$ be an invariant function of its eigenvalues $(\partial G_{(1)}/\partial G^0) = \partial G_{(1)}/\partial G^\alpha = 0)$, however, implies $C^\alpha = -1^\alpha$. Therefore, in an appropriate coordinate system, the eigenvalues of $G_{\alpha\beta}$ are

$$\begin{aligned} G^0 &= \zeta^0 1^\alpha \zeta_\alpha^2 - 1^\alpha \zeta_\alpha - n 1^\beta 1^\gamma \zeta_\beta \zeta_\gamma \\ &\qquad + m\zeta_1\zeta_2\zeta_3, \qquad \beta \neq \gamma, \\ G^\alpha &= 1^{(\alpha)}\zeta_\alpha. \end{aligned} \tag{15}$$

Retaining only terms in $\zeta_1\zeta_2\zeta_3$ for reasons clarified below, the invariants of $G_{\alpha\beta}$ can be written

$$\begin{aligned} G_{(1)}(\zeta^\alpha) &= m\zeta_1\zeta_2\zeta_3 + \cdots, \\ G_{(2)}(\zeta^\alpha) &= -3n 1^1 1^2 1^3 \zeta_1\zeta_2\zeta_3 + \cdots, \\ G_{(3)}(\zeta^\alpha) &= G_{(4)}(\zeta^\alpha) = 0 + \cdots, \\ G^2_{(1)} &= 0, \\ G^2_{\cdot(1)} &= G^\alpha_\beta G^\beta_\alpha = G^{\cdot 2}_{(1)} - 2G_{(2)} = -2G_{(2)}. \end{aligned} \tag{16}$$

Expressions (15) and (16) show that $\oint f\, d\zeta^0 = 0$ (where the path is the boundary of a connected "disk" on the ζ^0-tube) for all of the components, eigenvalues, and invariants of $G_{\alpha\beta}$. Given a spacelike, connected 3-polycylinder whose projections onto the ζ^1, ζ^2, ζ^3 surfaces have radii ϵ_1, ϵ_2, $\epsilon_3 > 0$, then the operator

$$\mathrm{res}_{(3)} = \lim_{\epsilon_\alpha \to 0} \frac{1}{(2\pi i)^3} \oint_{(3)} d^3\zeta$$

will annihilate any of the components, eigenvalues, or invariants of $G_{\alpha\beta}$ at any point z^α not on the intersection of all three independent singularities. On the singular intersection, however,

$$\begin{aligned} \mathrm{res}_{(3)} G^0 &= m, \\ \mathrm{res}_{(3)} G^\alpha &= 0 \end{aligned} \tag{17}$$

and

$$\begin{aligned} \mathrm{res}_{(3)} G_{(1)} &= m, \\ \mathrm{res}_{(3)} G_{(2)} &= -3n 1^1 1^2 1^3, \\ \mathrm{res}_{(3)} G_{(3)} &= \mathrm{res}_{(3)} G_{(4)} = 0, \\ \mathrm{res}_{(3)} G^2_{\cdot(1)} &= 6n 1^1 1^2 1^3. \end{aligned} \tag{18}$$

In the case of several singular intersections, $G_{\alpha\beta}$ can be written as a product $G_{\alpha\beta} = G^{(1)}_{\alpha\beta} G^{(2)}_{\alpha\beta} \cdots$ such that each factor has exactly one such intersection. Applying $\mathrm{res}_{(3)}$ at a point on one intersection to both the formal geodesic equation (5) and the differential equation relating $\Gamma^\alpha_{\beta\gamma}$ to $G_{\alpha\beta}$ is, then, equivalent to replacing $\Gamma_{\beta\gamma}$ by $(1/m)\Gamma'^\alpha_{\beta\gamma}$, where $\Gamma'^\alpha_{\beta\gamma}$ is the field "due to" all other intersections. Therefore, the invariant residue m must be identical with the *rest mass* of a particle and the singular intersection must be the corresponding "complex world line." Also, $\mathrm{res}_{(3)}$ is an operator which transforms a "field representation" into a "particle representation."

In General Relativity, Einstein assumed that $G_{ik} = 0$ in the absence of matter and $G_{ik} = T_{ik}$ in its presence. For a complex space, these equations correspond to $\mathrm{res}_{(3)}\ G_{\alpha\beta} = 0$, in the first, and $\mathrm{res}_{(3)} G_{\alpha\beta} = T_{\alpha\beta}$, in the second case. From the viewpoint of this study, the second equation is tautological since both $G_{\alpha\beta}$ and $T_{\alpha\beta}$ are geometrical entities. Instead, the problem is one of correlating the geometrically-defined entities with those which have been already been assigned names in particular experiments.

Equation (17) implies that the only eigenvector (of the extremum problem for $G_{\alpha\beta}w^\alpha w^\beta$) with an eigenvalue whose residue does not vanish at a "particle" must be timelike, $w^\alpha_{(0)} = \{w^0, 0, 0, 0\}$, in a "center-of-mass" coordinate system. Defining $\rho_\alpha = G_{\alpha\beta} w^\beta_{(0)}$, $\rho^\alpha = G^\alpha_\beta w^\beta_{(0)}$, then $\mathrm{res}_{(3)}\ \rho_\alpha w^\alpha_{(0)} = m$. If $w^\beta_{(0)}(\zeta^\gamma)$ is an analytic vector field in the neighborhood of the particle, the vector $p^\alpha = \mathrm{res}_{(3)}\ \rho^\alpha$ also must be timelike. Because $w^\alpha_{(0)}$ is timelike, its real part can be subjected to the subsidiary condition $u_\alpha u^\alpha = 1$; its imaginary part v^α cannot

[15] L. M. Graves, *Functions of Real Variables* (McGraw-Hill Book Company, Inc., New York, 1946).

satisfy such a condition. However, the extremum problem for $G_{\alpha\beta}u^\alpha u^\beta$ leads to the same eigenvalue with nonvanishing residue m. Defining $p^\alpha = k^\alpha + ih^\alpha$, then $k^\alpha = \text{res}_{(3)}\, G^\alpha_\beta u^\beta$ and $\text{res}_{(3)}\, G_{\alpha\beta}u^\alpha u^\beta = m$, or $k_\alpha k^\alpha = m^2$. Thus, k^α is the ordinary *real* 4-momentum of particle dynamics.

The definition $p^\alpha = k^\alpha + ih^\alpha = \text{res}_{(3)}\, \rho^\alpha$ implies $p^2 = \frac{1}{2}p_i p^i = \frac{1}{2}(p_\alpha p^\alpha + p_{\dot\alpha}p^{\dot\alpha}) = k_\alpha k^\alpha - h_\alpha h^\alpha = m_0^2$, where m_0 is the "absolute mass" as distinguished from the rest mass m. With the exception of the sign change for h^2 required by the non-Kaehlerian metric, the formula $m_d^2 = m^2 - h^2$ is the same as that suggested by Barut[16] for the mass-splitting of particle multiplets. However, the absolute mass m_0 can be directly related to the other invariant residue of $G_{\alpha\beta}$. The quantity $\text{res}_{(3)}\, \rho_\alpha \rho^\alpha$ must be equal to the eigenvalue with nonvanishing residue of the extremum problem for $G_{\alpha\gamma}G^\gamma_\beta w^\alpha w^\beta$. First applying the operator $\text{res}_{(3)}$, the nonzero residue must satisfy the eigenvalue equation $\mu^4 - \text{res}_{(3)}\, G^2_{\cdot(1)}\mu^3 + 0 \cdots = 0$, or $\mu = \text{res}_{(3)}\, G^2_{\cdot(1)}$. Thus, m_0 is given by $m_0^2 = \text{res}_{(3)}\, G^2_{\cdot(1)} = 6n1^1 1^2 1^3$.

[16] A. O. Barut, *Nuovo Cimento* **32**, 234(1964).

ACKNOWLEDGMENTS

We would like to express our appreciation for many discussions with Professor A. O. Barut during this study.

This study was partially supported financially by the National Bureau of Standards.

On the L–S Basis of the Poincaré Group for the Case of Nonzero Rest Mass

A. CHAKRABARTI
Centre de Physique Théorique de l'Ecole Polytechnique, Paris, France
(Received 22 June 1965)

The action of the generators of the Poincaré group on the basis functions spanning an irreducible representation $[m^2, s^2]$ $(m > 0)$ and labeled by the eigenvalues of the energy, orbital angular momentum and spin, and the projections on a fixed axis of the two latter are considered. The explicit canonical representation of the generators (1.1) are used. The formulas are derived first for the zero-spin case (Sec. 2) and then for the general case (Sec. 3). It is shown that our L–S basis permits a much more compact and systematic derivation of the formulas than the basis involving total angular momentum and helicity, which is considered by Lomont and Moses. The separation of the orbital and spin parts allows us also to display explicitly the role played by the "canonical" definition of spin.

1. INTRODUCTION

THE explicit forms of the infinitesimal generators of the Poincaré group in the unitary canonical representation[1] are (for nonzero rest mass $m > 0$)

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

They satisfy the well-known commutation relations of the group and an irreducible representation $[m^2, s^2]$ is specified by the eigenvalues of P^2 and $\mathbf{S}^2$. Here $\mathbf{S}$ are the usual $(2s + 1)(2s + 1)$ Hermitian spin matrices.

Let us now consider an irreducible representation whose basis vectors are characterized by the eigenvalues of the operators

$$P^2, \mathbf{S}^2; \mathbf{L}^2, L^3, S^3, P^0 \quad (1.2)$$

where

$$\mathbf{L} = -i\mathbf{P}\times(\partial/\partial\mathbf{P}) \equiv -\mathbf{P}\times\mathbf{X}[\mathbf{X} = i(\partial/\partial\mathbf{P})] \quad (1.3)$$

is the orbital part of the angular momentum operator. $\mathbf{L}$ is used to replace the operators $\mathbf{p}$ which is generally used to label the basis vectors. We propose to investigate, in the following sections, the action of the infinitesimal generators on such a basis.

Lomont and Moses[2] have derived the required results using a basis characterized by the eigenvalues of the operators

$$\mathbf{M}^2, M^s, \mathbf{P}\cdot\mathbf{M}/|\mathbf{P}|(=\mathbf{P}\cdot\mathbf{S}/|\mathbf{P}|). \quad (1.4)$$

Their derivations are however very long and complicated, particularly for the case $m > 0$, which, we propose to study. It will be seen that the use of an L–S basis permits us to derive the corresponding results in a very much more compact and systematic way. (Since the helicity operator does not commute with $\mathbf{L}$, the use of this latter is not indicated for particles of zero rest mass which are supposed to be always in eigenstates of helicity.)

We will first derive the results for the case of zero spin (Sec. 2). Both $\mathbf{P}$ and $\mathbf{N}$ act as tensor operators of rank one, and Wigner–Eckart theorem can be employed, leaving only the "reduced matrix elements" to be determined from other conditions to satisfied. This simplifies the task considerably. And once this part of the work is done the generalization necessary for the case of nonzero spin (Sec. 3) is trivial in our case, though the results for $\mathbf{N}$ contain more terms corresponding to the part

$$-(\mathbf{P}\times\mathbf{S})/(P^0 + m) \quad (1.5)$$

of (1.1).

Since eigenstates of the total angular momentum $\mathbf{M}^2$ (and M^3) may be constructed by coupling the L–S states [see (1.8)] as usual, we can write down explicit expressions for such states as

$$\varphi(Els; jj_3) = \sum_{l^3} (ll_3ss_3 \mid jj_3)\varphi(Ell_3ss_3). \quad (1.6)$$

Our technique also permits us to display explicitly the particular role played by the "canonical" definition of spin.

For the two fundamental inequivalent "spinor" representations, we obtain a representation of the generators on replacing in (1.1) only $\mathbf{N}$ by

$$\mathbf{N} = -iP^0(\partial/\partial\mathbf{P}) \pm i\mathbf{S}. \quad (1.6)$$

Thus when considering a representation with nonzero spin we have (if $\mathbf{S}$ is considered as the operator appearing in the spinor representation) to add to the

[1] A. Chakrabarti, J. Math. Phys. 4, 1215, 1223 (1963); 5, 922, 1747 (1964); These, Université de Paris.
[2] J. S. Lomont and H. E. Moses, J. Math. Phys. (I) 5, 294 (1964); (II) 5, 1438 (1964).

case of zero spin formulas for **N** only the effect of

$$\pm i\mathbf{S} \tag{1.7}$$

Whereas for the canonical definition of **S** we have to use (1.5), which lead to (3.2).

We start with a basis $\varphi(Ell_3s_3)$ with the normalization[3]

$$\langle\varphi(E'l'l_3's_3')\varphi(Ell_3s_3)\rangle = 2\,|\mathbf{p}'|\,\delta(E'-E)\delta ll'\delta l_3l_3'\delta s_3s_3' \tag{1.8}$$

and an inner product

$$\langle\psi\Phi\rangle = \sum_{l\,l_3\,s_3}\int\frac{dE}{2\,|\mathbf{p}|}\langle\psi\varphi(Ell_3s_3)\rangle\langle\varphi(Ell_3s_3)\Phi\rangle$$
$$= \sum_{l\,l_3\,s_3}\int dE\;\psi^*(Ell_3s_3)\Phi(Ell_3s_3). \tag{1.8$'$}$$

(We have suppressed the labels $[m^2, s^2]$ which specify the irreducible representation.) The matrix elements of P^0, **L**, and **S** are given by

$$P^0\varphi(Ell_3s_3) = E\varphi(Ell_3s_3),$$
$$\mathbf{L}^2\varphi(Ell_3s_3) = l(l+1)\varphi(Ell_3s_3),$$
$$L^3\varphi(Ell_3s_3) = l_3\varphi(Ell_3s_3),$$
$$(L^1 \pm iL^2)\varphi(Ell_3s_3)$$
$$= [(l\mp l_3)(l\pm l_3+1)]^{\frac{1}{2}}\varphi(Ell_3\pm 1s_3), \tag{1.9}$$
$$\mathbf{S}^2\varphi(Ell_3s_3) = s(s+1)\varphi(Ell_3s_3),$$
$$S_{(0)}\varphi(Ell_3s_3) = s_3\varphi(Ell_3s_3),$$
$$(S^1 \pm iS^2)\varphi(Ell_3s_3)$$
$$= [(s\mp s_3)(s\pm s_3+1)]^{\frac{1}{2}}\varphi(Ell_3s_3\pm 1).$$

The formulas for $\mathbf{L}^2$, $(L^1 \pm iL^2)$, $\mathbf{S}^2$, $(S^1 \pm iS^2)$ can be derived in a well-known fashion. Our real task to be carried out in the following sections, is the derivation of the matrix elements of **P** and **N**.

2. ANGULAR MOMENTUM BASIS FOR ZERO SPIN

In this case the basis characterized by the eigenvalues of $P^0, \mathbf{L}^2, L^3$ (where now $\mathbf{M} = \mathbf{L}$), is denoted as

$$\varphi(E, ll_3). \tag{2.1}$$

[3] We have chosen this normalization and scalar product in view of the relation

$$\int dE\,d\,|p|\,\delta(E^2-|p|^2-m^2) = \int\frac{dE}{2\,|\mathbf{p}|} = \int\frac{d\,|\mathbf{p}|}{2E}.$$
$$(m\le E<\infty, 0\le|\mathbf{p}|<\infty)$$

We may avoid the factor $|\mathbf{p}|$ in the denominator by replacing the integretation over dE by one over $d\,|\mathbf{p}|$. But in any case our operators **N**[as (2.31), (2.31′) show] are not well-defined at the point $E = m$ ($|\mathbf{p}| = 0$). See the final remarks (Sec. 4) in this connection.

A. Matrix elements of P

Let us consider the action of the generators

$$P_{(0)} = P^3, \qquad P_{(\pm1)} = \mp(1/\sqrt{2})(P^1 \pm iP^2). \tag{2.2}$$

These operators leave the eigenvalue E unaffected and act only on l, l_3. The relations

$$[L^i, P^j] = i\epsilon_{ijk}P^k \tag{2.3}$$

imply that the $P_{(q)}$'s are tensorial operators of rank 1. Hence, according to the Wigner–Eckart theroem, we must have

$$P_{(q)}\varphi(Ell_3) = \sum_{l'}(2l'+1)^{-\frac{1}{2}}(1q,\,ll_3\mid l'l_3')$$
$$\times(l'\,||P||\,l)\varphi(El'l_3') \quad (q=-1,0,+1), \tag{2.4}$$

where the Clebsch–Gordan coefficients imply that

$$l' = l,\ l\pm1 \quad\text{and}\quad l_3' = l_3+q. \tag{2.5}$$

Our essential task is the evaluation of the reduced matrix elements

$$(l'\,||P||\,l).$$

To this end we utilize the condition

$$\mathbf{P}^2\varphi(Ell_3) = (P_{(0)}^2 - 2P_{(+1)}P_{(-1)})\varphi(Ell_3)$$
$$= \mathbf{p}^2\varphi(Ell_3)\ (\mathbf{p}^2 = E^2-m^2). \tag{2.6}$$

This gives us the following equations:

$$(l+1\,||P||\,l+1)$$
$$= [l(2l+3)/(2l+1)(l+2)]^{\frac{1}{2}}(l\,||P||\,l) \tag{2.7}$$

[equating to zero the coefficient of $\varphi(l+1l_3)$ or $\varphi(l-1l_3)$];

$$a_{(l+1)} + a_{(l)} + b_{(l)} = (2l+1)\mathbf{p}^2 \tag{2.8}$$

[equating to $\mathbf{p}^2$ the coefficient free of l_3 of $\varphi(ll_3)$];

$$la_{(l+1)} - (l+1)a_{(l)} - b_{(l)} = 0 \tag{2.9}$$

[equating to zero the part of the coefficient of $\varphi(ll_3)$ containing l_3 as a factor].

In (2.8), (2.9) we have put

$$a_{(l)} \equiv -(l\,||P||\,l-1)(l-1\,||P||\,l) \tag{2.10}$$
$$= (l\,||P||\,l-1)(l\,||P||\,l-1)^* \tag{2.11}$$

[(2.11) follows from (2.10) as a consequence of the Hermiticity of **P**],

and

$$b_{(l)} \equiv (l\,||P||\,l)^2. \tag{2.12}$$

We easily find the solutions

$$a_{(l)} = l\mathbf{p}^2, \qquad b_{(l)} = 0, \tag{2.13}$$

i.e., $(l||P||\ l-1) = \pm l^{\frac{1}{2}}\ |\mathbf{p}|, \quad = -(l-1\ ||P||\ l)$

$$(l\ ||P||\ l) = 0. \tag{2.14}$$

The vanishing of $b_{(1)}$ might have been expected by considering the action of the parity operator P, such that

$$\mathcal{P}\varphi(Ell_3) = (-1)^l\varphi(Ell_3)$$

and

$$\mathcal{P}\mathbf{P}\mathcal{P}^{-1} = -\mathbf{P}. \tag{2.14'}$$

We can eliminate the ambiguity of sign eventually by a suitable phase convention. Keeping it for the present we can write explicitly

$$P_{(0)}\varphi(Ell_3) = \pm|\mathbf{p}|$$
$$\times\left[\left(\frac{(l-l_3+1)(l+l_3+1)}{(2l+3)(2l+1)}\right)^{\frac{1}{2}}\varphi(El+1l_3)\right.$$
$$\left.+\left(\frac{(l-l_3)(l+l_3)}{(2l+1)(2l-1)}\right)^{\frac{1}{2}}\varphi(El-1l_3)\right], \tag{2.15}$$

$$\sqrt{2}P_{(\pm1)}\varphi(Ell_3)$$
$$= \pm|\mathbf{p}|\left[\left(\frac{(l\pm l_3+1)(l\pm l_3+2)}{(2l+3)(2l+1)}\right)^{\frac{1}{2}}\right.$$
$$\times\varphi(El+1l_3\pm1) - \left(\frac{(l\mp l_3+1)(l\mp l_3)}{(2l+1)(2l-1)}\right)^{\frac{1}{2}}$$
$$\left.\times\varphi(El-1l_3\pm1)\right]. \tag{2.15'}$$

B. Matrix elements of N

As a first step towards the determination of the matrix elements of

$$\mathbf{N} = -iP^0(\partial/\partial\mathbf{P}), \tag{2.16}$$

we consider now those of the gradient

$$\mathbf{X} = i(\partial/\partial\mathbf{P}) \tag{2.17}$$

for *the rotation group only*. For this case, the basis is of the form

$$\psi(ll_3); \tag{2.18}$$

the eigenvalue E of P^0 no longer entering into consideration. For this restricted case our task is quite simple. The commutation relations of $\mathbf{X}$ with $\mathbf{L}$ imply [exactly as for (2.3) and (2.4)]

$$X_{(q)}\psi_{(ll_3)} = \sum_{l'}(2l'+1)^{-\frac{1}{2}}$$
$$\times(1q,\ ll_3\ |\ l'l_3')(l'\ ||X||\ l)\psi(l'l_3'). \tag{2.19}$$

The reduced matrix elements

$$(l'\ ||X||\ l)$$

can easily be evaluated by using the relations

$$(\mathbf{X}\times\mathbf{P})_{(q)}\psi(ll_3)$$
$$= L_{(q)}\psi(ll_3)\ (q=-1,\ 0,\ +1). \tag{2.20}$$

Thus, for example, from

$$i^{-1}(X_{(+1)}P_{(-1)} - X_{(-1)}P_{(+1)})\psi_{(113)} = l_3\psi(ll_3), \tag{2.21}$$

we have [using (2.14)] the equations

$$(l\ ||X||\ l) = 0 \tag{2.22}$$

and

$$\pm|\mathbf{p}|\left[\frac{-1}{(l+1)^{\frac{1}{2}}}(l\ ||X||\ l+1)\right.$$
$$\left.-\frac{1}{(l)^{\frac{1}{2}}}(l\ ||X||\ l-1)\right] = i(2l+1). \tag{2.23}$$

Here we have introduced $|\mathbf{p}|$ as a constant factor brought in by the definition of $(l'\ ||P||\ l)$, continuing to ignore the action of $\mathbf{X}$ on E. This artificial trick, however, helps us to arrive at the simple "ansatz" (2.25), since in analogy (though not in identity) with the "gradient formula" we may now expect only an additive extra term involving $\partial/\partial E$ for the full formula.

The solution of (2.23) is

$$(l\ ||X||\ l-1) = \mp il^{3/2}/|\mathbf{p}|$$
$$= (l-1)\ ||X||\ l). \tag{2.24}$$

It is to be noted that when $|\mathbf{p}|$ is zero in the denominator, l is also zero and we have $l^{\frac{1}{2}}$ as the numerator.

Our problem is, however, not yet completely solved. Since $\mathbf{X}$ (and $\mathbf{N}$) do not commute with P^0, some additional complications arise when we consider not only the rotation group (and the basis $\psi_{(ll_3)}$) but the Poincaré group and the basis $\varphi(Ell_3)$.

Since the tensorial properties of $\mathbf{X}$ (and $\mathbf{N}$) with respect to the subgroup formed by $\mathbf{L}$ remain inaltered, only a modification of the reduced elements $(l'\ ||X||\ l)$ is required.

As an ansatz let us write

$$N_{(0)}\varphi(Ell_3) = -P^0X_{(0)}\varphi(Ell_3)$$
$$= \pm i\frac{E}{|\mathbf{p}|}\left[\left(\frac{(l-l_3+1)(l+l_3+1)}{(2l+3)(2l+1)}\right)^{\frac{1}{2}}\right.$$
$$\times[(l+1)+\xi]\varphi(El+1l_3)$$
$$\left.-\left(\frac{(l-l_3)(l+l_3)}{(2l+1)(2l-1)}\right)^{\frac{1}{2}}(l+\xi')\varphi(El-1l_3)\right], \tag{2.25}$$

where the as yet unknown terms ξ, ξ' represent the necessary modification of (2.22) and (2.24). The same terms are also, of course, to be added to the expressions corresponding to $N_{(\pm1)}$.

We now evaluate the unknown terms using the relation

$$[N_{(0)}, P^0]\varphi(Ell_3) = (E - P^0)N_{(0)}\varphi(Ell_3) = -iP_{(0)}\varphi(Ell_3). \quad (2.26)$$

Using (2.15), (2.25) and the relation

$$(E - P^0)\varphi(Ell_3) = 0,$$

we obtain

$$\xi' = -\xi \quad (2.27)$$

and

$$(E/|\mathbf{p}|^2)(E - P^0)\xi\varphi(Ell_3) = -\varphi(Ell_3). \quad (2.28)$$

Lomont and Moses[2] (II) have noted that

$$(E - P^0)\frac{\partial}{\partial E}(Ell_3) = (E - P^0)\lim_{\Delta\to 0}\frac{\varphi(E + \Delta ll_3) - \varphi(Ell_3)}{\Delta} = \lim_{\Delta\to 0}\frac{-\Delta}{\Delta}\varphi(E + \Delta ll_3) = -\varphi(Ell_3). \quad (2.29)$$

(See the concluding remarks on this point.) Hence (2.26) is satisfied if we put

$$\xi = (|\mathbf{p}|^2/E)(\partial/\partial E). \quad (2.30)$$

Indeed, in light of the remark following (2.23) we might have expected just such a solution.

Now we can write down the matrix elements of **N** as

$$N_{(0)}\varphi(Ell_3) = \pm i\left[\left(\frac{(l - l_3 + 1)(l + l_3 + 1)}{(2l + 3)(2l + 1)}\right)^{\frac{1}{2}} \times \left((l + 1)\frac{E}{|\mathbf{p}|} + |\mathbf{p}|\frac{\partial}{\partial E}\right)\varphi(El + 1l_3) - \left(\frac{(l - l_3)(l + l_3)}{(2l + 1)(2l - 1)}\right)^{\frac{1}{2}}\left(l\frac{E}{|\mathbf{p}|} - |\mathbf{p}|\frac{\partial}{\partial E}\right)\varphi(El - 1l_3)\right], \quad (2.31)$$

$$\sqrt{2}N_{(\pm 1)}\varphi(Ell_3) = \pm i\left[\left(\frac{(l \pm l_3 + 1)(l \pm l_3 + 2)}{(2l + 3)(2l + 1)}\right)^{\frac{1}{2}} \times \left((l + 1)\frac{E}{|\mathbf{p}|} + |\mathbf{p}|\frac{\partial}{\partial E}\right)\varphi(El + 1l_3 \pm 1) + \left(\frac{(l \mp l_3 + 1)(l \mp l_3)}{(2l + 1)(2l - 1)}\right)^{\frac{1}{2}} \times \left(l\frac{E}{|\mathbf{p}|} - |\mathbf{p}|\frac{\partial}{\partial E}\right)\varphi(El - 1l_3 \pm 1)\right]. \quad (2.31')$$

The hermiticity of **N** is assured due to the combined effect of the change of sign (2.27) and our normalization (1.8) giving

$$\langle E'l'l'_3\ |\mathbf{N}|\ Ell_3\rangle = \langle Ell_3\ |\mathbf{N}|\ E'l'l'_3\rangle^*.$$

The ambiguity of sign is absent if we adopt a definite convention for $P_{(a)}$ in (2.15), no further assumption is required.

It may be verified by direct calculation that the sets (2.15) and (2.31) satisfy all the required commutation relations. The ambiguity of sign is not removed by the restrictions due to the commutation relations alone. We may, however, adopt the convention of retaining only one sign, for example, the plus sign. As might have been expected, these formulas agree with those of Lomont and Moses[2] (I) if one puts into their expressions

$$s = 0 = \alpha$$

and takes account of certain differences of conventions.

3. (*L–S*) BASIS FOR NONZERO SPIN

Given our chosen basis and the explicit representation of the generators (1.1), the generalization required to include the case of nonzero spin presents no problem at all. The basis

$$\varphi(Ell_3ss_3)$$

is now defined by the full system of Eq. (1.9).

The matrix elements for **P** (2.15), (2.15′) may be left inaltered, except for adding the same indices s, s_3 to the φ's on both sides.

As for **N** (2.31), (2.31′) we have to add the effect of the term

$$-(\mathbf{P} \times \mathbf{S})/(P^0 + m).$$

Thus, now

$$N_{(0)} = -P^0X_{(0)} + \frac{i(P_{(+1)}S_{(-1)} - P_{(-1)}S_{(+1)})}{P^0 + m},$$

$$N_{(\pm 1)} = -P^0X_{(\pm 1)} \pm \frac{i(P_{(\pm 1)}S_{(0)} - P_{(0)}S_{(\pm 1)})}{P^0 + m}. \quad (3.1)$$

We have

$$\frac{i(P_{(+1)}S_{(-1)} - P_{(-1)}S_{(+1)})}{P^0 + m}\varphi(Ell_3s_3) = \pm\frac{i\,|\mathbf{p}|}{2(E + m)}[(S + S_3)(S - S_3 + 1)]^{\frac{1}{2}} \times \left[\left(\frac{(l + l_3 + 1)(l + l_3 + 2)}{(2l + 3)(2l + 1)}\right)^{\frac{1}{2}}\varphi(El + 1l_3 + 1s_3 - 1) - \left(\frac{(l - l_3 + 1)(l - l_3)}{(2l + 1)(2l - 1)}\right)^{\frac{1}{2}}\varphi(El - 1l_3 + 1s_3 - 1)\right]$$

$$\pm \frac{i\,|\mathbf{p}|}{2(E+m)}\,[(S-S_3)(S+S_3+1)]^{\frac{1}{2}}\Big[\Big(\frac{(l-l_3+1)(l-l_3+2)}{(2l+3)(2l+1)}\Big)^{\frac{1}{2}}\varphi(El+1_{l_3}-1s_3+1)$$

$$-\Big(\frac{(l+l_3+1)(l+l_3)}{(2l+1)(2l-1)}\Big)^{\frac{1}{2}}\varphi(El-1l_3-1s_3+1)\Big]. \tag{3.2}$$

Similarly,

$$\frac{i(P_{(\pm1)}S_{(0)}-P_{(0)}S_{(\pm1)})}{P^0+m}\,\varphi(El_3s_3)$$

$$=\pm\frac{i\,|\mathbf{p}|}{\sqrt{2}(E+m)}\,S_3\Big[\Big(\frac{(l\pm l_3+1)(l\pm l_3+2)}{(2l+3)(2l+1)}\Big)^{\frac{1}{2}}\varphi(El+1l_3+1s_3)$$

$$-\Big(\frac{(l\mp l_3+1)(l\mp l_3)}{(2l+1)(2l-1)}\Big)^{\frac{1}{2}}\varphi(El-1l_3\pm1s_3)\Big]\mp\frac{i\,|\mathbf{p}|}{\sqrt{2}(E+m)}\,((S\mp S_3)(S\pm S_3+1))^{\frac{1}{2}}$$

$$\times\Big[\Big(\frac{(l-l_3+1)(l+l_3+1)}{(2l_3+3)(2l+1)}\Big)^{\frac{1}{2}}\varphi(El+1l_3s_3\pm1)+\Big(\frac{(l-l_3)(l+l_3)}{(2l+1)(2l-1)}\Big)^{\frac{1}{2}}\varphi(El-1l_3s_3\pm1)\Big]. \tag{3.2'}$$

Equations (3.2), (3.2′) represent the typical effects of the canonical definition of spin.

4. REMARKS

In conclusion we would like to add some comments about two points of the preceding analysis.

The first remark is about the Eq. (2.29). Since (dropping the indices ll_3 for brevity)

$$(P^0-E)(\partial/\partial E)\varphi(E)=[(P^0-E),(\partial/\partial E)]\varphi(E)$$

$$=[P^0,(\partial/\partial E)]\varphi(E)+\varphi(E), \tag{4.1}$$

we see that the inversion of the operations of P^0 and "limit $\Delta=0$" made in (2.29) is equivalent to

$$[P^0,(\partial/\partial E)]\varphi(E)=0. \tag{4.2}$$

(2.29) can also be proved by explicitly using the properties of δ-functions as follows.

Suppose we have a normalization

$$\langle\varphi(E')\varphi(E)\rangle=f(E')\,\delta(E'-E).$$

Then

$$\langle\varphi(E')(P^0-E)(\partial/\partial E)\varphi(E)\rangle=-(E'-E)f(E')$$

$$\times\,\delta'(E'-E)=f(E')\,\delta(E'-E)=\langle\varphi(E')\varphi(E)\rangle. \tag{4.3}$$

In fact the inversion of the operations of P^0 and "lim" tacitly admitted in (2.29) can only be justified by invoking the special properties of δ-functions defined as generalized functions.

Our second remark is about the singularity which the formulas (2.31), (2.31′) exhibit at the point

$$E=m \quad \text{or} \quad |\mathbf{p}|=0 \tag{4.4}$$

due to the factor $E/|\mathbf{p}|$.

It should be noted that at the point $E=m$ we have a φ which is simultaneously an eigenstate of all the components of the operators **P** and **L**. And such an exceptional situation can arise only at this point, when all the above eigenvalues are simultaneously zero.

Now **N**, which is the generator of pure Lorentz transformations, acting on an eigenstate **P**, must always give another eigenstate of **P**, which can be developed as a superposition of the different eigenstates of $\mathbf{L}^2$, L^3. But had the formulas (2.31), (2.31′) been perfectly well-defined at $E=m$, the generators **N** acting on the state $l=0$, would have given [according to (2.31), (2.31′)] one definite state of orbital angular momentum $l=1$. This of course would have been in contradiction with the situation previously described. Thus we see that there is a fundamental physical reason for the lack of definition of the matrix elements of **N** at the point $E=m$.

ACKNOWLEDGMENT

The author acknowledges with pleasure several helpful comments of Professor L. Michel and Dr. H. Epstein.

A New Method in the Theory of Potential Scattering

R. P. GILBERT*†
Institute for Fluid Dynamics and Applied Mathematics
AND
S. Y. SHIEH‡§
Department of Physics and Astronomy, University of Maryland, College Park, Maryland
(Received 26 July 1965)

A new method is given for investigating the position, number, and type of singularities the scattering amplitude may have as a function of complex momentum transfer. The advantage of this method over ones relying on the Watson transform is that one needs only to know the partial amplitudes for physical values of the angular momentum.

IN the theory of potential scattering there are two distinct approaches to the study of the scattering amplitude $<A(k, \cos\theta)$ as a function of both momentum k and scattering angle $\cos\theta$. One direct way is to study the Born[1] or Fredholm[2] series solution of the Lippmann–Schwinger equation for the scattering amplitude. With this method the analytic properties of the scattering amplitude as a function of the variable (momentum transfer) $t = -2k^2(1 - \cos\theta)$ for fixed k has been established for the Yukawa potential, but the nature of the branch point at $t = 4\mu^2$ (μ^{-1} is the range of the Yukawa potential) has not been investigated. The other approach consists of two steps; the first being to investigate the partial-wave amplitudes, and then to study the partial-wave expansion series for the whole amplitude. The study of partial-wave amplitudes is much easier and a great deal of work has been done. But the difficulty in this approach lies in the second step of studying the expansion series for the whole scattering amplitude. Regge[3] showed the usefulness of Watson transform and established the same result concerning the analytic properties of the scattering amplitude as a function of the variable t. But in order to make use of the Watson transform, he first had to generalize the partial-wave amplitude from physical integral angular momentum l to a complex one. We present a new method which requires only a knowledge of partial wave amplitudes for integral l. Furthermore, this method not only gives information on the existence of singularities but also determines the types of singularities.

The main ideas of this new method are as follows: First we observe that there exists an integral identity between the following two series

$$A(k, z) = \sum_{l=0}^{\infty} a_l(k) P_l(z), \tag{1}$$

$$B(k, z) = \sum_{l=0}^{\infty} a_l(k)(z)^l, \tag{2}$$

where $P_l(z)$ are Legendre polynomials ($z = \cos\theta$), and if we set $a_l(k) \equiv [(2l + 1)/2ik]A_l(k)$ [$A_l(k)$ are partial-wave amplitudes]. $A(k, z)$ becomes the scattering amplitude. The integral identity relating the two functions $A(k, z)$ and $B(k, z)$ is

$$A(k, z) = \frac{1}{2\pi i} \times \int_{|\zeta|=1} B[k, z + \tfrac{1}{2}(\zeta + \zeta^{-1})(z^2 - 1)^{\frac{1}{2}}] \frac{d\zeta}{\zeta}. \tag{3}$$

Using this formula we can study the singularities of $A(k, z)$ from those of $B(k, z)$. For example, if the singularities of $B(k, z)$ are poles, then the corresponding singularities of $A(k, z)$ will be algebraic branch points.

Now the series (2) is a power series and can be easily studied by the methods of Hadamard,[4] Mandelbrojt,[5] and Fabry.[6] In the following we list several theorems that are useful for this investigation. The first theorem is a slight modification of one established by Nehari[7] for Legendre series (see also Szegö[8]), and later extended by Gilbert using

* This research was supported in part by the National Science Foundation under Grants NSF GP-3937, NSF GP-5023 and in part by the Air Force Office of Scientific Research under Grant AFOSR 400-64.
† Present address: Department of Mathematics, Georgetown University, Washington, D. C.
‡ This research was supported in part by the Air Force Office of Scientific Research under Grant AFOSR 500-64.
§ Present address: Department of Physics, Lehigh University, Bethlehem, Pennsylvania.
[1] A. Klein, J. Math. Phys. 1, 41 (1960).
[2] R. Blankenbecker, M. L. Goldberger, M. N. Khuri, and S. B. Treiman, Ann. Phys. (N.Y.) 10, 62 (1960).
[3] T. Regge, Nuovo Cimento 14, 951 (1959).
[4] P. Dienes, *The Taylor Series* (Dover Publications, Inc., New York, 1957, 335.
[5] S. Mandelbrojt, Compt. Rend. (Paris) 204, 1456 (1937).
[6] See Ref. 4, p. 377.
[7] Z. Nehari, J. Ratl. Mech. Anal. 5, 987 (1956).
[8] G. Szegö, J. Ratl. Mech. Anal. 3, 561 (1954).

a different approach for Gegenbauer series. The remaining results are analogs of other theorems obtained by Gilbert (for the case of Gegenbauer series) in a study of a certain differential equation.[9] For the sake of brevity we state the theorems directly related to $A(k, z)$, namely, those theorems for the power series $B(k, z)$ which are transformed by the integral indentity (3).

First we list a fundamental result which connects the singularities of the series (1) and (2).

Theorem 1 [Nehari]: The necessary and sufficient conditions for $A(k, t)$ to be singular on the analytic manifold $t \equiv \frac{1}{2}\{\alpha(k) + [1/\alpha(k)]\} (t \neq \pm 1)$ is that $B(k, z)$ is singular on $z = \alpha(k)$. [The reader should consult the original paper by Nehari referred to above for a detailed explanation of the condition $t \neq \pm 1$, and an interesting example of how singularities may occur at $t = \pm 1$ which do not correspond to singularities of $B(k, z)$.]

Next we introduce some notation. Let

$$D_l^{(\mu)}(k) \equiv \begin{vmatrix} a_l(k) & a_{l+1}(k) & \cdots & a_{l+\mu}(k) \\ a_{l+1}(k) & \cdots & & a_{l+\mu+1}(k) \\ \cdot & & \cdot & \cdot \\ a_{l+\mu}(k) & a_{l+\mu+1}(k) & \cdots & a_{l+2\mu}(k) \end{vmatrix},$$

$$\mathfrak{L}_\mu \equiv \overline{\lim_{l\to\infty}} |D_l^{(\mu)}(k)|^{1/l},$$

and $E(\rho)$ is the ellipse

$$x^2(\rho + 1/\rho)^{-2} + y^2(\rho - 1/\rho)^{-2} = \tfrac{1}{4}.$$

Then we list the analog of the Hadamard criteria[10] for the number and type of singularities for series of the form (2). (The proof of this result is similar to ones given by Bergman[10] and by Gilbert[11] for a related series.)

Theorem 2: We have the following possibilities for each fixed value of k concerning the singularities of $A(k, z)$ in the z-plane:

(a) If there exists a ν such that for $\mu \geq \nu$, $\mathfrak{L}_\mu/\mathfrak{L}_{\mu-1} = 0$ (the ratios $\mathfrak{L}_\mu/\mathfrak{L}_{\mu-1}$ are monotone decreasing with respect to μ), then $A(k, z)$ has at most ν polelike branch points in the entire z plane. Furthermore, these branch points have algebraic ramifications.

(b) If $\mathfrak{L}_\mu/\mathfrak{L}_{\mu-1} \to 0$, $A(k, z)$ has just a finite number of polelike, algebraic, branch points in every compact set of the z-plane.

(c) If $\mathfrak{L}_\mu/\mathfrak{L}_{\mu-1} \to R^{-1}$, $A(k, z)$ has just a finite number of polelike algebraic branch points in the ellipse, $E(\rho)$ $(\rho < R)$, but an infinite number of polelike branch points in a neighborhood of $E(R)$.

(d) If $\mathfrak{L}_\mu/\mathfrak{L}_{\mu-1} \geq R^{-1}$ for $\mu \leq \nu$ and $\mathfrak{L}_\mu/\mathfrak{L}_{\mu-1} = R^{-1}$ for $\mu > \nu$, then (for each fixed k) $A(k, z)$ has in general a nonalgebraic branch point on the ellipse $E(R)$.

Remark 1: It is easy to see the connection between polar singularities of $B(k, z)$ and algebraic branch points of $A(k, z)$, namely, if

$$B(k, z) = \sum_{\nu=1}^{\infty} \sum_{\mu=1}^{m_\nu} \frac{M_{\mu\nu}(k)}{[z - b_\nu(k)]^\mu},$$

then by the residue theorem we have[10]

$$A(k, z) = \sum_{\nu=1}^{\infty} \sum_{\mu=1}^{m_\nu} \frac{(-1)^{\mu-1}}{(\mu - 1)!} M_{\mu\nu}(k) \times \frac{\partial^{\mu-1}}{\partial b_\nu^{\mu-1}} (b_\nu^2(k) - 2zb_\nu(k) + 1)^{-\frac{1}{2}}.$$

In general if the singularities of $B(k, z)$ are more complicated than poles a simple computation shows that $A(k, z)$ has a more involved behavior at its associated singular point.

Remark 2: In order to see what the connection between polar singularities of $A(k, z)$ (in the z-plane) are with those of the associated function $B(k, z)$ we may consider the following inverse integral relationship:[12]

$$B(k, z) = \tfrac{1}{2}(1 - z^2) \int_{-1}^{+1} \frac{A(k, \xi)\, d\xi}{(1 - 2\xi z + z^2)^{\frac{1}{2}}}.$$

Theorem 3: For each fixed k $A(k, z)$ converges uniformly and absolutely in any compact subset of the ellipse $E(\rho_0)$, where

$$\rho_0(k)^{-1} = \overline{\lim_{l\to\infty}} |a_l(k)|^{1/l}.$$

Remark 3: Having found the radius of convergence of the series for $B(k, z)$ [or which is the same thing the ellipse of convergence of $A(k, z)$] we are able to use either the Fabry or Mandelbrojt methods to determine the position of the singularity on the circle of convergence.

To illustrate these results we give several examples below.

Example 1: For the case where the scattering potential is a superposition of Yukawa potentials, i.e.,

$$V(r) = \int^{\infty} \rho(\mu) \frac{e^{-\mu r}}{r}\, d\mu, \qquad \mu_0 > 0,$$

[9] R. P. Gilbert, Arch. Ratl. Mech. Anal. 6, 171 (1960).
[10] R. P. Gilbert, J. Math. Phys. 5, 983 (1964).
[11] S. Bergman, Ergeb. Math. Grenzg., N.F. 23 (1961).
[12] See Ref. 7 for this result, and Ref. 9 for its extension to Gegenbauer series.

Newton has given the following asymptotic estimate[13]:

$$a_l(k) \approx \mathcal{O}(l^{-\frac{1}{2}}e^{-2l\alpha}),$$

where

$$\alpha = \ln\left[\frac{\mu_0}{2k} + \left(1 + \frac{\mu_0^2}{4k^2}\right)^{\frac{1}{2}}\right], \quad \text{and} \quad |k| < \mu_0.$$

From the Cauchy–Hadamard formula and the Fabry theorem we compute the singularity of $B(k, z)$ in the z plane to lie at

$$z = \{(\mu_0/2k) + [1 + (\mu_0^2/4k^2)]^{\frac{1}{2}}\}^2,$$

where this holds for those *complex* k such that $|k| < \mu_0$. We choose the branch of the square root so that it is positive for $k > 0$, i.e., for physical k.

Our Theorem 1 then shows that the scattering amplitude is singular at the point

$$\begin{aligned} z = \frac{1}{2}\Bigg\{&\left[\frac{\mu_0}{2k} + \left(1 + \frac{\mu_0^2}{4k^2}\right)^{\frac{1}{2}}\right]^2 \\ &+ \left[\frac{\mu_0}{2k} + \left(1 + \frac{\mu_0^2}{4k^2}\right)^{\frac{1}{2}}\right]^{-2}\Bigg\} \\ = \frac{1}{2}\Bigg\{&\left[\frac{\mu_0}{2k} + \left(1 + \frac{\mu_0^2}{4k^2}\right)^{\frac{1}{2}}\right]^2 \\ &+ \left[\frac{\mu_0}{2k^2} - \left(1 + \frac{\mu_0^2}{4k^2}\right)^{\frac{1}{2}}\right]^2\Bigg\} \\ = 1 + \frac{\mu_0^2}{2k^2}, \quad &|k| < \mu_0. \end{aligned}$$

This agrees with the known singularity in the momentum transfer plane at $t = 2k^2(z - 1) = \mu_0^2$, which initiates a cut. The nature of this singularity may be studied by using our Theorem 2; however, the computations are detailed and are being published separately.[14]

Example 2: In his book[15] Newton gives a general asymptotic expression for the partial waves which is valid providing the potential $V(r)$ decreases sufficiently fast as $r \to \infty$. It is not clear what the necessary conditions for this result are, but it is certainly true when $V(r)$ has compact support and is integrable. Let us assume the support of $V(r)$ lies in the interval $[0, R]$, then Newton's result becomes

$$a_l(k) \approx -i\pi\left(\frac{k}{2}\right)^{2l}\Gamma(l + 1)^{-2}\int_0^R V(r)r^{1+2l}\,dr.$$

By the mean value theorem for integrable functions[16] we have for some R_0, where $0 < R_0 < R$, that

$$\int_0^R V(r)r^{1+2l}\,dr = (R_0)^{1+2l}\int_0^R V(r)\,dr.$$

[An obvious variation is made when $V(r)$ is not integrable but $r^2V(r)$ is.] Hence,

$$\begin{aligned} \rho = (\overline{\lim_{l\to\infty}}\,|a_l(k)|^{1/l})^{-1} > \frac{1}{2}\left|\frac{2}{R_0k}\right|^2 \\ \times \lim_{l\to\infty}\left(\Gamma(l + 1)^2R_0^{-1}\left|\int_0^R V(r)\,dr\right|^{-1}\right)^{1/l}, \end{aligned}$$

which tends to ∞. From this we have that the scattering amplitudes $A(k, z)$, corresponding to integrable potentials with compact support are entire in the z plane.

Example 3: We consider the potentials

$$V(r) = \frac{1}{r^2 + \beta^2}, \qquad \beta > 0,$$

which lead to the partial wave estimate[17,18]

$$\begin{aligned} a_l(k) &\approx -i\pi\int_0^\infty (r^2 + \beta^2)^{-1}J_l(kr)^2r\,dr \\ &= -i\pi I_l(k\beta)K_l(k\beta), \end{aligned}$$

where $I(x)$, $K(x)$ are lth-order, modified Bessel functions of the first and third kind, respectively. After a short computation using known asymptotics for the modified Bessel functions,[19] we have

$$\begin{aligned} \frac{1}{l}\log I_l(k\beta) + \frac{1}{l}\log K_l(k\beta) \\ \approx \frac{2}{l}\log\left\{\Gamma(\tfrac{1}{2}) + \mathcal{O}\left(\frac{1}{k\beta}\right)\right\} \end{aligned}$$

or

$$\overline{\lim_{l\to\infty}}\,|I_l(k\beta)K_l(k\beta)|^{1/l} = 1 \quad \text{for physical } k.$$

From the Fabry theorem we conclude that the singularity of $B(k, z)$ is indeed at $z = 1$, and consequently the scattering amplitude is singular at $z = \frac{1}{2}[1 + \frac{1}{1}] = 1$. We remark, however, that in this case this is merely an estimate, since the general formula for $q_l(k)$, given above, is merely the Born approximation unless $V(r) \approx \mathcal{O}(e^{-\mu r})$, $\mu > 0$, as $r \to \infty$.

[13] R. Newton, *The Complex j-Plane* (W. A. Benjamin, Inc., New York, 1964), p. 44.

[14] R. P. Gilbert and H. C. Howard, Inst. Fluid Dynam. Appl. Math., University of Maryland, Technical Report 405 (1965).

[15] See Ref. 13, p. 43.

[16] M. E. Munroe, *Measure and Integration* (Addison-Wesley Publishing Company, Inc., Reading, Massachusetts, 1953), p. 184.

[17] See Ref. 13, p. 42.

[18] Bateman Manuscript Project, *Higher Transcendental Functions*, edited by A. Erdelyi (McGraw-Hill Book Company, Inc., New York, 1953), Vol. II, p. 96.

[19] See Ref. 18, p. 86.

On the Coupling and Recoupling of Relativistic States*

HAIM GOLDBERG†
Department of Physics and Astronomy, University of Rochester, Rochester, New York
(Received 5 October 1965)

Systems containing more than two relativistic particles are analyzed from the point of view of irreducible representations of the Poincaré group. Corresponding Clebsch–Gordan coefficients are calculated for three-particle systems, and recoupling functions (which are the analog of the Racah coefficients) are defined.

INTRODUCTION

THE pioneering work of Wigner[1] unveiled the nature of the unitary representations of the Poincaré group. The way then opened for the decomposition of direct products[2] and calculation of corresponding Clebsch–Gordan coefficients.[3] Naturally the next step is the calculation of the analog of the Racah coefficients.[4]

For the group R_3 of rotations in three dimensions, these coefficients are defined as follows: Given three tensors $\mathsf{T}^{(k)}$, $\mathsf{U}^{(l)}$, and $\mathsf{V}^{(m)}$, they may be coupled to yield a tensor of definite rank p according to the formula

$$X^{(p)}_{(n)\pi} = \sum_{\kappa\lambda\mu\nu} T^{(k)}_{\kappa} U^{(l)}_{\lambda} V^{(m)}_{\mu} (k\kappa l\lambda \mid n\nu)(n\nu m\mu \mid p\pi), \quad (1)$$

where $(a\alpha b\beta \mid c\gamma)$ are the usual[5] Clebsch–Gordan coefficients. In this case, $\mathsf{T}^{(k)}$ and $\mathsf{U}^{(l)}$ were coupled first, and then $\mathsf{V}^{(m)}$ was incorporated. Clearly, the alternative coupling

$$Y^{(p)}_{(n')\pi} = \sum_{\kappa\lambda\mu\nu'} T^{(k)}_{\kappa} U^{(l)}_{\lambda} V^{(m)}_{\mu} (k\kappa n'\nu' \mid p\pi)(l\lambda m\mu \mid n'\nu') \quad (2)$$

is as good as the former one. Each of the two sets $\{X^{(p)}_{(n)}\}$ and $\{Y^{(p)}_{(n')}\}$ (n and n' assuming all possible values) span the *same* linear space—the subspace of the direct product $\mathsf{T}^{(k)} \otimes \mathsf{U}^{(l)} \otimes \mathsf{V}^{(m)}$ which belongs to the eigenvalue $p(p+1)$ of J^2. Hence, the two bases are linearly dependent, and we may write

$$Y^{(p)}_{(n')} = \sum_{n} (-)^{k+l+m+p} [(2n'+1)(2n+1)]^{\frac{1}{2}} \times \begin{Bmatrix} k & l & n \\ m & p & n' \end{Bmatrix} X^{(p)}_{(n)}. \quad (3)$$

This implies

$$\sum_{\nu'} (k\kappa n'\nu' \mid p\pi)(l\lambda m\mu \mid n'\nu') = \sum_{n,\nu} (-)^{k+l+m+p} \times [(2n'+1)(2n+1)]^{\frac{1}{2}} \begin{Bmatrix} k & l & n \\ m & p & n' \end{Bmatrix} \times (k\kappa l\lambda \mid n\nu)(n\nu m\mu \mid p\pi). \quad (4)$$

Multiplying both sides of (4) by

$$\sum_{\bar{\nu}} (k\kappa l\lambda \mid \bar{n}\bar{\nu})(\bar{n}\bar{\nu} m\mu \mid p\pi),$$

summing over κ, λ, μ, and π, and using the orthogonality relations of the Clebsch–Gordan coefficients, one gets

$$(-)^{k+l+m+p} [(2\bar{n}+1)(2n'+1)]^{\frac{1}{2}} (2p+1) \begin{Bmatrix} k & l & \bar{n} \\ m & p & n' \end{Bmatrix} = \sum_{\kappa\lambda\mu\pi\nu'\bar{\nu}} (k\kappa l\lambda \mid \bar{n}\bar{\nu})(\bar{n}\bar{\nu} m\mu \mid p\pi) \times (k\kappa n'\nu' \mid p\pi)(l\lambda m\mu \mid n'\nu'). \quad (5)$$

Relations (3) and (4) reflect recoupling of tensors, whereas (5) enables to calculate the recoupling coefficients (or the Racah coefficients, or the 6-j symbols).

An equivalent consideration runs as follows: Given four tensors $\mathsf{T}^{(k)}$, $\mathsf{U}^{(l)}$, $\mathsf{V}^{(m)}$, $\mathsf{W}^{(p)}$, one may contract their direct product in two ways. Either

$$S_{(n)} = \sum_{\kappa\lambda\mu\pi\nu} T^{(k)}_{\kappa} U^{(l)}_{\lambda} V^{(m)}_{\mu} W^{(p)}_{\pi} \times (k\kappa l\lambda \mid n\nu)(m\mu p\pi \mid n-\nu) \quad (6)$$

or

$$\bar{S}_{(\bar{n})} = \sum_{\kappa\lambda\mu\pi\bar{\nu}} T^{(k)}_{\kappa} U^{(l)}_{\lambda} V^{(m)}_{\mu} W^{(p)}_{\pi} \times (k\kappa m\mu \mid \bar{n}\bar{\nu})(l\lambda p\pi \mid \bar{n}-\bar{\nu}). \quad (6')$$

* Supported in part by the U. S. Atomic Energy Commission.

† On leave of absence from the Soreq Nuclear Research Centre, Yavne, Israel.

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

[2] The decomposition of the direct product associated with many relativistic particles is described in the lectures of A. S. Wightman, "Invariance in Relativistic Quantum Mechanics," *Dispersion Relations and Elementary particles*, edited by C. deWitt and R. Omnes (John Wiley & Sons, Inc., New York, 1961), Sec. IV, p. 159. The present paper follows a different line of reasoning. For references dealing with two relativistic particles see Ref. 1 of A. J. Macfarlane, J. Math. Phys. 4, 490 (1963).

[3] A. J. Macfarlane, Rev. Mod. Phys. 34, 41 (1962). We shall generally refer to this paper as M, often making use of its results.

[4] U. Fano and G. Racah, *Irreducible Tensorial Sets* (Academic Press Inc., New York, 1959), Chap. 11.

[5] We write $(a\ \alpha\ b\ \beta \mid c\ \gamma)$ instead of $(a\ \alpha\ b\ \beta \mid a\ b\ c\ \gamma)$ as $a\ b$ on the right are obvious.

Again, the sets $\{S_{(n)}\}$ and $\{\bar{S}_{(\bar{n})}\}$ are linearly dependent, the dependence written in terms of 6-j symbols.

The last reasoning is analogous to that of crossing in the S-matrix theory, where four states are coupled into pairs in different ways and then contracted into scalars. Moreover, the coupling of three particles in different ways is closely related to the analysis of resonances which appear in elementary-particle reactions. This is illustrated by the following example. The cross sections for one-pion production in πN interactions are associated with the existence of N^* as well as of ρ. But it may well be that these resonances are related kinematically, namely, they may be derived from each other through the recoupling of the created pion. Otherwise phrased, it is the properties of the *same* invariant amplitudes which give rise both to N^* and to ρ. Thus, one is tempted to look for recoupling functions of the Poincaré group $\mathcal{P}$ as well.

The procedure in this case is, however, more complicated than that of R_3. First, infinite-dimensional representations are encountered, a fact which introduces into the problem integrations instead of summations. Secondly, two unitary irreducible representations are of different nature when the relevant little groups are different; and change in coupling may change the little groups involved. Finally, some irreducible representations which appear in a direct product may be equivalent; and the problem of choice of a scheme arises.

In the subsequent, only those cases will be treated in which all little groups involved are equivalent to R_3. In Sec. 2 we analyze systems containing more than two particles. An appropriate scheme of decomposition is suggested for corresponding direct products. This scheme is applied to systems of three particles (Sec. 3): Clebsch–Gordan coefficients are calculated, which do not emerge from those encountered in two-particle systems. Also, the recoupling functions are defined in Sec. 3.

In order not to interrupt the main line of the argument, some helpful calculations are given separately in two appendices.

REDUCTION OF DIRECT PRODUCTS

This subject was treated, among others, by Wightman[2] in his Les-Houches lectures. The idea runs as follows.

The states—in momentum-space representation—of a relativistic particle are essentially functionals defined over the hyperboloid $k^2 = \kappa^2$, $k_0 > 0$, where κ is the mass of the particle. The states of *two* particles are functionals over the direct product of the hyperboloids $k_1^2 = \kappa_1^2$, $k_2^2 = \kappa_2^2$; $k_{10}, k_{20} > 0$, where κ_1 and κ_2 are, respectively, the masses of the first and second particle. The manifold obtained is geometrically equivalent, from the point of view of $\mathcal{P}$, to the "direct product" of the ray $\kappa_1 + \kappa_2 \leq w \leq \infty$, the hyperboloid $r^2 = w^2$, and the "sphere" the points q of which fulfil $q \cdot r = q^2 + 1 = 0$. Quotation marks were inserted around the terms "direct product" and "sphere", since actually it is a direct *integral* of manifolds. For each w in the above ray there corresponds a hyperboloid $r^2 = w^2$; and to each point r on the hyperboloid a spheroid is attached, the points q of which fulfil $q \cdot r = q^2 + 1 = 0$. Thus, functionals over the direct product of the two hyperboloids are combinations of products of functionals over the ray, over the hyperboloids, and over corresponding spheroids.

The spheroid associated with the rest frame is a unit sphere. Since each spheroid is obtainable—through a one–one analytic function—from this sphere (e.g., using a pure Lorentz transformation), we may identify all spheroids with the unit sphere, labeling the points by the polar angles θ and φ. Moreover, the little groups of the spheroid and of the sphere become also identical. One may now use the Legendre functions $Y_m^l(\theta, \varphi)$ as an orthonormal basis for all functionals over the unit sphere. The advantage of this basis is, that it furnishes automatically bases for unitary irreducible representations (IRs) of the little group; and the spin indices may be coupled appropriately to these functions.

Adding a third particle with momentum k_3, $k_3^2 = \kappa_3^2$, $k_{30} > 0$, the above reasoning may be repeated: The hyperboloids $(k_1 + k_2)^2 = \epsilon^2$ and $k_3^2 = \kappa_3^2$ give rise to hyperboloids $(k_1 + k_2 + k_3)^2 = w^2 \geq \epsilon + \kappa_3$, and to corresponding spheres. Thus, from the point of view of the little group R_3 we have a direct product of *two* spheres: that one which is associated with particles 1 and 2, and the other which is encountered when particle 3 is added.

However, the direct product of two spheres is reducible under R_3, in a way which is made clear by the following considerations. Let $\mathbf{e}$ and $\mathbf{f}$ be the generic unit vectors of the first and second sphere, respectively, the angle in between being φ. $\cos\varphi = (\mathbf{e} \cdot \mathbf{f})$ is unchanged when the same rotation is applied both to $\mathbf{e}$ and to $\mathbf{f}$. Define the following right triad:

$$\mathbf{i} = [2 \sin(\varphi/2)]^{-1}\{\mathbf{e} - \mathbf{f}\};$$

$$\mathbf{j} = [2 \cos(\varphi/2)]^{-1}\{\mathbf{e} + \mathbf{f}\}; \mathbf{k} = [\sin\varphi]^{-1}[\mathbf{e} \times \mathbf{f}] \quad (7)$$

(We disregard the two singular cases $\varphi = 0$ and $\varphi = \pi$). This triad defines Euler angles—the angles associated with the transition from the basic triad $[\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z]$ to $[\mathbf{i}, \mathbf{j}, \mathbf{k}]$. At the same time, the Euler angles label the elements of the group R_3 of rotations in 3-dimensional space. Accordingly, it is more significant, geometrically, to regard the direct product of the two spheres as the direct product of the group R_3 with the segment $0 \leq \varphi \leq \pi$.[6] Rotations of the little group R_3 will leave each point in this segment invariant, whereas the points of R_3 will be effected: If a_0 is any element of R_3, rotation by R transforms it to Ra_0 (i.e., the little group appears here as the first parameter group).

We conclude that IRs which appear in the direct product $[\kappa_1\sigma_1] \otimes [\kappa_2\sigma_2] \otimes [\kappa_3\sigma_3]$ (σ_i referring to the spin of the ith particle) may be characterized by

$$\{[wJ](\epsilon, \varphi), \eta\alpha\}; \tag{8}$$

α stands for $\kappa_1\sigma_1\kappa_2\sigma_2\kappa_3\sigma_3$, and η for some other quantum numbers which specify the way of coupling the spin indices with the functions over R_3 to yield total angular momentum J (see next section).

Note that infinite multiplicity is associated with the angle φ, corresponding to the number of points in the segment $[0, \pi]$. Alternatively, since this segment is bounded, spaces which belong to equivalent IRs may be combined in such a way that only denumerably infinite multiplicity will occur through φ. Namely: let $\langle rM[wJ](\epsilon\varphi), \eta\alpha|$ denote a generic state of the IR (8), where r is the total 4-momentum and M the z-component of the total angular momentum. Then

$$\langle rM[wJ](\epsilon N), \eta\alpha| = \int_0^\pi d\varphi \sin\varphi [N + \tfrac{1}{2}]^{\frac{1}{2}} P_N(\cos\varphi) \times \langle rM[wJ](\epsilon\varphi), \eta\alpha| \tag{9}$$

is another state of the same transformation properties, characterized by the discrete value N rather than by φ.[7]

The addition of a fourth particle gives rise to another sphere, the angles $\theta_4\varphi_4$ of which may refer to the original triad $[\mathbf{i}, \mathbf{j}, \mathbf{k}]$. Thus, θ_4 and φ_4 remain unchanged under the operation of the little group R_3; and additional infinite degeneracy is encountered, this time corresponding to the points on a sphere. Again, similar to (9) we may write

$$\langle rM[wJ](\epsilon_2\epsilon_3; NQ_4q_4), \eta\alpha| = \int_0^\pi d\theta_4 \int_0^{2\pi} d\varphi_4 \sin\theta_4 Y^{Q_4}_{q_4}(\theta_4\varphi_4) \times \langle rM[wJ](\epsilon_2\epsilon_3; N\theta_4\varphi_4), \eta\alpha|. \tag{10}$$

In Eq. (10), $\epsilon_2^2 = (k_1 + k_2)^2$, $\epsilon_3^2 = (k_1 + k_2 + k_3)^2$, $w^2 = r^2 = (k_1 + k_2 + k_3 + k_4)^2$, $r = k_1 + k_2 + k_3 + k_4$, α stands for $\kappa_1\sigma_1\kappa_2\sigma_2\kappa_3\sigma_3\kappa_4\sigma_4$, and η specifies the way in which the spin indices and the functions over R_3 were coupled to total angular momentum J. In this way we have labeled the infinitely-many equivalent IRs by the discrete set NQ_4q_4.

The same procedure will be appropriate for any additional particle. Summarizing, the IRs which appear in the direct product

$$[\kappa_1\sigma_1] \otimes [\kappa_2\sigma_2] \otimes \cdots \otimes [\kappa_n\sigma_n] \qquad (n \geq 3) \tag{11}$$

may be characterized according to the scheme

$$\{[wJ](\epsilon_2\epsilon_3 \cdots \epsilon_{n-1}; NQ_4q_4 \cdots Q_nq_n), \eta\alpha\}. \tag{12}$$

In (12), $\epsilon_i^2 = (k_1 + k_2 + \cdots + k_i)^2$, $w^2 = (k_1 + k_2 + \cdots + k_n)^2$, J is the total angular momentum, N specifies the Legendre polynomial in $\cos\varphi$, (Q_iq_i) specify the Legendre function over the sphere which arises by the addition of the ith particle, α stands for $\kappa_1\sigma_1\kappa_2\sigma_2 \cdots \kappa_n\sigma_n$, and η defines the way in which the spin indices and the functions over R_3 were coupled to yield total angular momentum J.

THE CLEBSCH-GORDAN COEFFICIENTS FOR THREE PARTICLES AND THE RECOUPLING FUNCTIONS

The Clebsch–Gordan coefficients for two particles were calculated, among others, by A. J. Macfarlane.[3] The idea runs as follows:

Let $\langle k'\nu'k''\nu''|$ be the generic state of the direct product $[\kappa'\sigma'] \otimes [\kappa''\sigma'']$. If this is to be given as a combination of states of the form $\langle rM[wJ], \eta\alpha|$, then translational covariance implies

$$r = k' + k''. \tag{13}$$

In view of the relation (A12) it is plausible to construct the state

$$\langle (k'\nu'k''\nu'')r, \alpha| = \sum_{\mu'\mu''} D^{\sigma'}_{\nu'\mu'}\Big(R^{-1}(k', L^{-1}(r))\Big) \times D^{\sigma''}_{\nu''\mu''}\Big(R^{-1}(k'', L^{-1}(r))\Big) \langle k'\mu'k''\mu'', \alpha| \tag{14}$$

where $L(r)$ is a pure Lorentz transformation associated with the 4-momentum r, and for each $\Lambda \in \mathfrak{L}_+^\uparrow$, $R(k, \Lambda)$ is defined by M (3.4). The state

[6] The reason for the upper limit of φ to be π and not 2π is given in Appendix B.

[7] It is shown in Appendix B that the weight $\sin\varphi$ is attached to the segment $(0, \pi)$. The functions $(N + \frac{1}{2})^{\frac{1}{2}} P_N(\cos\varphi)$ were chosen for the sake of orthonormality.

(14) fulfils

$$\langle (k'\nu' k''\nu'')r, \alpha |\ U(0, \Lambda)$$
$$= \sum_{\mu'\mu''} D^{\sigma'}_{\nu'\mu'}(R(r, \Lambda)) D^{\sigma''}_{\nu''\mu''}(R(r, \Lambda))$$
$$\times \langle (\Lambda^{-1}k'\mu'\Lambda^{-1}k''\mu'')\Lambda^{-1}r, \alpha |. \quad (15)$$

Namely, the two spins undergo transformations which represent the *same* rotation $R(r, \Lambda)$.

Let q be, as in the previous section, a 4-vector coplanar with k' and k'', and $q\cdot r = q^2 + 1 = 0$. Let $\mathbf{e}$ be the spatial unit vector which fulfils

$$(0, \mathbf{e}) = L(r)q. \quad (16)$$

The six variables r, $\mathbf{e}$ may be used instead of the six variables k', k'' (with $k'^2 = \kappa'^2$, $k''^2 = \kappa''^2$) such that we may write

$$\langle (k'\nu' k''\nu'')r, \alpha | = \langle r\mathbf{e}\nu'\nu'', \alpha |. \quad (17)$$

It follows from (A7) that when k', $k'' \to \Lambda^{-1}k'$, $\Lambda^{-1}k''$, $\mathbf{e}$ undergoes the rotation

$$\mathbf{e} \to \mathbf{e}' = R^{-1}(r, \Lambda)\mathbf{e}. \quad (18)$$

Thus, one is led to project out of $\langle r\mathbf{e}\nu'\nu'', \alpha |$ the component parallel to $Y^{l*}_{l_z}(\mathbf{e})$, which transforms according to the representation l of $R(r, \Lambda)$. This projection is carried out by multiplying (17) with $Y^l_{l_z}(\mathbf{e})$ and integrating over the unit sphere.

The three angular momenta may now be coupled to some definite J; the final result will be a state of the form $\langle rM[wJ], \eta\alpha |$, where $M = \nu' + \nu'' + l_z$, $w^2 = r^2$, J is the total angular momentum, and η specifies the way in which σ', σ'' and l were coupled to J. From the above procedure we get[8]

$$\langle rM[wJ], \eta\alpha |\ k'\nu' k''\nu'', \alpha\rangle = 2w^{\frac{1}{2}}[\lambda(w^2, \kappa'^2, \kappa''^2)]^{-\frac{1}{4}}$$
$$\times \delta(w - [(k' + k'')^2]^{\frac{1}{2}})2r_0\delta(\vec{r} - \vec{k}' - \vec{k}'')$$
$$\times P(k'k'', \nu'\nu'', JM\eta) \quad (19)$$

where the function λ is defined by

$$\lambda(a, b, c) = a^2 + b^2 + c^2 - 2(ab + ac + bc), \quad (20)$$

and

$$P(k'k'', \nu'\nu'', JM\eta)$$
$$= \sum_{\mu'\mu''l_z s_z} D^{\sigma'}_{\mu'\nu'}\Big(R^{-1}(k', L^{-1}(r))\Big)$$
$$\times D^{\sigma''}_{\nu''\mu''}\Big(R^{-1}(k'', L^{-1}(r))\Big) Y^l_{l_z}(\mathbf{e})$$
$$\times (\sigma'\mu'\sigma''\mu'' \mid ss_z)(ll_z ss_z \mid JM). \quad (21)$$

As for the decomposition of representations associated with 3 particles, the scheme defined in section 2 differs from that of Macfarlane.[3] However, the above reasoning may be applied. As a first step define

$$\langle (k'\nu' k''\nu'' k'''\nu''')r, \alpha |$$
$$= \sum_{\mu'\mu''\mu'''} D^{\sigma'}_{\nu'\mu'}\Big(R^{-1}(k', L^{-1}(r))\Big)$$
$$\times D^{\sigma''}_{\nu''\mu''}\Big(R^{-1}(k'', L^{-1}(r))\Big)$$
$$\times D^{\sigma'''}_{\nu'''\mu'''}\Big(R^{-1}(k''', L^{-1}(r))\Big)$$
$$\times \langle k'\mu' k''\mu'' k'''\mu''', \alpha |, \quad (22)$$

where $r = k' + k'' + k'''$. In the second step an appropriate set of variables is chosen. The ones used in Ref. 3 are

$$w, \epsilon, r, \mathbf{e}_2 \text{ and } \mathbf{f}, \quad (23)$$

where $w^2 = r^2$, $\epsilon^2 = (k' + k'')^2$, $\mathbf{e}_2$ is the spatial unit vector associated with k' and k'', and $\mathbf{f}$ the unit vector associated with $k' + k''$ and k''' [see (A8)]. However, when k', k'', $k''' \to \Lambda^{-1}k'$, $\Lambda^{-1}k''$, $\Lambda^{-1}k'''$, $\mathbf{e}_2$ is rotated by $R^{-1}(k' + k'', \Lambda)$ rather than by $R^{-1}(r, \Lambda)$. Therefore define

$$\mathbf{e} = R^{-1}(k' + k'', L^{-1}(r))\mathbf{e}_2. \quad (24)$$

It follows from (A7) and (A12) that $\mathbf{e}$ and $\mathbf{f}$ undergo the same rotation $R^{-1}(r, \Lambda)$. We choose accordingly the set of variables

$$w, \epsilon, r, \mathbf{e}, \mathbf{f} \quad (25)$$

instead of the set (23). Again, let $\cos\varphi$, $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$ be defined in terms of $\mathbf{e}$ and $\mathbf{f}$ as in the previous section, and let ζ, θ, and ψ be the Euler angles associated with the triad $[\mathbf{i}, \mathbf{j}, \mathbf{k}]$. The set of variables

$$w, \epsilon, r, \varphi, \zeta, \theta, \text{ and } \psi \quad (26)$$

may be chosen instead of the set (25). Accordingly one may write

$$\langle (k'\nu' k''\nu'' k'''\nu''')r, \alpha | = \langle r\zeta\theta\psi(\epsilon\varphi)\nu'\nu''\nu''', \alpha |. \quad (27)$$

In the third step one projects out of (27) the component parallel to some tensor, the transformation properties of which are well known. In view of the above analysis, as well as that of Sec. 2, it is plausible to use $(4\pi)^{-1}[(2L+1)(2N+1)]^{\frac{1}{2}}P_N(\cos\varphi)D^L_{lm}(\zeta, \theta, \psi)$.

[8] See M (3.17), M (3.26)–M (3.29). We discussed here the form of the function P only; the origin of the other factors on right-hand side of (19) is made clear in Ref. (3). In particular, $2w^{\frac{1}{2}}[\lambda(w^2, k'^2, k''^2)]^{-\frac{1}{4}}$ is a normalization factor assuring the orthonormality of the right-hand side of (19).

Note that (21) differs from M (3.17) in that it involves $R^{-1}(k_i, L^{-1}(r))$ rather $R(k_i, L(r))$; these rotations do not equal each other.

The four angular momenta[9] σ', σ'', σ''', and L may then be coupled to some J.

It is shown in Appendix B that the transition from the variables (23) to the variables (26) does not effect the normalization used in Ref. (3), Sec. 6, in the case of 3 particles. Hence

$$\begin{aligned}&\langle 2M[wJ](\epsilon N),\ \eta\alpha \mid (k'\nu'k''\nu'')\epsilon,\ k'''\nu''',\ \alpha\rangle\\&= 2w^{\frac{1}{2}}[\lambda(w^2,\epsilon^2,\kappa'''^2)]^{-\frac{1}{4}}2\epsilon^{\frac{1}{2}}[\lambda(\epsilon^2,\kappa'^2,\kappa'''^2)]^{-\frac{1}{4}}\\&\quad\times\delta(w-[(k'+k''+k''')^2]^{\frac{1}{2}})\\&\quad\times 2r_0\delta(\vec{r}-\vec{k}'-\vec{k}''-\vec{k}''')\\&\quad\times Q(k'k''k'''(\epsilon N),\nu'\nu''\nu''',JM\eta)\end{aligned}\tag{28}$$

where[9]

$$\begin{aligned}&Q(k'k''k'''(\epsilon N),\nu'\nu''\nu''',JM\eta)\\&=\sum(4\pi)^{-1}[(2L+1)(2N+1)]^{\frac{1}{2}}\\&\quad\times D^{\sigma'}_{\mu',\nu'}(R^{-1}(k',L^{-1}(r)))D^{\sigma''}_{\mu'',\nu''}(R^{-1}(k'',L^{-1}(r)))\\&\quad\times D^{\sigma'''}_{\mu''',\nu'''}(R^{-1}(k''',L^{-1}(r)))D^{L}_{lm}(\zeta,\theta,\psi)\\&\quad\times P_N(\cos\varphi)(\sigma'\mu'\sigma''\mu''\mid ss_z)\\&\quad\times(\sigma'''\mu'''ss_z\mid tt_z)(Lltt_z\mid JM).\end{aligned}\tag{29}$$

The sum in (29) is over $\mu'\mu''\mu'''$, ls_zt_z; η stands for $stLm$. Evidently, numerous alternative ways of coupling may be as good.

The Q's fulfill the orthogonality relations[10]

$$\begin{aligned}&\sum_{\nu'\nu''\nu'''}\int d\varphi\,d\zeta\,d\theta\,d\psi\sin\varphi\sin\theta\\&\qquad\times Q^*(k'k''k'''(\epsilon N),\nu'\nu''\nu''',JM\eta)\\&\qquad\times Q(k'k''k'''(\epsilon N'),\nu'\nu''\nu''',J'M'\eta')\\&\quad=\delta_{NN'}\delta_{\eta\eta'}\delta_{JJ'}\delta_{MM'},\end{aligned}\tag{30}$$

$$\begin{aligned}&\sum_{\nu'\nu''\nu'''M}Q^*(k'k''k'''(\epsilon N),\nu'\nu''\nu''',JM\eta)\\&\quad\times Q(k'k''k'''(\epsilon N'),\nu'\nu''\nu''',JM\eta')\\&=(4\pi)^{-2}(2J+1)\delta_{\eta\eta'}[(2N+1)(2N'+1)]^{\frac{1}{2}}\\&\sum_K(2K+1)\begin{pmatrix}N&N'&K\\0&0&0\end{pmatrix}^2P_K(\cos\varphi).\end{aligned}\tag{31}$$

It is clear that the definition of $\mathbf{e}$ and $\mathbf{f}$ (or of φ, ζ, θ, and ψ) depends on the order of coupling. That is, reversing the order of k' and k'' yields a pair $\{\mathbf{e}',\mathbf{f}'\}$ of unit vectors related to $\{\mathbf{e},\mathbf{f}\}$ by

$$\mathbf{e}'=-\mathbf{e},\qquad \mathbf{f}'=\mathbf{f}.\tag{32}$$

This implies

$$\cos\varphi'=-\cos\varphi,\qquad \zeta'=\zeta,\qquad \theta'=\theta-\pi,\qquad \psi'=-(\psi+\tfrac{1}{2}\pi).\tag{33}$$

Inserting (33) in (29) one gets

$$\begin{aligned}Q(k''k'k'''(\epsilon N),\nu''\nu'\nu''',JM\eta)&=(-)^{\sigma'+\sigma''+s+N+L}i^m\\&\quad\times Q(k'k''k'''(\epsilon N),\nu'\nu''\nu''',JM\bar{\eta})\end{aligned}\tag{34}$$

where $\eta=[stLm]$, $\bar{\eta}=[stL-m]$.

The situation is more complicated in case k' and k''' are first coupled, and then k'' added. However, the associated vectors $\mathbf{e}'$, $\mathbf{f}'$ also undergo the rotation $R^{-1}(r,\Lambda)$ when k', k'', $k'''\to\Lambda^{-1}k'$, $\Lambda^{-1}k''$, $\Lambda^{-1}k'''$. This implies that the scalar products $(\mathbf{e}\cdot\mathbf{e}')$, $(\mathbf{e}\cdot\mathbf{f}')$, $(\mathbf{f}\cdot\mathbf{e}')$ and $(\mathbf{f}\cdot\mathbf{f}')$ are Lorentz-invariant. Hence, the triad $[\mathbf{i}',\mathbf{j}',\mathbf{k}']$ associated with $\{\mathbf{e}',\mathbf{f}'\}$ is obtained from $[\mathbf{i},\mathbf{j},\mathbf{k}]$ by a rotation the Euler angles ω, τ, ξ of which are Lorentz-invariant. Moreover: It is shown in Appendix B that $\cos\omega$, $\cos\tau$, and $\cos\xi$, as well as $\cos\varphi'$ are algebraic functions of the Lorentz-invariants

$$\kappa',\ \kappa'',\ \kappa''',\ w,\ \epsilon,\text{ and }\cos\varphi.\tag{35}$$

Thus one may write

$$D^L_{lm}(\zeta',\theta',\psi')=\sum_n D^L_{ln}(\zeta,\theta,\psi)D^L_{nm}(\omega,\tau,\xi)\tag{36}$$

where $D^L_{nm}(\omega,\tau,\xi)$ are algebraic functions of the variables (35).

In general, let $(k_ak_bk_c)$ be any permutation of $(k'k''k''')$. The recoupling functions—associated with the $(ab)c\to(ac)b$ recoupling—may be defined through the relations

$$\begin{aligned}&Q\Big(k_ak_ck_b(\epsilon(acb)N'),\nu_a\nu_c\nu_b,JM\eta(acb)\Big)\\&=\sum_{N\eta(abc)}R_1\Big(w\epsilon(abc)J\alpha;N'N\eta(acb)\eta(abc)\Big)\\&\quad\times Q\Big(k_ak_bk_c(\epsilon(abc)N),\nu_a\nu_b\nu_c,JM\eta(abc)\Big);\end{aligned}\tag{37}$$

(37) is solvable for the functions R_1 using (30), (36), the unitarity of the D's as well as the orthogonality and recoupling relations of the CGc's

[9] Since R_3 in this context is the first parametric group, a rotation R transforms $D_{lm}{}^L$ into $\sum_n D_{ln}(R)D_{nm}{}^L$. Accordingly, there are $(2L+1)$ independent linear bases for the representation L of R_3 in the space of functions over R_3; and we choose m, $-L\le m\le L$ as a label for these bases.

[10] Relations (30) and (31) follow from the unitarity of the D matrices, the orthogonality of the CGc's of R_3, and the relation

$$P_N(\cos\varphi)P_{N'}(\cos\varphi)=\textstyle\sum_K(2K+1)\begin{pmatrix}N&N'&K\\0&0&0\end{pmatrix}^2P_K(\cos\varphi)$$

where $\begin{pmatrix}N&N'&K\\0&0&0\end{pmatrix}$ is conventional $3-j$ symbol.

of R_3. The result reads

$$R_1(w\epsilon(abc)J\alpha;\, N'N\eta(acb)\eta(abc))$$
$$= (-)^{2t}(4\pi)^{-2}(2J+1)[(2s+1)(2s'+1)]^{\frac{1}{2}}$$
$$\times \begin{Bmatrix} \sigma_a & \sigma_b & s \\ t & \sigma_c & s' \end{Bmatrix} \delta(tt')\,\delta(LL')R(w\epsilon(abc)\alpha;\, m'mN'N) \quad (38)$$

where

$$R(w\epsilon(abc)J\alpha;\, m'mN'N)$$
$$= \int_0^\pi d\varphi(abc)\,\sin\varphi(abc)[(2N+1)(2N'+1)]^{\frac{1}{2}}$$
$$\times D^L_{m',m}(\omega,\tau,\xi)P_{N'}(\cos\varphi(acb))P_N(\cos\varphi(abc)) \quad (39)$$

In (38), $\eta(abc)$ stands for $[stLm]$ and $\eta(acb)$ for $[s't'L'm']$. According to the above, R is an integral of an algebraic function; however, we shall not deal in the present paper with its analytic properties, nor with its symmetries.

Note that the analogy between the functions R and the Racah coefficients leads one to look for certain relations among these functions. Namely, the product of the two recouplings

$$(ab)c \to (ac)b, \qquad (ac)b \to (ab)c$$

yields the original situation; whereas the product of

$$(ab)c \to (ac)b, \qquad (ac)b \to (bc)a$$

is equivalent to the recoupling $(ab)c \to (bc)a$. These imply for the R's analogs of (11.15) and (11.16) in Ref. (4). Similarly, one is tempted to derive something like the Bidenharn identity (I.3), Ref. (4), by defining first the analog of the $9j$-symbols. Again these subjects will not be treated here.

ACKNOWLEDGMENT

The author is indebted to Dr. P. Thurnauer for helpful discussions. The stimulating encouragement of Professor R. E. Marshak and Professor S. Okubo is gratefully acknowledged.

APPENDIX A

It is convenient, for explicit calculations, to use the 2×2 matrix-representation of 4-vectors[2] operated on left and right by the corresponding representation of the orthochronous proper Lorentz group. We shall use, in this Appendix, the $\times$ symbol to indicate matrix multiplication, a dot for Lorentz scalar product and a dot within round brackets—for bold-type vectors—in case of usual 3-dimensional scalar product. In cases of possible ambiguity, Lorentz scalar products will also appear within round brackets.

1. Let $a = a_0 + (\mathbf{A}\cdot\boldsymbol{\sigma})$; denote $\tilde{a} = a_0 - (\mathbf{A}\cdot\boldsymbol{\sigma})$. We have

$$a \times \tilde{a} = \tilde{a} \times a = a^2 = \alpha^2, \quad (A1)$$

$$a \times b = (a_0 + (\mathbf{A}\cdot\boldsymbol{\sigma})) \times (b_0 + (\mathbf{B}\cdot\boldsymbol{\sigma}))$$
$$= a\cdot\tilde{b} + a_0(\mathbf{B}\cdot\boldsymbol{\sigma}) + b_0(\mathbf{A}\cdot\boldsymbol{\sigma}) + i(\mathbf{A},\mathbf{B},\boldsymbol{\sigma}), \quad (A2)$$

$$(\mathbf{A}\cdot\mathbf{B}) = \tfrac{1}{2}\,\mathrm{tr}\,(\mathbf{A}\cdot\boldsymbol{\sigma})\times(\mathbf{B}\cdot\boldsymbol{\sigma}), \quad (A2')$$

$$a \times \tilde{b} \times c \times \tilde{b}$$
$$= 2(b\cdot c)\{a\cdot b + b_0(\mathbf{A}\cdot\boldsymbol{\sigma}) - a_0(\mathbf{B}\cdot\boldsymbol{\sigma}) - i(\mathbf{A},\mathbf{B},\boldsymbol{\sigma})\}$$
$$- b^2\{a\cdot c + c_0(\mathbf{A}\cdot\boldsymbol{\sigma}) - a_0(\mathbf{C}\cdot\boldsymbol{\sigma}) - i(\mathbf{A},\mathbf{C},\boldsymbol{\sigma})\}. \quad (A3)$$

Let $L(a)$ be a pure Lorentz transformation which transforms the 4-momentum a to its rest frame: $L(a) \times a \times L(a) = \alpha$. Then[2]

$$L(a) = N_a(\alpha + \tilde{a}); \qquad L^{-1}(a) = N_a(\alpha + a)$$
$$\text{where} \quad N_a = [2\alpha(\alpha + a_0)]^{-\frac{1}{2}}, \quad (A4)$$

$$L^2(a) = \alpha^{-1}\tilde{a}; \qquad L^{-2}(a) = \alpha^{-1}a. \quad (A5)$$

Direct calculation yields

$$a^\Delta = L(b) \times a \times L(b) = \beta^{-1}a\cdot b + (\mathbf{A}\cdot\boldsymbol{\sigma})$$
$$+ [\beta(\beta + b_0)]^{-1}[a\cdot b - a_0\beta](\mathbf{B}\cdot\boldsymbol{\sigma}) \quad (A6)$$

so that $a_0^\Delta = (a\cdot b)/\beta$ remains invariant when $a \to \Lambda^{-1} \times a \times \Lambda^{-1\dagger}$, $b \to \Lambda^{-1} \times b \times \Lambda^{-1\dagger}$. The transformation induced on the 4-vector a^Δ is, according to M(3.4),

$$a^\Delta \to a^{\Delta\prime} = L(\Lambda^{-1} \times b \times \Lambda^{-1\dagger}) \times \Lambda^{-1} \times L^{-1}(b)$$
$$\times a^\Delta \times L^{-1}(b) \times \Lambda^{-1\dagger} \times L(\Lambda^{-1} \times b \times \Lambda^{-1\dagger})$$
$$= D^{\frac{1}{2}}(R^{-1}(b,\Lambda)) \times a^\Delta \times D^{\frac{1}{2}\dagger}(R^{-1}(b,\Lambda)), \quad (A7)$$

i.e. the spatial component of a^Δ is rotated by $R^{-1}(b,\Lambda)$.

Denote $\epsilon^2 = (a+b)^2$; then (A.6) implies

$$L(a+b) \times a \times L(a+b)$$
$$= (2\epsilon)^{-1}\{[\epsilon^2 + \alpha^2 - \beta^2] + [\lambda(\epsilon^2,\alpha^2,\beta^2)]^{\frac{1}{2}}(\mathbf{e}\cdot\boldsymbol{\sigma})\} \quad (A8)$$

where λ is defined by (20) and $\mathbf{e}$ is a spatial unit vector. We refer to $\mathbf{e}$ as the unit vector associated with the 4-vectors a and b.

2. A useful relation reads [see M(3.4) and (10)]

$$D^{\frac{1}{2}}(R^{-1}(a, L^{-1}(b)))$$
$$= L(L(b) \times a \times L(b)) \times L(b) \times L^{-1}(a)$$
$$= \alpha N_{a^\Delta}[L(b) \times L^{-1}(a) + L^{-1}(b) \times L(a)]$$
$$= [2\beta(\alpha + a\cdot b/\beta)(\alpha + a_0)(\beta + b_0)]^{-\frac{1}{2}}$$
$$\times \{(\alpha + a_0)(\beta + b_0) - (\mathbf{A}\cdot\mathbf{B}) + i(\mathbf{A},\mathbf{B},\boldsymbol{\sigma})\}. \quad (A9)$$

Thus, $R^{-1}(a, L^{-1}(b))$ is a rotation around an axis

parallel to $[\mathbf{A}\times\mathbf{B}]$; and if φ is the angle of rotation, then

$$\cos\tfrac{1}{2}\varphi = [2\beta(\alpha + a\cdot b/\beta)(\alpha + a_0)(\beta + b_0)]^{-\frac{1}{2}} \times \{(\alpha + a_0)(\beta + b_0) - (\mathbf{A}\cdot\mathbf{B})\}. \quad \text{(A10)}$$

It is readily verified that

$$(\Lambda^{-1}\times a\times\Lambda^{-1\dagger})^{\sim} = \Lambda^{\dagger}\times\tilde{a}\times\Lambda. \quad \text{(A11)}$$

Relations (A5, 9, 11), as well as M(3.4), imply

$$D^{\frac{1}{2}\dagger}(R(b,\Lambda))\times D^{\frac{1}{2}}(R^{-1}(a,L^{-1}(b)))\times D^{\frac{1}{2}}(R(a,\Lambda)) = D^{\frac{1}{2}}(R^{-1}(\Lambda^{-1}\times a\times\Lambda^{-1\dagger}, L^{-1}(\Lambda^{-1}\times b\times\Lambda^{-1\dagger}))). \quad \text{(A12)}$$

3. Let a, b, and c be three 4-momenta. Denote[11]

$$x = a + b, \quad y = a + c, \quad r = a + b + c;$$
$$q(a,b) = \xi[\lambda(\xi^2,\alpha^2,\beta^2)]^{-\frac{1}{2}} \times \{a - b + [(\alpha^2-\beta^2)/\xi^2](a+b)\}, \quad \text{(A13)}$$

and similar definitions for $q(a,c)$, $q(x,c)$, and $q(y,b)$. The q's fulfill

$$q^2(a,b) + 1 = q^2(a,c) + 1 = q^2(x,c) + 1 = q^2(y,b) + 1 = 0,$$
$$q(a,b)\cdot x = q(a,c)\cdot y = q(x,c)\cdot r = q(y,b)\cdot r = 0. \quad \text{(A14)}$$

Let $\mathbf{e}_2$ be the spatial unit vector associated with a and b, $\mathbf{e}_2'$ with a and c, $\mathbf{f}$ with x and c, and $\mathbf{f}'$ with y and b. These vectors are related to the respective q's through relations similar to (16). Denote

$$(\mathbf{e}\cdot\boldsymbol{\sigma}) = D^{\frac{1}{2}}(R^{-1}(x,L^{-1}(r)))\times(\mathbf{e}_2\cdot\boldsymbol{\sigma})\times D^{\frac{1}{2}\dagger}(R^{-1}(x,L^{-1}(r))) = \rho\xi[(\rho+\xi)^2-\gamma^2]^{-1} \times\{L(r)\times L^{-1}(x) + L^{-1}(r)\times L(x)\}\times L(x)\times q(a,b)\times L(x) \times\{L(x)\times L^{-1}(r) + L^{-1}(x)\times L(r)\}, \quad \text{(A15)}$$

$$(\mathbf{e}'\cdot\boldsymbol{\sigma}) = D^{\frac{1}{2}}(R^{-1}(y,L^{-1}(r)))\times(\mathbf{e}_2'\cdot\boldsymbol{\sigma})\times D^{\frac{1}{2}\dagger}(R^{-1}(y,L^{-1}(r))) = \rho\eta[(\rho+\eta)^2-\beta^2]^{-1} \times\{L(r)\times L^{-1}(y) + L^{-1}(r)\times L(y)\}\times L(y)\times q(a,c)\times L(y) \times\{L(y)\times L^{-1}(r) + L^{-1}(y)\times L(r)\}. \quad \text{(A16)}$$

[11] Greek letters denote masses or invariant energies of corresponding 4-momenta; e.g., $\gamma^2 = c^2$, $\xi^2 = x^2$, etc.

By definition

$$(\mathbf{f}\cdot\boldsymbol{\sigma}) = L(r)\times q(x,c)\times L(r), \quad \text{(A17)}$$
$$(\mathbf{f}'\cdot\boldsymbol{\sigma}) = L(r)\times q(y,b)\times L(r). \quad \text{(A18)}$$

Using (A2'), (A5), and some simple manipulations we have

$$\cos\varphi = (\mathbf{e}\cdot\mathbf{f}) = \rho\xi[2((\rho+\xi)^2-\gamma^2)]^{-1} \operatorname{tr}\{q(x,c)\times[\rho^{-1}\tilde{r}+\xi^{-1}\tilde{x}] \times q(a,b)\times[\rho^{-1}\tilde{r}+\xi^{-1}\tilde{x}]\}, \quad \text{(A19)}$$

$$\cos\varphi' = (\mathbf{e}'\cdot\mathbf{f}') = \rho\eta[2((\rho+\eta)^2-\beta^2)]^{-1} \operatorname{tr}\{q(y,b)\times[\rho^{-1}\tilde{r}+\eta^{-1}\tilde{y}] \times q(a,c)\times[\rho^{-1}\tilde{r}+\eta^{-1}\tilde{y}]\}, \quad \text{(A20)}$$

$$(\mathbf{e}\cdot\mathbf{e}') = \rho^2\xi\eta[2((\rho+\xi)^2-\gamma^2)((\rho+\eta)^2-\beta^2)]^{-1} \operatorname{tr}\{\rho^{-1}r\times[\rho^{-1}\tilde{r}+\xi^{-1}\tilde{x}]\times q(a,b)\times[\rho^{-1}\tilde{r}+\xi^{-1}\tilde{x}] \times\rho^{-1}r\times[\rho^{-1}\tilde{r}+\eta^{-1}\tilde{y}]\times q(a,c)\times[\rho^{-1}\tilde{r}+\eta^{-1}\tilde{y}]\}, \quad \text{(A21)}$$

$$(\mathbf{e}\cdot\mathbf{f}') = \rho\xi[2((\rho+\xi)^2-\gamma^2)]^{-1} \operatorname{tr}\{q(y,b)\times[\rho^{-1}\tilde{r}+\xi^{-1}\tilde{x}] \times q(a,b)\times[\rho^{-1}\tilde{r}+\xi^{-1}\tilde{x}]\}, \quad \text{(A22)}$$

$$(\mathbf{e}'\cdot\mathbf{f}) = \rho\eta[2((\rho+\eta)^2-\beta^2)]^{-1} \operatorname{tr}\{q(x,c)\times[\rho^{-1}\tilde{r}+\eta^{-1}\tilde{y}] \times q(a,c)\times[\rho^{-1}\tilde{r}+\eta^{-1}\tilde{y}]\}, \quad \text{(A23)}$$

$$(\mathbf{f}\cdot\mathbf{f}') = (2\rho^2)^{-1}\operatorname{tr} q(x,c)\times\tilde{r}\times q(y,b)\times\tilde{r}. \quad \text{(A24)}$$

It follows from (A.11) that $\operatorname{tr}\{a_1\times\tilde{a}_2\times a_3\times\tilde{a}_4\times\cdots\times\tilde{a}_{2n}\}$ is invariant under the transformation $a_i\to\Lambda^{-1}\times a_i\times\Lambda^{-1\dagger}$ $(i = 1, 2, \cdots, 2n)$. This proves the invariance of (A19–24) under Lorentz transformations.

Using (A.3), it is possible to solve (A.19) for η^2:

$$\eta^2 = (2\xi^2)^{-1}\{[\lambda(\rho^2,\xi^2,\gamma^2)\lambda(\xi^2,\alpha^2,\beta^2)]^{\frac{1}{2}}\cos\varphi + [\xi^2(\alpha^2+\beta^2+\gamma^2+\rho^2-\xi^2) - (\alpha^2-\beta^2)(\rho^2-\gamma^2)]\} \quad \text{(A25)}$$

Inspection on relations (A.20–24) implies that their right-hand sides are expressible as algebraic functions of the variables

$$\alpha^2,\ \beta^2,\ \gamma^2,\ \rho^2,\ \xi^2, \text{ and } \cos\varphi. \quad \text{(A26)}$$

Let $\mathbf{i}'$, $\mathbf{j}'$, and $\mathbf{k}'$ be defined in terms of $\mathbf{e}'$ and $\mathbf{f}'$ in a way analogous to (7). It is clear from the above that the scalar products $(\mathbf{i}\cdot\mathbf{i}')$, $(\mathbf{i}\cdot\mathbf{j}')$, $\cdots$, $(\mathbf{k}\cdot\mathbf{k}')$ are algebraic functions of the variables (A.26). At the same time, these scalar products are the matrix elements of the rotation which transforms

the triad [i, j, k,] onto [i′, j′, k′]. Since the cosines of the associated Euler angles are algebraic functions of these matrix elements, we conclude that they are algebraic functions of the variables (A26) as well.

APPENDIX B

1. The 4-momentum $k' + k''$ (as well as k''') is expressible in terms of the variables $\epsilon^2 = (k'+k'')^2$, $\kappa'''^2 = k'''^2$, r and $\mathbf{f}$. Accordingly, the rotation $R^{-1}(k' + k'', L^{-1}(r))$ is independent of $\mathbf{e}_2$.

Let α_2, β_2 be the polar angles of $\mathbf{e}_2$; α, β those of $\mathbf{e}$; and

$$\mathbf{e} = R^{-1}(k' + k'', L^{-1}(r))\mathbf{e}_2. \tag{B1}$$

It is readily checked that $\partial(\alpha_2, \beta_2)/\partial(\alpha, \beta) = \sin\alpha/\sin\alpha_2$, and therefore

$$\sin\alpha_2\, d\alpha_2\, d\beta_2 = \sin\alpha\, d\alpha\, d\beta, \quad \text{or} \quad d\Omega(\mathbf{e}_2) = d\Omega(\mathbf{e}). \tag{B2}$$

2. Let $\mathbf{e}$ and $\mathbf{f}$ be two independent unit vectors; α, β the polar angles of $\mathbf{e}$; γ, δ those of $\mathbf{f}$. Then

$$\begin{aligned}\mathbf{e} &= (\sin\alpha\cos\beta, \sin\alpha\sin\beta, \cos\alpha),\\ \mathbf{f} &= (\sin\gamma\cos\delta, \sin\gamma\sin\delta, \cos\gamma).\end{aligned} \tag{B3}$$

The unit vector

$$\lambda = [\cos^2\alpha + \cos^2\gamma - 2\cos\alpha\cos\gamma\cos\varphi]^{-\frac{1}{2}} \times \{\mathbf{e}\cdot\cos\gamma - \mathbf{f}\cdot\cos\alpha\} \tag{B4}$$

lies on the intersection of the (e, f) plane π with the xy plane, since its z-component vanishes. In (B4), as before, $\cos\varphi = (\mathbf{e}\cdot\mathbf{f})$.

The unit vector

$$\mathbf{y} = (\sin\varphi)^{-1}[\cos^2\alpha + \cos^2\gamma - 2\cos\alpha\cos\gamma\cos\varphi]^{-\frac{1}{2}} \times \{\mathbf{e}(\cos\gamma\cos\varphi - \cos\alpha) + \mathbf{f}(\cos\alpha\cos\varphi - \cos\gamma)\} \tag{B5}$$

is perpendicular to λ and lies in the plane π.

Let i, j, and k be defined as in (7). By definitions, the z axis, k, and $\mathbf{y}$ are perpendicular to λ; hence they are coplanar. Also, $(\mathbf{y}\cdot\mathbf{k}) = 0$.

Three Euler angles ζ, θ, and ψ associated with [i, j, k] are defined as follows: First, rotate around $\mathbf{e}_z$ through an angle ζ such that $\mathbf{e}_y$ coincides with λ; then rotate around λ through an angle θ such that $\mathbf{e}_z$ coincides with k; finally, rotate around k through ψ such that λ coincides with j.

It follows that

$$\sin\theta = \mu_2 = (\sin\varphi)^{-1} \times [\cos^2\alpha + \cos^2\gamma - 2\cos\alpha\cos\gamma\cos\varphi]^{\frac{1}{2}}, \tag{B6}$$

$$\cot\zeta = \frac{\lambda_y}{\lambda_x} = \frac{\cos\gamma\sin\alpha\sin\beta - \cos\alpha\sin\gamma\sin\delta}{\cos\gamma\sin\alpha\cos\beta - \cos\alpha\sin\gamma\cos\delta}, \tag{B7}$$

and therefore

$$\cos\theta = \pm(\sin\varphi)^{-1}[1 + 2\cos\alpha\cos\gamma\cos\varphi - \cos^2\alpha - \cos^2\gamma - \cos^2\varphi]^{\frac{1}{2}}, \tag{B6′}$$

$$\cos\psi = (\lambda\cdot\mathbf{j}) = [\cos^2\alpha + \cos^2\gamma - 2\cos\alpha\cos\gamma\cos\varphi]^{-\frac{1}{2}}(\cos\gamma - \cos\alpha)\cos\tfrac{1}{2}\varphi. \tag{B8}$$

Evidently,

$$\cos\varphi = \cos\alpha\cos\gamma + \sin\alpha\sin\gamma\cos(\beta - \delta). \tag{B9}$$

In order to calculate the volume element

$$d\Omega(\mathbf{e})\, d\Omega(\mathbf{f}) = \sin\alpha\sin\gamma\, d\alpha\, d\beta\, d\gamma\, d\delta \tag{B10}$$

in terms of angles φ, ζ, θ, and ψ, it is helpful to define

$$\begin{aligned}\bar\alpha &= \alpha + \gamma, \qquad \bar\gamma = \alpha - \gamma,\\ \bar\beta &= \tfrac{1}{2}(\beta + \delta), \qquad \bar\delta = \tfrac{1}{2}(\beta - \delta).\end{aligned} \tag{B11}$$

It is readily checked that the Jacobian of this transformation is 1, and

$$d\Omega(\mathbf{e})\, d\Omega(\mathbf{f}) = \tfrac{1}{2}(\cos\bar\gamma - \cos\bar\alpha)\, d\bar\alpha\, d\bar\beta\, d\bar\gamma\, d\bar\delta. \tag{B12}$$

Also,

$$\cos\varphi = r(\bar\alpha, \bar\gamma, \bar\delta) = \cos\bar\gamma\cos^2\bar\delta + \cos\bar\alpha\sin^2\bar\delta, \tag{B9′}$$

$$\sin\theta = \left[\frac{s(\bar\alpha, \bar\gamma, \bar\delta)}{1 - r^2(\bar\alpha, \bar\gamma, \bar\delta)}\right]^{\frac{1}{2}} = \left[\frac{1 - (\cos^2\bar\gamma\cos^2\bar\delta + \cos^2\bar\alpha\sin^2\bar\delta)}{1 - (\cos\bar\gamma\cos^2\delta + \cos\bar\alpha\sin^2\bar\delta)^2}\right]^{\frac{1}{2}}, \tag{B6″}$$

$$\cot\zeta = \frac{\sin\bar\alpha\tan\bar\beta + \sin\bar\gamma\cot\bar\delta}{\sin\bar\alpha - \sin\bar\gamma\tan\bar\beta\cot\bar\delta}, \tag{B7′}$$

$$\cos\psi = \left[\frac{u(\bar\alpha, \bar\gamma)[1 + r]}{s}\right]^{\frac{1}{2}} = \left[\frac{\tfrac{1}{2}(1 - \cos\bar\alpha)(1 - \cos\bar\gamma)(1 + \cos\bar\gamma\cos^2\bar\delta + \cos\bar\alpha\sin^2\bar\delta)}{1 - (\cos^2\bar\alpha\sin^2\upsilon + \cos^2\bar\gamma\cos^2\bar\delta)}\right]^{\frac{1}{2}}. \tag{B8′}$$

The meaning of r, s, and u is obvious.

It is readily checked that

$$\left|\frac{\partial(\alpha,\beta,\gamma,\delta)}{\partial(\zeta,\theta,\psi,\varphi)}\right| = \left|\frac{\partial(\cot\zeta,\sin\theta,\cos\psi,\cos\varphi)}{\partial(\bar\beta,\bar\alpha,\bar\gamma,\bar\delta)}\right|^{-1} \times \frac{1}{\sin^2\zeta}\,|\cos\theta\sin\psi\sin\varphi| \tag{B13}$$

(vertical lines indicate absolute value).
Since θ, φ, and ψ do not depend on $\bar\beta$, we have

$$\left|\frac{\partial(\cot\zeta,\sin\theta,\cos\psi,\cos\varphi)}{\partial(\bar\beta,\bar\alpha,\bar\gamma,\bar\delta)}\right| = \left|\frac{\partial\cot\zeta}{\partial\bar\beta}\cdot\frac{\partial(\sin\theta,\cos\psi,\cos\varphi)}{\partial(\bar\alpha,\bar\gamma,\bar\delta)}\right|. \tag{B14}$$

Let x be any of the variables $\bar\alpha$, $\bar\gamma$, $\bar\delta$; then

$$(\cos\varphi)_x = r_x; \qquad (\sin\theta)_x = \frac{1}{2}\sin\theta\left\{\frac{s_x}{s} + \frac{2rr_x}{1-r^2}\right\};$$

$$(\cos\psi)_x = \frac{1}{2}\cos\psi\left\{\frac{u_x}{u} + \frac{r_x}{1+r} - \frac{s_x}{s}\right\}. \tag{B15}$$

Hence

$$\left|\frac{\partial(\sin\theta,\cos\psi,\cos\varphi)}{\partial(\bar\alpha,\bar\gamma,\bar\delta)}\right| = \frac{1}{4}\left|\frac{\sin\theta\cos\psi}{su}\cdot\frac{\partial(r,s,u)}{\partial(\bar\alpha,\bar\gamma,\bar\delta)}\right|$$

$$= |(8su)^{-1}\sin\theta\cos\psi(1-\cos\varphi)\sin\bar\alpha\sin\bar\gamma \times (\cos\bar\gamma-\cos\bar\alpha)^2\sin 2\bar\delta|, \tag{B16}$$

$$\left|\frac{\partial\cot\zeta}{\partial\bar\beta}\right| = \frac{s}{\sin^2\bar\delta(\sin\bar\alpha\cos\bar\beta - \sin\bar\gamma\sin\bar\beta\cot\bar\delta)^2} = \frac{1}{\sin^2\zeta}. \tag{B17}$$

Hence

$$\left|\frac{\partial(\alpha,\beta,\gamma,\delta)}{\partial(\zeta,\theta,\psi,\varphi)}\right| = \left|\frac{8su\cos\theta\sin\psi\sin\varphi}{\sin\theta\cos\psi(1-\cos\varphi)\sin\bar\alpha\sin\bar\gamma(\cos\bar\gamma-\cos\bar\alpha)^2\sin 2\bar\delta}\right|. \tag{B18}$$

The following relations hold (according to B.6″, 7′, 8′, 9′):

$$s = \sin^2\theta\sin^2\varphi$$

$$= 1 - (\cos\bar\alpha+\cos\bar\gamma)\cos\varphi + \cos\bar\alpha\cos\bar\gamma, \tag{B19}$$

$$u = \tfrac{1}{2}(1-\cos\bar\alpha)(1-\cos\bar\gamma)$$

$$= \sin^2\theta\cos^2\psi(1-\cos\varphi), \tag{B20}$$

$$(\cos\bar\gamma-\cos\bar\alpha)\sin 2\bar\delta = 2\cos\theta\sin\varphi, \tag{B21}$$

$$\cos\bar\alpha+\cos\bar\gamma = \sin^2\theta(1+\cos\varphi-2\cos^2\psi). \tag{B22}$$

Hence

$$|\sin\bar\alpha\sin\bar\gamma| = |[2u(1+\cos\bar\alpha)(1+\cos\bar\gamma)]^{\frac{1}{2}}| = |4\sin^2\theta\sin\psi\cos\psi\sin\varphi|, \tag{B23}$$

and therefore

$$\left|\frac{\partial(\alpha,\beta,\gamma,\delta)}{\partial(\zeta,\theta,\psi,\varphi)}\right| = \left|\frac{\sin\theta\sin\varphi}{\cos\bar\gamma-\cos\bar\alpha}\right|. \tag{B24}$$

Geometrically, φ varies from 0 to 2π, or from $-\pi$ to π. However, rotation around the bisector of **e** and **f** through an angle of π changes φ into $-\varphi$. Accordingly, we limit the variation of φ from 0 to π, and multiply the Jacobian by 2. Therefore, by (B.12),

$$d\Omega\,(\mathbf{e})\,d\Omega\,(\mathbf{f}) = \sin\theta\sin\varphi\,d\zeta\,d\theta\,d\psi\,d\varphi = d\Omega\,(R_3)\,d\,(-\cos\varphi) \tag{B25}$$

where $0 \le \theta, \varphi \le \psi, 0 \le \zeta, \psi \le 2\pi$.

Integrating over the two spheres of **e** and **f**—using either (B.10) or (B.25)—yields $16\pi^2$, which checks the relation (B.25).

3. (B.2) and (B.25) imply that the normalization of the CGc's in the scheme of Sec. 2 is the same as the one used by Macfarlane,[3] provided the projection functions are orthonormal (using the weight function $\sin\theta\sin\varphi$), as are the products of two $P(k_1k_2,\ \nu_1\nu_2,\ JM\eta)$ (when the weight function $\sin\alpha\sin\gamma$ is used).

Some Spatially Homogeneous Anisotropic Relativistic Cosmological Models

R. KANTOWSKI* AND R. K. SACHS†
The University of Texas, Austin, Texas
(Received 20 July 1965)

Some solutions of the Einstein field equations for a dust source are given in explicit form. They are spatially homogeneous, irrotational, and anisotropic. They can be characterized as those spatially homogeneous expanding models that do not permit a simply transitive three-parameter group of motions. The models are compared in detail with observations and with the Friedmann models. In a few instances slightly longer time scales are obtained with the present models than from the corresponding Friedmann models.

INTRODUCTION

SOME relativistic dust models are given that are spatially homogeneous, have shear, and have no rotation. Because the metrics have very simple explicit forms they are useful for judging what effects shear has on cosmological evolution. The models that have, at some time, an expansion, acceleration parameter, and shear consistent with current observations are qualitatively rather similar to Friedmann models. Despite their simplicity the models apparently have not appeared explicitly in the literature.[1]

We consider solutions of the Einstein field equations without cosmological constant for a dust source

$$G_b^a = -\rho u^a u_b. \tag{1}$$

Here u^a is a timelike vector field normalized by $u^a u_a = 1$. We use the units $c = 1$, $8\pi k = 1$ throughout. Suppose the space–time permits an isometry group transitive on three-dimensional spacelike hypersurfaces. Most such metrics belong to the Taub–Heckmann–Schücking class[2] because there exists at least one 3-parameter isometry group which is simply transitive on the hypersurfaces. We here consider only metrics which permit no simply transitive 3-parameter groups so that the methods of Ref. 2 are not applicable. From the results of Petrov[3] it follows that there are only two Lie Algebras to consider, both four-dimensional:

Case 1:

$$[Z_1, Z_2] = Z_3, \quad [Z_2, Z_3] = Z_1, \quad [Z_3, Z_1] = Z_2, \quad [Z_4, Z_i] = 0 \quad (i = 1, 2, 3,); \tag{2}$$

Case 2:

$$[Z_1, Z_2] = Z_1, \quad [Z_2, Z_3] = Z_3, \quad [Z_3, Z_1] = 2Z_2, \quad [Z_4, Z_i] = 0 \quad (i = 1, 2, 3,). \tag{3}$$

The operators Z_i, $(i = 1, 2, 3)$, must operate on two-dimensional subspaces.

EQUATIONS AND THEIR SOLUTIONS

The resulting fields are spatially homogeneous and nonisotropic. Because of the assumed symmetries there exist, locally, coordinates

$$x^0 = t, \quad x^1 = r, \quad x^2 = \varphi, \quad x^3 = \theta$$

such that

$$u^a = \alpha\delta_0^a + \beta\delta_1^a,$$

and the metric has the respective forms[3]:

Case 1:

$$ds^2 = dt^2 - X^2(t)\,dr^2 - Y^2(t)\,d\Omega^2, \quad \text{where } d\Omega^2 = \sin^2\theta\,d\varphi^2 + d\theta^2; \tag{4}$$

Case 2:

$$ds^2 = dt^2 - X^2(t)\,dr - Y^2(t)\,d\tau^2, \quad \text{where } d\tau^2 = \sinh^2\theta\,d\varphi^2 + d\theta^2. \tag{5}$$

The group generators are[3]

Case 1:

$$Z_4 = \frac{\partial}{\partial r}, \quad Z_1 = \frac{\partial}{\partial\varphi},$$
$$Z_2 = \sin\varphi\frac{\partial}{\partial\theta} + \cot\varphi\cos\varphi\frac{\partial}{\partial\varphi}, \tag{6}$$
$$Z_3 = \cos\varphi\frac{\partial}{\partial\theta} - \cot\theta\sin\varphi\frac{\partial}{\partial\varphi};$$

---

* NSF Coop Fellow.

† Research supported by Aerospace Research Labs, OAR, AF-33(615) 1029.

[1] See the review of matter solutions by J. Ehlers in the proceedings of the London (1965) Conference on General Relativity.

[2] L. Witten, *Gravitation* (John Wiley & Sons, Inc., New York, 1962), Chap. 11.

[3] A. S. Petrov, *Einstein-Raume* (Akademie-Verlag, Berlin, 1964). Cases 1 and 2 are classified as G_4-VIII and VII, respectively, by Petrov, pp. 60, 225, and 226.

Case 2:

$$Z_4 = \frac{\partial}{\partial r}, \quad Z_1 = -\cos\varphi \frac{\partial}{\partial \theta} + (\coth\theta \sin\varphi - 1)\frac{\partial}{\partial \varphi},$$

$$Z_2 = \sin\varphi \frac{\partial}{\partial \theta} + \coth\theta \cos\varphi \frac{\partial}{\partial \varphi}, \tag{7}$$

$$Z_3 = \cos\varphi \frac{\partial}{\partial \theta} - (\coth\theta \sin\varphi + 1)\frac{\partial}{\partial \varphi}.$$

The field equations become

Case 1:

$$T_0^0 = 2\frac{\dot{X}\dot{Y}}{XY} + \frac{1+\dot{Y}^2}{Y^2}, \tag{8a}$$

$$T_1^1 = 2\frac{\ddot{Y}}{Y} + \frac{1+\dot{Y}^2}{Y^2}, \tag{9a}$$

$$T_2^2 = T_3^3 = \frac{\ddot{X}}{X} + \frac{\ddot{Y}}{Y} + \frac{\dot{X}\dot{Y}}{XY}, \tag{10a}$$

$$T_j^i = 0 \qquad (i \neq j); \tag{11a}$$

Case 2:

$$T_0^0 = \frac{\dot{X}\dot{Y}}{XY} - \frac{1-\dot{Y}^2}{Y^2}, \tag{8b}$$

$$T_1^1 = 2\frac{\ddot{Y}}{Y} - \frac{1-\dot{Y}^2}{Y^2}, \tag{9b}$$

$$T_2^2 = T_3^3 = \frac{\ddot{X}}{X} + \frac{\ddot{Y}}{Y} + \frac{\dot{X}\dot{Y}}{XY}, \tag{10b}$$

$$T_j^i = 0 \qquad (i \neq j). \tag{11b}$$

The field equations for Case 1 are given by Tolman, Bondi, and others[4]; however, the further computations of these authors do not apply in our case due to the fact that $\partial Y/\partial r = 0$.

Equations (1), (11a), and (11b) require $u^a = \delta_0^a$. The remaining equations are

$$T_0^0 = \rho, \qquad T_j^i = 0 \qquad (i, j = 1, 2, 3). \tag{12}$$

They are quite easily solved by first solving (9a) and (9b) and then treating Y as an independent variable in (10a) and (10b). The solutions are best expressed in terms of a function $\eta(t)$:

Case 1: Closed Solution,

$$X = \epsilon + (\epsilon\eta + b)\tan\eta, \tag{13}$$

$$Y = a\cos^2\eta, \tag{14}$$

$$t - t_0 = a(\eta + \tfrac{1}{2}\sin 2\eta), \tag{15}$$

$$\rho = \epsilon \sec^4\eta/a^2[1 + (\eta + b)\tan\eta]; \tag{16}$$

here b, a, and ϵ are constants satisfying

$$\epsilon = 0, 1, \qquad -\infty < a < \infty, \qquad a \neq 0,$$
$$-\tfrac{1}{2}\pi \leq b < 0; \tag{17}$$

Case 2: Open Solution (a),

$$X = \epsilon - (\epsilon\eta + b)\tanh\eta, \tag{18}$$

$$Y = a\cosh^2\eta, \tag{19}$$

$$t - t_0 = a(\eta + \tfrac{1}{2}\sinh 2\eta), \tag{20}$$

$$\rho = -\epsilon \operatorname{sech}^4\eta/a^2[1 - (\eta + b)\tanh\eta]; \tag{21}$$

Open Solution (b),

$$X = \epsilon - (\epsilon\eta + b)\coth\eta, \tag{22}$$

$$Y = a\sinh^2\eta, \tag{23}$$

$$t - t_0 = a(\eta - \tfrac{1}{2}\sinh 2\eta), \tag{24}$$

$$\rho = -\epsilon \operatorname{csch}^4\eta/a^2[1 - (\eta + b)\coth\eta]. \tag{25}$$

For case 2, a and ϵ have the same ranges as for case 1, but b satisfies $0 \leq b < \infty$.

Computing the conformal tensor shows that these metrics are type D with principal null directions in the r, t plane.[5] For $\epsilon = 0$, Case 1 is the vacuum Schwarzschild solution inside the "singularity":

$$ds^2 = dY^2/(aY^{-1} - 1) - (aY^{-1} - 1)\,dr^2 - Y^2\,d\Omega^2, \qquad 1 \leq a/Y. \tag{26}$$

For $\epsilon = 0$ Case 2 reduces to the known vacuum solutions[6]:

Open solution (a),

$$ds^2 = dY^2/(1 - aY^{-1}) - (1 - aY^{-1})\,dr^2 - Y^2\,d\tau^2, \qquad 0 \leq a/Y < 1; \tag{27}$$

Open solution (b),

$$ds^2 = dY^2/(1 + aY^{-1}) - (1 + aY^{-1})\,dr^2 - Y^2\,d\tau^2, \qquad 0 < a/Y. \tag{28}$$

COSMOLOGICAL CONSIDERATIONS

For $\epsilon \neq 0$, Cases 1 and 2 resemble the Friedmann solutions[7] for vanishing pressure. In the Friedmann case there are both open and closed solutions; however, our closed metric, Case 1, is not periodic in time as is the corresponding closed Friedmann universe. Our open models suffer the same difficulty as the open Friedmann models—the magnitude of

[4] H. Bondi, M. N. **107**, 401 (1947), and references listed there.

[5] Compare F. Pirani, *Brandeis Lectures on General Relativity* (Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1964), Vol. 1, p. 321.

[6] Reference 3, p. 173.

[7] L. D. Landau and E. Lifshitz, *Classical Theory of Fields* (Addison-Wesley Publishing Company, Inc., Reading, Massachusetts, 1962), p. 375.

TABLE I. Representative "good" points. Shear σ, expansion h_1, acceleration parameter α, proper time T since last singularity, and time T_F for a corresponding Friedmann model since the last singularity are listed for some selected values of a, b, and proper time t. 1 H = 10^{10} years in units for which $c = 8\alpha k = 1$.

a	b	t	σ	h_1	$-\alpha$	T	T_F
0.4 H	−1.9	1.0 H	.25 H^{-1}	1.09 H^{-1}	2.0 H^{-2}	.38 H	.47 H
1.0	−1.7	2.05	.06	1.03	.91	.48	.58
−0.7	−1.6	−.52	.57	1.01	1.56	.58	.52
−0.8	−1.55	−.67	.40	1.00	1.19	.57	.55
−0.7	−1.5	−.57	.36	1.00	1.24	.53	.54
−0.6	−1.4	−.47	.24	1.02	1.28	.47	.53
−0.3	−1.0	−.15	.16	1.01	2.66	.32	.43
−0.22	−.8	−.09	.10	1.02	4.27	.26	.37

the acceleration parameter is too small for consistency with current observations.

A more detailed comparison of the models in Case 1 with current observations can be made if we take some numerical values[8,9] for the Hubble constant h_1 and the acceleration parameter $\alpha = h_2/h_1^2$:

$$h_1 = 10^{-10}(\text{years})^{-1}, \tag{29}$$

$$\alpha \simeq -2. \tag{30}$$

Moreover, the fact that the red shift vs luminosity curve shows no gross anisotropies gives an upper limit on the shear σ, defined by

$$2\sigma^2 = (u_{i;j} - h_{ij}u^k{}_{;k}/3)(u^{i;j} - h^{ij}u^m{}_{;m}/3), \tag{31}$$

where

$$h_{ij} = g_{ij} - u_iu_j.$$

Current observations suggest $\sigma \leq .35 \times 10^{-10}$ (years)$^{-1}$. With σ small it is reasonable to identify the observed values of h_1 and α with the theoretical values averaged over angles on the celestial sphere. Then[9]

$$h_1 = u^a{}_{;a}/3, \tag{32}$$

$$\alpha = -(\rho/6 + 2\sigma^2)/h_1^2. \tag{33}$$

Because current observational estimates of ρ vary so widely we do not assume an independent value for it.

Machine calculations were used to get values of a, b, and t in the closed models for which α, h_1, and σ are fairly close to the above values. Some of the points are given in Table I. This table also contains the proper time T since the last singularity and the corresponding proper time T_F for that pressure-free Friedmann model which has the same h_1 and α. Of course all our models have singularities, as follows from the Raychaudhuri equation.[10] The singularities for the closed models are of two kinds. We may visualize the two kinds by suppressing the angle φ in the metric and visualizing three-space as the surface of a cylinder. In one kind of singularity the cylinder squashes to a disk, in the other it contracts to a line. If ρ and h_1

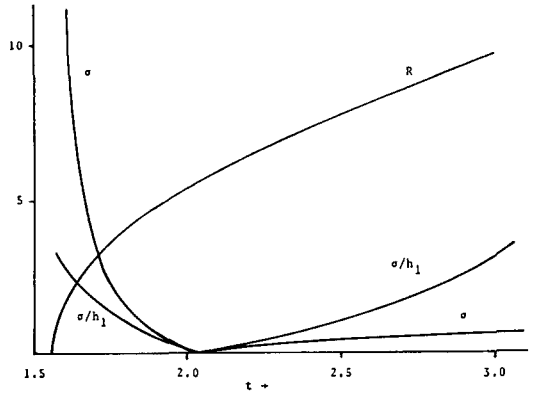

FIG. 1. Plots of R, σ, and σ/h_1 as functions of time for a = 1.0 H b = −1.7, and t_0 = 0. σ is discontinuous at σ = 0 because σ is defined as a positive square root. The units of R, σ, and t are 0.1 $H^{\frac{1}{2}}$, H^{-1}, and H, respectively. σ/h_1 is unitless and 1 H = 10^{10} years.

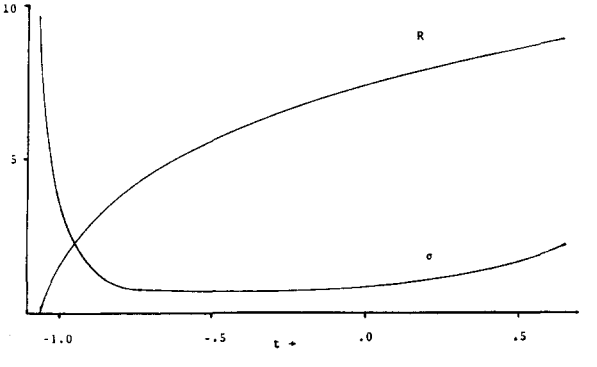

FIG. 2. Plots of R and σ as functions of t for a = −0.7 H, b = −1.6, and t_0 = 0. The units of R, σ, and t are 0.1 $H^{\frac{1}{2}}$, H^{-1}, and H, respectively, where 1 H = 10^{10} years.

[8] G. C. McVittie, *Encyclopedia of Physics* (Springer-Verlag, Berlin, 1959), Vol. LIII, p. 445.

[9] J. Kristian and R. K. Sachs, Ap. J., Feb. 1966.

[10] Raychaudhuri, Phys. Rev. 98, 1123 (1955).

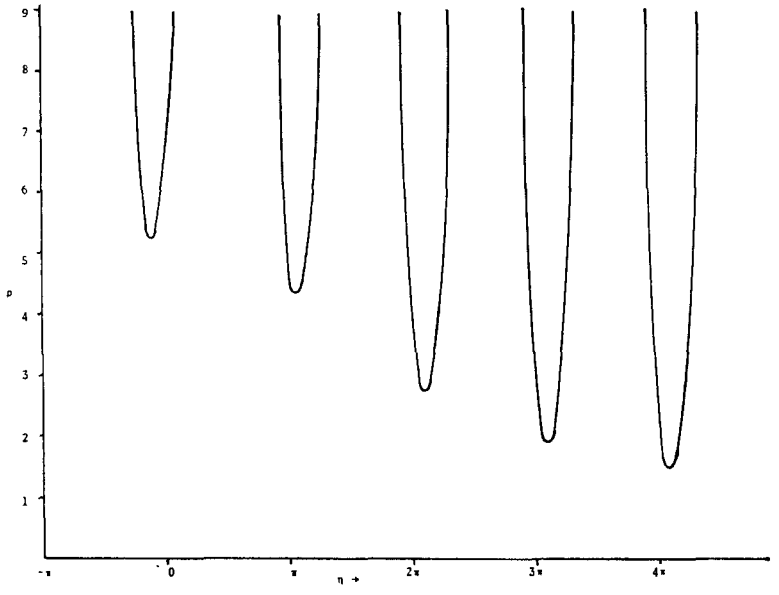

FIG. 3. A plot of ρ as a function of η for $a = 0.4$ H, $b = -1.18$. ρ is in units of H^{-2} where $H = 10^{10}$ years.

are given $T_F > T$, but with α and h_1 given it can happen that $T > T_F$.

Also given below are some graphs of R, σ, σ/h_1, and ρ for Case 1. Figures 1 and 2 indicate the tendency of σ to become very large for small values of R, and Fig. 3 shows the overall structure of ρ as a function of η. $R = \rho^{-\frac{1}{2}}$.

The main qualitative conclusion to be drawn from the foregoing data is that a shear which is presently small does not really have a large effect on the available time scale or on the observed values of h_1 and α. However, at times much earlier or later than the present, σ becomes much larger than h_1 in all our models, as is indicated in the graphs. Large values of σ at early stages in the evolution of the universe would critically affect the mass density and with it all other physical parameters for these early stages.

Case 1 has been found independently by Kip S. Thorne in his thesis, *Geometrodynamics of Cylindrical Systems*, Princeton University, May 1965, (unpublished).

ACKNOWLEDGMENTS

We are indebted to Professor R. Kerr, Professor J. Ehlers, and Professor E. Schücking for their advice during the course of this work.

Special Functions of Mathematical Physics from the Viewpoint of Lie Algebra

BRURIA KAUFMAN
Department of Applied Mathematics, Weizmann Institute, Rehovot, Israel
(Received 24 May 1965)

Families of special functions, known from mathematical physics, are defined here by their recursion relations. The operators which raise and lower indices in these functions are considered as generators of a Lie algebra. The general element of the corresponding Lie group thus operates on the function in two ways: on the one hand it shifts the argument of the function; on the other hand it produces an infinite sum of functions (at the unchanged argument) with shifted indices. Equating the two results of the operation gives us "addition theorems," hitherto derived by analytical methods. The present paper restricts itself to the study of 2- and 3-parameter Lie groups.

I. INTRODUCTION

A.

THE study of special functions has always made use of techniques from the theory of analytic functions. And indeed, there are properties of the special functions which by their nature can only be handled analytically (e.g., integral representations). However, it will be seen that a great many of the properties which appear in the form of summations (expansions, addition theorems) may be derived without any recourse to analytic methods. These methods are replaced here by the powerful concepts of the Lie algebra (or Lie group) related to the special function under consideration. The method by which addition theorems are obtained is extremely simple and straightforward; many other expansion theorems are then derived from the addition theorems.

The basic idea of the Lie group is that from an "infinitesimal operator" $\mathfrak{M}$, which shifts the point $s \equiv (x, y, z, \cdots)$ to a neighboring point $(s + ds)$, one may generate a "finite operator" $\exp \alpha\mathfrak{M}$, which shifts the point s into a point s' at a finite distance along the "path curve" of the one-parameter group $\exp \alpha\mathfrak{M}$. If the same operator is applied to a function of the coordinates $F(s)$, it will operate on the argument of the function:

$$\exp \alpha\mathfrak{M}\cdot F(s) = F(s') \equiv F(\exp \alpha\mathfrak{M}\cdot s). \tag{1}$$

In what follows, we consider those infinitesimal differential operators which appear in the recursion-relations for the various special functions, and generate from them finite operators. Any pair of recursion relations for the special functions may be rewritten in the form

$$\begin{aligned} \mathfrak{R}\cdot F_n &= \rho_n F_{n+1}, \\ \mathfrak{L}\cdot F_n &= \lambda_n F_{n-1}, \end{aligned} \tag{2}$$

where $\mathfrak{R}$ and $\mathfrak{L}$ are differential operators ("raising" and "lowering" operators for the index n) and ρ_n, λ_n are constants. Hence $\mathfrak{R}^m\cdot F_n$ is defined for any integral m, so that one may apply the operators $\exp \alpha\mathfrak{R}$, $\exp \alpha\mathfrak{L}$ to the functions $F_n(s)$:

$$\begin{aligned} \exp \alpha\mathfrak{R}\cdot F_n(s) &= \sum_{m=0}^{\infty} \frac{\alpha^m}{m!} (\rho_n\rho_{n+1} \cdots \rho_{n+m-1}) F_{n+m}(s), \\ \exp \alpha\mathfrak{L}\cdot F_n(s) &= \sum_{m=0}^{\infty} \frac{\alpha^m}{m!} (\lambda_n\lambda_{n-1} \cdots \lambda_{n-m+1}) F_{n-m}(s). \end{aligned} \tag{3}$$

These series may or may not terminate, depending on the particular values of ρ_n, λ_n.

Equations (1) and (3) may now be combined in various ways. To begin with we have two "addition theorems," one due to the operator $\mathfrak{R}$:

$$F_n(\exp \alpha\mathfrak{R}\cdot s) = F_n(s') = \sum_{m=0}^{\infty} \frac{\alpha^m}{m!} \left\{ \prod_{i=0}^{m-1} \rho_{n+i} \right\} F_{n+m}(s), \tag{4a}$$

and a similar one due to the operator $\mathfrak{L}$:

$$F_n(\exp \alpha\mathfrak{L}\cdot s) = F_n(s'') = \sum_{m=0}^{\infty} \frac{\alpha^m}{m!} \left\{ \prod_{i=0}^{m-1} \lambda_{n-i} \right\} F_{n-m}(s). \tag{4b}$$

Further, we can try to combine $\mathfrak{R}$ and $\mathfrak{L}$, e.g., in the form $\exp (\alpha\mathfrak{R} + \beta\mathfrak{L})$. Now, if $\mathfrak{R}$ and $\mathfrak{L}$ commute, i.e., $[\mathfrak{R}, \mathfrak{L}] = 0$, we have the simple "composition law"

$$\exp \alpha\mathfrak{R}\cdot \exp \beta\mathfrak{L} = \exp (\alpha\mathfrak{R} + \beta\mathfrak{L});$$

however, when $[\mathfrak{R}, \mathfrak{L}] \neq 0$ then the left-hand side differs from the right-hand side, and depends upon the successive commutators of $\mathfrak{R}$ and $\mathfrak{L}$. The set of such operators which is closed under commutation constitutes the Lie algebra generated by $\mathfrak{R}$, $\mathfrak{L}$. Any member $\mathfrak{M}$ of the algebra generates a finite operator $\exp \alpha\mathfrak{M}$; the products of these generate the Lie group corresponding to this Lie algebra. The composition law within the Lie group is uniquely

determined by the commutation rules of the Lie algebra generated by $\mathfrak{R}$ and $\mathfrak{L}$. It will be seen that the algebras for all functions considered below are quite simple and immediately identifiable.

B.

We consider any set of functions as *defined* by its recursion relations, and derive all properties of the functions from these relations. This is similar to the point of view taken by Infeld and Hull in their paper on the factorization method.[1] In their work Infeld and Hull show that (under very mild restrictions) every second-order differential equation of the form

$$[(d/d\theta)\{p(\theta)(d/d\theta)\} + q(\theta) + \lambda\rho(\theta)]\cdot P(\theta) = 0$$

may be brought to the form

$$[(d^2/dx^2) + r(x, m) + \lambda]\cdot Y(x) = 0$$

(m is a parameter which takes on integral values). If certain conditions are met, this differential equation may be replaced by two first-order equations

$$\begin{aligned}\mathfrak{R}_m\cdot Y^m &\equiv [k(x, m+1) - (d/dx)]\cdot Y^m \\ &= [\lambda - L(m+1)]^{\frac{1}{2}} Y^{m+1}, \\ \mathfrak{L}_m\cdot Y^m &\equiv [k(x, m) + (d/dx)]\cdot Y^m \\ &= [\lambda - L(m)]^{\frac{1}{2}} Y^{m-1},\end{aligned} \tag{5}$$

these being the recursion relations. Infeld and Hull investigate all the functions $k(x, m)$, $L(m)$ which permit this factorization of the differential equation, and classify these into six interdependent types, treating many examples of families of special functions as special cases of these types, and determining the functions together with their eigenvalues quite elegantly.

Suggestive for our present work is the proof which Infeld and Hull give for the fact that $L(m)$ is *at most* quadratic in m.[2]

$$L(m) = am^2 + bm + c. \tag{6}$$

In order to draw conclusions from Eq. (6), we prefer to introduce[3] differential operators $\mathfrak{R}$, $\mathfrak{L}$ which do not refer to the index m of the function being operated upon. Equations (5) are then replaced by

$$\begin{aligned}\mathfrak{R}\cdot Y^m &= [\lambda - L(m+1)]^{\frac{1}{2}} Y^{m+1}, \\ \mathfrak{L}\cdot Y^m &= [\lambda - L(m)]^{\frac{1}{2}} Y^{m-1},\end{aligned} \tag{7}$$

from which two forms of the differential equation are obtained, by applying either $\mathfrak{R}.\mathfrak{L}.$ or $\mathfrak{L}.\mathfrak{R}.$ to Y^m. Subtracting the two forms from each other gives:

$$\begin{aligned}[\mathfrak{R}, \mathfrak{L}]\cdot Y^m &\equiv [\mathfrak{R}\mathfrak{L} - \mathfrak{L}\mathfrak{R}]\cdot Y^m \\ &= \{L(m) - L(m+1)\} Y^m.\end{aligned} \tag{8}$$

If we now consider a family of functions for which $L(m) = \text{const}$ [i.e., $a = b = 0$ in Eq. (6)], then $[\mathfrak{R}, \mathfrak{L}]$. $Y^m = 0$ for all m. This suggests that $[\mathfrak{R}, \mathfrak{L}] = 0$, and indeed for the case of the Bessel functions [where $L(m) = 0$] we find that the differential operators commute.

Suppose however that $L(m)$ is linear in m (as is the case for the Hermite functions); then $\{L(m) - L(m+1)\} = -b$ is independent of m. We write $[\mathfrak{R}, \mathfrak{L}] \equiv \mathfrak{M}$, and ask for the effect of $[\mathfrak{M}, \mathfrak{R}]$, $[\mathfrak{L}, \mathfrak{M}]$ on Y^m:

$$\begin{aligned}[\mathfrak{M}\mathfrak{R} - \mathfrak{R}\mathfrak{M}]\cdot Y^m &= -b[\lambda - L(m+1)]^{\frac{1}{2}} Y^{m+1} \\ &\quad + b[\lambda - L(m+1)]^{\frac{1}{2}} Y^{m+1} = 0,\end{aligned}$$

and similarly for $\mathfrak{L}$. This suggests $[\mathfrak{M}, \mathfrak{R}] = [\mathfrak{M}, \mathfrak{L}] = 0$, and there is no difficulty in verifying this for the differential operators of the Hermite functions. The Lie group generated by $\mathfrak{R}$, $\mathfrak{L}$ is thus a three-parameter group; the composition law in this group is known ("Weyl's identity"), and is rederived by use of matrix representations in the Appendix.

In the most general case, where $L(m)$ is quadratic in m, similar reasoning leads us to find that $\mathfrak{R}$, $\mathfrak{L}$ generate a three-parameter Lie group which (after normalization) is locally isomorphic with the rotation group in three-dimensions, 0_3. (Some useful decompositions of the general element of this group are derived in the Appendix.) Thus, all functions treated by Infeld and Hull turn out to be related to *at most* three-parameter groups. However, larger groups make their appearance as soon as more than one independent variable is admitted (for example, see Sec. IIE below, where it is shown how the shift operators generate the group U_n).

C.

The close relationship of Infeld and Hull's work with Lie groups, which has been discussed above, has also been noted by Miller.[4] In his Memoir he recasts the classification into types, introduced by Infeld and Hull, into a classification of Lie groups, all of them being special cases of a master group with four parameters. Our paper makes no attempt at an over-all classification, but rather takes known families of special functions and their known recursion relations as *given*. From each set of recursion

[1] L. Infeld and T. E. Hull, Rev. Mod. Phys. **23**, 21(1951).

[2] This holds for the four types denoted as *A-D* by Infeld and Hull; the other two types (*E,F*) are transformable into *A-D*.

[3] To do this, we introduce a spurious variable, say ϕ, and and consider functions $F^m(x,\phi) \equiv G^m(\phi)Y^m(x)$ such that

$$k(x,m+1)\cdot[G^m(\phi)Y^m(x)] \equiv K(x)(\partial/\partial\phi)[G^m(\phi)Y^m(x)].$$

Variations on this procedure occur in each one of the sections below. See, in this connection, the similar steps taken by L. Weisner, Pacific J. Math. **5**, 1033(1955).

[4] W. Miller, Mem. Am. Math. Soc. No. 50 (1964).

relations the corresponding Lie algebra is generated and various expansions are derived, again with no attempt at exhausting the long lists of such expansions which are available in the literature. The examples in this paper are presented to support the point of view that all expansions and addition theorems are induced by the operations of elements of the appropriate Lie algebra. We have restricted ourselves to two- and three-parameter groups in this paper, indicating in several places which higher groups remain to be treated in a later publication.

II. SPECIAL FUNCTIONS RELATED TO TWO- AND THREE-PARAMETER GROUPS

A. Bessel Functions

The recursion relations for Bessel functions are

$$\begin{aligned}[(n/r) - (d/dr)]\cdot J_n(r) &= J_{n+1}(r)\\ [(n/r) + (d/dr)]\cdot J_n(r) &= J_{n-1}(r)\end{aligned}\quad (\rho_n = \lambda_{n+1} = 1), \tag{9}$$

from which the Bessel differential equation follows:

$$\left[\frac{d^2}{dr^2} + \frac{1}{r}\frac{d}{dr} - \frac{n^2}{r^2} + 1\right]\cdot J_n(r) = 0. \tag{10}$$

The differential equation will not play any role in the following considerations (consistent with the fact that $\mathfrak{R}\mathfrak{L}$, $\mathfrak{L}\mathfrak{R}$ do not belong to the Lie algebra), and is presented only as an aid in identifying the function.

In order to obtain $\mathfrak{R}$ and $\mathfrak{L}$ in a form which does not refer to the index of the Bessel function, we interpret n as the result of operating with $-i(\partial/\partial\phi)$ upon the function $e^{in\phi}$. Our recursion equations become

$$\begin{aligned}e^{i\phi}\left(-\frac{i}{r}\frac{\partial}{\partial\phi} - \frac{\partial}{\partial r}\right)\cdot[e^{in\phi}J_n(r)] &= [e^{i(n+1)\phi}J_{n+1}(r)],\\ e^{-i\phi}\left(-\frac{i}{r}\frac{\partial}{\partial\phi} + \frac{\partial}{\partial r}\right)\cdot[e^{in\phi}J_n(r)] &= [e^{i(n-1)\phi}J_{n-1}(r)].\end{aligned} \tag{11}$$

Our raising and lowering operators are thus (expressed in polar coordinates)

$$\begin{aligned}\mathfrak{R} &= e^{i\phi}\left(-\frac{i}{r}\frac{\partial}{\partial\phi} - \frac{\partial}{\partial r}\right),\\ \mathfrak{L} &= e^{-i\phi}\left(-\frac{i}{r}\frac{\partial}{\partial\phi} + \frac{\partial}{\partial r}\right).\end{aligned} \tag{12}$$

Expressed in Cartesian coordinates these become

$$\begin{aligned}\mathfrak{R} &= -\frac{\partial}{\partial x} - i\frac{\partial}{\partial y},\\ \mathfrak{L} &= +\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}.\end{aligned} \tag{13}$$

In this form one sees directly that $\mathfrak{R}$ and $\mathfrak{L}$ commute:

$$[\mathfrak{R}, \mathfrak{L}] = 0; \tag{14}$$

hence, the composition law within the Lie group is additive:

$$\exp\alpha\mathfrak{R}\cdot\exp\beta\mathfrak{L} = \exp(\alpha\mathfrak{R} + \beta\mathfrak{L}). \tag{15}$$

We now proceed to construct addition theorems for the functions

$$F_n \equiv [e^{in\phi}J_n(r)],$$

which we consider alternately as functions of polar or of Cartesian coordinates. For dealing with the polar coordinates it proves convenient to introduce the variables

$$u \equiv x + iy \equiv re^{i\phi}, \qquad v \equiv x - iy \equiv re^{-i\phi}, \tag{16}$$

in terms of which

$$\mathfrak{R} = -2\frac{\partial}{\partial v}, \qquad \mathfrak{L} = 2\frac{\partial}{\partial u}. \tag{17}$$

In the space of u and v, the finite operators $\exp\alpha\mathfrak{R}$, $\exp\beta\mathfrak{L}$ are translation operators. $\exp\alpha\mathfrak{R}$ sends v into $(v - 2\alpha)$, and leaves u invariant, while $\exp\beta\mathfrak{L}$ sends u into $(u + 2\beta)$, leaving v invariant. Operating on the functions of (u, v) we have, on the one hand,

$$\begin{aligned}\exp\alpha\mathfrak{R}\cdot F_n(u, v) &= \exp[-2\alpha(\partial/\partial v)]\cdot F_n(u, v)\\ &= F_n(u, v - 2\alpha)\end{aligned}$$

and, on the other hand,

$$\exp\alpha\mathfrak{R}\cdot F_n(u, v) = \sum_{m=0}^{\infty}\frac{\alpha^m}{m!}F_{n+m}(u, v);$$

or, combining these equations:

$$F_n(u, v - 2\alpha) = \sum_{m=0}^{\infty}\frac{\alpha^m}{m!}F_{n+m}(u, v). \tag{18}$$

We may now rewrite this addition theorem in the original variables (r, ϕ) so that it is brought into a better-known form. To do this we note the relations

$$e^{i\phi} = (u/v)^{\frac{1}{2}}, \quad (uv)^{\frac{1}{2}} = r, \quad u = re^{i\phi}, \quad v = re^{-i\phi}; \tag{19}$$

thus

$$\begin{aligned}F_n(u, v - 2\alpha) &\equiv [e^{in\phi'}J_n(r')]\\ &= \left(\left(\frac{u}{v - 2\alpha}\right)^{n/2}J_n\{[u(v - 2\alpha)]^{\frac{1}{2}}\}\right)\\ &= \left\{\left(\frac{u^2}{uv - 2\alpha u}\right)^{n/2}J_n[(uv - 2\alpha u)^{\frac{1}{2}}]\right\}.\end{aligned} \tag{20}$$

Putting $-2\alpha u \equiv h$ we have, from (18), (19), and (20),

$$(r^2+h)^{-\frac{1}{2}n}J_n[(r^2+h)^{\frac{1}{2}}] = \sum_{m=0}^{\infty}\left(\frac{-h}{2u}\right)^m\frac{u^{-n}}{m!}[e^{i(n+m)\phi}J_{n+m}(r)] \quad (21a)$$

$$= \sum_{m=0}^{\infty}\frac{1}{m!}\left(\frac{-h}{2}\right)^m r^{-n-m}J_{n+m}(r).$$

Similarly, using the operator $\mathfrak{L}$,we would obtain the companion equation

$$(r^2+h)^{\frac{1}{2}n}J_n[(r^2+h)^{\frac{1}{2}}] = \sum_{m=0}^{\infty}\frac{1}{m!}\left(\frac{h}{2}\right)^m r^{n-m}J_{n-m}(r). \quad (21b)$$

Both of these are precisely the forms listed in the Bateman project.[5]

Next, consider the combination

$$-\tfrac{1}{2}(\mathfrak{R}-\mathfrak{L}) \equiv \partial/\partial x.$$

This is a translation operator in the plane of (x, y):

$$\exp[\alpha(\partial/\partial x)]\cdot x = x+\alpha, \qquad \exp[\alpha(\partial/\partial x)]\cdot y = y.$$

By the previous reasoning, we find

$$\{\exp\tfrac{1}{2}\alpha(\mathfrak{L}-\mathfrak{R})\}\cdot F_n(x,y) = \exp[\alpha(\partial/\partial x)]\cdot F_n(x,y) = F_n(x+\alpha, y).$$

But, on the other hand,

$$\{\exp\tfrac{1}{2}\alpha(\mathfrak{L}-\mathfrak{R})\}\cdot F_n(x,y) = \exp(-\tfrac{1}{2}\alpha\mathfrak{R})\cdot\exp(\tfrac{1}{2}\alpha\mathfrak{L})\cdot F_n(x,y)$$

$$= \sum_{t=0}^{\infty}\sum_{s=0}^{\infty}\frac{1}{t!\,s!}\left(\frac{-\alpha}{2}\right)^t\left(\frac{\alpha}{2}\right)^s F_{n-s+t}(x,y)$$

$$= \sum_{t-s=m=-\infty}^{\infty} F_{n+m}(x,y)\left[\sum_{t=0}^{\infty}\frac{(-1)^t}{t!\,(t-m)!}\left(\frac{\alpha}{2}\right)^{2t-m}\right].$$

Equating these results, and returning to polar coordinates, we arrive at

$$[e^{in\phi'}J_n(r')] = \sum_{m=-\infty}^{\infty}[e^{i(n+m)\phi}J_{n+m}(r)]\cdot\left[\sum_{t=0}^{\infty}\frac{(-1)^t}{(t-m)!\,t!}\left(\frac{\alpha}{2}\right)^{2t-m}\right]. \quad (22)$$

This addition theorem contains the well-known series expansion of $J_n(r)$ as a special case. If we specialize (22) for the case $r = \phi = 0$ (which implies $r' = \alpha$, $\phi' = 0$) we obtain

$$J_n(\alpha) = \sum_{m=-\infty}^{\infty}J_{n+m}(0)\left[\sum_{t=0}^{\infty}\frac{(-1)^t}{(t-m)!\,t!}\left(\frac{\alpha}{2}\right)^{2t-m}\right]. \quad (23)$$

[5] *Higher Transcendental Functions*, (McGraw-Hill Book Company, Inc., New York, 1953), Vol. II. Equations (21a) and (21b) appear on p. 100.

But from the recursion relations it is easy to see that $J_k(0) = 0$ for every k, except $k = 0$. Hence, if n is an integer, we find [normalizing $J_0(0) = 1$]

$$J_n(\alpha) = \sum_{t=0}^{\infty}\frac{(-1)^t}{(t+n)!\,t!}\left(\frac{\alpha}{2}\right)^{2t+n}. \quad (24)$$

Introducing $J_n(\alpha)$ in (22) instead of its series expansion, we find Graf's addition theorem[6]

$$e^{in(\phi'-\phi)}J_n(r') = \sum_{m=-\infty}^{\infty}e^{im\phi}\cdot J_{n+m}(r)J_{-n}(\alpha). \quad (25)$$

To find further expansions, we make use of the most general operator available to us: $(\alpha\mathfrak{L}-\beta\mathfrak{R})$. Introducing

$$\tfrac{1}{2}z \equiv (\alpha\beta)^{\frac{1}{2}}, \qquad \gamma \equiv (\alpha/\beta)^{\frac{1}{2}},$$

we have

$$\exp(\alpha\mathfrak{L}-\beta\mathfrak{R}) = \exp\left[\frac{z}{2}\left(\gamma\mathfrak{L}-\frac{1}{\gamma}\mathfrak{R}\right)\right] = \sum_{s,t=0}^{\infty}\frac{(-)^t}{s!\,t!}\left(\frac{z}{2}\right)^{s+t}\gamma^{s-t}\mathfrak{L}^s\mathfrak{R}^t. \quad (26)$$

We can now make use of the fact that the operations of $\mathfrak{R}$ and $\mathfrak{L}$ on the functions $F_m(u,v) \equiv [e^{im\phi}J_m(r)]$ cancel each other out: $\mathfrak{R}\cdot\mathfrak{L}\cdot F_m = F_m = \mathfrak{L}\cdot\mathfrak{R}\cdot F_m$. Hence, (for operations on the functions F_m) the operator $\mathfrak{R}$ may be replaced by $\mathfrak{L}^{-1}$, out of which

$$\exp(\alpha\mathfrak{L}-\beta\mathfrak{R}) = \sum_{\substack{k=-\infty\\ k=s-t}}^{\infty}\left[\sum_{t=0}^{\infty}\frac{(-)^t}{(k+t)!\,t!}\left(\frac{z}{2}\right)^{k+2t}\right]\gamma^k\mathfrak{L}^k,$$

or, by Eq. (24),

$$\exp(\alpha\mathfrak{L}-\beta\mathfrak{R}) = \sum_{k=-\infty}^{\infty}\gamma^k J_k(z)\mathfrak{L}^k. \quad (27)$$

Using this operator equation we obtain the addition theorem

$$F_m(u',v') = \exp(\alpha\mathfrak{L}-\beta\mathfrak{R})\cdot F_m(u,v) = \sum_{k=-\infty}^{\infty}\gamma^k J_k(z)F_{m-k}(u,v), \quad (28)$$

where

$$u' = u+2\alpha, \qquad v' = v+2\beta.$$

Translating back from the coordinates (u, v) to the polar coordinates, Eq. (28) takes on the form

$$\left(\frac{re^{i\phi}+2\alpha}{re^{-i\phi}+2\beta}\right)^{\frac{1}{2}m}J_m\{[(re^{i\phi}+2\alpha)(re^{-i\phi}+2\beta)]^{\frac{1}{2}}\} = \sum_{k=-\infty}^{\infty}\left(\frac{\alpha}{\beta}\right)^{\frac{1}{2}k}J_k[2(\alpha\beta)^{\frac{1}{2}}]\,e^{i(m-k)\phi}J_{m-k}(r). \quad (29)$$

[6] Reference 5, p. 45.

We may now make various specializations of our general addition theorem. Two special choices are detailed here.

Choosing $\alpha = \beta \equiv -\frac{1}{2}z$ gives

$$[(re^{i\phi} + 2\alpha)(re^{-i\phi} + 2\beta)]^{\frac{1}{2}} \equiv Z = (r^2 + z^2 - 2rz\cos\phi)^{\frac{1}{2}}$$

and

$$2(\alpha\beta)^{\frac{1}{2}} = 2\alpha = -z.$$

Equation (29) becomes

$$\left(\frac{re^{i\phi} - z}{re^{-i\phi} - z}\right)^{\frac{1}{2}m} J_m(Z) = \sum_{k=-\infty}^{\infty} (-)^k J_k(z) e^{i(m-k)\phi} J_{m-k}(r). \tag{30}$$

In particular,[7] for $m = 0$,

$$J_0(Z) = \sum_{k=-\infty}^{\infty} e^{-ik\phi} \cdot J_k(z) \cdot J_k(r). \tag{31}$$

The choice $\alpha = \frac{1}{2}re^{i\phi}t^2$, $\beta = \frac{1}{2}re^{-i\phi}t^2$ results in

$$(te^{i\phi})^m J_m\left[r\left(t + \frac{1}{t}\right)\right] = \sum_{k=-\infty}^{\infty} (t^2 e^{i\phi})^k J_k(r) e^{i(m-k)\phi} J_{m-k}(r)$$

or[8]

$$t^m J_m\left[r\left(t + \frac{1}{t}\right)\right] = \sum_{k=-\infty}^{\infty} t^{2k} J_k(r) J_{m-k}(r). \tag{32}$$

It is of interest to note that the operator equation (27) may be applied to *any* functions on which $\mathfrak{R}$ and $\mathfrak{L}$ cancel each other. Thus

$$(\tfrac{1}{2}\mathfrak{L}) \cdot \{\exp(av - u/a)\} = a\{\exp(av - u/a)\}$$

while

$$(\tfrac{1}{2}\mathfrak{R}) \cdot \{\exp(av - u/a)\} = (1/a)\{\exp(av - u/a)\},$$

so that $(\frac{1}{2}\mathfrak{R})$ acts as the reciprocal of $(\frac{1}{2}\mathfrak{L})$ on the function $\exp(av - u/a)$. Hence we may use the operator equation (27) to obtain

$$[\exp \tfrac{1}{2}\alpha(\mathfrak{L} - \mathfrak{R})] \cdot \left\{\exp\left(av - \frac{u}{a}\right)\right\} = \sum_{k=-\infty}^{\infty} J_k(2\alpha)\left(\frac{\mathfrak{L}}{2}\right)^k \left\{\exp\left(av - \frac{u}{a}\right)\right\} = \left\{\sum_{k=-\infty}^{\infty} J_k(2\alpha) a^k\right\}\left\{\exp\left(av - \frac{u}{a}\right)\right\}.$$

On the other hand, the transformation in the (u, v) plane gives

$$[\exp \tfrac{1}{2}\alpha(\mathfrak{L} - \mathfrak{R})] \cdot \{\exp(av - u/a)\} = \exp\{a(v + \alpha) - (u + \alpha)/a\} = \{\exp \alpha(a - 1/a)\}\{\exp(av - u/a)\}.$$

Equating these two results, we find the Bessel generating function

$$\exp \alpha\left(a - \frac{1}{a}\right) = \sum_{k=-\infty}^{\infty} a^k \cdot J_k(2\alpha). \tag{33}$$

B. Hermite Functions

The recursion relations which define Hermite functions are

$$\left.\begin{aligned} &[x - (d/dx)]h_n(x) \\ &= [2(n+1)]^{\frac{1}{2}} h_{n+1}(x), \\ &[x + (d/dx)]h_n(x) \\ &= (2n)^{\frac{1}{2}} h_{n-1}(x), \end{aligned}\right\} \rho_{n,r} = \lambda_{n+1} = [2(n+1)]^{\frac{1}{2}}, \tag{34}$$

leading to the differential equation

$$[x^2 - (d^2/dx^2)] \cdot h_n(x) = (2n+1)h_n(x).$$

It follows from (34) that

$$[x + (d/dx)]h_0(x) = 0 \tag{35}$$

so that (with a choice of normalization)

$$h_0(x) = e^{-\frac{1}{2}x^2} \quad \text{and therefore} \quad h_0(0) = 1. \tag{36}$$

By adding the two equations of (34), and taking $x = 0$ we find

$$(n+1)^{\frac{1}{2}} h_{n+1}(0) = -n^{\frac{1}{2}} h_{n-1}(0) \tag{37}$$

from which, in particular,

$$h_1(0) = 0. \tag{38}$$

Equation (37) leads to an explicit form of $h_n(0)$:

$$\left.\begin{aligned} h_{2m}(0) &= (-)^m\{[(2m)!]^{\frac{1}{2}}/2^m \cdot m!\}, \\ h_{2m+1}(0) &= 0. \end{aligned}\right\} \tag{39}$$

As in the case of the Bessel functions, we introduce a further variable in order to bring the raising and lowering operators into the standard form of Lie elements: the operation $x.$ on $h_n(x)$ is interpreted as $x(\partial/\partial t)$ operating on $[e^t h_n(x)]$. Thus, the operators which generate the Lie algebra take on the form

$$\left.\begin{aligned} \mathfrak{R} &\equiv x\,\partial/\partial t - \partial/\partial x, \\ \mathfrak{L} &\equiv x\,\partial/\partial t + \partial/\partial x. \end{aligned}\right\} \tag{40}$$

The commutation laws are

$$[\mathfrak{R}, \mathfrak{L}] \equiv \mathfrak{M} = -2\,\partial/\partial t, \qquad [\mathfrak{R}, \mathfrak{M}] = [\mathfrak{L}, \mathfrak{M}] = 0. \tag{41}$$

[7] Reference 5, p. 101.
[8] Reference 5, p. 102.

Within the plane of (x, t) we find

$$\exp(p\mathfrak{R}):\begin{cases}x \to x - p,\\ t \to t + px - \frac{1}{2}p^2.\end{cases}$$
$$\exp(q\mathfrak{L}):\begin{cases}x \to x + q,\\ t \to t + qx + \frac{1}{2}q^2.\end{cases} \tag{42}$$

The composition law in the corresponding Lie group is (see the Appendix)

$$\exp(p\mathfrak{R} + q\mathfrak{L}) = \exp(p\mathfrak{R})\cdot\exp(q\mathfrak{L})\cdot\exp[(-\tfrac{1}{2}pq)\mathfrak{M}]$$
$$= \exp(q\mathfrak{L})\cdot\exp(p\mathfrak{R})\cdot\exp[(\tfrac{1}{2}pq)\mathfrak{M}].$$

We now derive the generating function for Hermite functions by applying the operator $\exp(z\mathfrak{R})$ to $h_0(x)$,

$$[\exp(z\mathfrak{R})]\cdot[e^t h_0(x)] = \exp(t + zx - \tfrac{1}{2}z^2)h_0(x - z)$$
$$= \exp(t + 2zx - z^2 - \tfrac{1}{2}x^2).$$

On the other hand,

$$[\exp(z\mathfrak{R})]\cdot[e^t h_0(x)] = e^t \sum_{n=0}^{\infty}\frac{z^n}{n!}[\mathfrak{R}^n h_0(x)]$$
$$= e^t \sum_{n=0}^{\infty} z^n\left(\frac{2^n}{n!}\right)^{\frac{1}{2}} h_n(x).$$

Equating the two results, we obtain the generating function for Hermite polynomials[8]

$$\exp(2xz - z^2) = \sum_{n=0}^{\infty}\frac{z^n}{n!}H_n(x), \tag{43}$$

where $H_n(x) \equiv [2^n n!]^{\frac{1}{2}}\exp(\frac{1}{2}x^2)h_n(x)$.

Next, we apply $\exp(\alpha\mathfrak{L})$ to $[e^t h_n(x)]$. On the one hand,

$$[\exp(\alpha\mathfrak{L})]\cdot[e^t h_n(x)] = \exp(t + \alpha x + \tfrac{1}{2}\alpha^2)h_n(x + \alpha),$$

and on the other hand,

$$[\exp(\alpha\mathfrak{L})]\cdot[e^t h_n(x)]$$
$$= e^t \sum_{m=0}^{n}\frac{\alpha^m}{m!}\left[2^m\frac{n!}{(n-m)!}\right]^{\frac{1}{2}} h_{n-m}(x).$$

Equating these last two results gives us our addition theorem, rewritten in terms of Hermite polynomials:

$$H_n(x + \alpha) = \sum_{m=0}^{n}\frac{(2\alpha)^m n!}{m!\,(n-m)!}H_{n-m}(x). \tag{44}$$

This addition theorem is not commonly known.

Letting $x = 0$, and making use of Eq. (39), we find

$$h_n(\alpha) = \left(\frac{n!}{2^n}\right)^{\frac{1}{2}} e^{-\frac{1}{2}\alpha^2}\left[\sum_{k=0}^{[\frac{1}{2}n]}\frac{(2\alpha)^{n-2k}(-)^k}{(n-2k)!\,k!}\right], \tag{45}$$

where the finite sum is the well-known explicit form for the Hermite polynomial.

Further addition formulas for Hermite functions involve products of these functions, and we see (Sec. E below) that product functions with n factors belong to the algebra of U_n.

C. Gegenbauer Polynomials

From the recursion relations for the functions $G_l^\alpha(\cos\theta)$,

$$[-(l + 2\alpha)\cos\theta - \sin\theta(d/d\theta)]G_l^\alpha$$
$$= -[(l + 1)(l + 2\alpha)]^{\frac{1}{2}}G_{l+1}^\alpha,$$
$$[-l\cos\theta + \sin\theta(d/d\theta)]G_l^\alpha \tag{46}$$
$$= -[l(l + 2\alpha - 1)]^{\frac{1}{2}}G_{l-1}^\alpha,$$

one obtains the differential equation of the Gegenbauer polynomials[9]

$$[(d^2/d\theta^2) + 2\alpha\cot\theta(d/d\theta) + l(l + 2\alpha)]$$
$$\cdot G_l^\alpha(\cos\theta) = 0. \tag{47}$$

For $\alpha = \frac{1}{2}$ this reduces to the equation of the Legendre polynomials $P_l(\cos\theta)$. For $\alpha = 0$, (47) is the differential equation of $\cos n\theta \equiv T_n(\cos\theta)$, or of $\sin n\theta \equiv \sin\theta U_{n-1}(\cos\theta)$, where T_n, U_n are Tchebichef polynomials of the first and second kind,[10] respectively.

For $l = 0$, the lowering equation in (46) becomes $\sin\theta(d/d\theta)G_0^\alpha(\cos\theta) = 0$, so that G_0^α is independent of θ, and may be normalized to unity: $G_0^\alpha(\cos\theta) = 1$. Adding the equations of (46), and setting $\cos\theta = 1$, gives us a recurrence relation for the $G_l^\alpha(1)$, from which we find

$$G_l^\alpha(1) = [\Gamma(2\alpha + l)/\Gamma(2\alpha)l!]^{\frac{1}{2}}. \tag{48}$$

In the literature the Gegenbauer polynomials are commonly normalized so that[9]

$$C_l^\alpha(1) = \frac{\Gamma(2\alpha + l)}{\Gamma(2\alpha)l!}, \quad \text{i.e.,}$$
$$C_l^\alpha(\cos\theta) \equiv \left[\frac{\Gamma(2\alpha + l)}{\Gamma(2\alpha)l!}\right]^{\frac{1}{2}}G_l^\alpha(\cos\theta). \tag{49}$$

Setting $\cos\theta = 0$ in the recursion equations gives us $G_l^{(\alpha)}(0) = 0$ for odd l, while for $l = 2k$ we find

$$C_{2k}^\alpha(0) = \left[\frac{\Gamma(2\alpha + 2k)}{\Gamma(2\alpha)(2k)!}\right]^{\frac{1}{2}}G_{2k}^\alpha(0) = (-)^k\frac{\Gamma(\alpha + k)}{\Gamma(\alpha)k!}. \tag{50}$$

[9] Reference 5, p. 174ff.
[10] Reference 5, p. 183ff.

In order to eliminate the indices (l, α) from the recursion relations, we consider functions

$$F_l^\alpha(\theta, t) \equiv e^{(l+\alpha)t}(\sin\theta)^\alpha G_l^\alpha(\cos\theta), \tag{51}$$

and operators

$$\begin{aligned} \mathfrak{R} &\equiv e^t\left[-\cos\theta\frac{\partial}{\partial t} - \sin\theta\frac{\partial}{\partial\theta}\right], \\ \mathfrak{L} &\equiv e^{-t}\left[-\cos\theta\frac{\partial}{\partial t} + \sin\theta\frac{\partial}{\partial\theta}\right]. \end{aligned} \tag{52}$$

It follows from (46) that

$$\begin{aligned} \mathfrak{R}\cdot F_l^\alpha(\theta, t) &= -[(l+1)(l+2\alpha)]^{\frac{1}{2}}F_{l+1}^\alpha(\theta, t), \\ \mathfrak{L}\cdot F_l^\alpha(\theta, t) &= -[l(l+2\alpha-1)]^{\frac{1}{2}}F_{l-1}^\alpha(\theta, t). \end{aligned} \tag{53}$$

To find the algebra generated by $\mathfrak{R}$ and $\mathfrak{L}$, we first compute their commutator:

$$[\mathfrak{R}, \mathfrak{L}] \equiv \mathfrak{M} = -2\,\partial/\partial t;$$

and then

$$[\mathfrak{M}, \mathfrak{R}] = -2\mathfrak{R}, \qquad [\mathfrak{M}, \mathfrak{L}] = 2\mathfrak{L}.$$

The algebra is isomorphic to that of the three-dimensional rotation group, and the composition laws derived in Appendix B may be applied.

To find the effect of $\mathfrak{R}$ and $\mathfrak{L}$ on the variables t and θ, it proves convenient to introduce the combinations

$$\begin{aligned} u_1 &\equiv e^{-t}\cos\theta, \qquad v_1 \equiv e^{-t}\sin\theta; \\ u_2 &\equiv e^{t}\cos\theta, \qquad v_2 \equiv e^{t}\sin\theta. \end{aligned} \tag{54}$$

We find

$$\mathfrak{R}\cdot u_1 = 1, \qquad \mathfrak{R}\cdot v_1 = 0,$$

so that

$$[\exp\gamma\mathfrak{R}]\cdot u_1 = u_1 + \gamma \equiv u_1' \quad [\exp\gamma\mathfrak{R}]\cdot v_1 = v_1 \equiv v_1'. \tag{55}$$

In the (u_1, v_1) space, $[\exp\gamma\mathfrak{R}]$ is a translation in the u_1 direction. Reverting back to the original variables (θ, t) we have

$$\begin{aligned} [\exp\gamma\mathfrak{R}]\cdot e^{-t} &= [\exp\gamma\mathfrak{R}]\cdot[u_1^2 + v_1^2]^{\frac{1}{2}} \\ &= [(u_1+\gamma)^2 + v_1^2]^{\frac{1}{2}} \\ &= [e^{-2t} + 2\gamma e^{-t}\cos\theta + \gamma^2]^{\frac{1}{2}} \equiv e^{-t'} \end{aligned}$$

or

$$e^{t-t'} = [1 + 2\gamma e^t\cos\theta + \gamma^2 e^{2t}]^{\frac{1}{2}} \equiv R. \tag{56}$$

Similarly it follows from

$$[\exp\gamma\mathfrak{R}]\cdot\cos\theta = [\exp\gamma\mathfrak{R}]\cdot[e^t u_1] = e^{t'}u_1'$$

and

$$[\exp\gamma\mathfrak{R}]\cdot\sin\theta = e^{t'}v_1'$$

that

$$\begin{aligned} \cos\theta' &= [\cos\theta + (\gamma e^t)]\cdot R^{-1}, \\ \text{and} \qquad \sin\theta' &= [\sin\theta]\cdot R^{-1}. \end{aligned} \tag{57}$$

We are now in position to apply $[\exp\gamma\mathfrak{R}]$ to any function of (θ, t). For example, we obtain the generating function of the $C_l^\alpha(\cos\theta)$ by applying the raising operator to $F_0^\alpha(\theta, t)$. On the one hand [taking into account $G_0^\alpha(x) = 1$],

$$\begin{aligned} [\exp\gamma\mathfrak{R}]\cdot F_0^\alpha(\theta, t) &= (e^{t'}\sin\theta')^\alpha G_0^\alpha(\cos\theta') \\ &= (e^t\sin\theta R^{-2})^\alpha. \end{aligned}$$

On the other hand,

$$\begin{aligned} &[\exp\gamma\mathfrak{R}]\cdot F_0^\alpha(\theta, t) \\ &= \sum_{k=0}^\infty \frac{\gamma^k}{k!}(-)^k\left[\frac{\Gamma(2\alpha+k)k!}{\Gamma(2\alpha)}\right]^{\frac{1}{2}}F_k^\alpha(\theta, t) \\ &= \sum_{k=0}^\infty (-\gamma e^t)^k\left[\frac{\Gamma(2\alpha+k)}{\Gamma(2\alpha)k!}\right]^{\frac{1}{2}}G_k^\alpha(\cos\theta)(e^t\sin\theta)^\alpha. \end{aligned}$$

Equating these last two, and setting $(-\gamma e^t) \equiv x$, we find

$$\begin{aligned} &\sum_{k=0}^\infty x^k\left[\frac{\Gamma(2\alpha+k)}{\Gamma(2\alpha)k!}\right]^{\frac{1}{2}}G_k^\alpha(\cos\theta) \\ &\qquad = [1 - 2x\cos\theta + x^2]^{-\alpha}, \end{aligned} \tag{58}$$

or, with the notation common in the literature,

$$\sum_{k=0}^\infty x^k C_k^\alpha(z) = [1 - 2xz + z^2]^{-\alpha}. \tag{59}$$

To find the effect of $\mathfrak{L}$ on (θ, t) it is convenient to use the combinations (u_2, v_2) of (54). We have

$$\mathfrak{L}\cdot u_2 = -1, \qquad \mathfrak{L}\cdot v_2 = 0,$$

so that

$$[\exp\gamma\mathfrak{L}]\cdot u_2 = u_2 - \gamma = u_2', \quad [\exp\gamma\mathfrak{L}]v_2 = v_2 = v_2'. \tag{60}$$

Applying $\exp\gamma\mathfrak{L}$ to $F_l^\alpha(\theta, t)$ gives us, on the one hand,

$$\begin{aligned} &[\exp\gamma\mathfrak{L}]\cdot[e^t\sin\theta)^\alpha(e^t)^l G_l^\alpha(\cos\theta)] \\ &\qquad = (e^t\sin\theta)^\alpha(e^{t'})^l G_l^\alpha(\cos\theta'). \end{aligned}$$

On the other hand,

$$\begin{aligned} &[\exp\gamma\mathfrak{L}]\cdot F_l^\alpha(\theta, t) \\ &\quad = \sum_{k=0}^l \frac{\gamma^k}{k!}(-)^k\left[\frac{l!}{(l-k)!}\frac{\Gamma(l+2\alpha)}{\Gamma(l+2\alpha-k)}\right]^{\frac{1}{2}} \\ &\qquad \cdot(e^t\sin\theta)^\alpha(e^t)^{l-k}G_{l-k}^\alpha(\cos\theta). \end{aligned}$$

Hence, using the common notation,

$$C_l^\alpha(\cos\theta') = (-\gamma e^{-t'})^l \sum_{s=l-k=0}^{l} \frac{(-\gamma e^{-t})^{-s}}{(l-s)!}\frac{\Gamma(l+2\alpha)}{\Gamma(s+2\alpha)} C_s^\alpha(\cos\theta).$$

From the geometry γ/e^t and $\gamma/e^{t'}$ may be expressed in terms of the sines of the angles so that[11]

$$C_l^\alpha(\cos\theta') = \left[\frac{\sin(\theta-\theta')}{\sin\theta}\right]^l \cdot \sum_{s=0}^{l}\left[\frac{\sin\theta'}{\sin(\theta-\theta')}\right]^s \frac{\Gamma(2\alpha+l)}{\Gamma(2\alpha+s)(l-s)!} C_s^\alpha(\cos\theta). \tag{61}$$

If we choose $\theta = \frac{1}{2}\pi$, and set $\cos\theta' \equiv x$, this becomes an explicit polynomial expression for $C_l^\alpha(x)$:

$$C_l^\alpha(x) = \sum_{s=0}^{l} x^{l-s}(1-x^2)^{\frac{1}{2}s}\frac{\Gamma(2\alpha+l)}{\Gamma(2\alpha+s)(l-s)!} C_s^\alpha(0),$$

or, using (50),

$$C_l^\alpha(x) = x^l \sum_{m=0}^{[\frac{1}{2}l]}\left[\frac{\Gamma(\alpha+m)}{m!\,\Gamma(\alpha)}\frac{\Gamma(2\alpha+l)}{(l-2m)!\,\Gamma(2\alpha+2m)}\right]\cdot\left(\frac{1-x^2}{x^2}\right)^m. \tag{62}$$

If we choose $\theta' = \frac{1}{2}\pi$ we see that the left-hand side in (61) has to be treated differently for odd and for even l. When l is *odd*

$$\sum_{s=0}^{l}\left[\frac{\Gamma(2\alpha+l)}{(l-s)!\,\Gamma(2\alpha+s)}\right]\frac{C_s^\alpha(x)}{(-x)^s} = 0; \tag{63a}$$

when l is *even* ($l \equiv 2m$) we find

$$\sum_{s=0}^{l}\left[\frac{\Gamma(2\alpha+l)}{(l-s)!\,\Gamma(2\alpha+s)}\right]\frac{C_s^\alpha(x)}{(-x)^s} = \left[\frac{\Gamma(\alpha+m)}{m!\,\Gamma(\alpha)}\right]\left(\frac{x^2-1}{x^2}\right)^m. \tag{63b}$$

Since the operators $\mathfrak{R}$ and $\mathfrak{L}$ act as translations in the planes of (u_1, v_1) and (u_2, v_2), respectively, it is natural to ask how these operators are related to the raising and lowering operators of the Bessel functions, which also induce translations in the plane. The group which is generated by all of these operators then leads to addition theorems in which products of Bessel and Gegenbauer functions are involved. The enlarged group will be dealt with in a later publication.

D. Associated Legendre Polynomials

It is well known that the associated Legendre polynomials are related to the three-dimensional rotation group through the operators which shift the index m of these polynomials. One forms the spherical harmonics

$$Y_l^m(\theta,\phi) \equiv e^{im\phi}P_l^m(\cos\theta) \tag{64}$$

and one finds the raising and lowering operators on Y_l^m to be

$$\begin{aligned}\mathfrak{R} &\equiv e^{i\phi}[-i\cot\theta(\partial/\partial\phi) - (\partial/\partial\theta)],\\ \mathfrak{L} &\equiv e^{-i\phi}[-i\cot\theta(\partial/\partial\phi) + (\partial/\partial\theta)],\end{aligned} \tag{65}$$

corresponding to the recursion relations

$$[m\cot\theta \mp (d/d\theta)]P_l^m(\cos\theta) = [(l\mp m)(l\pm m+1)]^{\frac{1}{2}}P_l^{m\pm1}(\cos\theta). \tag{66}$$

These lead to the differential equation of the Legendre polynomials

$$\frac{1}{\sin\theta}\frac{d}{d\theta}\left(\sin\theta\frac{dP}{d\theta}\right) - \frac{m^2}{\sin^2\theta}P + l(l+1)P = 0. \tag{67}$$

If we add the two equations of (66) and choose $\theta = 0$, we find

$$P_l^m(1) = 0, \quad \text{for} \quad m \neq 0. \tag{68}$$

The operators $\mathfrak{R}$ and $\mathfrak{L}$ generate the algebra of 0_3:

$$\begin{aligned}[\mathfrak{R},\mathfrak{L}] &\equiv \mathfrak{M} = -2i(\partial/\partial\phi),\\ [\mathfrak{M},\mathfrak{R}] &= 2\mathfrak{R}, \qquad [\mathfrak{M},\mathfrak{L}] = -2\mathfrak{L}.\end{aligned} \tag{69}$$

The combinations $\frac{1}{2}i(\mathfrak{R}+\mathfrak{L})$, $\frac{1}{2}(\mathfrak{R}-\mathfrak{L})$ are infinitesimal operators for rotations about the x and y axis, respectively, while $\frac{1}{2}i\mathfrak{M}$ generates the rotations about the z axis.

If a rotation by α about the y axis sends the point (θ,ϕ) into (θ',ϕ'), we may write

$$Y_l^m(\theta',\phi') = [\exp\tfrac{1}{2}\alpha(\mathfrak{R}-\mathfrak{L})]\cdot Y_l^m(\theta,\phi), \tag{70}$$

and we may apply the decomposition rule derived in the Appendix [Equation (A2.6)] to the operator in (70):

$$\begin{aligned}Y_l^m(\theta',\phi') &= \exp(\tan\tfrac{1}{2}\alpha\mathfrak{R})\cdot\exp(-\sin\tfrac{1}{2}\alpha\cos\tfrac{1}{2}\alpha\mathfrak{L})\\ &\quad\cdot\exp(-\log\cos\tfrac{1}{2}\alpha\mathfrak{M})\cdot Y_l^m(\theta,\phi)\\ &= (\cos\tfrac{1}{2}\alpha)^{-2m}\exp\left(\frac{\sin\frac{1}{2}\alpha}{\cos\frac{1}{2}\alpha}\mathfrak{R}\right)\\ &\quad\cdot\exp(-\sin\tfrac{1}{2}\alpha\cos\tfrac{1}{2}\alpha\mathfrak{L})\cdot Y_l^m(\theta,\phi)\\ &= (\cos\tfrac{1}{2}\alpha)^{-2m}\sum_{s,t}\frac{(-)^t(\sin\frac{1}{2}\alpha)^{t+s}(\cos\frac{1}{2}\alpha)^{t-s}}{s!\,t!}\\ &\quad\cdot\mathfrak{R}^s\mathfrak{L}^t\cdot Y_l^m(\theta,\phi)\end{aligned}$$

or

$$Y_l^m(\theta',\phi') \equiv \sum_{m'=m-t+s} Y_l^{m'}(\theta,\phi)\,\mathfrak{D}_{mm'}^{(l)}(0,\alpha,0), \tag{71}$$

[11] E. D. Rainville, *Special Functions* (Macmillan Company, New York, 1960), p. 280.

where $\mathfrak{D}^{(l)}(0, \alpha, 0)$ is the standard notation for the matrix representing a rotation by α about the y axis in the $(2l + 1)$-dimensional basis of $Y_l^m(\theta, \phi)$. From the above calculation it follows that

$$\mathfrak{D}^{(l)}_{mm'}(0, \alpha, 0) = [(l+m)!\,(l-m)!\,(l+m')!\,(l-m')!]^{\frac{1}{2}} \cdot \sum_t \frac{(-)^t (l-m+t)!(\sin \tfrac{1}{2}\alpha)^{m'-m+2t}(\cos \tfrac{1}{2}\alpha)^{-m-m'}}{(l-m)!\,(l-m')!\,(l+m-t)!\,t!\,(m'-m+t)!}. \quad (72)$$

To see that this expression is identical with the one first derived by Wigner,[12] where the sum has the form

$$\sum_x \frac{(-)^x(\sin \tfrac{1}{2}\alpha)^{m'-m+2x}(\cos \tfrac{1}{2}\alpha)^{-m'+m+2l-2x}}{(l-m'-x)!(l+m-x)!x!(m'-m+x)!},$$

set

$$(\cos \tfrac{1}{2}\alpha)^{-m'+m+2l-2x} = (\cos \tfrac{1}{2}\alpha)^{-m'-m}(1-\sin^2 \tfrac{1}{2}\alpha)^{l+m-x},$$

expand, and make use of the binomial identity

$$\sum_x \binom{t}{x}\binom{l-m}{l-m'-x} = \binom{l-m+t}{l-m'}.$$

If starting from the point $(\theta, \phi) = (0, 0)$ we make a rotation about the y axis, we come to the point $(\theta, 0)$, so that Eq. (71) reads

$$\begin{aligned} Y_l^m(\theta, 0) &= P_l^m(\cos\theta) \\ &= \sum_{m'} Y_l^{m'}(0, 0)\, \mathfrak{D}^{(l)}_{mm'}(0, \theta, 0) \\ &= P_l(1)\, \mathfrak{D}^{(l)}_{m0}(0, \theta, 0), \end{aligned}$$

since $Y_l^{m'}(0, 0) = 0$ when $m' \neq 0$. This leads to the addition theorem for Legendre polynomials $(m = 0)$:

$$P_l(1)P_l(\cos\theta') = \sum_{m'=-l}^{l} e^{im'\phi} P_l^{m'}(\cos\theta) P_l^{m'}(\cos\alpha). \quad (73)$$

When we adjoin the l-shifting operations to the m-shifting ones, we find that the algebra generated is that of $0_4 \simeq 0_3 \times 0_3$. This enlarged group will be treated in a later publication.

E. Products of Hermite Functions

We consider two different groups generated by operators which shift indices in the product functions $[h_a(x)h_b(y)]$, h_m being the Hermite functions of Sec. IIB. We denote

$$\mathfrak{R}_x \equiv \frac{1}{\sqrt{2}}\left(x - \frac{\partial}{\partial x}\right) \qquad \mathfrak{L}_x \equiv \frac{1}{\sqrt{2}}\left(x + \frac{\partial}{\partial x}\right), \quad (74)$$

and similarly for $\mathfrak{R}_y$, $\mathfrak{L}_y$.

[12] E. P. Wigner, *Group Theory and its Application to Quantum Mechanics*, Academic Press Inc., New York, 1959), p. 216.

The operator which raises both indices of the product function simultaneously is

$$\varrho \equiv i\mathfrak{R}_x\mathfrak{R}_y = \tfrac{1}{2}i[x - (\partial/\partial x)][y - (\partial/\partial y)], \quad (75)$$

so that, e.g.,

$$\varrho \cdot [h_a(x)h_b(y)] = i[(a+1)(b+1)]^{\frac{1}{2}}[h_{a+1}(x)h_{b+1}(y)].$$

In particular

$$\exp(-it\varrho)\cdot[h_0(x)h_0(y)] = \sum_{n=0}^{\infty} t^n [h_n(x)h_n(y)]. \quad (76)$$

We also introduce

$$\lambda \equiv i\mathfrak{L}_x\mathfrak{L}_y, \quad (77)$$

the operator which lowers both indices of the product function, so that $\lambda \cdot [h_0(x)h_0(y)] = 0$. We find

$$[\varrho, \lambda] \equiv \nu = 1 + \mathfrak{R}_x\mathfrak{L}_x + \mathfrak{R}_y\mathfrak{L}_y, \quad (78)$$

and

$$[\nu, \varrho] = 2\varrho, \qquad [\nu, \lambda] = -2\lambda; \quad (79)$$

that is to say, ϱ and λ generate the algebra of 0_3.

We note that

$$i(\varrho - \lambda) = [x(\partial/\partial y) + y(\partial/\partial x)] \quad (80)$$

so that this operator sends $x \to y$, $y \to x$, and

$$[\exp(i\alpha)(\varrho - \lambda)]\cdot\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x\cosh\alpha + y\sinh\alpha \\ x\sinh\alpha + y\cosh\alpha \end{pmatrix}. \quad (81)$$

Equation (81) helps us evaluate the left-hand side of (76), since it is not quite as simple to derive the effect of ϱ, or λ, separately on the variables $(x. y)$

From Eq. (A2.6) in the Appendix we have

$$\begin{aligned} &[\exp(-i\phi)(\varrho - \lambda)]\cdot[h_0(x)h_0(y)] \\ &\quad = [\exp(-i\tanh\phi\varrho)][\exp(i\sinh\phi\cosh\phi\lambda) \\ &\quad\quad \cdot[\exp(-\log\cosh\phi\nu)]\cdot[h_0(x)h_0(y)] \\ &\quad = (\cosh\phi)^{-1}[\exp(-i\tanh\phi\varrho)]\cdot[h_0(x)h_0(y)] \\ &\quad = (\cosh\phi)^{-1}\sum_n (\tanh\phi)^n [h_n(x)h_n(y)]. \end{aligned}$$

Since $[h_0(x)h_0(y)] = \exp[-\frac{1}{2}(x^2 + y^2)]$, we have, using (81),

$$\begin{aligned} &[\exp(-i\phi)(\varrho - \lambda)] \\ &\quad = \exp[-\tfrac{1}{2}(x\cosh\phi - y\sinh\phi)^2 \\ &\quad\quad - \tfrac{1}{2}(-x\sinh\phi + y\cosh\phi)^2] \\ &\quad = \exp[-\tfrac{1}{2}(x^2 + y^2)(\cosh 2\phi) + xy\sinh 2\phi]. \end{aligned}$$

Equating these two results (and setting $\tanh\phi \equiv t$) we find

$$\sum_n t^n[h_n(x)h_n(y)] = (1-t^2)^{-\frac{1}{2}} \cdot \exp\left\{-\tfrac{1}{2}(x^2+y^2)\frac{1+t^2}{1-t^2} + 2xy\frac{t}{1-t^2}\right\}. \quad (82)$$

This "generating function" for the product functions is Mehler's formula[13]; it is listed in Truesdell's monograph on the special functions[14] in the section dealing with questions unanswered by his methods.

A second group is generated by the operators

$$\mathfrak{A} \equiv \tfrac{1}{2}\mathfrak{R}_x\mathfrak{L}_y, \qquad \mathfrak{B} \equiv \tfrac{1}{2}\mathfrak{L}_x\mathfrak{R}_y. \quad (83)$$

Their commutator is

$$[\mathfrak{A}, \mathfrak{B}] \equiv \mathfrak{C} = \tfrac{1}{2}(\mathfrak{L}_x\mathfrak{R}_x - \mathfrak{L}_y\mathfrak{R}_y), \quad (84)$$

and we find

$$[\mathfrak{C}, \mathfrak{A}] = 2\mathfrak{A}, \qquad [\mathfrak{C}, \mathfrak{B}] = -2\mathfrak{B}. \quad (85)$$

Letting these operate on the product functions, we have

$$\begin{aligned}\mathfrak{A}\cdot[h_a(x)h_b(y)] &= [(a+1)b]^{\frac{1}{2}}[h_{a+1}(x)h_{b-1}(y)],\\ \mathfrak{B}\cdot[h_a(x)h_b(y)] &= [a(b+1)]^{\frac{1}{2}}[h_{a-1}(x)h_{b+1}(y)], \quad (86)\\ \mathfrak{C}\cdot[h_a(x)h_b(y)] &= \tfrac{1}{2}(a-b)[h_a(x)h_b(y)].\end{aligned}$$

The sum $(a+b)$ remains invariant under these operations; for any value of

$$l \equiv \tfrac{1}{2}(a+b) \quad (87)$$

there are $2l+1 = a+b+1$ functions which are transformed into each other under $\mathfrak{A}$, $\mathfrak{B}$, $\mathfrak{C}$. Using the notation

$$m = \tfrac{1}{2}(a-b), \qquad \eta_l^m(x,y) \equiv [h_a(x)h_b(y)], \quad (88)$$

the relations in (86) take on the form

$$\begin{aligned}\mathfrak{A}\cdot\eta_l^m &= [(l+m+1)(l-m)]^{\frac{1}{2}}\eta_l^{m+1},\\ \mathfrak{B}\cdot\eta_l^m &= [(l+m)(l-m+1)]^{\frac{1}{2}}\eta_l^{m-1}, \quad (89)\\ \mathfrak{C}\cdot\eta_l^m &= m\eta_l^m,\end{aligned}$$

showing that $\mathfrak{A}$, $\mathfrak{B}$, $\mathfrak{C}$ are represented in the basis $\eta_l^m(x,y)$ by the same matrices as $\mathfrak{R}$, $\mathfrak{L}$, $\mathfrak{M}$ in the basis $P_l^m(\cos\theta)$ (Sec. IID). (Note that while l is necessarily integral in the P_l^m, it may take on half-integer values as well, in the η_l^m basis; thus the even-dimensioned representations $\mathfrak{D}^{(l)}$ appear, in addition to the odd-dimensioned ones.)

In the complete analogy with Sec. IID, we apply the operation $[\exp\frac{1}{2}\alpha(\mathfrak{A}-\mathfrak{B})]$ to $\eta_l^m(x,y)$. Here

$$\tfrac{1}{2}(\mathfrak{A}-\mathfrak{B}) = [y(\partial/\partial x) - x(\partial/\partial y)] \quad (90)$$

so that

$$[\exp\tfrac{1}{2}\alpha(\mathfrak{A}-\mathfrak{B})]\cdot\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}x\cos\alpha + y\sin\alpha\\ -x\sin\alpha + y\cos\alpha\end{pmatrix}. \quad (91)$$

Thus

$$\begin{aligned}&[\exp\tfrac{1}{2}\alpha(\mathfrak{A}-\mathfrak{B})]\cdot\eta_l^m(x,y)\\ &= h_{l+m}(x\cos\alpha + y\sin\alpha)\cdot h_{l-m}(-x\sin\alpha + y\cos\alpha)\\ &= \sum_{m'}\mathfrak{D}^{(l)}_{mm'}(0,\alpha,0)h_{l+m'}(x)h_{l-m'}(y),\end{aligned}$$

or, in terms of Hermite polynomials,

$$\begin{aligned}&H_{l+m}(x\cos\alpha + y\sin\alpha)H_{l-m}(-x\sin\alpha + y\cos\alpha)\\ &\quad= \left[\frac{(l+m)!\,(l-m)!}{(l+m')!\,(l-m')!}\right]^{\frac{1}{2}}\\ &\quad\cdot\sum_{m'}\mathfrak{D}^{(l)}_{mm'}(0,\alpha,0)H_{l+m'}(x)H_{l-m'}(y). \quad (92)\end{aligned}$$

One may generalize the above considerations to products of n Hermite functions $h_{a_1}(x_1)h_{a_2}(x_2)\cdots h_{a_n}(x_n)$, with operators $(\mathfrak{R}_{x_i}\mathfrak{L}_{x_j})$ which leave invariant the sum $\sum_{i=1}^n a_i$. The commutation rules

$$[\mathfrak{R}_{x_i}\mathfrak{L}_{x_j}, \mathfrak{R}_{x_m}\mathfrak{L}_{x_n}] = \delta_{jm}\mathfrak{R}_{x_i}\mathfrak{L}_{x_n} - \delta_{in}\mathfrak{R}_{x_m}\mathfrak{L}_{x_j} \quad (93)$$

are identical with those of the operators $x_i(\partial/\partial x_j)$, which generate the algebra of U_n. Hence representation of U_n (or SU_n) may be built using the product functions as bases. The product functions and their relationship with spherical harmonics will be dealt with in a later publication.

ACKNOWLEDGMENT

The author wishes to thank Dr. C. Noack for his suggestion that the work of Infeld and Hull might be interpretable in terms of Lie algebras.

APPENDIX

COMPOSITION LAWS IN THE THREE-PARAMETER LIE GROUPS

1. The Group of the Hermite Functions

The commutation laws here are

$$[\mathbf{A}, \mathbf{B}] = \mathbf{C}, \qquad [\mathbf{A}, \mathbf{C}] = [\mathbf{B}, \mathbf{C}] = 0. \quad (A1.1)$$

A suitable three-dimensional representation for the algebra is

$$\mathbf{A} = \begin{bmatrix}0&1&0\\0&0&0\\0&0&0\end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix}0&0&0\\0&0&1\\0&0&0\end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix}0&0&1\\0&0&0\\0&0&0\end{bmatrix}. \quad (A1.2)$$

[13] Reference 5, p. 194.

[14] C. Truesdell, "An Essay Toward a Unified Theory of Special Functions," Ann. Math. Studies, No. 18, (Princeton University Press, Princeton, New Jersey, 1948).

The corresponding elements in the group are

$$\exp \alpha\mathbf{A} = \begin{bmatrix} 1 & \alpha & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad \exp \beta\mathbf{B} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & \beta \\ 0 & 0 & 1 \end{bmatrix},$$

$$\exp \gamma\mathbf{C} = \begin{bmatrix} 1 & 0 & \gamma \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{A1.3}$$

A general element of the algebra is $\mathbf{X} \equiv p\mathbf{A} + q\mathbf{B} + r\mathbf{C}$ and it must be possible to express $\exp \mathbf{X}$ as a product of three factors from the group:

$$\exp (p\mathbf{A} + q\mathbf{B} + r\mathbf{C}) = \exp \alpha\mathbf{A} \cdot \exp \beta\mathbf{B} \cdot \exp \gamma\mathbf{C}, \tag{A1.4}$$

where α, β, γ are to be determined in terms of p, q, r.

We find

$$\begin{bmatrix} 1 & p & r + \frac{1}{2}pq \\ 0 & 1 & q \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & \alpha & \gamma + \alpha\beta \\ 0 & 1 & \beta \\ 0 & 0 & 1 \end{bmatrix},$$

from which: $\alpha = p$, $\beta = q$, $\gamma = r - \frac{1}{2}pq$.

Recalling that $\mathbf{C}$ commutes with $\mathbf{A}$, $\mathbf{B}$, we may split off the factor $\exp r\mathbf{C}$ on the left side of Eq. (A1.4) and obtain

$$\exp (p\mathbf{A} + q\mathbf{B}) = \exp p\mathbf{A} \cdot \exp q\mathbf{B} \cdot \exp (-\tfrac{1}{2}pq)\mathbf{C}. \tag{A1.5}$$

Similarly we could find

$$\exp (p\mathbf{A} + q\mathbf{B}) = \exp q\mathbf{B} \cdot \exp p\mathbf{A} \cdot \exp (+\tfrac{1}{2}pq)\mathbf{C}. \tag{A1.6}$$

This gives us a rule for the interchange of order of the operators $\mathbf{A}$, $\mathbf{B}$:

$$\exp (p\mathbf{A}) \cdot \exp (q\mathbf{B}) = \exp (q\mathbf{B}) \cdot \exp (p\mathbf{A}) \cdot \exp (pq\mathbf{C}). \tag{A1.7}$$

2. The Three-Dimensional Rotation Group

The matrices

$$\varrho = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad \lambda = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad \nu = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \tag{A1.2}$$

satisfy the commutation rules

$$[\varrho, \lambda] = \nu, \quad [\nu, \varrho] = 2\varrho, \quad [\nu, \varrho] = -2\varrho. \tag{A2.2}$$

For the general member of the algebra

$$\mathbf{X} \equiv p\varrho + q\lambda + r\nu = \begin{bmatrix} r & p \\ q & -r \end{bmatrix}$$

we have the corresponding group member

$$\exp \mathbf{X} = \cosh z \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \frac{\sinh z}{z} \begin{bmatrix} r & p \\ q & -r \end{bmatrix}, \tag{A2.3}$$

with $z = (r^2 + pq)^{\frac{1}{2}}$.

The group member $\exp \mathbf{X}$ may be expressed as a product:

$$\exp (p\varrho + q\lambda + r\nu) = \exp \alpha\varrho \cdot \exp \beta\lambda \cdot \exp \gamma\nu, \tag{A2.4}$$

where α, β, γ, are to be determined.

One finds readily

$$\alpha = \frac{p \tanh z}{z - r \tanh z} \qquad \beta = \frac{q}{z} \sinh z \cosh z - \frac{qr}{z^2} \sinh^2 z,$$

$$e^{-\gamma} = \cosh z - (r/z) \sinh z. \tag{A2.5}$$

In particular, for the important case where

$$p = -q \equiv \phi \quad \text{and} \quad r = 0,$$

we find ($z = i\phi$)

$$\exp [\phi(\varrho - \lambda)] = \exp (\tan \phi\varrho) \cdot \exp (-\sin \phi \cos \phi\lambda) \cdot \exp (-\log \cos \phi\nu). \tag{A2.6}$$

It is, of course, possible to have the factors on the right-hand side of (A2.4) appear in a different order. Thus, for example, if it is desired to have the order of operators, $\exp a\lambda \cdot \exp b\varrho \cdot \exp c\nu$; one is led to

$$a = \frac{q \tanh z}{z - r \tanh z}, \quad b = \frac{p}{z} \sinh z \cosh z - \frac{pr}{z^2} \sinh^2 z,$$

$$e^{c} = \cosh z - (r/z) \sinh z. \tag{A2.7}$$

For the special case of $p = -q \equiv \phi$, $r = 0$, this means

$$a = -\tan \phi, \quad b = \sin \phi \cos \phi, \quad c = \log \cos \phi \tag{A2.8}$$

Using the same method, one easily obtains a rule for the interchange of order of the operators ϱ, λ:

$$\exp (p\varrho) \cdot \exp (q\lambda) = \exp \left[\frac{q}{1 + pq} \lambda\right] \cdot \exp [p(1 + pq)\varrho] \cdot \exp [\log (1 + pq)\nu]. \tag{A2.9}$$

Multipole Moments in Einstein's Gravitational Theory

D. J. LAMB
Department of Mathematics, King's College, London
(Received 5 August 1965)

Janis and Newman have proposed definitions of multipole moments of axially symmetric gravitational fields in terms of the initial data on a characteristic surface. This paper extends their definitions to source distributions without symmetry by considering the electromagnetic and linear gravitational fields. The global definition of four-momentum agrees, for sandwich waves, with the objects constructed by Møller and Cornish in their attempts to provide a local definition of four-momentum.

INTRODUCTION

THE multipole structure of the linearized gravitational field and its relation to a source distribution has been known for many years.[1] It bears a striking analogy with the multipole structure of the electromagnetic field in that the poles are of two types. The "electric" type poles are due to deviation of the mass distribution from spherical symmetry whilst the "magnetic" type poles are due to differential rotations of the source.

Janis and Newman[2] reobtained these results recently by using the spin-coefficient formalism. The advantage of this technique is that the multipole structure is defined directly in terms of components of the Weyl tensor (the vacuum Riemann tensor) whose physical significance was pointed out by Pirani.[3] Similar results were obtained by Bondi *et al.*[4] using a luminosity distance.

JN found that the angular dependence of multipoles in the (linear) gravitational field was specified by the associated Legendre polynomial of second order (P_n^2). These results were obtained by imposing suitable conditions of regularity on the field components in the neighborhood of the polar axis. In the Einstein field, this is equivalent to ensuring the existence of a Minkowski tangent space.

When the assumption of axial symmetry is dropped, the angular dependence, albeit more complicated, is found to be a natural generalization of the associated Legendre polynomials. The results may be obtained by linear operations on the spherical harmonics. They transform as second-rank tensors under $U(1)$.[5]

With this angular dependence we define the multipole moments of the linearized field. We then use these definitions with slight modification in the full Einstein theory. It is shown that the definitions of mass and linear momentum possess certain desirable properties obtained by others.[6,7]

I. MAXWELL THEORY

Throughout this paper we use the spin-coefficient formalism of Newman and Penrose.[8] In Minkowski space we introduce a null tetrad $\{l^\mu, n^\mu, m^\mu, \bar{m}^\mu\}$ satisfying the relations

$$l^\mu n_\mu = 1 = -m^\mu \bar{m}_\mu,$$

all other inner products being zero.

The Minkowski line element is chosen in the form

$$ds^2 = du^2 + 2\,du\,dr - r^2(d\theta^2 + \sin^2\theta\,d\varphi^2),$$

where u is a retarded time coordinate, r an affine parameter along the null geodesic lying in the surfaces u = constant, θ and φ the usual spherical angles. The surfaces u = constant are the light cones emanating from the origin $r = 0$.

The tetrad system thus has the form

$$l^\mu = \delta_1^\mu, \tag{1.2a}$$

$$n^\mu = \delta_0^\mu - \tfrac{1}{2}\delta_1^\mu, \tag{1.2b}$$

$$m^\mu = \frac{1}{\sqrt{2}r}\left(\delta_2^\mu + \frac{i}{\sin\theta}\,\delta_3^\mu\right), \tag{1.2c}$$

where $(u, r, \theta, \varphi) = (x^0, x^1, x^2, x^3)$.

Introduce the following definitions:

$$D = l^\mu\,\partial/\partial x^\mu$$

$$\Delta = n^\mu\,\partial/\partial x^\mu$$

$$\delta = m^\mu\,\partial/\partial x^\mu$$

[1] R. Sachs and P. G. Bergmann, Phys. Rev. **112**, 674 (1958).
[2] A. I. Janis and E. T. Newman, J. Math. Phys. **6**, 902 (1965). Hereinafter referred to as JN.
[3] F. A. E. Pirani, Acta Phys. Polon. **15**, 389 (1956).
[4] H. Bondi, M. G. J. van der Burg, and A. W. K. Metzner, Proc. Roy. Soc. (London) **A269**, 21 (1962).
[5] E. T. Newman (private communication). $U(1)$ is the unitary group in one dimension (the group of phase transformations).
[6] C. Møller, Kgl. Danske Videnskab. Selskab Mat. Fys. Medd. **34**, No. 3 (1964).
[7] F. H. J. Cornish, Proc. Roy. Soc. (London) **A286**, 270 (1965).
[8] E. Newman and R. Penrose, J. Math. Phys. **3**, 566 (1962).

$$\Phi_0 = F_{\mu\nu}l^\mu m^\nu$$

$$\Phi_1 = \tfrac{1}{2}F_{\mu\nu}(l^\mu n^\nu + \bar{m}^\mu m^\nu)$$

$$\Phi_2 = F_{\mu\nu}\bar{m}^\mu m^\nu$$

$$\rho = -1/r, \qquad \alpha = -(1/2\sqrt{2}r)\cot\theta,$$

$$\beta = (1/2\sqrt{2}r)\cot\theta, \qquad \mu = -1/2r.$$

Then Maxwell's equations can be written in the form

$$D\Phi_1 - \bar{\delta}\Phi_0 = 2\rho\Phi_1 - 2\alpha\Phi_0 \tag{1.3a}$$

$$D\Phi_2 - \bar{\delta}\Phi_1 = \rho\Phi_2 \tag{1.3b}$$

$$\delta\Phi_1 - \Delta\Phi_0 = \mu\Phi_0 \tag{1.3c}$$

$$\delta\Phi_2 - \Delta\Phi_1 = 2\mu\Phi_1 - 2\beta\Phi_2. \tag{1.3d}$$

We shall find it convenient to let

$$\mathfrak{D} = \frac{\partial}{\partial\theta} + \frac{i}{\sin\theta}\frac{\partial}{\partial\varphi}.$$

The characteristic initial-value problem may be solved as follows. Let us give Φ_0 as an arbitrary function of r, θ, and φ on an initial null hypersurface u = constant ($\mathfrak{N}_0$ say). The first two equations are readily integrated and yield two "constants" of integration which we shall call $\Phi_1^0(u, \theta, \varphi)$ and $\Phi_2^0(u, \theta, \varphi)$. After solving the final field equations we find that the time development is completely specified once $\Phi_2^0(u, \theta, \varphi)$ is given as an arbitrary function of time. Let us call this world tube $\mathfrak{s}$. Hence the characteristic initial problem is solved by specifying Φ_0 on $\mathfrak{N}_0$, Φ_2^0 on $\mathfrak{s}$ and Φ_1^0 on $\mathfrak{s} \cap \mathfrak{N}_0$.

Assuming that $\Phi_0 = O(r^{-3})$, $\mathfrak{D}\Phi_0$, $\mathfrak{D}^2\Phi_0 = O(r^{-3})$, then $\Phi_1 = O(r^{-2})$ and $\Phi_2 = O(r^{-1})$; the peeling theorem in electromagnetic theory. Following Janis and Newman[2] we now assume that

$$\Phi_0(u, r, \theta, \varphi) = \sum_{n=1}^{\infty} \frac{\Phi_0^{n-1}(u, \theta, \varphi)}{r^{n+2}}, \tag{1.4}$$

corresponding to the most general multipole structure.[9] We solve the field equations to find that

$$\Phi_1 = \frac{\Phi_1^0}{r^2} - \frac{1}{\sqrt{2}}(\bar{\mathfrak{D}} + \cot\theta)\sum_{n=1}^{\infty}\Phi_0^{n-1}/nr^{n+2} \tag{1.5a}$$

$$\Phi_2 = \frac{\Phi_2^0}{r} - \frac{1}{\sqrt{2}}\frac{\bar{\mathfrak{D}}\Phi_1^0}{r^2} + \tfrac{1}{2}\bar{\mathfrak{D}}(\bar{\mathfrak{D}} + \cot\theta)\sum_{n=1}^{\infty}\Phi_0^{n-1}/n(n+1)r^{n+2} \tag{1.5b}$$

$$\dot{\Phi}_1^0 = (1/\sqrt{2})(\mathfrak{D} + \cot\theta)\Phi_2^0 \tag{1.5c}$$

$$\dot{\Phi}_0^0 = (1/\sqrt{2})\mathfrak{D}\Phi_1^0 \tag{1.5d}$$

$$\dot{\Phi}_0^n = -\tfrac{1}{2}(n+1)\Phi_0^{n-1} - (1/2n)\mathfrak{D}(\bar{\mathfrak{D}} + \cot\theta)\Phi_0^{n-1}, \quad n \geq 1. \tag{1.5e}$$

We assume that Φ_0, Φ_1, Φ_2 do not have angular singularities (which correspond to nonisolated sources extending to spatial infinity). Let us now determine what news (Φ_2^0) yields a finite multipole structure, viz., Φ_0^l, $l > N$, are zero. If all Φ_0^l, $l > N$, are initially zero and $\dot{\Phi}_0^{N+1} = 0$ then they remain zero. Hence the angular dependence of Φ_0^N is determined by

$$\mathfrak{D}(\bar{\mathfrak{D}} + \cot\theta)\Phi_0^N + (N+1)(N+2)\Phi_0^N = 0. \tag{1.6}$$

The solution may be found by noting that Φ_0^N is periodic in φ (otherwise the solution would be multiple-valued). Thus

$$\Phi_0^N(u, \theta, \varphi) = A_{N+1,m}(u)e^{im\varphi}\chi(\cos\theta),$$

where m is a real integer and χ satisfies

$$\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\chi}{\partial\theta}\right) + \left\{(N+1)(N+2) - \frac{m^2 + 2m\cos\theta + 1}{\sin^2\theta}\right\}\chi = 0. \tag{1.7}$$

The general solution[10] is given by

$$\chi(\mu) = (1+\mu)^{\frac{1}{2}(m-1)}(1-\mu)^{\frac{1}{2}(m+1)} \times F(n+1+p, p-n; m+2; \tfrac{1}{2}(1-\mu)), \tag{1.8}$$

where $\mu = \cos\theta$, $2p = M + k$, $M^2 = (m-1)^2$, $k^2 = (m+1)^2$, and $n = N + 1$. F denotes the usual hypergeometric series defined by

$$F(a, b; c; x) = 1 + \frac{a\cdot b}{1\cdot c}x + \frac{a(a+1)b(b+1)}{1\cdot 2\cdot c(c+1)}x^2 + \cdots.$$

If $n - p$ is a positive integer the series terminates and if n is also a positive integer, we obtain a solution in the form of a polynomial.

Let ${}_1Y_n^m(\theta, \varphi) = e^{im\varphi}\chi(\cos\theta)$. Then the solution for Φ_0^N may be written

$$\Phi_0^N(u, \theta, \varphi) = \sum_m A_{N+1,m}(u)\,{}_1Y_{N+1}^m(\theta, \varphi), \tag{1.9}$$

where m is restricted by the above considerations to lie in the range $|m| \leq N + 1$. Let us take a particular value of m and find the remaining Φ_0^l's. Since the field equations are linear, it follows that the new angular dependence introduced by the complemen-

[9] Contrary to previous assertions this does *not* exclude incoming radiation fields. The subsequent angular considerations do, however, restrict the results to outgoing radiation fields in the linear theory.

[10] H. Bateman, *Partial Differential Equations of Mathematical Physics*, (Cambridge University Press, Cambridge, England, 1932).

tary function can be obtained by linear superposition of particular integrals to lower-order multipoles. Hence

$$\Phi_0^0(u, \theta, \varphi) = \alpha_{Nm}(d^N/du^N)A_{N+1,m}(u)\,{}_1Y^m_{N+1}(\theta, \varphi) \tag{1.10a}$$

$$\Phi_1^0(u, \theta, \varphi) = \beta_{Nm}(d^{N+1}/du^{N+1})A_{N+1,m}(u)\,Y^m_{N+1}(\theta, \varphi) \tag{1.10b}$$

$$\Phi_2^0(u, \theta, \varphi) = \gamma_{Nm}(d^{N+2}/du^{N+2})A_{N+1,m}(u)\,{}_1\bar{Y}^{-m}_{N+1}(\theta, \varphi), \tag{1.10c}$$

where α_{Nm}, β_{Nm}, γ_{Nm} are numerical constants and $Y^m_{N+1}(\theta, \varphi)$ are the spherical harmonics. The most general form of Φ_2^0 may be obtained by summing over m and N.

Hence

$$\Phi_2^0(u, \theta, \varphi) = \sum_{k=0}^{\infty}\sum_{m=-k}^{+k}\gamma_{km}\frac{d^{k+2}}{du^{k+2}}A_{k+1,m}(u)\,{}_1\bar{Y}^{-m}_{k+1}(\theta, \varphi). \tag{1.11}$$

We shall call ${}_1Y^m_n(\theta, \varphi)$ the vector spherical harmonics.[11]

A more useful expression for (1.11) may be obtained with the aid of the $\eth$ operator. Assuming that Φ_1^0 is expanded as a series of spherical harmonics, it follows from (1.5d) that

$${}_1Y^m_n = \eth Y^m_n, \tag{1.12}$$

where we have chosen the constant of proportionality to be unity. Note that when $m = 0$, we obtain the usual relation between $P^0_n(\cos\theta)$ and $P^1_n(\cos\theta)$, viz.,

$$P^1_n = dP^0_n/d\theta.$$

The most general news function may thus be specified as

$$\Phi_2^0(u, \theta, \varphi) = \sum_{m,n}\gamma_{mn}\eth Y^m_n, \tag{1.13}$$

where γ_{mn} is a function of time.

When $\Phi_0 = 0$ we have a multipole solution which is

$$\Phi_1 = A_0/r^2, \qquad \Phi_2 = 0, \qquad \dot{A}_0 = 0.$$

The last equation corresponds to the law of conservation of charge (i.e., there can be no monopole news). In order to exclude magnetic monopoles we assume A_0 is real.

We now affect the following definitions. Let $\mathcal{Q}$ be the total charge of the system under consideration and $\mathcal{E}^m_N$ the mth component of the 2^N − pole moment. Then

$$\mathcal{Q} \equiv \int_{S^2}\Phi_1^0\,d\omega \tag{1.14}$$

$$\mathcal{E}^m_N \equiv \alpha_{Nm}\int_{S^2}\Phi^N_{0\,1}\bar{Y}^m_{N+1}\,d\omega \tag{1.15}$$

where S^2 is the 2-sphere, α_{Nm} are normalization constants and $d\omega$ is the surface element of $S^2(d\omega = \sin\theta d\theta d\varphi)$. In the case of axial symmetry, they reduce to the definitions given by JN.

The quadrupole solution is given in Appendix I and the octupole news is given in Appendix II.

II. GRAVITATIONAL THEORY

In this section we again follow the notation of Newman and Penrose. We assume that space–time is flat and that our coordinate and tetrad system is chosen as in I. Then if we define the gravitational field components as

$$\Psi_0 = -C_{\mu\nu\rho\sigma}l^\mu m^\nu l^\rho m^\sigma,$$
$$\Psi_1 = -C_{\mu\nu\rho\sigma}l^\mu n^\nu l^\rho m^\sigma,$$
$$\Psi_2 = -C_{\mu\nu\rho\sigma}\bar{m}^\mu n^\nu l^\rho m^\sigma,$$
$$\Psi_3 = -C_{\mu\nu\rho\sigma}\bar{m}^\mu n^\nu l^\rho n^\sigma,$$
$$\Psi_4 = -C_{\mu\nu\rho\sigma}\bar{m}^\mu n^\nu \bar{m}^\rho n^\sigma,$$

where $C_{\mu\nu\rho\sigma}$ is a linearized Weyl tensor, the Bianchi identities[12] become

$$D\Psi_1 - \bar{\delta}\Psi_0 = 4\rho\Psi_1 - 4\alpha\Psi_0, \tag{2.1a}$$
$$D\Psi_2 - \bar{\delta}\Psi_1 = 3\rho\Psi_2 - 2\alpha\Psi_1, \tag{2.1b}$$
$$D\Psi_3 - \bar{\delta}\Psi_2 = 2\rho\Psi_3, \tag{2.1c}$$
$$D\Psi_4 - \bar{\delta}\Psi_3 = \rho\Psi_4 + 2\alpha\Psi_3, \tag{2.1d}$$
$$\Delta\Psi_0 - \delta\Psi_1 = -\mu\Psi_0 - 2\beta\Psi_1, \tag{2.1e}$$
$$\Delta\Psi_1 - \delta\Psi_2 = -2\mu\Psi_1, \tag{2.1f}$$
$$\Delta\Psi_2 - \delta\Psi_3 = -3\mu\Psi_2 - 2\bar{\alpha}\Psi_3, \tag{2.1g}$$
$$\Delta\Psi_3 - \delta\Psi_4 = -4\mu\Psi_3 + 4\beta\Psi_4. \tag{2.1h}$$

The differential operators D, Δ, and δ and the spin coefficients ρ, α, β, and μ are the same as those introduced in the Maxwell theory.

The characteristic initial-value problem may be solved in a similar manner to the electromagnetic theory. Ψ_0 is specified on an arbitrary initial hypersurface (say $\mathfrak{N}_0$) as a function of r, θ, and φ. Ψ_4^0 (for the notation see JN) is specified on a world tube $\mathcal{J}$ as an arbitrary function of u, θ, and φ, thus

[11] I owe this suggestion to R. Penrose. An alternative name could be the associated spherical harmonics of first order since they are analogous to the associated Legendre polynomials of first order in axially symmetric theory.

[12] These can be regarded as the field equations of $C_{\mu\nu\rho\sigma}$.

corresponding to the news $\Phi_2^0(u, \theta, \varphi)$ in Maxwell theory. [Note that in the Einstein theory Ψ_4^0 is essentially the second time derivative of the news $\sigma^0(u, \theta, \varphi)$.] Ψ_1^0, Ψ_2^0, and Ψ_3^0 are specified as arbitrary functions of θ and φ on $\mathscr{I} \cap \mathfrak{N}_0$.

Assuming that $\Psi_0 = O(r^{-5})$ and that enough $\mathfrak{D}$ derivatives are $O(r^{-5})$ we have that $\Psi_1 = O(r^{-4})$, $\Psi_2 = O(r^{-3})$, $\Psi_3 = O(r^{-2})$, and $\Psi_4 = O(r^{-1})$; the peeling theorem in gravitational theory.[13] The most general multipole expansion is then given by

$$\Psi_0(u, r, \theta, \varphi) = \sum_{n=1}^{\infty} \frac{\Psi_0^{n-1}(u, \theta, \varphi)}{r^{n+4}}. \tag{2.2}$$

Solving the field equations (2.1), we find

$$\Psi_1 = \frac{\Psi_1^0}{r^4} - \frac{1}{\sqrt{2}}(\bar{\mathfrak{D}} + 2\cot\theta)\sum_{n=1}^{\infty}\Psi_0^{n-1}/nr^{n+4}, \tag{2.3a}$$

$$\Psi_2 = \Psi_2^0/r^3 - \frac{1}{\sqrt{2}}(\bar{\mathfrak{D}} + \cot\theta)\frac{\Psi_1^0}{r^4} + \tfrac{1}{2}(\bar{\mathfrak{D}} + \cot\theta)(\bar{\mathfrak{D}} + 2\cot\theta) \times \sum_{n=1}^{\infty}\Psi_0^{n-1}/n(n+1)r^{n+4}, \tag{2.3b}$$

$$\Psi_3 = \frac{\Psi_3^0}{r^2} - \frac{1}{\sqrt{2}}\bar{\mathfrak{D}}\frac{\Psi_2^0}{r^3} + \tfrac{1}{4}\bar{\mathfrak{D}}(\bar{\mathfrak{D}} + \cot\theta)\frac{\Psi_1^0}{r^4} - \frac{1}{2\sqrt{2}}\bar{\mathfrak{D}}(\bar{\mathfrak{D}} + \cot\theta)(\bar{\mathfrak{D}} + 2\cot\theta) \times \sum_{n=1}^{\infty}\Psi_0^{n-1}/n(n+1)(n+2)r^{n+4}, \tag{2.3c}$$

$$\Psi_4 = \frac{\Psi_4^0}{r} - \frac{1}{\sqrt{2}}(\bar{\mathfrak{D}} - \cot\theta)\frac{\Psi_3^0}{r^2} + \tfrac{1}{4}(\bar{\mathfrak{D}} - \cot\theta)\bar{\mathfrak{D}}\frac{\Psi_2^0}{r^3} - \frac{1}{12\sqrt{2}}(\bar{\mathfrak{D}} - \cot\theta)\bar{\mathfrak{D}}(\bar{\mathfrak{D}} + \cot\theta)\frac{\Psi_1^0}{r^4} + \tfrac{1}{4}(\bar{\mathfrak{D}} - \cot\theta)\bar{\mathfrak{D}}(\bar{\mathfrak{D}} + \cot\theta)(\bar{\mathfrak{D}} + 2\cot\theta) \times \sum_{n=1}^{\infty}\Psi_0^{n-1}/n(n+1)(n+2)(n+3)r^{n+4}, \tag{2.3d}$$

$$\dot{\Psi}_3^0 = (1/\sqrt{2})(\mathfrak{D} + 2\cot\theta)\Psi_4^0, \tag{2.3e}$$

$$\dot{\Psi}_2^0 = (1/\sqrt{2})(\mathfrak{D} + \cot\theta)\Psi_3^0, \tag{2.3f}$$

$$\dot{\Psi}_1^0 = (1/\sqrt{2})\mathfrak{D}\Psi_2^0, \tag{2.3g}$$

$$\dot{\Psi}_0^0 = (1/\sqrt{2})(\mathfrak{D} - \cot\theta)\Psi_1^0, \tag{2.3h}$$

$$\dot{\Psi}_0^n = -\tfrac{1}{2}(n+3)\Psi_0^{n-1} - (1/2n)(\mathfrak{D} - \cot\theta) \times (\bar{\mathfrak{D}} + 2\cot\theta)\Psi_0^{n-1}, \quad n \geq 1. \tag{2.3j}$$

Making the same assumptions as in Sec. I we find that the angular dependence of Ψ_0^N is determined by

$$(\mathfrak{D} - \cot\theta)(\bar{\mathfrak{D}} + 2\cot\theta)\Psi_0^N + (N+1)(N+4)\Psi_0^N = 0. \tag{2.4}$$

In the axially symmetric case, $\mathfrak{D} = \bar{\mathfrak{D}} = \partial/\partial\theta$ and the above equation reduces to

$$\left[\frac{\partial^2}{\partial\theta^2} + \cot\theta\frac{\partial}{\partial\theta} + (N+2)(N+3) - \frac{4}{\sin^2\theta}\right]\Psi_0^N = 0$$

which gives $\Psi_0^N(u, \theta) = A_{N+2}(u)P_{N+2}^2(\cos\theta)$ as found previously in JN.

As in Maxwell theory we assume that the angular dependence of Ψ_0^N is of the form $e^{im\varphi}$ where m is a real integer. If we let

$$\Psi_0^N(u, \theta, \varphi) = A_{N+2,m}(u)e^{im\varphi}\chi(\cos\theta)$$

we see that

$$\left[\frac{\partial^2}{\partial\theta^2} + \cot\theta\frac{\partial}{\partial\theta} + (N+2)(N+3) - \frac{m^2 + 4m\cos\theta + 4}{\sin^2\theta}\right]\chi = 0. \tag{2.5}$$

The solution is readily found,[10]

$$\chi(\mu) = (1+\mu)^{\frac{1}{2}(m-2)}(1-\mu)^{\frac{1}{2}(m+2)} \times F(n+1+p, p-n; m+3; \tfrac{1}{2}(1-\mu)), \tag{2.6}$$

where $\mu = \cos\theta$, $2p = M + k$, $M^2 = (m-2)^2$, $k^2 = (m+2)^2$ and $n = N + 2$. Written in this form it is seen as a natural generalization of the multipole solution of Maxwell's theory. Hence we write the most general multipole ($N \geq 2$) as

$$\Psi_0^N(u, \theta, \varphi) = \sum_m A_{N+2,m}(u)\,{}_2Y_{N+2}^m(\theta, \varphi), \tag{2.7}$$

where ${}_2Y_{N+2}^m(\theta, \varphi) = e^{im\varphi}\chi(\cos\theta)$. We shall call ${}_2Y_{N+2}^m$ the tensor spherical harmonics (or associated spherical harmonics of the second order).

Since the theory is linear, without further ado, we have that

$$\Psi_0^0(u, \theta, \varphi) = k_0(N, m)(d^N/du^N)A_{N+2,m}(u)\,{}_2Y_{N+2}^m(\theta, \varphi), \tag{2.8a}$$

$$\Psi_1^0(u, \theta, \varphi) = k_1(N, m)(d^{N+1}/du^{N+1})A_{N+2,m}(u)\,{}_1Y_{N+2}^m(\theta, \varphi), \tag{2.8b}$$

$$\Psi_2^0(u, \theta, \varphi) = k_2(N, m)(d^{N+2}/du^{N+2})A_{N+2,m}(u)\,Y_{N+2}^m(\theta, \varphi), \tag{2.8c}$$

$$\Psi_3^0(u, \theta, \varphi) = k_3(N, m)(d^{N+3}/du^{N+3})A_{N+2,m}(u)\,{}_1\bar{Y}_{N+2}^{-m}(\theta, \varphi), \tag{2.8d}$$

[13] E. T. Newman and T. W. J. Unti, J. Math. Phys. 3, 891 (1962).

$$\Psi_4^0(u, \theta, \varphi) = k_4(N, m)(d^{N+4}/du^{N+4})A_{N+2,m}(u)\,{}_2\bar{Y}^{-m}_{N+2}(\theta, \varphi), \quad (2.8e)$$

where the k_i's are constants depending on N and m. The most general news functions Ψ_4^0 is obtained by linear superposition of the above. Hence

$$\Psi_4^0(u, \theta, \varphi) = \sum_{l=0}^{\infty}\sum_{m=-l}^{+l} k_4(l, m) \times \frac{d^{l+4}}{du^{l+4}} A_{l+2,m}(u)\,{}_2\bar{Y}^{-m}_{l+2}(\theta, \varphi). \quad (2.9)$$

Using the same argument as in Sec. I we can write

$$\begin{aligned} {}_2Y^m_n &= \sin\theta\,\eth({}_1Y^m_n/\sin\theta) \\ &= \sin\theta\,\eth(\eth Y^m_n/\sin\theta). \end{aligned} \quad (2.10)$$

When $m = 0$, we obtain the well-known formula

$$\begin{aligned} P^2_n &= \sin\theta\frac{d}{d\theta}\left(\frac{P^1_n}{\sin\theta}\right) \\ &= \sin\theta\frac{d}{d\theta}\left(\frac{dP_n/d\theta}{\sin\theta}\right). \end{aligned}$$

The most general news function can thus be written

$$\Psi_4^0(u, \theta, \varphi) = \sum_{m,n}\beta_{mn}\sin\theta\,\eth\left(\frac{\eth Y^m_n}{\sin\theta}\right), \quad (2.11)$$

where β_{mn} are functions of time. The form (2.11) is most useful for general relativity. From Newman and Unti[13] equation (40i) shows us that

$$\ddot{\sigma}^0(u, \theta, \varphi) = \sum_{m,n}\ddot{\sigma}^0_{mn}\sin\theta\,\eth\left(\frac{\eth Y^m_n}{\sin\theta}\right), \quad (2.12)$$

where $\ddot{\sigma}^0_{mn}$ are functions of time. A news function specified in this manner ensures the existence of a Minkowski tangent space in the neighborhood of the polar axis. Physically this means that the coordinate system is valid at large distances from isolated sources (assuming that the Ψ_i's peel).

Returning to the linear theory we see that the two cases $\Psi_0 = 0$ and $\Psi_0 = \Psi_1 = 0$ have already been solved in the axially symmetric theory. Since they correspond to a monopole and a dipole distribution it follows that the full linear theory can yield no new information. Conservation of mass and conservation of linear momentum and intrinsic angular momentum are thus assured as in JN.

Just as in electromagnetic theory we can introduce definitions of the mass $\mathfrak{M}$, the components of dipole moment—intrinsic angular momentum $\mathfrak{L}^m$ and multipole moments $\mathfrak{S}^m_N$; viz.

$$\mathfrak{M} \equiv -\int_{s^2}\Psi_2^0\,d\omega, \quad (2.13a)$$

$$\mathfrak{L}^m \equiv \beta_m\int_{s^2}\Psi^0_1\,{}_1\bar{Y}^m_1\,d\omega, \quad (2.13b)$$

$$\mathfrak{S}^m_N \equiv \beta_{Nm}\int_{s^2}\Psi^N_0\,{}_2\bar{Y}^m_{N+2}\,d\omega, \quad N \geq 2, \quad (2.13c)$$

where β_m and β_{Nm} are constants. The integrals are defined over the same surfaces as in electromagnetic theory.

In the general theory of relativity we may make similar definitions. Introducing appropriate normalization constants we have

$$\mathfrak{M} \equiv -\frac{1}{4\pi}\int_{s^2}\mathfrak{Re}\,\Psi_2^0\,d\omega, \quad (2.14a)$$

$$\mathcal{P}_x \equiv -\frac{1}{4\pi}\int_{s^2}\mathfrak{Re}\,\Psi_2^0\sin\theta\cos\varphi\,d\omega, \quad (2.14b)$$

$$\mathcal{P}_y \equiv -\frac{1}{4\pi}\int_{s^2}\mathfrak{Re}\,\Psi_2^0\sin\theta\sin\varphi\,d\omega, \quad (2.14c)$$

$$\mathcal{P}_z \equiv -\frac{1}{4\pi}\int_{s^2}\mathfrak{Re}\,\Psi_2^0\cos\theta\,d\omega, \quad (2.14d)$$

with the other moments defined as previously. The $\mathcal{P}_i$'s are components of the linear momentum.

From Newman and Unti[13] we use Eq. (42c) to determine the time dependence of the above quantities. Then we have

$$\frac{d}{du}\mathfrak{M} = -\frac{1}{4\pi}\int_{s^2}|\dot{\sigma}^0|^2\,d\omega + \frac{1}{8\pi}\int_{s^2}\frac{\partial^2}{\partial u^2}|\sigma^0|^2\,d\omega, \quad (2.15a)$$

$$\frac{d}{du}\mathcal{P}_x = -\frac{1}{4\pi}\int_{s^2}|\dot{\sigma}^0|^2\sin\theta\cos\varphi\,d\omega + \frac{1}{8\pi}\int_{s^2}\frac{\partial^2}{\partial u^2}|\sigma^0|^2\sin\theta\cos\varphi\,d\omega, \quad (2.15b)$$

$$\frac{d}{du}\mathcal{P}_y = -\frac{1}{4\pi}\int_{s^2}|\dot{\sigma}^0|^2\sin\theta\sin\varphi\,d\omega + \frac{1}{8\pi}\int_{s^2}\frac{\partial^2}{\partial u^2}|\sigma^0|^2\sin\theta\sin\varphi\,d\omega, \quad (2.15c)$$

$$\frac{d}{du}\mathcal{P}_z = -\frac{1}{4\pi}\int_{s^2}|\dot{\sigma}^0|^2\cos\theta\,d\omega + \frac{1}{8\pi}\int_{s^2}\frac{\partial^2}{\partial u^2}|\sigma^0|^2\cos\theta\,d\omega. \quad (2.15d)$$

The total change of 4-momentum is obtained by integrating the above expressions over all time. If we consider "sandwich" processes then $\dot{\sigma}^0\,|_{u=-\infty} = \dot{\sigma}^0\,|_{u=+\infty} = 0$ and we see that the above expressions suffer the same change in 4-momentum given by previous authors.[6,7]

Thus, with the above assumptions,

$$\Delta\mathfrak{M} = -\frac{1}{4\pi}\iint_{s^2} |\dot{\sigma}^0|^2\, d\omega\, du,$$

$$\Delta\mathcal{P}_x = -\frac{1}{4\pi}\iint_{s^2} |\dot{\sigma}^0|^2 \sin\theta\cos\varphi\, d\omega\, du,$$

$$\Delta\mathcal{P}_y = -\frac{1}{4\pi}\iint_{s^2} |\dot{\sigma}^0|^2 \sin\theta\sin\varphi\, d\omega\, du,$$

$$\Delta\mathcal{P}_z = -\frac{1}{4\pi}\iint_{s^2} |\dot{\sigma}^0|^2 \cos\theta\, d\omega\, du.$$

These expressions form a 4-vector under the GBM transformations.[14] When a physical process is news-free (i.e., when $\dot{\sigma}^0 = 0$), the strong form of the expressions [i.e., Eqs. (2.14)] transform as a 4-vector (see Møller[6]).

ACKNOWLEDGMENTS

I wish to thank Dr. E. T. Newman for his helpful discussions and advice and Professor H. Bondi for his useful criticism of an earlier draft of the manuscript.

APPENDIX I

Electromagnetic Quadrupole

$$\begin{aligned}\Phi_0^1(u,\theta,\varphi) = {} & A_{2,2}(u)e^{2i\varphi}\sin\theta(1-\cos\theta) \\ & + A_{2,1}(u)e^{i\varphi}(\cos 2\theta - \cos\theta) \\ & + A_{2,0}(u)\sin\theta\cos\theta \\ & + A_{2,-1}(u)e^{-i\varphi}(\cos 2\theta + \cos\theta) \\ & + A_{2,-2}(u)e^{-2i\varphi}\sin\theta(1+\cos\theta),\end{aligned}$$

$$\Phi_0^0(u,\theta,\varphi) = \tfrac{1}{2}\dot{\Phi}_0^1(u,\theta,\varphi),$$

$$\begin{aligned}\Phi_1^0(u,\theta,\varphi) = {} & -\frac{\ddot{A}_{2,2}}{2\sqrt{2}}e^{2i\varphi}\sin^2\theta \\ & - \frac{\ddot{A}_{2,1}}{\sqrt{2}}e^{i\varphi}\sin\theta\cos\theta - \frac{\ddot{A}_{2,0}}{6\sqrt{2}}(3\cos^2\theta - 1) \\ & + \frac{\ddot{A}_{2,-1}}{\sqrt{2}}e^{-i\varphi}\sin\theta\cos\theta + \frac{\ddot{A}_{2,-2}}{2\sqrt{2}}e^{-2i\varphi}\sin^2\theta,\end{aligned}$$

$$\begin{aligned}\Phi_2^0(u,\theta,\varphi) = {} & \tfrac{1}{6}A_{2,2}^{(3)}e^{2i\varphi}\sin\theta(1+\cos\theta) \\ & + \tfrac{1}{6}A_{2,1}^{(3)}e^{i\varphi}(\cos 2\theta + \cos\theta) - \tfrac{1}{6}A_{2,0}^{(3)}\sin\theta\cos\theta \\ & + \tfrac{1}{6}A_{2,-1}^{(3)}e^{-i\varphi}(\cos\theta - \cos 2\theta) \\ & + \tfrac{1}{6}A_{2,-2}^{(3)}e^{-2i\varphi}\sin\theta(1-\cos\theta).\end{aligned}$$

Gravitational Quadrupole

$$\begin{aligned}\Psi_0^0(u,\theta,\varphi) = {} & B_{2,2}(u)e^{2i\varphi}[-\sin^2\theta + 2(1-\cos\theta)] \\ & + B_{2,1}(u)e^{i\varphi}\sin\theta(1-\cos\theta) + B_{2,0}(u)\sin^2\theta \\ & + B_{2,-1}(u)e^{-i\varphi}\sin\theta(1+\cos\theta) \\ & + B_{2,-2}(u)e^{-2i\varphi}[-\sin^2\theta + 2(1+\cos\theta)],\end{aligned}$$

$$\begin{aligned}\Psi_1^0(u,\theta,\varphi) = {} & -\sqrt{2}\dot{B}_{2,2}e^{2i\varphi}\sin\theta(1-\cos\theta) \\ & - \frac{1}{\sqrt{2}}\dot{B}_{2,1}e^{i\varphi}(\cos\theta - \cos 2\theta) \\ & - \sqrt{2}\dot{B}_{2,0}\sin\theta\cos\theta \\ & + \frac{1}{\sqrt{2}}\dot{B}_{2,-1}e^{-i\varphi}(\cos\theta + \cos 2\theta) \\ & + \sqrt{2}\dot{B}_{2,-2}e^{-2i\varphi}\sin\theta(1+\cos\theta),\end{aligned}$$

$$\begin{aligned}\Psi_2^0(u,\theta,\varphi) = {} & \ddot{B}_{2,2}e^{2i\varphi}\sin^2\theta \\ & + \ddot{B}_{2,1}e^{i\varphi}\sin\theta\cos\theta + \tfrac{1}{3}\ddot{B}_{2,0}(3\cos^2\theta - 1) \\ & + \ddot{B}_{2,-1}e^{-i\varphi}\sin\theta\cos\theta + \ddot{B}_{2,-2}e^{-2i\varphi}\sin^2\theta,\end{aligned}$$

$$\begin{aligned}\Psi_3^0(u,\theta,\varphi) = {} & -\frac{\sqrt{2}}{3}B_{2,2}^{(3)}e^{2i\varphi}\sin\theta(1+\cos\theta) \\ & - \frac{\sqrt{2}}{6}B_{2,1}^{(3)}e^{i\varphi}(\cos\theta + \cos 2\theta) \\ & + \frac{\sqrt{2}}{3}B_{2,0}^{(3)}\sin\theta\cos\theta \\ & + \frac{\sqrt{2}}{6}B_{2,-1}^{(3)}e^{-i\varphi}(\cos\theta - \cos 2\theta) \\ & + \frac{\sqrt{2}}{3}B_{2,-2}^{(3)}e^{-2i\varphi}\sin\theta(1-\cos\theta),\end{aligned}$$

$$\begin{aligned}\Psi_4^0(u,\theta,\varphi) = {} & \tfrac{1}{6}B_{2,2}^{(4)}e^{2i\varphi}[-\sin^2\theta + 2(1+\cos\theta)] \\ & - \tfrac{1}{6}B_{2,1}^{(4)}e^{i\varphi}\sin\theta(1+\cos\theta) + \tfrac{1}{6}B_{2,0}^{(4)}\sin^2\theta \\ & + \tfrac{1}{6}B_{2,-1}^{(4)}e^{-i\varphi}\sin\theta(1-\cos\theta) \\ & + \tfrac{1}{6}B_{2,-2}^{(4)}e^{-2i\varphi}[-\sin^2\theta + 2(1+\cos\theta)].\end{aligned}$$

APPENDIX II

Octupole Electromagnetic News

$$\begin{aligned}\Phi_2^0(u,\theta,\varphi) = {} & A_{3,3}(u)e^{3i\varphi}\sin^2\theta(1+\cos\theta) \\ & + A_{3,2}(u)e^{2i\varphi}\sin\theta(1+\cos\theta)(1-3\cos\theta) \\ & + A_{3,1}(u)e^{i\varphi}(1+\cos\theta)(15\cos^2\theta - 10\cos\theta - 1) \\ & + A_{3,0}(u)\sin\theta(5\cos^2\theta - 1) \\ & + A_{3,-1}(u)e^{-i\varphi}(1-\cos\theta)(15\cos^2\theta + 10\cos\theta - 1) \\ & + A_{3,-2}(u)e^{-2i\varphi}\sin\theta(1-\cos\theta)(1+3\cos\theta) \\ & + A_{3,-3}(u)e^{-3i\varphi}\sin^2\theta(1-\cos\theta).\end{aligned}$$

Octupole Gravitational News

$$\begin{aligned}\Psi_4^0(u,\theta,\varphi) = {} & B_{3,3}(u)e^{3i\varphi}\sin\theta(1+\cos\theta)^2 \\ & + B_{3,2}(u)e^{2i\varphi}(1+\cos\theta)^2(2-3\cos\theta) \\ & + B_{3,1}(u)e^{i\varphi}\sin\theta(1+\cos\theta)(1-3\cos\theta) \\ & + B_{3,0}(u)\sin^2\theta\cos\theta \\ & + B_{3,-1}(u)e^{-i\varphi}\sin\theta(1-\cos\theta)(1+3\cos\theta) \\ & + B_{3,-2}(u)e^{-2i\varphi}(1-\cos\theta)^2(2+3\cos\theta) \\ & + B_{3,-3}(u)e^{-3i\varphi}\sin\theta(1-\cos\theta)^2.\end{aligned}$$

[14] R. K. Sachs, Phys. Rev. 128, 2851 (1962).

A General Theory of First-Passage Distributions in Transport and Multiplicative Processes

J. E. MOYAL
Argonne National Laboratory, Argonne, Illinois
(Received 7 July 1965)

The "Milne problem," expressed in probabilistic terms, is solved for general transport and multiplicative processes. If a particle initially in a given state at a given position inside a surface τ is multiply scattered while traveling through a fixed medium, then given the scattering cross sections and, if required, the probability distribution for a change of state between collisions (e.f., by diffusion or ionization), the problem is to obtain the probability that the particle eventually effects a first passage through a specified position on the surface τ and in a specified state. In the case of a multiplicative process, the problem is, given in addition the rates of creation and annihilation of particles (considering the nature of the particle as a state variable), to obtain the probability that eventually n particles will emerge for the first time through specified positions on τ and in specified states (with $n = 0, 1, 2, \cdots$). A general solution is given in the form of a convergent series whose terms are obtained by iteration; this solution is unique if and only if the probability θ_∞ of an infinity of atomic events before a first passage (which is the limit of a certain nonincreasing sequence) is identically zero; in the multiplicative case $\theta_\infty \not\equiv 0$ may be taken to mean that the process is "supercritical." The mathematical theory which leads to this solution is a generalization of the corresponding theory for time-dependent Markov processes in which the time variable is replaced by a set of surfaces ordered by inclusion of their "insides" and is valid for Euclidean space of any number of dimensions. Applying it to the 4-dimensional space of special relativity with ordered sets of spacelike surfaces, one obtains a Lorentz-invariant formulation of the theory of physical Markov processes. A few examples are given.

1. INTRODUCTION

THERE exists a variety of physical processes having the Markovian character where a particle suffers a succession of independent random scatterings while travelling through some medium. Examples are: the diffuse scattering of light, where the particle is a photon, the diffusion of neutrons, the multiple scattering of charged particles. In connection with such processes one is often interested in the probability distribution of the state variables of the particle (velocity, energy, spin, etc.) and of the position at which it emerges on its first passage through some surface independently of time; the position probability will yield by integration the average flux density in the case of a source or beam of particles. The theory of such *first-passage distributions* goes by the name of theory of radiative transfer in the case of the scattering of light (c.f. Chandrasekhar,[1] Sobolev[2]; see also Wing[3], where such problems are discussed in a more general context), and the problem of obtaining the first-passage distributions given the microscopic scattering laws is essentially the well-known Milne problem, which has been solved exactly under rather restrictive simplifying assumptions by means of the Wiener–Hopf integral equation technique (c.f. Busbridge[4]). The purpose of the present paper is to present a general theory for the solution of such first-passage problems. We show that a solution always exists and present it in the form of a convergent series whose terms are obtained by iteration, and we give necessary and sufficient conditions for this solution to be unique. Furthermore, we see that the theory developed for this purpose can also be made to yield a Lorentz-invariant formulation of the basic equations of physical Markov processes. Finally, we show that the theory generalizes to first-passage problems in the case of processes involving the creation and annihilation of particles, such as the multiplication of neutrons in fissionable material and electron–photon or mucleon cascades.

2. FIRST-PASSAGE PROCESSES

The theory we are going to develop is based on a generalization of the concept of a Markov process: such a process is usually defined on a linearly ordered set (the time axis) and the required generalization consists in extending this definition to a partially ordered set. Let T be such a set, and let $\tau \geq \tau_0$ denote the ordering relation. For the purposes of the present theory, T will consist of a set of two-sided, oriented, continuous and simply-connected surfaces in some finite-dimensional Euclidean space

[1] S. Chandrasekhar, *Radiative Transfer* (Oxford University Press, New York, 1950).
[2] V. V. Sobolev, *A Treatise on Radiative Transfer* (D. Van Nostrand, Inc., Princeton, N. J., 1963).
[3] G. M. Wing, *An Introduction to Transport Theory* (John Wiley & Sons, Inc., New York, 1962).
[4] I. W. Busbridge, *The Mathematics of Radiative Transfer* (Cambridge University Press, New York, 1960).

X. Each such surface τ partitions X into two disjoint sets: the "inside" X_τ of τ (which conventionally will include τ) and the "outside" X'_τ of τ. We partially order T by inclusion of the "insides" of its elements: i.e., we set $\tau \geq \tau_0$ whenever $X_\tau \supseteq X_{\tau_0}$. In the applications to scattering processes $X = R_3$; in the Lorentz-invariant formulation of Markov process theory $X = R_4$; the whole theory can be extended to more general topological spaces than R_n, but we do not consider this generalization here. To each $\tau \in T$ is assigned a space Ω_τ of elementary events ω_τ and a σ-field F_τ of measurable subsets Γ_τ of Ω_τ. If $\mathfrak{A}$ is the set of all possible states α of the particle (velocity, energy, spin, etc.) and if x_τ denotes the position at which it emerges on its first passage through the surface τ, then $\omega_\tau = (\alpha, x_\tau)$ and $\Omega_\tau = \mathfrak{A} \times \tau$, which is a subset of the space $\Omega = \mathfrak{A} \times X$. We assume given σ-fields B_A of sets A in $\mathfrak{A}$ and B_X of sets S in X such that $T \subseteq B_X$ (i.e., each surface τ is measurable). It follows that the class of all measurable subsets S_τ of τ is itself a σ-field B_τ, and we set $F_\tau = B_A \times B_\tau$, which is a subfield of the σ-field $F = B_A \times B_X$. Suppose, furthermore, that for each ordered pair $\tau \geq \tau_0$ of elements of T we have defined a function P on $F_\tau \times \Omega_{\tau_0}$ such that for each fixed $\omega_{\tau_0} \in \Omega_{\tau_0}$, $P(\cdot \mid \omega_{\tau_0})$ is a measure on F_τ satisfying the normalization condition

$$\kappa(\tau \mid \omega_{\tau_0}) = P(\Omega_\tau \mid \omega_{\tau_0}) \leq 1, \tag{2.1}$$

and for each fixed set $\Gamma_\tau \in F_\tau$, $P(\Gamma_\tau \mid \cdot)$ is a measurable function on Ω_{τ_0}: then P has the character of an incomplete conditional probability distribution (incomplete in the sense that it is not normalized to unity). For our present purposes, P is interpreted as follows: given that the particle is initially at the point $x_{\tau_0} \in X_\tau$ in the state α_0, $P(A \times S_\tau \mid \alpha_0, x_{\tau_0})$ is the probability that it effects a first passage through some point of the set $S_\tau \subseteq \tau$ in some state $\alpha \in A$: hence, $\kappa(\tau \mid \alpha_0, x_{\tau_0})$ is the total probability of a first passage through τ and $\eta = 1 - \kappa$ is the probability that the particle never passes through τ, which will in general be the sum of a *stopping probability* σ (due to slowing down of the particle in the scattering medium) and an *escape probability* ε (due to the particle escaping to infinity if τ is not closed). For this reason we shall call P a *first-passage distribution* and η a *no-passage probability*; it is precisely because we have to allow for processes with nonzero η that we do not require P to be normalized to unity.

We now say that the family of all first-passage distributions defines a *generalized Markov process* over the partially ordered set T if its elements P satisfy the Chapman–Kolmogorov relation

$$P(\Gamma_\tau \mid \omega_{\tau_0}) = \int_{\Omega_{\tau_1}} P(\Gamma_\tau \mid \omega_{\tau_1}) P(d\omega_{\tau_1} \mid \omega_{\tau_0}) \tag{2.2}$$

for every ordered triple $\tau \geq \tau_1 \geq \tau_0$. It follows from this definition that the transition distribution P defines an ordinary incomplete Markov process over every linearly ordered chain in T. We may therefore regard such a generalized process as a family of ordinary Markov processes over the chains in T mutually related by (2.2). It is clear that this family does not define a stochastic process over T in the usual sense, because the conditional distributions P are not defined for pairs of elements of T which are not related by ordering. For the purposes of the present paper, where P has the interpretation outlined above, we call such a generalized Markov process a *first-passage process*.

3. LORENTZ-INVARIANT FORMULATION OF MARKOV PROCESS EQUATIONS

Suppose now that X is the Minkowski 4-dimensional space-time of special relativity R_4, and take B_X to be the Borel field of subsets of R_4. If we choose T to be the set of all *space-like* surfaces in R_4 satisfying the assumptions made above, then it is clear that the *definition above yields a relativistically invariant formulation of the concept of a temporal particle Markov process*, in the sense that the Chapman–Kolmogorov relation (2.2) is then invariant under Lorentz transformations, and the same will be true, as we see later, of other basic equations, such as the integral equation (4.2) and the "backward" integro–differential equation (5.6). If we choose a particular Galileian frame of reference L in R_4, and consider the linear chain T_l in T consisting of all *flat* spacelike surfaces τ normal to the time axis in L, then clearly we can assimilate τ to the time coordinate and T_l to the time axis in L; x_τ is then a point x, S_τ a Borel subset S of 3-dimensional space R_3 at the time τ, and the first-passage distribution P, restricted to T_l, defines an ordinary temporal Markov process over T_l, with τ as parameter, satisfying the usual Chapman–Kolmogorov relation

$$\begin{aligned} &P(A \times S; \tau \mid \alpha_0, x_0; \tau_0) \\ &= \int_{\mathfrak{A} \times R_3} P(A \times S; \tau \mid \alpha_1, x_1; \tau_1) \\ &\quad \times P(d\alpha_1\, dx_1; \tau_1 \mid \alpha_0, x_0; \tau_0)(\tau \geq \tau_1 \geq \tau_0). \end{aligned} \tag{3.1}$$

The no-passage probability η defined in Sec. 2, if it is not identically zero, must clearly be inter-

preted in this context as the cumulative distribution of the *lifetime* of the particle: i.e., $\eta(\tau \mid \alpha_0, x_0; \tau_0)$ is the probability that the particle initially in state α_0 and position x_0 at time τ_0 is annihilated at some time $t \leq \tau$. We remark that a Lorentz-invariant formulation of this kind is not appropriate in the case of multiple scattering processes of the type studied in the sections that follow, because for such processes there exists an obviously preferred class of reference systems, namely, those where the scattering medium is at rest.

4. DISCONTINUOUS FIRST-PASSAGE PROCESSES

In the present section, we extend to first-passage processes the theory of discontinuous Markov processes developed in Moyal[5,6] (referred to henceforth, respectively, as I and II). The type of process we have in mind is one where the particle suffers multiple collisions, each causing an instantaneous change in its state. We assume that *between* collisions the process is governed by a known first-passage distribution P_0; more precisely, $P_0(A \times S_\tau | \alpha_0, x_{\tau_0})$ is the probability that the particle, initially in a state α_0 at the point $x_{\tau_0} \in X_\tau$ (the "interior" of τ) effects a first passage through S_τ in some state $\alpha \in A$ *before* it has suffered any collision. We assume that P_0 satisfies the Chapman–Kolmogorov relation (2.2), so that it defines a first-passage process *dependent on no collisions*. This formulation has the virtue of including processes where not only the position, but also the state of the particle can change between collisions: for example, P_0 may characterize changes of velocity by diffusion, or loss of energy by ionization. If between collisions the particle moves in a straight line with its velocity and other state variables remaining constant, then we have a *purely discontinuous* (or *pure multiple scattering*) process: this case is dealt with in greater detail in Sec. 5. The effect of the collisions is assumed to be specified by a known *first collision and consequent state distribution* Q, which is a conditional distribution on $\mathfrak{F} \times \Omega$. $Q(A \times S; \tau \mid \alpha_0, x_{\tau_0})$ is interpreted as the probability that the particle, initially in state α_0 at $x_{\tau_0} \in X_\tau$, suffers its first collision at some point $x \in S$ (where S is a measurable subset of X) *before* it has made a first passage through the surface τ, and that its state immediately consequent to this first collision is some $\alpha \in A$. It follows from this definition that $Q(A \times S; \tau \mid \alpha_0, x_{\tau_0}) \equiv Q(A \times (S \cap X_\tau); \tau | \alpha_0, x_{\tau_0})$. The first-passage distribution P_0 and the conditional distribution Q are assumed to be related as follows: for every ordered triple $\tau \geq \tau_1 \geq \tau_0$,

$$Q(A \times S; \tau \mid \alpha_0, x_{\tau_0}) = Q(A \times S; \tau_1 \mid \alpha_0, x_{\tau_0}) + \int_{\alpha \times \tau} Q(A \times S; \tau \mid \alpha_1, x_{\tau_1}) P_0(d\alpha_1\, dx_{\tau_1} \mid \alpha_0, x_{\tau_0}). \tag{4.1}$$

The intuitive meaning of this relation is that the first collision and consequent state probability in its left-hand side, which depends on the particle not making a first passage through τ, is the sum of two first collision and consequent state probabilities, the first (which is the first term in its right-hand side) dependent on no first passage through $\tau_1 \leq \tau$, the second (which is the second term in its right-hand side) dependent on a first passage through τ_1 without collision and no first passage through τ.

The first-passage distribution P we are seeking must then satisfy the following integral equation:

$$P(A \times S_\tau \mid \alpha_0, x_{\tau_0}) = P_0(A \times S_\tau \mid \alpha_0, x_{\tau_0}) + \int_{\alpha \times X_\tau} P(A \times S_\tau \mid \alpha, x) Q(d\alpha\, dx \mid \alpha_0, x_{\tau_0}), \tag{4.2}$$

which we also write in the abbreviated notation $P = P_0 + P * Q$, where the symbol $*$ stands for the integration operation which occurs in the second term in the right-hand side of (4.2). The intuitive meaning of this equation is that the first-passage distribution P in its left-hand side is the sum of two first-passage distributions: the first P_0 with no collisions, and the second (which is the second term in its right-hand side) with at least one collision. The problem that now confronts us is that of the *existence* of a solution of (4.2) which is a first-passage distribution satisfying the Chapman–Kolmogorov relation (2.2), and of the conditions under which this solution is unique.

We define as in I and II two sequences $\{Q_n\}$ $\{P_n\}$, where $Q_1 = Q$, $Q_{n+1} = Q_n * Q$ and $P_n = P_0 * Q_n$, $n = 1, 2, \cdots$. If the particle is initially in the state α_0 at $x_{\tau_0} \in X_\tau$, then $Q_n(A \times S; \tau \mid \alpha_0, x_{\tau_0})$ represents the probability that the n-th collision occurs in S with consequent state $\alpha \in A$ before the particle effects a first passage through τ; $\theta_n(\tau \mid \alpha_0, x_{\tau_0}) = Q_n(\mathfrak{A} \times X; \tau \mid \alpha_0, x_{\tau_0})$ therefore represents the probability that the particle suffers at least n collisions in X_τ before it effects a first passage through τ; $P_n(A \times S_\tau \mid \alpha_0, x_{\tau_0})$ represents the probability that the particle effects a first passage through S_τ while in a state $\alpha \in A$ and after suf-

[5] J. E. Moyal, Acta Math. Stockholm **98**, 221 (1957).
[6] J. E. Moyal, J. Appl. Probability **2**, 69 (1965).

fering *exactly* n collisions in X_τ; $\kappa_n(\tau \mid \alpha_0, x_{\tau_0}) = P_n(\mathcal{A} \times \tau \mid \alpha_0, x_{\tau_0})$ therefore represents the probability of a first passage through τ after exactly n collisions in X_τ. Let $\eta_0 = 1 - \kappa_0 - \theta_1$, and let $\eta_n = \eta_0 * Q_n$, $n = 1, 2, \cdots$. It is then easy to see that $\eta_n(\tau \mid \alpha_0, x_{\tau_0})$ is the probability that the particle suffer *exactly* n collisions without ever making a first passage through τ. One can then show almost exactly as in the proof of (2.4) in II that

$$\theta_n = 1 - \sum_{i=0}^{n-1} (\kappa_i + \eta_i) \leq 1, \tag{4.3}$$

and that the sequence $\{\theta_n\}$ is nondecreasing. Let

$$\theta_\infty = \lim_{n\to\infty} \theta_n = 1 - \sum_{i=0}^{\infty} (\kappa_i + \eta_i) \leq 1. \tag{4.4}$$

We see from (4.4) that $\sum \kappa_n$ and $\sum \eta_n$ both converge. The series $\sum P_n$ is obviously majorized by $\sum \kappa_n$ and is, hence, convergent. Let

$$P_\tau = \sum_{n=0}^{\infty} P_n. \tag{4.5}$$

We then prove almost precisely as in the proof of Theorem 6.1 of I that P_τ is a first-passage distribution satisfying the Chapman–Kolmogorov relation (2.2) and the integral equation (4.2); we call P_τ the *regular solution* of (4.2).

One can show as in the proof of Theorem 8.3 of I that the regular solution P_τ is the minimal nonnegative solution of (4.2). Whether it is also its unique solution hinges on the values of θ_∞: it is unique if and only if θ_∞ is identically zero; this is shown as in the proof of the corollary to Theorem 6.2 of I. We call the process *stable* when $\theta_\infty \equiv 0$. These results are summarized in the following theorem:

Theorem 4.1: *The series* $\sum_{n=0}^{\infty} P_n$ *converges to a first-passage distribution* P_τ *which satisfies the Chapman–Kolmogorov relation* (2.2) *and is the minimal nonnegative solution of the integral equation* (4.2); *moreover,* P_τ *is the unique solution of* (4.2) *if and only if* $\theta_\infty \equiv 0$.

Let $\kappa_\tau = \sum_0^\infty \kappa_n$ and $\eta_\tau = \sum_0^\infty \eta_n$; then clearly $\kappa_\tau(\tau \mid \alpha_0, x_{\tau_0}) = P_\tau(\mathcal{A} \times \tau \mid \alpha_0, x_{\tau_0})$, and it follows from (4.4) that the corresponding no-passage probability $\eta = 1 - \kappa_\tau = \eta_\tau + \theta_\infty$. Thus, if the particle is initially in the state α_0 at x_{τ_0}, then the no-passage probability $\eta(\tau \mid \alpha_0, x_{\tau_0})$ is the sum of two probabilities, $\eta_\tau(\tau \mid \alpha_0, x_{\tau_0})$ and $\theta_\infty(\tau \mid \alpha_0, x_{\tau_0})$. Clearly, η_τ represents the stopping or escape probability after a finite number of collisions in X_τ, and hence, θ_∞ must be interpreted as the stopping or escape probability after an infinite number of collisions. Note that the process is stable if and only if $\eta = \eta_\tau$. It is also clear that $\eta_\tau \equiv 0$ if and only if the no-passage probability with no collisions $\eta_0 \equiv 0$. Hence, the total no-passage probability $\eta \equiv 0$ if and only if both $\eta \equiv 0$ and $\theta_\infty \equiv 0$.

5. FIRST-PASSAGE DISTRIBUTIONS IN PURE MULTIPLE SCATTERING PROCESSES

In the present section, we apply the theory outlined in Sec. 4 to the special case of a pure multiple scattering process in R_3, where only the position of the particle changes between collisions, the other state variables remaining constant. Let μ stand for the unit vector in the direction of motion of the particle. If the particle initially at x suffers no collisions while traveling a length of path s, then its position vector becomes $x + \mu s$. It is convenient to distinguish μ from the remaining state variables γ; thus $\alpha = (\gamma, \mu)$ and $\mathcal{A} = G \times M$, where G is the set of all γ and M is the set of all directions μ. We assume that the probability that a particle at x and in the state (γ, μ) suffers a collision while traveling a small distance δs is $\lambda(\gamma, \mu, x)\delta s + o(\delta s)$, and the probability of more than one collision is of order $o(\delta s)$; thus the *collision rate* per unit distance traveled λ is seen to be the inverse of the mean free path of the particle. We also assume known the transition probability ϕ for the particle state conditional on a collision: that is, $\phi(A \mid \gamma_0, \mu_0, x_0)$ is the probability, given a collision at x_0, of a transition from the state (γ_0, μ_0) to some state $(\gamma, \mu) \in A$. Let

$$R_\tau(\mu_0, x_0) = \min \{s \mid x_0 + \mu_0 s \in \tau\}, \tag{5.1}$$

where $x_0 \in X_\tau$ and we set $R_\tau(\mu_0, x_0) = \infty$ if $x_0 + \mu_0 s$ does not lie on the surface τ for any finite s. Then it is easy to see that

$$P_0(A \times S_\tau \mid \gamma_0, \mu_0, x_0) = \exp\left\{-\int_0^{R_\tau(\mu_0, x_0)} \lambda(x_0 + \mu_0 s)\, ds\right\} \times \delta(A \mid \gamma_0, \mu_0)\delta_L S_\tau \mid x_0 + \mu_0 R_\tau(\mu_0, x_0)], \tag{5.2}$$

where

$$\delta(A \mid \gamma_0, \mu_0) = \begin{cases} 1 & \text{if } (\gamma_0, \mu_0) \in A, \\ 0 & \text{otherwise} \end{cases}$$

$$\delta(S_\tau \mid x_0) = \begin{cases} 1 & \text{if } x_0 \in S_\tau, \\ 0 & \text{otherwise} \end{cases}$$

and we have written for brevity x_0 for x_{τ_0}, $\lambda(x_0 + \mu_0 s)$ for $\lambda(\gamma_0, \mu_0, x_0 + \mu_0 s)$. Similarly, one shows that

$$Q(A \times S; \tau \mid \gamma_0, \mu_0, x_0) = \int_0^{R(\mu_0, x_0)} \phi(A \mid \gamma_0, \mu_0, x_0 + \mu_0 s)\delta(S \mid x_0 + \mu_0 s) \exp\left\{-\int_0^s \lambda(x_0 + \mu_0\sigma)\, d\sigma\right\}\lambda(x_0 + \mu_0 s)\, ds. \tag{5.3}$$

We assert that

Lemma 5.1: The distributions P_0 and Q defined, respectively, by (5.2) *and* (5.3) *satisfy the consistency relation* (4.1).

Proof. Suppose that $\tau \geq \tau_1 \geq \tau_0$ and write x_0 for $x_{\tau 0}$, x_1 for x_{τ_1}; then

$$\begin{aligned}
&\int_{A\times\tau_1} Q(A \times S; \tau \mid \gamma_1, \mu_1, x_1)P_0(d\gamma_1, d\mu_1, dx_1 \mid \gamma_0, \mu_0, x_0) \\
&\quad = Q[A \times S; \tau \mid \gamma_0, \mu_0, x_0 + \mu_0 R_{\tau_1}(\mu_0, x_0)] \exp\left\{-\int_0^{R_{\tau_1}(\mu_0, x_0)} \lambda(x_0 + \mu_0 s)\, ds\right\} \\
&\quad = \int_0^{R_\tau[\mu_0, x_0 + \mu_0 R_{\tau_1}(\mu_0, x_0)]} \phi\{A \mid \gamma_0, \mu_0, x_0 + \mu_0[R_{\tau_1}(\mu_0, x_0) + s]\}\,\delta\{Sx_0 + \mu_0[R_{\tau_1}(\mu_0, x_0) + s]\} \\
&\qquad \times \exp\left\{-\int_0^s \lambda(x_0 + \mu_0\sigma)\, d\sigma - \int_0^{R_{\tau_1}(\mu_0, x_0)} \lambda(x_0 + \mu_0\sigma)\, d\sigma\right\}\lambda\{x_0 + \mu_0[R_{\tau_1}(\mu_0, x_0) + s]\}\, ds.
\end{aligned} \tag{5.4}$$

It is immediate from the definition of R_τ that

$$R_\tau[\mu_0, x_0 + \mu_0 R_{\tau_1}(\mu_0, x_0)] = R_\tau(\mu_0, x_0) - R_{\tau_1}(\mu_0, x_0).$$

Substituting this relation in the last line of (5.4) and changing the variables of integration to $\xi = R_{\tau_1}(\mu_0, x_0) + s$, $\zeta = R_{\tau_1}(\mu_0, x_0) + \sigma$, we transform the integral there to

$$\begin{aligned}
&\int_{R_{\tau_1}(\mu_0, x_0)}^{R_\tau(\mu_0, x_0)} \phi(A \mid \gamma_0, \mu_0, x_0 + \mu_0\xi)\delta(S \mid x_0 + \mu_0\xi) \\
&\qquad \times \exp\left\{-\int_{R_{\tau_1}(\mu_0, x_0)}^{\xi} \lambda(x_0 + \mu_0\zeta)\, d\zeta \right. \\
&\qquad \left. - \int_0^{R_{\tau_1}(\mu_0, x_0)} \lambda(x_0 + \mu_0\zeta)\, d\zeta\right\}\lambda(x_0 + \mu_0\xi)\, d\xi \\
&\quad = Q(A \times S; \tau \mid \gamma_0, \mu_0, x_0) \\
&\qquad - Q(A \times S; \tau_1 \mid \gamma_0, \mu_0, x_0).
\end{aligned}$$

This completes the proof of the lemma.

The integral equation (4.2) which the first-passage distribution P must satisfy becomes in the present case

$$\begin{aligned}
&P(A \times S_\tau \mid \gamma_0, \mu_0, x_0) = P_0(A \times S_\tau \mid \gamma_0, \mu_0, x_0) \\
&+ \int_A \int_0^{R_\tau(\mu_0, x_0)} P(A \times S_\tau \mid \gamma, \mu, x_0 + \mu_0 s) \\
&\times \phi(d\gamma\, d\mu \mid \gamma_0, \mu_0, x_0 + \mu_0 s) \\
&\times \exp\left\{-\int_0^s \lambda(x_0 + \mu_0\sigma)\, d\sigma\right\}\lambda(x_0 + \mu_0 s)\, ds.
\end{aligned} \tag{5.5}$$

If P satisfies (5.5) then one shows by an elementary calculation similar to that leading to the so-called "backward" equation (4.26) in I that P must satisfy the integro-differential equation

$$\begin{aligned}
&\mu_0 \frac{\partial}{\partial x_0} P(\gamma_0, \mu_0, x_0) = \lambda(x_0) \\
&\times \left\{P(\gamma_0, \mu_0, x_0) - \int_a P(\gamma, \mu, x_0)\phi(d\gamma\, d\mu \mid \gamma_0, \mu_0, x_0)\right\},
\end{aligned} \tag{5.6}$$

where we have suppressed the variables $A \times S_\tau$ in the notation for P, and where $\mu_0(\partial P/\partial x_0)$ is the derivative of P in the direction μ_0: i.e., $\mu_0(\partial P/\partial x_0) = \sum_1^3 \mu_0^{(i)}\partial P/\partial x_0^{(i)}$, $\mu_0^{(i)}$ and $x_0^{(i)}$ being the components of μ_0 and x_0, respectively, in some orthogonal reference frame for R_3.

The iteration relations which define the sequences $\{Q_n\}$, $\{P_n\}$, and hence, the regular solution P_r, take a simple form in the present case. Thus we have for the Q_n:

$$\begin{aligned}
&Q_{n+1}(A \times X; \tau \mid \gamma_0, \mu_0, x_0) \\
&= \int_a \int_0^{R_\tau(\mu_0, x_0)} Q_n(a \times X; \tau \mid \gamma, \mu, x_0 + \mu_0 s) \\
&\quad \times \phi(d\gamma\, d\mu \mid \gamma_0, \mu_0, x_0 + \mu_0 s) \\
&\quad \times \exp\left\{-\int_0^s \lambda(x_0 + \mu_0\sigma)\, d\sigma\right\}\lambda(x_0 + \mu_0 s)\, ds.
\end{aligned} \tag{5.7}$$

Using the fact that $P_n = P_0 * Q_n$, we find a similar iteration relation for the P_n:

$$P_{n+1}(A \times S_\tau \mid \gamma_0, \mu_0, x_0)$$
$$= \int_{\mathfrak{A}} \int_0^{R_\tau(\mu_0, x_0)} P_n(A \times S_\tau \mid \gamma, \mu, x_0 + \mu_0 s)$$
$$\times \phi(d\gamma\, d\mu \mid \gamma_0, \mu_0, x_0 + \mu_0 s)$$
$$\times \exp\left\{-\int_0^s \lambda(x_0 + \mu_0\sigma)\, d\sigma\right\}\lambda(x_0 + \mu_0 s)\, ds. \quad (5.8)$$

6. MULTIPLICATIVE AND CASCADE FIRST-PASSAGE PROCESSES

We now turn to the consideration of first-passage distributions in processes involving the creation and annihilation of particles as well as their scattering; we use the term *atomic event* as a generic name for all of these. We restrict ourselves here to processes involving only one type of particle (e.g., neutron multiplication, nucleon cascades neglecting meson production); the generalization to processes involving several types of particles (e.g., electron–photon cascades, nucleon cascades with meson production) is immediate and may be effected simply by considering the type of particle as an additional state variable. Suppose that the process starts with k particles (the "ancestors") in specified states and positions inside the surface τ, and let $\alpha^{(k)} = (\alpha_{(1)}, \cdots, \alpha_{(k)})$, $x_{\tau_0}^{(k)} = (x_{\tau_0(1)}, \cdots, x_{\tau_0(k)})$ stand, respectively, for these states and positions. As a result of multiplications and annihilations, after an infinite lapse of time n particles will emerge for the first time through the positions $x_\tau^{(n)} = (x_{\tau(1)}, \cdots, x_{\tau(n)})$ on τ and in the states $\alpha^{(n)} = (\alpha_{(1)}, \cdots, \alpha_{(n)})$, where n can be any integer 0, 1, 2, $\cdots$ and $n = 0$ means that no particle effects a first passage through τ; note that particles may have been created which are either annihilated or stop, or else go to infinity before passing through τ. The first-passage distributions we are interested in here are precisely those which yield the probability that n particles will effect a first-passage through specified subsets of τ in specified sets of states. Let $\Omega_\tau = \mathfrak{A} \times \tau$ be the set of all ordered pairs $\omega_\tau = (\alpha, x_\tau)$. In the present context, an elementary event $\boldsymbol{\omega}_\tau = \omega_\tau^{(\cdot)}$ is an ordered set $\omega_\tau^{(n)} = (\omega_{\tau(1)}, \cdots, \omega_{\tau(n)})$, and hence, the space of elementary events assigned to τ is $\boldsymbol{\Omega}_\tau = \bigcup_{n=0}^\infty \Omega_\tau^{(n)}$, where $\Omega_\tau^{(n)}$ is the n-fold Cartesian product of Ω_τ with itself and $\Omega_\tau^{(0)}$ corresponds to 0 particles. For a fixed initial $\omega_{\tau_0}^{(k)}$, the first-passage distribution $P(\cdot \mid \omega_{\tau_0}^{(k)})$ is then a probability measure on a suitably defined σ-field of subsets of $\boldsymbol{\Omega}_\tau$ normalized as in (2.1); we assume that P is symmetric: i.e., it is invariant under permutations of the coordinates $\omega_{\tau(1)}, \cdots, \omega_{\tau(n)}$ of $\omega_\tau^{(n)}$ for both the initial and the final states, which is equivalent to the assumption that the particles are indistinguishable. We are thus dealing with a *stochastic population process* (see Moyal[7] for the general theory of such processes). We also assume here that the process is *multiplicative* (or a *branching process*: c.f. Harris[8] and Moyal[9]; the term "*cascade process*" is used for a multiplicative process whose total energy is nonincreasing) in the sense that the k ancestors propagate independently of each other, so that the first-passage distribution $P(\cdot \mid \omega_{\tau_0}^{(1)})$ relative to a single ancestor will suffice to characterize the process. This condition may be expressed more precisely in terms of the *probability generating functional* (p.g.f. for short) G of P: let ζ be a bounded measurable function on Ω_τ and let $\boldsymbol{\zeta}$ be the measurable function on $\boldsymbol{\Omega}_\tau$ whose restriction to $\Omega_\tau^{(n)}$ is

$$\zeta^{(n)}(\omega_\tau^{(n)}) = \zeta(\omega_{\tau(1)}) \cdots \zeta(\omega_{\tau(n)});$$

then G is the expectation of $\boldsymbol{\zeta}$ relative to P: i.e.,

$$G[\zeta, \tau \mid \omega_{\tau_0}^{(k)}] = \mathcal{E}(\boldsymbol{\zeta} \mid \omega_{\tau_0}^{(k)}) = \int_{\boldsymbol{\Omega}_\tau} \boldsymbol{\zeta}(\boldsymbol{\omega}_\tau) P(d\boldsymbol{\omega}_\tau \mid \omega_{\tau_0}^{(k)})$$
$$= \sum_{n=0}^\infty \int_{\Omega_\tau^{(n)}} \zeta(\omega_{\tau(1)}) \cdots \zeta(\omega_{\tau(n)}) P^{(n)}(d\omega_\tau^{(n)} \mid \omega_{\tau_0}^{(k)}), \quad (6.1)$$

where $P^{(n)}$ is the restriction of P to $\Omega_\tau^{(n)}$. The process is multiplicative if and only if

$$G[\zeta, \tau \mid \omega_{\tau_0}^{(k)}] = \prod_{i=1}^k G[\zeta, \tau \mid \omega_{\tau_0(i)}]. \quad (6.2)$$

We can now extend to these *multiplicative first-passage processes* all of the considerations of Secs. 2–5. Using (6.2), the Chapman–Kolmogorov relation (2.2) becomes in terms of the p.g.f.:

$$G[\zeta, \tau \mid \omega_{\tau_0}]$$
$$= \sum_{n=0}^\infty \int_{\Omega^{(n)}} \prod_{i=1}^n G[\zeta, \tau \mid \omega_{\tau_1(i)}] P^{(n)}(d\omega_{\tau_1}^{(n)} \mid \omega_{\tau_0})$$
$$= G\{G[\zeta, \tau \mid \cdot] \mid \omega_{\tau_0}\}, \qquad \tau \geq \tau_1 \geq \tau_0. \quad (6.3)$$

Similarly, the integral equation (4.2) becomes in terms of p.g.f.

$$G[\zeta, \tau \mid \omega_{\tau_0}] = G_0[\zeta, \tau \mid \omega_{\tau_0}]$$
$$+ \sum_{n=0}^\infty \int_{\Omega^{(n)}} \prod_{i=1}^n G[\zeta, \tau \mid \omega_{\tau_1}(i)] Q^{(n)}(d\omega_{\tau_1}^{(n)} \mid \omega_{\tau_0}), \quad (6.4)$$

---

[7] J. E. Moyal, Acta Math. Stockh. **108**, 1 (1962).
[8] T. E. Harris, *The Theory of Branching Processes* (Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1964).
[9] J. E. Moyal, J. Appl. Probability **1**, 267 (1964).

where G_0 is the p.g.f. of the first-passage distribution P_0 with no atomic events and Q is the first atomic event position and consequent state distribution; for fixed ω_{τ_0}, $Q(\cdot \mid \omega_{\tau_0})$ is a probability distribution on a suitably defined σ-field of subsets of the space $\Omega = \bigcup_{n=0}^{\infty} (\mathcal{A} \times X)^{(n)}$ ($Q^{(0)}$ corresponds to annihilation, $Q^{(1)}$ to scattering and $Q^{(n)}$ with $n \geq 2$ to creation). The sequences $\{Q_n\}$, $\{P_n\}$ and the regular solution P_τ are defined as in Sec. 4; we write $G_n = G_0 * Q_n$ for the p.g.f. of P_n (the first-passage distribution with n atomic events); the p.g.f. of P_τ is then $G_\tau = \sum_{n=0}^{\infty} G_n$.

We generalize Sec. 5 in the same way. Let $\mathbf{A} = \bigcup_{n=0}^{\infty} \mathcal{A}^{(n)}$; then the transition probability $\phi(\cdot \mid \omega_\tau)$ conditional on an atomic event is a probability measure on a suitably defined σ-field of subsets of $\mathbf{A}$ with $\phi^{(0)}$ corresponding to annihilation, $\phi^{(1)}$ to scattering, and $\phi^{(n)}$, $n \geq 2$ to creation. The integral equation (5.5) takes the form

$$G[\omega_{\tau_0}] = G_0[\omega_{\tau_0}] + \sum_{n=0}^{\infty} \int_{\mathcal{A}^{(n)}} \int_0^{R(\mu_0, x_0)} \prod_{i=1}^{n} G[\gamma_{(i)}, \mu_{(i)}, x_0 + \mu_0 s] \times \phi^{(n)}(d\gamma^{(n)}\, d\mu^{(n)} \mid \gamma_0, \mu_0, x_0 + \mu_0 s) \times \exp\left\{-\int_0^s \lambda(x_0 + \mu_0 \sigma)\, d\sigma\right\} \lambda(x_0 + \mu_0 s)\, ds, \quad (6.5)$$

where for the sake of brevity we have suppressed the variables ζ, τ in G and G_0 and the variables γ_0, μ_0 in λ. The "backward" integro-differential equation (5.6) becomes

$$\mu_0 \frac{\partial}{\partial x_0} G[\gamma_0, \mu_0, x_0] = \lambda(x_0)\left\{G[\gamma_0, \mu_0, x_0] - \sum_{n=0}^{\infty} \int_{\mathcal{A}^{(n)}} \prod_{i=1}^{n} G[\gamma_{(i)}, \mu_{(i)}, x_0] \times \phi^{(n)}(d\gamma^{(n)}\, d\mu^{(n)} \mid \gamma_0, \mu_0, x_0)\right\}. \quad (6.6)$$

We obtain similar generalizations of the iteration relations (5.7), (5.8) and so on.

We remark that in the applications of the foregoing theory, the first-passage distribution with no atomic events P_0 will as a rule conserve the total number of particles [i.e., for a single "ancestor" $P_0(\cdot \mid \omega_{\tau_0}) \equiv P_0^{(1)}(\cdot \mid \omega_{\tau_0})$]. If, however, there is a nonvanishing probability $\eta_0(\tau \mid \omega_{\tau_0})$ of the particle stopping or escaping to infinity inside τ before the occurrence of the first atomic event, and if we wish the solution G of the integral equation (6.4) to yield the distribution of all particles that eventually effect a first-passage through τ, then we must assimilate η_0 to an annihilation probability and include it either in the expression for Q or in that for P_0: that is, we must either set $Q^{(0)} = q_0 + \eta_0$, where q_0 is the "true" annihilation probability, or we must set $P_0^{(0)} = \eta_0$, and hence, $G_0[\zeta, \tau \mid \omega_{\tau_0}] = \eta_0(\tau \mid \omega_{\tau_0}) + \int_{\Omega_\tau} \zeta(\omega_\tau) P_0^{(1)}(d\omega_\tau \mid \omega_{\tau_0})$. We also remark that in physical applications such as a neutron multiplication, where τ represents the boundary of the body where the multiplication occurs, the case where the probability of an infinity of atomic events $\theta_\infty \not\equiv 0$ will usually mean that the process is "supercritical," since it will usually imply an infinite outgoing flux of particles in the steady state for a constant source inside the body or a constant incoming flux (see, e.g., example 2 in Sec. 7).

7. EXAMPLES

Probably the simplest nontrivial examples one can construct to illustrate the foregoing theory are "one-dimensional" ones where the particles move on a line and the "surfaces" τ are the end points of intervals: thus we take the set of all surfaces T to be the set of all pairs of real numbers $\{a, b\}$ with $a < b$, where the "interior" of the surface $\{a, b\}$ is the closed interval $[a, b]$, and possibly in addition the set of all real numbers a with "interior" $(-\infty, a]$.

Example 1. The first and simplest example we consider is that of a particle moving with constant absolute velocity, so that the only state variable is the direction of motion μ, which can only take two values: $\mu = 1$ for motion to the right and $\mu = -1$ for motion to the left. We assume a constant mean free path λ^{-1} and a reversal of the direction of motion at each collision (see Brockwell and Moyal[10] for a more thorough treatment of this example). We may without loss of generality take the "surfaces" to be the set of all pairs $\{-a, a\}$, where $a > 0$. Let $P(\pm a \mid \mu, x)$ be the probability that the particle initially at x and moving in the direction μ makes a first passage through $\pm a$. Clearly, by symmetry

$$P(-a \mid \mu, x) = P(a \mid -\mu, -x), \quad (7.1)$$

so that one need only determine $P(a \mid \mu, x)$. The integral equation (5.5) and the corresponding "backward" equation (5.6) become in this case, respectively,

$$P(a \mid \mu, x) = e^{-\lambda(a - \mu x)} + \int_0^{a-\mu x} P(a \mid -\mu, x + \mu s) e^{-\lambda s} \lambda\, ds, \quad (7.2)$$

[10] P. J. Brockwell and J. E. Moyal, Nuovo Cimento 33, 776 (1965).

$$\mu \frac{\partial}{\partial x} P(a \mid \mu, x) = \lambda\{P(a \mid \mu, x) - P(a \mid -\mu, x)\}. \tag{7.3}$$

The solution of (7.1) is (c.f. Brockwell and Moyal,[10] p. 15):

$$P(a \mid \mu, x) = [\tfrac{1}{2}(1 + \mu) + \lambda(a + x)][1 + 2\lambda a]^{-1}. \tag{7.4}$$

It follows by (7.2) that $P(a \mid \mu, x) + P(-a \mid \mu, x) = P(a \mid \mu, x) + P(a \mid -\mu, -x) = 1$; hence, the process is stable and the solution (7.4) is unique. (7.4) may be obtained either via the iteration relation (5.8), which here takes the form

$$P_{n+1}(a \mid \mu, x) = \int_0^{a-\mu x} P_n(a \mid -\mu, x + \mu s) e^{-\lambda s} \lambda \, ds,$$

or by solving (7.3) with the boundary conditions $P(a \mid 1, a) = 1$ and $P(a \mid -1, -a) = 0$.

Example 2. We construct an example of a multiplicative first-passage process by modifying the previous example as follows: Instead of just a reversal at each collision, we assume that there is a probability q_{ij} that the particle splits into $i + j$ particles , with i moving in the same, j in the reverse direction and all with the same constant absolute velocity as the "parent" particle, such that

$$\sum_{i=0}^{\infty} \sum_{j=0}^{\infty} q_{ij} = 1$$

(note that q_{00} is the probability of annihilation of the "parent," q_{10} that of a continuation of and q_{01} of a reversal of direction without splitting). Let $P^{(n)}_{k,n-k}(\mu, x)$ be the probability conditional on a single initial "ancestor" at x and moving in the direction μ that eventually n particles will cross the boundaries a first time, with k moving to the right through a and $n - k$ moving to the left through $-a$. In this case the function ζ in the p.g.f. can take only two values: ζ_1 for a passage through a and ζ_{-1} for a passage through $-a$, and the expression (6.1) for the p.g.f. becomes

$$G[\zeta \mid \mu, x] = \sum_{n=0}^{\infty} \sum_{k=0}^{n} \zeta_1^k \zeta_{-1}^{n-k} P^{(n)}_{k,n-k}(\mu, x).$$

Let $g(\zeta_1, \zeta_{-1}) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \zeta_1^i \zeta_{-1}^j q_{ij}$; then the integral euqation (6.5) and the corresponding "backward" equation (6.6) become, respectively,

$$G[\zeta \mid \mu, x] = e^{-\lambda(a-\mu x)} + \int_0^{a-\mu x} \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} G^i[\zeta \mid \mu, x + \mu s] \times G^j[\zeta \mid -\mu, x + \mu s] e^{-\lambda s} \lambda \, ds = e^{-\lambda(a-\mu x)} + \int_0^{a-\mu x} g(G[\zeta \mid \mu, x + \mu s] \, G[\zeta \mid -\mu, x + \mu s]) e^{-\lambda s} \lambda \, ds, \tag{7.5}$$

$$\mu \frac{\partial}{\partial x} G[\zeta \mid \mu, x] = \lambda\{G[\zeta \mid \mu, x] - g(G[\zeta \mid \mu, x], G[\zeta \mid -\mu, x])\}. \tag{7.6}$$

If $q_{-1} = 1$, then Eq. (7.6) becomes (setting $\lambda = 1$, as we can do without loss in generality, since this merely amounts to taking the mean free path as the unit of length)

$$\mu \frac{\partial}{\partial x} G[\zeta \mid \mu, x] = G[\zeta \mid \mu, x]\{1 - G[\zeta \mid \mu, x] G[\zeta \mid -\mu, x]\}, \tag{7.7}$$

which can be solved explicitly (the solution is due to Brockwell). With the boundary conditions $G[\zeta \mid \mu, \mu a] = \zeta_\mu$, the solution is

$$G[\zeta \mid \mu, x] = \zeta_\mu \exp\{-[1 - \alpha(\zeta)](a - \mu x)\},$$

where for $0 \leq \zeta_\mu \leq 1$, $\alpha(\zeta)$ is the minimal nonnegative solution of the functional equation

$$\alpha(\zeta) \exp\{2a[1 - \alpha(\zeta)]\} = \zeta_1 \zeta_{-1}.$$

The probability of an infinity of atomic events is then

$$\theta_\infty(\mu, x) = 1 - \exp\{-(1 - \gamma)(a - \mu x)\},$$

where $\gamma = \alpha(1)$ is the smallest nonnegative solution of $\gamma \exp[2a(1 - \gamma)] = 1$; hence, $\gamma = 1$ and the process is stable (i.e., $\theta_\infty \equiv 0$) if and only if $2a \leq 1$; otherwise, $\gamma < 1$ and the process is unstable. We may therefore interpret $2a = 1$ as the "criticality" condition for this process and say that it is "subcritical" when $2a < 1$, "critical" when $2a = 1$, and "supercritical" when $2a > 1$. The total mean number of particles which effect a first passage through either a or $-a$ is (setting $\zeta_1 = \zeta_{-1} = z$)

$$m(\mu, x) = \left\{\frac{\partial}{\partial_z} G[z \mid \mu, x]\right\}_{z=1} = \left\{1 + \frac{2(a - \mu x)}{1 - 2a\gamma}\right\} \exp\{-(1 - \gamma)(a - \mu x)\}. \tag{7.8}$$

We see that this mean is infinite when the process is critical, or more precisely, that $m \to +\infty$ if $\mu x \neq a$ and $2a \to 1$ from the left. If we set $\gamma = 1$ in (7.8), then

$$m(\mu, x) = (1 - 2\mu x)/(1 - 2a), \tag{7.9}$$

which is the solution of the equation for the mean

$$\mu(\partial/\partial x)m(\mu, x) = -m(\mu, x) - m(-\mu, x) \quad (7.10)$$

with boundary conditions $m(\mu, \mu a) = 1$, for all values of a. The reason for the discrepancy between (7.8) and (7.9) is simply that Eq. (7.10) is obtained by setting $\zeta_1 = \zeta_{-1} = z$ in (7.7), differentiating both sides with respect to z and then setting $z = 1$ and $G[1 \mid \mu, x] \equiv 1$. It therefore ceases to be valid in the supercritical case where $G[1 \mid \mu, x] = \kappa_R(\mu, x) = 1 - \theta_\infty(\mu, x) \not\equiv 1$, and this is reflected by the fact that m in (7.9) can take negative values when $2a > 1$.

Example 3. We now exhibit an example of an unstable one-dimensional purely scattering process. We assume that the mean free path is v/α, where α is a constant and v is the absolute velocity of the particle, that at each collision there is a constant probability p that the direction of motion μ is reversed and $1 - p$ that it continues the same and that the absolute velocity u after a collision is uniformly distributed between 0 and the velocity v before the collision, independently of whether the direction of motion is reversed or continued. Let $P(u, \pm a \mid v, \mu, x)du$ be the probability that the particle, initially at x with velocity v and direction of motion μ, makes a first passage through $\pm a$ with velocity between u and $u + du$; clearly, $P = 0$ when $u > v$ and P satisfies (7.1). The integral and "backward" equations for this process are, respectively,

$$P(u, a \mid \sigma, \mu, x) = \exp\left[-\frac{\alpha}{v}(a - x)\right] + \int_0^{a-\mu x} \exp\left[-\frac{\alpha s}{v}\right]\frac{\alpha}{v}\,ds \times \int_u^v \{pP(u, a \mid w, -\mu, x + \mu s) + (1 - p)P(u, a \mid w, \mu, x + \mu s)\}\frac{dw}{v},$$

$$\mu\frac{\partial}{\partial x}P(u, a \mid v, \mu, x) = \frac{\alpha}{v}P(u, a \mid v, \mu, x) - \int_u^v \{pP(u, a \mid w, -\mu, x) + (1 - p)P(u, a \mid w, \mu, x)\}\frac{dw}{v}.$$

Let $\theta_\infty^{(p)}$ be the probability of an infinity of collisions for a given p. In the degenerate case $p = 0$, $\theta_\infty^{(0)}$ can be obtained explicitly (see I, p. 259):

$$\theta_\infty^{(0)}(v, \mu, x) = 1 - \left[1 + \frac{\alpha}{v}(a - \mu x)\right] \times \exp\left[-\frac{\alpha}{v}(a - \mu x)\right] \equiv 0.$$

It can be shown that $\theta_\infty^{(p)}(v, 1, x) \geq \theta_\infty^{(0)}(v, 1x)$ for $x \geq 0$ and $0 \leq p \leq 1$, which proves that this process is unstable for all p.

The collision rate per unit time of this process is the constant α, so that the probability of n collisions in a finite time interval t is $p_n(t) = (\alpha t)^n e^{-\alpha t}/n!$; hence, $\sum_0^\infty p_n(t) = 1$, which means that the probability of an infinite number of collisions is zero in any finite time interval. However, each collision slows the particle down and thereby decreases its mean free path, thus creating the possibility that it will not reach either boundary in any finite time. The probability that the particle is thus stopped is precisely θ_∞, and since $\lim_{t\to\infty} \sum_0^n p_i(t) = 0$ for all n, we see that θ_∞ is also the probability that the particle will suffer an infinite number of collisions.

Example 4. The Milne Problem. We now consider briefly the Milne problem from the point of view of the present paper; the treatment is very similar to that of Sobolev,[2] Chap. 6. We are concerned with a particle moving with constant velocity and constant mean free path (which we can take equal to unity) in R_3, and suffering isotropic scattering at each collision. We take T to see the set of all ordered pairs of planes normal to the x axis, and we can without loss of generality assume that they cut this axis at $\pm a$, where $a > 0$. Let θ be the angle between the direction of motion and the x axis; let $\mu = |\cos\theta|$ and $\sigma = \operatorname{sgn}(\cos\theta)$. We see by symmetry considerations that the first-passage probability density for the planes $\pm a$ depends only on the initial values σ, μ, k, the final values σ_a, μ_a, and on $r_a = [(y_a - y)^2 + (z_a - z)^2]^{\frac{1}{2}}$, where P refers to a first passage through the planes $\pm a$ according as $\sigma_a = \pm 1$, and y_a, z_a are the coordinate of the point of passage. Let

$$B(\sigma_a, \mu_a, r_a \mid x) = \frac{1}{2}\sum_{\sigma=-1}^{1}\int_0^1 P(\sigma_a, \mu_a, r_a \mid \sigma, \mu, x)d\mu$$

and let B_n be similarly related to P_n. The integral equation (5.5) becomes

$$P(\sigma_a, \mu_a, r_a \mid \sigma, \mu, x) = P_0(\sigma_a, \mu_a, r_a \mid \sigma, \mu, x) + \int_0^{(a-\sigma x)/\mu} B(\sigma_a, \mu_a, r_a \mid x + \sigma s)e^{-s}\,ds, \quad (7.11)$$

where

$$P_0(\sigma_a, \mu_a, r_a \mid \sigma, \mu, x) = \delta(\mu_a - \mu)\delta[r_a - (a - \sigma x)(1 - \mu^2)^{\frac{1}{2}}]e^{-(a-\sigma x)/\mu}.$$

It follows that B satisfies the integral equation

$$B(\sigma_a, \mu_a, r_a \mid x) = B_0(\sigma_a, \mu_a, r_a \mid x) + \int_{-a}^{a} B(\sigma_a, \mu_a, r_a \mid \xi)Ei(\xi - x)\, d\xi, \qquad (7.12)$$

where

$$Ei(x) = \int_0^1 e^{-|x|/\mu}\frac{d\mu}{\mu}.$$

It then follows from the iteration relation (5.8) that

$$B_{n+1}(\sigma_a, \mu_a, r_a \mid x) = \int_{-a}^{a} B_n(\sigma_a, \mu_a, r_a \mid \xi)Ei(\xi - x)\, d\xi,$$

and that $B = \sum_{n=0}^{\infty} B_n$ is the solution of (7.12). (It is easy to see that the series converges.) The first-passage distribution P is then obtained from B by (7.11).

ACKNOWLEDGEMENT

I am much indebted to P. J. Brockwell for many stimulating discussions and for his help and collaboration in connection with the one-dimensional examples discussed in Sec. 7.

Approximate Solution of Hill's Equation

HERBERT PAYNE
Aerospace Research Laboratories, Wright-Patterson Air Force Base, Ohio
(Received 1 February 1965)

A modified WKB approximation, amenable to successive corrections, to the solution of the linear differential equation of second order having a periodic coefficient in normal form is presented. Considered as an application of the related equation method, it uses the free-particle wave equation rather than, as in an alternative approach, Mathieu's equation. Particular attention is given the instances, where two simple turning points appear in the period and where there are no turning points. With respect to the one-dimensional crystal, it is shown how the energy band structure can be gleaned directly from the given periodic potential.

I. INTRODUCTION

HILL'S equation has been of importance in physics from its original application to celestial mechanics to its description of the motion of electrons in a one-dimensional crystal. In the latter context it appears as

$$(d^2\psi/dx^2) + p(E, x)\psi = 0, \tag{1}$$

where

$$p(E, x) = (2m/\hbar^2)[E - V(x)]$$

with

$$p(E, x + a) = p(E, x). \tag{2}$$

The additional condition that p be symmetric about some point x_1, i.e.,

$$p(E, -x + 2x_1) = p(E, x), \tag{3}$$

although not essential to the development below, is imposed (in keeping with one definition of Hill's equation) because it considerably simplifies the analysis and is ordinarily true in physical problems. Then,

$$p(E, -x + 2x_2) = p(E, x), \tag{4}$$

where

$$x_2 = x_1 + \tfrac{1}{2}a$$

follows from Eqs. (2) and (3).

According to Floquet's theorem, Eq. (1) with condition (2) yields solutions in the form

$$\psi = e^{ik(E)x}u(E, x), \tag{5}$$

where

$$u(E, x + a) = u(E, x).$$

For ψ to be bounded, as normally required, k must be real, whereupon it is termed the wavenumber. Ranges of values of E which render k real are called energy bands.

This paper endeavors to relate ψ and, in particular, k directly to the given function $p(E, x)$ to an approximation which can be successively improved upon. In this it extends the WKB method to the special case of the periodic coefficient. To do this, however, a function associated with the WKB method, the local momentum, is assigned the periodicity of $p(E, x)$. Then the effective momentum $\hbar k$ becomes simply the average local momentum. A formula for transferring from one expression for the local momentum to another is provided to enable one to employ existing iteration procedures for developing the local momentum, but not necessarily as a periodic function, from $p(E, x)$.

Two such procedures, applicable where a single simple turning point occurs in the interval (x_1, x_2) and where there are no turning points, are considered. The success of these procedures requires some restriction on the magnitudes of the derivatives of $V(x)$—in the latter case the stipulation

$$|p^{-\frac{1}{2}}\partial^2 p^{-\frac{1}{2}}/\partial x^2| \ll 1.$$

The occurrence of multiple turning points in the interval (x_1, x_2) or turning points of higher orders calls for treatments different from these.

Another means of serving the purpose of this paper has been discussed by Moriguchi.[1] Instead of the free-particle wave equation, which serves the WKB method as the related equation, it employs Mathieu's equation. It is believed, however, that the method presented here is simpler.

II. RELATION BETWEEN THE EFFECTIVE AND LOCAL MOMENTA

With respect to a function $z(E, x)$ governed by

$$\left(\frac{\partial z}{\partial x}\right)^2 + \frac{1}{2}\left(\frac{\partial z}{\partial x}\right)^{-1}\frac{\partial^3 z}{\partial x^3} - \frac{3}{4}\left[\left(\frac{\partial z}{\partial x}\right)^{-1}\frac{\partial^2 z}{\partial x^2}\right]^2 = p, \tag{6}$$

[1] H. Moriguchi, J. Phys. Soc. Japan 14, 1771 (1959).

the local momentum[2] may be defined as $\hbar\partial z/\partial x$. Since $\partial z/\partial x$ can satisfy arbitrary boundary conditions associated with this second-order differential equation, the local momentum is not thus uniquely specified.[3] If for a particular solution $\partial Z/\partial x$ these boundary conditions are chosen as

$$\partial^2 Z/\partial x^2 = 0 \quad \text{for} \quad x = x_1 \quad \text{and} \quad x = x_2, \tag{7}$$

it follows from Eqs. (2), (3), (4), and (6) that

$$\partial Z(E, x + a)/\partial x = \partial Z(E, x)/\partial x. \tag{8}$$

It may be verified that

$$\psi = (\partial z/\partial x)^{-\frac{1}{2}} e^{iz}$$

satisfies Eq. (1). A comparison of this and Eq. (5) upon setting $z = Z$ so as to utilize Eq. (8) reveals

$$-i \log [\psi(x + a)/\psi(x)] = ak(E) = Z(E, x + a) - Z(E, x) = \int_x^{x+a} dy\, \partial Z(E, y)/\partial y, \tag{9}$$

or from the above symmetry considerations

$$\tfrac{1}{2}ak(E) = Z(E, x_2) - Z(E, x_1), \tag{10}$$

i.e.,

$$\hbar k = (x_2 - x_1)^{-1} \int_{x_1}^{x_2} dx\, \hbar\, \partial Z/\partial x.$$

Also, one has

$$u(E, x) = (\partial Z/\partial x)^{-\frac{1}{2}} \times \exp\{i[(x + a)Z(E, x) - xZ(E, x + a)]/a\}.$$

A relation for the density of states derives from

$$\frac{\partial^2 z}{\partial E\, \partial x} + \frac{1}{4}\frac{\partial}{\partial x}\left[\left(\frac{\partial z}{\partial x}\right)^{-1} \frac{\partial^2}{\partial E\, \partial x} \log \frac{\partial z}{\partial x}\right] = \frac{m}{\hbar^2}\left(\frac{\partial z}{\partial x}\right)^{-1},$$

obtained through differentiating both sides of Eq. (6) with respect to E. Applying this to the result of a similar differentiation upon Eq. (9) and taking Eq. (8) into consideration give

$$dk/dE = (m/a\hbar^2) \int_{x'}^{x'+a} dx (\partial Z/\partial x)^{-1}, \tag{11}$$

which is consistent with Kramers' relation[4] for $2(d/dE) \cos ak$.

III. FORMULA FOR THE WAVENUMBER FROM THE LOCAL MOMENTUM IN GENERAL

Having found an arbitrary solution $\partial z/\partial x$ to Eq. (6), one may verify that

$$\partial Z/\partial x = (\sigma^2 - 1)^{\frac{1}{2}}[\sigma + \sin 2(z + \zeta)]^{-1}\, \partial z/\partial x \tag{12}$$

is also a solution. (The additional parameter makes this more general than the relation proposed by Ballinger and March[3] and discussed by Hecht and Mayer.[5]) The values of σ and ζ needed to satisfy conditions (7) are

$$\sigma = \frac{\sin 2(z_2 - z_1 + \eta_2 - \eta_1)}{[\sin^2 2\eta_1 + \sin^2 2\eta_2 - 2 \sin 2\eta_1 \sin 2\eta_2 \cos 2(z_2 - z_1 + \eta_2 - \eta_1)]^{\frac{1}{2}}} \tag{13}$$

and

$$\zeta = \frac{1}{2} \tan^{-1} \frac{\sin 2\eta_1 \cos 2(z_2 + \eta_2) - \sin 2\eta_2 \cos 2(z_1 + \eta_1)}{\sin 2\eta_1 \sin 2(z_2 + \eta_2) - \sin 2\eta_2 \sin 2(z_1 + \eta_1)}, \tag{14}$$

where

$$z_i = z(E, x_i)$$

and

$$2\eta_i = -\tan^{-1}[\tfrac{1}{2}\, \partial(\partial z/\partial x)^{-1}/\partial x] \quad \text{for} \quad x = x_i. \tag{15}$$

It will henceforth be assumed that z is real. Furthermore, inasmuch as Eq. (6) forbids $\partial z/\partial x$ to vanish (unless it does so for all x) and admits $-\partial z/\partial x$ as a solution, it is sufficient to specify

$$\partial z/\partial x > 0 \tag{16}$$

and

$$-\tfrac{1}{2}\pi < 2\eta_i < \tfrac{1}{2}\pi. \tag{17}$$

Accordingly, Eqs. (12), (13), and (14) permit

$$\partial Z/\partial x > 0 \quad \text{for} \quad |\sigma| > 1, \tag{18}$$

but not Z (and, therefore, k) to be real for $|\sigma| < 1$.

Determining the right side of Eq. (10) by integrating both sides of Eq. (12) over the interval (x_1, x_2) and applying Eqs. (13) and (14) yields

$$\tfrac{1}{2}ak = \tan^{-1}\left[\frac{\sin(z_2 - z_1)\sin(z_2 - z_1 + 2\eta_2 - 2\eta_1)}{\cos(z_2 - z_1 + 2\eta_2)\cos(z_2 - z_1 - 2\eta_1)}\right]^{\frac{1}{2}}. \tag{19}$$

[2] L. A. Young, Phys. Rev. **38**, 1612 (1931).
[3] R. A. Ballinger and N. H. March, Proc. Phys. Soc. (London) **A67**, 378 (1954).
[4] H. A. Kramers, Physica **2**, 483 (1935).
[5] C. E. Hecht and J. E. Mayer, Phys. Rev. **106**, 1156 (1957).

To depict the energy band structure consider

$$\lambda = (-1)^n[\tfrac{1}{2}(n - \tfrac{1}{2})\pi + z_1 - z_2 + \eta_1 - \eta_2], \quad (20)$$

where the integer n is determined by stipulating

$$-\tfrac{1}{4}\pi \leq \lambda \leq \tfrac{1}{4}\pi.$$

Relation (16) renders n positive. It thus serves as the band number. From Eqs. (19) and (20) one obtains

$$\tan^2 \tfrac{1}{2}ak = \frac{\cos 2(\eta_2 - \eta_1) + \sin 2\lambda}{\cos 2(\eta_2 + \eta_1) - \sin 2\lambda}. \quad (21)$$

With relation (17) in mind it is seen that for k to be real, i.e., for the right side of Eq. (21) to be nonnegative, requires

$$-\cos 2(\eta_2 - \eta_1) \leq \sin 2\lambda \leq \cos 2(\eta_2 + \eta_1). \quad (22)$$

Relations (11) and (18) make clear that

$$dk/dE > 0 \quad \text{for} \quad |\sigma| > 1.$$

From this one surmises that, if $k=0$ marks the lower limit of the first band, $\sin 2\lambda - \cos 2(\eta_2+\eta_1)=0$ and $\sin 2\lambda + \cos 2(\eta_2 - \eta_1) = 0$ represent the upper (lower) and lower (upper) limits, respectively, of the nth band when n is odd (even) and that these upper and lower limits correspond to $ak = n\pi$ and $ak = (n - 1)\pi$, respectively.

IV. APPROXIMATIONS TO THE ENERGY BAND STRUCTURE

A. Two Simple Turning Points per Period

If for a given E one turning point ξ of first order occurs in the interval (x_1, x_2), one may employ the iteration procedure developed by Hecht and Mayer[5] to determine z. No loss of generality arises from designating $p(E, x_1) < p(E, x_2)$, i.e., $p < 0$ for $x_1 \leq x < \xi$ and $p > 0$ for $\xi < x \leq x_2$.

Let

$$\Omega^{(\nu)}(E, x) = 1 + 2\int_x^{\xi^{(\nu)}} dy\,[-p^{(\nu)}(E, y)]^{\frac{1}{2}}$$

and

$$Q^{(\nu)}(E, x) = \tfrac{1}{4}\pi + \int_{\xi^{(\nu)}}^x dy\,[p^{(\nu)}(E, y)]^{\frac{1}{2}},$$

where

$$p^{(\nu)} = p + \gamma^{(\nu)} \quad \text{and} \quad p^{(\nu)}(E, \xi^{(\nu)}) = 0.$$

The nonnegative integer ν denotes the order of the approximation $\gamma^{(\nu)}$ of

$$\gamma = -\tfrac{1}{4}z^{-2}\left(\frac{\partial z}{\partial x}\right)^2 - \frac{1}{2}\left(\frac{\partial z}{\partial x}\right)^{-1}\frac{\partial^3 z}{\partial x^3} + \frac{3}{4}\left(\frac{\partial z}{\partial x}\right)^{-2}\left(\frac{\partial^2 z}{\partial x^2}\right)^2,$$

which evolves, after starting with $\gamma^{(0)} = 0$, by substituting in this equation the resulting successive expressions $z^{(\nu+1)}$. For a particular ν, the function $z^{(\nu+1)}$ is implicitly related to $\Omega^{(\nu)}$ for $x_1 \leq x < \xi^{(\nu)}$ and to $Q^{(\nu)}$ for $\xi^{(\nu)} < x \leq x_2$. In the ensuing discussion the superscripts are dropped.

Letting

$$\omega(E) = \Omega(E, x_1) = 1 + 2\int_{x_1}^{\xi} dx\,[-p(E, x)]^{\frac{1}{2}}$$

and

$$q(E) = Q(E, x_2) = \tfrac{1}{4}\pi + \int_{\xi}^{x_2} dx\,p^{\frac{1}{2}}(E, x), \quad (23)$$

one may employ the expansions

$$z_1 = e^{-\omega} + e^{-3\omega}g(\omega), \quad (24)$$

where

$$g(\omega) = 1 + 3e^{-2\omega} + (37/3)e^{-4\omega} + \cdots$$

and

$$z_2 = q[1 - \tfrac{1}{8}q^{-2} - (7/384)q^{-4} - (83/15360)q^{-6} + \cdots] \quad (25)$$

for sufficiently large ω and q. Equations (15) become

$$2\eta_1 = \tfrac{1}{2}\pi - 2e^{-\omega}[1 - (13/3)e^{-2\omega} + (31/5)e^{-4\omega} - (1949/21)e^{-6\omega} + \cdots] \quad (26)$$

and

$$2\eta_2 = -\tfrac{1}{8}q^{-3}[1 + \tfrac{5}{8}q^{-2} + (47/128)q^{-4} + \cdots]. \quad (27)$$

Designating the energy corresponding to the top of the nth band as $E^{(n)}_{\max}$, one infers from relations (20), (22), (23), (24), (25), (26), and (27) that

$$\int_{\xi}^{x_2} dx\,p^{\frac{1}{2}}(E^{(n)}_{\max}, x) = \tfrac{1}{2}(n - \tfrac{1}{2})\pi + e^{-\omega} + e^{-3\omega}g(\omega) + \tfrac{1}{8}q^{-1}f(q),$$

where

$$f(q) = 1 + [(31/48) - \tfrac{1}{2}(-1)^n]q^{-2} + [(683/1920) - \tfrac{5}{16}(-1)^n]q^{-4} + \cdots.$$

Similarly, for the energy $E^{(n)}_{\min}$ corresponding to the bottom of the band one has

$$\int_{\xi}^{x_2} dx\,p^{\frac{1}{2}}(E^{(n)}_{\min}, x) = \tfrac{1}{2}(n - \tfrac{1}{2})\pi - e^{-\omega} + (29/3)e^{-3\omega}h(\omega) + \tfrac{1}{8}q^{-1}f(q),$$

where

$$h(\omega) = 1 - (141/145)e^{-2\omega} + (4231/261)e^{-4\omega} + \cdots.$$

It may be noted that if $\omega \to \infty$ (making this a potential well problem as the energy bands narrow to become lines) these equations reduce to one involving the nth eigenvalue $E_0^{(n)}$,

$$\int_{\xi}^{x_2} dx\, p^{\frac{1}{2}}(E_0^{(n)}, x) = \tfrac{1}{2}(n - \tfrac{1}{2})\pi + \tfrac{1}{8}q^{-1}f(q),$$

which extends, for a symmetric potential, the WKB statement of the Bohr–Sommerfeld quantization rule.

B. No Turning Points

Since ψ is unbounded if $p < 0$ for all x, only the circumstance $p > 0$ is considered. This would seem to justify the well known expansion

$$\frac{\partial z}{\partial x} \sim p^{\frac{1}{2}}\left\{1 + \left[\frac{5}{32}p^{-3}\left(\frac{\partial p}{\partial x}\right)^2 - \frac{1}{8}p^{-2}\frac{\partial^2 p}{\partial x^2}\right] + \cdots\right\}.$$

Because this series exhibits the periodicity of p, it would serve as an expression for $\partial Z/\partial x$, yielding according to Eq. (10)

$$\tfrac{1}{2}ak \sim \int_{x_1}^{x_2} dx\, p^{\frac{1}{2}}\{1 + \cdots\}.$$

However, this relation fails to provide for the existence of bands, attributable to the above series' being only an asymptotic expansion.

An alternative approach stems from

$$\partial z/\partial x = p^{\frac{1}{2}}e^{\phi}. \tag{28}$$

To comply with Eq. (6) ϕ must satisfy the integral equation

$$\phi(E, x) = \alpha \cos\left[2\int_{x_1}^{x} dy\, p^{\frac{1}{2}}(E, y)\right] + \beta \sin\left[2 \times \int_{x_1}^{x} dy\, p^{\frac{1}{2}}(E, y)\right] + \int_{x_1}^{x} dy \sin\left[2\int_{y}^{x} dy'\, p^{\frac{1}{2}}(E, y')\right] \times [p^{-\frac{1}{2}}(E, y)\, \partial^2 p^{-\frac{1}{2}}(E, y)/\partial y^2 + r(E, y)], \tag{29}$$

where $r(E, x) = \frac{1}{4}p^{-\frac{1}{2}}(\partial\phi/\partial x)^2 + p^{\frac{1}{2}}(1 + 2\phi - e^{2\phi})$. The solution of this equation by the method of successive approximations, with $r = 0$ in the initial approximation, is suggested by the assumed smallness of ϕ and $\partial\phi/\partial x$. From Eqs. (15) and (28) and the implication of Eqs. (3) and (4) that $\partial p/\partial x = 0$ at $x = x_1$ and $x = x_2$ one obtains

$$2\eta_i = \tan^{-1}\tfrac{1}{2}(p^{-\frac{1}{2}}e^{-\phi}\,\partial\phi/\partial x) \quad \text{for} \quad x = x_i. \tag{30}$$

Setting $\beta = 0$ makes $2\eta_1 = 0$, whereupon, according to relations (20) and (22), k cannot be real if

$$\tfrac{1}{2}n\pi > z_2 - z_1 > \tfrac{1}{2}n\pi - 2\eta_2 \text{ for } \tan 2\eta_2 > 0 \tag{31a}$$

or

$$\tfrac{1}{2}n\pi - 2\eta_2 > z_2 - z_1 > \tfrac{1}{2}n\pi \text{ for } \tan 2\eta_2 < 0, \tag{31b}$$

which defines the forbidden region at the top of the nth band.

It might be supposed that α can be chosen so as to enable $\partial\phi/\partial x = 0$ at $x = x_2$, giving by Eq. (30) $2\eta_2 = 0$. However, with the aid of

$$z_2 - z_1 = \int_{x_1}^{x_2} dx\, p^{\frac{1}{2}}e^{\phi} \approx \int_{x_1}^{x_2} dx\, p^{\frac{1}{2}} \tag{32}$$

from Eq. (28) it is seen that, in the expression for $\partial\phi/\partial x$ at $x = x_2$ derived from Eq. (29), the coefficient of α vanishes about where conditions (31) prevail. For nearby values of E the resulting size of α would destroy the limitation on ϕ. Setting $\alpha = 0$ produces from Eqs. (29) and (30)

$$\tan 2\eta_2 = e^{-\phi(E, x_2)}\int_{x_1}^{x_2} dx \cos\left(2\int_{x}^{x_2} dy\, p^{\frac{1}{2}}\right) \times (p^{-\frac{1}{2}}\,\partial^2 p^{-\frac{1}{2}}/\partial x^2 + r) \approx \int_{x_1}^{x_2} dx \cos\left(2\int_{x}^{x_2} dy\, p^{\frac{1}{2}}\right)p^{-\frac{1}{2}}\,\partial^2 p^{-\frac{1}{2}}/\partial x^2. \tag{33}$$

It is instructive to compare the width ϵ_n of the forbidden region at the top of the nth band, i.e., $\epsilon_n = E_{\min}^{(n+1)} - E_{\max}^{(n)}$, obtained from relations (31), (32), and (33) with that from second-order perturbation theory for the case of nearly free electrons, i.e.,

$$V(x) = V_0 + v(x)$$

with

$$|v(x)| \ll E - V_0 = W(E).$$

The latter result is

$$\epsilon_n = (4/a)\left|\int_{x_1}^{x_2} dx\, v(x) \cos[2\pi n(x_2 - x)/a]\right|, \tag{34}$$

which agrees with the lowest-order approximation from the above relations.

The smallness of v (and its derivatives) justifies

$$p^{-\frac{1}{2}}\,\partial^2\, p^{-\frac{1}{2}}/\partial x^2 \approx \tfrac{1}{4}\hbar(2mW^3)^{-\frac{1}{2}}\, d^2\, v/dx^2$$

and

$$\int_{x}^{x_2} dy\, p^{\frac{1}{2}}(E, y) \approx (2mW)^{\frac{1}{2}}(x_2 - x)/\hbar. \tag{35}$$

Then, Eq. (33) becomes

$$\tan 2\eta_2 \approx \tfrac{1}{4}\hbar(2mW^3)^{-\frac{1}{2}} \times \int_{x_1}^{x_2} dx\,(d^2v/dx^2)\cos[2(2mW)^{\frac{1}{2}}(x_2 - x)/\hbar]$$
$$= (2W)^{-1}v(x_1)\sin[2(2mW)^{\frac{1}{2}}(x_2 - x_1)/\hbar] - (2m/W)^{\frac{1}{2}}\hbar^{-1}\int_{x_1}^{x_2} dx\,v(x)\cos[2(2mW)^{\frac{1}{2}}(x_2 - x)/\hbar]$$

upon noting that $dv/dx = 0$ at $x = x_1$ and $x = x_2$. When conditions (31) are satisfied one has according to relations (32) and (35)

$$\tfrac{1}{2}n\pi \approx z_2 - z_1 \approx (2mW)^{\frac{1}{2}}(x_2 - x_1)/\hbar = (mW/2)^{\frac{1}{2}}a/\hbar, \qquad (36)$$

yielding

$$\tan 2\eta_2 \approx -(2m/W)^{\frac{1}{2}}\hbar^{-1} \times \int_{x_1}^{x_2} dx\,v(x)\cos[2\pi n(x_2 - x)/a]. \qquad (37)$$

Conditions (31) also give

$$|2\eta_2| = [z_2(E_{\min}^{(n+1)}) - z_1(E_{\min}^{(n+1)})] - [z_2(E_{\max}^{(n)}) - z_1(E_{\max}^{(n)})],$$

or, after applying relation (36),

$$|\tan 2\eta_2| \approx |2\eta_2| \approx (m/2)^{\frac{1}{2}}a \times [W^{\frac{1}{2}}(E_{\min}^{(n+1)}) - W^{\frac{1}{2}}(E_{\max}^{(n)})]/\hbar \approx \tfrac{1}{4}(2m/W)^{\frac{1}{2}}\hbar^{-1}a\epsilon_n.$$

Eq. (34) follows from this and relation (37).

Geometrization of a Complex Scalar Field. I. Algebra

R. PENNEY
Scientific Laboratory, Ford Motor Company, Dearborn, Michigan
(Received 9 June 1965)

The necessary and sufficient conditions that a Ricci tensor should algebraically represent the energy tensor of a complex scalar field are given.

I. INTRODUCTION

IN the present analysis, we should like to extend the Rainich[1] procedure to the case of the complex massless scalar field. That is, we wish to exhibit purely geometric conditions on the Ricci tensor of a Riemann space, such that we may uniquely determine that such a space represents a classical, charged, massless "meson."

The philosphy, methods, and notations have been adequately treated previously,[2-4] and we refer the reader to the literature for details. We remark that our local metric will be (-11111), using $x^0=ct$.

II. SPECIFICATION OF THE PROBLEM

We consider a field specified by a Lagrangian density in a Riemann space, viz.,

$$L = g^{\beta\alpha}\varphi_{|\alpha}{}^{*}\varphi_{|\beta}. \tag{1}$$

The usual action principle leads, by variation of φ or φ^* to the field equations

$$\varphi_{|\alpha}{}^{|\alpha} = 0, \qquad {}^{*}\varphi_{|\alpha}{}^{|\alpha} = 0. \tag{2}$$

As a consequence of the field equations, the energy–momentum tensor

$$2T_{\alpha\beta} = \varphi_{|\alpha}\varphi^{*}_{|\beta} + \varphi^{*}_{|\alpha}\varphi_{|\beta} - g_{\alpha\beta}\varphi_{|\rho}{}^{*}\varphi^{|\rho} \tag{3}$$

has a vanishing divergence.

The usual field equations relating the geometry and the φ-field are

$$R_{\mu\nu} - \tfrac{1}{2}Rg_{\mu\nu} = T_{\mu\nu}, \tag{4}$$

as is well known.

Suppose now that we write $\varphi_{|\alpha}$ in terms of real vectors,

$$\varphi_{|\alpha} \equiv \mu_\alpha + i\omega_\alpha, \tag{5}$$

and write $R_{\mu\nu}$ in terms of μ_α, ω_β. We easily see than that

$$R_{\alpha\beta} = \mu_\alpha\mu_\beta + \omega_\alpha\omega_\beta. \tag{6}$$

The problem of geometrization, therefore is to specify algebraic and differential conditions on $R_{\mu\nu}$ such that $R_{\mu\nu}$ is of the form of the sum of vector products given by Eq. (6), where

$$\mu_{\alpha|\rho} - \mu_{\rho|\alpha} = 0, \tag{7}$$

$$\omega_{\alpha|\rho} - \omega_{\rho|\alpha} = 0, \tag{8}$$

$$\mu_\alpha{}^{|\alpha} = 0, \tag{9}$$

$$\omega_\beta{}^{|\beta} = 0. \tag{10}$$

In the present analysis, we shall solve the algebraic problem, deferring the differential problem to a later date. We wish, however, to remark that the present problem exhibits the features of the electromagnetic problem solved by Rainich,[1] in that $R_{\alpha\beta}$ is invariant under duality rotations[5] of the vectors μ_α, ω_β.

III. NECESSARY CONDITIONS

Let us define the following entities:

$$R_{\mu\nu} \equiv A_\mu A_\nu + B_\mu B_\nu, \tag{11}$$

$$x \equiv A_\alpha A^\alpha, \tag{12}$$

$$y \equiv B_\alpha B^\alpha, \tag{13}$$

$$z \equiv A_\alpha B^\alpha, \tag{14}$$

and note that the invariants x, y, z are expressible in terms of the lowest-order invariants of $R_{\mu\nu}$, viz.

$$R = g^{\mu\nu}R_{\mu\nu}, \tag{15}$$

$$R^{\alpha\beta}R_{\alpha\beta}, \tag{16}$$

$$R_{\alpha\beta}R^{\alpha\gamma}R^{\beta}_{\gamma}. \tag{17}$$

The mere fact that one must use the cubic invariant of $R_{\mu\nu}$ in order to express the scalars formed from A_α, B_β indicates that $R_{\mu\nu}$ must obey a cubic equation.

[1] G. Y. Rainich, Trans. Am. Math. Soc. 27, 106 (1925).
[2] J. A. Wheeler, *Geometrodynamics* (Academic Press Inc., New York, 1962), p. 225.
[3] *Gravitation—An Introduction to Current Research*, edited by L. Witten (John Wiley & Sons, Inc., New York, 1962).
[4] R. Penney, Phys. Letters 11, 228 (1964).
[5] R. Penney, J. Math. Phys. 5, 1431 (1964).

Indeed, with a little algebra, one easily finds that $R_{\mu\nu}$ obeys the indentity

$$R_{\alpha\beta}R^{\alpha\gamma}R_{\gamma\rho} - RR_{\lambda\beta}R^{\lambda}_{\rho} \equiv \tfrac{1}{2}R_{\beta\rho}[R^{\lambda\alpha}R_{\lambda\alpha} - R^2], \quad (18)$$

which is the primary necessary condition on $R_{\mu\nu}$.

In addition to the above condition, we know that $R_{\mu\nu}$ must obey certain inequalities in order to represent the complex scalar field.

First of all, since $R_{\mu\nu}$ is to represent a physical field, the energy density must be positive-definite. This condition is expressed by

$$G_{00} = R_{00} - \tfrac{1}{2}Rg_{00} > 0. \quad (19)$$

Next, since we want $R_{\mu\nu}$ to represent a *complex* scalar field, we must be sure that two vectors are necessary; hence we need

$$R^{\alpha\beta}R_{\alpha\beta} - R^2 \neq 0. \quad (20)$$

Lastly, since we are attempting to represent a physical energy tensor, we must have that the trace of the energy tensor be positive-definite,[6] or

$$G = -R > 0. \quad (21)$$

IV. SUFFICIENCY OF THE CONDITIONS

Let us suppose that we are given certain information concerning the Ricci tensor as follows:

$$R_{\alpha\beta}R^{\alpha\gamma}R_{\gamma\rho} - RR_{\lambda\beta}R^{\lambda}_{\rho} = \tfrac{1}{2}R_{\beta\rho}[R^{\lambda\alpha}R_{\lambda\alpha} - R^2], \quad (18)$$

$$R_{00} - \tfrac{1}{2}Rg_{00} > 0, \quad (19)$$

$$R^{\alpha\beta}R_{\alpha\beta} - R^2 \neq 0, \quad (20)$$

$$-R > 0. \quad (21)$$

Then, at a given point, we may suppose the metric to have its Euclidean values (-1111) and treat the algebraic problem posed. If we write out the ten equations, we readily find that we may take $R_{\mu\nu}$ as diagonal, at the given point.

Next, let us denote

$$r_0 = R_{00}, \quad r_1 = R_{11}, \quad r_2 = R_{22}, \quad r_3 = R_{33}; \quad (22)$$

$$2\lambda \equiv R^{\alpha\beta}R_{\alpha\beta} - R^2, \quad (23)$$

in terms of which our ten equations reduce to the following:

$$r_0^3 + Rr_0^2 = \lambda r_0, \quad (24)$$

$$r_1^3 - Rr_1^2 = \lambda r_1, \quad (25)$$

$$r_2^3 - Rr_2^2 = \lambda r_2, \quad (26)$$

$$r_3^3 - Rr_3^2 = \lambda r_3. \quad (27)$$

[6] See, e.g., L. D. Landau and E. Lifshitz, *The Classical Theory of Fields* (Addison-Wesley Publishing Company, 1959), 2nd ed., p. 89.

Further simplification, by addition and subtraction *alone*, leads to

$$r_0(r_1r_2 + r_1r_3 + r_2r_3) = 0, \quad (28)$$

$$r_1(r_0r_2 + r_0r_3 - r_2r_3) = 0, \quad (29)$$

$$r_2(r_0r_1 + r_0r_3 - r_1r_3) = 0, \quad (30)$$

$$r_3(r_1r_0 + r_0r_2 - r_1r_2) = 0. \quad (31)$$

Using these latter equations, we can easily show that at least one of the quantities r_0, r_1, r_2, r_3 vanishes, by assuming the contrary.

Knowing that at least *one* component vanishes, it is easy to see that at least one other must vanish, by direct solution.

Now, we must use the inequalities. First of all, we may not have that all four components vanish, obviously. Next, we note that if only r_0, say, were nonzero, we would violate Eq. (20).

Thus, we know at this point that a pair of the components vanish. We can readily deduce that r_0 cannot vanish. Suppose r_0 and r_1 vanished. Then we would have

$$G_{00} = \tfrac{1}{2}(r_2 + r_3) > 0, \quad (32)$$

$$G = -(r_2 + r_3) > 0, \quad (33)$$

and hence a contradiction.

We know, then, that one of the sets (r_1r_2), (r_2r_3), (r_1r_3) vanishes. However, we do not have any way of deciding which pair should vanish, and, indeed, the same result is obtained for any choice.

Suppose, arbitrarily, that (r_2r_3) vanish. Then, we note that $R_{\mu\nu}$ may be expressed as

$$R_{\mu\nu} = A_\mu A_\nu + B_\mu B_\nu, \quad (11)$$

where

$$A_\mu = (a, b, 0, 0), \quad (34)$$

$$B_\mu = (c, d, 0, 0), \quad (35)$$

provided that

$$a^2 + c^2 = r_0, \quad (36)$$

$$b^2 + d^2 = r_1, \quad (37)$$

$$ab + cd = 0. \quad (38)$$

Since the expression of $R_{\mu\nu}$ so obtained is covariant, and true in one reference frame, it is generally true.

Note that, if we wish, we may pick A_μ, B_μ to be orthogonal by taking

$$-ac + bd = 0, \quad (39)$$

corresponding to the fact that $R_{\mu\nu}$ is invariant under the transformation

$$A_\mu \to A'_\mu = A_\mu \cos\theta + B_\mu \sin\theta, \tag{40}$$

$$B_\mu \to B'_\mu = -A_\mu \sin\theta + B_\mu \cos\theta. \tag{41}$$

V. THE DEGENERATE CASE

If we relax the restriction of Eq. (20), but maintain the other conditions of Sec. IV, we readily find that only the component r_0 may be nonzero. In this highly degenerate case $(r_1 r_2 r_3)$ vanish, and $R_{\mu\nu}$ is expressible in terms of a single vector field.

Indeed, referring to Eqs. (34), (35), we may take B_μ to vanish, and A_μ to be of the form

$$A_\mu = (a, 000). \tag{34a}$$

We have not treated the degenerate case in concert with the general case because of the special problems which arise when one discusses the *differential* characterization of such a field.

VI. NULL FIELDS

If, in the degenerate case of the preceding section, one allows R to vanish, the resulting "complex scalar field" will be algebraically indistinguishable from a null electromagnetic field.

However, if we allow R to vanish, we are, in fact, outside the realm of physical applicability of the complex scalar field.[6] Indeed, the null degenerate case should be interpreted as a null electromagnetic field, in the opinion of the author. The vanishing trace of the energy–momentum tensor always implies a coalescence of the classical fields of physics within the scheme of geometrization.

Only by the differential properties may the null fields of electromagnetism, scalar fields, or, generally, massless wave fields be distinguished properly within the context of geometrization schemes.

VII. CONCLUSIONS

We have shown that the necessary and sufficient conditions for $R_{\mu\nu}$ to represent a nondegenerate complex scalar field are as follows:

$$R_{\alpha\beta}R^{\alpha\gamma}R_{\gamma\rho} - RR_{\lambda\beta}R^{\lambda}_{\rho} = \tfrac{1}{2}R_{\beta\rho}[R^{\lambda\alpha}R_{\lambda\alpha} - R^2], \tag{18}$$

$$R_{00} - \tfrac{1}{2}Rg_{00} > 0, \tag{19}$$

$$R^{\alpha\beta}R_{\alpha\beta} - R^2 \neq 0, \tag{20}$$

$$-R > 0. \tag{21}$$

In a future publication, we hope to treat the differential problem associated with the complex scalar field, which exhibits several interesting features analogous to the electromagnetic field problem.[2]

VIII. ACKNOWLEDGMENT

The author is indebted to an anonymous referee for calling attention to the degenerate and null cases treated in Secs. V and VI by a specific example of such fields.

Functional Differential Calculus of Operators. II

F. ROHRLICH AND M. WILNER
Department of Physics, Syracuse University, Syracuse, New York
(Received 13 August 1965)

The calculus of operators developed in a previous paper is here carried through in momentum space. The differential quotient, i.e., the derivative of an operator with respect to a free-field operator is expressed algebraically by means of generalized functions and certain commutators. No recourse is made to configuration space but the resultant calculus in p-space is related to the previously developed one in x-space by means of Fourier transforms. The calculus is presented for spins 0, $\frac{1}{2}$, and 1. The difficulties which were encountered before in the calculus for charged vector fields are resolved by working with a field equation which incorporates the supplementary condition needed to eliminate the spin-zero components.

I. INTRODUCTION

THE complexity of quantum field theory demands the development of a suitable and sufficiently powerful mathematical formalism. As was recognized some time ago,[1-3] the functional differential calculus appears to be part of such a formalism.

This calculus was first used as a c-number calculus and therefore permitted differentiation with respect to unquantized external fields only. Or, it was developed with respect to the "smearing functions" that are used to convert operator-valued distributions into operators. It led to difficulties when applied to anticommuting fields and resulted in such absurdities as "anticommuting c-numbers." As a consequence some authors suggested the complex formalism of hyperquantization as an alternative.[4]

The development of the asymptotic formulation of quantum field theory[5] clarified the important role of the free fields (in- and out-fields) and the subordinate role of the interpolating field.[6] It became clear that the only derivative with respect to operators that is required is that with respect to free-field operators. But these operators have commutation relations that are known for all space–time points and therefore offer an easy way for defining this calculus on an algebraic basis.[7] This reduction to an operator algebra seems to be essential to the successful construction of a derivative with respect to an operator. The inhomogeneous commutator, defined in I, thus formed the basis of the explicit definition of the operator derivative.

The completeness of the free fields permits an expansion of any operator in terms of them. Differentiation will then move certain space–time points off the mass shell. Since this process is of fundamental importance in the dynamics of the quantum field theory, it is clumsy and inefficient to have to refer to a space–time point x as on or off the mass shell. A momentum-space formulation is therefore demanded.

The importance of momentum space is also apparent from the physical representation of the initial and final states of the scattering matrix. These are obviously specified in p-space, referring to particles rather than to fields. While the customary formulations of quantum field theory exhibit an equivalence of x-space and p-space in the sense of Fourier transforms, it is not obvious that this equivalence is here to stay. For this reason we shall develop the fundamental operator calculus in p-space and shall remain independent of x-space. The equivalence will be seen to continue to hold, but it is not invoked in the logical development of the subject.

In Sec. II, the basic ideas are presented in detail for the scalar field. The Dirac field with the complication brought about by the anticommutation relations is discussed in Sec. III. The vector field forms the subject matter of Sec. IV where it is also shown how the difficulties can be overcome that prevented an explicit definition of the derivative in I. Section V is devoted to the massless vector field (electromagnetic field) which cannot easily be treated as a special case of vector field with mass.

[1] J. S. Schwinger, Proc. Natl. Acad. Sci. U. S. **37**, 452, 455 (1951).
[2] K. Symanzik, Z. Naturforsch. **10**, 809 (1954).
[3] N. N. Bogoliubov and D. C. Shirkov, *Introduction to the Theory of Quantized Fields*, English translation by G. M. Volkoff (Interscience Publishers, Inc., New York, 1959).
[4] J. M. Jauch, Helv. Phys. Acta **29**, 287 (1956). Further references to this formalism and to the difficulties that initiated it can be found in this paper.
[5] H. Lehmann, K. Symanzik, and W. Zimmermann, Nuovo Cimento **1** 205 (1955); *ibid.* **6**, 319 (1957).
[6] R. E. Pugh, Ann. Phys. (N.Y.) **23**, 335 (1963) and J. Math. Phys. **6**, 740 (1965).
[7] F. Rohrlich, J. Math. Phys. **5**, 324 (1964). In the following this last paper will be referred to as I.

The most important results are summarized in the last Sec. VI. The appendices present details whose inclusion in the main text would detract from the the logical development.

II. SCALAR FIELDS (REAL AND COMPLEX)

In configuration space (see I), a free-field operator is defined by the field equation

$$K(x)A(x) = 0, \qquad K(x) = \Box_x - m^2, \tag{II.1}$$

and by the commutation relation

$$[A(x), A^*(y)]_- = -i\,\Delta(x-y) = \int \frac{d^4p}{(2\pi)^3} e^{ip\cdot(x-y)}\epsilon(p)\,\delta(p^2+m^2). \tag{II.2}$$

In addition, if the field is real, $A^*(x) = A(x)$, and if the field is complex,

$$[A(x), A(y)] = 0. \tag{II.3}$$

If F is a functional of A (and A^*), the operator derivative of F is defined implicitly by

$$[A(x), F]_- = -i\int \Delta(x-y)\frac{\delta F}{\delta A^*(y)}\,d^4y \tag{II.4}$$

and

$$[A^*(x), F]_- = -i\int \Delta(x-y)\frac{\delta F}{\delta A(y)}\,d^4y,$$

and explicitly by

$$[A(x), F]^{\mathrm{I}}_- = -i\int \Delta_{\mathrm{I}}(x-y)\frac{\delta F}{\delta A^*(y)}\,d^4y, \tag{II.5}$$

$$[A^*(x), F]^{\mathrm{I}}_- = -i\int \Delta_{\mathrm{I}}(x-y)\frac{\delta F}{\delta A(y)}\,d^4y.$$

These last two equations can be solved to yield

$$i\,\delta F/\delta A^*(x) = K(x)[A(x), F]^{\mathrm{I}}_-, \qquad i\,\delta F/\delta A(x) = K(x)[A^*(x), F]^{\mathrm{I}}_-. \tag{II.6}$$

$[A(x), F]^{\mathrm{I}}_-$ is an inhomogeneous commutator, defined in I. $\Delta_{\mathrm{I}}(x)$ is any of the invariant functions (or any linear combination of them) that satisfies

$$K(x)\,\Delta_{\mathrm{I}}(x) = -\delta_4(x). \tag{II.7}$$

In momentum space we may define a free-field operator $A_m(p)$ by the mass-shell condition

$$(p^2+m^2)A_m(p) = 0, \tag{II.8}$$

which is satisfied by

$$A_m(p) = \delta(p^2+m^2)A(p), \tag{II.9}$$

$A(p)$ having no singularity at $p^2 = -m^2$, and by the commutation relation

$$[A_m(p), A^*_m(q)]_- = -i\,\Delta(p)\,\delta_4(p-q). \tag{II.10}$$

If the field A_m is to describe a neutral particle, we impose the condition

$$A^*_m(p) = A_m(-p) \tag{II.11}$$

and if the field is to describe a charged particle, then we add to (II.10) the commutation relation

$$[A_m(p), A_m(q)]_- = 0. \tag{II.12}$$

[It is convenient not to use Eqs. (II.11) and (II.12) too soon in the discussion. Thus, if no distinction is made, the equations we derive will apply to both neutral and charged cases. If it is necessary to specialize to either case, the appropriate equation will be imposed.]

The function $\Delta(p)$ in the commutation relation (II.10) is so far not determined. It is easily found, however, from the decomposition of $A_m(p)$ into the creation and annihilation operator a^* and a for particles and b^* and b for antiparticles which satisfy the standard defining relations

$$[a(\mathbf{p}), a^*(\mathbf{q})] = \omega\,\delta(\mathbf{p}-\mathbf{q}), \tag{II.13a}$$

$$[b(\mathbf{p}), b^*(\mathbf{q})]_- = \omega\,\delta(\mathbf{p}-\mathbf{q}), \tag{II.13b}$$

$$[a, a]_- = [b, b]_- = [a, b]_- = 0, \tag{II.13c}$$

$$[a, b^*]_- = 0, \tag{II.13d}$$

where $\omega = (\mathbf{p}^2+m^2)^{\frac{1}{2}}$. The decomposition of $A_m(p)$ is

$$A_m(p) = [(4\pi)^{\frac{1}{2}}/2\omega][a(\mathbf{p})\,\delta(p^0-\omega) + b^*(-\mathbf{p})\,\delta(p^0+\omega)], \tag{II.14}$$

i.e., it is linear and invariant under the simultaneous interchanges $a \leftrightharpoons b^*$, $\mathbf{p} \leftrightharpoons -\mathbf{p}$. This decomposition is assured by the invariant properties usually imposed on the quantum theory of free fields. In particular, (II.14) assures the validity of the substitution law in the perturbation solution and it lies at the basis of crossing symmetry.

Substitution of (II.13) and (II.14) into (II.10) yields

$$\Delta(p) = 2\pi i\epsilon(p)\,\delta(p^2+m^2), \tag{II.15}$$

i.e., $\Delta(p)$ is the momentum-space representation of the usual Δ-function. The constant in (II.14) was so chosen that (II.15) conforms with the usual notation.[8]

[8] The commutation relations (II.13) are consistent with (II.10) and (II.12) via (II.14). For neutral particles, when (II.12) is replaced by (II.11), $a = b$ and (II.13d) is to be omitted.

The operator $A(p)$ of (II.9) which is not restricted to the mass shell has the following mass-shell form as a consequence of (II.14):

$$A(\mathbf{p}, \omega) = (4\pi)^{\frac{1}{2}} a(\mathbf{p}), \tag{II.16a}$$

$$A(-\mathbf{p}, -\omega) = (4\pi)^{\frac{1}{2}} b^*(-\mathbf{p}). \tag{II.16b}$$

Off the mass shell we can write for timelike momenta

$$A(p) = (4\pi)^{\frac{1}{2}}[a(p)\theta(p) + b^*(-p)\theta(-p)], \tag{II.17}$$

with

$$a(p) = a(\mathbf{p}, p^0), \qquad a(\mathbf{p}, \omega) = a(\mathbf{p}), \qquad \text{etc.}$$

We conclude these considerations by the remark that the definition of $A(x)$,

$$A(x) \equiv \frac{1}{(2\pi)^2} \int d^4p\, e^{ip\cdot x} A_m(p), \tag{II.18}$$

recovers all the usual x-space relations from the above p-space equations.

After these preliminaries we are now ready to define the inhomogeneous commutator

$$[A_m(p), A_m^*(q)]_-^{\mathrm{I}} \equiv -i\,\Delta_{\mathrm{I}}(p)\,\delta_4(p-q) = -i(p^2 + m^2)_{\mathrm{I}}^{-1}\,\delta_4(p-q);$$

and

$$[A_m(p), A_m(q)]_-^{\mathrm{I}} = [A_m^*(p), A_m^*(q)]_-^{\mathrm{I}} = 0, \tag{II.19}$$

if charged particles are to be described.

The distribution $\Delta_{\mathrm{I}}(p)$ is defined as the solution of

$$(p^2 + m^2)\,\Delta_{\mathrm{I}}(p) = 1. \tag{II.20}$$

For example, $\Delta_{\mathrm{I}}(p) = \Delta_{\mathrm{R}}(p)$ is such a solution,

$$\Delta_{\mathrm{R}}(p) = P(p^2 + m^2)^{-1} + i\pi\epsilon(p)\,\delta(p^2 + m^2).$$

The symbol P denotes the Cauchy principle value in the p^0 integration over the test function.

For products of operators we have

$$[A_m(p), A_m^*(q_1)\cdots A_m^*(q_n)]_-^{\mathrm{I}} \equiv \sum_{i=1}^{n} [A_m(p), A_m^*(q_i)]_-^{\mathrm{I}} A_m^*(q_1)\cdots\Lambda_i\cdots A_m^*(q_n), \tag{II.21}$$

where Λ_i indicates that the factor $A_m^*(q_i)$ is missing. This relation is a slightly weaker defining equation for the inhomogeneous commutator than that given in I (I.7), though still sufficient for our purpose. The relation between these definitions and their internal consistency is discussed in Appendix A.

Now consider any functional $F(A, A^*)$ of the alternative forms

$$\begin{aligned} F &= \sum_{r,s=0}^{\infty} \int d^4p_1\cdots d^4p_r\, d^4q_1\cdots d^4q_s \\ &\quad \times f_{rs}(p_1\cdots p_r, q_1\cdots q_s) \\ &\quad \times A_m(p_1)\cdots A_m(p_r) A_m^*(q_1)\cdots A_m^*(q_s) \\ &= \sum_{r,s=0}^{\infty} \int d^4x_1\cdots d^4x_r\, d^4y_1\cdots d^4y_s \\ &\quad \times \tilde{f}_{rs}(x_1\cdots x_r, y_1\cdots y_s) \\ &\quad \times A(x_1)\cdots A(x_r) A^*(y_1)\cdots A^*(y_s). \end{aligned} \tag{II.22}$$

The inhomogeneous commutator $[A_m(p), F]_-^{\mathrm{I}}$ is well defined according to (II.21). We can therefore define the functional operator derivative by

$$-i\,\Delta_{\mathrm{I}}(p)\,\delta F/\delta A^*(p) \equiv [A_m(p), F]_-^{\mathrm{I}}, \tag{II.23a}$$

or equivalently by

$$i\,\delta F/\delta A^*(p) \equiv -(p^2 + m^2)[A_m(p), F]_-^{\mathrm{I}}. \tag{II.23b}$$

It is easily verified that the derivative in x-space is related to (II.23) by

$$\frac{\delta F}{\delta A^*(x)} = \frac{1}{(2\pi)^2} \int d^4p\, e^{ipx} \frac{\delta F}{\delta A^*(p)}, \tag{II.24}$$

provided the two alternative expansions (II.22) hold and the Fourier transform (II.18) exists.

The derivative (II.23b) is of course not restricted to the mass shell as is obvious when F is expressed as in (II.22). One has

$$\begin{aligned} i\frac{\delta F}{\delta A^*(p)} &= \sum_{\substack{r=0\\ s=1}}^{\infty} s \int d^4p_1\cdots d^4p_r\, d^4q_1\cdots d^4q_{s-1} \\ &\quad \times f_{rs}(p_1\cdots p_r, p, q_1\cdots q_{s-1}) \\ &\quad \times A_m(p_1)\cdots A_m(p_r) A_m^*(q_1)\cdots A_m^*(q_{s-1}). \end{aligned}$$

In a manner identical with that of the above discussion, we can define the inhomogeneous commutator

$$[A_m(q_1)\cdots A_m(q_n), A_m^*(p)]_-^{\mathrm{I}} \equiv \sum_{i=1}^{n} [A_m(q_i), A_m^*(p)]_-^{\mathrm{I}} A_m(q_1)\cdots\Lambda_i\cdots A_m(q_n). \tag{II.25}$$

Then for any functional $F(A, A^*)$ defined as in (II.22), the inhomogeneous commutator $[F, A_m^*(p)]_-^{\mathrm{I}}$ is well defined, and so is the functional derivative $\delta F/\delta A(p)$:

$$-i\,\Delta_{\mathrm{I}}(p)\,\delta F/\delta A(p) \equiv [F, A_m^*(p)]_-^{\mathrm{I}}, \tag{II.26a}$$

or

$$i\,\delta F/\delta A(p) \equiv -(p^2 + m^2)[F, A_m^*(p)]_-^{\mathrm{I}}. \tag{II.26b}$$

As before, the x-space results of I are recovered by

$$\frac{\delta F}{\delta A(x)} = \frac{1}{(2\pi)^2}\int d^4p\, e^{-ip\cdot x}\frac{\delta F}{\delta A(p)}. \tag{II.27}$$

The apparent sign contradiction between (II.26b) and (II.6) implied by (II.27) is resolved in Appendix A.

A more explicit check on the consistency of this argument is to consider the special cases $F = A_m(q)$ and $F = A_m^*(q)$. Then from (II.23) and (II.26) we have

$$\begin{aligned}\frac{\delta A_m^*(q)}{\delta A^*(p)} &= \delta(q^2+m^2)\frac{\delta A^*(q)}{\delta A^*(p)} = \delta_4(p-q)\\ &= \int \frac{d^4x\, d^4y}{(2\pi)^4} e^{-ip\cdot x+iq\cdot y}\,\delta_4(x-y)\\ &= \int \frac{d^4x\, d^4y}{(2\pi)^4} e^{-ip\cdot x+iq\cdot y}\frac{\delta A^*(y)}{\delta A^*(x)},\end{aligned}$$

and

$$\begin{aligned}\frac{\delta A_m(q)}{\delta A(p)} &= \delta(q^2+m^2)\frac{\delta A(q)}{\delta A(p)} = \delta_4(p-q)\\ &= \int \frac{d^4x\, d^4y}{(2\pi)^4} e^{ip\cdot x-iq\cdot y}\frac{\delta A(y)}{\delta A(x)}.\end{aligned} \tag{II.28}$$

If $A_m(p)$ represents charged particles, (II.19) gives us

$$\delta A_m(q)/\delta A^*(p) = \delta A_m^*(q)/\delta A(p) = 0. \tag{II.29}$$

Equations (II.21) and (II.25) can be used to derive various formulas of our differential calculus. From (II.22), we have

$$\begin{aligned}[A_m(p), \lambda F]_-^{\mathrm{I}} &= \lambda[A_m(p), F]_-^{\mathrm{I}}\\ [A_m(p), F+G]_-^{\mathrm{I}} &= [A_m(p), F]_-^{\mathrm{I}} + [A_m(p), G]_-^{\mathrm{I}}\\ [A_m(p), FG]_-^{\mathrm{I}} &= [A_m(p), F]_-^{\mathrm{I}}G + F[A_m(p), G]_-^{\mathrm{I}}.\end{aligned} \tag{II.30}$$

F and G are operator functionals of the form (II.22) and λ is not an operator. Similarly, from (II.25) we have

$$\begin{aligned}[\lambda F, A_m^*(p)]_-^{\mathrm{I}} &= \lambda[F, A_m^*(p)]_-^{\mathrm{I}}\\ [F+G, A_m^*(p)]_-^{\mathrm{I}} &= [F, A_m^*(p)]_-^{\mathrm{I}} + [G, A_m^*(p)]_-^{\mathrm{I}}\\ [FG, A_m^*(p)]_-^{\mathrm{I}} &= [F, A_m^*(p)]_-^{\mathrm{I}}G + F[G, A_m^*(p)]_-^{\mathrm{I}}.\end{aligned} \tag{II.31}$$

Finally, the rules of differentiation are

$$\begin{aligned}\frac{\delta\lambda F}{\delta A^*(p)} &= \lambda\frac{\delta F}{\delta A^*(p)}, \qquad \frac{\delta\lambda F}{\delta A(p)} = \lambda\frac{\delta F}{\delta A(p)},\\ \frac{\delta(F+G)}{\delta A^*(p)} &= \frac{\delta F}{\delta A^*(p)} + \frac{\delta G}{\delta A^*(p)},\\ \frac{\delta(F+G)}{\delta A(p)} &= \frac{\delta F}{\delta A(p)} + \frac{\delta G}{\delta A(p)},\\ \frac{\delta(FG)}{\delta A^*(p)} &= \frac{\delta F}{\delta A^*(p)}G + F\frac{\delta G}{\delta A^*(p)},\\ \frac{\delta(FG)}{\delta A(p)} &= \frac{\delta F}{\delta A(p)}G + F\frac{\delta G}{\delta A(p)}.\end{aligned} \tag{II.32}$$

The last two equations of (II.32) can also be written as

$$\left[\frac{\delta}{\delta A^*(p)}, F\right]_- = \frac{\delta F}{\delta A^*(p)}, \quad \left[\frac{\delta}{\delta A(p)}, F\right]_- = \frac{\delta F}{\delta A(p)}. \tag{II.33}$$

As a special case, we have

$$\left[\frac{\delta}{\delta A^*(p)}, A_m^*(q)\right]_- = \left[\frac{\delta}{\delta A(p)}, A_m(q)\right]_- = \delta_4(p-q).$$

In addition, from (II.11) and (II.19)

$$\left[\frac{\delta}{\delta A^*(p)}, A_m(q)\right]_- = \begin{cases} 0, & \text{for charged particles}\\ \delta_4(p+q), & \text{for neutral particles.}\end{cases} \tag{II.34}$$

The Jacobi indentities do not hold for inhomogeneous commutators, but one special case is valid as is proven in Appendix B, viz.

$$[A_m(q), [A_m(p), F]_-^{\mathrm{I}}]_-^{\mathrm{I}} = [A_m(p), [A_m(q), F]_-^{\mathrm{I}}]_-^{\mathrm{I}}, \tag{II.35a}$$

$$[[F, A_m^*(p)]_-^{\mathrm{I}}, A_m^*(q)]_-^{\mathrm{I}} = [[F, A_m^*(q)]_-^{\mathrm{I}}, A_m^*(p)]_-^{\mathrm{I}}, \tag{II.35b}$$

and

$$[A_m(p), [F, A_m^*(q)]_-^{\mathrm{I}}]_-^{\mathrm{I}} = [[A_m(p), F]_-^{\mathrm{I}}, A_m^*(q)]_-^{\mathrm{I}}. \tag{II.35c}$$

These relations are of importance for the double derivative; one finds

$$\begin{aligned}\left[\frac{\delta}{\delta A^*(p)}, \frac{\delta}{\delta A^*(q)}\right]_- F = 0,\\ \left[\frac{\delta}{\delta A(p)}, \frac{\delta}{\delta A(q)}\right]_- F = 0,\end{aligned} \tag{II.36}$$

and

$$\left[\frac{\delta}{\delta A(p)}, \frac{\delta}{\delta A^*(q)}\right]_- F = 0.$$

Using (II.33) and (II.36) we can prove a translation theorem similar to the one for configuration space operators, namely:

$$e^{\lambda\delta/\delta A(p)} F(A_m, A_m^*) e^{-\lambda\delta/\delta A(p)} = F(A_m(q) + 1\lambda\delta_4(p-q), A_m^*(q)), \quad \text{(II.37)}$$

$$e^{\lambda\delta/\delta A^*(p)} F(A_m, A_m^*) e^{-\lambda\delta/\delta A^*(p)} = F(A_m, A_m^*(q) + 1\lambda\delta_4(p-q)).$$

The proof follows the same line as in x-space (cf. I). However, it may be preferable to regard the right side of (II.37) simply as defined by these equations. In any case, this relation indicates how a connection can be established between our operator calculus and that defined by a limiting procedure [cf. I (IV.1)]. We do not find the latter definition satisfactory. This point will be taken up again in connection with the Dirac field.

III. DIRAC SPINOR FIELDS

The four-component spinor fields in momentum space are characterized by the mass-shell condition

$$(i\gamma\cdot p + m)\psi_m(p) = \bar{\psi}_m(p)(i\gamma\cdot p + m) = 0 \quad \text{(III.1)}$$

and the commutation relations

$$[\psi_m(p), \bar{\psi}_m(q)]_+ = iS(p)\,\delta_4(p-q) = 2\pi(m - i\gamma\cdot p)\varepsilon(p) \times \delta(p^2 + m^2)\,\delta_4(p-q), \quad \text{(III.2a)}$$

$$[\psi_m(p), \psi_m(q)]_+ = 0. \quad \text{(III.2b)}$$

The derivation of the second equality in (III.2a) proceeds exactly as for (II.15).

We can define an inhomogeneous commutator as follows. For two field operators

$$[\psi_m(p), \bar{\psi}_m(q)]_+^{\mathrm{I}} \equiv iS_{\mathrm{I}}(p)\,\delta_4(p-q) = [-i(m - i\gamma\cdot p)/(p^2 + m^2)_{\mathrm{I}}] \times \delta_4(p-q), = [-i/(m + i\gamma\cdot p)_{\mathrm{I}}]\,\delta_4(p-q), \quad \text{(III.3a)}$$

$$[\psi_m(p), \psi_m(q)]_+^{\mathrm{I}} = [\bar{\psi}_m(p), \bar{\psi}_m(q)]_+^{\mathrm{I}} = 0. \quad \text{(III.3b)}$$

The ordinary (anti) commutator of a product of field operators has the decomposition

$$[\psi_m(p), \bar{\psi}_m(q_1)\cdots\bar{\psi}_m(q_n)]_{-s(n)} = \sum_{i=1}^{n} (-1)^{i+1}[\psi_m(p), \bar{\psi}_m(q_i)]_+ \times \bar{\psi}_m(q)\cdots\Lambda_i\cdots\bar{\psi}_m(q_n), \quad \text{(III.4)}$$

where $s(n) = (-1)^n$. This states that on the left-hand side of (III.4), we take the commutator if n is even, and the anticommutator if n is odd. We therefore define the corresponding inhomogeneous (anti)commutator to be

$$[\psi_m(p), \bar{\psi}_m(q_1)\cdots\bar{\psi}_m(q_n)]_{-s(n)}^{\mathrm{I}} \equiv \sum_{i=1}^{n} (-1)^{i+1}[\psi_m(p), \bar{\psi}_m(q)]_+^{\mathrm{I}} \times \bar{\psi}_m(q_1)\cdots\Lambda_i\cdots\bar{\psi}_m(q_n), \quad \text{(III.5)}$$

and check the consistency of this definition in the same way as we do for Eq. (II.21) in Appendix A. [If any of the factors $\bar{\psi}_m(q_i)$ is really a $\psi_m(q_i)$, the anticommutator with $\psi_m(p)$ will give zero, but they must be included in the sum in order to assure the correct sign factor.]

If $F(\psi, \bar{\psi})$ is a functional of ψ and $\bar{\psi}$, for example of the form

$$F = \sum_{r,s=0}^{\infty} \int d^4p_1\cdots d^4p_r\, d^4q_1\cdots d^4q_s \times f_{rs}(p_1\cdots p_r q_1\cdots q_s) \times \bar{\psi}_m(p_1)\cdots\bar{\psi}_m(p_r)\psi_m(q_1)\cdots\psi_m(q_s), \quad \text{(III.6)}$$

and if every term in this representation of F has only an even or only an odd number of operator factors, that is, if F transforms according to some definite (spinor or tensor) representation of the Lorentz group, such as, for example, the S operator and its functional derivatives, then the operator $[\psi_m(p), F]_-^{\mathrm{I}}$, is well defined. The subscript $s = +1$ if F has an even number of spinor operator factors and $s = -1$ if F has an odd number of spinor operator factors. The operator derivative of F with respect to $\bar{\psi}$ is then defined (with spinor indices written explicitly) as

$$i\,\delta F/\delta\bar{\psi}_\alpha(p) \equiv -(m + i\gamma\cdot p)_{\alpha\lambda}[\psi_\lambda^m(p), F]_{-s}^{\mathrm{I}}. \quad \text{(III.7)}$$

As a special case, we have

$$[\delta/\delta\bar{\psi}_\lambda(p)](\phi_{\alpha_1}^m(q_1)\cdots\phi_{\alpha_n}^m(q_n)) = \sum_{i=1}^{n} (-1)^{i+1}\phi_{\alpha_1}^m(q_1)\cdots\frac{\delta\phi_{\alpha_i}^m(q_i)}{\delta\bar{\psi}_\lambda(p)}\cdots\phi_{\alpha_n}^m(q_n) \quad \text{(III.8)}$$

where $\phi_{\alpha_i}^m(q_i)$ is either $\bar{\psi}_{\alpha_i}^m(q_i)$ or $\psi_{\alpha_i}^m(q_i)$,

and

$$\delta\bar{\psi}^m_\alpha(q)/\delta\bar{\psi}_\lambda(p) = \delta_{\lambda\alpha}\,\delta_4(p-q),$$

and

$$\delta\psi^m_\alpha(q)/\delta\bar{\psi}_\lambda(p) = 0.$$

From (III.5) and (III.7) a number of rules of differentiation follow. Thus, we have immediately

$$\begin{aligned}
[\psi_m(p), \lambda F]^{\mathrm{I}}_{-s} &= \lambda[\psi_m(p), F]^{\mathrm{I}}_{-s}\\
[\psi_m(p), O_1 + O_2]^{\mathrm{I}}_{+} &= [\psi_m(p), O_1]^{\mathrm{I}}_{+} + [\psi_m(p), O_2]^{\mathrm{I}}_{+}\\
[\psi_m(p), E_1 + E_2]^{\mathrm{I}}_{-} &= [\psi_m(p), E_1]^{\mathrm{I}}_{-} + [\psi_m(p), E_2]^{\mathrm{I}}_{-}.
\end{aligned} \tag{III.9}$$

E_i and O_i are even and odd functionals, respectively. That is,

$$s(E) = +1, \qquad s(O) = -1.$$

Slightly less immediate are

$$\begin{aligned}
[\psi_m(p), OE]^{\mathrm{I}}_{+} &= [\psi_m(p), O]^{\mathrm{I}}_{+}E - O[\psi_m(p), E]^{\mathrm{I}}_{-},\\
[\psi_m(p), EO]^{\mathrm{I}}_{+} &= [\psi_m(p), E]^{\mathrm{I}}_{-}O + E[\psi_m(p), O]^{\mathrm{I}}_{+},\\
[\psi_m(p), O_1O_2]^{\mathrm{I}}_{-} &= [\psi_m(p), O_1]^{\mathrm{I}}_{+}O_2 - O_1[\psi_m(p), O_2]^{\mathrm{I}}_{+},\\
[\psi_m(p), E_1E_2]^{\mathrm{I}}_{-} &= [\psi_m(p), E_1]^{\mathrm{I}}_{-}E_2 + E_1[\psi_m(p), E_2]^{\mathrm{I}}_{-}.
\end{aligned} \tag{III.10}$$

The proof of the first of Eqs. (III.10) is as follows. Every term of OE is of the form

$$\phi_1 \cdots \phi_{2n+1}\tilde{\phi}_1 \cdots \tilde{\phi}_{2m},$$

where each ϕ and $\tilde{\phi}$ can be either ψ_m or $\bar{\psi}_m$. The ϕ subscript refers both to momentum and spinor index [$\phi_i = \phi^m_{\alpha_i}(q_i)$, and $\tilde{\phi}_i = \phi_{2n+1+i}$]. Then

$$\begin{aligned}
&[\psi, \phi_1 \cdots \phi_{2n+1}\tilde{\phi}_1 \cdots \phi_{2m}]^{\mathrm{I}}_{+}\\
&= \sum_{i=1}^{2n+1} (-1)^{i+1}[\psi, \phi_i]^{\mathrm{I}}_{+}\phi_1 \cdots \Lambda_i \cdots \phi_{2n+1}\tilde{\phi}_1 \cdots \tilde{\phi}_{2m}\\
&\quad + (-1)^{2n+1}\phi_1 \cdots \phi_{2n+1}\\
&\quad \times \sum_{i=1}^{2m} (-1)^{i+1}[\psi, \tilde{\phi}_i]^{\mathrm{I}}_{+}\tilde{\phi}_1 \cdots \Lambda_i \cdots \tilde{\phi}_{2m}\\
&= [\psi, \phi_1 \cdots \phi_{2n+1}]^{\mathrm{I}}_{+}\tilde{\phi}_1 \cdots \tilde{\phi}_{2m}\\
&\quad - \phi_1 \cdots \phi_{2n+1}[\psi, \tilde{\phi}_1 \cdots \tilde{\phi}_{2m}]^{\mathrm{I}}_{-}.
\end{aligned} \tag{III.11}$$

The other three equations of (III.10) are proved in the same way. From these commutation relations follow the linearity of the derivative and the rules of differentiation of products exactly as in I (III.13) to (III.15). In particular,

$$\begin{aligned}
\frac{\delta(OF)}{\delta\bar{\psi}} &= \frac{\delta O}{\delta\bar{\psi}}F - O\frac{\delta F}{\delta\bar{\psi}}\\
\text{and}\qquad \frac{\delta(EF)}{\delta\bar{\psi}} &= \frac{\delta E}{\delta\bar{\psi}}F + E\frac{\delta F}{\delta\bar{\psi}}.
\end{aligned} \tag{III.12}$$

The double derivative satisfies identities which are determined by Jacobi-type identities of the inhomogeneous commutator. As is proven in Appendix B,

$$[\psi_m(p), [\psi_m(q), F]^{\mathrm{I}}_{-s}]^{\mathrm{I}}_{+s} + [\psi_m(q), [\psi_m(p), F]^{\mathrm{I}}_{-s}]^{\mathrm{I}}_{+s} = 0. \tag{III.13}$$

Then, from the definition (III.7), follows immediately that

$$\left[\frac{\delta}{\delta\bar{\psi}(p)}, \frac{\delta}{\delta(\bar{\psi}q)}\right]_{+} F = 0. \tag{III.14}$$

Let us now return to (III.3), and obtain the operator derivative with respect to $\psi(p)$. Following the decomposition of the ordinary (anti)commutator

$$\begin{aligned}
&[\phi_1 \cdots \phi_n, \bar{\psi}_m(p)]_{-s(n)}\\
&= \sum_{i=1}^{n} (-1)^{n-i}[\phi_i, \bar{\psi}_m(p)]_{+}\phi_1 \cdots \Lambda_i \cdots \phi_n,
\end{aligned} \tag{III.15}$$

we define the inhomogeneous (anti)commutator

$$\begin{aligned}
&[\phi_1 \cdots \phi_n, \bar{\psi}_m(p)]^{\mathrm{I}}_{-s(n)}\\
&\equiv \sum_{i=1}^{n} (-1)^{n-i}[\phi_i, \bar{\psi}_m(p)]^{\mathrm{I}}_{+}\phi_1 \cdots \Lambda_i \cdots \phi_n.
\end{aligned} \tag{III.16}$$

Then $[F.\ \bar{\psi}_m(p)]^{\mathrm{I}}_{-s}$ is well defined for any functional of the form (III.6) that has a definite signature s, and the operator derivative of F with respect to $\psi(p)$ is defined as

$$i\,\delta F/\delta\psi_\alpha(p) = -[F, \bar{\psi}^m_\lambda(p)]^{\mathrm{I}}_{-s}(m + i\gamma\cdot p)_{\lambda\alpha}. \tag{III.17}$$

As a special case,

$$\begin{aligned}
&\frac{\delta}{\delta\psi_\lambda(p)}(\phi^m_{\alpha_1}(q_1) \cdots \phi^m_{\alpha_n}(q_n))\\
&\qquad = \sum_{i=1}^{n} (-1)^{n-i}\phi^m_{\alpha_1}(q_1) \cdots \frac{\delta\phi^m_{\alpha_i}(q_i)}{\delta\psi_\lambda(p)} \cdots \phi^m_{\alpha_n}(q_n),
\end{aligned}$$

where

$$\frac{\delta\phi^m_\alpha(q)}{\delta\psi_\lambda(p)} = \begin{cases} \delta_{\lambda\alpha}\,\delta_4(p-q), & \text{if } \phi^m_\alpha(q) = \psi^m_\alpha(q)\\ 0, & \text{if } \phi^m_\alpha(q) = \bar{\psi}^m_\alpha(q). \end{cases} \tag{III.18}$$

The following relations between inhomogeneous commutators are a consequence of (III.16):

$$\begin{aligned}
[\lambda F, \bar{\psi}_m(p)]^{\mathrm{I}}_{-s} &= \lambda[F, \bar{\psi}_m(p)]^{\mathrm{I}}_{-s},\\
[O_1 + O_2, \bar{\psi}_m(p)]^{\mathrm{I}}_{+} &= [O_1, \bar{\psi}_m(p)]^{\mathrm{I}}_{+} + [O_2, \bar{\psi}_m(p)]^{\mathrm{I}}_{+},\\
[E_1 + E_2, \bar{\psi}_m(p)]^{\mathrm{I}}_{-} &= [E_1, \bar{\psi}_m(p)]^{\mathrm{I}}_{-} + [E_2, \bar{\psi}_m(p)]^{\mathrm{I}}_{-},\\
[OE, \bar{\psi}_m(p)]^{\mathrm{I}}_{+} &= O[E, \bar{\psi}_m(p)]^{\mathrm{I}}_{-} + [O, \bar{\psi}_m(p)]^{\mathrm{I}}_{+}E,\\
[EO, \bar{\psi}_m(p)]^{\mathrm{I}}_{+} &= E[O, \bar{\psi}_m(p)]^{\mathrm{I}}_{+} - [E, \bar{\psi}_m(p)]^{\mathrm{I}}_{-}O,\\
[O_1O_2\ \bar{\psi}_m(p)]^{\mathrm{I}}_{-} &= O_1[O_2, \bar{\psi}_m(p)]^{\mathrm{I}}_{+} - [O_1, \bar{\psi}_m(p)]^{\mathrm{I}}_{+}O_2,\\
[E_1E_2, \bar{\psi}_m(p)]^{\mathrm{I}}_{-} &= E_1[E_2, \bar{\psi}_m(p)]^{\mathrm{I}}_{-} + [E_1, \bar{\psi}_m(p)]^{\mathrm{I}}_{-}E_2,
\end{aligned} \tag{III.19}$$

and from these one obtains the rules of differentiation given in I (III.22) to (III.24). In particular, we have

$$\frac{\delta(FE)}{\delta\psi} = F\frac{\delta E}{\delta\psi} + \frac{\delta F}{\delta\psi}E,$$
$$\frac{\delta(FO)}{\delta\psi} = F\frac{\delta O}{\delta\psi} - \frac{\delta F}{\delta\psi}O. \tag{III.20}$$

In addition to the above, we have from (III.16) the commutation relations analogous to (III.13),

$$[[F, \bar{\psi}_m(q)]^{\mathrm{I}}_{-s}, \bar{\psi}_m(p)]^{\mathrm{I}}_{+s} + [[F, \bar{\psi}_m(p)]^{\mathrm{I}}_{-s}, \bar{\psi}_m(q)]^{\mathrm{I}}_{+s} = 0, \tag{III.21}$$

and

$$[[\psi_m(p), F]^{\mathrm{I}}_{-s}, \bar{\psi}_m(q)]^{\mathrm{I}}_{+s} - [\psi_m(p), [F, \bar{\psi}_m(q)]^{\mathrm{I}}_{-s}]^{\mathrm{I}}_{+s} = 0. \tag{III.22}$$

These identities are also proven in Appendix B. They imply

$$\left[\frac{\delta}{\delta\psi(p)}, \frac{\delta}{\delta\psi(q)}\right]_+ F = 0, \tag{III.23}$$

and

$$\left[\frac{\delta}{\delta\bar{\psi}(p)}, \frac{\delta}{\delta\psi(q)}\right]_- F = 0. \tag{III.24}$$

We collect here a few special cases of Eqs. (III.12) and (III.20) which are of importance.

$$\begin{aligned}
[\delta/\delta\psi_\alpha(p), \psi^m_\beta(q)]_- G &= s(G)\,\delta_{\alpha\beta}\,\delta_4(p-q)G,\\
G[\bar{\psi}^m_\beta(q), \overleftarrow{\delta}/\delta\bar{\psi}_\alpha(p)]_- &= s(G)G\,\delta_{\alpha\beta}\,\delta_4(p-q),\\
[\delta/\delta\bar{\psi}_\alpha(p), \psi^m_\beta(q)]_+ G &= 0,\\
G[\bar{\psi}^m_\beta(q), \delta/\delta\psi_\alpha(p)]_+ &= 0,\\
[\delta/\delta\bar{\psi}_\alpha(p), \bar{\psi}^m_\beta(q)]_+ G &= \delta_{\alpha\beta}\,\delta_4(p-q)G,\\
G[\psi^m_\beta(q), \overleftarrow{\delta}/\delta\psi_\alpha(p)]_+ &= G\,\delta_{\alpha\beta}\,\delta_4(p-q),\\
G[\psi^m_\beta(q), \overleftarrow{\delta}/\delta\bar{\psi}_\alpha(p)]_- &= 0,\\
[\delta/\delta\psi_\alpha(p), \bar{\psi}^m_\beta(q)]_- &= 0.
\end{aligned} \tag{III.25}$$

Finally we make contact with x-space by observing that if

$$\psi(x) = \int \frac{d^4p}{(2\pi)^2} e^{ip\cdot x}\psi_m(p),$$
$$\bar{\psi}(x) = \int \frac{d^4p}{(2\pi)^2} e^{-ip\cdot x}\bar{\psi}_m(p), \tag{III.26}$$

then

$$\frac{\delta F}{\delta\psi(x)} = \int \frac{d^4p}{(2\pi)^2} e^{-ip\cdot x}\frac{\delta F}{\delta\psi(p)},$$
$$\frac{\delta F}{\delta\bar{\psi}(x)} = \int \frac{d^4p}{(2\pi)^2} e^{ip\cdot x}\frac{\delta F}{\delta\bar{\psi}(p)}, \tag{III.27}$$

and every formula in I involving derivatives with respect to $\psi(x)$ and $\bar{\psi}(x)$ is the Fourier transform of a corresponding formula in this chapter involving derivatives with respect to $\psi(p)$ and $\bar{\psi}(p)$. Note that in (III.27) the p-space integration is not restricted to the mass shell, which is why we left off the subscript "m" in $\delta F/\delta\psi(p)$.

IV. VECTOR FIELDS ($m \neq 0$)

As was pointed out in Ref. 1, an explicit definition of an operator derivative cannot be given if the inhomogeneous commutators are to satisfy a Klein–Gordon-type equation analogous to (II.7). It was indicated that this difficulty was related to the supplementary condition every vector field has to satisfy.

In the following we shall show that this difficulty can be overcome when the fundamental equations of the vector field are not a Klein–Gordon equation and a supplementary condition but one single field equation (which is of course equivalent to those two).

Consider the equation

$$[g_{\mu\nu}K(x) - \partial_\mu\partial_\nu]\phi^\nu(x) = 0 \tag{IV.1}$$

for the vector field $\phi_\mu(x)$. The divergence of this field equation yields

$$m^2\partial_\mu\phi^\mu(x) = 0 \tag{IV.2}$$

and therefore implies the supplementary condition, provided the mass m is assumed not to vanish. Both equations then combine to yield

$$K(x)\phi^\mu(x) = 0. \tag{IV.3}$$

Thus, (IV.2) and (IV.3) are implied in (IV.1). But it is also clear that (IV.1) is a consequence of the last two equations, so that a complete equivalence exists.

If one now defines the inhomogeneous commutator satisfying

$$[\phi_\mu(x), \phi^*_\nu(y)]^{\mathrm{I}}_- = -i(g_{\mu\nu} - \partial_\mu\partial_\nu/m^2)\Delta_{\mathrm{I}}(x-y), \tag{IV.4}$$

the operator which determines the field equation (IV.1) yields

$$(g^{\mu\alpha}K(x) - \partial^\mu\partial^\alpha)[\phi_\alpha(x), \phi^*_\nu(y)]^{\mathrm{I}}_- = i\,\delta_4(x-y)\,\delta^\mu_\nu, \tag{IV.5}$$

as follows from the operator relation

$$(g^{\mu\alpha}K - \partial^\mu\partial^\alpha)(g_{\alpha\nu} - \partial_\alpha\partial_\nu/m^2) = K\,\delta^\mu_\nu. \tag{IV.6}$$

Therefore, the implicit definition of the derivative can be made explicit by means of

$$[F, \phi^*_\mu(x)]^{\mathrm{I}}_{-}$$

$$= -i \int (g_{\mu\alpha} - \partial_\mu \partial_\alpha / m^2)\, \Delta_{\mathrm{I}}(x - y)\, \frac{\delta F}{\delta \phi_\alpha(y)}\, d^4y.$$

corresponding to (II.16) of Ref. 1. Using (IV.6) this yields the explicit form of the derivative[9]

$$i[\delta F / \delta \phi_\mu(x)] = (g^{\mu\alpha} K - \partial^\mu \partial^\alpha)[F, \phi^*_\alpha(x)]^{\mathrm{I}}_{-}. \qquad \text{(IV.7)}$$

We note that an immediate consequence of this relation is

$$i\, \partial_\mu[\delta F / \delta \phi_\mu(x)] = -m^2\, \partial^\mu [F, \phi^*_\mu(x)]^{\mathrm{I}}_{-}. \qquad \text{(IV.8)}$$

This expression does not vanish in general.

In momentum space we can start with the field equation

$$[g_{\mu\nu}(p^2 + m^2) - p_\mu p_\nu] \phi^\nu_m(p) = 0 \qquad \text{(IV.9)}$$

and the commutation relation

$$[\phi^m_\mu(p), \phi^{m*}_\nu(q)]_-$$

$$= 2\pi(g_{\mu\nu} + p_\mu p_\nu / m^2) \epsilon(p)\, \delta(p^2 + m^2)\, \delta_4(p - q)$$

$$= -i(g_{\mu\nu} + p_\mu p_\nu / m^2)\, \Delta(p)\, \delta_4(p - q). \qquad \text{(IV.10)}$$

If we put[10]

$$\phi^m_\mu(p) = \delta(p^2 + m^2)(g_{\mu\nu} + p_\mu p_\nu / m^2) \phi^\nu(p), \qquad \text{(IV.11)}$$

where $\phi_\mu(p)$ is free of singularities, then

$$p^\mu \phi^m_\mu(p) = 0, \qquad \text{(IV.12a)}$$

and

$$(p^2 + m^2) \phi^m_\mu(p) = 0. \qquad \text{(IV.12b)}$$

The last two equations are obviously equivalent to (IV.9).

The commutation relation (IV.10) reduces to the usual ones if one uses (IV.11) and

$$\phi_\mu(p)\,|_{p^0 = \omega} = (4\pi\omega)^{\frac{1}{2}} a_\mu(\mathbf{p}), \qquad \text{(IV.13a)}$$

$$\phi_\mu(-p)\,|_{p^0 = \omega} = (4\pi\omega)^{\frac{1}{2}} b^*_\mu(\mathbf{p}). \qquad \text{(IV.13b)}$$

For a charged vector field (IV.10) must be supplemented by

$$[\phi^m_\mu(p), \phi^m_\nu(q)]_- = 0, \qquad \text{(IV.14)}$$

for a neutral field by

$$\phi^*_\mu(p) = \phi_\mu(-p), \qquad \text{(IV.15)}$$

corresponding to $a_\mu(\mathbf{p}) = b_\mu(\mathbf{p})$ in (IV.13).

The inhomogeneous commutator follows from (IV.10) to be

$$[\phi^m_\mu(p), \phi^{m*}_\nu(q)]^{\mathrm{I}}_{-} = -i(g_{\mu\nu} + p_\mu p_\nu / m^2)\, \Delta_{\mathrm{I}}(p)\, \delta_4(p - q)$$

$$= -i\, \frac{g_{\mu\nu} + p_\mu p_\nu / m^2}{(p^2 + m^2)_{\mathrm{I}}}\, \delta_4(p - q). \qquad \text{(IV.16)}$$

From here on the argument proceeds in complete analogy to the scalar case.

The right side of (IV.16) is reduced to a δ-function by means of

$$[g_{\mu\alpha}(p^2 + m^2) - p_\mu p_\alpha][\delta^\alpha_\nu + p^\alpha p_\nu / m^2]\, \Delta_{\mathrm{I}}(p)$$

$$= (p^2 + m^2) g_{\mu\nu}\, \Delta_{\mathrm{I}}(p) = g_{\mu\nu}, \qquad \text{(IV.17)}$$

in analogy to (IV.6). Therefore,

$$[g_{\mu\alpha}(p^2 + m^2) - p_\mu p_\alpha][\phi^\alpha_m(p), \phi^{\nu*}_m(q)]^{\mathrm{I}}_{-} = -i\, \frac{\delta \phi^{\nu*}_m(q)}{\delta \phi^{\mu*}(p)}, \qquad \text{(IV.18)}$$

and generally

$$-i\, \delta F / \delta \phi^{\mu*}(p) = [g_{\mu\alpha}(p^2 + m^2) - p_\mu p_\alpha][\phi^\alpha_m(p), F]^{\mathrm{I}}_{-}, \qquad \text{(IV.19a)}$$

$$-i\, \delta F / \delta \phi^\mu(p) = [F, \phi^{\alpha*}_m(p)]^{\mathrm{I}}_{-} [g_{\alpha\mu}(p^2 + m^2) - p_\alpha p_\mu]. \qquad \text{(IV.19b)}$$

V. THE ELECTROMAGNETIC FIELD

The free electromagnetic field operator satisfies the zero-mass condition

$$p^2 A^{(0)}_\mu(p) = 0, \qquad \text{(V.1)}$$

the zero-charge condition

$$A^{(0)*}_\mu(p) = A^{(0)}_\mu(-p), \qquad \text{(V.2)}$$

and the commutation relation

$$[A^{(0)}_\mu(p), A^{(0)*}_\nu(q)] = -i g_{\mu\nu}\, D(p)\, \delta_4(p - q), \qquad \text{(V.3)}$$

$$D(p) = 2\pi i \epsilon(p)\, \delta(p^2).$$

If we define the inhomogeneous commutator as

$$[A^{(0)}_\mu(p), A^{(0)*}_\nu(q)]^{\mathrm{I}}_{-} \equiv -i g_{\mu\nu}\, D_{\mathrm{I}}(p)\, \delta_4(p - q), \qquad \text{(V.4)}$$

$$D_{\mathrm{I}}(p) = 1/(p^2)_{\mathrm{I}},$$

then

$$p^2 [A^{(0)}_\mu(p), A^{(0)*}_\nu(q)]^{\mathrm{I}}_{-} = -i g_{\mu\nu}\, \delta_4(p - q). \qquad \text{(V.5)}$$

[9] This result shows how the correspondence between non-renormalizable theories and fields which do not have derivatives can be broken. The essential point here seems to be the incorporation of the supplementary condition into the field equation. Furthermore, the Bogoliubov causality condition can now be applied also for vector mesons with mass, even if the source current is *not* conserved [F. Rohrlich, Bull. Am. Phys. Soc. 10, 87 (1965)].

[10] The functions $\phi_\mu(p)$ and $\phi_\mu(x)$ are of course different. No confusion can arise as is clear from the argument and the context.

If F is a functional of A_μ, expressible as a sum of products of the field operators, then the inhomogeneous commutator,

$$[A_\mu^{(0)}(p), F]_-^{\mathrm{I}},$$

is defined as in Sec. II, and the functional derivative of F is given by

$$i\,\delta F/\delta A_\mu(p) = -p^2[A_\mu^{(0)*}(p), F]_-^{\mathrm{I}}. \qquad \text{(V.6)}$$

Commutators of the functional derivative operator with itself and with A_μ are given by formulas identical in form with Eq. (II.13) of I, and the translation theorem stated in Sec. II for scalar fields can be generalized to vector fields.

This formulation assumes that the supplementary condition is not fulfilled as an operator equation but holds only in the subspace of physical state vectors in the well-known manner. If one uses another gauge, suitable definitions can also be given.

VI. CONCLUSIONS

The functional operator calculus can be developed in momentum space without reference to configuration space. Crucial for this development is the possibility of formulating the quantum theory of free fields entirely in momentum space. The operator calculus succeeds because it is based on algebraic relationships, in particular on the inhomogeneous commutator which can be defined entirely in p-space. The algebraic basis of this calculus can be avoided by a rather daring limiting process for the case of tensor fields [cf. (II. 37) and I (IV.1)]. But such a limiting process is *not* possible for the spinor case because of the underlying anticommutation relations. The cherished relation of classical mathematics which defines the derivative by a limit must here be abandoned.

Related to this fact is the breakdown of the "translation theorem" which is essentially a way of expressing a Taylor espansion. This matter is elaborated in Appendix C.

The operator calculus carries through without difficulties for spins 0, $\frac{1}{2}$, and 1 for either neutral or charged fields. For higher spins, the lesson learned for spin 1 will undoubtedly be crucial: the supplementary condition in the charged vector field case must be combined with the field equation into one single equation. Otherwise the operator derivative cannot be defined explicitly. It is precisely at this point where the renormalizability of the usual formulation of quantum field theory ceases to hold. But the operator calculus developed here is no longer limited in this way and therefore promises to be useful also for nonrenormalizable theories.

APPENDIX A

We want to discuss two problems of consistency:

(1) how the definitions of inhomogeneous commutators given by (II.21) and (II.25) can be generalized in a natural fashion to conform to the definition given in I (I.7), and

(2) what consistency requirement is imposed on (II.21) and on similar expressions by the connection with configuration space made with the Fourier transform, (II.18).

(1) First, we note that in (II.21) any operator $A_m^*(q_i)$ may be replaced by $A_m(q_i)$, and in (II.25), any $A_m(q_i)$ by $A_m^*(q_i)$, for if the field operators are to describe neutral particles, we impose (II.11), and if charged particles are to be described, we postulate that

$$[A_m(p), A_m(q)]_-^{\mathrm{I}} = [A_m^*(p), A_m^*(q)]_-^{\mathrm{I}} = 0. \qquad \text{(A1)}$$

In configuration space, among the known inhomogeneous commutators, the following relation holds:

$$[A^*(y), A(x)]_-^{\mathrm{I}} = [A(x), A^*(y)]_-^{\mathrm{I}'}. \qquad \text{(A2)}$$

The correct I' corresponding to a given I is obtained from the following table:

$I =$	R	A	P	$1R$	$1A$
$I' =$	A	R	P	$1R$	$1A$

; (A3)

(A2) and (A3) can be verified in each case using the definitions of the inhomogeneous commutators, the functions $\Delta_{\mathrm{I}}(x)$ given in I (I.9), and the fact that

$$\begin{aligned}[A(x), A^*(y)]_-^{1R,1A} &= -i\,\Delta_{1R,1A}(x-y) \\ &= \langle T_{+,-}(A(x)A^*(y))\rangle_0.\end{aligned} \qquad \text{(A4)}$$

$T_{+,-}$ signifies the positively (negatively) time-ordered product. We may take the momentum-space equation corresponding to (A2) as a definition:

$$[A_m^*(q), A_m(p)]_-^{\mathrm{I}} \equiv [A_m(p), A_m^*(q)]_-^{\mathrm{I}'}. \qquad \text{(A5)}$$

(A2) or (A5) enables us to define

$$[A^*, B_1 \cdots B_n]_-^{\mathrm{I}} \equiv \sum_{i=1}^{n} [A^*, B_i]_-^{\mathrm{I}} B_1 \cdots \Lambda_i \cdots B_n,$$

and

$$[B_1 \cdots B_n, A]_-^{\mathrm{I}} \equiv \sum_{i=1}^{n} [B_i, A]_-^{\mathrm{I}} B_1 \cdots \Lambda_i \cdots B_n. \qquad \text{(A6)}$$

Each B_i is a single operator, either an A or an A^*.

If $F(A, A^*)$ is a functional of the form (II.22), then

$$[A^*, F]^{\mathrm{I}}_{-} = [F, A^*]^{\mathrm{I}'}_{-}$$

and

$$[F, A]^{\mathrm{I}}_{-} = [A, F]^{\mathrm{I}'}_{-}. \tag{A7}$$

In this light, it is clear that (II.27) implies no sign contradiction between (II.6) and (II.26b), since any inhomogeneous commutator can be used in those equations, and, moreover,

$$[A^*, F]^{\mathrm{I}}_{-} \neq -[F, A^*]^{\mathrm{I}}_{-}. \tag{A8}$$

If $G(A, A^*)$ is also a functional of the form (II.22), the generalized inhomogeneous commutator $[F, G]^{I}_{-}$ is defined by extending the definitions (II.21), (II.25), and (A6) to

$$[B_1 \cdots B_n, F]^{\mathrm{I}}_{-} \equiv \sum_{i=1}^{n} B_1 \cdots B_{i-1}[B_i, F]^{\mathrm{I}}_{-}B_{i+1} \cdots B_n \tag{A9}$$

and

$$[F, B_1 \cdots B_n]^{\mathrm{I}}_{-} \equiv \sum_{i=1}^{n} B_1 \cdots B_{i-1}[F, B_i]^{\mathrm{I}}_{-}B_{i+1} \cdots B_n. \tag{A10}$$

From (A9) and (A10) the commutation relations in I (I.7) can be derived. Parenthetically, we emphasize that

$$[F, G]^{\mathrm{I}}_{-} \neq -[G, F]^{\mathrm{I}}_{-}. \tag{A11}$$

These considerations can be applied without difficulty to tensor fields of higher order. They need not be generalized to the case of spinor fields, however, since no such generalized inhomogeneous (anti)-commutator of spinor functionals, $[F, G]^{\mathrm{I}}_{\pm}$, is used in I.

(2) If we require that (II.21) (and similar expressions) be, according to (II.18), the Fourier transform of an analogous expression in configuration space,

$$[A(x), A^*(y_1) \cdots A^*(y_n)]^{\mathrm{I}}_{-} = \sum_{i=1}^{n} [A(x), A^*(y_i)]^{\mathrm{I}}_{-}A^*(y_1) \cdots \Lambda_i \cdots A^*(y_n), \tag{A12}$$

then (II.21) must satisfy a consistency condition which is the Fourier transform of the consistency condition which (A12) satisfies. For example, for equal times $y_i^0 = y^0$, $i = 1, \cdots, n$, (A12), considered as a retarded commutator, becomes an identity for ordinary commutators, multiplied on both sides by the step function $\theta(x - y)$. In momentum space, this condition would be

$$\begin{aligned}
&\int_{-\infty}^{\infty} dq_2^0 \cdots \int_{-\infty}^{\infty} dq_n^0 \\
&\times [A_m(p), A_m^*(\mathfrak{q}_1, q_1^0 - q_2^0)A_m^*(\mathfrak{q}_2, q_2^0 - q_3^0) \cdots A_m^*(q_n)]^{R}_{-} \\
&= \frac{1}{2\pi i}\int_{-\infty}^{\infty} \frac{d\lambda}{\lambda - i\epsilon} \int_{-\infty}^{\infty} dq_2^0 \cdots \int_{-\infty}^{\infty} dq_n^0 \\
&\times [A_m(\mathfrak{p}, p^0 + \lambda), A_m^*(\mathfrak{q}_1, q_1^0 - q_2^0 + \lambda) \cdots A_m^*(\mathfrak{q}_n, q_n^0 + \lambda)]_- \\
&= \sum_{i=1}^{n} \int_{-\infty}^{\infty} dq_2^0 \cdots \int_{-\infty}^{\infty} dq_n^0 [A_m(p), A_m^*(\mathfrak{q}_i, q_i^0 - q_{i+1}^0)]^{R}_{-} \\
&\times A_m^*(\mathfrak{q}_1, q_1^0 - q_2^0) \cdots \Lambda_i \cdots A_m^*(\mathfrak{q}_n, q_n^0), \\
&\qquad q_{n+1}^0 \equiv 0,
\end{aligned} \tag{A13}$$

which can be verified by using the relation

$$[A_m(p), A_m^*(q)]^{R}_{-} = \frac{1}{2\pi i}\int_{-\infty}^{\infty} \frac{d\lambda}{\lambda - i\epsilon} [A_m(\mathfrak{p}, p^0 + \lambda), A_m^*(\mathfrak{q}, q^0 + \lambda)]_-. \tag{A14}$$

(A14) follows from the identity

$$\Delta_R(p) = \frac{1}{2\pi i}\int_{-\infty}^{\infty} \frac{d\lambda}{\lambda - i\epsilon} \Delta(p, p^0 + \lambda), \tag{A15}$$

with $\Delta_R(p)$ and $\Delta(p)$ given in (II.20) and (II.15), respectively.

The consistency of the advanced commutator and the "P" commutator can be exhibited in an identical manner, with only a change in the contour of the λ integration. That of the $1R$ and $1A$ commutators is somewhat more complicated, but can be carried through in much the same way. Of course, if we do not require a connection with configuration space via the Fourier transform, then there would be no problem of consistency raised by (II.18), and there would be no need to impose (A13).

APPENDIX B

The following identities and their proofs have exactly the same form in x-space and in p-space. Since they have not been presented in I they are here given in x-space.

We first prove the following two identities:

$$[\phi_\mu(x), [\phi_\nu(y), F]^{\mathrm{I}}_{-}]^{\mathrm{I}}_{-} - [\phi_\nu(y), [\phi_\mu(x), F]^{\mathrm{I}}_{-}]^{\mathrm{I}}_{-} = 0, \tag{B1}$$

and

$$[\psi(x), [\psi(y), F]^{\mathrm{I}}_{-s}]^{\mathrm{I}}_{+s} + [\psi(y), [\psi(x), F]^{\mathrm{I}}_{-s}]^{\mathrm{I}}_{+s} = 0. \tag{B2}$$

The identity (B1) implies that there exists also an identity for the scalar field,

$$[A(x), [A(y), F]^{I}_{-}]^{I}_{-} - [A(y), [A(x), F]^{I}_{-}]^{I}_{-} = 0, \quad \text{(B3)}$$

which need not be proven separately.

If F is expanded as in (II.22) the linearity of (B1) in F implies that it is sufficient to prove (B1) when F is a product of n vector fields and that this be true for arbitrary n.

The proof of (B1) now uses induction. Clearly, for $F = \phi_\lambda(z)$ or $\phi^*_\lambda(z)$, Eq. (B1) is identically satisfied because the inhomogeneous commutator of two field operators is a c-number, and the inhomogeneous commutator of an operator and a c-number vanishes.

Assume now that (B1) holds when F is a product of n field operators. Let us use the symbolic notation x for $\phi_\mu(x)$, y for $\phi_\nu(y)$, etc. Then the first term in (B1) when F is a product of $n+1$ field operators becomes

$$\begin{aligned}
&[x, [y, z_1 \cdots z_{n+1}]^{I}_{-}]^{I}_{-} \\
&= [x, [y, z_1 \cdots z_n]^{I}_{-} z_{n+1}]^{I}_{-} + [x, z_1 \cdots z_n [y, z_{n+1}]^{I}_{-}]^{I}_{-} \\
&= [x, [y, z_1 \cdots z_n]^{I}_{-}]^{I}_{-} z_{n+1} + [y, z_1 \cdots z_n]^{I}_{-} [x, z_{n+1}]^{I}_{-} \\
&\quad + [x, z_1 \cdots z_n]^{I}_{-} [y, z_{n+1}]^{I}_{-} + z_1 \cdots z_n [x, [y, z_{n+1}]^{I}_{-}]^{I}_{-}.
\end{aligned} \quad \text{(B4)}$$

The last term vanishes because it involves an inhomogeneous commutator with a c-number. The second term of (B1) differs from the first only in the interchange $x \rightleftharpoons y$. Since, furthermore, (B1) is assumed to hold for $F = z_1 \cdots z_n$, the first term of (B4) will cancel out and we are left with

$$\begin{aligned}
&[x, [y, z_1 \cdots z_{n+1}]^{I}_{-}]^{I}_{-} - (x \rightleftharpoons y) \\
&\quad = [y, z_1 \cdots z_n]^{I}_{-} [x, z_{n+1}]^{I}_{-} \\
&\qquad + [x, z_1 \cdots z_n]^{I}_{-} [y, z_{n+1}]^{I}_{-} - (x \rightleftharpoons y) = 0.
\end{aligned}$$

Equation (B1) therefore does hold for $F = z_1 \cdots z_{n+1}$ if it holds for $F = z_1 \cdots z_n$. Tjis completes the proof by induction of (B1).

The proof for (B2) is completely analogous. It is obviously satisfied for $F = \psi(z)$ or $\bar{\psi}(z)$. Assume it is now also valid for $F = z_1 \cdots z_n$. We then make use of

$$[x, Gx']^{I}_{\mp s} = [x, G]^{I}_{\pm s} x' \mp s G [x, x']^{I}_{+}, \quad \text{(B5)}$$

where x stands for either ψ or $\bar{\psi}$ (permitting also mixed relations), and find

$$\begin{aligned}
&[x, [y, z_1 \cdots z_{n+1}]^{I}_{-s}]^{I}_{s} \\
&\quad = [x, [y, z_1 \cdots z_n]^{I}_{s} z_{n+1}]^{I}_{s} - s[x, z_1 \cdots z_n [y, z_{n+1}]^{I}_{+}]^{I}_{s} \\
&\quad = [x, [y, z_1 \cdots z_n]^{I}_{s}]^{I}_{-s} z_{n+1} + s[y, z_1 \cdots z_n]^{I}_{s} [x, z_{n+1}]^{I}_{+} \\
&\qquad - s[x, z_1 \cdots z_n]^{I}_{s} [y, z_{n+1}]^{I}_{+}.
\end{aligned}$$

A last term which vanishes identically was omitted. The first term will again cancel by assumption when substituted into (B2). Therefore,

$$\begin{aligned}
&[x, [y, z_1 \cdots z_{n+1}]^{I}_{-s}]^{I}_{s} + (x \rightleftharpoons y) \\
&= s\{[y, z_1 \cdots z_n]^{I}_{s} [x, z_{n+1}]^{I}_{+} - [x, z_1 \cdots z_n]^{I}_{s} [y, z_{n+1}]^{I}_{+}\} \\
&\qquad + (x \rightleftharpoons y) = 0.
\end{aligned}$$

In the two equations following (B5), $s = s(n+1) = (-1)^{n+1}$. Equation (B2) is therefore established.

It is obvious that (B1) continues to hold if one or both of the operators ϕ are replaced by their adjoints ϕ^*, since the Lorentz-transformation properties are thereby not affected and the inhomogeneous commutator of any two of them continues to be a c-number. Equations (II.35b) and (II.35c) are therefore also valid.

In Sec. III we also need the p-space analogy of the relations

$$[[F, \bar{\psi}(x)]^{I}_{-s}, \bar{\psi}(y)]^{I}_{s} + [[F, \bar{\psi}(y)]^{I}_{-s}, \bar{\psi}(x)]^{I}_{s} = 0 \quad \text{(B6)}$$

and

$$[[\psi(x), F]^{I}_{-s}, \bar{\psi}(y)]^{I}_{s} = [\psi(x), [F, \bar{\psi}(y)]^{I}_{-s}]^{I}_{s}. \quad \text{(B7)}$$

The proof of these relations proceeds in exactly the same way as for (B2) except that in addition to (B5) we need also

$$[Fx, x']^{I}_{\mp s} = F[x, x']^{I}_{+} - [F, x']^{I}_{\pm s} x. \quad \text{(B8)}$$

Again, x and x' can each stand for ψ or $\bar{\psi}$. Further details need not be given here.

Equations (B2), (B6), and (B7) are the Fourier transforms of the equations (III.13), (III.21), and (III.22) of the main text.

APPENDIX C

For c-number functionals we have the Taylor series[11]

$$\begin{aligned}
&F[A + \varphi] \\
&\quad = \sum_{n=0}^{\infty} \frac{1}{n!} \int \frac{\delta^n F[A]}{\delta A(y_1) \cdots \delta A(y_n)} \varphi(y_1) \cdots \varphi(y_n)(dy).
\end{aligned} \quad \text{(C1)}$$

[11] V. Volterra, *Theory of Functionals and of Integral and Integro-Differential Equations* (Dover Publications, Inc., New York, 1959).

If one assumes that this expansion continues to be valid when A is a (neutral scalar) operator field while φ remains an ordinary function (times the unit operator), we have in particular

$$\begin{aligned} F[A(y) + 1\lambda\ \delta(x-y)] &= \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} \int \frac{\delta^n F}{\delta A(y_1) \cdots \delta A(y_n)} \\ &\quad \times \delta(x-y_1) \cdots \delta(x-y_n)(dy) \\ &= \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} \frac{\delta^n F}{\delta A^n(x)} \\ &= \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} \left[\frac{\delta}{\delta A(x)}, F\right]_n \\ &= e^{\lambda\delta/\delta A(x)} F\, e^{-\lambda\delta/\delta A(x)}, \end{aligned} \tag{C2}$$

in the notation of I (I.15). Of course, such an assumption is rather questionable and must be regarded as purely formal. From (C2) follows

$$\lim_{\lambda\to 0} \frac{F[A(y) + 1\lambda\ \delta(x-y)] - F[A(y)]}{\lambda} = \frac{\delta F}{\delta A(x)}. \tag{C3}$$

This equation appears repeatedly in the literature as the definition of the operator derivative. As we have indicated in I, this definition does seem to lead to the same results as our algebraic definition, despite its questionable structure.

When one deals with spinor fields, however, the situation is quite different. First one finds by analogy as above

$$\begin{aligned} F[\psi(y) + \lambda\ \delta(x-y)1] &= \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} \int \frac{\delta^n F}{\delta\psi(y_1) \cdots \delta\psi(y_n)} \\ &\quad \times \delta(x-y_1) \cdots \delta(x-y_n)(dy), \end{aligned}$$

but it is easily seen that this equation is meaningless. The integrand is an antisymmetric function of the arguments $y_1, \cdots, y_n$ multiplied by a symmetric function, because

$$\delta^2 F/\delta\psi(x_1)\ \delta\psi(x_2) = -\delta^2 F/\delta\psi(x_2)\ \delta\psi(x_1). \tag{C4}$$

All terms with $n > 1$ on the right side vanish therefore identically. The translation theorem [last equality of (C2)] is thus not valid for spinors,[12] nor can the quantity $\delta\psi$ be defined. The limit definition of the operator derivative is therefore not tenable for spinors.

The above arguments can also be carried out in momentum space and lead to the same conclusion.

[12] In I, Eqs. (III.33) to (III.35), an erroneous statement to the contrary was made.

Model of an Oscillating Cosmos which Rejuvenates during Contraction

HELMUT SCHMIDT*
Boeing Scientific Research Laboratories, Seattle, Washington
(Received 25 August 1965)

For an oscillating cosmos we assume time-symmetric initial conditions, invariant under $t \to -t$ and $t \to T + t$, where $t = 0, \pm T, \cdots$ are the times of maximal cosmos contraction. In particular we specify the same value U_0 for the cosmos' "internal energy" U at all these times. The expectation value of an Heisenberg operator becomes now

$$\langle Q(t) \rangle = \mathrm{Tr}\ \{Q(t)\, D\}, \text{ with } D = [A(T), A(0)]_+/2\ \mathrm{Tr}\ \{A(T)\, A(0)\}.$$

Here $A(0)$ is the projector into the space of the eigenvectors of U with eigenvalue U_0, and $A(t)$ has the time dependence of an Heisenberg operator.

The time symmetry implies that the cosmos' oscillations are periodic and that expansions and contractions occur (except for local statistical fluctuations) time-symmetrically to each other: If the entropy increases during expansion then it must decrease during contraction.

The general theory is applied to a highly simplified cosmos model in which no star condensation occurs. Assuming here that the initial state contains equal numbers of particles and antiparticles, the theory predicts a slightly inhomogeneous distribution of these particles such that during the cosmos' expansion not all particles annihilate, but a realistic density of particles and antiparticles (in different space regions) survives the expansion.

INTRODUCTION

EINSTEIN's equations of general relativity lead to a small number of possible cosmos types. One of those is represented by a cosmos which oscillates between a very dense contracted state and an expanded state of very low density. Astronomers cannot yet decide to which type the actual cosmos belongs. For the following discussion, however, we will assume that we live in an oscillating world.

R. C. Tolman has studied the statistical thermodynamics of such a cosmos. He assumes the validity of the second law and comes to the conclusion that the oscillations of this cosmos are unstable. Due to the second law, the entropy increases from cycle to cycle and so does the amplitude of the oscillations.

This instability and the closely related time-unsymmetry of this model seem sufficiently objectionable to justify the study of other possibilities, in particular, the possibility of a violation of the second law:

It is well known that the second law is valid only for systems which satisfy certain initial conditions which, for laboratory situations, are overwhelmingly probable. Since, *a priori*, we know nothing about the cosmos' initial conditions, it seems reasonable to try several particularly plausible or mathematically attractive initial conditions and to see how the corresponding cosmos model would agree with reality.

The most accepted form of initial conditions may be obtained by the following picture: Imagine that the cosmos was created in the manner in which a sample is prepared in a laboratory. Assume, in particular, that at some time $t = 0$ where the cosmos was in a state of maximal contraction, the cosmos' macroscopic state was specified, leaving the microscopic state unspecified such that all microstates consistent with the macrostate A are equally probable. Then the second law would hold.

We will explore, however, another possibility. We assume that the cosmos was not made like a sample in the laboratory. We postulate, instead, that the cosmos' initial conditions are (like all microscopic laws of physics) invariant under time reversal and that, furthermore, they do not distinguish one particular cycle in the cosmos' history.

In order to specify the initial conditions mathematically, we assume that at the time $t = 0$ a macroscopic state A of the cosmos is given. Now, however, we consider only those microstates of the cosmos as admissible which are with the macrostate A consistent, not only at $t = 0$, but at all times $t = 0, \pm T, \pm 2T, \cdots$ of maximal cosmos contraction.

The history of this cosmos is (except for statistical fluctuations) invariant under the time reflection $t \to -t$ and under the time displacement $t \to t + T$. Here the expansions and contractions occur time-symmetrically to each other. Therefore, if the entropy increases during expansion, then it must decrease during the following contraction. A physicist living in the contracting stage of the cosmos, how-

* This work was begun at the University of Cologne/Germany.

ever, could not realize this. His "subjective" time direction would be the direction of increasing entropy and he would have the impression of living in an expanding cosmos.

I. THE PROBLEM OF AN ISOLATED OSCILLATING COSMOS

We want to study questions connected with the arrow of time in an oscillating cosmos of finite volume. In order to do this we consider the cosmos as an isolated finite quantum mechanical system.

This approach leads to many difficult unsolved problems. We first discuss how we may circumvent some of those problems such as to obtain a self-consistent cosmos model in which the arrow of time is the "only" difficult problem.

A. The Singularity of an Oscillating Cosmos During Passage Through Its State of Highest Density

We consider a spacially isotropic cosmos[1] with finite volume ($R_0^2 > 0$) and vanishing cosmological constant $\Lambda = 0$. Then the line element can be written as

$$ds^2 = c^2\,dt^2 - R(t)^2[1 + \tfrac{1}{4}r^2]^{-2}(dx^2 + dy^2 + dz^2), \tag{I1}$$

with

$$r^2 = x^2 + y^2 + z^1.$$

Here the physical (x, y, z) space equals the three-dimensional surface of a four-dimensional real sphere, i.e., we can write

$$ds^2 = c^2\,dt^2 - R(t)^2[dz_1^2 + dz_2^2 + dz_3^2 + dz_4^2], \tag{I2}$$

with

$$z_1^2 + z_2^2 + z_3^2 + z_4^2 = 1.$$

The total volume of this cosmos is

$$V(t) = 2\pi^2 R(t)^3. \tag{I3}$$

Now one can derive the following relation between the "world radius" $R(t)$, the energy density $u(t)$, and the pressure $p(t)$ of the cosmos:

$$(d/dt)(uR^3) + p(d/dt)(R^3) = 0, \tag{I4}$$

$$(dR/dt)^2 = \tfrac{1}{3}\kappa c^2 uR^2 - c^2, \tag{I5}$$

with

$$\kappa = 2.073\ 10^{-48}\ \ \mathrm{sec}^2/\mathrm{g\ cm}.$$

If we assume nonnegative values for u and p, then (I4) and (I5) imply, as is easily seen:

(a) An expanding cosmos cannot expand forever. After reaching some $R = R_{\max}$, the cosmos starts contracting.

(b) This contraction continues until $R \to 0$ and $u \to \infty$, i.e., the cosmos reaches an unphysical, singular state of infinite energy density.

Lifshitz and Khalatnikov[2] have pointed out that this singularity might disappear if the cosmos is not homogeneously filled with matter and radiation. In order to have a simple model, however, we neglect inhomogeneities and compensate the singularity by the following device: We calculate "as if" the cosmos would contain some compensating medium which interacts with the rest of the cosmos only via gravitation. Let us assume for example that pressure and energy density of this medium are given by

$$\left.\begin{aligned} p' = u' = 0 \quad & \text{for } R > R_{\min}, \\ p' = -A \quad & \\ u' = +A(1 - R_{\min}^3/R^3) \quad & \text{for } R < R_{\min}; \end{aligned}\right\} \tag{I6}$$

here A is a very large positive pressure, $A \to \infty$, and $R_{\min}$ is a finite constant, the "minimal world radius." Note that (I4) is satisfied by $u = u'$ and $p = p'$.

In (I5) the conventional energy density u has now to be replaced by the total energy density $u + u'$:

$$(dR/dt)^2 = \tfrac{1}{3}\kappa c^2(u + u')R^2 - c^2. \tag{I5'}$$

If now in a contracting cosmos $R(t)$ falls below $R_{\min}$ then the compensating medium comes into play: If A is large enough then $\dot{R}$ will go to zero for some value of R slightly below $R_{\min}$ and the contraction goes over into expansion. In the limit $A \to \infty$ the effect of the compensating medium is just to change the sign of $\dot{R}$, i.e., to convert the contraction into an expansion as soon as R reaches $R_{\min}$.

We can therefore use (I5) instead of (I5') if we supplement (I5) by the boundary condition

$$\dot{R}(t) \text{ changes sign if } R \text{ reaches the value } R_{\min}. \tag{I7}$$

Let us consider briefly the example of a light cosmos: From (I4) we obtain with $u = 3p$

$$u = u_0(R_{\min}/R)^4, \tag{I8}$$

where u_0 is the energy density in the state of maximal contraction. Now (I5) becomes

$$(dR/dt)^2 = (\alpha^2/R^2) - c^2 \text{ with } \alpha^2 = \tfrac{1}{3}\kappa c^2 u_0 R_{\min}^4. \tag{I9}$$

[1] R. C. Tolman, *Relativity, Thermodynamics and Cosmology* (Oxford University Press, New York, 1946).

[2] E. M. Lifshitz and I. M. Khalatnikov, Advan. Phys. 12, 185 (1963).

The radius at maximal expansion is

$$R_{\max} = \alpha/c. \tag{I10}$$

The solution of (I9) with the boundary condition (I7) is given by

$$R(t)^2 = R_{\max}^2 - c^2(t - \tfrac{1}{2}T)^2 \text{ for } 0 \leq t \leq T, \tag{I11a}$$

$$R(t) = R(t + T) = R(-t) \text{ for all } t. \tag{I11b}$$

Here T, the time for one cosmos oscillation, follows from the requirement that $R(0)^2 = R_{\min}^2$, i.e.,

$$R_{\min}^2 = R_{\max}^2 - \tfrac{1}{4}c^2T^2. \tag{I12}$$

We see that the cosmos oscillates periodically between states of high and low energy density and that these oscillations distinguish no time direction.

We will be particularly interested in the states of fairly high energy density, for example in the region $|t| \ll T$. Assuming that

$$R_{\min} \ll R_{\max}, \tag{I13}$$

we obtain from (I11), (I12),

$$R(t)^2 = R_{\min}^2[(t_0 + |t|)/t_0] \quad \text{for} \quad |t| \ll T, \tag{I14}$$

with

$$t_0^{-2} = \tfrac{4}{3}\kappa c^2 u_0. \tag{I15}$$

For the energy density u, (I14) and (I8) give

$$\tfrac{4}{3}\kappa c^2 u(t) = (|t| + t_0)^{-2}. \tag{I16}$$

Assuming finally black-body radiation where energy density u and temperature ϑ are related by

$$u = \frac{\pi^2}{15\hbar^3 c^3}(k\vartheta)^4, \tag{I17}$$

we obtain for the temperature

$$k\vartheta = \lambda(|t| + t_0)^{-\frac{1}{2}} \tag{I18}$$

with

$$\lambda^4 = \frac{45}{4\pi^2}\frac{\hbar^3 c}{\kappa}. \tag{I19}$$

(The parameter t_0 is related to the initial temperature ϑ_0 at $t = 0$ by $k\vartheta_0 = \lambda t_0^{-\frac{1}{2}}$. For $k\vartheta_0 = Mc^2$ with $M = 2\cdot 10^{-24}$ g, we would obtain $t_0 = 10^{-6}$ sec. Fortunately, the value of t_0 does not enter into the results of this paper.)

B. The Interpretation of Quantum Theory in a Closed System.

Some features of the arrow of time in cosmology could be discussed within the frame of classical physics. We, however, use quantum statistics since we want to discuss later quantum processes like conversion between matter and radiation.

One might object to applying quantum theory to the whole closed cosmos for the following reason: Conventional quantum theory describes a system mathematically by a state vector, but the physical interpretation of this state vector requires the existence of an "external observer," which does certainly not exist for a system comprising the whole cosmos.

This objection, however, is a purely academic one in the following sense: If one applies quantum theory naively, without philosophical qualms, to a large macroscopic closed system, then all predictions of the theory which can be checked experimentally are quite independent of the existence or nonexistence of an external observer. Furthermore, Everett[3] has shown how the quantum theory of a closed macroscopic system can be interpreted consistently by redefining the concept of physical reality.

Another possibility maintaining a more conventional concept of physical reality has been discussed by the author.[4] In this theory the world's state vector behaves as if the macroscopic state of the cosmos were subject to frequent external observations. These macroscopic observations need not disturb the cosmos appreciably, but they guarantee that the macroscopic state of the cosmos is always well defined.

C. The Quantization of the Gravitational Field

We circumvent this difficult problem by using the following approximations:

(a) We describe the cosmos by the world radius $R(t)$ which characterizes the "gravitational background" and by a state vector $\zeta(t)\rangle$ which depends on the particles and fields (including local gravitational fields), embedded in the gravitational background.

(b) We treat the world radius $R(t)$ as a classical quantity [this seems safe since $R(t)$ is truely macroscopic] which satisfies (I5), where u is the expectation value of the space-averaged energy density.

(c) We assume that the state vector $\zeta(t)\rangle$ satisfies a Schrödinger equation

$$i\hbar\dot{\zeta}(t)\rangle = H(t)\zeta(t)\rangle, \tag{I20}$$

where $H(t)$ depends on the time t only via the world radius, for example,

$$H(t) = H[R(t), \dot{R}(t)]. \tag{I21}$$

[3] H. Everett, Rev. Mod. Phys. 29, 454 (1957).
[4] H. Schmidt, Z. Naturforsch. 18a, 265 (1963).

In the framework of the quantum theory discussed in Ref. 4 this coupling between the macroscopic classical quantity $R(t)$ and the quantum mechanical state vector $\zeta(t)\rangle$ and operator u can be discussed without conceptual difficulties.

II. TWO DIFFERENT STATISTICAL BOUNDARY CONDITIONS FOR AN OSCILLATING COSMOS

A. Different Boundary Conditions for Systems in a Laboratory

For a reasonable description of a physical system we need, beside the equations of motion, certain "boundary conditions." These boundary conditions may, but need not, have the form of "initial conditions." The question of the proper boundary conditions is closely related to the question how the system was "prepared."

Let us illustrate this by the following example:

Consider a system whose state vector obeys the Schrödinger equation

$$i\hbar\dot{\zeta}(t)\rangle = H\zeta(t)\rangle. \tag{II1}$$

Now we will generally not be interested in all solutions of (II1), but only in those which satisfy certain restrictions (boundary conditions). To specify these boundary conditions let us assume (for the discussion of our example) the existence of some observable B with eigenvalues b_n and eigenvectors $\beta_n\rangle$,

$$B\beta_n\rangle = b_n\beta_n\rangle \text{ with } b_n \neq b_m \quad \text{for} \quad n \neq m, \tag{II2}$$

such that B can be measured exactly, which implies that the state vector after a B-measurement is completely known.

Suppose now that we have primarily a very wide ensemble S of systems with the same hamiltonian H, but with all possible initial conditions. In order to make reasonable measurements, we "prepare" the system somehow for the experiment, i.e., we select from S some subensemble of systems with certain specified properties.

We mention here two such possibilities:

(1) We consider only those systems as properly prepared for which the B measurement at some time $t = 0$ gave the value $B = b_0$. The ensemble S_I of these systems is specified by the initial condition for the state vector

$$\zeta(0)\rangle = \beta_0\rangle. \tag{II3}$$

The expectation value of some operator Q at time $t > 0$ is here

$$Q_I(t) = \langle\beta_0 e^{iHt/\hbar}Q\, e^{-iHt/\hbar}\beta_0\rangle. \tag{II4}$$

(2) We consider only those systems as properly prepared for which the B-measurement at $t = 0$ as well as the B-measurement at $t = T$ give the value $B = b_0$. The ensemble S_{II} of these systems is determined by a "double-time boundary condition."

To calculate in this ensemble the expectation value $Q_{II}(t)$ of an operator Q at time t (with $0 < t < T$), let us assume for convenience that Q is a projection operator, i.e., $Q^2 = Q$.

Now, as first step, we can easily calculate the probability that, provided $\zeta(0)\rangle = \beta_0\rangle$, we measure $Q = 1$ at time t, and $B = b_0$ at time T:

$$G(B = b_0, T; Q = 1, t \mid B = b_0, 0) = |\langle\beta_0 U(T, t)QU(t, 0)\beta_0\rangle|^2, \tag{II5}$$

where

$$U(t, t') = \exp\{-iH(t - t')/\hbar\}. \tag{II6}$$

The corresponding probability for measuring $Q = 0$ is

$$G(B = b_0, T; Q = 0, t \mid B = b_0, 0) = |\langle\beta_0 U(T, t)\{1 - Q\}U(t, 0)\beta_0\rangle|^2. \tag{II7}$$

Since (II5) and (II6) give relative probabilities for measuring $Q = 1$ and $Q = 0$ in our ensemble S_{II}, the average value of Q becomes

$$Q_{II}(t) = M^{1-}\,|\langle\beta_0 U(T, t)QU(t, 0)\beta_0\rangle|^2 \tag{II8}$$

with $M = |\langle\beta_0 U(T, t)QU(t, 0)\beta_0\rangle|^2 + |\langle\beta_0 U(T, t)\{1 - Q\}U(t, 0)\beta_0\rangle|^2$

Equations (II4) and (II8), referring to systems with single-time and double-time boundary conditions, respectively, can both be checked experimentally. To check (II4) we select only such systems for which $B = b_0$ at $t = 0$ and measure here the value of Q at some time $t > 0$. Equation (II4) does here *pre*dict (statistically) the result of the Q-measurement. To check (II8) we take into account the values of Q at time t ($0 < t < T$) for those systems only which give $B = b_0$ at the times $t = 0$ and $t = T$. Equation (II8) cannot be used here to *pre*dict the outcome of the Q-measurement, since we have to wait until $t = T$ in order to decide whether the particular system belongs to the class S_{II} which alone we want to consider. This explains why the first situation, where the ensemble is specified by a (single-time) initial condition, has a much wider physical interest.

B. Cosmos with a Single-Time or Double-Time Boundary Condition

We have seen that we can study in the laboratory systems with single-time and double-time boundary conditions by properly selecting the systems we take into account.

If the system under observation is the whole cosmos, we have no choice in selecting the boundary conditions, we cannot prepare the systems in a specified way. This "preparation" is done already by nature.

We do not know which boundary conditions are realized in nature. Therefore, it seems reasonable to compare different possibilities. One possibility, where the boundary conditions are such as to guarantee the second law, has already been the subject of many investigations. This possibility seems particularly plausible if one imagines that the cosmos was created in a way a sample is prepared in a laboratory. We call the corresponding cosmos a I-cosmos. Another, perhaps not less reasonable, possibility is a cosmos with time-symmetric initial conditions, such that no time direction and no particular cosmos cycle plays a distinguished role. This cosmos, we call it a II-cosmos, can essentially be described by the double-time boundary conditions mentioned above.

C. The I-Cosmos

Let us assume that the cosmos is at time $t = 0$ maximally contracted and that here the cosmos' state is given. In principle the cosmos' state vector might be specified at $t = 0$. We prefer to characterize the initial state by some density matrix D_0.

For simplicity we will assume that D_0 can be written as

$$D_0 = N_0^{-1} A_0, \tag{II9}$$

where A_0 is a projection operator with trace N_0:

$$A_0^2 = A_0, \qquad \mathrm{Tr}\,\{A_0\} = N_0. \tag{II10}$$

To be more specific, let A_0 be the projector into the space spanned by the eigenvectors of $H(O)$ with energy eigenvalues in the narrow energy band ΔE around some energy E_0.

Then the density matrix D_0 describes an initial state where the total energy [as measured by $H(O)$] is prescribed within $\pm\frac{1}{2}\Delta E$, and where all the N_0 microstates which satisfy this macroscopic requirement are equally likely.

Let us discuss here only qualitatively what the history of such a cosmos might be. We will assume that at $t = 0$ the cosmos is very hot and filled with radiation and thermally excited particles and antiparticles. Furthermore we have here, in a I-cosmos, to assume the existence of uncompensated baryons in the initial state. In the case of equal numbers of particles and antiparticles in the initial state there would be practically no particles left today.[5]

This cosmos would expand very rapidly, the temperature would decrease, most particle–antiparticle pairs would annihilate, the remaining uncompensated particles would first fill the cosmos homogeneously, later galaxies and stars would form and the present cosmos state would be reached. Then the cosmos would continue to expand until in some distant future the cosmos contraction begins.

In this cosmos we could expect the second law to hold during expansion and contraction. After one oscillation, due to irreversible processes, the cosmos' entropy would be increased. This implies an increase of the cosmos' internal energy, and furthermore an increase of the amplitudes of the R-oscillations.[1] [Equation (I5) gives the conservation law

$$2\pi^2 R(t)^3 u(t) - \frac{6\pi^2 R(t)}{\kappa}\,(1 + \dot{R}(t)^2/c^2) = 0.$$

Here the first term is the cosmos' internal energy (energy density times volume), and the second negative term can be considered as the energy of the gravitational background field. We see that this conservation law permits an unlimited increase of the cosmos' internal energy.]

D. The II-Cosmos

Here we assume that the cosmos behaves as if it had been selected at random from an ensemble S_{II} of many cosmos copies.

To define the ensemble S_{II} we start from an ensemble S_I, described by the density matrix D_0 at time $t = 0$. Now we select from S_I those systems for which a measurement of A_0 $(= N_0 D_0)$ at the time T gives the value $A_0 = 1$. (T is the time required for one cosmos oscillation. Due to statistical fluctuations T will vary slightly among different systems, but we neglect this effect.)

The ensemble S_{II} thus defined contains only systems which have at the times $t = 0$ and $t = T$, the same internal energy E_0, within $\pm\frac{1}{2}\Delta E$.

In a similar way we can define an ensemble S_∞ of systems which happen to have at all the times $t = 0, \pm T, \pm 2T, \cdots$ the same internal energy E_0 $(\pm\frac{1}{2}\Delta E)$. This ensemble is, as we shall show in Sec. III invariant under time reflection, $t \to -t$,

[5] R. A. Alpher, J. W. Follin, and R. C. Herman, Phys. Rev. 92, 1347 (1953).

and time displacement, $t \to t + T$. That means all the cosmos oscillations are (apart from individual statistical fluctuations) equal, and the contractions are time-symmetric to the expansions.

Assuming that the entropy in such a cosmos increases during expansion, the entropy must decrease during the following contraction. Here the entropy would increase/decrease with increasing/decreasing world radius. If we furthermore assume that the subjective time direction for an observer in the cosmos is the direction of increasing entropy, then this subjective time direction is always the direction of increasing R, and all observers (no matter in which period of the cosmos' history they live) would see the cosmos expanding, never contracting.

As long as we have in mind the naive picture of a world creation at some time $t = 0$, similar to the preparation of some sample in the laboratory, we would say that it is very "improbable" for a cosmos with initial energy E_0 to reach the same energy after one oscillation (since for almost all systems with initial density matrix D_0 and energy E_0 the energy is increased after one oscillation, due to irreversible processes which increase the entropy and the internal energy).

This argument that a cosmos describable by the ensemble S_{II} or S_∞ be "improbable", however may not be applicable. We must admit that we do not know any thing about the cosmos' statistical boundary conditions, and it might well be that these boundary conditions are determined by symmetry principles rather than by similarities to laboratory conditions.

The ensemble S_{II} with boundary conditions at $t = 0$ and $t = T$ has a lower symmetry than the ensemble S_∞. We will see later, however, that in the region $0 < t < T$ the ensembles S_{II} and S_∞ cannot be distinguished experimentally. It is therefore sufficient if we study in the following a II-cosmos determined by S_{II}.

III. GENERAL MATHEMATICAL THEORY OF II-COSMOS

A. Microscopic Time-Inversion Invariance

The world radius $R(t)$ and the internal state vector $\zeta(t)\rangle$ of the cosmos are in our theory coupled by the Eqs. (I5), (I20), and (I21).

Let us now consider for the moment the different case where $\zeta(t)\rangle$ satisfies (I20) but where $R(t)$ is "given" in such a way that

$$R(t) = R(-t), \tag{III1}$$

$$R(t) = R(t + T). \tag{III2}$$

This implies

$$R(t) = R(T - t). \tag{III3}$$

We shall show later that these symmetry properties of $R(t)$ are self-consistent.

In this time-symmetric background we postulate now microscopic reversibility in the following sense: To every solution $\zeta(t)\rangle$ of the Schrödinger equation there shall exist a time-reflected solution $\hat{\zeta}(t)\rangle$ such that

$$\hat{\zeta}(t)\rangle = S(t)\zeta^*(-t)\rangle, \tag{III4}$$

$$i\hbar\dot{\zeta}(t)\rangle = H(t)\zeta(t)\rangle, \tag{III5a}$$

$$i\hbar\dot{\hat{\zeta}}(t)\rangle = H(t)\hat{\zeta}(t)\rangle. \tag{III5b}$$

Here S is a unitary operator,

$$S(t)^\dagger = S(t)^{-1}, \tag{III6}$$

and $\zeta^*(t)\rangle$ is the conjugate complex (i replaced by $-i$) to $\zeta(t)\rangle$.

If we want to interpret $\hat{\zeta}(t)\rangle$ as the time reflected state then we must have for a "time-symmetric" Schrödinger operator Q (e.g., position, energy density, but not momentum):

$$\langle\zeta(t)Q\zeta(t)\rangle = \langle\hat{\zeta}(-t)Q\hat{\zeta}(-t)\rangle, \tag{III7}$$

for time-symmetric Q.

We postulate a corresponding equation for the time-dependent Hamiltonian:

$$\langle\zeta(t)H(t)\zeta(t)\rangle = \langle\hat{\zeta}(-t)H(-t)\zeta(-t)\rangle. \tag{III8}$$

Next we observe that double-time inversion leads to the original state, i.e., $\hat{\hat{\zeta}}(t)\rangle = e^{i\varphi(t)}\zeta(t)\rangle$, or with (III4),

$$S(t)S^*(-t) = e^{i\varphi(t)}. \tag{III9}$$

[With (III6) and (III12) it follows that here only two possibilities exist: $e^{i\varphi(t)} = \pm 1$.] Here S^* is the operator conjugate complex (i replaced by $-i$) to S, to be distinguished from the Hermitian conjugate $S^\dagger = S^{*\text{transp}}$. With (III4) and (III9) we can write (III7) and (III8) as

$$Q = S(t)Q^*S^\dagger(t), \quad \text{for time-symmetric } Q, \tag{III10}$$

$$H(t) = S(t)H^*(-t)S^\dagger(t). \tag{III11}$$

Substituting (III11) and (III4) into (III5) we see that $\dot{S}(t) = 0$, or

$$S(t) = S \quad \text{is time-independent.} \tag{III12}$$

From (III2) and (I, 21) we obtain the further symmetry property

$$H(t) = H(t + T). \tag{III13}$$

B. Macroscopic Time Symmetry of II-Cosmos

We introduce the projection operator

$$A_0 = \sum_{n=1}^{N_0} \alpha_n\rangle\langle\alpha_n = N_0 D_0, \qquad \text{(III14)}$$

where $\alpha_1\rangle$, $\alpha_2\rangle$, $\cdots$ $\alpha_{N_0}\rangle$ are all (orthonormal) eigenvectors of $H(0)$ with eigenvalues in the interval $(E_0 - \Delta E/2,\ E_0 + \Delta E/2)$. From (III4), (III11), (III12) we see that the vectors

$$\hat{\alpha}_n = S\alpha_n^*\rangle \qquad \text{(III15)}$$

span the same space as the $\alpha_1\rangle$, $\cdots$ $\alpha_{N_0}\rangle$; i.e., we can write

$$A_0 = \sum_{n=1}^{N_0} \hat{\alpha}_n\rangle\langle\hat{\alpha}_n, \quad \text{or with (III15)},$$

$$A_0 = SA_0^*S^\dagger. \qquad \text{(III16)}$$

Let us calculate now the expectation value $Q_{II}(t)$ of some operator Q at time t, for the case of a II-cosmos [c.f. II(d)]. We will assume for convenience that Q is a projection operator, i.e.,

$$Q^2 = Q, \qquad Q = Q^\dagger. \qquad \text{(III17)}$$

This is no severe restriction, since the measurement of any observable is equivalent to the simultaneous measurement of several projection operators. Considering a system where the operator A_0 is measured at $t = 0$ and $t = T$, and where Q is measured at some intermediate time t, we introduce the relative probability that, provided the system's density matrix at $t = 0$ was D_0, we measure $Q = 1$ at time t and $A_0 = 1$ at time T:

$$G(A_0 = 1, T; Q = 1, t \mid D_0 \text{ at } 0) = G_I(A_0 = 1, T; Q = 1, t). \qquad \text{(III18)}$$

The right-hand side of (III18) is only an abbreviated notation which we will use for convenience. According to the definition of II-cosmos in Sec. II(d) we have now

$$Q_{II}(t) = \frac{G_I(A_0=1,T;Q=1,t)}{G_I(A_0=1,T;Q=1,t)+G_I(A_0=1,T;Q=0,t)}. \qquad \text{(III19)}$$

In order to evaluate the expressions (III18), (III19) we introduce the time-displacement operator $U(t, t')$ by

$$i\hbar(\partial/\partial t)U(t, t') = H(t)U(t, t'), \quad U(t, t) = 1. \qquad \text{(III20)}$$

From (III20), (III11), (III12), and (III13) we obtain the symmetry relations

$$U^*(t, t') = S^\dagger U(-t, -t')S, \qquad \text{(III21a)}$$

$$U(t, t') = U(t + T, t' + T). \qquad \text{(III21b)}$$

In particular (III21a, b) give

$$U^*(0, T - t) = S^\dagger U(T, t)S, \qquad U^*(T - t, T) = S^\dagger U(t, 0)S, \qquad \text{(III22)}$$

and two corresponding equations where the arguments in U are interchanged. In order to calculate (III18) assume first that the system was at $t = 0$ in a particular state $\alpha_i\rangle$. Then the probability for measuring $Q = 1$ at time t and $A_0 = 1$ at time T can obviously be written (Q and A_0 are projection operators) as the norm of a vector:

$$\text{Norm } \{|A_0U(T, t)QU(t, 0)\alpha_i\rangle\} = \text{Tr } \{U(0, t)QU(t, T)A_0U(T, t)QU(t, 0)\alpha_i\rangle\langle\alpha_i\}.$$

Since the initial state (density matrix D_0) can be formed, with equal probabilities, by any of the vectors $\alpha_1\rangle$, $\cdots$, $\alpha_{N_0}\rangle$, we obtain generally

$$G_I(A_0 = 1, T; Q = 1, t) = N_0^{-1} \text{Tr } \{Q(t)A_0(T)Q(t)A_0\}. \qquad \text{(III23a)}$$

Here we have introduced

$$Q(t) = U(0, t)QU(t, 0), \qquad A_0(T) = U(0, T)A_0U(T, 0). \qquad \text{(III24)}$$

Equation (III23a) holds for an arbitrary projection operator. Substituting $Q \to 1 - Q$ or $Q \to$ I (unit operator), we obtain

$$G_I(A_0 = 1, T; Q = 0, t) = N_0^{-1} \text{Tr } \{[1 - Q(t)]A_0(T)[1 - Q(t)]A_0\}, \qquad \text{(III23b)}$$

$$G_I(A_0 = 1, T) = N_0^{-1} \text{Tr } \{A_0(T)A_0\}. \qquad \text{(III23c)}$$

Equation (III23c) gives the probability that in the ensemble S_I the value $A_0 = 1$ is measured at time T, this is the probability that a member of S_I belongs also to the ensemble S_{II}.

Remembering (III16) and (III22) we obtain from (III23a) the symmetry relation

$$G_I(A_0 = 1, T; Q = 1, t) = G_I(A_0 = 1, T; \hat{Q} = 1, T - t) \qquad \text{(III25)}$$

with

$$\hat{Q} = SQ^*S^\dagger. \qquad \text{(III26)}$$

In particular we see that for a time-symmetric operator $Q = \hat{Q}$ [c.f. Eq. (III10)], both sides of (III23a) and (III23b) are invariant under the substitution $t \to T - t$. Now the same symmetry applies to (III19), i.e.,

$$Q_{II}(t) = Q_{II}(T - t) \quad \text{for} \quad Q = \hat{Q} \text{ and } 0 < t < T. \tag{III27}$$

Thus in our II-cosmos the expectation value of a time-symmetric operator Q has the same value at some time t (with $0 < t < \frac{1}{2}T$) during expansion and at the corresponding time $T - t$ during contraction, i.e., the expansion and the contraction occur time-symmetrically to each other. (Note that we have proved this symmetry only for the oscillation in the time interval $0 < t < T$, "between" the two boundary conditions.)

C. Cosmos with Multiple-Time Boundary Conditions

Let us consider first a cosmos with a triple-time boundary condition where $A_0 = 1$ is required at the times 0, T, $2T$. The corresponding ensemble S_{III} is specified as that subensemble of S_I for which $A_0 = 1$ is observed at the times $t = T$ and $t = 2T$.

To calculate here the expectation value $Q_{III}(t)$ of a time-symmetric projection operator Q, we have to replace in (III19) the function G, assuming, for example, $0 < t < T$, by

$$G_I(A_0 = 1, 2T; A_0 = 1, T; Q = 1, t) \quad \text{for} \quad 0 < t < T. \tag{III28}$$

In order to simplify this expression we will make the important assumption that, during passage through a state of maximal contraction, the cosmos remembers only the internal energy and forgets all other details of its past history.

To be more specific, consider a contracting cosmos which reaches a state of maximal density at the time $t = T$. Let here the cosmos' state vector be

$$\zeta(T)\rangle = \sum_{n=1}^{N_0} c_n \alpha_n\rangle.$$

This implies that here $A_0 = 1$.

We assume now that at $t = T$ the cosmos is sufficiently hot and that here all observable coherences between states are quickly destroyed, such that the cosmos' history at $t > T$ is practically independent of the values of the c_n's in the expression for $\zeta(T)\rangle$. This means we can calculate in the region $t > T$ as if the cosmos' density matrix at the time T were D_0.

We refer to this assumption as "the cosmos' memory loss at maximal contraction".

Now we can replace the expression (III28) by

$$G(A_0 = 1, 2T \mid D_0 \text{ at } T) \cdot G_I(A_0 = 1, T; Q = 1, t). \tag{III29}$$

This shows immediately that here

$$Q_{III}(t) = Q_{II}(t) \quad \text{for} \quad 0 < t < T \tag{III30a}$$

and a similar argument gives

$$Q_{III}(t + T) = Q_{II}(t) \quad \text{for} \quad 0 < t < T. \tag{III30b}$$

In this cosmos the two oscillations in the interval $0 < t < 2T$ ("between" the boundary conditions) are equal and the contractions are time-symmetric to the expansions.

In a similar way we can define more generally an ensemble S_{2M+1} of cosmos copies for which $A_0 = 1$ is satisfied at the times $t = 0, \pm T, \pm 2T, \cdots \pm MT$. Finally taking M very large ($M \to \infty$) we obtain the ensemble S_∞ which is invariant under the symmetry operations

$$\begin{aligned} t &\to t + T, \\ t &\to -t. \end{aligned} \tag{III31}$$

That means the expectation value $Q_\infty(t)$ of a time-symmetric projection operator Q satisfies here

$$\begin{aligned} Q_\infty(t) &= Q_\infty(t + T), \\ Q_\infty(t) &= Q_\infty(-t). \end{aligned} \tag{III32}$$

Here, except for individual statistical fluctuations, all cosmos oscillations are equal and the contractions are time-symmetric to the expansions. In particular (III32) implies for the average energy density $u(t)$ in S_∞:

$$u(t) = u(-t) = u(t + T). \tag{III33}$$

We see now that in the ensemble S_∞ the symmetry assumptions (III1), (III2) about $R(t)$ are self consistent: If $R(t)$ has the symmetry (III31), then $u(t)$, as we have just proved, has the same symmetry. On the other hand, if $u(t)$ has the symmetry (III31), then by (I5) also $R(t)$ must have this symmetry.

From our assumption about the cosmos' memory loss at maximal contraction it follows that in the interval $0 < t < T$ the ensembles S_{II} and S_∞ cannot be distinguished. It is therefore sufficient if we study this interval in the II-cosmos. All oscillations in the ensemble S_∞ are equal to this particular oscillation of the II-cosmos.

D. The Effective Density Matrix for the II-Cosmos

Equations (III19), (III23a), and (III23b) give the expectation value $Q_{II}(t)$ of the operator Q in the II-cosmos. We can simplify these results under the assumption that the projection operator Q is a "macroscopic observable", which means that a measurement of Q need not disturb the system apprec-

iably. In particular we require that Q satisfies the equation

$$G_{\mathrm{I}}(A_0 = 1, T; Q = 1, t) + G_{\mathrm{I}}(A_0 = 1, T; Q = 0, t) = G_{\mathrm{I}}(A_0 = 1, T). \quad \text{(III34)}$$

This equation says: The probability that a cosmos with density matrix D_0 at $t = 0$ leads to $A_0 = 1$ at time T is the same, whether the operator Q was measured [left-hand side of Eq. (III34)] or not measured (right-hand side) at time t. Substituting (III23a, b) into (III34) we obtain the identity

$$2\,\mathrm{Tr}\,\{Q(t)A_0(T)Q(t)A_0\} = \mathrm{Tr}\,\{Q(t)A_0(T)A_0 + Q(t)A_0A_0(T)\}. \quad \text{(III35)}$$

Now (III19) can be written with (III23, 34, 35) as

$$Q_{\mathrm{II}}(t) = \mathrm{Tr}\,\{Q(t)\tilde{D}_0\} \quad \text{for} \quad 0 < t < T \quad \text{(III36)}$$

with

$$\tilde{D}_0 = \frac{[A_0, A_0(T)]_+}{2\,\mathrm{Tr}\,\{A_0A_0(T)\}}. \quad \text{(III37)}$$

Comparing this with the expectation value of Q in the ensemble S_{I} with initial density matrix D_0 at $t = 0$,

$$Q_{\mathrm{I}}(t) = \mathrm{Tr}\,\{Q(t)D_0\}, \quad \text{(III38)}$$

we see that our ensemble S_{II} behaves for $0 < t < T$ like an ensemble with the single-time boundary condition specified by the effective density matrix $\tilde{D}_0$ from (III37) at $t = 0$.

IV. CONVERSION BETWEEN MATTER AND RADIATION IN A SIMPLE MODEL OF A II-COSMOS WITHOUT STAR CONDENSATION.

A. Introduction

The initial state of a I-cosmos has been discussed for example by Alpher, Follin, and Herman.[5] These authors assume that at some time the cosmos was very hot, such as to contain a high concentration of thermally excited particle–antiparticle pairs. Due to the cosmos' expansion the temperature must have decreased rapidly such that very soon practically all particle–antiparticle pairs were annihilated. In order to account for the presently existing matter in the cosmos the authors have to assume that the initial state contained a certain concentration of noncompensated baryons.

The statistical treatment of the early cosmos seems comparatively easy since here all interesting reactions (for example the formation of higher elements from protons and neutrons) involve only a small number of particles. Much more difficult is the statistical treatment of the later cosmos stages where the formation of galaxies and stars occurs. Thus its seems natural that we start our discussion of a II-cosmos by considering its early stage.

There arises, however, the following difficulty: Due to the second boundary condition at $t = T$, the cosmos history at $t \approx 0$ depends on the cosmos history in the whole interval $0 < t < T$. Therefore we can no longer study just the initial state without considering simultaneously the whole cosmos history.

We hope that we will finally overcome this difficulty by approximating the later cosmos stages by more or less crude models. At present we will study only the (unrealistic) model of a II-cosmos without cooperative processes, i.e., we will assume that the expanding cosmos starts contracting before the formation of galaxies and stars begins. In particular we shall study the initial state of this cosmos model.

Let the cosmos in its initial state be filled with "hot vacuum", that is thermal radiation and thermally excited particle–antiparticle pairs. Let, furthermore, this state be macroscopically specified by its internal energy $H(0)$. Then the effective density matrices for the corresponding I-cosmos and II-cosmos respectively are D_0 and [cf. (III14, 37)]

$$\tilde{D}_0 = \frac{[D_0, D_0(T)]_+}{2\,\mathrm{Tr}\,\{D_0D_0(T)\}}. \quad \text{(IV1)}$$

The density matrix D_0 describes an ensemble where the cosmos is (with overwhelming probability) filled nearly homogeneously with radiation, thermal particle–antiparticle pairs and noncompensated baryons. Here practically all particle–antiparticle pairs recombine during the cosmos expansion.

The density matrix D_0, however, describes (in the model considered) an ensemble where each cosmos shows large-scale fluctuations of the baryon distribution such that, with high probability, one half of the cosmos contains more baryons and the other half more antibaryons. Now not all particles can annihilate during cosmos expansions because the two cosmos halves with excess baryons or antibaryons are spacially separated.

In particular our model gives the right order of magnitude for the present density of baryons in our half of the cosmos. Whether this is just a coincidence or has deeper significance will have be decided after we can apply the double-time boundary conditions to more realistic cosmos models.

B. The Cosmos Model: Approximations

We consider a cosmos which contains light (γ) and one type of baryon–antibaryon pair (N, $\bar{N}$).

We assume that there is only one reaction possible,

$$N + \bar{N} \leftrightarrow 2\gamma,$$

and that the cross section for N–$\bar{N}$ annihilation is constant, $\sigma_0 = 10^{-24}$ cm^2. In the following we will use for simplicity the "light-cosmos approximation", i.e., we approximate the world radius $R(t)$ and the temperature ϑ in our cosmos by the corresponding values in a light cosmos, in particular [cf. (I.14, 18)]:

$$\left.\begin{aligned} R(t) &= R_{\min}\left(\frac{|t| + t_0}{t_0}\right)^{\frac{1}{2}} \\ k\vartheta(t) &= \lambda(|t| + t_0)^{-\frac{1}{2}} \end{aligned}\right\} \quad \text{for} \quad |t| \ll T. \qquad \text{(IV1b)}$$

This approximation does not change the order of magnitude of the results.

In order to show that the two-time boundary condition favours states with large-scale inhomogeneities we consider two halves R and L of the cosmos separately. Let the R-half and the L-half be specified in such a way that the boundary surface between R and L is as small as possible. Then the particles in L are far separated from the particles in R, except for the boundary region whose influence may be neglected for a large cosmos.

The distinction of two cosmos halves may seem to be a very arbitrary procedure. In fact, we might imagine the cosmos divided into more than two regions which are spacially far separated. It seems, however, that this generalization will not change our main result concerning the density of the surviving baryons in our part of the cosmos. Considering now the initial cosmos state we introduce the number Z of noncompensated baryons in the R-half of the cosmos. (Z = number of baryons minus number of antibaryons in R).

In order to describe the statistical fluctuations of Z among different cosmos copies we define $P_{\mathrm{I}}[Z]$ and $P_{\mathrm{II}}[Z]$ as the probability for finding a certain value of Z in the ensembles S_{I} and S_{II}, respectively. The function $P_{\mathrm{I}}[Z]$ can be easily calculated (Appendix I):

$$P_{\mathrm{I}}[Z] = \frac{1}{(2\pi)^{\frac{1}{2}}\rho} \exp\left\{-\frac{1}{2}\frac{Z^2}{\rho^2}\right\} \qquad \text{(IV2)}$$

with

$$\rho^2 = \frac{V_0}{12}\left(\frac{k\vartheta_0}{\hbar c}\right)^3. \qquad \text{(IV3)}$$

(Here V_0 and ϑ_0 are volume and temperature in the initial cosmos.) This confirms that the fluctuations in a I-cosmos "practically never" give rise to an appreciable density of noncompensated particles in one half of the cosmos.

The calculation of $P_{\mathrm{II}}[Z]$ requires more effort. Let us start from the general formalism developed in Sec. III. Let us write A_0 as

$$A_0 = \sum_{Z=-\infty}^{\infty} B_Z, \qquad \text{(IV4)}$$

where B_Z is the projector into the space of those microstates for which the total internal energy is E_0 ($\pm\Delta\frac{1}{2}E$) *and* for which the number of noncompensated baryons in the R-half is Z. Then

$$B_Z A_0 = A_0 B_Z = B_Z. \qquad \text{(IV5)}$$

Note that in a I-cosmos the probability for a given Z is proportional to the number of microstates realizing that Z-value, i.e.,

$$P_{\mathrm{I}}[Z] = \text{const}\cdot\mathrm{Tr}\,\{B_Z\}. \qquad \text{(IV6)}$$

("const" means here and in the following a quantity which does not depend on Z.) From (IV5, 6) and (III36, 37) we get now

$$P_{\mathrm{II}}[Z] = \mathrm{Tr}\,\{B_Z \tilde{D}_0\} = \text{const}\cdot\mathrm{Tr}\,\{B_Z A_0(T)\} \qquad \text{(IV7)}$$

or

$$P_{\mathrm{II}}[Z] = \text{const}\cdot P_{\mathrm{I}}[Z]\,\frac{\mathrm{Tr}\,\{B_Z A_0(T)\}}{\mathrm{Tr}\,\{B_Z\}} \qquad \text{(IV8)}$$

or

$$P_{\mathrm{II}}[Z] = \text{const}\, P_{\mathrm{I}}[Z] G(A_0 = 1,\, T \mid B_Z \text{ at } 0). \qquad \text{(IV9)}$$

Here $G(A_0 = 1,\, T \mid B_Z \text{ at } 0)$ is the probability that $A_0 = 1$ is measured at time T in a cosmos with initial (at $t = 0$) density matrix $B_Z/\mathrm{Tr}\,B_Z$. This is the probability that a cosmos with Z noncompensated baryons in one half has not changed its internal energy ($\pm\Delta\frac{1}{2}E$) after one oscillation.

C. Calculation of $G(A_0 = 1,\, T \mid B_Z \text{ at } 0)$

Consider an ensemble with density matrix $B_Z/\mathrm{Tr}\,\{B_Z\}$ at $t = 0$. Here the R-half and the L-half of the cosmos are both filled practically homogeneously with particles and radiation, but the R-half contains $2Z$ baryons more than the L-half. Since the two cosmos halves meet only on a boundary which is negligible for a large cosmos, the number Z remains practically constant in time.

After one oscillation, at time T, the internal energy of the cosmos will generally have changed. Let $P(E, Z)dE$ be the probability that this energy change is E, within dE. Then clearly

$$G(A_0 = 1,\, T \mid B_Z \text{ at } 0) = \text{const}\, P(0, Z). \qquad \text{(IV10)}$$

We will approximate $P(E, Z)$ by a normal distribution:

$$P(E, Z) = \frac{1}{(2\pi)^{\frac{1}{2}}\sigma} \exp\left\{-\frac{(E-\bar{E})^2}{2\sigma^2}\right\}, \qquad \text{(IV11)}$$

where $\bar{E}$ and σ^2 depend on Z.

Next we make the serious assumption that the approximation (IV11) can be used at $E = 0$. (A careful study of the error implied would be desirable.) This gives

$$P(0, Z) = \text{const}\ \sigma^{-1} \exp\left\{-\frac{\bar{E}^2}{2\sigma^2}\right\}, \qquad \text{(IV12)}$$

and with (IV2, 9, 10) we obtain

$$P_{II}[Z] = \text{const}\ \sigma^{-1} \exp\left\{-\frac{1}{2}\left(\frac{Z^2}{\rho^2} - \frac{\bar{E}^2}{\sigma^2}\right)\right\}. \qquad \text{(IV13)}$$

Now we have to calculate $\bar{E}(Z)$ and $\sigma^2(Z)$.

D. Calculation of $\bar{E}(Z)$ and $\sigma^2(Z)$

We introduce the parameter

$$\Theta = Mc^2/k\vartheta, \qquad \Theta_0 = Mc^2/k\vartheta_0 \qquad \text{(IV14)}$$

where ϑ is the cosmos temperature and ϑ_0 its value at $t = 0$.

Using the light-cosmos approximation, the cosmos volume can be written as [cf. (I8, 17)]

$$V(\Theta) = V_0 \cdot (\Theta/\Theta_0)^3. \qquad \text{(IV15)}$$

Next we need the density $C(\Theta)$ of thermally excited $N - \bar{N}$ pairs in thermal equilibrium at the temperature ϑ. This is calculated in Appendix II, (AII31):

$$C(\Theta) = 8\pi\left(\frac{Mc}{h}\right)^3 \frac{1}{\Theta^3} K(\Theta), \qquad \text{(IV16)}$$

$$K(\Theta) = \int_\Theta^\infty d\eta\, \frac{\eta(\eta^2 - \Theta^2)^{\frac{1}{2}}}{1 + e^\eta}. \qquad \text{(IV17)}$$

In particular we have

$$K(\Theta) = (\tfrac{1}{2}\pi)^{\frac{1}{2}}\Theta^{\frac{1}{2}}e^{-\Theta} \quad \text{for} \quad \Theta \gg 1. \qquad \text{(IV17a)}$$

Now consider a cosmos which has at $t = 0$ in the R-half Z (>0) noncompensated baryons. We want to compare here the density $2Z/V(\Theta)$ of these baryons with the density of thermal pairs $C(\Theta)$. [The expression (IV16) has to be changed slightly if noncompensated particles are present, but this change is here negligible.]

Let us assume that at $t = 0$ the pair density is much higher than the density of noncompensated baryons. During expansion the number of pairs decreases; the number Z, however, remains practically constant. Thus at a certain value $\Theta = \Theta_Z$ (during expansion), the density of thermal pairs becomes equal to the density of noncompensated baryons in the R-half:

$$C(\Theta_Z) = 2Z/V(\Theta_Z). \qquad \text{(IV18)}$$

With (IV14–16) this gives the following relationship between the original density $2Z/V_0$ of noncompensated particles in each cosmos half and the value Θ_Z of Θ at which the thermal density of pairs equals the density of noncompensated particles:

$$2Z/V_0 = 8\pi(k\vartheta_0/hc)^3 K(\Theta_Z). \qquad \text{(IV19)}$$

In Appendix II we shall derive (under certain simplifying assumptions) the following expressions for $\bar{E}$ and σ^2 in terms of Θ_Z [see Eqs. (A.II29), (A.II38), Appendix II]:

$$\bar{E} = \frac{4}{7}\left(\frac{2}{\pi}\right)^{\frac{1}{2}} \frac{(k\vartheta_0)^2}{Mc^2} \frac{V_0}{ct_0\sigma_0} \Theta_Z^{7/2}, \qquad \text{(IV20)}$$

$$\sigma^2 = \overline{(E - \bar{E})^2} = 4k\vartheta_0\bar{E}. \qquad \text{(IV21)}$$

At this point we want to explain only qualitatively why $\bar{E}$ increases with increasing Θ_Z.

The average increase $\bar{E}$ of the cosmos energy (for the ensemble S_I) during one oscillation is due to the finite rate of the reaction $N + \bar{N} \leftrightarrow 2\gamma$. During expansion, for example, the annihilation of the N–$\bar{N}$ pairs lags behind, i.e., there are more pairs present than in the case of an reversible (infinitely slow) expansion. And since particle pairs contribute less to the pressure p in the cosmos than light quanta of the corresponding energy, the pressure in the expanding cosmos is lower than it would be in the case of a reversible expansion. Thus, in particular, the pressure during expansion is lower than the pressure in the corresponding state during contraction. This implies that the cosmos' internal energy has increased after one oscillation.

This energy increase stems mainly (as the calculation in Appendix II shows) from the region of low density ($\Theta \gg 1$). Here each particle must travel very long before it can annihilate, i.e., the relaxation time is very long. Therefore, during expansion, the few existing particles lag very far behind the equilibrium situation, and so a large contribution to $\bar{E}$ can result from this region.

The presence of uncompensated particles favors the annihilation of the minority particles. Therefore, with increasing number of uncompensated particles, i.e., with decreasing Θ_Z, the deviation from the equilibrium state decreases, and so does $\bar{E}$.

E. The Most Probable Value of Z

The most probable value of Z is given by the maximum of (IV13), which is essentially the minimum of

$$f(\Theta) = \frac{\bar{E}^2}{\sigma^2} + \frac{Z^2}{\rho^2}. \tag{IV22}$$

With (IV21, 20, 3, 19) and (I18, 19) we get

$$f(\Theta_z) = \text{const}\ \{K(\Theta_z)^2 + \mu\Theta_z^{7/2}\} \tag{IV23}$$

with

$$\mu = \frac{\pi^3}{63\cdot 5^{\frac{1}{2}}} \frac{h}{M\sigma_0} \left(\frac{h\kappa}{c}\right)^{\frac{1}{2}}. \tag{IV24}$$

Choosing for example

$$M = 2\cdot 10^{-24}\ \text{g}, \qquad \sigma_0 = 10^{-24}\ \text{cm}^2 \tag{IV25}$$

we obtain

$$\mu = 4.9\cdot 10^{-22}. \tag{IV26}$$

From (IV17) we see that $K(\Theta)$ decreases monotonically. Therefore $f(\Theta_z)$ has just one minimum, at $\Theta_z = \Theta_{\max}$. Furthermore one sees easily that $\Theta_{\max} \gg 1$, since μ is so small. Therefore, to determine the minimum of $f(\Theta_z)$ we may use the approximation (IV17a). This gives for $\Theta_{\max}$ the condition

$$\exp\{2\Theta_{\max}\} = \frac{2\pi}{7\mu}\Theta_{\max}^{\frac{1}{2}} \tag{IV27}$$

or with (IV26)

$$\Theta_{\max} \approx 25. \tag{IV28}$$

Inserting this value $\Theta_z = \Theta_{\max}$ into (IV19), we obtain the most probable number $Z = Z_{\max}$ of noncompensated particles in one cosmos half.

F. The Density of Surviving Baryons in our Part of the Cosmos

In the light cosmos approximation the relation between the time t and the parameter Θ is [cf. (IV14) and (I18)] for $|t| \gg t_0$

$$\Theta = \frac{Mc^2}{\lambda}\sqrt{t}. \tag{IV29}$$

Then $t_{\max}$ (corresponding to $\Theta_{\max}$) is the time where the pair density in our part of the cosmos falls below the density of noncompensated baryons.

We can assume that at this stage the energy density of the radiation is much higher than the energy density of the noncompensated baryons. Only at some later time $t = \hat{t}$ (corresponding to $\hat{\Theta}$, will the energy density of the radiation fall below the energy density of the matter (which is then present only in the form of noncompensated particles).

We want now to calculate the time $\hat{t}$, which can be used as a measure for the density of surviving baryons. The energy density of the noncompensated baryons in the R-half is for $k\vartheta \ll Mc^2$ [cf. (IV19, 15, 17a)]:

$$Mc^2 \frac{2\cdot Z_{\max}}{V(t)} = 2(2\pi)^{\frac{1}{2}} \frac{M^4c^5}{h^3\Theta^3}\Theta_{\max}^{\frac{1}{2}} e^{-\Theta_{\max}}. \tag{IV30}$$

The energy density of the radiation is [cf. (IV14) and (I17)]

$$u_{\text{rad}} = \frac{8\pi^5}{15}\frac{M^4c^5}{h^3\Theta^4}. \tag{IV31}$$

Equating (IV30) and (IV31) we obtain for $\hat{\Theta}$

$$\hat{\Theta} = \frac{\pi^3(2\pi)^{\frac{1}{2}}}{15}\Theta_{\max}^{-\frac{1}{2}} e^{\Theta_{\max}}. \tag{IV32}$$

This gives with (IV24, 25, 27, 29) and (I19) the result

$$\hat{t} = \frac{3}{10}\pi^2(2\pi)^{\frac{1}{2}}\frac{\sigma_0}{Mc^3\kappa}\Theta_{\max}^{-5/2} = 0.7\cdot 10^6\ \text{years}. \tag{IV33}$$

G. Gamow [*The Creation of the Universe* (The Viking Press, New York, 1961)] estimates for $\hat{t}$ the value $\hat{t}_{\text{Gamow}} = 3\cdot 10^8$ years. This differs from our $\hat{t}$ by a factor 400, which means [cf. (I14–18)] that our value for the density of noncompensated baryons in the early cosmos is by a factor 20 larger than Gamow's estimate.

This "discrepancy" might not seem too bad in view of the fact that we used in our model only one type of particles and assumed a fairly arbitrary value for σ_0.

We must bear in mind, however, that the direct application of our results to the actual cosmos is not justified, because our model does not take into account cooperative processes like star formation, which might (in the ensemble S_{II}) have a strong influence on the initial state.

Therefore the discussion of a more refined model which could account for star condensation (at least in a crude way) would be desirable.

APPENDIX I: DERIVATION OF EQ. (IV 2).

Let us consider here a cosmos which is at $t = 0$ in a state of thermal equilibrium (temperature ϑ_0, volume V_0). We are interested in the fluctuations of the numbers of baryons and antibaryons in the cosmos halves R and L.

Let $(M + M_R)$ and $(M + M_L)$ be the numbers

of baryons in R and L respectively, and $(M + \tilde{M}_R)$, $(M + \tilde{M}_L)$ the corresponding numbers of antibaryons. Here M is the average value for all four numbers. Let us assume temporarily that R and L can exchange energy and particles with some large reservoir of temperature ϑ_0.

Then, considering small deviations from equilibrium only, we may approximate the probability for a particular set of the above four numbers by a normal distribution:

$$P[M_R, M_L, \tilde{M}_R, \tilde{M}_L] = \frac{1}{(2\pi)^2 \rho^4} \exp\left\{-\frac{1}{2\rho^2}(M_R^2 + M_L^2 + \tilde{M}_R^2 + \tilde{M}_L^2)\right\}. \tag{A.I1}$$

To calculate ρ, consider the volume $V_0/2$ in a large heat bath of temperature ϑ_0, with particle exchange.

Let ϵ_n and n_n be the energy and occupation number of the baryon state $\mathbf{n}$. (We consider the baryons as spin-$\frac{1}{2}$ particles with mass M. Then the symbol $\mathbf{n}$ must characterize momentum and spin.)

Now

$$M = \sum_n \bar{n}_n,$$

$$\rho^2 = \sum_n \overline{(n_n - \bar{n}_n)^2} = \sum_n \bar{n}_n(1 - \bar{n}_n) = \sum_n \frac{e^{\epsilon_n/k\vartheta_0}}{(1 + e^{\epsilon_n/k\vartheta_0})^2}. \tag{A.I2a}$$

A short calculation gives

$$\rho^2 = \frac{V_0}{2\pi^2}\left(\frac{Mc}{\hbar}\right)^3 \frac{1}{\Theta_0^3} \int_{\Theta_0}^{\infty} d\eta\, \eta(\eta^2 - \Theta_0^2)^{\frac{1}{2}} \frac{e^\eta}{(1 + e^\eta)^2} \tag{A.I2b}$$

with

$$\Theta_0 = Mc^2/k\vartheta_0. \tag{A.I3}$$

For $\Theta_0 \ll 1$ this gives

$$\rho^2 = \frac{V_0}{12}\left(\frac{Mc}{\hbar}\right)^3 \frac{1}{\Theta_0^3} = \frac{V_0}{12}\left(\frac{k\vartheta_0}{\hbar c}\right)^3. \tag{A.I4}$$

We are interested now in the case where V_0 contains equal numbers of baryons and antibaryons. (Particle exchange between R and L, but no particle exchange with the reservoir). This gives the restriction

$$M_R + M_L = \tilde{M}_R + \tilde{M}_L. \tag{A.I5}$$

Here Eq. (A.I1) gives still the correct relative probabilities for our system with restriction (A.I5).

We introduce

$$Z = M_R - \tilde{M}_R = \tilde{M}_L - M_L,$$
$$R = M_R + M_L = \tilde{M}_R + \tilde{M}_L, \tag{A.I6}$$
$$W = M_R - \tilde{M}_L = \tilde{M}_R - M_L.$$

Now the probability for a particular configuration Z, R, W becomes

$$P[Z, R, W] = \text{const} \exp\left\{-\frac{1}{2\rho^2}(Z^2 + R^2 + W^2)\right\}. \tag{A.I7}$$

This gives the probability of a particular Z-value

$$P_I[Z] = \frac{1}{(2\pi)^{\frac{1}{2}}\rho} \exp\left\{-\frac{1}{2\rho^2} Z^2\right\}. \tag{A.I8}$$

This is the result stated in (IV2).

APPENDIX II

A. Cosmos without Relaxation

We consider a cosmos of volume $V(t)$ which is filled homogeneously with radiation and N–$\bar{N}$ pairs. The generalization to the case where two cosmos halves have different baryon concentrations will be discussed in (E). Let the baryons have mass M and spin $\frac{1}{2}$ and let σ_0 be the cross section for the annihilation reaction $N + \bar{N} \to 2\gamma$.

Neglect for the moment relaxation effects ($\sigma_0 \to \infty$). Then volume $V(t)$, energy density $u(t)$, temperature $\vartheta(t)$, and pressure $p(t)$ are well defined and can be calculated.

We characterize the possible baryon states by the wave vector $\mathbf{k}$, or by the number vector $\mathbf{n}$, with

$$\mathbf{k} = \frac{2\pi}{V(t)^{\frac{1}{3}}}\mathbf{n}, \qquad \mathbf{n} = (n_1, n_2, n_3) \quad \text{with integer } n_i. \tag{A.II1}$$

The corresponding baryon energy is

$$\epsilon_n = [M^2c^4 + c^2\hbar^2k^2]^{\frac{1}{2}}. \tag{A.II2}$$

The occupation numbers $n_n(t)$ and $m_n(t)$ for N and $\bar{N}$ have the equilibrium average values

$$\bar{n}_n(t) = \bar{m}_n(t) = (1 + e^{\epsilon_n/k\vartheta})^{-1}. \tag{A.II3}$$

The pressure contribution of a N or $\bar{N}$ with energy ϵ_n is

$$\Pi_n^N = \frac{1}{3V(t)} \frac{\epsilon_n^2 - M^2c^4}{\epsilon_n}. \tag{A.II4a}$$

The pressure of a light quantum with energy ϵ is

$$\Pi_\epsilon = \frac{1}{3V(t)}\epsilon. \tag{A.II4b}$$

If the reaction $\gamma + \gamma \rightarrow N + \bar{N}$ occurs, where ϵ_n, ϵ_m are the energies of N and $\bar{N}$, then according to (A.II4) the pressure is changed by

$$\delta p = \Pi_n + \Pi_m, \tag{A.II5a}$$

with

$$\Pi_n = -\frac{M^2c^4}{3V(t)\epsilon_n}, \tag{A.II5b}$$

because particles furnish less pressure than light with the same energy.

B. Increase E of Cosmos Energy after one Oscillation, with Relaxation

Let us first "switch on" the relaxation only in the short time interval $t_1 < t < t_1 + \Delta t$, where Δt is long, however, compared with the relaxation time.

We write the occupation numbers for the particle states in this time interval as

$$\begin{aligned} n_n(t) &= \bar{n}_n(t) + \delta n_n(t) \\ m_n(t) &= \bar{m}_n(t) + \delta m_n(t). \end{aligned} \tag{A.II6}$$

The fluctuations are restricted by the particle conservation

$$\sum_n \delta n_n(t) = \sum_n \delta m_n(t). \tag{A.II7}$$

For simplicity we will furthermore assume for the fluctuations

$$\delta n\hat{\mathfrak{n}}(t) = \delta m_n(t) \qquad \text{for all } \mathbf{n}. \tag{A.II8}$$

This does not change the order of magnitude of the results. [We might as well neglect the restriction (A.II7) and consider $\delta n_n(t)$ and $\delta m_n(t)$ as independent. This gives the same results, except for a factor $\frac{1}{2}$ in the expression for $\overline{(E - \bar{E})^2}$.]

The fluctuations (A.II8) give rise to the pressure change (A.II5b)

$$\delta p(t) = 2 \sum_n \Pi_n \, \delta n_n(t). \tag{A.II9}$$

Due to this changed pressure, the gravitational field pumps in $(t_1, t_1 + \Delta t)$ the additional energy into the cosmos:

$$\Delta E = -\int_{t_1}^{t_1+\Delta t} \delta p(t) \dot{V}(t)\, dt. \tag{A.II10}$$

This additional energy will increase the light energy and the particle energy in the cosmos. We will assume for simplicity (light-cosmos approximation) that the added energy is transformed into light only. Then this additional light undergoes blue shift during cosmos contraction. Therefore, by adding the energy ΔE at time t, the cosmos energy at the end of the oscillation is increased by

$$\Delta E' = \Delta E \left(\frac{V(t)}{V_0}\right)^{\frac{1}{3}}. \tag{A.II11}$$

Here we can approximate $V(t)$ in those regions where nearly all relaxation processes occur [cf. (I14)] by

$$\begin{aligned} V(t) &= V_0 \left(\frac{|t| + t_0}{t_0}\right)^{\frac{1}{3}} \qquad \text{for } t \approx 0, \\ V(t) &= V_0 \left(\frac{|T - t| + t_0}{t_0}\right)^{\frac{1}{3}} \quad \text{for } t \approx T. \end{aligned} \tag{A.II12}$$

Due to the relaxation processes in the interval $(t_1, t_1 + \Delta t)$, the internal cosmos energy at the end of the cycle is increased by [cf. (A.II10–12)]

$$\Delta E' = -\int_{t_1}^{t_1+\Delta t} \delta p(t) \dot{V}(t) \left(\frac{|t| + t_0}{t_0}\right)^{\frac{1}{3}} dt \quad \text{for } t_1 \approx 0 \tag{A.II13}$$

(corresponding expression for $t_1 \approx T$).

The total energy increase of the world during one oscillation (with relaxation switched on all the time) is now, taking only the terms linear in $\delta\Pi$,

$$\begin{aligned} E &= E^+ + E^- \\ &= \int_0^{T/2} dt\, L(t) \sum_n \Pi_n(t) \{\delta n_n(t) - \delta n_n(T - t)\} \end{aligned} \tag{A.II14}$$

with

$$L(t) = -3 \frac{V_0}{t_0^2} (t + t_0). \tag{A.II15}$$

C. Average Energy Increase $\bar{E}$.

The energy increase E is a random variable, depending on the statistical fluctuations $\delta n_n(t)$. We shall assume here that E has a normal distribution, i.e., that the probability to find E in some interval dE is given by $P(E)dE$, with

$$P(E) = \frac{1}{(2\pi)^{\frac{1}{2}}\sigma} \exp\{-(E - \bar{E})^2/2\sigma^2\}. \tag{A.II16}$$

In order to calculate $\bar{E}$ let us introduce

$$\left.\begin{aligned} \overline{n_n(t)} &= \text{average occupation number for the baryon state } \mathbf{n}; \\ \alpha_n(t) &= \tau_n(t)^{-1} = \text{probability for annihilation of a baryon; out of the state } \mathbf{n}, \text{ per second} \\ \beta_n(t) &= \text{probability for creation of a baryon into } \mathbf{n}, \text{ per second.} \end{aligned}\right\} \tag{A.II17a}$$

Then we have

$$(d/dt)\overline{n_n(t)} = -\alpha_n(t)\overline{n_n(t)} + \beta_n(t)\{1 - \overline{n_n(t)}\}. \quad \text{(A.II17b)}$$

Next we consider for comparison an infinitely slowly expanding cosmos, without relaxation. We can write here [cf. (A.II12)] the volume as

$$V(\tilde{t}) = V_0[(\tilde{t} + t_0)/t_0]^{\frac{3}{2}} \text{ for positive } \tilde{t} \approx 0, \quad \text{(A.II18)}$$

where now $\tilde{t}$ is not the time, but some slowly with t increasing parameter. For the cosmos temperature $\tilde{\vartheta}$ we have here [cf. (I18)]

$$k\tilde{\vartheta}(\tilde{t}) = \lambda(\tilde{t} + t_0)^{-\frac{1}{2}}. \quad \text{(A.II18a)}$$

Writing the average occupation number, lifetime, and the transition probabilities in this case $\tilde{n}_n(t)$, $\tilde{\tau}_n(\tilde{t})$, $\tilde{\alpha}_n(\tilde{t})$, $\tilde{\beta}_n(\tilde{t})$ [note that $\tilde{n}_n(t)$ is the average value in a cosmos without relaxation, or in a infinitely slowly expanding cosmos, while $\overline{n_n(t)}$ refers to the actual cosmos], we have (in thermal equilibrium)

$$\tilde{\beta}_n(\tilde{t})/[\tilde{\alpha}_n(\tilde{t}) + \tilde{\beta}_n(\tilde{t})] = \tilde{n}_n(t) = (1 + e^{\epsilon_n/k\tilde{\vartheta}})^{-1}. \quad \text{(A.II19)}$$

Now we assume that the transition rates in the actual cosmos can be approximated by the equilibrium transition rates, i.e.,

$$\alpha_n(t) = \tilde{\alpha}_n(t), \qquad \beta_n(t) = \tilde{\beta}_n(t), \qquad \tau_n(t) = \tilde{\tau}_n(t). \quad \text{(A.II20)}$$

Next we solve (A.II17b) by writing

$$\overline{n_n(t)} = \tilde{n}_n(t) + \overline{\delta n_n(t)} \quad \text{(A.II20a)}$$

and neglecting the time derivative of the last term,

$$(d/dt)\overline{\delta n_n(t)} = 0. \quad \text{(A.II20b)}$$

With $\vartheta(t) \approx \tilde{\vartheta}(t)$ and (A.II17a–20b) we obtain then

$$\overline{\delta n_n(t)} = -\tilde{\tau}_n(t)\{1 - \tilde{n}_n(t)\}(d/dt)\tilde{n}_n(t). \quad \text{(A.II21)}$$

Furthermore (A.II1–3, 12) and (I18) give

$$\frac{d}{dt}\tilde{n}_n(t) = -\frac{M^2c^4}{2\lambda}(t + t_0)^{-\frac{1}{2}}\frac{1}{\epsilon_n}\tilde{n}_n(t)\{1 - \tilde{n}_n(t)\}. \quad \text{(A.II22)}$$

From (A.II14) we obtain (E^+ and E^- give the same contribution)

$$\bar{E} = 2\int_0^{T/2} dt\, L(t) \sum_n \Pi_n(t)\, \overline{\delta n_n(t)}. \quad \text{(A.II23)}$$

With (A.II15, 5b, 21, 22) this gives the result

$$\bar{E} = \frac{(Mc^2)^4}{\lambda t_0^{\frac{3}{2}}}\int_0^{T/2}\frac{dt}{t + t_0}\sum_n \tilde{\tau}_n(t)\frac{1}{\epsilon_n^2}\tilde{n}_n(t)[1 - \tilde{n}_n(t)]^2. \quad \text{(A.II24)}$$

D. Fluctuations

Equation (A.II14) gives

$$E^+ - \overline{E^+} = \int_0^{T/2} dt\, L(t) \sum_n \Pi_n(t)\{n_n(t) - \overline{n_n(t)}\} \quad \text{(A.II25)}$$

and

$$\overline{(E^+ - \overline{E^+})^2} = \iint_0^{T/2} dt\, dt'\, L(t)\, L(t') \times \sum_{n,m} \Pi_n(t)\Pi_m(t')\{n_n(t) - \overline{n_n(t)}\}\{n_m(t') - \overline{n_m(t')}\}. \quad \text{(A.II26)}$$

To calculate the correlation in (A.II26) we can neglect the cosmos expansion, i.e., we can assume thermal equilibrium. Here one easily finds

$$\overline{\{n_n(t) - \tilde{n}_n(t)\}\{n_m(t') - \tilde{n}_m(t')\}} = \delta_{n,m}\tilde{n}_n(t)[1 - \tilde{n}_n(t)] \exp\{-|t - t'|(\alpha_n + \beta_n)\}. \quad \text{(A.II27)}$$

Using this correlation in (II.A26) we obtain with (A.II17a, 19)

$$\overline{(E^+ - \overline{E^+})^2} = 2\int_0^{T/2} dt\, L(t)^2 \sum_n \Pi_n^2 \tau_n \tilde{n}_n(t)[1 - \tilde{n}_n(t)]^2.$$

With (A.II5b, 14, 15) and the statistical independence of E^+ and E^- this gives

$$\overline{(E - \bar{E})^2} = \frac{4(Mc^2)^4}{t_0}\int_0^{T/2}\frac{dt}{t + t_0}\sum_n \tau_n \frac{1}{\epsilon_n^2}\tilde{n}_n(t) \times [1 - \tilde{n}_n(t)]^2. \quad \text{(A.II28)}$$

Comparison of (A.II24) and (A.II28) gives with (I18)

$$\overline{(E - \bar{E})^2}/\bar{E} = 4k\vartheta_0. \quad \text{(A.II29)}$$

E. The Explicit Expression for E

We have assumed so far that the cosmos is filled homogeneously with $N - \bar{N}$ pairs and radiation. In this case we can approximate the lifetime of a particle in state $\mathfrak{n}$ by

$$\tilde{\tau}_n = [\sigma_0 C(t) v_n]^{-1}, \quad \text{(A.II30)}$$

where $C(t)$ is the density of thermal N–$\bar{N}$ pairs, v_n is the velocity of a particle in state $\mathfrak{n}$, and σ_0 is the annihilation cross section which we assume to be velocity-independent.

Here the density of thermal pairs,

$$C(t) = V(t)^{-1}\sum_n \tilde{n}_n(t),$$

can be written with (A.II12, 1–3) as

$$C(t) = 8\pi\left(\frac{Mc}{h}\right)^3\frac{1}{\Theta^3}\int_\Theta^\infty d\eta\,\frac{\eta(\eta^2 - \Theta^2)^{\frac{1}{2}}}{1 + e^\eta} \quad \text{(A.II31)}$$

where we have introduced

$$\eta = \epsilon_n/k\vartheta, \qquad \Theta = Mc^2/k\vartheta. \tag{A.II32}$$

Furthermore

$$v_n^{-1} = \eta[c(\eta^2 - \Theta^2)^{\frac{1}{2}}]^{-1}. \tag{A.II33}$$

Now (A.II24) becomes with (A.II3, 30–33)

$$\bar{E} = \frac{(k\vartheta_0)^2}{Mc^2}\frac{2V_0}{t_0 c\sigma}\int_{\Theta_0}^{\Theta'} d\Theta\,\Theta^4\frac{J(\Theta)}{K(\Theta)}, \tag{A.II34}$$

where

$$J(\Theta) = \int_\Theta^\infty d\eta\,\frac{e^{2\eta}}{(1+e^\eta)^3} \tag{A.II35}$$

$$K(\Theta) = \int_\Theta^\infty d\eta\,\frac{\eta(\eta^2-\Theta^2)^{\frac{1}{2}}}{1+e^\eta}. \tag{A.II36}$$

Note that for $\Theta \gg 1$

$$J(\Theta) \to e^{-\Theta}, \qquad K(\Theta) \to (\tfrac{1}{2}\pi)^{\frac{1}{2}}\Theta^{\frac{1}{2}}e^{-\Theta}. \tag{A.II37}$$

The upper limit Θ' in the integral in (A.II34) is (in the cosmos without noncompensated particles) the Θ-value at maximal cosmos extension. From (A.II37) we see that the integral diverges for $\Theta' \to \infty$, i.e., the cosmos stages with very large Θ-values contribute considerably to $\bar{E}$.

Next we want to generalize the previous calculations to the case where one half-cosmos contains excess baryons and the other half excess antibaryons. In order to calculate $\bar{E}$ (and σ) here, we may consider each half-cosmos separately.

Let Θ_z be that Θ-value for which, in the cosmos half considered, the density of pairs equals the density of excess baryons. Then the previous calculations apply to the region with $\Theta \ll \Theta_z$, since here the concentration of the excess baryons is relatively small.

In the region $\Theta \gg \Theta_z$, however, the noncompensated particles drastically reduce the density of thermal pairs. Here the lifetime of an antibaryon is limited by these noncompensated particles.

Now the region $\Theta \gg \Theta_z$ gives practically no contribution to $\bar{E}$. Therefore we can take the influence of the noncompensated particles into account summarily (in good approximation) by choosing in (A.II34) the value $\Theta' = \Theta_z$ as upper limit for the integral.

We can assume that $\Theta_z \gg 1$. Then the main contribution to the integral in (A.II34) comes from the region with $\Theta \gg 1$, and we can use the approximations (A.II37) for calculating the integral. This gives

$$\bar{E} = \frac{4}{7}\left(\frac{2}{\pi}\right)^{\frac{1}{2}}\frac{(k\vartheta_0)^2}{Mc^2}\frac{V_0}{c\sigma_0 t_0}\Theta_z^{7/2}. \tag{A.II38}$$

Euclidean Quantum Field Theory. I. Equations for a Scalar Model*

K. SYMANZIK
Courant Institute of Mathematical Sciences, New York University, New York, New York
(Received 2 September 1965)

The analytic continuations to imaginary time of the Green's functions of local quantum field theory define Euclidean Green's functions. Use of the proper-time method allows to represent these functions as multiple Wiener integrals of functionals that obey infinite systems of coupled integral equations which are similar to, and for the particular model of a complex scalar field in quadrilinear self-interaction considered here a limiting case of, systems studied in quantum statistical mechanics by Ginibre. As a consequence, the Euclidean Green's functions can for this model be obtained by a limiting process, with temperature and density going to infinity, from the reduced density matrices of a nonrelativistic Bose gas. Reduced functionals are defined and their equations determined as a prepartory step to renormalization in the super-renormalizable cases of two and three dimensions.

INTRODUCTION

It is well known[1] that quantum field theory in Minkowski space (MQFT), if a Lagrangian is given, can be cast in the form of an infinite system of coupled integral equations[2] for the infinite set of Green's functions. These systems of equations have so far been of little help except for studying certain formal properties of Green's functions (e.g., properties under gauge transformations in quantum electrodynamics[3] or how to define a Bethe–Salpeter kernel without recourse to perturbation theory[1,4,5]). The main obstacle to a nonformal use of those systems is our inaptitude to formulate properly the boundary conditions on such systems to make them mathematically meaningful. Prescriptions on how to break such systems off have been given at times but seem so far lacking in convincing justification as well as success.

One feature of those equations is the already poor formulation of each single equation. For example, in their momentum-space form, one encounters even under the most favorable of circumstances only conditionally convergent integrals, while in coordinate space they involve products of distributions. Dyson[6] has shown that this difficulty is overcome in perturbation theory by a rotation of the paths of integration. More generally, one may use simultaneous analytic continuation[7] of all functions in the equations and define the original functions by the boundary values of their continuations. Continuing to imaginary times respectively energies yields the Euclidean Green's functions studied in their own right by Schwinger[8] and Nakano.[9] These functions can be defined even without reference to a Lagrangian and may be associated with a Euclidean quantum field theory (EQFT) whose characteristic symmetry group is not the Lorentz group but the orthogonal group in four dimensions.

EQFT is of no particular interest in an axiomatic framework,[10] since the axioms are formulated directly in MQFT terms and all of EQFT is secondary. If, however, a Lagrangian is given, the situation is quite different. Then the investigation of the resulting particular system of integral equations for EQFT Green's functions becomes useful since the existence of a solution is a necessary condition for a corresponding MQFT to exist, provided one supposes the MQFT to possess a lowest energy state as is done generally, and may be easier[11] to prove or disprove.

Compared to MQFT Green's functions systems, the ones for EQFT functions have the following advantages: (1) The EQFT functions are singular

* The research reported in this paper was supported by the Ford Foundation.

[1] J. Schwinger, Proc. Natl. Acad. Sci. U. S. **37**, 452, 455 (1951).

[2] We are not concerned here with the systems of equations that use no Lagrangian but intermediate-state insertions and the asymptotic condition, e.g., H. Lehmann, K. Symanzik, and W. Zimmermann, Nuovo Cimento **1**, 205 (1955).

[3] N. N. Bogoliubov and D. Shirkov, *Introduction to the Theory of Quantized Fields* (Interscience Publishers, Inc., New York, 1959), Sec. 40; B. Zumino, J. Math. Phys. **1**, 1 (1960).

[4] H. Umezawa and A. Visconti, Nuovo Cimento **1**, 1079 (1955).

[5] K. Symanzik, in *Lectures on High Energy Physics*, edited by B. Jakšić (Federal Nuclear Energy Commission of Yugoslavia, Zagreb, 1961), pp. 485–517.

[6] F. J. Dyson, Phys. Rev. **75**, 1736 (1949), cf. also N. Nakanishi, Progr. Theoret. Phys. (Kyoto) **17**, 401 (1958).

[7] F. J. Dyson, Phys. Rev. **83**, 608 (1951); G. C. Wick, *ibid.* **96**, 1124 (1954).

[8] J. Schwinger, Proc. Natl. Acad. Sci. U. S. **44**, 956 (1958); Phys. Rev. **115**, 721 (1959).

[9] T. Nakano, Progr. Theoret. Phys. (Kyoto) **21**, 241 (1959).

[10] A. S. Wightman, Phys. Rev. **101**, 860 (1956); R. Haag, B. Schroer, J. Math. Phys. **3**, 248 (1962).

[11] K. Symanzik, in *Analysis in Function Space*, edited by W. T. Martin and I. Segal (MIT Press, Cambridge, Massachusetts, 1964), p. 197; *A modified model of Euclidean Quantum Field Theory*, IMM-NYU 327 (CIMS, New York University, New York, June 1964).

only at coinciding arguments and do not have the light-cone singularities which are part of the origin of the ambiguities mentioned before. (2) Elliptic systems are more easily handled than hyperbolic systems. This expresses itself in the fact that for the analysis of EQFT relatively efficient mathematical tools are available, and this is the point of this series of papers. (3) There is an interesting and suggestive direct relation between EQFT and nonrelativistic quantum equilibrium statistical mechanics, which has no counterpart in MQFT. (4) The EQFT metric is always positive-definite even if the MQFT metric, as in manifestly covariant quantum electrodynamics, is not.

Ultraviolet difficulties are the same in EQFT as in MQFT, but in the first case manifest themselves in terms of divergent, instead of meaningless, integrals. Renormalization of the coupled system of integral equations may be performed either with the help of limiting processes[12] or by renouncing manifest locality.[13,5] Both these ways are not suitable for our present purpose. We shall first construct a formal solution of the coupled system of EQFT integral equations and will invoke an *ad hoc* regularization of this solution wherever this seems to be illuminating. Our procedure is to derive from this formal solution new integral equations which are renormalized by eliminating the renormalization constants and which are the basis for a constructive existence proof for a nonformal solution.

EQFT and MQFT are on a comparable level as far as phenomena like spontaneous symmetry breakdown and vacuum degeneracy are concerned: about these one learns as much from EQFT as from MQFT Green's functions, although the question whether, e.g., a symmetry breakdown holds for the S matrix also is directly answerable only on the basis of MQFT functions.

EQFT is clearly in a great disadvantage with respect to questions about observables, e.g., if there exists a particle interpretation and asymptotic completeness holds, what the scattering amplitudes are, etc. However, as Lagrangian MQFT has resisted so far any attempt to extract from it such information (except in terms of most untrustworthy[14] perturbation expansions, or for models with scattering amplitudes identically zero) the indirect EQFT approach can be defended.

[12] J. Valatin, Proc. Roy. Soc. (London) **A225**, 534 (1954); H. Rollnik, B. Stech, and E. Nunnemann, Z. Physik **159**, 482 (1960); K. Wilson, Phys. Rev. (to be published), W. Zimmermann (to be published).

[13] J. Schwinger, *Theory of coupled fields*, Harvard University, 1954 (unpublished).

[14] E.g., A. Jaffé, Commun. Math. Phys. **1**, 127 (1965).

In Sec. 1 we show in a general way how EQFT Green's functions are related to MQFT functions, and thereby derive some of their expected properties. This connection is, however, not needed for a study of EQFT itself.

In Sec. 2 we derive the starting equations for the model of a scalar complex field with quadrilinear self-coupling and review at this example an operator formulation of EQFT given elsewhere.[11]

In Sec. 3 we solve the equations formally, introduce an auxiliary intermediary field, and thereby obtain a form of solution which, if expanded, gives an expansion in increasing numbers of closed loops.

In Sec. 4 we obtain analogs of the Kirkwood–Salsburg and Mayer–Montroll integral equations[15,16] for distribution functions in classical statistical mechanics. The equations obtained closely resemble, and are for the model treated in this paper a limiting case of, equations used in nonrelativistic quantum statistical mechanics by Ginibre.[17] This is discussed in detail in Appendix A while in Sec. 4 itself the analogy to classical statistical mechanics is shown and exploited. Appendices B and C illustrate our equations and their properties in the lowest-dimensional cases of zero and one dimension, respectively, where no renormalization is needed.

In Sec. 5 we introduce, as preparatory to renormalization, reduced functionals, whereby in the super-renormalizable cases of two and three dimensions all terms that need be renormalized are collected in one simple equation. The renormalization of this equation by adaption and extension of a method due to Nelson[18] will be presented in the next paper of this series, together with the closely related treatment of derivative couplings as occur in scalar and two-component spinor quantum electrodynamics.

1. AXIOMATIC FORMULATION OF EQFT

For this chapter, we adopt the axiomatic approach to relativistic quantum field theory developed by Wightman.[19] We consider the theory of one Hermitian scalar field $A(x)$ only.

Due to the stability of the vacuum (denoted by

[15] T. L. Hill, *Statistical Mechanics* (McGraw-Hill Book Company, Inc., New York, 1956).

[16] J. L. Lebowitz and J. K. Percus, J. Math. Phys. **4**, 1495 (1963).

[17] J. Ginibre, J. Math. Phys. **6**, 238, 252, 1432 (1965). The author is indebted to Dr. Ginibre for preprints of his interesting work.

[18] E. Nelson, in *Analysis in Function Space*, edited by W. T. Martin and I. Segal (MIT Press, Cambridge, Massachusetts, 1964), p. 87.

[19] Reference 10 and: R. F. Streater and A. S. Wightman, *PCT, Spin and Statistics, and All That* (W. A. Benjamin, Inc., New York, 1964).

⟨ and ⟩), the spectrum condition, and their assumed temperedness as distributions, the vacuum expectation values

$$\langle A(x_0)A(x_1)\cdots A(x_n)\rangle, \qquad x_i = (x_i^0, x_i^1, x_i^2, x_i^3)$$

of products of field operators are, as functions of $\xi_i = x_{i-1} - x_i (i = 1 \cdots n)$, boundary values of analytic functions $W_n((\zeta))$, $(\zeta) = (\zeta_1 \cdots \zeta_n)$, with analyticity domain the tube

$$R_n = \{(\zeta) : \operatorname{Im} \zeta_i \in V^+, \forall i\},$$

i.e., $\operatorname{Im} \zeta_i^0 > 0$, $(\operatorname{Im} \zeta_i)^2 > 0$, with $g_{\mu\nu} = -\delta_{\mu\nu}(-1)^{\delta_{\mu 0}}$. Due to relativistic invariance, $W_n((\zeta))$ is analytic and single-valued in the extended tube

$$R_n' = \{(\zeta) : \exists\, \Lambda_+(C), (\zeta) = (\Lambda_+(C)\zeta'), (\zeta') \in R_n\}$$

where $\Lambda_+(C)$ is a proper homogeneous complex Lorentz transformation:

$$\Lambda_+(C)^T g\, \Lambda_+(C) = g, \qquad \det \Lambda_+(C) = 1.$$

Due to local commutativity, $W_n((\zeta))$ is analytic and single valued in the permuted extended tube

$$R_n'' = \bigcup_{\text{all } P} PR_n'$$

with

$$PR_n' = \{(\zeta) : (\zeta) = (P\zeta'), (\zeta') \in R_n'\}$$

where $P \in S_{n+1}$ is a permutation

$$P : (0, 1 \cdots n) \to (P(0)P(1) \cdots P(n)),$$

and if $\zeta_i = z_{i-1} - z_i$, then $P\zeta_i = z_{P(i-1)} - z_{P(i)}$. In R_n'', $W_n((P\zeta)) = W_n((\zeta))$ due to our use of one field only. The Schwinger points

$$(\zeta_S) : \operatorname{Re} \zeta_i^0 = 0, \qquad \operatorname{Im} \zeta_i^{1,2,3} = 0, \forall i$$

lie in the interior of R_n'' if $\zeta_{s_{i+1}} + \cdots + \zeta_{s_k} \neq 0$ for all $1 \leq i + 1 \leq k \leq n$. We may write

$$\zeta_{s_i} = x_{i-1} - x_i, \quad x = (x^1, x^2, x^3, x^4) \text{ real}, \quad x^4 = ix^0,$$

and introduce the Schwinger functions[8]

$$W_n((\zeta_s)) \equiv S(x_0 x_1 \cdots x_n).$$

These Euclidean Green's functions are symmetric functions of $(n + 1)$ 4-vector arguments, invariant under the proper inhomogeneous orthogonal group in four dimensions (here called the Euclidean group), and real-analytic except at points of coincidence of some arguments. (Their analytic continuations are invariant under the complex Euclidean group and are the original Wightman functions in different notation). They satisfy

$$S(x_0 \cdots x_n) = S(x_0^T \cdots x_n^T)^* = S(x_0^s \cdots x_n^s)^* = S(-x_0 \cdots -x_n) \qquad (1.1)$$

where

$$x^T = (x^1, x^2, x^3, -x^4), \; x^s = (-x^1, -x^2, -x^3, x^4)$$

and are, therefore, real if the theory is invariant under time reversal or space reflection.

The Green's functions

$$F(x_0 \cdots x_n) \equiv \langle TA(x_0) \cdots A(x_n)\rangle,$$

where T is the symbol for operator ordering with increasing times from right to left, are for noncoindicing arguments symmetric tempered Lorentz-invariant distributions. Assuming that these functions can be extended[20] to such distributions for all arguments, Ruelle[21] has shown that the Fourier transforms

$$\tilde{F}(p_1 \cdots p_n) = \int dx_1 \cdots dx_n\, e^{i\sum x_i p_i} F(0x_1 \cdots x_n)$$

are boundary values of analytic functions[22] which are invariant under the proper homogeneous Lorentz group. The Schwinger points $(p_s) : \operatorname{Im} p_i^{1,2,3} = 0$, $\operatorname{Re} p_i^0 = 0$, $p_i^0 = -ip_i^4$, $\forall i$ lie inside the analyticity domain except for points where a nonempty partial sum of the vectors p_{s_i} vanishes. We shall write

$$i^n \tilde{F}((p_s)) \equiv \tilde{S}(p_0 p_1 \cdots p_n), \quad p_0 = -p_1 - \cdots - p_n.$$

Then

$$(2\pi)^4 \tilde{S}(p_0 \cdots p_n)\, \delta(p_0 + \cdots + p_n) = \int dx_0 \cdots dx_n\, e^{-i\sum x_i p_i} S(x_0 \cdots x_n),$$

where $x_i p_i = x_i^1 p_i^1 + \cdots + x_i^4 p_i^4$. If truncated Wightman functions[24] W^T and truncated Green's functions[21] F^T are introduced, the functions $\tilde{F}^T$ have no singularities at Schwinger points. Therefore, the functions $\tilde{S}^T(p_0 \cdots p_n) = \tilde{F}^T((p_s))$ are symmetric real-analytic functions, invariant under the homogeneous proper orthogonal group, and satisfy

$$\tilde{S}^T(p_0 \cdots p_n) = \tilde{S}^T(p_0^T \cdots p_n^T)^* = \tilde{S}^T(p_0^s \cdots p_n^s)^* = \tilde{S}^T(-p_0 \cdots -p_n), \qquad (1.2)$$

with definitions analogous to those in (1). They possess analytic continuations into the tube

[20] See Ref. 21 for a precise statement of the assumption.
[21] D. Ruelle, thesis, Bruxelles (1959); Nuovo Cimento 19, 356 (1961).
[22] For descriptions of the domain of analyticity, see Refs. 21 and 23.
[23] H. Araki, J. Math. Phys. 2, 163 (1961).
[24] R. Haag, Phys. Rev. 112, 668 (1958).

$$(\mathrm{Im}\, p) \in D^m \equiv \bigcap_{\text{all } I} D_I^m, \qquad \text{with}$$

$$D_I^m = \left\{(\mathrm{Im}\, p) : \sum_{k=1}^{4} \left(\sum_{i\in I} \mathrm{Im}\, p_i^k\right)^2 < m^2\right\},$$

where I is a proper subset of $\{0, 1 \cdots n\}$ and $m > 0$ is the lower bound of the mass spectrum (except for the vacuum) of the theory. It follows that, provided

$$\min_{i\neq j} (x_i - x_j)^2 \geq \epsilon > 0,$$

$$\lim_{D\to\infty} S^T(x_0 \cdots x_n) \exp\left(\sum_{i=0}^{n} \alpha_i x_i\right) = 0, \qquad (\alpha) \in D^m, \tag{1.3}$$

where

$$\alpha_i x_i = \alpha_i^1 x_i^1 + \cdots + \alpha_i^4 x_i^4,\ D = (\max_{i,j}(x_i - x_j)^2)^{\frac{1}{2}}.$$

(1.3) shows the exponential decrease of $S^T(x_0 \cdots x_n)$ for increasing distance between its arguments.

Having established the existence of Euclidean Green's functions in every theory that satisfies Wightman's postulates, we will further on proceed more heuristically, which seems justified as no physically nontrivial example of a Wightman theory is known. Our goal is, in fact, to construct models for which the axiomatic assumptions can be verified and on this basis perhaps be sharpened.

2. A SCALAR MODEL

We consider the theory of one non-Hermitian scalar field in d space–time dimensions corresponding to the Lagrangian density

$$L = \partial^\mu B^\dagger\, \partial_\mu B - m^2 B^\dagger B - \tfrac{1}{2} g (B^\dagger B)^2 + \alpha B^\dagger B. \tag{2.1}$$

Here $\hbar = c = 1$, m is a finite mass that need not be the mass of a particle, g the positive coupling constant, and

$$\alpha = 2gG_0(0) + \delta m^2 \tag{2.2}$$

where $G_0(0)$ a (for $d \geqq 2$, infinite[25]) constant obtained from (2.6), and δm^2 another (for $d \geqq 3$, negative-infinite) constant determined in our next paper. The nonvanishing canonical commutators derived from (1) are

$$[B(\mathbf{x}, x^0), \dot{B}^\dagger(\mathbf{x}', x^0)] = [B^\dagger(\mathbf{x}, x^0), \dot{B}(\mathbf{x}', x^0)] = i\,\delta(\mathbf{x} - \mathbf{x}'). \tag{2.3}$$

We denote the Euclidean Green's functions derived from

$$\langle TB(x_1) \cdots B(x_m) B^\dagger(y_1) \cdots B^\dagger(y_n)\rangle$$

as described in Sec. 1 by $S(x_1 \cdots x_m, y_1 \cdots y_n)$, and their generating functional[26] by

$$S[\bar{J}, J] = \sum_{m=n=0}^{\infty} (m!\, n!)^{-1} \int \cdots \int dx_1 \cdots dx_m\, dy_1 \cdots dy_n \times \bar{J}(x_1) \cdots \bar{J}(x_n) J(y_1) \cdots J(y_n) \times S(x_1 \cdots x_m, y_1 \cdots y_n), \tag{2.4}$$

where $\bar{J}(x)$ and $J(x)$ are independent functions with algebraic meaning only. One can show[11] that[27]

$$S[\bar{J}, J] = \left\langle T_\tau \exp \int_{-\infty}^{+\infty} d\tau \times \int dx [\bar{J}(\mathbf{x}, \tau) B(\mathbf{x}, \tau) + B^\dagger(\mathbf{x}, \tau) J(\mathbf{x}, \tau)] \right\rangle, \tag{2.5}$$

where

$$B(\mathbf{x}, \tau) = e^{H\tau} B(\mathbf{x}, 0) e^{-H\tau},$$
$$B^\dagger(\mathbf{x}, \tau) = e^{H\tau} B^\dagger(\mathbf{x}, 0) e^{-H\tau} \neq [B(\mathbf{x}, \tau)]^\dagger,$$

and T_τ means ordering with increasing τ from right to left. (Note that the orthogonal invariance of the left-hand side of (2.5) is not manifest on the right.)

Field equations and canonical commutation relations to (1) give differential equations for the Green's functions, and analytic continuation results in the functional differential equations (we suppress the common argument x)

$$(-\Delta + m^2)[\delta/(\delta J)]S + g[\delta^3/(\delta J^2\, \delta \bar{J})]S - \alpha[\delta/\delta J]S = \bar{J}\, S, \tag{2.6a}$$

$$(-\Delta + m^2)[\delta/(\delta \bar{J})]S + g[\delta^3/(\delta J\, \delta \bar{J}^2)]S - \alpha[\delta/\delta \bar{J}]S = J\, S, \tag{2.6b}$$

where Δ is the Laplacian in d dimensions. Integrating (6) with the elliptic Green's function[28]

$$G_0(x - y) = (2\pi)^{-d} \int dk\, e^{ikx} (k^2 + m^2)^{-1} \tag{2.7}$$

(here and in the following we use the scalar product

[25] At this point we shall not discuss possibilities to formulate the theory less objectionably. See Introduction and Sec. 4.

[26] If the formal invariance of (2.1) under gauge transformation of the first kind is not violated by the vacuum, Green's functions with $m \neq n$ are zero. No assumption to this effect is needed in this section.

[27] In this and other formulas that do not make reference to (2.1) or to canonical commutation relations, the field operators should be the renormalized ones whenever the amplitude renormalization is not finite; then the formulas remain valid also for $d = 4$.

[28] Integration with the Green's function that satisfies periodic boundary conditions in one coordinate would lead to finite-temperature Green's functions; see, e.g., A. A. Abrikosov, L. P. Gorkov, and I. E. Dzyaloshinski, *Methods of Quantum Field Theory in Statistical Physics* (Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1963).

$kx = \sum_{\alpha=1}^{d} k_\alpha x_\alpha$) gives vanishing boundary terms according to (1.3).

Interpreting J and $\bar{J}$ realistically as numerically-valued functions conjugate complex to each other, we may introduce the Hermitian operator

$$H_\tau = H - \int d\mathbf{x}[\bar{J}(\mathbf{x}, \tau)B(\mathbf{x}, 0) + B^\dagger(\mathbf{x}, 0)J(\mathbf{x}, \tau)]$$

where H is the canonical Hamiltonian to (1), adjusted such that $H\,\rangle = 0$, or, more generally, the time-displacement operator. Then

$$S[\bar{J}, J] = \left\langle T_\tau \exp\left[-\int_{-\infty}^{+\infty} H_\tau\, d\tau\right]\right\rangle \tag{2.8}$$

and

$$\left\|T_\tau \exp\left[-\int_{-\infty}^{+\infty} H_\tau\, d\tau\right]\right\| \leq \exp\left[-\int_{-\infty}^{+\infty} d\tau E_0(J_\tau)\right],$$

where $E_0(J_\tau)$ is the ground-state energy to H_τ.

Using arguments based on functional integration, it can be shown[11] that (provided certain limiting processes are reasonably behaved as they are in renormalized perturbation theory) for two sets J_i and J_i' of k complex functions and k complex constants C_i

$$\sum_{i,j=1}^{k} \bar{C}_i C_j S[\bar{J}_i + \bar{J}_j', J_j + J_i'] \geq 0. \tag{2.9}$$

This property[29] allows[30] to give an operator formulation of EQFT, which for the present model takes the form: choose two pairs of canonically conjugate Hermitian field operators in d dimensions ($i = 1, 2$) $q_i(x)$, $p_i(x)$ such that

$$[q_i(x), q_j(x')] = [p_i(x), p_j(x')] = 0,$$
$$[q_i(x), p_j(x')] = i\,\delta_{ij}\,\delta(x - x'),$$

and (with summation convention) the Hamiltonian

$$H = \frac{1}{2}\int dx\, C_i^\dagger(x) C_i(x) \tag{2.10}$$

where

$$C_i = p_i - \tfrac{1}{2} i(-\Delta + m^2)q_i - \tfrac{1}{4} i g q_i q_i q_i + \tfrac{1}{2} i\alpha q_i.$$

Then, with $>$ the state of lowest energy [which satisfies $C_i(x)\,> = 0$], we have

$$S[\bar{J}, J] = \left\langle \exp\left(2^{-\frac{1}{2}}\int dx[\bar{J}(q_1 + iq_2) + J(q_1 - iq_2)]\right)\right\rangle \tag{2.11}$$

as the generating functional of the equal-time ground-state expectation values of operator products. We have the momentum operators ($\alpha = 1 \cdots d$)

$$P_\alpha = \int dx\, p_i(x)\,\partial_\alpha q_i(x) = -\int dx\, q_i(x)\,\partial_\alpha p_i(x)$$

and (at least formally) the charge operator

$$Q = \int dx[q_1(x)p_2(x) - q_2(x)p_1(x)]$$

with the properties

$$[P_\alpha, q_i(x)] = -i\,\partial_\alpha q_i(x), \quad [P_\alpha, p_i(x)] = -i\,\partial_\alpha p_i(x),$$
$$[Q, q_{1,2}(x)] = \pm i q_{2,1}(x), \quad [Q, p_{1,2}(x)] = \pm i p_{2,1}(x),$$
$$[Q, H] = [Q, P_\alpha] = [H, P_\alpha] = [P_\alpha, P_\beta] = 0,$$

and the usual (Euclidean) invariance properties of the ground state.

This Hamiltonian theory[31] is of the general type studied by Araki,[32] and the cluster property (1.3) (extended to two fields) finds here its natural place.[33] Due to absence of Lorentz invariance, locality holds only "nonrelativistically". This theory is subjected to Haag's theorem[34] even for $d = 1$, when the corresponding "MQFT" describes only the anharmonic oscillator and thus is physically trivial. We discuss this theory in Appendix C.

3. FORMAL SOLUTION AND LOOP EXPANSION

Equation (2.6), integrated with (2.7), is formally solved by[35]

$$S[\bar{J}, J] = C \exp\left(-\tfrac{1}{2} g[\delta^4/(\delta J)^2(\delta\bar{J})^2] + \alpha[\delta^2/\delta J\,\delta\bar{J}]\right) \times \exp([\bar{J}G_0 J]) \tag{3.1}$$

where we suppress the obvious integrations over d-dimensional space in the square-bracketed terms.

[29] By special choices of the C_i and J_i one can derive from (2.9) inequalities that resemble the positive-definiteness conditions for Wightman functions.[19] However, there is no simple interrelation since the present inequalities do not require the MQFT metric to be positive-definite. Because of their generality, they might hold even in axiomatically introduced EQFT but we have no proof of this.

[30] M. A. Neumark, *Normierte Algebren* (VEB Deutscher Verlag der Wissenschaften, Berlin, 1959).

[31] The original canonical commutation relations (2.3) find their expression only in the discontinuities of (space) derivatives of the Euclidean Green's functions for coinciding arguments; see Ref. 11.

[32] H. Araki, J. Math Phys. 1, 492 (1960).

[33] H. Araki, Ann. Phys. (N. Y.) 11, 260 (1960).

[34] R. Haag, Kgl. Danske Videnskab. Selskab, Mat. Fys. Medd. 29, No. 12 (1955). See also: R. F. Streater and A. S. Wightman, Ref. 19.

[35] This form will be seen to exclude Green's functions that do not obey gauge invariance of the first kind. To admit the nonvanishing of those, one would in (3.1) have to introduce the fictitious symmetry-violating terms used by Bogoliubov (Rochester Conference, 1960). We will not pursue this possibility here. See also the discussion of symmetry-breaking in Sec. 4.

C has to be chosen such that $S[0, 0] = 1$. We now use

$$\exp(-\tfrac{1}{2}g[\delta^2/\delta\psi^2])\exp([f\psi])\,|_{\psi=0} = \exp(-[\tfrac{1}{2}f^2]) \tag{3.2}$$

to write (1) as

$$S[\bar{J}, J] = C\exp(-\tfrac{1}{2}g[\delta^2/\delta\psi^2]) \times \exp([(\psi+\alpha)\,\delta^2/(\delta\bar{J}\,\delta J)])\exp([\bar{J}G_0J])\,|_{\psi=0}$$

which can be evaluated[36] to be

$$S[\bar{J}, J] = C\exp(-\tfrac{1}{2}g[\delta^2/\delta\psi^2]) \times \exp([\bar{J}(G_0^{-1}-\alpha-\psi)^{-1}J] - \operatorname{Tr}\ln(1-(\alpha+\psi)G_0))\,|_{\psi=0}. \tag{3.3}$$

In order to be able to use (2) again, we introduce the representations

$$A^{-1} = \frac{1}{2}\int_0^\infty ds\exp(-\tfrac{1}{2}sA) \tag{3.4a}$$

and

$$\ln A - \ln B = \int_0^\infty s^{-1}\,ds \times [\exp(-\tfrac{1}{2}Bs) - \exp(-\tfrac{1}{2}As)], \tag{3.4b}$$

where the parameter s is called "proper time".[37] The solution that vanishes in infinity of the parabolic differential equation

$$(\partial/\partial t)U(x, y, t) = [\tfrac{1}{2}\Delta - V(x, t)]U(x, y, t); \tag{3.5a}$$

$$U(x, y, +0) = \delta(x-y) \tag{3.5b}$$

is the Wiener integral[38]

$$U(x, y, t) = \int P^t_{xy}(d\omega)\exp\left[-\int_0^t d\tau V(x(\tau), \tau)\right], \tag{3.6}$$

where $P^t_{xy}(d\omega)$ is the conditional Wiener measure on continuous paths $x(\tau)$ starting at $\tau = 0$ at y, ending at $\tau = t$ at x, and parametrized by ω.

$$\int P^t_{xy}(d\omega) = (2\pi t)^{-\frac{1}{2}d}\exp[-\tfrac{1}{2}t^{-1}(x-y)^2] \tag{3.7}$$

is the fundamental solution of the heat equation in d-dimensional space.[39]

Expanding (3) and using (4), (6), and (2) we find $S(x_1\cdots x_m, y_1\cdots y_n) = 0$ unless $m = n$, and

$$S(x_1\cdots x_n, y_1\cdots y_n) = \sum_{\Pi\in S_n} S(x_1y_{\Pi(1)}, \cdots, x_ny_{\Pi(n)}), \tag{3.8a}$$

where

$$S(x_1y_1, \cdots, x_ny_n) = 2^{-n}\int_0^\infty\cdots\int_0^\infty ds_1\cdots ds_n \times \exp(-\tfrac{1}{2}m^2(s_1+\cdots+s_n)) \times \int\cdots\int P^{s_1}_{x_1y_1}(d\omega_1)\cdots P^{s_n}_{x_ny_n}(d\omega_n) \times n(\omega_1\cdots\omega_n) \tag{3.8b}$$

with

$$n(\omega_1\cdots\omega_n) = C'\exp(-\tfrac{1}{2}g[\delta^2/\delta\psi^2])\exp\left(\tfrac{1}{2}\alpha(s_1+\cdots+s_n) + \sum_{i=1}^n\frac{1}{2}\int_0^{s_i}\psi(x_i(\sigma_i))\,d\sigma_i + \int_0^\infty t^{-1}\,dt \times \exp(-\tfrac{1}{2}m^2t + \tfrac{1}{2}\alpha t)\int dz\int P^t_{zz}(d\omega)\exp\left[\frac{1}{2}\int_0^t\psi\{z(\tau)\}\,d\tau\right]\right)\Bigg|_{\psi=0}, \tag{3.9a}$$

and by further expansion

$$n(\omega_1\cdots\omega_n) = C'\sum_{l=0}^\infty (l!)^{-1}\prod_{i=1}^l\left(\int_0^\infty t_i^{-1}\,dt_i\exp(-\tfrac{1}{2}m^2t_i + \tfrac{1}{2}\alpha t_i)\int dz_i\int P^{t_i}_{z_iz_i}(d\tilde{\omega}_i)\right) \times \exp\left(\tfrac{1}{2}\alpha(s_1+\cdots+s_n) - \frac{g}{8}\sum_{i=1}^n\int_0^{s_i}\int_0^{s_i}d\sigma_i\,d\sigma_i'\,\delta\{x_i(\sigma_i)-x_i(\sigma_i')\} - \frac{g}{4}\sum_{i<i'}\int_0^{s_i}\int_0^{s_{i'}}d\sigma_i\,d\sigma_{i'}\,\delta\{x_i(\sigma_i)-x_{i'}(\sigma_{i'})\} - \frac{g}{8}\sum_{i=1}^l\int_0^{t_i}\int_0^{t_i}d\tau_i\,d\tau_i'\,\delta\{z_i(\tau_i)-z_i(\tau_i')\} - \frac{g}{4}\sum_{i<i'}\int_0^{t_i}\int_0^{t_{i'}}d\tau_i\,d\tau_{i'}\,\delta\{z_i(\tau_i)-z_{i'}(\tau_{i'})\} - \frac{g}{4}\sum_{i,i}\int_0^{s_i}\int_0^{t_i}d\sigma_i\,d\tau_i\,\delta\{x_i(\sigma_i)-z_i(\tau_i)\}\right). \tag{3.9b}$$

[36] See, e.g., K. Yamazaki, Progr. Theoret. Phys. (Kyoto) 7, 449 (1952).
[37] V. Fock, Phys. Z. Sowjetunion 12, 404 (1937).
[38] See, e.g., Ref. 17 and further literature cited there.
[39] The Wiener process periodic in one coordinate (see Ref. 28) seems not to have been considered so far.
Note added in proof. E. Nelson has pointed out to the author that this is conditional Brownian motion on a torus.

The sum in (3.8a) goes over the $n!$ permutations of n elements, and C' differs from C by a factor and must be such that $n(\phi) = 1$. We shall refer to the sum in (3.9b) as the loop expansion (LE) and, combining in the last exponent the first two terms, abbreviate that exponent as

$$\sum (-\tfrac{1}{2}V_{ii} - V_{ii'} - \tfrac{1}{2}V_{jj} - V_{jj'} - V_{ij}).$$

The interpretation of (3.8) with (3.9b) in terms of n open arcs and l closed loops, which are contact-connected in themselves and with each other in all possible ways (if also the last exponential is expanded), is the same as given in quantum electrodynamics (QED) by Feynman.[40] While in the latter case the intermediary field is the electromagnetic one, here it is the field ψ that mediated contact interactions only and permitted us to obtain the combinatorics of QED, which is simpler than that of the quadrilinear case.

The principal difference between (3.9b) and the expressions used by Feynman is that (3.9) contains the perfectly well-defined Wiener integral and thus allows a rigorous discussion, while the "integrals" used instead in QED or, as we could attempt here, in scalar MQFT [by letting in (3.8) and (3.9) $x_4 \to ix^0$, $s \to is$, $t \to it$ etc.] are not integrals. Rather, they would be symbols associated with an exponential representation of the solution of the differential equation (Δ being here the Laplacian in $d - 1$ dimensions)

$$(\partial/\partial t)U(x, y, t) = i[\tfrac{1}{2}\Delta - \tfrac{1}{2}\partial_{x^0}^2 - V(x, t)]U(x, y, t),$$

with $U(x, y, 0) = \delta(x - y)$, where for $V \equiv 0$

$$U(x, y, t) = i(2\pi it)^{-\frac{1}{2}d} \times \exp\{i(2t)^{-1}[(\mathbf{x} - \mathbf{y})^2 - (x^0 - y^0)^2]\}. \quad (3.10)$$

Since this fundamental solution is not positive, no measure can be obtained from it[41] nor useful estimates derived.

If instead of (1) we would have chosen the Lagrangian density for a Hermitian scalar field

$$L = \tfrac{1}{2}[\partial^\mu A\, \partial_\mu A - m^2A^2 - \tfrac{1}{2}gA^4 + \alpha A^2] \quad (3.11)$$

with $\alpha = 3gG_0(0) + \delta m^2$, we would, under slight change of (3.8a) but none of (3.8b), have obtained (3.9) again with g replaced by $2g$ and a factor one-half for each t-integral. Thus, all equations we shall derive can be immediately transcribed[35] into this case. In contrast, if, e.g., a Lagrangian density for two scalar fields in trilinear interaction

$$L = \partial^\mu B^\dagger\, \partial_\mu B + \tfrac{1}{2}\,\partial^\mu A\, \partial_\mu A - M^2B^\dagger B - \tfrac{1}{2}m^2A^2 - gAB^\dagger B + \alpha B^\dagger B + \tfrac{1}{2}\beta A^2 + \gamma A \quad (3.12)$$

is chosen, for the Green's functions that do not contain the A field, similar formulas as before are obtained, with the replacements

$$-\frac{g}{8}\int_0^s \int_0^s d\sigma\, d\sigma'\, \delta\{x(\sigma) - x(\sigma')\} \to \tfrac{1}{2}g^2 \int_0^s \int_0^s d\sigma\, d\sigma' G_0\{x(\sigma) - x(\sigma')\},$$

etc., i.e., instead of a positive contact interaction we obtain a negative singular finite-range interaction. We show in Appendix A that this excludes at least the possibility of obtaining such EQFT by a limiting process from nonrelativistic quantum statistical mechanics (QSM) as described there. In fact, the MQFT to (3.12) is suspected[42] (and for $d = 2$ proved[43]) not to possess a translation-invariant lowest energy state,[44] which also deprives the formally corresponding EQFT of its basis,[45] since for the transition from MQFT to EQFT the spectral condition[10] is crucial.

The LE behaves very differently from perturbation theory obtained by expanding (3.1). As given in (3.9b) the LE has no meaning yet.[46] If, however, we introduce an *ad hoc* regularization that (a) replaces the delta functions by smooth integrable functions, (b) replaces[47] the lower limits zero of the s- and t- integrals by $\epsilon > 0$, (c) replaces the upper limits by $E < \infty$, and (d) replaces the infinite z-integration volume by a finite one, then the LE converges like an exponential series since all terms are then trivally majorized. In contrast, the usual perturbation expansion does not converge even under such drastic modifications[11] of the model.

The modifications just described give

$$n(\omega_1 \cdots \omega_n) > 0.$$

[40] R. P. Feynman, Phys. Rev. **80**, 440 (1950).

[41] The formal integral is also not a Feynman path integral, but using semigroup theory as for the latter [see E. Nelson, J. Math. Phys. 5, 332 (1964)] a convergent approximation procedure can still be given for a certain class of integrands, which does not, however, include the case at hand.

[42] G. Baym, Phys. Rev. **117**, 886 (1960).

[43] A. S. Wightman, *Introduction to Some Aspects of the Relativistic Dynamics of Quantized Fields* [Institut des Hautes Etudes Scientifiques, Bures-sur-Yvette, France, 1964]. Cf. also: A. Galindo, Proc. Natl. Acad. Sci. U. S. **48**, 1128 (1962).

[44] K. Mano [Progr. Theoret. Phys. (Kyoto) **14**, 434 (1955)] has argued that the energy of the system of one relativistic scalar nucleon coupled to a scalar meson field is, in spite of usual renormalization, not bounded below. In the present discussion based on the relation to QSM, closed loops are not neglected and the number of dimensions (it may even be one, cf. Appendix C) plays no role.

[45] The same holds if in (2.1) g is chosen negative instead of positive.

[46] E.g., we prove in the next paper that for $d \geq 2$ V_{jj} is almost everywhere infinite on Wiener space.

[47] Such regularization is actually gauge-invariant in QED.

If, upon gradual removal of the modifications, the functions $n(\omega_1 \cdots \omega_n)$ approach the ones of the unmodified model, $n(\omega_1 \cdots \omega_n) \geq 0$ must hold for these and because of (8) for $S(x_1 \cdots x_n, y_1 \cdots y_n)$ also. That such approach takes place is probable on the basis of earlier results[11] and the (at least in perturbation theory) known insensitivity of renormalizable theories against the manner of regularization. Another class of regularizations applicable to the present model is that described in Appendix A and gives the same result $n(\omega_1 \cdots \omega_n) \geq 0$.

As the space–time volume goes to infinity, C' becomes infinite (or zero, depending on the details of the regularization) since it depends on that volume exponentially.[11] The problem to show that the $n(\omega_1 \cdots \omega_n)$ have limits is, however, essentially the same as to show in classical statistical mechanics that the thermodynamical limit exists for distribution functions. The classical methods can, therefore, be applied.

4. KIRKWOOD–SALSBURG AND MAYER–MONTROLL INTEGRAL EQUATIONS

We single out in (3.9a) the path ω_1 from the others. Using, with the abbreviations introduced after (3.9b),

$$\exp\left(-\tfrac{1}{2}g[\delta^2/(\delta\psi)^2]\exp\frac{1}{2}\int_0^{s_1} d\sigma_1\ \psi[x_1(\sigma_1)]\right)$$

$$= \exp\left(-\tfrac{1}{2}\alpha s_1\right)\exp\left(\frac{1}{2}\int_0^{s_1} d\sigma_1\ \psi[x_1(\sigma_1)]\right)\exp\left(-\tfrac{1}{2}g[\delta^2/(\delta\psi)^2] - \tfrac{1}{2}g\int_0^{s_1} d\sigma_1\ \delta/\delta\psi\ [x_1(\sigma_1)] - \tfrac{1}{2}V_{11}\right),$$

$$\exp\left(-\tfrac{1}{2}g\int_0^{s_1} d\sigma_1\ \delta/\delta\psi\ [x_1(\sigma_1)]\right)\exp\left(\frac{1}{2}\int_0^{s_i} d\sigma_i\ \psi[x_i(\sigma_i)]\right) = \exp\left(\frac{1}{2}\int_0^{s_i} d\sigma_i\ \psi[x_i(\sigma_i)] - V_{1i}\right),$$

and

$$\exp\left(-\tfrac{1}{2}g\int_0^{s_1} d\sigma_1\ \delta/\delta\psi\ [x_1(\sigma_1)]\right)\exp\left(\frac{1}{2}\int_0^{t} d\tau\ \psi[z(\tau)]\right)$$

$$= \exp\left(\frac{1}{2}\int_0^{t} d\tau\ \psi[z(\tau)]\right)[1 - K(\omega_1, \bar{\omega})]\exp\left(-\tfrac{1}{2}g\int_0^{s_1} d\sigma_1\ \delta/\delta\psi\ [x_1(\sigma_1)]\right),$$

where $K(\omega, \bar{\omega})$ is the "bond" functional

$$K(\omega,\bar{\omega}) = 1 - \exp\left(-\frac{g}{4}\int_0^{s} d\sigma \int_0^{t} d\tau\ \delta[x(\sigma) - z(\tau)]\right), \tag{4.1}$$

we find, by expanding in powers of the bond functional and comparing with the definition (3.9a) of $n(\omega_1 \cdots \omega_n)$, with $n(\phi) = 1$ and empty products being one,

$$n(\omega_1 \cdots \omega_n) = \exp\left(-\tfrac{1}{2}V_{11} - \sum_{i=2}^{n} V_{1i}\right)$$
$$\times \sum_{l=0}^{\infty}(-1)^l(l!)^{-1}\prod_{i=1}^{l}\left[\int_0^{\infty} t_i^{-1}\,dt_i\right.$$
$$\times \exp\left(-\tfrac{1}{2}m^2 t_i\right)\int dz_i \int P^{t_i}_{z_i z_i}(d\bar{\omega}_i)K(\omega_1, \bar{\omega}_i)\Big]$$
$$\times n(\omega_2 \cdots \omega_n\, \bar{\omega}_1 \cdots \bar{\omega}_l), \tag{4.2}$$

the analog of the equation of Kirkwood and Salsburg (KS)[15] for distribution functions in the grand canonical ensemble of classical statistical mechanics (CSM).

Either iterating (4.2) $n - 1$ times or, more conveniently, dealing with all trajectories $\omega_1 \cdots \omega_n$ as done above with ω_1, we obtain the analog of the equation of Mayer and Montroll (MM)[16]

$$n(\omega_1 \cdots \omega_n) = \exp\left[-\frac{1}{2}\sum_{i=1}^{n} V_{ii} - \sum_{i<i'} V_{ii'}\right]$$
$$\times \sum_{l=0}^{\infty}(-1)^l(l!)^{-1}\prod_{i=1}^{l}\left[\int_0^{\infty} t_i^{-1}\,dt_i\right.$$
$$\times \exp(-\tfrac{1}{2}m^2 t_i)\int dz_i \int P^{t_i}_{z_i z_i}(d\bar{\omega}_i)K(\omega_1 \cdots \omega_n, \bar{\omega}_i)\Big]$$
$$\times n(\bar{\omega}_1 \cdots \bar{\omega}_l), \tag{4.3}$$

where

$$K(\omega_1 \cdots \omega_n, \bar{\omega}_i) = 1 - \exp\left[-\sum_{i=1}^{n} V_{ii}\right]. \tag{4.4}$$

The relation of (4.2) and (4.3) to equations derived in QSM by Ginibre[17] is discussed in Appendix

A. Here we rather draw a parallel to CSM. In (4.2) and (4.3) the Wiener integration is partly redudant.[48] We may define

$$\int_0^\infty t^{-1}\,dt\,\exp(-\tfrac{1}{2}m^2t)\int dz\int P^t_{zz}(d\omega)$$
$$\times F\Big\{\int_0^t d\tau\;\delta[\cdot - z(\tau)]\Big\} \equiv \int Q(db)\,F\{b\}, \qquad (4.5a)$$

where $F\{b\}$ is a functional defined on the space of nonnegative integrable functions[49]

$$b(z') = \int_0^t d\tau\;\delta(z' - z(\tau)), \qquad (4.5b)$$

which are such that $b(z')dz'$ is the time the Brownian particle executing the closed-path motion $z(\tau)$ spends in the volume dz', and $Q(db)$ is a (nonfinite[50]) measure of that space. We will refer to the "blob" picture and the "blob" measure. The functionals $n(\cdots)$ depend only on blobs and may be written as functionals $\bar{n}(b_1 \cdots b_n)$, or briefly $n(1 \cdots n)$, of blobs, since

$$s_i = \int dx\;b_i(x) \equiv ||b_i||, \quad V_{ij} = \frac{g}{4}\int dz\;b_i(z)b_j(z), \qquad (4.5c)$$

i.e., the "interaction potential" between blobs is determined by the degree of overlap. If we consider a blob as an internal degree of freedom of a particle, (4.2) and (4.3) become (except for the self-potential of a blob) identical with the KS and MM equations of CSM, with the integration over the internal degree of freedom.

We shall find in the next paper that renormalization can be simply expressed only in the Wiener picture (although the final formulas can be transcribed into the blob picture) since a divergence arises[51] only if the Brownian particle stays in the environment of a point instead of only returning to it at a later time. However, the blob picture is sometimes convenient and is the basis of our discussing (4.2) and (4.3) in the following in terms of CSM.

The derivation of the KS and MM equations from the modified LE described in Sec. 3 can be criticized on the following counts: (a) In QFT the volume is intrinsically infinite; (b) the *ad hoc* regularizations are not justified but (presumedly) correctly renormalized MQFT should be started from; (c) the infinite sums in (4.2) and (4.3) may not converge absolutely after partial (or later full) removal of the modifications; (d) (3.1) and (3.9) are formally summed perturbation theory and, therefore, not a satisfactory starting point if nonperturbation theoretical phenomena, like symmetry breakdown or bound states,[52] are to be expected.

(a) and (b) cannot be answered convincingly at this stage but one may argue: (a) The infinite-volume problem is present in CSM in reverse: actual systems are finite. Nevertheless, condensation, mathematically possible only in an infinite volume, is observable in finite containers. Thus, what is observable does not depend qualitatively on the size of the system (provided it is large enough), except for the following: in strictly infinite space, the observer, at a fixed point, will (with probability one) stay forever either in gas or in liquid even if both coexist, and any convex linear combination of the distribution functions solves the KS equations. (b) It is possible to build more familiar regularizations into the equations. This requires the introduction of noncontinuous paths, however, since the continuity of paths is consistent only with the Wiener measure and thus leads to $G_0(0)$-divergence due to "tadpoles." It is not obvious how to extend the renormalization procedure that will be given later to such cases. The same applies to renormalization by limiting processes in field equations themselves.[12] Therefore we shall be content to examine a constructive solution of the renormalized equations on its relation to renormalized EQFT[53] later. (c) This must be checked later and, if necessary, a summation prescription be given and justified.[54] In CSM, the corresponding summation is not considered a problem: often a hard-core potential (for which perturbation theory is meaningless altogether) of strictly finite range is assumed or

[48] The corresponding redundancy in QSM, Ref. 17, is slight.

[49] $b(z')$ is a functional of the Wiener process $z(\tau)$, specifically, the occupation-time distribution.

[50] The integral (4.5a) will exist (i.e., not be "ultraviolet divergent") only if $F\{b\}$ vanishes sufficiently strongly for $||b|| \to 0$.

[51] The divergence mentioned in Ref. 46 implies that (in blob measure) almost no blob function is square-integrable.

[52] This objection was raised by H. Stumpf at the Seminar on Unified Theories of Elementary Particles (Feldafing, July 1965) where the material of this paper was presented.

[53] Note that a formulation of renormalized EQFT by the usual coupled integral equations for Green's functions need not define the theory uniquely while the KS and MM equations may incorporate the missing boundary conditions. This is in fact so for the model discussed in Appendix B.

[54] Such possibilities are suggested, e.g., by Ref. 16 and: O. Penrose, J. Math Phys. 4, 1488 (1963). For $d = 0$, Appendix B, we find absolute convergence always, and for $d = 1$, Appendix C, absolute convergence as far as we can show unique solvability.

Note added in proof. For $d = 1$ absolute convergence holds under wider conditions. For $d = 2$ and $d = 3$, only the equations of Sec. 5 can be used.

at least considered[55] for almost all phenomena a permissible approximation, and then the sums are finite. (d) The last remark is relevant also here: the finite-sum KS equations are (at least for finite volume) rigorous[56] and unproblematic, and nevertheless expected[55] to describe phase transitions like condensation and crystallization, i.e., have in the latter case a "broken-symmetry" solution. Thus, noninvariant[57] and, more generally, nonperturbative behavior of solutions is not incompatible with an invariance an equation may possess, as was first proposed in field theoretical context by Heisenberg and is now generally accepted. Of course, the attempted method of solution must not exclude such behavior.

In the CSM interpretation of (4.2) and (4.3) the coupling constant g stands at the place of $\beta = (\mathrm{kT})^{-1}$, and $-m^2$ plays a role comparable to that of the chemical potential. Thus, from results on the analytic properties of distribution functions in CSM[58] one infers analyticity in both these variables, especially, for fixed real m^2 analyticity in g in an environment of the positive real axis. This inference will be verified for the simple cases $d = 0$ and $d = 1$ in Appendices B and C, respectively.

Equations (4.2) and (4.3) may be written

$$N = N_0 + \mathrm{Op}\, N, \tag{4.6}$$

where N is the vector $(n(\omega_1), n(\omega_1, \omega_2), \cdots)$ or $(n(1), n(12), \cdots)$, Op a (Wiener or blob) integral operator, and N_0 the contribution from $n(\phi) = 1$, written separately. Simple estimates now show that under the modifications described in Sec. 3, but with infinite instead of finite space–time, (4.2) and (4.3), or (4.6), have in a suitable Banach space unique solutions that can be obtained from the inhomogenous term of (4.6) by iteration. Since the technique hereto is also implicit in Appendix C, we will not discuss this further.

For $d = 0$ and $d = 1$, α is to be chosen finite and the KS or MM equations can be used as they stand. This is done in Appendices B and C. For $d = 2$ and $d = 3$, the only divergence expected is that α will have to be infinite as (2.2) suggests. For $d = 4$, also amplitude- and coupling-constant renormalization will be necessary. Since these renormalizations are more complicated, we confine ourselves now to $d = 3$ where renormalization leads to equations that are relatively simple and can therefore be discussed conclusively. This discussion will be given in a later paper of this series.

5. REDUCED FUNCTIONALS

The primitively divergent diagrams for $d = 3$ are shown in Fig. 1, together with the diagrams that they separate into by our introduction of the intermediary ψ-field, whose lines are broken. Inspection of (2) and (3) shows that the contributions A and C, as far as they arise anew on the right-hand sides, are isolated by introducing "reduced" functionals that do not contain factors corresponding to bare arcs. Since we have to allow for reducible diagrams also, we introduce an unknown functional $f(\omega)$ of one trajectory and define (arguments in square brackets are omitted)

$$n_r(\omega_1 \cdots \omega_n) = n(\omega_1 \cdots \omega_n) - \sum_i f(\omega_i) n(\omega_i \cdots [\omega]_i \cdots \omega_n) + \sum_{i<i'} f(\omega_i) f(\omega_{i'}) n(\omega_1 \cdots [\omega]_i \cdots [\omega]_{i'} \cdots \omega_n) - + \cdots, \tag{5.1a}$$

whence

$$n(\omega_1 \cdots \omega_n) = n_r(\omega_1 \cdots \omega_n) + \sum_i f(\omega_i) n_r(\omega_1 \cdots [\omega]_i \cdots \omega_n) + \cdots + \sum_{i<i'} f(\omega_i) f(\omega_{i'}) n(\omega_1 \cdots [\omega]_i \cdots [\omega]_{i'} \cdots \omega_n) + \cdots. \tag{5.1b}$$

Solving (4.2) and (4.3) for n_r requires lengthy formulas. Therefore, we take advantage of the blob picture and introduce the generating functional of functionals of blobs, the expansion element being a general functional $J(b)$ of a blob,

$$N\{J\} = \sum_{n=0}^{\infty} (n!)^{-1} \int \cdots \int Q(db_1) \cdots Q(db_n) \times J(b_1) \cdots J(b_n) \exp\left(-\tfrac{1}{2} \sum_i V_{ii} - \sum_{i<i'} V_{ii'}\right).$$

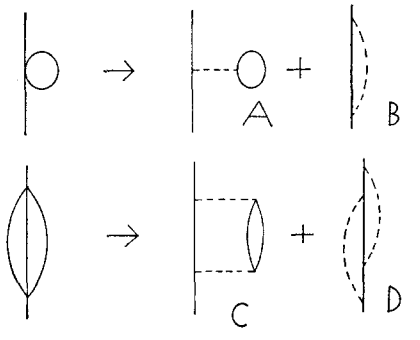

FIG. 1. Primitively divergent diagrams for $d = 3$.

[55] E.g., Ref. 15 and: A. Münster, "Statistische Thermodynamik kondensierter Phasen" in *Encyclopedia of Physics*, edited by S. Flügge (Springer-Verlag, Berlin, 1962), Vol. XIII.

[56] The rigorous results for infinite volume [see: D. Ruelle, in *Lectures in Theoretical Physics*, Vol. VI (University of Colorado Press, Boulder, 1964, p. 37)] cover only the domain of sufficiently small activity, i.e., at most the gaseous phase.

[57] Equations (4.2) and (4.3) describe only the Green's functions that possess gauge invariance of the first kind. See Ref. 35.

[58] D. Ruelle, Rev. Mod. Phys. **36**, 580 (1964).

The generating functional of the n-functionals is, according to (3.9b), $n\{J\} = N\{1\}^{-1}N\{J+1\}$. The KS equation for the N-functional is [the index b at N and n denotes functional differentiation $\delta/\delta J(b)$]

$$N_b\{J\} = e^{-\frac{1}{2}V_{bb}}N\{\exp(-V_{b\cdot})J(\cdot)\}$$

such that for the n-functional it is

$$n_b\{J\} = e^{-\frac{1}{2}V_{bb}}n\{e^{-V_{b\cdot}}J(\cdot) + (e^{-V_{b\cdot}} - 1)\} \tag{5.2}$$

while the MM equations are

$$N_{1\cdots n}\{J\} = \exp\left(-\frac{1}{2}\sum_{i=1}^{n}V_{ii} - \sum_{i<i'}V_{ii'}\right) \times N\left\{\exp\left(-\sum_{i=1}^{n}V_{i\cdot}\right)J(\cdot)\right\}$$

and therefrom

$$n(1\cdots n) = \exp\left(-\frac{1}{2}\sum_{i=1}^{n}V_{ii} - \sum_{i<i'}V_{ii'}\right) \times n\left\{\left[\exp\left(-\sum_{i=1}^{n}V_{i\cdot}\right) - 1\right]\right\}. \tag{5.3}$$

The definition $n_r\{J\} = \exp\{-\int Q(db)J(b)f(b)\}\, n\{J\}$ gives a rather unsymmetric KS equation for n_r, while the MM equation remains manifestly symmetric:

$$\begin{aligned} &n_r(1\cdots n) + \sum_{i=1}^{n}f(i)n_r(1\cdots[i]\cdots n) + \cdots \\ &= \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}V_{ii} - \sum_{i<i'}V_{ii'} + \int Q(db)f(b)\left[\exp\left(-\sum_{i=1}^{n}V_{ib}\right) - 1\right]\right\} \\ &\times n\left\{\left[\exp\left(-\sum_{i=1}^{n}V_{i\cdot}\right) - 1\right]\right\} \equiv s(1\cdots n) \end{aligned} \tag{5.4a}$$

such that

$$n_r(1\cdots n) = s(1\cdots n) - \sum_{i=1}^{n}f(i)s(1\cdots[i]\cdots n) \pm \cdots. \tag{5.4b}$$

We now define

$$\rho(1\cdots n) \equiv n_r(1\cdots n)f(1)^{-1}\cdots f(n)^{-1}, \tag{5.5}$$

and with

$$\rho(1\cdots n) = \sigma(1\cdots n) - \sum_{i=1}^{n}\sigma(1\cdots[i]\cdots n) + \sum_{i<i'}\sigma(1\cdots[i]\cdots[i]'\cdots n) - +\cdots \tag{5.6}$$

it is seen that the following two choices of f are convenient:

$$f(b) = \exp\left(-\tfrac{1}{2}V_{bb} - \int Q(db')K(b, b')\right) \equiv 1 + \epsilon(b) \tag{5.7a}$$

which leads to

$$\begin{aligned} &\sigma(1\cdots n) \\ &= \exp\left(-\sum_{i<i'}V_{ii'} - \int Q(db)\left[\epsilon(b)K(b_1\cdots b_n, b) - \sum_{i=1}^{n}K(b_i, b) + K(b_1\cdots b_n, b)\right]\right) \\ &\times \rho(-K(b_1\cdots b_n, \cdot)f(\cdot)) \end{aligned} \tag{5.7b}$$

and

$$f(b) = \exp\left(-\tfrac{1}{2}V_{bb} - \int Q(db')K(b, b')f(b')\right) \tag{5.8a}$$

which leads to

$$\begin{aligned} \sigma(1\cdots n) = \exp&\left(-\sum_{i<i'}V_{ii'} + \int Q(db)\left[\sum_{i=1}^{n}K(b_i, b) - K(b_1\cdots b_n, b)\right]f(b)\right) \\ &\times \rho(-K(b_1\cdots b_n, \cdot)f(\cdot)), \end{aligned} \tag{5.8b}$$

where the K are the blob transcriptions of (4.1) and (4.4) and $\rho(\cdots)$ is the generating functional of the ρ-functionals.

In both cases self-interaction and inter-blob interaction are separated; (5.7) and (5.8) differ only in the manner in which higher corrections that do not matter for renormalization are distributed. While (5.7a) gives f explicitly for use in (5.7b) with (5.6), (5.8a) is an integral equation in blob space of the Hammerstein type. No self-consistency problem is here involved, however, since (5.8a) is likely always to have a solution if (5.7a) is finite. Moreover, there is no such problem in the separation (5.7) and no consistency problem is expected physically.

We note that[59]

$$\begin{aligned} 0 &\leq \sum_{i=1}^{n}K(b_i, b) - K(b_1\cdots b_n, b) \\ &\leq \sum_{i<i'}K(b_i, b)K(b_{i'}, b) \end{aligned} \tag{5.9}$$

such that, because of, effectively, $K(b_i, b) = O(||b_i||\,||b||)$, the integrals in the exponents in (5.7b) and (5.8b) will be found to converge (for $d = 3$)[50] due to (3.7) and (4.5) if

$$\epsilon(b) = O(||b||). \tag{5.10}$$

[59] E. H. Lieb, J. Math. Phys. 4, 671 (1963).

The same is true for the integrals implicit in the last factors of (5.7b) and (5.8b) if, for orientation, one inserts in (5.6) for σ [with $\sigma(\phi) = 1$ always] the first approximation

$$\sigma(1 \cdots n) = \exp\left(-\sum_{i<i'} V_{ii'}\right), \qquad \sigma(1) = 1.$$

This shows that if (5.7a) turns out to be finite with (5.10) satisfied, the present renormalization is likely to have been successful for $d = 3$ but to be insufficient for $d = 4$ where more than mass renormalization is required.

APPENDIX A: RELATION TO QUANTUM STATISTICAL MECHANICS

The connection between EQFT of charged scalar particles and QSM of nonrelativistic neutral Bose particles[60] is best established by comparing (3.8) and (4.2) with G(5.4) and G(6.8), G(6.9) of Ginibre. The QSM combinatorics is dealt with simpler, however, by using for the generating functional of reduced density matrices

$$Z_\beta\{\bar{J}, J\} = \sum_{n=0}^{\infty} (n!)^{-2} \int \cdots \int dx_1 \cdots dx_n\, dy_1 \cdots dy_n \times \bar{J}(x_1) \cdots \bar{J}(x_n) J(y_1) \cdots J(y_n) \rho_\beta(x_1 \cdots x_n,\ y_1 \cdots y_n) \tag{A1}$$

the expression

$$Z_\beta\{\bar{J}, J\} = C \exp\left\{-2 \int_0^\beta d\tau \times \iint' dx\, dy V(x-y)\, \delta^2/(\delta\psi(x,\tau)\, \delta\psi(y,\tau))\right\} \times \exp\left[\bar{J}K(1-K)^{-1}J - \mathrm{Tr}\ln(1-K)\right]|_{\psi=0}, \tag{A2}$$

where K is the integral operator[61]

$$K(x,y) = z \int P^\beta_{xy}(d\omega) \times \exp\left[\tfrac{1}{2}\beta V(0) + \frac{1}{2}\int_0^\beta d\tau\, \psi(x(\tau),\tau)\right] \tag{A3}$$

and C such that $Z_\beta\{0, 0\} = 1$. Expanding the last exponential in (2) in powers of K and using (3.2) gives G(5.4) with substitution of G(5.8).

Equation (A.2) is related to (3.3) as follows. If we let

$$V(x-y) = \frac{g}{4}\beta\, \delta^\beta_{\mathrm{reg}}(x-y) \tag{A4}$$

where $\delta^\beta_{\mathrm{reg}}$ is a β-dependent, regularized delta function with $\delta^\beta_{\mathrm{reg}} \to \delta$ as $\beta \to 0$, and imagine β to become very small, the implicit τ-dependence of ψ in (A3) may be neglected, and the explicit τ-dependence has a trivial effect such that it is then equivalent to write the ψ in (A3) as $2^{-1}[\psi(x) + \psi(y)]$ and the first exponent in (A2) as $-2^{-1}g[\delta^2/\delta\psi^2]$. With $z = \exp\beta\mu$ and keeping μ finite, we have for very small β

$$K(1-K)^{-1} \approx \sum_{s=1}^{\infty} \int P^{\beta s}_{xy}(d\omega) \times \exp\left[\beta\mu s + \frac{1}{8}g\beta^2\, \delta^\beta_{\mathrm{reg}}(0)s + \frac{1}{2}\int_0^{\beta s} d\sigma\, \psi(x(\sigma))\right]$$

$$\approx \beta^{-1} \int_0^\infty ds \int P^s_{xy}(d\omega) \times \exp\left[-\tfrac{1}{2}m^2 s + \tfrac{1}{2}\alpha s + \frac{1}{2}\int_0^s d\sigma\, \psi(x(\sigma))\right] \tag{A5a}$$

and

$$-\mathrm{Tr}\ln(1-K) \approx \sum_{t=1}^{\infty} t^{-1} \int dz \int P^{\beta t}_{zz}(d\omega) \times \exp\left[\beta\mu t + \frac{1}{8}g\beta^2\, \delta^\beta_{\mathrm{reg}}(0)t + \frac{1}{2}\int_0^{\beta t} d\tau\, \psi(x(\tau))\right]$$

$$\approx \int_0^\infty t^{-1}\, dt \int dz \int P^t_{zz}(d\omega) \times \exp\left[-\tfrac{1}{2}m^2 t + \tfrac{1}{2}\alpha t + \frac{1}{2}\int_0^t d\tau\, \psi(z(\tau))\right] \tag{A5b}$$

provided we set

$$\mu = -\tfrac{1}{2}m^2 + \tfrac{1}{2}\alpha - \tfrac{1}{8}g\,\beta\,\delta^\beta_{\mathrm{reg}}(0)\,, \tag{A6}$$

which, however, is meaningless as it stands and not generally correct even for $d = 1$ (see Appendix C).

Actually, the "approximation" (A5) mainly shows that the combinatorics is correct but not what a suitable sequence of values for μ is as $\beta \to 0$. Such sequence can be found by comparing (5.7a) and (5.8a) with their QSM analogs [which resemble (4–12) of the last paper of Ref. 17]. This will be done in our next paper.

However, comparison of (A5) with (3.8) and (3.9b) gives, for a suitable choice of μ as function of β,

$$\lim_{\beta\to 0} 2^{-n}\beta^n \rho_\beta(x_1 \cdots x_n, y_1 \cdots y_n) = S(x_1 \cdots x_n, y_1 \cdots y_n) \tag{A7}$$

provided $M = 1$ and the potential is chosen as in (A4). More generally, (A4), (A6) and (A7) would read

$$V(x-y) = (4M^2)^{-1} g\beta\hbar^2\, \delta^\beta_{\mathrm{reg}}(x-y), \tag{A4'}$$

[60] J. Ginibre, J. Math. Phys. 6, 238 (1965). We shall refer to equations of this paper by G.

[61] The definition (3.7) corresponds to particle mass $M = 1$ instead of 2 as in G(2.3).

$$\mu = -(2M)^{-1}m^2c^2 + \Delta\mu(\beta), \qquad (A6')$$

$$\lim_{\beta\to 0}[\beta\hbar^3/(2M)]^n\rho_\beta(x_1 \cdots x_n, y_1 \cdots y_n) = S(x_1 \cdots x_n, y_1 \cdots y_n), \qquad (A7')$$

where $\Delta\mu(\beta) \to \infty$ as $\beta \to 0$ if $d \geq 2$.[62]

For the case without Wiener integrals, $d = 0$, (A7) can be verified directly as is done in Appendix B. The no longer elementary case $d = 1$ is briefly discussed in Appendix C.

For $d = 2$ and $d = 3$, it should be noted that a rigorous delta-function pair-potential has no physical effect,[60] such that $\delta^\beta_{\text{reg}} \not\to \delta$ for $\beta > 0$ is essential. However, the density goes to infinity (at least for $d \leq 3$) more strongly than β^{-1} (at least for $d \geq 2$) such that an effect of the potential also in the delta-limit is plausible. The statements on the density result from (A7) with $n = 1$ and the following bounds derived by other methods[11]

$$-\sup\{0, -(2g)^{-1}\,\delta m^2\} \leq S(0,0) - G_0(0) \leq G_0(0) + g^{-1}\,\delta m^2 + \{[G_0(0) + g^{-1}\,\delta m^2]^2 + G_0(0)^2\}^{\frac{1}{2}}$$

whereof, according to the remarks to (2.2), the first inequality is meaningful for $d \leq 2$ and the second for $d \leq 1$.

For $d = 4$ one expects also amplitude-and coupling-constant renormalization to be necessary; in (A4'), (A6'), and (A7') there are as many multipliers as necessary to take these effects into account. Thus it seems that the rules (A4'), (A6'), (A7') provide at least a regularization[63] of EQFT (of charged particles). From this viewpoint it is significant that a recent result of Fisher and Ruelle[64] excludes rigorously the possibility of the existence of the thermodynamical limit for the QSM Hamiltonians one would associate with the EQFT to the Lagrangian density (3.12) since the potential would have to be essentially negative to approach singular attraction.

Moreover, the quantitative rules (A4'), (A6'), (A7') suggest speculation about counterparts in EQFT and, by inference, MQFT of collective phenomena in QSM as exhibited in condensed phases.

[62] This will be required, due to tadpole divergence, for the limit (A7) to exist even if $\delta^\beta_{\text{reg}}$ is not made to become a delta function, cf. Sec. 4.

[63] This may be regarded even as a regularization of MQFT if one uses a $\delta^\beta_{\text{reg}}$ that is a delta function in one coordinate direction, which would correspond to the original, MQFT, time. However, MQFT not regularized further would still be ultraviolet-divergent, see Ref. 62.

[64] M. Fisher and D. Ruelle, IHES preprint, Theorem III and Remark.

APPENDIX B: $d = 0$: A NUMERICAL MODEL

We write here the formulas of Secs. 2–4 for zero dimensions such that all Wiener integrals are absent, since then the exact solution is known and the KS and MM equations are elementary but not trivial although the model is.

Equation (2.6) takes the form

$$S_x + g\,S_{xxy} - \alpha\,S_x = y\,S \qquad (B1a)$$

$$S_y + g\,S_{xyy} - \alpha\,S_y = x\,S \qquad (B1b)$$

for a function $S(x, y)$. All solutions of this system depend on $xy \equiv U$ only, i.e., preserve gauge invariance of the first kind. With $S(x, y) = S(U)$, (1) takes the form

$$S' + g(2\,S'' + U\,S''') - \alpha\,S' = S. \qquad (B2)$$

The three solutions, all of which are nonalytic in g at $g = 0$, can be written

$$S_1(U) = C\int_{\curvearrowright} dz(1 - \alpha - ig^{\frac{1}{2}}z)^{-1} \times \exp[-\tfrac{1}{2}z^2 + U(1 - \alpha - ig^{\frac{1}{2}}z)^{-1}] \qquad (B3)$$

and

$$S_{2,3}(U) = C\,\text{Re, Im}\oint dz(1 - \alpha - ig^{\frac{1}{2}}z)^{-1} \times \exp[-\tfrac{1}{2}z^2 + U(1 - \alpha - ig^{\frac{1}{2}}z)^{-1}].$$

Only $S_1(U)$ satisfies the analog of property (2.9),

$$\sum_{i,j=1}^{k} \bar{C}_i C_j\, S((\bar{a}_i + \bar{b}_i)(a_j + b_i)) > 0, \qquad (B4)$$

and possesses for $\alpha < 1$ an asymptotic expansion for $g \to 0$ from $|\arg g| < \pi$, while being analytic in the cut g-plane. It is also analytic in g and m^2 simultaneously in a certain domain. (B3) is the analog of (3.3). Using (3.4) gives the analog of (3.8b)

$$S_1(U) = \sum_{n=0}^{\infty}(n!)^{-1}\,U^n 2^{-n}\int_0^\infty \cdots \int ds_1 \cdots ds_n \times \exp[-\tfrac{1}{2}(s_1 + \cdots + s_n)]\,n(s_1 \cdots s_n) \qquad (B5)$$

where, explicitly,

$$n(s_1 \cdots s_n) = n(s_1 + \cdots + s_n) \qquad (B6)$$

with

$$n(s) = C\int_{\curvearrowright} dz(1 - \alpha - ig^{\frac{1}{2}}z)^{-1} \times \exp[-\tfrac{1}{2}z^2 + \tfrac{1}{2}\alpha s + \tfrac{1}{2}ig^{\frac{1}{2}}zs] = C'\,e^{s/2}\,\text{Erfc}\,[2^{-\frac{1}{2}}g^{\frac{1}{2}}s + (2g)^{-\frac{1}{2}}(1 - \alpha)] \qquad (B7)$$

where C (or C') such that $n(0) = 1$. The LE analogous to (3.9b) is again meaningless unless, e.g., the lower limits of the t-integrations are raised and then converges exponentially. The MM equation shows explicitly that (B6) holds and thus permits to sum up all terms on its right-hand side to obtain the integral equation

$$n(s) = \exp\left(\tfrac{1}{2}\alpha s - \tfrac{1}{8}gs^2\right) \times \left[1 - \frac{gs}{4}\int_0^\infty dt\, n(t) \exp\left(-\frac{1}{2}t - \frac{gst}{4}\right)\right] \quad \text{(B8)}$$

which is solved by (B7). The sum of terms converges absolutely since, summing absolutely and replacing $n(s)$ by $m(s)$, we obtain

$$m(s) = \exp\left(\tfrac{1}{2}\alpha s - (\tfrac{1}{8}g)s^2\right) \times \left[1 + \frac{gs}{4}\int_0^\infty dt\, m(t) \exp\left(-\tfrac{1}{2}t\right)\right] \quad \text{(B9)}$$

and with $n(t)$ for $m(t)$ the integral in (B9) converges for (B7). Using (B6) we obtain for (B5)

$$S_1(u) = 1 + u\int_0^\infty ds\, e^{-s/2}(2us)^{-\frac{1}{2}} I_1((2us)^{\frac{1}{2}}) n(s) \quad \text{(B10)}$$

which converges absolutely for all u and is the solution[65] (B3) of (B2). Thus, the intermediate use of the formal LE to arrive at the MM equation has introduced no error and $n(s)$ is positive as we expect it to be also for $d > 0$.

The construction of a solution of (B8) by iteration from the inhomogeneous term is certainly possible if also the iteration solution of (B9) converges, which is the case if and only if $\alpha < 1$. The natural value for α, according to (2.2), may be taken as $2g$, and then g should not be too large. To consider instead the convergence of the iteration solution of (B8) presupposes that the cancellations between the terms of alternating algebraic signs are brought to bear since if $\alpha > 1$ not all iterative approximations to $n(s)$, starting from the inhomogeneous term, are everywhere nonnegative. Even if this cancellation is observed, the iteration solution of (B8) does not converge for α too large. However, it does converge for α smaller than a g-dependent bound greater than one and gives, for these α, the unique solution of (B8). We have not shown that (B8) does not possess homogeneous solutions for α large enough. [It should be stressed that the boundary conditions on the determination of $n(s)$ are only that it be nonnegative and that the integrals in (B8), and in (B10) for an infinitesimal environment of the origin, converge.]

The KS equation does not make it manifest that $n(s_1 \cdots s_n)$ obeys (B6). However, since the MM equation is a consequence of the KS equation and implies (B6), we will first for simplicity consider only the solutions of this form. Then again the summation of all terms on the right-hand side of the KS equation can be performed and leads to

$$n(s + s') = \exp\left(\frac{1}{2}\alpha s - \frac{g}{8}s^2 - \frac{g}{4}ss'\right) \times \left[n(s') - \frac{gs}{4}\int_0^\infty dt\, n(t + s') \exp\left(-\frac{t}{2} - \frac{g}{4}st\right)\right].$$

This implies

$$n''(s) + \tfrac{1}{4}(-2 - 2\alpha + gs)\, n'(s) + \tfrac{1}{8}(2\alpha - gs)\, n(s) = 0 \quad \text{(B11)}$$

whereof the solution besides (7) is $\exp(\frac{1}{2}s)$ which does not solve (11) while (B7) does. Thus, (B11) has only the solution (B7).

The discussion of the convergence of an iteration solution of the KS equation is lengthier since, e.g., starting the iteration from $n(\phi) = 1$ all approximations give functions that depend only on the sum of their arguments but no longer one universal function for all n in a given step of iteration. However, in this special case the discussion can be reduced to the former one of the MM equation and the final result is the same.

The discussion of the MM and KS equations for reduced functionals is more complicated since the reduced functions depend no longer on the sum of arguments only, and has not been done.

The analog of the theory described by (3.11) has instead of (B1)

$$S_x + g\, S_{xxx} - \alpha\, S_x = x\, S. \quad \text{(B12)}$$

The change in the KS and MM equations is the same as that described for (3.11), and the discussion of (B12)[66] leads in every detail to similar results as were obtained for (B1).

The model (B1) stands in the relation described in Appendix A to the model given by

$$H = \tfrac{1}{2}V a^\dagger a^\dagger a\, a, \qquad [a, a^\dagger] = 1.$$

The "n-particle distribution function" is

$$\rho_n(\beta) = [\mathrm{Tr} \exp(-\beta H + \beta\mu N)]^{-1} \times \mathrm{Tr}\, \{(a^\dagger)^n a^n \exp(-\beta H + \beta\mu N)\}$$

[65] The solutions $S_{2,3}(u)$ do not admit the representation (B10).

[66] This model was also studied from a different point of view by E. R. Caianiello, A. Campolattaro, and M. Marinaro, Nuovo Cimento **38**, 1777 (1965).

where $N = a^\dagger a$. The generating function is

$$S_\beta(x) = \sum_{n=0}^{\infty} (n!)^{-2} x^n \rho_n(\beta)$$
$$= C \operatorname{Tr} \{\exp(-\beta H + \beta\mu N)\, e^{xa^\dagger e^a}\}$$

which is easily calculated to be

$$S_\beta(x) = C' \int_{\curvearrowright} dz \{1 - \exp[\beta\mu + \tfrac{1}{2}\beta V + iz(\beta V)^{\frac{1}{2}}]\}^{-1}$$
$$\times \exp\{-\tfrac{1}{2}z^2 + x$$
$$\times [-1 + \exp(-\beta\mu - \tfrac{1}{2}\beta V - iz(\beta V)^{\frac{1}{2}}]^{-1}\}.$$

The substitution corresponding to (A4), (A6), and (A7),

$$V = \tfrac{1}{4}\beta g, \qquad x = \tfrac{1}{2}\beta U, \qquad \mu = -\tfrac{1}{2} + \tfrac{1}{2}\alpha,$$

leads, with $C' \propto \beta$, for $\beta \to 0$ to (B3).

APPENDIX C. $d = 1$: THE ANHARMONIC OSCILLATOR

For $d = 1$, (2.1) becomes

$$L = \dot{B}^\dagger \dot{B} - m^2 B^\dagger B - \tfrac{1}{2} g (B^\dagger B)^2 + \alpha B^\dagger B. \qquad \text{(C1)}$$

With $B = r \exp(i\phi)$, $B^\dagger = r \exp(-i\phi)$ the Hamiltonian becomes

$$H = 4^{-1}[-\partial^2/(\partial r)^2 - r^{-1}\,\partial/\partial r$$
$$- r^{-2}\,\partial^2/(\partial\phi)^2 + 4(m^2 - \alpha)r^2 + 2gr^4] \qquad \text{(C2)}$$

which commutes with $-i\partial/\partial\phi$. The Green's functions

$$\langle T\, B(t_1) \cdots B(t_n) B^\dagger(t_1') \cdots B^\dagger(t_n')\rangle$$

can be continued analytically as described in Sec. 1 and an EQFT be based on them as described in Sec. 2.

The Hamiltonian (2.10) is a typical field theoretical one and as such beset with the familiar features concomitant with an infinite volume. There are no ultraviolet divergences, however. The energy spectrum is now continuous except for the vacuum state which may be separated by a gap[67] from the onset of the continuum. The Green's functions are the vacuum expectation values of the equal-time field-operator products. The theory differs from a two-dimensional MQFT by not possessing relativistic invariance, but is likely to admit a particle interpretation as it does for $g = 0$ with, however, an unusual energy–momentum relation. The eigenvalues and selection rules of the two-dimensional oscillator (C2) manifest themselves in the familiar asymptotic decrease (1.3) of the equal-time vacuum expectation values for large distances. Gauge invariance of the first kind is, of course, not broken.

The volume would be finite, with periodic boundary conditions, if we had taken the finite-temperature expectation[28] of the time-ordered operator product instead of the ground-state expectation. At least in the interaction-free case $g = 0$, the energy spectrum is now also discrete since it is then simply related to the discrete momentum spectrum.

All formulas of Sec. 3 hold with Wiener trajectories in one-dimensional space and the formulas of Sec. 4 concerning blobs hold now with one-dimensional blobs). As all diagrams are now convergent, we need not introduce reduced functionals and discuss the KS and MM equations directly, closely following Ginibre.[17] The linear vector space of sequences of Wiener-integrable functionals $n(\omega_1 \cdots \omega_n)$, $n = 1, 2, \cdots$ is made a Banach space by choosing as norm

$$||N|| = \operatorname*{Sup}_n \operatorname*{ess\,sup}_{\omega_1 \cdots \omega_n} |n(\omega_1 \cdots \omega_n)|\, g(\omega_1)^{-1} g(\omega_n)^{-1} \qquad \text{(C3)}$$

with a functional $g(\omega)$ to be suitably chosen, and completing. The operator Op of (4.6) has for the KS and MM equation a norm bounded by $c < 1$ if

$$g(\omega)\, c > \exp\left[\tfrac{1}{2}\alpha s + \int_0^\infty t^{-1}\, dt\, e^{-\frac{1}{2}m^2 t} \int dz \right.$$
$$\left. \times \int P_{zz}^t(d\tilde{\omega}) K(\omega, \tilde{\omega}) g(\tilde{\omega}) \right], \qquad \text{(C4)}$$

where we have used that $V_{ij} \geq 0$ and, in the MM case, (5.9). We may restrict ourselves to translation-invariant $g(\omega)$. Then, with (4.1),

$$\int dz \int P_{zz}^t(d\tilde{\omega}) K(\omega, \tilde{\omega}) g(\tilde{\omega}) = \int P_{00}^t(d\tilde{\omega}) g(\omega) \int dz$$
$$\times K(\omega, \tilde{\omega} + z) < 4^{-1} g s t \int P_{00}^t(d\tilde{\omega}) g(\tilde{\omega}).$$

Thus, setting

$$\int_0^\infty dt\, e^{-\frac{1}{2}m^2 t} \int P_{00}^t(d\tilde{\omega}) g(\tilde{\omega}) = a$$

we satisfy (C4) by choosing

$$g(\omega)\, c = \exp[4^{-1}(2\alpha s + g s a)],$$

from which follows

$$ac = (m^2 - \alpha - \tfrac{1}{2}ag)^{-\frac{1}{2}}$$

which can be solved for a with $c < 1$ provided

$$m^2 - \alpha > 3(\tfrac{1}{4}g)^{\frac{1}{3}}, \qquad \text{(C5}$$

[67] If $g = 0$ and $m^2 = 0$ there is no gap since the theory is then the usual one of free nonrelativitistic particles. If $g = 0$ and $m^2 > 0$ there is a gap.

i.e., the anharmonicity should not be too large for the iteration solution of (4.6) to converge to the then unique solution according to this estimate. Expecially, $\alpha < m^2$ would be required which would correspond to $\alpha < 1$ from (B9). It is likely that this restriction is due to the present crude estimate only since it did not apply to (B8), although on the basis of the results of Appendix B one would not expect the iteration from the N_0 in (4.6) to converge for arbitrarily large g and α.

For complex g the iteration solution converges uniformly if $\mathrm{Re}\, g > 0$ and in (C5) g is replaced by $|g|$, and therefore this solution is an analytic function of g in the open semicircle.[68] It is in addition analytic in $m^2 - \alpha$ provided

$$\mathrm{Re}\,(m^2 - \alpha) > 3(|g|/4)^{\frac{1}{3}}, \qquad \mathrm{Re}\, g > 0,$$

where it should be kept in mind that in view of (2.2) it is natural to have α depend on g.

The QSM model corresponding to this EQFT in the sense of Appendix A is the quantum gas of nonrelativistic neutral scalar bosons in one dimension with repulsive delta-function pair potential. This model has been solved exactly by Lieb and Liniger,[69] who observed nonanalytic behaviour of, e.g., the ground state energy at $g = 0$ and no phase transition in accordance with Landau and Lifshitz.[70] It is interesting that already in this case, which is free of ultraviolet divergences, (A6) should not be taken as it stands since we may set $\delta^{\beta}_{\mathrm{reg}} = \delta$ already for $\beta > 0$. This will be cleared up by "QSM regularization" of (5.7a) in our next paper.

[68] The method of Ref. 11, which does not rely on iteration, can be used to prove that the EQFT functions in finite one-dimensional volume with periodic boundary conditions, see Ref. 28, are analytic in g in the right g half-plane, and that the perturbation expansion is an asymptotic one.

[69] E. H. Lieb and W. Liniger, Phys. Rev. 130, 1605 (1963).

[70] L. D. Landau and E. M. Lifshitz, *Statistical Physics* (Pergamon Press Ltd., London, 1958), Sec. 149.

On Variational Principles for Electromagnetic Theory*

L. N. TAO
Illinois Institute of Technology, Chicago, Illinois
(Received 7 July 1965)

New variational principles for electromagnetic theory are established. A functional consisting of the field vectors is defined through the use of a convolution, and it is shown that the variation of this functional subject to appropriate constraints is completely equivalent to Maxwell's equations, Ohm's law, and the constitutive equations, together with appropriate boundary and initial conditions. The present formulation does not have the defects of the classical variational principle for electromagnetic theory since it does not require the introduction of scalar and vector potentials and *a priori* knowledge of the field vectors at the final stage. Two variational formulations for the electric and magnetic field vectors alone are also presented.

1. INTRODUCTION

ELECTROMAGNETIC phenomena are characterized by four field vectors: the electric field **E**, the magnetic field **H**, the electric displacement **D**, and the magnetic induction **B**. These four field vectors are indeed related to each other and to the distribution of current density **J**. The vector **J** must, of course, satisfy the conservation of charge equation. The five vectors, **E**, **H**, **D**, **B**, and **J**, are governed by the field equations, Maxwell's equations. Furthermore, for a given medium there exist specified constitutive equations between **E** and **D** and between **H** and **B**, and Ohm's law, between **J** and **E**. For a homogeneous medium, these pairs are linearly related with their ratios of proportionality specified by the physical nature of the medium. In principle, then, an electromagnetic problem constitutes the determination of these five vectors, i.e., the determination of solutions to Maxwell's equations subject to proper boundary and initial conditions.

In classical mechanics it is known that equations of motion can be formulated either through a set of differential equations, Lagrange equations, or through a variational principle, Hamilton's principle. In electromagnetic field theory, the set of field equations in differential forms, Maxwell's equations, have been well established. On the other hand, the variational formulation is far from complete. One approach to the variational formulation, but not without defects, is known.[1] It assumes that there exist a vector potential **A** and a scalar potential ϕ. By treating these two potentials as generalized coordinates of a Lagrangian represented by Hamilton's principle, it is possible to derive half of Maxwell's equations. In other words, with half of Maxwell's equations through which potentials **A** and ϕ are defined, together with a functional, the other half of Maxwell's equations are obtained by the variation of a functional. In this approach, the functional is specified over a fixed time interval, and the variations of the generalized coordinates are taken to be zero at end points of the time interval. This supposes *a priori* knowledge of these quantities at the final state, which unfortunately is not available in advance.

In this paper a new variational principle of electromagnetic theory will be established without the above mentioned defects. We intend to derive Maxwell's equations, as well as Ohm's law and constitutive equations, from a single functional subject to certain constraints, and without the knowledge of the field vectors at the final state. In other words, we intend to establish variational principles for electromagnetic theory which are equivalent to the field equations in differential forms.

The approach of this variational formulation is based on the use of a convolution, which has recently been applied by Gurtin[2,3] to some variational formulations in other fields of mathematical physics. In the following section we first detail some preliminary definitions and remarks. In Sec. 3 we discuss the variational formulation which is completely equivalent to Maxwell's equations with appropriate initial and boundary conditions. Based on this variational formulation the set of Maxwell's equations, Ohm's law, and constitutive equations, as well as the initial and boundary conditions can be derived. This formulation is then simplified to the case where field vectors satisfy the boundary conditions. Finally in Sec. 4 two variational formulations for the electric and magnetic field vectors are presented.

* This research was supported by the National Science Foundation through grant NSF GP-2258.

[1] See, for example, H. Goldstein, *Classical Mechanics* (Addison-Wesley Publishing Company, Inc., New York, 1950).

[2] M. E. Gurtin, Quart. Appl. Math. **22**, 252 (1964).

[3] M. E. Gurtin, Arch. Ratl. Mech. Anal. **16**, 34 (1964).

2. PRELIMINARY DEFINITIONS AND REMARKS

Throughout the paper the rationalized mks units will be used. The field vectors, **E**, **H**, **D**, **B**, and **J** and the charge density ρ are taken to be finite in the domain of physical interest and are real-valued continuous functions of space and time with continuous derivatives as necessary. Discontinuities in these field quantities may occur, however, on boundaries which mark an abrupt change in the physical properties of the medium.

The present variational formulation is based on the use of the convolution $f*g$, defined for two scalar functions $f(\mathbf{x}, t)$ and $g(\mathbf{x}, t)$, by

$$f*g(\mathbf{x}, t) = \int_0^t f(\mathbf{x}, t - \tau)g(\mathbf{x}, \tau)\, d\tau. \tag{2.1}$$

It is known[4] that this convolution is commutative, associative and distributive, and that $f*g = 0$ implies either $f = 0$ or $g = 0$. If $\mathbf{A}(\mathbf{x}, t)$ and $\mathbf{B}(\mathbf{x}, t)$ denote two vector functions,

$$\begin{aligned}
[f*\mathbf{A}](\mathbf{x}, t) &= \int_0^t f(\mathbf{x}, t - \tau)\mathbf{A}(\mathbf{x}, \tau)\, d\tau, \\
[\mathbf{A}*\cdot\mathbf{B}](\mathbf{x}, t) &= \int_0^t \mathbf{A}(\mathbf{x}, t - \tau)\cdot\mathbf{B}(\mathbf{x}, \tau)\, d\tau, \\
[\mathbf{A}*\times\mathbf{B}](\mathbf{x}, t) &= \int_0^t \mathbf{A}(\mathbf{x}, t - \tau)\times\mathbf{B}(\mathbf{x}, \tau)\, d\tau,
\end{aligned} \tag{2.2}$$

are the definitions of the convolutions of the product of a scalar and a vector, and of scalar and vector products of two vectors, respectively. And we write

$$\begin{aligned}
\left[\frac{\partial \mathbf{A}}{\partial t}*\cdot\mathbf{B}\right](\mathbf{x}, t) &= \left[\mathbf{B}*\cdot\frac{\partial \mathbf{A}}{\partial t}\right](\mathbf{x}, t) \\
&= \int_0^t \frac{\partial \mathbf{A}(\mathbf{x}, t - \tau)}{\partial(t - \tau)}\cdot\mathbf{B}(\mathbf{x}, \tau)\, d\tau \\
&= \int_0^t \mathbf{B}(\mathbf{x}, t - \tau)\cdot\frac{\partial \mathbf{A}(\mathbf{x}, \tau)}{\partial \tau}\, d\tau, \\
\left[\frac{\partial \mathbf{A}}{\partial t}*\cdot\frac{\partial \mathbf{B}}{\partial t}\right](\mathbf{x}, t) &= \left[\frac{\partial \mathbf{B}}{\partial t}*\cdot\frac{\partial \mathbf{A}}{\partial t}\right](\mathbf{x}, t) \\
&= \int_0^t \frac{\partial \mathbf{A}(\mathbf{x}, t - \tau)}{\partial(t - \tau)}\cdot\frac{\partial \mathbf{B}(\mathbf{x}, \tau)}{\partial \tau}\, d\tau \\
&= \int_0^t \frac{\partial \mathbf{B}(\mathbf{x}, t - \tau)}{\partial(t - \tau)}\cdot\frac{\partial \mathbf{A}(\mathbf{x}, \tau)}{\partial \tau}\, d\tau
\end{aligned} \tag{2.3}$$

for the convolutions of time derivatives of two vectors. Also the variation of a functional $I(F)$ is defined in the usual manner by

$$\delta I(F) = \frac{d}{d\epsilon} I(F + \epsilon F')\,|_{\epsilon=0} \tag{2.4}$$

for any finite real-valued quantity F'.

For future reference we list here the fundamental system of equations for classical electromagnetic theory. Maxwell's equations connecting the field vectors are

$$\begin{aligned}
\nabla\times\mathbf{E} &= -\partial\mathbf{B}/\partial t, \\
\nabla\times\mathbf{H} &= \mathbf{J} + \partial\mathbf{D}/\partial t, \\
\nabla\cdot\mathbf{D} &= \rho, \\
\nabla\cdot\mathbf{B} &= 0.
\end{aligned} \tag{2.5}$$

Ohm's law for a stationary medium is

$$\mathbf{J} = \sigma\mathbf{E}. \tag{2.6}$$

The constitutive equations are

$$\mathbf{D} = \epsilon\mathbf{E}, \qquad \mathbf{B} = \mu\mathbf{H}. \tag{2.7}$$

Also, the conservation of charge equation is

$$\partial\rho/\partial t + \nabla\cdot\mathbf{J} = 0. \tag{2.8}$$

The field vectors are of course subject to proper initial and boundary conditions. The appropriate boundary conditions are

$$\begin{aligned}
\mathbf{n}\times[\![\mathbf{E}]\!] = 0, \qquad \mathbf{n}\times[\![\mathbf{H}]\!] = \kappa, \qquad \mathbf{n}\cdot[\![\mathbf{D}]\!] = \omega, \\
\mathbf{n}\cdot[\![\mathbf{B}]\!] = 0,
\end{aligned} \tag{2.9}$$

where $\mathbf{n}$ is the unit normal at the boundary and the double bracket represents the jump at the boundary. κ and ω are the surface current density and surface charge density, respectively. It is known[5] that $\kappa = 0$ when both media have finite conductivity. The initial conditions, though arbitrary, must satisfy the static field equations, i.e.,

$$\begin{aligned}
\nabla\times\mathbf{E}_0 = 0, \qquad \nabla\times\mathbf{H}_0 = \mathbf{J}_0, \qquad \nabla\cdot\mathbf{D}_0 = \rho_0, \\
\nabla\cdot\mathbf{B}_0 = 0,
\end{aligned} \tag{2.10}$$

where the subscript 0 has been used to indicate the initial condition or the static field. Though other boundary and initial conditions may be stated, they are essentially the consequence of Ohm's law and constitutive equations in integral forms.

The electromagnetic problem is completely characterized by these field vectors which must satisfy the governing equations as well as the boundary and initial conditions. For simplicity in later discussion we use the term *an electromagnetic state* $F = (\mathbf{E}, \mathbf{H}, \mathbf{D}, \mathbf{B}, \mathbf{J})$ to be an ordered array of well-

[4] J. Mikusinski, *Operational Calculus* (Pergamon Press, Inc., New York, 1959).

[5] J. A. Stratton, *Electromagnetic Theory* (McGraw-Hill Book Company, Inc., New York, 1941).

defined vector functions **E**, **H**, **D**, **B**, and **J**, and *an admissible state* to be an electromagnetic state which satisfies the boundary conditions. We use the term *a solution of an electromagnetic field* to mean an admissible state which satisfies the governing equations and initial conditions.

3. VARIATIONAL FORMULATIONS FOR ELECTROMAGNETIC THEORY

We now proceed to state the variational principle. For this principle the electromagnetic states are not required to satisfy any field equations or initial or boundary conditions. First define[6] a functional of the electromagnetic state F by

$$I(F) = \int_R \Big[1 * (\nabla \times \mathbf{H}) *\cdot \mathbf{E} - 1 * \mathbf{J} *\cdot \mathbf{E} - \mathbf{D} *\cdot \mathbf{E} - \mathbf{B} *\cdot \mathbf{H} + \mathbf{D}_0 *\cdot \mathbf{E} + \mathbf{B}_0 *\cdot \mathbf{H} + \frac{1}{2\sigma} * \mathbf{J} *\cdot \mathbf{J} + \frac{1}{2\epsilon} \mathbf{D} *\cdot \mathbf{D} + \frac{1}{2\mu} \mathbf{B} *\cdot \mathbf{B} \Big] (\mathbf{x}, t) dv - \int_Z [\mathbf{n} \cdot \mathbf{E}_b * \times \mathbf{H}](\mathbf{x}, t)\, dS. \tag{3.1}$$

The first integral is a volume integral over the physical space of interest, and the second integral is a surface integral over the bounding surface of that domain. Subscript b denotes the other medium. This functional contains two initial conditions $\mathbf{D}_0$ and $\mathbf{B}_0$ which must satisfy the constraints

$$\nabla \cdot \mathbf{D}_0 = \rho_0 \quad \text{and} \quad \nabla \cdot \mathbf{B}_0 = 0. \tag{3.2}$$

An additional constraint, which is purely *kinematic* and is different in nature from the field equations and constitutive equations, is the conservation of charge

$$\partial \rho / \partial t + \nabla \cdot \mathbf{J} = 0. \tag{3.3}$$

Taking the variation of $I(F)$ and using the properties of convolution and divergence theorems, we obtain

$$\delta I(F) = \int_R \Big\{ [1 * (\nabla \times \mathbf{H}) - 1 * \mathbf{J} - \mathbf{D} + \mathbf{D}_0] *\cdot \delta \mathbf{E} - [1 * (\nabla \times \mathbf{E}) + \mathbf{B} - \mathbf{B}_0] *\cdot \delta \mathbf{H} - \Big(1 * \mathbf{E} - \frac{1}{\sigma} * \mathbf{J}\Big) *\cdot \delta \mathbf{J} - (\mathbf{E} - \mathbf{D}/\epsilon) *\cdot \delta \mathbf{D} - (\mathbf{H} - \mathbf{B}/\mu) *\cdot \delta \mathbf{B} \Big\} (\mathbf{x}, t)\, dv + \int_Z [\mathbf{n} \times (\mathbf{E} - \mathbf{E}_b) *\cdot \delta \mathbf{H}](\mathbf{x}, t)\, dS. \tag{3.4}$$

[6] Some minor modifications of this functional are possible. As an example the first term of the volume intergral may be replaced by $-1*(\nabla \times \mathbf{E}) *\cdot \mathbf{H}$ and the integrand of the surface integral by $\mathbf{n} \cdot \mathbf{H}_{b} * \times \mathbf{E} + \kappa *\cdot \mathbf{E}$.

Clearly, if $\delta I(F) = 0$, then

$$1 * (\nabla \times \mathbf{H}) - 1 * \mathbf{J} - \mathbf{D} + \mathbf{D}_0 = 0,$$

$$1 * (\nabla \times \mathbf{E}) + \mathbf{B} - \mathbf{B}_0 = 0, \quad 1 * (\mathbf{E} - \mathbf{J}/\sigma) = 0,$$

$$\mathbf{H} - \mathbf{B}/\mu = 0, \quad \mathbf{E} - \mathbf{D}/\epsilon = 0, \quad \mathbf{n} \times [\![\mathbf{E}]\!] = 0. \tag{3.5}$$

Differentiation of the first three of this set of equations with respect to time yields

$$\nabla \times \mathbf{H} - \mathbf{J} - \partial \mathbf{D}/\partial t = 0,$$

$$\nabla \times \mathbf{E} + \partial \mathbf{B}/\partial t = 0, \quad \mathbf{E} - \mathbf{J}/\sigma = 0. \tag{3.6}$$

Furthermore, by taking the divergence of the first two, we have

$$\nabla \cdot \mathbf{J} + \partial (\nabla \cdot \mathbf{D})/\partial t = 0, \quad \partial (\nabla \cdot \mathbf{B})/\partial t = 0. \tag{3.7}$$

Now using the constraints $\nabla \cdot \mathbf{D}_0 = \rho_0$, $\nabla \cdot \mathbf{B}_0 = 0$, and $\partial \rho / \partial t + \nabla \cdot \mathbf{J} = 0$ results in

$$\nabla \cdot \mathbf{D} = \rho \quad \text{and} \quad \nabla \cdot \mathbf{B} = 0. \tag{3.8}$$

These equations form a complete set of the field equations of electromagnetism. It is noted they were derived without any *a priori* knowledge of Maxwell's equations, constitutive equations or Ohm's law. Boundary conditions which have not been obtained in this derivation can be easily established from the derived equations.

It may be argued that we have assumed the knowledge of three constraints. Two are completely *static* in nature. If necessary, they may be easily established by some simple variational formulation. Since the interest of this analysis is in dynamic problems, these constraints are of no concern. The third constraint is the conservation of charge, i.e., the relationship between the distributions of charge and the current density which gives rise to the electromagnetic field. In contrast to other equations which are *dynamic* in nature, the conservation of charge is a purely *kinematic* relationship, in a manner similar to the law of conservation of mass (continuity equation) in continuum mechanics.

It can easily be shown that the inverse of the above variational principle applies, i.e., if an electromagnetic state F is a solution, then $\delta I(F) = 0$ and the constraints are satisfied. By integrating the two Maxwell's vector equations and Ohm's law, and substituting these results into Eq. (3.4), $\delta I(F) = 0$ is readily obtained. In addition, the three constraints are automatically satisfied. As a summary of these developments, the following theorem is specified:

Theorem 3.1: Let $F = (\mathbf{E}, \mathbf{H}, \mathbf{D}, \mathbf{B}, \mathbf{J})$ be an electromagnetic state and $I(F)$ a real-valued functional defined by Eq. (3.1). Then F is a solution of electromagnetic theory, if and only if $\delta I(F) = 0$ subject to the constraints, Eqs. (3.2) and (3.3).

If, as a stronger condition F is an admissible state, i.e., an electromagnetic state satisfying the boundary conditions, then the surface integral of Eq. (3.1) need not be included in the definition of the functional. At the boundary or interfaces the tangential components of $\mathbf{E}$ and $\mathbf{H}$ are continuous. With a bounding medium of finite conductivity, $\mathbf{n} \times [\![\mathbf{H}]\!] = 0$ or $\mathbf{n} \times \delta\mathbf{H} = 0$. For an infinitely conducting bounding medium the field intensities $\mathbf{E}_b$ and $\mathbf{H}_b$ are zero. This implies that at the boundary $\mathbf{n} \times \mathbf{E} = \mathbf{n} \times \mathbf{E}_b = 0$. In either case the surface integral of Eq. (3.4) vanishes,

$$\int_\Sigma [\mathbf{n} \times \mathbf{E} *\cdot \delta\mathbf{H}](\mathbf{x}, t)\, dS = 0. \tag{3.9}$$

Therefore, we may conclude:

Theorem 3.2: Let $F = (\mathbf{E}, \mathbf{H}, \mathbf{D}, \mathbf{B}, \mathbf{J})$ be an admissible state and $I(F)$ a real-valued functional defined by

$$I(F) = \int_R \Big[1 * (\nabla \times \mathbf{H}) *\cdot \mathbf{E} - 1 * \mathbf{J} *\cdot \mathbf{E} - \mathbf{D} *\cdot \mathbf{E} - \mathbf{B} *\cdot \mathbf{H} + \mathbf{D}_0 *\cdot \mathbf{E} + \mathbf{B}_0 *\cdot \mathbf{H} + \frac{1}{2\sigma} \mathbf{J} *\cdot \mathbf{J} + \frac{1}{2\epsilon} \mathbf{D} *\cdot \mathbf{D} + \frac{1}{2\mu} \mathbf{B} *\cdot \mathbf{B} \Big](\mathbf{x}, t)\, dv. \tag{3.10}$$

Then F is a solution of electromagnetic theory, if and only if $\delta I(F) = 0$ subject to the constraints, Eqs. (3.2) and (3.3).

Two variational principles for electromagnetic theory which are equivalent to the field equations in differential forms have been established. Many simplifications can be achieved, if the electromagnetic state or admissible state satisfies, in addition, one or more of the constitutive equations and Ohm's law. To illustrate this point, we give a typical corollary which is the consequence of the above theorems.

Corollary 3.1: Let $F = (\mathbf{E}, \mathbf{H}, \mathbf{J})$ be an admissible state satisfying the constitutive equations $\mathbf{D} = \epsilon\mathbf{E}$, and $\mathbf{B} = \mu\mathbf{H}$ and $I(F)$ be a real-valued functional defined by

$$I(F) = \int_R \Big[1 * (\nabla \times \mathbf{H}) *\cdot \mathbf{E} - 1 * \mathbf{J} *\cdot \mathbf{E} - \frac{\epsilon}{2} \mathbf{E} *\cdot \mathbf{E} - \frac{\mu}{2} \mathbf{H} *\cdot \mathbf{H} + \epsilon\mathbf{E}_0 *\cdot \mathbf{E} - \mu\mathbf{H}_0 *\cdot \mathbf{H} + \frac{1}{2\sigma} * \mathbf{J} *\cdot \mathbf{J} \Big](\mathbf{x}, t)\, dv. \tag{3.11}$$

Then F is a solution of electromagnetic theory, if and only if $\delta I(F) = 0$ subject to the constraints

$$\nabla \cdot \mathbf{E}_0 = 0, \qquad \nabla \cdot \mathbf{H}_0 = 0,$$

and

$$\partial\rho/\partial t + \nabla \cdot \mathbf{J} = 0. \tag{3.12}$$

This is the direct consequence of Theorem 3.2 with the substitutions of the constitutive equations $\mathbf{D} = \epsilon\mathbf{E}$, and $\mathbf{B} = \mu\mathbf{H}$ into Eq. (3.10).

4. ADMISSIBLE ELECTRIC AND MAGNETIC FIELDS

In many applications it is often the case that the solution of one of the field vectors, commonly $\mathbf{E}$ or $\mathbf{H}$, is first determined. With the use of either of these field vectors and with Maxwell's equations other field vectors, if necessary, may be found. $\mathbf{E}(\mathbf{x}, t)$ and $\mathbf{H}(\mathbf{x}, t)$ are governed by

$$\mu\epsilon \frac{\partial^2 \mathbf{E}}{\partial t^2} + \mu\sigma \frac{\partial \mathbf{E}}{\partial t} + \nabla \times (\nabla \times \mathbf{E}) = 0, \tag{4.1}$$

$$\mu\epsilon \frac{\partial^2 \mathbf{H}}{\partial t^2} + \mu\sigma \frac{\partial \mathbf{H}}{\partial t} + \nabla \times (\nabla \times \mathbf{H}) = 0, \tag{4.2}$$

subject to the initial conditions

$$\frac{\partial \mathbf{E}}{\partial t}(\mathbf{x}, 0) = \dot{\mathbf{E}}_0(\mathbf{x}), \qquad \mathbf{E}(\mathbf{x}, 0) = \mathbf{E}_0(\mathbf{x}), \tag{4.3}$$

$$\frac{\partial \mathbf{H}}{\partial t}(\mathbf{x}, 0) = \dot{\mathbf{H}}_0(\mathbf{x}), \qquad \mathbf{H}(\mathbf{x}, 0) = \mathbf{H}_0(\mathbf{x}), \tag{4.4}$$

and boundary conditions

$$\mathbf{n} \times [\![\mathbf{E}]\!] = 0, \qquad \mathbf{n} \cdot [\![\epsilon\mathbf{E}]\!] = \omega, \tag{4.5}$$

$$\mathbf{n} \times [\![\mathbf{H}]\!] = \kappa, \qquad \mathbf{n} \cdot [\![\mu\mathbf{H}]\!] = 0. \tag{4.6}$$

In this section we present variational formulations which are equivalent to these differential formulations. As in the previous section several alternate formulations may be given. The specific one presented is the one in which both initial and boundary conditions are satisfied as preliminary conditions. This one would seem to be most appropriate to applications. Introduce the term *an initially admissible electric field* to mean the electric field satisfying both the initial and boundary conditions. Also, define

a functional of an initially admissible electric field by

$$I(\mathbf{E}) = \int_R \left[\frac{\mu\epsilon}{2} \frac{\partial \mathbf{E}}{\partial t} *\cdot \frac{\partial \mathbf{E}}{\partial t} + \frac{\mu\sigma}{2} \frac{\partial \mathbf{E}}{\partial t} *\cdot \mathbf{E} - \frac{1}{2} (\nabla \times \mathbf{E}) *\cdot (\nabla \times \mathbf{E}) - \mu\epsilon \dot{\mathbf{E}}_0 \cdot \mathbf{E} - \frac{\mu\sigma}{2} \mathbf{E}_0 \cdot \mathbf{E} \right](\mathbf{x}, t)\, dv. \tag{4.7}$$

The variation of $I(\mathbf{E})$ with the use of the divergence theorem and the properties of convolution leads to

$$\delta I(\mathbf{E}) = \int_R \left\{ \left[\mu\epsilon \frac{\partial^2 \mathbf{E}}{\partial t^2} + \mu\sigma \frac{\partial \mathbf{E}}{\partial t} + \nabla \times (\nabla \times \mathbf{E}) \right] *\cdot \delta\mathbf{E} \right\}(\mathbf{x}, t)\, dv + \int_R \left\{ \mu\epsilon \left[\frac{\partial \mathbf{E}}{\partial t}(\mathbf{x}, 0) - \dot{\mathbf{E}}_0 \right] + \frac{\mu\sigma}{2} [\mathbf{E}(\mathbf{x}, 0) - \mathbf{E}_0] \right\} \cdot \delta\mathbf{E}(\mathbf{x}, t)\, dv - \int_R \left[\mu\epsilon \frac{\partial \mathbf{E}}{\partial t} - \frac{\mu\sigma}{2} \mathbf{E} \right](\mathbf{x}, t) \cdot \delta\mathbf{E}(\mathbf{x}, 0)\, dv + \int_\Sigma [\mathbf{n} \times (\nabla \times \mathbf{E}) *\cdot \delta\mathbf{E}](\mathbf{x}, t)\, dS. \tag{4.8}$$

Clearly, for an initially admissible electric field, the second and third volume integrals vanish. Also, at the boundary, $\mathbf{n} \times [\![\mathbf{E}]\!] = 0$ implies $\mathbf{n} \times \delta\mathbf{E} = 0$. This shows that the surface integral vanishes. Therefore, if $\mathbf{E}$ is a solution of electromagnetic theory, $\delta I(\mathbf{E}) = 0$. Conversely, if $\delta I(\mathbf{E}) = 0$, from Eq. (4.8), we see that Eq. (4.1) holds. Therefore, we may conclude:

Theorem 4.1: Let $\mathbf{E}$ be an initially admissible electric field vector and $I(\mathbf{E})$ a real-valued functional defined by Eq. (4.7). Then $\mathbf{E}$ is a solution of electromagnetic theory if and only if $\delta I(\mathbf{E}) = 0$.

Similarly, if we define a functional of an initially admissible magnetic field vector $\mathbf{H}(\mathbf{x}, t)$ by

$$I(\mathbf{H}) = \int_R \left[\frac{\mu\epsilon}{2} \frac{\partial \mathbf{H}}{\partial t} *\cdot \frac{\partial \mathbf{H}}{\partial t} + \frac{\mu\sigma}{2} \frac{\partial \mathbf{H}}{\partial t} *\cdot \mathbf{H} - \tfrac{1}{2} (\nabla \times \mathbf{H}) *\cdot (\nabla \times \mathbf{H}) - \mu\epsilon \dot{\mathbf{H}}_0 \cdot \mathbf{H} - \frac{\mu\sigma}{2} \mathbf{H}_0 \cdot \mathbf{H} \right](\mathbf{x}, t)\, dv, \tag{4.9}$$

then the variation of $I(\mathbf{H})$ yields

$$\delta I(\mathbf{H}) = \int_R \left\{ \left[\mu\epsilon \frac{\partial^2 \mathbf{H}}{\partial t^2} + \mu\sigma \frac{\partial \mathbf{H}}{\partial t} + \nabla \times (\nabla \times \mathbf{H}) \right] *\cdot \delta\mathbf{H} \right\}(\mathbf{x}, t)\, dv + \int_R \left\{ \mu\epsilon \left[\frac{\partial \mathbf{H}}{\partial t}(\mathbf{x}, 0) - \dot{\mathbf{H}}_0 \right] + \frac{\mu\sigma}{2} [\mathbf{H}(\mathbf{x}, 0) - \mathbf{H}_0] \right\} \cdot \delta\mathbf{H}(\mathbf{x}, t)\, dv - \int_R \left[\mu\epsilon \frac{\partial \mathbf{H}}{\partial t} - \frac{\mu\sigma}{2} \mathbf{H} \right](\mathbf{x}, t) \cdot \delta\mathbf{H}(\mathbf{x}, 0)\, dv + \int_\Sigma [\mathbf{n} \times (\nabla \times \mathbf{H}) *\cdot \delta\mathbf{H}](\mathbf{x}, t)\, dS. \tag{4.10}$$

The second- and third-volume integrals vanish by virtue of the initial conditions. The surface integral also vanishes since at the boundary $\mathbf{n} \times [\![\mathbf{H}]\!] = 0$ or $\mathbf{n} \times \delta\mathbf{H} = 0$ for a bounding medium with finite conductivity. For an infinitely conducting bounding medium, the tangential component of $\nabla \times \mathbf{H}$ approaches zero, i.e., $\mathbf{n} \times (\nabla \times \mathbf{H}) = 0$. Therefore, if $\mathbf{H}$ is a solution, $\delta I(\mathbf{H}) = 0$. The converse is also true. If $\delta I(\mathbf{H}) = 0$, where $\mathbf{H}$ is an initially admissible magnetic field vector, then Eq. (4.2) holds. Hence:

Theorem 4.2: Let $\mathbf{H}$ be an initially admissible magnetic field vector and $I(\mathbf{H})$ a real-valued functional defined by Eq. (4.9). Then $\mathbf{H}$ is a solution of electromagnetic theory if and only if $\delta I(\mathbf{H}) = 0$.

Phase Transition of a Two-Dimensional Continuum—Ising Model*

COLIN J. THOMPSON†
Department of Physics, University of California, San Diego, La Jolla, California
(Received 28 June 1965)

We decorate a plane two-dimensional Ising model by placing on each bond of the lattice a continuum spin which is allowed to interact only with the Ising spins at the ends of the bonds. The continuum spins are chosen to be either Gaussian (i.e., normally distributed with zero mean and unit variance) or spherical (i.e., constrained to lie on the surface of a sphere) after Berlin and Kac, and by integrating out the continuum spins, the partition function of the decorated lattice is expressed in terms of the Onsager partition function of the plane two-dimensional Ising model. The critical behavior of the model is as follows: For the Gaussian case, as for the plane Ising model, the specific heat has a logarithmic singularity at the critical point T_c^G given by

$$2 \tanh^2 [2(J/kT_c^G)^2] = 1$$

and as $t = T_c^G - T \to 0^+$, the spontaneous magnetization goes to zero like $t^{1/8}$. For the spherical case, the specific heat is continuous and has a cusp at the critical point T_c^s given by $J/kT_c^s = [z_c + (2 + 2^{\frac{1}{2}}) z_c^2]^{\frac{1}{4}}$, $2 \tanh^2 (2z_c) = 1$, with slope going to $\pm \infty$ like $t^{-1} [\ln |t|\ ]^{-2}$ as $t = T_c^s - T \to 0^{\pm}$, and as $t \to 0^+$, the spontaneous magnetization goes to zero like $[t/\ln t]^{1/8}$.

1. INTRODUCTION

THE theoretical difficulties involved in the study of phase transitions, the consequent lack of any general rigorous approach, and the inadequacies of approximate treatments of the problem, have led people to formulate models, no matter now far removed from reality, which can be analyzed exactly and which exhibit phase transitions, in the hope that in so doing some light may be shed on the intricacies of the transition region.

The most instructive model in this respect, and one which has received a good deal of attention over the years, is the Ising model, for which we have the now classical Onsager solution in two dimensions.[1,2] The three-dimensional model has so far not yielded to exact analysis, but a good deal of information has accumulated in recent years as a result of the successive approximation work of Domb, Sykes, Fisher, Baker,[3] etc. Various modifications of the two-dimensional model have been treated, perhaps the most interesting of which are the decorated models of Fisher,[4] and more recently of Syozi.[5] The decoration in each case is performed by placing on each bond of the Ising lattice, another spin which is allowed to interact only with the Ising spins at the ends of the bonds. By summing over the spins of the decorating system, the partition function of the decorated lattice can be reduced to that of the underlying Ising lattice with modified coupling constant. Fisher decorated with spin one-half's (i.e., Ising spins $\mu = \pm 1$), and by a judicious choice of coupling, he was able to obtain the partition function for his decorated lattice in an arbitrary magnetic field from the Onsager solution. This, of course, supplemented the original Onsager solution which was obtained in the absence of external fields. As for the ordinary plane Ising model, Fisher's model has a logarithmic singularity in the specific heat. Syozi's modification was obtained by decorating with spin-one's (i.e., $\mu = -1, 0, +1$), and by adjusting the sum over decorated sites $\sum \mu^2$, Syozi finds that the logarithmic singularity in the specific heat can be replaced by a cusp with vertical tangent at the critical point.

In the present work we consider a plane Ising model decorated with continuum spins rather than with discrete spins. As for the discrete case, the partition function is expressed in terms of the Onsager solution by first integrating out the continuum spins. The model is defined in the following section and its critical behavior, which is essentially the same as for Syozi's model, is examined in Secs. 3, 4, and 5.

2. THE MODEL

The underlying lattice is two-dimensional with n rows and n columns, and at the intersection of the ith row and jth column [vertex (i, j)] there is an Ising spin $\mu_{i,j} = \pm 1$, and between vertices (i, j) and $(i, j + 1)$, and (i, j) and $(i + 1, j)$ are

* Supported by AF contract No. AF-AFOSR-610-64, Theory of Solids.
† Present address: The Rockefeller University, New York, N.Y.
[1] L. Onsager, Phys. Rev. 65, 117 (1944).
[2] For reviews, see G. F. Newell and E. W. Montroll, Rev. Mod. Phys. 25, 353 (1953), and C. Domb, Adv. Phys. 9, 151 (1960).
[3] For reviews, see M. E. Fisher, J. Math. Phys. 4, 278 (1963), and C. Domb, Ref. 2.
[4] M. E. Fisher, Proc. Roy. Soc. (London) A254, 66 (1960); A256, 502(1960).
[5] I. Syozi, Progr. Theoret. Phys. (Kyoto) 34, 189 (1965).

continuum spins $-\infty < x_{i,j} < \infty$, and $-\infty < y_{i,j} < \infty$, respectively. We distinguish two cases: Gaussian: the $x_{i,j}$ and $y_{i,j}$ are normally distributed with zero mean and unit variance, and spherical: the $x_{i,j}$ and $y_{i,j}$ are constrained to lie on the surface of a $2N$-dimensional sphere, i.e.,

$$\sum_{ij} (x_{i,j}^2 + y_{i,j}^2) = 2N, \tag{1}$$

where $N = n^2$ is the total number of Ising spins (and one-half of the total number of continuum spins), after Berlin and Kac,[6] who have considered models made up solely of these continuum spins (the Gaussian and spherical models, respectively).

We allow only nearest-neighbor interactions between Ising and continuum spins so that the model Hamiltonian is given by

$$H = -J \sum (x_{i,j}[\mu_{i,j} + \mu_{i,j+1}] + y_{i,j}[\mu_{i,j} + \mu_{i+1,j}]), \tag{2}$$

where J is some coupling constant. For convenience we choose periodic boundary conditions

$$\mu_{i+n,j+n} = \mu_{i,j}; \quad x_{i+n,j+n} = x_{i,j}; \quad y_{i+n,j+n} = y_{i,j}.$$

The partition function for the system is given by

$$Z_{2N} = \sum_{\mu_{i,j}=\pm 1} A_{2N}^{-1} \Big(\prod \int_{-\infty}^{\infty} dx_{i,j}\, dy_{i,j} \Big) f(x, y) e^{-\beta H}, \tag{3}$$

where for the Gaussian case

$$A_{2N}^{G} = (2\pi)^N; \quad f(x, y) = \exp\left[-\tfrac{1}{2} \sum (x_{i,j}^2 + y_{i,j}^2)\right] \tag{4}$$

and for the spherical case

$$A_{2N}^{s} = 2(\pi)^N (2N)^{N-\frac{1}{2}}/\Gamma(N); \qquad f(x, y) = \delta(2N - \sum (x_{i,j}^2 + y_{i,j}^2)). \tag{5}$$

We are ultimately interested in the free energy per (continuum) spin, F in the limit $N \to \infty$, given by

$$-\beta F = \lim_{N\to\infty} (2N)^{-1} \ln Z_{2N} \tag{6}$$

and the resulting thermodynamical quantities; the internal energy per spin defined by

$$U = J\, d(\beta F)/dK \tag{7}$$

and the specific heat per spin by

$$C = -kK^2\, d^2(\beta F)/dK^2, \tag{8}$$

where $K = \beta J$.

In the following section we consider the Gaussian case. This serves to illustrate the basic method and enables us to summarize the Ising model results which are needed in Sec. 4 for the treatment of the more interesting spherical case.

[6] T. H. Berlin, and M. Kac, Phys. Rev. **86**, 821 (1952).

3. GAUSSIAN CASE

The analog of Fisher's summation over the decorating system here is the evaluation of the $x_{i,j}$, $y_{i,j}$ integrals in (3), which are readily performed using

$$\int_{-\infty}^{\infty} dx\, e^{-qx^2+bx} = \left(\frac{\pi}{q}\right)^{\frac{1}{2}} e^{b^2/4q} \tag{9}$$

with the result (remembering that $\mu_{i,j}^2 = 1$)

$$Z_{2N} = \exp\left[2NK^2 + \ln Z_N^{\mathrm{I}}(K^2)\right], \tag{10}$$

where

$$Z_N^{\mathrm{I}}(z) = \sum_{\mu_{i,j}=\pm 1} \exp\left[z \sum \mu_{i,j}(\mu_{i,j+1} + \mu_{i+1,j})\right] \tag{11}$$

is the partition function for the plane Ising model (with coupling constant z/β).

The free energy per spin by (6) is given by

$$-\beta F = K^2 - \tfrac{1}{2}\beta F_N^{\mathrm{I}}(K^2) \tag{12}$$

and the internal energy (7) and specific heat (8) per spin by

$$U = -K[J + U_{\mathrm{I}}(K^2)] \tag{13}$$

and

$$C = 2kK^2 - (kK^2/J)U_{\mathrm{I}}(K^2) + 2C_{\mathrm{I}}(K^2), \tag{14}$$

respectively, where $F_{\mathrm{I}}(z)$, $U_{\mathrm{I}}(z)$, and $C_{\mathrm{I}}(z)$ are the usual plane Ising-model free energy, internal energy, and specific heat per spin respectively, given by[1]

$$-\beta F_{\mathrm{I}}(z) = \lim_{N\to\infty} N^{-1} \ln Z_N^{\mathrm{I}}(z) = \ln(2\cosh 2z) + \frac{1}{2\pi} \int_0^{\pi} \ln \tfrac{1}{2}(1 + \sqrt{1 - \kappa^2 \sin^2\theta})\, d\theta, \tag{15}$$

$$U_{\mathrm{I}}(z) = J\, d(\beta F_{\mathrm{I}}(z))/dz = -J \coth(2z)[1 + (2/\pi)\kappa' K(\kappa)], \tag{16}$$

and

$$C_{\mathrm{I}}(z) = -kz^2\, d^2(\beta F_{\mathrm{I}}(z))/dz^2 = (2kz^2/\pi)(\coth 2z)^2\{2K(\kappa) - 2E(\kappa) - (1 - \kappa')[\tfrac{1}{2}\pi + \kappa' K(\kappa)]\}, \tag{17}$$

where

$$\kappa = 2\sinh 2z/\cosh^2 2z, \tag{18}$$

$$\kappa' = 2\tanh^2 2z - 1, \qquad \kappa^2 + \kappa'^2 = 1, \tag{19}$$

and $K(\kappa)$, $E(\kappa)$ are the complete elliptic integrals of the first and second kind given by

$$K(\kappa) = \int_0^{\pi/2} (1 - \kappa^2 \sin^2\theta)^{-\frac{1}{2}}\, d\theta \tag{20}$$

and

$$E(\kappa) = \int_0^{\pi/2} (1 - \kappa^2 \sin^2 \theta)^{\frac{1}{2}}\, d\theta. \tag{21}$$

It is immediately obvious that the critical behavior of the Gaussian case is the same as that of the plane Ising model, the latter coming from the singularity

$$K(\kappa) \sim \ln (4/\kappa') \quad \text{as} \quad \kappa' \to 0 \tag{22}$$

in the elliptic integral of the first kind. Thus the critical temperature $T_c^G = J/kK_c^G$ of our model is determined from

$$\kappa' = 2 \tanh^2 (2K_c^{G*}) - 1 = 0, \tag{23}$$

which is the same as for the plane Ising model with K^2 in place of K, and in the neighborhood of T_c^G we have

$$-U/2J \sim \sqrt{2}(1 + at \ln |t|) + K_c^G \tag{24}$$

and

$$C \sim \ln |t| \tag{25}$$

as $t = T - T_c^G \to 0$, where a is a constant.

4. SPHERICAL CASE

If we employ the representation of the delta function

$$\delta(x) = \frac{1}{2\pi i} \int_{-i\infty}^{i\infty} dq\, e^{qx} \tag{26}$$

for the spherical constraint (5) and substitute it into (3), the $x_{i,j}$, $y_{i,j}$ integrals can be performed using (9) of the previous section with the result

$$Z_{2N} = \pi^N (A_{2N}^s)^{-1} \frac{1}{2\pi i} \int_{\alpha-i\infty}^{\alpha+i\infty} dq \cdot \exp\left[2Nq - N \ln q + \frac{NK^2}{q} + \ln Z_N^I\left(\frac{K^2}{2q}\right)\right], \tag{27}$$

where α is chosen so that all the singularities of the integrand as a function of q are to the left of the line $q = \alpha$, and $Z_N^I(z)$ is the plane Ising partition function defined by (11). In the limit $N \to \infty$ the integral in (27) can be performed by the method of steepest descents with the result

$$-\beta F = \tfrac{1}{2}[G(q_s) - 1 - \ln 2], \tag{28}$$

where

$$G(q) = 2q - \ln q + K^2/q - \beta F_I(K^2/2q), \tag{29}$$

$F_I(z)$ is the free energy per spin of the plane Ising model, given by (15) and the saddle point q_s (if it exists) is determined from

$$(\partial G/\partial q)_{q_s} = 0, \qquad (\partial^2 G/\partial q^2)_{q_s} > 0. \tag{30}$$

In terms of the quantity $z = K^2/2q$, the saddle point is given by

$$\left(\frac{\partial G}{\partial z}\right)_{z_s} = -\frac{K^2}{z_s^2} + \frac{1}{z_s} + 2 - \frac{1}{J} U_I(z_s) = 0, \tag{31}$$

where $U_I(z)$ is the internal energy per spin of the plane Ising model given by (16). As a function of $1/z$, $-U_I(z)/2J$ is antisymmetric about the origin with

$$\lim_{1/z \to 0^\pm} \left[-\frac{1}{2J} U_I(z)\right] = \pm 1,$$

and as $1/z \to \infty$ it decreases monotonically to zero. It is therefore clear that (31) has a unique positive solution z_s which is a continuous function of K for all finite K. In addition, it is easily shown that $z_s > 2K^2$, and that as $T \to 0^+$, $z_s^{-1} \to 0^+$ and $K \to \infty$ in such a way that $K/z_s \to 2$, and that as $T \to \infty$, $z_s^{-1} \to \infty$, and $K \to 0$ in such a way that $K^2/z_s \to 1$.

Using (31) and (7), we find that the internal energy per spin is given by

$$U = -J[K/z_s - K^{-1}] \tag{32}$$

and since z_s is a continuous function of temperature, so is U. The specific heat per spin C is obtained straightforwardly from (8) and (32) with the result

$$C = kK^2\left[\frac{1}{z_s} + \frac{1}{K^2} - \frac{K}{z_s^2}\frac{dz_s}{dK}\right], \tag{33}$$

where, on differentiating (31) with respect to K and using the definition (17) of the plane Ising model specific heat $C_I(z)$,

$$dz_s/dK = 2Kz_s/(2K^2 - z_s + z_s C_I(z_s)k^{-1}). \tag{34}$$

Since $2K^2 - z_s$, z_s, and $z_s C_I(z_s)/k$ are all positive, the denominator of (34) is always positive. $C_I(z_s)$ moreover is a continuous function of z_s and consequently of temperature which goes to infinity like $\ln |z_s - z_{sc}|$ [cf. (25)] in the neighborhood of the Ising critical point defined by

$$2 \tanh^2 (2z_{sc}) = 1, \quad \text{i.e.,} \quad z_{sc} = 0.4406868. \tag{35}$$

At this point $dz_s/dK \to 0$ continuously and it follows that the specific heat given by (33) is a finite continuous function of temperature.

It remains to examine the slope of the specific-heat curve as a function of temperature in the neighborhood of the critical point defined by (35). Substituting (35) into (31) and using (16) we find first of all that the critical temperature T_c^s is given by

$$K_c^s = J/kT_c^s = (z_{sc} + (2 + 2^{\frac{1}{2}})z_{sc}^2)^{\frac{1}{2}} \tag{36}$$

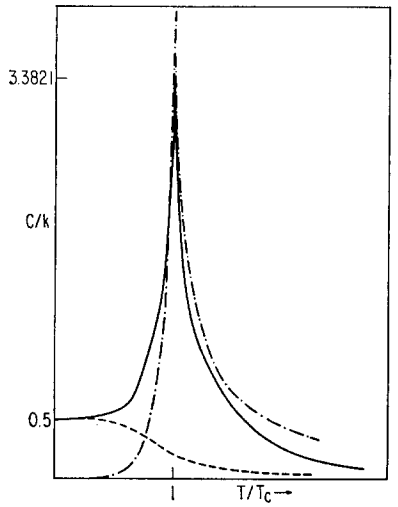

FIG. 1. Specific heat versus reduced temperature.————: Spherical–Ising;- - - - - -: Spherical;—·—·—·—: Ising.

and that in the neighborhood of T_c^s, using [cf. (24)]

$$-U_I(z)/J \sim \sqrt{2}(1 + ax \ln |x|) \tag{37}$$

in (31),

$$x \sim t/\ln |t|, \tag{38}$$

where $x = z_{sc} - z_s$, and $t = T_c^s - T$. Using (38), it is then a straightforward matter to show that

$$\frac{C(z_s)}{k} \sim \left[\frac{K_c^{s*}}{z_{sc}} + 1\right] - \frac{b}{\ln |t|} \quad \text{as} \quad t \to 0, \tag{39}$$

where $b > 0$ is a constant, and as a consequence that the specific heat has a cusp at T_c^s with slope going to $\pm\infty$ like $1/t[\ln |t|]^2$ as $t \to 0^{\pm}$. A rough sketch of $C(z_s)/k$ versus reduced temperature T/T_c^s is given in Fig. 1 where we have also shown for comparison in broken lines the corresponding curves for the two-dimensional spherical and Ising models.

5. SPONTANEOUS MAGNETIZATION

With an external magnetic field B acting on both Ising and continuum spins (which we assume for simplicity to have identical magnetic moments μ), the term

$$H' = -\mu B \sum (x_{i,j} + y_{i,j} + \mu_{i,j}) \tag{40}$$

is added to the Hamiltonian (2), and with the modified Hamiltonian in place of (2), the partition function $Z_{2N}(K, B, \mu)$ is given by (3), the free energy per particle $F(K, B, \mu)$ by (6), and the spontaneous magnetization is defined by

$$M(K, \mu) = \lim_{B \to 0} \left\{-\frac{\partial}{\partial B}[\beta F(K, B, \mu)]\right\}. \tag{41}$$

To evaluate $Z_{2N}(K, B, \mu)$ we proceed exactly as before, evaluating the $x_{i,j}$, $y_{i,j}$ integrals using (9), with the results

$$Z_{2N}(K, B, \mu) = \exp [N(2K^2 + L^2) + \ln Z_N^I(K^2, B, \mu(1 + 4K))] \tag{42}$$

for the Gaussian case, and

$$Z_{2N}(K, B, \mu) = \pi^N (A_{2N}^s)^{-1} \frac{1}{2\pi i} \int_{\alpha - i\infty}^{\alpha + i\infty} dq \cdot \exp\left[2Nq - N \ln q + \frac{N}{2q}(2K^2 + L^2) + \ln Z_N^I\left(\frac{K^2}{2q}, B, \mu\left(1 + \frac{2K}{q}\right)\right)\right] \tag{43}$$

for the spherical case, where $L = \beta B$, and $Z_N^I(z, B, \bar{\mu})$ is the plane Ising model partition function in a magnetic field B, with $\bar{\mu}$ the magnetic moment per Ising spin. It follows immediately that the spontaneous magnetization for the Gaussian case is given by

$$M_G = \tfrac{1}{2} M_I(K^2, \mu(1 + 4K)) \tag{44}$$

and for the spherical case by

$$M_s = \tfrac{1}{2} M_I(z_s, \mu(1 + 4z_s/K)) \tag{45}$$

where $M_I(z, \bar{\mu})$ is the spontaneous magnetization of the plane Ising model[7]

$$M_I(z, \bar{\mu}) = \bar{\mu}\left[\frac{\cosh^2 (2z)}{\sinh^4 (2z)}(\sinh^2 (2z) - 1)\right]^{1/8} \tag{46}$$

and z_s is given by (31). It then follows that M_G has the same critical behavior as M_I, namely

$$M_G \sim t^{1/8} \quad \text{as} \quad t = T_c^G - T \to 0^+, \tag{47}$$

with T_c^G defined by (23), but for the spherical case, using (38),

$$M_s \sim x^{1/8} \sim (t/\ln |t|)^{1/8} \text{ as } t = T_c^s - T \to 0^+ \tag{48}$$

with T_c^s given by (36), and $x = z_s - z_{sc}$, with z_{sc} given by (35).

ACKNOWLEDGMENTS

The author is grateful to Professor T. Matsubara for drawing his attention to Professor Syozi's work to Professor F. J. Dyson for helpful discussions and comments, and to Professor H. Suhl for the hospitality extended to him at the University of California, San Diego.

[7] C. N. Yang, Phys. Rev. 85, 809 (1952).

Theorem on Gravitational Fields with Geodesic Rays*

T. W. J. UNTI AND R. J. TORRENCE
Physics Department, University of Pittsburgh, Pittsburgh, Pennsylvania
(Received 4 March 1965; final manuscript received 15 August 1965)

If one of the Ruse vectors of a field is assumed to be a geodesic having nonvanishing divergence θ, curl ω, and complex shear σ, the only vacuum metrics that exist are found to be of the cylindrical type, where the geodesic rays obey $\theta^2 + \omega^2 = \sigma\bar{\sigma}$.

I. INTRODUCTION

A congruence of null geodesics in a hyperbolic Riemannian space–time can be characterized by its divergence, curl, and complex shear, here denoted by θ, ω, and σ, respectively.[1,2] Such a congruence is defined by any empty-space solution of the Einstein field equations possessing a family of geodesic rays (i.e., principal null directions of the Riemann tensor, also known as Ruse vectors). It is convenient to classify these solutions according to the vanishing or nonvanishing of the divergence θ, the curl ω, and the shear σ of the geodesic rays. Table I exhibits the various classes.[3–6]

TABLE I. Metrics with geodesic rays.[a]

Number	Class			Reference
	θ	ω	σ	
1	0	0	0	Kundt[3]
2	0	0	$\neq 0$	Not allowed by field equations
3	0	$\neq 0$	0	Not allowed by field equations
4	$\neq 0$	0	0	Robinson and Trautman[4]
5	0	$\neq 0$	$\neq 0$	*
6	$\neq 0$	0	$\neq 0$	Newman and Tamburino[6]
7	$\neq 0$	$\neq 0$	0	Kerr[6]
8	$\neq 0$	$\neq 0$	$\neq 0$	*

[a] The cases specially considered in this paper are denoted by an asterisk. A few authors are also listed, however their analyses were not meant to be exhaustive in every case.

The first classification given in Table I, that of divergence, curl, and shear-free metrics, was solved by Kundt in 1961.[3] By removing the restriction divergence = 0, Robinson and Trautman[4] found closed form solutions for the class of metrics admitting a family of hypersurface orthogonal shear-free null curves. However, line singularities occur in the radiation solutions. It was expected that the removal of two restrictions, divergence = 0 and shear = 0, would lead to an interesting generalization of the Robinson-Trautman solutions. Newman and Tamburino[5] calculated the curl-free metrics, but found only a very restricted class of solutions, containing no arbitrary functions. Kerr[6] removed the conditions divergence = 0, curl = 0 and found the class of shear-free metrics.

The removal of all three restrictions corresponds in principle to a broad subclass of Petrov type I nondegenerate fields. The specialization, suggested by Sachs,[2] consists in the assumption that one of the Ruse vectors of the Riemann tensor is a geodesic. Sachs derived some theorems concerning this subset of the algebraically general metrics, and showed the class is not empty by exhibiting a solution. He points out, however, that the solution merely serves to show there exist algebraically general metrics with geodesic rays, belonging to class 8, Table I. The example itself is not physically interesting, since the metrics given are static and cylindrically symmetric, in which the rays obey $\theta^2 + \omega^2 = \sigma\bar{\sigma}$.

The authors, in attempting to solve for all metrics having nonvanishing divergence, curl, and shear, but with geodesic rays, have found that this class is "almost empty." The fields all belong to the special (and uninteresting) category in which the rays are restricted by $\theta^2 + \omega^2 = \sigma\bar{\sigma}$.

II. THE FIELD EQUATIONS

The calculations were done using a form of the empty space Einstein field equations developed by Newman–Penrose[1] (hereafter referred to as NP). The NP equations are essentially linear combinations of the equations for the Riemann tensor expressed in terms of either Ricci rotation coefficients, or in terms of the spinor affine connection. For a complete derivation and discussions of the NP equations and examples of their use, the reader is referred

* Supported by Aerospace Research Labs., Office of Aerospace Research, U. S. Air Force.
[1] E. Newman and R. Penrose, J. Math. Phys. **3**, 366 (1962).
[2] R. Sachs, Proc. Roy. Soc. (London) **264**, 309 (1961). This interesting paper develops many theorems and geometrical concepts relevant to the present calculations.
[3] W. Kundt, Z. Physik **163**, 77 (1961).
[4] I. Robinson and A. Trautman, Proc. Roy. Soc. (London) **265**, 463 (1962).
[5] E. Newman and L. Tamburino, J. Math. Phys. **3**, 902 (1962).
[6] R. Kerr, Phys. Rev. Letters **11**, 237 (1963).

to Ref. 1 and to summaries in later publications.[5,7] A brief summary of the necessary results will be given here.

A family of null hypersurfaces $u = x^0 = \text{const}$ is introduced into the normal hyperbolic Riemannian manifold, so that[8]

$$g^{\mu\nu}u_{,\mu}u_{,\nu} = 0. \tag{2.1}$$

The second coordinate is an affine parameter $r = x^1$ along the null geodesics which lie in each hypersurface. Two parameters x^2, x^3 are necessary to label these null geodesics. A tetrad may be associated with the coordinate system. The first tetrad vector l^μ is tangent to the congruence of null geodesics. Another null vector n^μ may be chosen, pointing out of the u hypersurfaces, normalized by $l_\mu n^\mu = 1$. Finally, two spacelike vectors a^μ and b^μ are chosen, orthogonal to l^μ, n^μ, and each other. Instead of the real spacelike vectors a^μ and b^μ it is convenient to use the complex "null" vector

$$m^\mu = (a^\mu - ib^\mu)/\sqrt{2}. \tag{2.2}$$

The four null tetrad vectors l^μ, n^μ, m^μ, and $\bar{m}^\mu$ satisfy the orthonormality relations

$$\begin{aligned} l_\mu n^\mu &= -m_\mu \bar{m}^\mu = 1, \\ l_\mu l^\mu &= n_\mu n^\mu = m_\mu m^\mu = \bar{m}_\mu \bar{m}^\mu = l_\mu m^\mu \\ &= l_\mu \bar{m}^\mu = n_\mu m^\mu = n_\mu \bar{m}^\mu = 0, \end{aligned} \tag{2.3}$$

and therefore have the form

$$\begin{aligned} l^\mu &= \delta^\mu_1, \qquad m^\mu = w\delta^\mu_1 + \xi^A \delta^\mu_A, \\ n^\mu &= U\delta^\mu_1 + X^A \delta^\mu_A. \end{aligned} \tag{2.4}$$

The completeness relation

$$g^{\mu\nu} = l^\mu n^\nu + n^\mu l^\nu - m^\mu \bar{m}^\nu - \bar{m}^\mu m^\nu \tag{2.5}$$

permits us to express the metric in terms of the tetrad components. The following freedom in the choice of tetrad will be used later to simplify solutions: the spatial rotation

$$l^{\mu'} = l^\mu, \; n^{\mu'} = n^\mu, \; m^{\mu'} = m^\mu e^{iC}, \tag{2.6}$$

which depends on one real parameter $C(x^A)$, the combined tetrad-coordinate transformation

$$l^{\mu'} = [b(x^A)]^{-1} l^\mu, \; r' = b(x^A) r, \tag{2.7}$$

and the so-called null rotation

$$\begin{aligned} l^{\mu'} &= l^\mu, \quad m^{\mu'} = m^\mu + B(x^A) l^\mu, \\ n^{\mu'} &= n^\mu + \bar{B}(x^A) m^\mu \\ &\quad + B(x^A)\bar{m}^\mu + B(x^A)\bar{B}(x^A) l^\mu. \end{aligned} \tag{2.8}$$

In NP, 12 complex functions ρ, σ, α, $\cdots$ named "spin coefficients" are defined in terms of the tetrad. In this paper the following five spin coefficients are used:

$$\begin{aligned} &\rho = l_{\mu;\nu} m^\mu \bar{m}^\nu, \quad \sigma = l_{\mu;\nu} m^\mu m^\nu, \quad \tau = l_{\mu;\nu} m^\mu n^\nu, \\ &\alpha = \tfrac{1}{2}(l_{\mu;\nu} n^\mu \bar{m}^\nu - m_{\mu;\nu} \bar{m}^\mu \bar{m}^\nu), \\ &\beta = \tfrac{1}{2}(l_{\mu;\nu} n^\mu m^\nu - m_{\mu;\nu} \bar{m}^\mu m^\nu). \end{aligned}$$

Five complex tetrad components of the Weyl tensor are defined as

$$\begin{aligned} \psi_0 &= -C_{\alpha\beta\gamma\delta} l^\alpha m^\beta l^\gamma m^\delta, \\ \psi_1 &= -C_{\alpha\beta\gamma\delta} l^\alpha n^\beta l^\gamma m^\delta, \\ \psi_2 &= -\tfrac{1}{2} C_{\alpha\beta\gamma\delta} (l^\alpha n^\beta l^\gamma n^\delta + l^\alpha n^\beta m^\gamma \bar{m}^\delta), \\ \psi_3 &= C_{\alpha\beta\gamma\delta} l^\alpha n^\beta n^\gamma \bar{m}^\delta, \\ \psi_4 &= -C_{\alpha\beta\gamma\delta} n^\alpha \bar{m}^\beta n^\gamma \bar{m}^\delta. \end{aligned} \tag{2.9}$$

The Maxwell stress tensor is also resolved on the tetrad:

$$\begin{aligned} &\phi_0 = F_{\mu\nu} l^\mu m^\nu, \quad \phi_1 = \tfrac{1}{2} F_{\mu\nu}(l^\mu n^\nu + \bar{m}^\mu m^\nu), \\ &\phi_2 = F_{\mu\nu} \bar{m}^\mu n^\nu. \end{aligned} \tag{2.10}$$

The ψ's, ϕ's, spin coefficients and components of the tetrad constitute the variables of the NP equations. For convenience in our calculations, we group these variables into the following matrices:

$$\begin{aligned} &\mathbf{M} = \begin{bmatrix} \rho & \sigma \\ \bar{\sigma} & \bar{\rho} \end{bmatrix}, \quad \tilde{\mathbf{M}} = \text{transpose of } \mathbf{M}, \\ &\mathbf{T} = \begin{bmatrix} \tau \\ \bar{\tau} \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} \bar{\beta} \\ \bar{\alpha} \end{bmatrix}, \\ &\mathbf{E}^A = \begin{bmatrix} \xi^A \\ \bar{\xi}^A \end{bmatrix}, \quad \mathbf{W} = \begin{bmatrix} \bar{w} \\ w \end{bmatrix}, \quad \mathbf{F} = \begin{bmatrix} \psi_1 \\ \bar{\psi}_1 \end{bmatrix}, \quad \mathbf{G} = \begin{bmatrix} 0 \\ \psi_1 \end{bmatrix}. \end{aligned} \tag{2.11}$$

Not all of the NP equations are needed for the calculations in this paper. The equations that will be used here may be written as

$$\partial \mathbf{M}/\partial r = \mathbf{M}^2, \tag{2.12a}$$

$$\partial \mathbf{T}/\partial r = \mathbf{MT} + \mathbf{F}, \tag{2.12b}$$

$$\partial \mathbf{A}/\partial r = \tilde{\mathbf{M}}\mathbf{A} + \mathbf{G}, \tag{2.12c}$$

$$\partial \mathbf{E}^A/\partial r = \tilde{\mathbf{M}}\mathbf{E}^A, \tag{2.12d}$$

$$\partial \mathbf{W}/\partial r = \tilde{\mathbf{M}}\mathbf{W} - (\mathbf{A} + \mathbf{B}), \tag{2.12e}$$

[7] E. Newman and T. Unti, J. Math. Phys. 3, 891 (1962).

[8] Range and summation convention: lower-case Greek letters (μ, ν, $\cdots$), 0, 1, 2, 3; lower-case latin letters (k, m, $\cdots$), 2, 3; upper case latin (A, B, $\cdots$) 0, 2, 3. A comma denotes partial differentiation, a semicolon denotes covariant differentiation, and a bar denotes complex conjugate.

$$\partial\psi_1/\partial r = 4\rho\psi_1 \tag{2.13a}$$

$$\partial\psi_2/\partial r = 3\rho\psi_2 + \bar{w}(\partial\psi_1/\partial r) + \bar{\xi}^A(\partial\psi_1/\partial x^A) - 2\alpha\psi_1 + 2\rho\phi_1\bar{\phi}_1, \tag{2.13b}$$

$$\partial\phi_1/\partial r = 2\rho\phi_1, \tag{2.14}$$

$$\xi^A(\partial\psi_1/\partial x^A) - 2\beta\psi_1 + 4\psi_1(\rho w - \tau) + 3\sigma\psi_2 + 2\sigma\phi_1\bar{\phi}_1 = 0. \tag{2.15}$$

Equations (2.12) are first-order differential equations for the radial dependence of the spin coefficients. Equations (2.13) are Bianchi identities, Eq. (2.14) is a Maxwell field equation, and Eq. (2.15) may be considered a consistency relation which must be obeyed by the variables. [In NP it is shown that $\psi_0 = 0$ is the necessary and sufficient condition that l^μ is tangent to a geodesic ray. This has been taken into account in converting the NP equations into Eqs. (2.12)–(2.15).]

The spin coefficient ρ can be expressed in terms of the divergence and curl of l^μ, as[1]

$$\rho = \tfrac{1}{2}(-l^\mu_{;\mu} + i \text{ curl } l_\mu) \tag{2.16}$$

where

$$\text{curl } l_\mu \equiv [(l_{\mu;\nu} - l_{\nu;\mu})l^{\mu;\nu}]^{\frac{1}{2}}.$$

For simplicity of notation, we define

$$\rho = \theta + i\omega, \tag{2.17}$$

and call θ the "divergence" and ω the "curl" of the geodesic congruence. Strictly speaking, θ and ω are only proportional to the divergence and curl, respectively, with the proportionality factors given by

$$\theta = -\tfrac{1}{2}(\text{div } l_\mu), \qquad \omega = \tfrac{1}{2}(\text{curl } l_\mu). \tag{2.18}$$

In the following section, Eqs. (2.12)–(2.15) are solved under the general assumption that the determinant of $\mathbf{M}$ is nonzero. For the other (degenerate) case

$$\det \mathbf{M} = \rho\bar{\rho} - \sigma\bar{\sigma} = 0, \quad \theta^2 + \omega^2 = \sigma\bar{\sigma}. \tag{2.19}$$

This we refer to as the "cylindricity condition," since when $\theta^2 + \omega^2 = \sigma\bar{\sigma} \neq 0$, there is a single direction in which the rays do not spread out.[2]

The authors were originally interested in finding all the nondegenerate (i.e., $\det \mathbf{M} \neq 0$) solutions of class 8, Table I. As shown in Section III, no such metrics exist. For completeness, the other "algebraically general metrics with geodesic rays" of Table I were then examined. Case 6 is of no physical interest since the metrics contain no arbitrary constants. For case 5, Eq. (2.12a) reduces to the coupled equations

$$i\,\partial\omega/\partial r = -\omega^2 + \sigma\bar{\sigma} = 0, \qquad \partial\sigma/\partial r = 2i\omega\sigma,$$

which implies the cylindricity condition (2.19) since $\theta = 0$. Case 7 is algebraically special, as are cases 4 and 1. It may be emphasized that the authors did *not* solve for the fields constrained by the cylindricity condition, $\det \mathbf{M} = 0$, since the far field of a cylindrical metric does not correspond to that of a radiative source.[2]

The equations (2.12)–(2.15) were first solved in the absence of a Maxwell field. When it was found that the only vacuum fields with geodesic rays, having θ, ω, $\sigma \neq 0$, were of the cylindrical type, Maxwell fields of the class $\phi_0 = 0$ were included for greater generality. But in this case also interesting solutions were found not to exist.

III. CALCULATIONS

All the integrations of Eqs. (2.12), (2.13a), and (2.14) are with respect to r, so the "constants" of integration will be arbitrary functions of u, x^2, and x^3. A zero superscript will be used to indicate that the function is independent of r.

The general solution of Eqs. (2.12b)–(2.12e) can be written down, since by Eq. (2.12a)

$$\partial\mathbf{M}^{-1}/\partial r = -\mathbf{1}, \tag{3.1}$$

where $\mathbf{1}$ is the 2×2 unit matrix. Hence if Eq. (2.12b) is multiplied by $\mathbf{M}^{-1}$ we have

$$\mathbf{M}^{-1}(\partial\mathbf{T}/\partial r) = \partial(\mathbf{M}^{-1}\mathbf{T})/\partial r + \mathbf{T} = \mathbf{T} + \mathbf{M}^{-1}\mathbf{F}, \qquad \mathbf{T} = \mathbf{M}\left(\int \mathbf{M}^{-1}\mathbf{F}\,dr + \mathbf{T}^0\right). \tag{3.2}$$

Since Eq. (3.1) also holds for the transpose of $\mathbf{M}$, Eqs. (2.12c)–(2.12e) are solved similarly.

The integration constant $\mathbf{T}^0(x^A)$ may be transformed away by the null rotation Eq. (2.8). It is seen by the definition of τ in terms of the tetrad that

$$\mathbf{T}' = \mathbf{T} + \mathbf{M}\mathbf{B}, \quad \text{where} \quad \mathbf{B} \equiv \begin{bmatrix} B \\ \bar{B} \end{bmatrix}$$

under the null rotation. Applying this to Eq. (3.2) we may eliminate

$$\mathbf{T}^0 \equiv \begin{bmatrix} \tau^0 \\ \bar{\tau}^0 \end{bmatrix}.$$

The solution of Eq. (2.12a) is

$$\mathbf{M}^{-1} = -\begin{bmatrix} r + \bar{\rho}^0 & \sigma^0 \\ \bar{\sigma}^0 & r + \rho^0 \end{bmatrix}. \tag{3.3}$$

The real part of ρ^0 may be eliminated by shifting the origin, $r' = r + R^0(x^4)$. Using the spatial rotation Eq. (2.6) to make σ^0 pure imaginary, and introducing a convenient parameter $b(x^4)$ by Eq. (2.7), we obtain from Eq. (3.3)

$$\rho = -\frac{r + ibc}{r^2 + b^2(c^2 - s^2)}, \qquad \sigma = \frac{ibs}{r^2 + b^2(c^2 - s^2)}, \tag{3.4}$$

where c and s are arbitrary real functions of x^4.

The problem may now be divided into three parts: group (1) $c = s \neq 0$, group (2) $s < c$, group (3) $s > c$. If either c or s vanish everywhere the field reduces to one of the first seven cases listed in Table I and already discussed. For group (1) the integrations are easily performed with $b = 1$, and it is found that Eq. (2.15), the "consistency relation," cannot be satisfied. Therefore no such solutions exist. In group (2) the parameter b was set equal to $(c^2 - s^2)^{-\frac{1}{2}}$, while in group (3) it was set equal to $(s^2 - c^2)^{-\frac{1}{2}}$. It was a calculational aid to substitute

$$r = \tan f \tag{3.5}$$

for group (2). For group (3) the equivalent substitutions are

$$r = \tanh f \quad \text{for} \quad 0 < r \leq 1, \qquad r = \coth f \quad \text{for} \quad 1 < r. \tag{3.6}$$

In the remainder of the paper, the calculations will be performed only for group (2). For group(3) the calculations were done in an exactly analogous way to the steps for group (2), and the same result was found.

For group (2), then, we make two more convenient definitions:

$$p = \frac{c}{(c^2 - s^2)^{\frac{1}{2}}}, \qquad q = \frac{s}{(c^2 - s^2)^{\frac{1}{2}}}, \qquad p^2 - q^2 = 1. \tag{3.7}$$

[The "constants" $p(x^4)$, $q(x^4)$ may be considered trigonometric functions of an angle $\varphi(x^4)$.]

When the above substitutions are used in Eq. (3.4), the result is

$$\rho = -\cos f(\sin f + ip \cos f), \qquad \sigma = iq \cos^2 f. \tag{3.8}$$

Equation (3.2), the prototype for the solutions of Eqs. (2.12b)–(2.12e), may also be written in terms of the "angle" $f(x^4)$, since by definition (3.5), $dr = \sec^2 f \, df$. The integrations of Eqs. (2.12b)–(2.12e), (2.13) and (2.14) are then straightforward but tedious. The spin coefficients, tetrad components, etc., are found to be rather unwieldly functions of f, p, q, and "constants" of integration. The simplest solutions are those of the scalar equations (2.13a) and (2.14), yielding

$$\psi_1 = \psi_1^0 e^{-4ipf} \cos^4 f, \qquad \phi_1 = \phi_1^0 e^{-2ipf} \cos^2 f.$$

The other variables have more terms, ranging from three terms for ξ^4 to about seventy terms for ψ_2. When the variables obtained from the integrations are substituted into the consistency relation Eq. (2.15), it is found that Eq. (2.15) can only be satisfied if ψ_1 and ϕ_1 vanish.

IV. DISCUSSION

If $\psi_0 = 0$ (by assumption, so that the field has geodesic rays) and $\psi_1 = 0$ [by Eq. (2.15)], then the space is algebraically special by the Goldberg–Sachs theorem.[1] Hence the generality is lost. In attempting to find fields with geodesic rays belonging to classes 5 or 8, Table I, it is discovered that the only such fields that exist are restricted by $\theta^2 + \omega^2 = \sigma\bar{\sigma}$. This "cylindrical condition" makes them physically uninteresting (hence the authors have not solved for these metrics explicitly). In summary: Except for the Newman–Tamburino[5] metrics which contain no arbitrary functions, algebraically general fields with geodesic rays all obey the restriction $\theta^2 + \omega^2 = \sigma\bar{\sigma}$.

V. ACKNOWLEDGMENTS

We wish to acknowledge helpful discussions with E. Newman, A. Janis, R. Sachs, and J. Stachel. We are grateful to Richard Roman for checking some of the equations.

Overlap Integrals between Atomic Orbitals*

KLAUS RUEDENBERG AND KIYOSI O-OHATA†
Department of Chemistry, The Johns Hopkins University, Baltimore, Maryland and Department of Chemistry and Physics, Iowa State University, Ames, Iowa
AND
D. G. WILSON
Department of Physics, Iowa State University, Ames, Iowa
(Received 9 June 1964)

General formulas are developed for overlap integrals between Slater-type atomic orbitals of arbitrary integer quantum numbers (nlm) and parameter values $\rho = \frac{1}{2}R(\zeta + \zeta')$, $\tau = (\zeta - \zeta')/(\zeta + \zeta')$. The overlap integrals are expressed as finite trilinear forms of powers of ρ and certain auxiliary functions $I_{\alpha\beta}(\tau\rho)$, $J_{\mu\nu}^{\lambda}(\tau)$. The auxiliary functions are related to confluent hypergeometric functions and Jacobi polynomials, respectively, and stable recurrence procedures for their evaluation are given. The method is practical for use with an electronic computer and, even for high orbital quantum numbers, the loss in significant figures is found to be small for all parameter values. The overlap integral between two atomic orbitals $(Anlm)$ and $(Bn'l'm')$ is shown to be proportional to $R_{AB}{}^{|l-l'|}$.

INTRODUCTION

OVERLAP integrals are of importance in molecular quantum mechanics, when electronic wavefunctions are expressed in terms of atomic orbitals. Not only do they determine the metric of all algebraic operations in this nonorthogonal basis, but they also play an important role for the understanding of chemical binding and, frequently, are used in calculating various interatomic energy contributions. An adequate mathematical understanding as well as reliable and rapid methods of evaluation are therefore essential.

Although closed expressions are well known for the simpler cases,[1] generally applicable formulas or recurrence schemes would be preferable. Particularly desirable would be methods avoiding such defects of certain closed expressions as the loss of significant figures in certain argument ranges. For Slater-type atomic orbitals, two general approaches have been suggested previously, one by Roothaan,[2] another by Miller and Browne.[3] But there remains considerable room for improvement.

In the present note it is shown that overlap integrals between Slater-type atomic orbitals of arbitrary integer quantum numbers (nlm) can be reduced to confluent hypergeometric functions and certain polynomials, similar to Jacobi polynomials. A method of evaluating is outlined which has proved practical for use with an electronic computer. The derived expressions contain $R^{|l-l'|}$ as a factor in the overlap integral S, and this circumstance makes them convenient source functions for the Poisson equation $\Delta C = -4\pi S$ which, in the subsequent paper, is used to find expressions for the Coulomb integrals C.

1. OVERLAP INTEGRALS BETWEEN s ORBITALS

1.1 Reduction to Auxiliary Functions

Let

$$(An00) = (2\zeta)^{n+\frac{1}{2}}[4\pi(2n)!]^{-\frac{1}{2}}r_A^{n-1}\exp(-\zeta r_A) \qquad (1.1)$$

be a normalized, Slater-type s-atomic orbital, with orbital exponent ζ, centered on point A (r_A being the distance from A). Let $(Bn'00)$ be another atomic orbital of the same type, with orbital exponent ζ' and centered on a point B located at a distance R from A. The overlap integral

$$S_{nn'}^{000} = \int dV(An00)(Bn'00)$$

is most conveniently evaluated in elliptic coordinates, given by

$$\begin{aligned} \xi &= (r_A + r_B)/R,\ \eta = (r_A - r_B)/R, \\ \varphi &= \text{azimuthal angle around } A - B \text{ axis}, \qquad (1.2) \\ dV &= (R/2)^3 d\varphi d\xi d\eta(\xi^2 - \eta^2), \end{aligned}$$

* Work performed in part in the Ames Laboratory of the Atomic Energy Commission, Contribution No. 1550. This work was supported by National Science Foundation, Grant G 10351 and Grant GP 129.

† On leave of absence from Tokyo Institute of Technology, Tokyo, Japan.

[1] See for example: R. S. Mulliken, J. Chem. Phys. 17, 1248 (1949); C. C. J. Roothaan, *ibid.* 19, 1445 (1951); K. Ruedenberg, C. C. J. Roothaan, and W. Jaunzemis, *ibid.* 24, 201 (1956); and Technical Report, Laboratory of Molecular Structure and Spectra, Department of Physics, The University of Chicago (1952–1953), Part 2; M. Kotani, A. Amemiya, E. Ishiguro, and T. Kimura, *Table of Molecular Integrals* (Maruzen, Tokyo, 1955); H. Preuss, *Integraltafeln zur Quantenchemie* (Springer-Verlag, Berlin, 1956, 1957, 1960), Vols. I, II, and IV; A. Lofthus, Molecular Phys. 5, 105 (1962).

[2] C. C. J. Roothaan, J. Chem. Phys. 24, 947 (1956).

[3] I. Miller, and I. C. Browne, Technical Report, Molecular Physics Group, The University of Texas (1962).

and can be written as

$$S^{000}_{nn'} = (1+\tau)^{n+\frac{1}{2}}(1-\tau)^{n'+\frac{1}{2}} t_{n00} t_{n'00} T_{nn'}(\rho, \tau), \quad (1.3)$$

where

$$t_{n00} = 2^n n! \, [(2n)!]^{-\frac{1}{2}}, \quad (1.4)$$

$$T_{nn'}(\rho, \tau) = (n!\, n'!)^{-1} (\tfrac{1}{2}\rho)^{n+n'+1} \times \int_1^\infty d\xi \int_{-1}^1 d\eta (\xi+\eta)^n (\xi-\eta)^{n'} e^{-\rho\xi - \tau\rho\eta}, \quad (1.5)$$

and

$$\rho = \tfrac{1}{2}R(\zeta + \zeta'), \qquad \tau = (\zeta - \zeta')/(\zeta + \zeta'). \quad (1.6)$$

The functions T, while being of the same order of magnitude as the overlap integrals S, are more convenient for mathematical and numerical manipulation. Similar functions have been used previously by Roothaan,[2] who also found, using integration by parts, the recurrence relation

$$T_{\alpha\beta} = \tfrac{1}{2}T_{\alpha-1,\beta} + \tfrac{1}{2}T_{\alpha,\beta-1} + \Delta_{\alpha\beta}, \quad (1.7)$$

$$\Delta_{\alpha\beta} = [\rho^{\alpha+\beta}/(\alpha+\beta+1)!]\, e^{-\rho} I_{\alpha\beta}(\tau\rho), \quad (1.8)$$

where

$$I_{\alpha\beta}(x) = [(\alpha+\beta+1)!/\alpha!\,\beta!\,2^{\alpha+\beta+1}] \times \int_{-1}^1 d\eta\, e^{-x\eta}(1+\eta)^\alpha (1-\eta)^\beta. \quad (1.9)$$

An explicit form can be derived as follows. Making the substitution $s = \xi - 1$, we obtain

$$T_{nn'} = (n!\, n'!)^{-1}(\tfrac{1}{2}\rho)^{n+n'+1} \sum_{\alpha=0}^{n} \sum_{\beta=0}^{n'} \binom{n}{\alpha}\binom{n'}{\beta} e^{-\rho} \times \int_0^\infty ds\, s^{n+n'-\alpha-\beta} e^{-\rho s} \int_{-1}^1 d\eta (1+\eta)^\alpha (1-\eta)^\beta e^{-\tau\rho\eta}$$

$$T_{nn'} = e^{-\rho} \sum_{\alpha=0}^{n} \sum_{\beta=0}^{n'} \binom{n+n'-\alpha-\beta}{n-\alpha} \times \frac{\rho^{\alpha+\beta}}{(\alpha+\beta+1)!\, 2^{n+n'-\alpha-\beta}} I_{\alpha\beta}(\tau\rho), \quad (1.10)$$

an expression which indeed satisfies the recurrence formula (1.7).

The evaluation of the s-type integrals is therefore reduced to that of the auxiliary functions $I_{\alpha\beta}(x)$ of Eq. (1.9). We note that these functions are related to the confluent hypergeometric functions[4] $\Phi(a, c; x)$ by the relation

$$I_{\alpha\beta}(x) = e^{-x}\Phi(\beta+1, \alpha+\beta+2; 2x), \quad (1.11)$$

so that use can be made of the many recurrence relations known for the latter. It is however advantageous to formulate the overlap integral in terms of the functions $I_{\alpha\beta}$.

1.2 Properties of Auxiliary Functions

The following properties are readily verified by reference to the confluent hypergeometric functions[4] or, also, by direct mathematical methods:

$$I_{\alpha\beta}(0) = 1, \qquad e^{-x} \le I_{\alpha\beta}(x) \le e^x, \quad (1.12)$$

$$I_{\alpha\beta}(x) = I_{\beta\alpha}(-x), \quad (1.13)$$

$$I_{\alpha\alpha}(x) = [(2\alpha+1)!/\alpha!\,2^{\alpha+1}][2\pi/x^{2\alpha+1}]^{\frac{1}{2}} I_{\alpha+\frac{1}{2}}(x), \quad (1.14)$$

where I_μ is the modified Bessel function of the first kind, $I_\mu(x) = (i)^{-\mu} J_\mu(ix)$. Introducing the series expansion of the latter yields

$$I_{\alpha\alpha}(x) = \sum_{k=0}^{\infty} x^{2k} \binom{k+\alpha}{\alpha} \frac{(2\alpha+1)!}{(2\alpha+2k+1)!}. \quad (1.15)$$

This series can be considered as a special case of the general series[5]

$$I_{\alpha\beta}(x) = e^{-\gamma x} \sum_{k=0}^{\infty} c^k_{\alpha\beta} \left[\frac{2x}{(\alpha+\beta)}\right]^k \frac{(\alpha+\beta+1)!}{(\alpha+\beta+k+1)!},$$

$$\gamma = (\alpha-\beta)/(\alpha+\beta), \quad (1.16)$$

$$c^k_{\alpha\beta} = \sum_{i=0}^{k} (-1)^i \alpha^{k-i} \beta^i \binom{\alpha+j}{\alpha}\binom{\beta+k-j}{\beta},$$

because

$$c^k_{\alpha\alpha} = \begin{cases} 2^\alpha \binom{\alpha+n}{\alpha}, & \text{if } k = \text{even}, \\ 0, & \text{if } k = \text{odd}. \end{cases} \quad (1.17)$$

It results by expanding the exponential in definition (1.9) in terms of powers of $(\eta - \gamma)$, which is suggested by the fact that $\eta = \gamma$ is the maximum of the function which multiplies the exponential in the integrand of that equation. Other special cases are

$$I_{\alpha 0}(x) = e^{-x} \sum_{k=0}^{\infty} (2x)^k \frac{(\alpha+1)!}{(\alpha+k+1)!}, \quad (1.18)$$

$$I_{0\beta}(x) = e^{x} \sum_{k=0}^{\infty} (-2x)^k \frac{(\beta+1)!}{(\beta+k+1)!}. \quad (1.18')$$

There exist infinitely many distinct recurrence relations for the functions $I_{\alpha\beta}$. If the latter are represented by grid points, as in Fig. 1, the use of these recurrence relations can be indicated by diagrams. The figure illustrates the following selected formulas:

$$I_{\alpha\beta} = (I_{\alpha-1,\beta} - I_{\alpha,\beta-1})(\alpha+\beta+1)/2x, \quad (1.19)$$

[4] Bateman Manuscript Project, *Higher Transcendented Functions* edited by A. Erdelyi, W. Magnus, F. Oberhettinger, and F. G. Tricomi (McGraw-Hill Book Company, Inc., New York, 1953), Vol. I, p. 248.

[5] This series is not given in Ref. 4.

$$I_{\alpha\beta} = I_{\alpha,\beta-1} + I_{\alpha+1,\beta}2x(\alpha+1)/(\alpha+\beta+1)(\alpha+\beta+2), \quad (1.20)$$

$$I_{\alpha\beta} = I_{\alpha-1,\beta} - I_{\alpha,\beta+1}2x(\beta+1)/(\alpha+\beta+1)(\alpha+\beta+2), \quad (1.21)$$

$$I_{\alpha\beta} = [(\alpha+1)I_{\alpha+1,\beta}+(\beta+1)I_{\alpha,\beta+1}]/(\alpha+\beta+2), \quad (1.22)$$

$$I_{\alpha\beta} = [1-2x/(\alpha+\beta+2)]I_{\alpha,\beta+1} + I_{\alpha,\beta+2}2x(\beta+2)/(\alpha+\beta+2)(\alpha+\beta+3), \quad (1.23)$$

$$I_{\alpha\beta} = [1+2x/(\alpha+\beta+2)]I_{\alpha+1,\beta} - I_{\alpha+2,\beta}2x(\alpha+2)/(\alpha+\beta+2)(\alpha+\beta+3), \quad (1.24)$$

$$I_{\alpha\beta} = [1+2x(\beta-\alpha)/(\alpha+\beta+2)(\alpha+\beta+4)]I_{\alpha+1,\beta+1} + \frac{4x^2(\alpha+2)(\beta+2)}{(\alpha+\beta+3)(\alpha+\beta+4)^2(\alpha+\beta+5)} I_{\alpha+2,\beta+2}. \quad (1.25)$$

Furthermore one has the special cases

$$I_{\alpha 0} = (I_{\alpha-1,0} - e^{-x})(\alpha + 1)/2x, \quad (1.26)$$

$$I_{0\beta} = (e^{x} - I_{0,\beta-1})(\beta + 1)/2x, \quad (1.27)$$

$$I_{\alpha\alpha} = I_{\alpha+1,\alpha+1} + I_{\alpha+2,\alpha+2}x^2/(2\alpha + 3)(2\alpha + 5). \quad (1.28)$$

1.3. Evaluation of Auxiliary Functions

Required is, for a fixed value of x, a complete table of $I_{\alpha\beta}$'s for $0 \leq \alpha \leq A$, $0 \leq \beta \leq B$, where A and B are certain maximum values. By using Eq. (1.13) it is always possible to give x a positive value and this is assumed to have been done in the following description of a procedure which does not loose significant figures.

In any case, the *last row*, given by $\alpha = A$, $\beta = 0$ to B, is constructed first, whereupon the rest of the table is obtained without subtraction by vertical backward recurrence using alternately Eqs. (1.19) and (1.22). For the calculation of the last row two different paths have to be followed depending upon whether β is larger or smaller than $[C(x) - A - 1]$, where $C(x)$ is the integer obtained by rounding off $2x$, i.e., the integer lying between $(2x - \frac{1}{2})$ and $(2x + \frac{1}{2})$.

First Case: $\beta \leq C(x) - A - 1$. If there are β values of this kind, then $\beta = 0$ is one of these, so that necessarily $A + 1 \leq C(x)$. The first step is then to find I_{A0} by upwards recurrence, using Eq. (1.26) and starting with

$$I_{00} = \sinh x/x. \quad (1.29)$$

No figures are lost since $e^{-x} \ll I_{\alpha 0}$ for $0 \leq \alpha \leq A \leq C(x) - 1 \leq 2x - \frac{1}{2}$. Next the remaining $I_{A,\beta}$, for $1 \leq \beta \leq C(x) - A - 1$ are obtained by horizontal recurrence using Eq. (1.23) in the rearranged form:

$$I_{\alpha\beta} = \{[1 - (\alpha + \beta)/2x]I_{\alpha,\beta-1} + [(\alpha + \beta)/2x]I_{\alpha,\beta-2}\}(\alpha + \beta + 1)/\beta, \quad (1.30)$$

where $I_{A,-1}$ must be replaced by e^{-x}, when the formula is applied for $\beta = 1$. Since $A + \beta \leq C(x) - 1 \leq (2x - \frac{1}{2})$ holds in this case, one has

$$[1 - (A + \beta)/2x] \geq 1/4x, \quad (1.31)$$

indicating that the coefficients in Eq. (1.30) are always positive so that the only concellation occurs in calculating the coefficient of the first term. It can, however, be seen that the contribution from this term becomes smaller, as the cancellation inside the coefficient grows, so that in fact no error arises.

Second Case: $\beta \geq C(x) - A - 1$. If there are β values of this type, then $\beta = B$ is one of them so that necessarily $A + B + 1 \geq C(x)$. The first step is now to calculate I_{AB} and $I_{A,B-1}$, and we return to this point presently. After these two functions have been obtained, the remaining $I_{A,\beta}$, for $C(x) - A - 1 \leq \beta \leq B$ are found by the backward horizontal recurrence of Eq. (1.23). Since now $\beta + A + 2 \geq C(x) + 1 \geq 2x + \frac{1}{2}$, one has

$$[2(\alpha + \beta + 2)]^{-1} \leq [1 - 2x/(A + \beta + 2)] \leq 1, \quad (1.32)$$

which, in a manner similar to that discussed in the first case, again ensures that no error arises. The initial terms of this recurrence, i.e., the functions

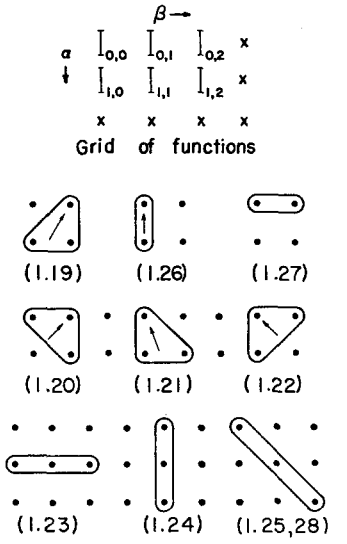

FIG. 1. Illustration of recurrence relations on function grid. The small arrows in some of the diagrams indicate the direction in which a particular relation can be used without subtraction for positive x.

I_{AB} and $I_{A,B-1}$, are arrived at in the following four steps.

Step 1 consists of finding the diagonal terms I_{MM} and $I_{M-1,M-1}$ by means of the series (1.15). Here M is an integer, at least as large as A and B, but moreover large enough that convergence of the series is as rapid as desired. Step 2 (required only if M is chosen larger than A and B) consists of recurring backward on the diagonal, using Eq. (1.28), until I_{NN} and $I_{N-1,N-1}$ are found, where N = maximum (A, B). In step 3, $I_{N,N-1}$ and $I_{N-1,N}$ are obtained by application of Eqs. (1.21) and (1.20), respectively. A subtraction occurs only in the former which becomes

$$I_{N,N-1} = I_{N-1,N-1} - I_{NN}x/(2N + 1).$$

No significant figures are lost, however, since (i) one has $I_{NN} < I_{N-1,N-1}$, as Eq. (1.28) shows, and (ii) one has $x/(2N + 1) \leq \frac{1}{2}$ because of $2x \leq A + B + \frac{1}{2} \leq 2N + \frac{1}{2}$. Step 4 finally leads to I_{AB} and $I_{A,B-1}$. If $N = B$, they are obtained from $I_{N-1,N-1}$ and $I_{N-1,N}$ without subtractions by vertical recurrence using Eq. (1.19) and Eq. (1.22). If $N = A$, they are obtained from $I_{N,N}$ and $I_{N-1,N}$ by horizontal recurrence using Eq. (1.23), which has already been discussed. If $N = A = B$, step 4 is, of course, superfluous.

2. OVERLAP INTEGRALS BETWEEN (s) ORBITALS AND ($l\sigma$) ORBITALS

2.1. Reduction to Auxiliary Functions

A general σ-type atomic orbital on center B is given by the expression

$$(Bn'l0) = (2\zeta')^{n'+\frac{1}{2}}[(2l + 1)/4\pi(2n)!]^{\frac{1}{2}} r_B^{n'-1} \times e^{-\zeta' r_B} P_l(\cos\theta_B), \quad (2.1)$$

where θ_B is the angle with the internuclear axis, pointing to atom A, and $P_l(x)$ is the Legendre polynomial, whose definition can be written in the form

$$P_l(x) = (2^l l!)^{-1}(d/dx)^l(x - 1)^l(x + 1)^l, \quad = 2^{-l}\sum_{\nu=0}^{l}\binom{l}{\nu}^2 (x - 1)^\nu (x + 1)^{l-\nu}. \quad (2.2)$$

The overlap integral between this atomic orbital and the s-atomic orbital of Eq. (1.1) can be written

$$S_{nn'}^{0l0} = \int dV(An00)(Bn'l0) = (1 + \tau)^{n+\frac{1}{2}}(1 - \tau)^{n'+\frac{1}{2}} t_{n00} t_{n'l0} T^l_{n,n'-l}(\rho, \tau), \quad (2.3)$$

where t_{n00} was given by Eq. (1.4) and

$$t_{rl0} = 2^r(n - l)!\, l!\, [(2l + 1)/(2n)!]^{\frac{1}{2}}, \quad (2.4)$$

and

$$T^l_{ik}(\rho, \tau) = (j!\, k!\, l!)^{-1}(\tfrac{1}{2}\rho)^{i+k+l+1} \times \int_1^\infty d\xi \int_{-1}^1 d\eta(\xi + \eta)^i(\xi - \eta)^{k+l} \times P_l(1 - \xi\eta/\xi - \eta)\exp(-\rho\xi - \tau\rho\eta). \quad (2.5)$$

Defining

$$\sigma = \tau\rho, \quad (2.6)$$

$$\partial_+ = -\tfrac{1}{2}(\partial/\partial\rho + \partial/\partial\sigma), \quad \partial_- = -\tfrac{1}{2}(\partial/\partial\rho - \partial/\partial\sigma), \quad (2.7)$$

and

$$T^l(\rho, \sigma) = (l!)^{-1}(\tfrac{1}{2}\rho)^{l+1}\int_1^\infty d\xi \int_{-1}^1 d\eta(\xi - \eta)^l \times P_l(1 - \xi\eta/\xi - \eta)\exp(-\rho\xi - \sigma\eta), \quad (2.8)$$

one can express T^l_{ik} as

$$T^l_{ik} = (j!\, k!)^{-1}\rho^{i+k+l+1}\partial_+^i\partial_-^k[\rho^{-l-1}T^l(\rho, \sigma)]. \quad (2.9)$$

Now the following identity can be derived from the definition of Eq. (2.2):

$$(\xi - \eta)^l P_l(1 - \xi\eta/\xi - \eta) = \sum_{\lambda=0}^{l}\binom{l}{\lambda}(\lambda!)^{-1}[\tfrac{1}{2}(\xi - 1)]^\lambda\left(\frac{d}{d\eta}\right)^\lambda[(1 + \eta)^\lambda(1 - \eta)^l]. \quad (2.10)$$

Substitution of this identity in Eq. (2.8), integration with respect to $s = (\xi - 1)$, and repeated integrations by parts with respect to η, yield

$$\rho^{-l-1}T^l = e^{-\rho}\sum_{\lambda=0}^{l}\frac{l!}{(l - \lambda)!\,(l + \lambda + 1)!}\frac{\sigma^\lambda}{\rho^{\lambda+1}} I_{\lambda l}(\sigma), \quad (2.11)$$

where the $I_{\lambda l}$ are the functions defined in Eq. (1.9). Substitution of this expression in Eq. (2.9) and further differential rearrangements then result in

$$T^l_{ik} = (j!\, k!)^{-1}\rho^{i+k+l+1} \times e^{-\rho}\sum_{\lambda=0}^{l}\sum_{r=0}^{j}\sum_{s=0}^{k}\binom{j}{r}\binom{k}{s}\frac{l!}{(l - \lambda)!\,(l + \lambda + 1)!} \times [\partial_+^{i-r}\partial_-^{k-s}\sigma^\lambda/\rho^{\lambda+1}][(\tfrac{1}{2} + \partial_+)^r(\tfrac{1}{2} + \partial_-)^s I_{\lambda l}(\sigma)]. \quad (2.12)$$

Since $I_{\lambda l}$ is independent of ρ, one has

$$(\tfrac{1}{2} + \partial_+)^r(\tfrac{1}{2} + \partial_-)^s I_{\lambda l} = \frac{(l + \lambda + 1)!\,(l + s)!\,(\lambda + r)!}{l!\,\lambda!\,(l + \lambda + r + s + 1)!} I_{\lambda+r,l+s}. \quad (2.13)$$

and, thus, one finally obtains

$$T^l_{ik} = e^{-\rho} \sum_{\lambda=0}^{l} \sum_{r=0}^{j} \sum_{s=0}^{k} k^{l\lambda}_{rs} \rho_{l+r+s} J^{\lambda}_{i-r,k-s}(\tau) I_{\lambda+r,l+s}(\tau\rho), \tag{2.14}$$

where

$$k^{l\lambda}_{rs} = \binom{l}{\lambda}\binom{l+s}{l}\binom{\lambda+r}{\lambda} \Big/ \binom{l+\lambda+r+s+1}{\lambda+1}(\lambda+1), \tag{2.15}$$

$$\rho_\alpha = \rho^\alpha/\alpha!, \tag{2.16}$$

and

$$J^{\lambda}_{\mu\nu}(\tau) = (\rho^{\mu+\nu+1}/\mu!\,\nu!)\partial^\mu_+ \partial^\nu_-(\sigma^\lambda/\rho^{\lambda+1}). \tag{2.17}$$

It will be shown in the next section that $J^{\lambda}_{\mu\nu}$ depends indeed upon $\tau = \sigma/\rho$ only. Since the functions $I_{\alpha\beta}$ have been discussed in the preceding sections, the functions $J^{\lambda}_{\mu\nu}$ are in fact the only new auxiliary functions needed for these integrals.

It should be noted that the functions T^l_{ik}, and hence the corresponding overlap integrals are proportional to the factor ρ^l.

2.2. Discussion of Auxiliary Functions

Using the definition (2.7) it is readily seen that

$$\rho^{\mu+\nu+1}(2\partial_+)^\mu(2\partial_-)^\nu(\sigma^\lambda/\rho^{\lambda+1}) = \sum_{\alpha=0}^{\mu}\sum_{\beta=0}^{\nu}(-1)^\alpha\binom{\mu}{\alpha}\binom{\nu}{\beta}\frac{(\lambda+\mu+\nu-\alpha-\beta)!}{(\lambda-\alpha-\beta)!}\tau^{\lambda-\alpha-\beta}, \tag{2.18}$$

where $\tau = \sigma/\rho$ and α, β are restricted by $\alpha + \beta \leq \lambda$. Using this result in the definition (2.17) of the functions $J^{\lambda}_{\mu\nu}$, the following two representations are obtained:

$$J^{\lambda}_{\mu\nu}(\tau) = (\tfrac{1}{2})^{\mu+\nu}\binom{\mu+\nu}{\mu} \times \sum_{\substack{\alpha=0 \\ (\alpha+\beta\leq\lambda)}}^{\mu}\sum_{\beta=0}^{\nu}\binom{\mu}{\alpha}\binom{\nu}{\beta}\binom{\lambda+\mu+\nu-\alpha-\beta}{\mu+\nu}\tau^{\lambda-\alpha-\beta}(-1)^\alpha \tag{2.19}$$

and

$$J^{\lambda}_{\mu\nu}(\tau) = (2^{\mu+\nu}\mu!\,\nu!)^{-1}(d/d\tau)^{\mu+\nu}[\tau^\lambda(\tau-1)^\mu(\tau+1)^\nu]. \tag{2.20}$$

The last equation exhibits a similarity to the Jacobi polynomials.[6]

Starting from that equation it is readily verified that the following relations hold:

$$J^{\lambda}_{\mu\nu}(-\tau) = (-1)^\lambda J^{\lambda}_{\nu\mu}(\tau), \tag{2.21}$$

[6] Reference 4, Vol. II, p. 168.

$$J^0_{\mu\nu}(\tau) = (\tfrac{1}{2})^{\mu+\nu}\binom{\mu+\nu}{\nu}, \tag{2.22}$$

$$J^{\lambda}_{00}(\tau) = \tau^\lambda, \tag{2.23}$$

$$J^{\lambda}_{\mu\nu} = \tau J^{\lambda-1}_{\mu\nu} + (J^{\lambda-1}_{\mu-1,\nu} - J^{\lambda-1}_{\mu-1,\nu})(\mu+\nu)/2\mu, \tag{2.24}$$

$$J^{\lambda}_{\mu\nu} = \tau J^{\lambda-1}_{\mu\nu} + (J^{\lambda-1}_{\mu,\nu-1} + J^{\lambda-1}_{\mu,\nu-1})(\mu+\nu)/2\nu. \tag{2.25}$$

By means of the two recurrence formulas (2.24), (2.25) the entire table of $J^{\lambda}_{\mu\nu}$ can be easily constructed from the initial functions given by Eqs. (2.22) and (2.23). The following procedure appeared satisfactory: Construct first the matrix for $\lambda = 0$, then for $\lambda = 1$, then for $\lambda = 2$, etc. For each matrix, use first Eq. (2.24) to construct the column $J^{\lambda}_{\mu 0}$, then use Eq. (2.25) to construct the rows $J^{\lambda}_{\mu\nu}$. The cancellation of figures was found not to be serious, the situation appeared to be similar to that occurring in the recurrence scheme of the Legendre polynomials.

3. GENERAL OVERLAP INTEGRALS

3.1. Reduction to Auxiliary Functions

Let

$$(Anlm) = (2\zeta)^{n+\frac{1}{2}}(2n!)^{-\frac{1}{2}} r_A^{n-1} e^{-\zeta r_A} Y_{lm}(\theta_A, \varphi), \tag{3.1}$$

$$(Bn'l'm') = (2\zeta')^{n'+\frac{1}{2}}(2n'!)^{-\frac{1}{2}} r_B^{n'-1} e^{-\zeta' r_B} Y_{l'm'}(\theta_B, \varphi), \tag{3.2}$$

be two Slater-type atomic orbitals on centers A and B, respectively. The z_A axis points to B, the z_B axis points to A, x_A and x_B are parallel, so are y_A and y_B. The $Y_{lm}(\theta\varphi)$ may be real or complex spherical harmonics. The overlap integral is

$$\int dV(Anlm)(Bn'l'm') = S^{ll'm}_{nn'}\,\delta_{mm'}$$

and has the same value for $(+m)$ and $(-m)$. In what follows, *m is therefore taken to be nonnegative.* Using the elliptic coordinates defined in Eq. (1.2) one obtains

$$S^{ll'm}_{nn'} = (1+\tau)^{n+\frac{1}{2}}(1-\tau)^{n'+\frac{1}{2}}[(2n)!\,(2n')!]^{-\frac{1}{2}}\rho^{n+n'+1} \times \int_1^\infty d\xi \int_{-1}^{1} d\eta(\xi+\eta)^n(\xi-\eta)^{n'} \times e^{-\rho\xi-\tau\rho\eta}\Theta^m_l(\cos\theta_A)\Theta^m_{l'}(\cos\theta_B), \tag{3.3}$$

where

$$\Theta^m_l(x) = [\tfrac{1}{2}(2l+1)(l-m)!/(l+m)!]^{\frac{1}{2}}P^m_l(x) \tag{3.4}$$

are the normalized associated Legendre functions, and

$$\cos\theta_A = (1+\xi\eta)/(\xi+\eta), \quad \cos\theta_B = (1-\xi\eta)/(\xi-\eta). \tag{3.5}$$

This integral can be reduced to several of the type treated in Sec. 2 by means of the following two transformations. The first is[7]

$$(-1)^m\left(\frac{r_A}{R}\right)^l\Theta_l^m(\cos\theta_A) = \sum_{k=m}^{l}(-1)^k\left(\frac{r_B}{R}\right)^k\Theta_k^m(\cos\theta_B)$$
$$\times\left[\frac{(2l+1)(l+m)!\,(l-m)!}{(2k+1)(k+m)!\,(k-m)!}\right]^{\frac{1}{2}}[(l-k)!]^{-1}, \tag{3.6}$$

and transfers all Legendre functions on to center B. The second is

$$(-1)^m\Theta_k^m(\cos\theta_B)\Theta_{l'}^m(\cos\theta_B)$$
$$= \sum_L \Theta_L^0(\cos\theta_B)\begin{pmatrix} kl'L \\ 00\,0\end{pmatrix}\begin{pmatrix} kl'L \\ m\;-m\,0\end{pmatrix}$$
$$\times[\tfrac{1}{2}(2k+1)(2l'+1)(2L+1)]^{\frac{1}{2}}, \tag{3.7}$$

and arises from expanding the product of spherical harmonics $Y_{k,m}Y_{l',-m}$ in terms of the spherical harmonics $Y_{L,0}$.[8] In Eq. (3.7),

$$\begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3\end{pmatrix}$$

represents the Wigner $3j$ symbol and, hence, the summation over L covers only the values

$$L = |l'-k|,\ |l'-k|+2,$$
$$|l'-k|+4,\ \cdots,\ (l'+k). \tag{3.8}$$

Combination of the two relations furnishes the expansion

$$(r_A/R)^l\Theta_l^m(\cos\theta_A)\Theta_{l'}^m(\cos\theta_B)$$
$$= [(2l+1)(l+m)!\,(l-m)!\,(2l'+1)(l'+m)!\,(l'-m)!]^{\frac{1}{2}}$$
$$\times\sum_{k=m}^{l}\sum_{L=|l'-k|}^{l'+k}\left(\frac{r_B}{R}\right)^k P_L(\cos\theta_B)(-1)^k$$
$$\times\frac{(2L+1)}{2(l-k)!}\begin{pmatrix} kl'L \\ 00\,0\end{pmatrix}\begin{pmatrix} kl'L \\ m\;-m\,0\end{pmatrix}$$
$$\times[(k+m)!\,(k-m)!\,(l'+m)!\,(l'-m)!]^{-\frac{1}{2}} \tag{3.9}$$

and insertion of this expression in Eq. (3.3) for the overlap integral finally yields

$$S_{nn'}^{ll'm} = (1+\tau)^{n+\frac{1}{2}}(1-\tau)^{n'+\frac{1}{2}}t_{nlm}t_{n'l'm}$$
$$\times\sum_{k=m}^{l}\sum_{L=|l'-k|}^{l'+k}B_{kL}^{n'l'm}\rho_{l-k}T_{n-l,n'+k-L}^{L}(\rho,\tau), \tag{3.10}$$

where

$$t_{nlm} = 2^n(n-l)!\,[(2l+1)(l+m)!\,(l-m)!/(2n)!]^{\frac{1}{2}}$$
$$= 2^n\left[(2l+1)\Big/\binom{2n}{2l}\binom{2l}{l+m}\binom{2(n-l)}{(n-l)}\right]^{\frac{1}{2}}, \tag{3.11}$$

$$B_{kL}^{n'l'm} = (-1)^k(2L+1)$$
$$\times\frac{(n'+k-L)!\,L!}{(n'-l')!}\begin{pmatrix} kl'L \\ 00\,0\end{pmatrix}\begin{pmatrix} kl'L \\ m\;-m\,0\end{pmatrix}$$
$$\times[(k+m)!\,(k-m)!\,(l'+m)!\,(l'-m)!]^{-\frac{1}{2}} \tag{3.12}$$

and the functions ρ_{l-k} and T_{mn}^L are defined by Eqs. (2.16) and (2.5), respectively.

From the last sentence in Section 2.1 follows that the general term in the summation of Eq. (3.10) is proportional to ρ^{l-k+L}. Taking into account the limits of the two summations in that equation it is readily seen that, in any case, the lowest occurring power of ρ is $|l-l'|$. *Hence the integral $S_{nn'}^{ll'm}$ is proportional $\rho^{|l-l'|}$.*

It is now possible to insert the expression (2.14) for T_{ik}^{λ} into Eq. (3.10) and thereby express the overlap integral in terms of the previously defined auxiliary functions J and I, whence

$$S_{nn'}^{ll'm} = (1+\tau)^{n+\frac{1}{2}}(1-\tau)^{n'+\frac{1}{2}}t_{nlm}t_{n'l'm}$$
$$\times e^{-\rho}\sum_{k=m}^{l}\sum_{L=|l'-k|}^{l'+k}\sum_{r=0}^{n-l}\sum_{s=0}^{n'+k-L}\sum_{\lambda=0}^{L}$$
$$\times B_{kL}^{n'l'm}k_{rs}^{L\lambda}\rho_{l-k}\rho_{L+r+s}J_{n-l-r,n'+k-L-s}^{\lambda}(\tau)I_{\lambda+r,L+s}(\tau\rho).$$

Making now the substitution $t = s + L$ and interchanging the order of the summations, one finally obtains

$$S_{nn'}^{ll'm} = (1+\tau)^{n+\frac{1}{2}}(1-\tau)^{n'+\frac{1}{2}}t_{nlm}t_{n'l'm}$$
$$\times e^{-\rho}\sum_{k=m}^{l}\sum_{r=0}^{n-l}\sum_{t=|l'-k|}^{n'+k}\sum_{\lambda}$$
$$\times K_{rt}^{k\lambda}\rho_{l-k+r+t}J_{n-l-r,n'+k-t}^{\lambda}(\tau)I_{\lambda+r,t}(\tau\rho), \tag{3.13}$$

with

$$K_{rt}^{k\lambda} = \binom{l-k+r+t}{l-k}\sum_L B_{kL}^{n'l'm'}k_{r,t-L}^{L\lambda}. \tag{3.13'}$$

In Eq. (3.13), λ runs from 0 to min $(t^*,\ l'+k)$, where $t^* = t$ or $(t-1)$, whichever of the two has the same partity as $(l'+k)$. In Eq. (3.13'), L runs from max $(\lambda,\ |l'-k|)$ to min $(t,\ l'+k)$ and, more-

[7] E. W. Hobson, *Theory of Spherical and Ellipsoidal Harmonics* (Cambridge University Press, London, 1931). See also R. J. Buehler and J. O. Hirschfelder, Phys. Rev. **83**, 628 (1951).

[8] A. R. Edmonds, *Angular Momentum in Quantum Mechanics* (Princeton University Press, Princeton, New Jersey, 1957), p. 63. M. E. Rose, *Elementary Theory of Angular Momentum* (John Wiley & Sons, Inc., New York, 1957), p. 61. M. Rotenberg, R. Bivins, N. Metropolis, and J. K. Wooten, *The 3-j and 6-j symbols* (Technology Press, Cambridge, Massachusetts, 1959), p. 9.

over, is restricted to the same parity as $(l' + k)$. It should be noted that $K_{rt}^{k\lambda}$ also depends on (l, n', l', m).

A convenient form for calculating $k_{r,t-L}^{L\lambda}$ was given in Eq. (2.15). A convenient form for calculating $B_{kL}^{n'l'm}$ is discussed in the next section. There it will also be shown that the coefficients K for the terms having the extreme values $t = (n' + k)$ and $\lambda = 0$ reduce to

$$K_{r,n'+k}^{k0} = (-1)^k \binom{n' + k}{l' + k}\binom{l' + k}{k}\binom{n' + l + r}{l - k} \times (r + n' + k + 1)^{-1}\, \delta_{m0}, \qquad (3.14)$$

showing that these terms must be omitted if $m \neq 0$.

3.2. Expressions for $K_{rt}^{k\lambda}$ and $B_{kL}^{n'l'm'}$

Inserting the definitions (3.12) and (2.15) in Eq. (3.13′) one obtains

$$K_{rt}^{k\lambda} = (-1)^k \binom{n' + k}{l' + k}\binom{l' + k}{\iota'} \times\left[\binom{2k}{k + m}\binom{2l'}{l' + m}\Big/\binom{2k}{k}\binom{2l'}{l'}\right]^{\frac{1}{2}} \times\left[\binom{l-k+r+t}{l-k}\binom{r+t}{r}\binom{t}{\lambda}\Big/\binom{r+t+\lambda}{t}(r+t+\lambda+1)\right] \times \sum_L (2L+1)\left[\binom{t-\lambda}{L-\lambda}\Big/\binom{n'+k}{L}\right]\binom{klL}{000}\binom{kl'L}{m-m0}. \qquad (3.15)$$

Putting $t = n' + k$ and $\lambda = 0$, the sum in L reduces to the orthogonality relation for the $3 - j$ symbols, viz.,

$$\sum_{L=|l'-k|}^{l'+k} (2L + 1)\binom{klL}{000}\binom{kl'L}{m\ -m0} = \delta_{m0}, \qquad (3.16)$$

so that in fact the expression (3.14) results.

Except for this special case, $K_{rt}^{k\lambda}$ can be conveniently calculated from Eq. (3.13′), using for $B_{kL}^{n'l'm'}$ the following expression:

$$B_{kL}^{n'l'm} = (-1)^e \times\left[\left(\frac{2L + 1}{2a + 1}\right)\binom{n' + k}{l' + k}\binom{L}{c}\binom{a}{d}\Big/\binom{n' + k}{L}\binom{2a}{L}\right] \times \sum_\alpha (-1)^\alpha \binom{2b}{l' - m - \alpha}\binom{2c}{k - m - \alpha}\binom{2d}{\alpha}, \qquad (3.17)$$

where

$$a = \tfrac{1}{2}(k + l' + L), \quad b = a - k, \quad c = a - l', \quad d = a - L, \qquad (3.18)$$

and α runs from max $(0, k - L - m, l' - L - m)$ to min $(k - m, l' - m, k + l' - L)$.

The equation for B is a consequence of Polo's[9] expression for the $3 - j$ symbols, viz.,

$$\binom{j_1\ j_2\ j_3}{m_1 m_2 m_3} = \delta_{0,m_1+m_2+m_3}\left[\frac{(j_1 + m_1)!\,(j_1 - m_1)!\,(j_2 + m_2)!\,(j_2 - m_2)!\,(j_3 + m_3)\ (j_3 - m_3)!}{(-j_1 + j_2 + j_3)!\,(j_1 - j_2 + j_3)!\,(j_1 + j_2 - j_3)!\,(j_1 + j_2 + j_3 + 1)!}\right]^{\frac{1}{2}} \times \sum_\alpha (-1)^{\alpha+j_1-j_2-m_3}\binom{j_1 + j_2 - j_3}{\alpha}\binom{j_1 - j_2 + j_3}{j_1 - m_1 - \alpha}\binom{-j_1 + j_2 + j_3}{j_1 + m_2 - \alpha} \qquad (3.19)$$

and

$$\binom{j_1 j_2 j_3}{000} = (-1)^g\left[\frac{(-j_1 + j_2 + j_3)!\,(j_1 - j_2 + j_3)!\,(j_1 + j_2 - j_3)!}{(j_1 + j_2 + j_3 + 1)!}\right]^{\frac{1}{2}}\binom{g}{g - j_1}\binom{j_1}{g - j_2} \qquad (3.20)$$

with

$$g = \tfrac{1}{2}(j_1 + j_2 + j_3), \qquad (3.21)$$

whence

$$\binom{j_1 j_2 j_3}{000}\binom{j_1\ j_2\ j_3}{m_1 m_2 m_3}[(j_1 + m_1)!\,(j_1 - m_1)!\,(j_2 + m_2)! \times (j_2 - m_2)!\,(j_3 + m_3)!\,(j_3 - m_3)!]^{-\frac{1}{2}} = \delta_{0,m_1+m_2+m_3}(-1)^{g+j_1-j_2-m_3} \times[(2g + 1)!]^{-1}\binom{g}{g - j_1}\binom{j_1}{g - j_2} \times \sum_\alpha (-1)^\alpha \binom{j_1+j_2-j_3}{\alpha}\binom{j_1-j_2+j_3}{j_1-m_1-\alpha}\binom{-j_1+j_2+j_3}{j_2+m_2-\alpha}.$$

As pointed out by Polo, use of binomial coefficients permits the exact calculation of the sum in double-precision fixed-point representation for quite high values of $j_1\ j_2\ j_3$.

3.3. Remarks on Evaluation

The evaluation of an overlap integral consists of the following steps: calculation of the functions of τ, ρ, and $\sigma = \rho\tau$, such as $J_{\mu\nu}^{\lambda}(\tau)$, ρ_k, $I_{\alpha\beta}(\sigma)$; calculation of all coefficients; and then calculation of the integral.

[9] S. R. Polo, *Studies on Crystal Field Theory* (RCA Laboratories, Princeton, New Jersey, 1961), Vol. I, Eqs. (8.14) and (8.22).

It can be seen that, for the functions $I_{\alpha\beta}$, it is necessary to construct the matrix consisting of all functions with indices

$$\begin{aligned} 0 &\leq \alpha \leq n + l', \\ 0 &\leq \beta \leq n' + l. \end{aligned} \tag{3.23}$$

For the functions $J^{\lambda}_{\mu\nu}$, the following set is required:

$$\begin{aligned} 0 &\leq \lambda \leq l + l', \\ 0 &\leq \mu \leq n - l, \\ 0 &\leq \nu \leq n' - l + 2 \min(l, l'). \end{aligned} \tag{3.24}$$

The functions ρ_k are needed for the index values

$$0 \leq k \leq (n + n'). \tag{3.25}$$

In the calculation of the integral positive and negative values occur for the coefficients $B^{n'l'm}_{kL}$ and the functions $J^{\lambda}_{\mu\nu}(\tau)$. Only small losses of significant figures, were observed however, when the exact cancellation expressed by Eq. (3.14) was avoided by taking into account that relationship explicitly. A check upon the accuracy can be obtained by using the following alternate formulation. Interchange of the centers A and B yields

$$S^{ll'm}_{nn'}(\rho, \tau) = S^{l'lm}_{n'n}(\rho, -\tau). \tag{3.26}$$

Taking into account the relations (1.13), (2.21) for $I_{\alpha\beta}$ and $J^{\lambda}_{\mu\nu}$, respectively, one finds from Eq. (3.13) the expression

$$\begin{aligned} S^{ll'm}_{nn'} &= (1 - \tau)^{n'+\frac{1}{2}}(1 + \tau)^{n+\frac{1}{2}} t_{nlm} t_{n'l'm} e^{-\rho} \\ &\times \sum_{k=m}^{l'} \sum_{r=0}^{n'-l'} \sum_{t=|l-k|}^{n+k} \sum_{\lambda} \bar{K}^{k\lambda}_{rt} \rho_{l'-k+r+t} J^{\lambda}_{n+k-t,n'-l'-r}(\tau) \\ &\times I_{t,\lambda+r}(\tau\rho), \end{aligned} \tag{3.27}$$

$$\bar{K}^{k\lambda}_{rt} = (-1)^{\lambda} \binom{l' - k + r + t}{l' - k} \sum_{L} B^{nlm}_{kL} k^{L\lambda}_{r,t-L}. \tag{3.28}$$

In these equations, λ runs from 0 to min $(t^*, l + k)$, where $t^* = t$ or $(t - 1)$, whichever of the two has the same parity as $(l + k)$, and L runs from max $(\lambda, |l - k|)$ to min $(t, l + k)$ and, moreover, is restricted to the same parity as $(l + k)$.

ACKNOWLEDGMENT

The authors wish to thank Kenneth Miller for his valuable assistance and many helpful comments.

Two-Center Coulomb Integrals between Atomic Orbitals*

KIYOSI O-OHATA† AND KLAUS RUEDENBERG
Department of Chemistry, The Johns Hopkins University, Baltimore, Maryland, and Departments of Chemistry and Physics, Iowa State University, Ames, Iowa
(Received 9 June 1964)

General analytical formulas are developed for the two-center Coulomb integrals between Slater-type atomic orbitals of arbitrary integral quantum numbers (nlm). The Coulomb integral C is obtained by integrating the Poisson equation $\Delta C = -4\pi S$, where S is the corresponding overlap integral. The Coulomb integrals are expressed as trilinear forms of powers and certain auxiliary functions $G_{\alpha\beta}{}^{\gamma}$, $H_{\alpha\beta}{}^{\gamma}$, $J_{\mu\nu}{}^{\lambda}$, which are related to confluent hypergeometric functions and Jacobi polynomials respectively. Stable recurrence procedures for their evaluation are given, and use with an electronic computer program showed the loss in numerical accuracy to be small for all argument values as well as for high quantum numbers.

INTRODUCTION

THE steepest mathematical obstacles of unrelativistic molecular quantum mechanics arise from the repulsions between different electrons. The energy contributions from these interactions are very hard to evaluate and it is for this reason that it has not yet been possible to extend exact self-consistent-field calculations to large molecules, let alone to implement any of the many schemes designed to deal with the correlation problem for such systems.

If the wavefunctions are expanded in terms of Slater-type atomic orbitals, the problem reduces to that of evaluating electron interaction integrals between orbitals on one, two, three, and four different atomic centers. Next to the one-center integrals, the two-center Coulomb integrals are the simplest and, therefore, one of the authors has pursued an approach by which all other many-center integrals are reduced to them.[1] However, even for the two-center Coulomb integrals there exists no general analytical treatment. Only numerical integration schemes have been available for the general case. One method consists in using the procedure developed by one of the authors for the two-center exchange integrals.[2] In another, the integration over the first electron is worked out analytically, whereupon the integration over the second is performed numerically.[3] Explicit formulas have been given, for the special cases involving $1s$, $2s$, and $2p$ orbitals,[4] but these closed expressions suffer from severe loss of significant figures in certain argument regions.

The present investigation contains a general analysis of the two-center Coulomb integrals between Slater-type atomic orbitals and yields analytical expressions for the general case, expressions which moreover are free of subtractive loss of numerical accuracy. The approach is based on the observation that a Coulomb integral C and a corresponding overlap integral S are related to each other by the Poisson equation $\Delta C = -4\pi S$ so that the former can be expressed as the potential integral $C = \int d\tau S/r$. Using the results obtained for the overlap integrals in the preceding paper,[5] it proves possible to carry out this integration analytically. The Coulomb integral finally appears in terms of a set of auxiliary functions, integrals over confluent hypergeometric functions, for which stable recurrence schemes are derived. The resulting expressions have proved practical for the use with an electronic computer.

In the following, repeated reference will be made to the results reported in Ref. 5, to be quoted as "the preceding paper."

1. COULOMB INTEGRAL AS PONTENTIAL OF OVERLAP INTEGRAL

1.1. Coulomb Integral between Basic Change Distribution

Let $f^*(\mathbf{r}_A)$ be an electrostatic charge distribution given explicitly in terms of $\mathbf{r}_A = \mathbf{r} - \mathbf{A}$, the position vector relative to a point $\mathbf{A}$. Let $g(\mathbf{r}_B)$ be another

* Work performed in part in the Ames Laboratory of the Atomic Energy Commission, Contribution No. 1549. This work was supported by National Science Foundation, Grant G 10351 and Grant GP 129.

† On leave of absence from Tokyo Institute of Technology, Tokyo, Japan.

[1] K. Ruedenberg (to be published).

[2] K. Ruedenberg, in *Molecular Orbitals in Chemistry, Physics, and Biology*, edited by B. Pullman and P. O. Löwdin (Academic Press Inc., New York, 1964), p. 215.

[3] A. C. Wahl, P. E. Cade, and C. C. J. Roothaan, Technical Report 1962-3, part 2, p. 148 (Laboratory of Molecular Structure and Spectra, The University of Chicago).

[4] C. C. J. Roothaan, J. Chem. Phys. 19, 1445 (1951); K. Ruedenberg, C. C. J. Roothaan, and W. Jaunzemis, J. Chem. Phys. 24, 201 (1956). Also M. Kotani, A. Amemiya, E. Ishiguro, and T. Kimura, *Tables of Molecular Integrals* (Maruzen, Tokyo, 1955); H. Preuss, *Integraltafeln zur Quantenchemie* (Springer–Verlag, Berlin, 1956, 1957, 1960).

[5] K. Ruedenberg, K. O-Ohata, and D. G. Wilson, J. Math. Phys. 7, 539 (1966) (preceding paper).

charge distribution given explicitly in terms of $\mathbf{r}_B = \mathbf{r} - \mathbf{B}$, the position vector relative to a point **B**. The electrostatic interaction energy between these two charge distributions,

$$C(\mathbf{R}) = \int dV_1 \int dV_2 \, f^*(\mathbf{r}_{A1}) g(\mathbf{r}_{B2})/r_{12} = [f^* \mid g], \tag{1.1}$$

depends on $\mathbf{R} = \mathbf{B} - \mathbf{A}$, the location of **B** relative to **A**. This integral represents a "basic two-center Coulomb integral" if $f(\mathbf{r})$ and $g(\mathbf{r})$ are the "basic charge distributions"

$$f(\mathbf{r}_A) = [Anlm] = f(r_A) Y_{lm}(\theta_A \varphi_A), \tag{1.2}$$

$$g(\mathbf{r}_B) = [Bn'l'm'] = g(r_B) Y_{l'm'}(\theta_B \varphi_B), \tag{1.2'}$$

where[6]

$$f(r) = [(2l+1)/4\pi]^{\frac{1}{2}} [2^{n+1}(n+1)!]^{-1} (2\zeta)^{n+2} r^{n-1} \times \exp(-2\zeta r), \tag{1.3}$$

$$g(r) = [(2l'+1)/4\pi]^{\frac{1}{2}} [2^{n'+1}(n'+1)!]^{-1} (2\zeta')^{n'+2} r^{n'-1} \times \exp(-2\zeta' r), \tag{1.3'}$$

and $Y_{lm}(\theta\varphi)$ are real or complex spherical harmonics.

The *basic charge distribution* $[nlm]$ is related to the normalized Slater-type atomic orbital (nlm), defined in Eq. (3.1) of the previous paper, by the equality

$$[nlm] = \{\zeta^{\frac{1}{2}} [(2n)!\,(2l+1)]^{\frac{1}{2}} / 2^{2n+1} (n+1)!\, \pi^{\frac{1}{2}}\} (nlm)_{2\zeta}, \tag{1.4}$$

where $(nlm)_{2\zeta}$ means that in the Slater orbital (nlm), the orbital exponent ζ is to be replaced by (2ζ). The complex spherical harmonics are given by[7]

$$Y_{lm}(\theta\varphi) = \Theta_l^m(\cos\theta)(2\pi)^{-\frac{1}{2}} e^{im\varphi}, \tag{1.5}$$

where

$$\Theta_l^m(z) = \epsilon_m [\tfrac{1}{2}(2l+1)(l-|m|)!/(l+|m|)!]^{\frac{1}{2}} P_l^{|m|}(z) \tag{1.6}$$

with

$$\epsilon_m = \begin{cases} (-1)^m, & \text{for } m \geq 0, \\ 1, & \text{for } m \leq 0. \end{cases} \tag{1.7}$$

The real spherical harmonics are defined by

$$\mathcal{Y}_{lm}(\theta, \varphi) = \epsilon_m \Theta_l^m(\cos\theta)[(1+\delta_{m0})\pi]^{-\frac{1}{2}} \cos m\varphi, \quad m \geq 0, \tag{1.8}$$

$$\mathcal{Y}_{lm}(\theta, \varphi) = \Theta_l^m(\cos\theta)\pi^{-\frac{1}{2}} \sin m\varphi, \quad m < 0, \tag{1.8'}$$

and related to the complex spherical harmonics by

$$\mathcal{Y}_{l,m} = [Y_{l,-m} + (-1)^m Y_{l,m}]/\sqrt{2}, \quad m > 0, \tag{1.9}$$

$$\mathcal{Y}_{l,m} = [Y_{l,m} - (-1)^m Y_{l,-m}]i/\sqrt{2}, \quad m < 0, \tag{1.9'}$$

$$\mathcal{Y}_{l,0} = Y_{l0}. \tag{1.9''}$$

Since r_{12} is a function of $\cos(\varphi_1 - \varphi_2)$ only, the φ-dependent part of the integral (1.1) results in the factor

$$\begin{aligned} &\int_0^{2\pi} d\varphi_1 \int_0^{2\pi} d\varphi_2 \phi_m^*(\varphi_1) \phi_{m'}(\varphi_2) F[\cos(\varphi_1 - \varphi_2)], \\ &= \int_0^{2\pi} d\varphi F(\cos\varphi) \int_0^{2\pi} d\varphi_1 \phi_m^*(\varphi_1) \phi_{m'}(\varphi_1 + \varphi), \\ &= \delta_{mm'} \int_0^{2\pi} d\varphi \cos m\varphi F(\cos\varphi), \end{aligned} \tag{1.10}$$

which has the same value for the real and the complex case. The same holds therefore for the Coulomb integral

$$[[Anlm]^* \mid [Bnlm]] = \delta_{mm'} C_{nn'}^{ll'|m|}, \tag{1.11}$$

and we can therefore use the complex harmonics to find the general expression for $C_{nn'}^{ll'm}$ $(m \geq 0)$.

When molecular electronic wavefunctions are expressed in terms of atomic orbitals, there occur two-center Coulomb integrals between products of normalized Slater-type atomic orbitals. The reduction of these Coulomb integrals to the basic Coulomb integrals of Eq. (1.11) are discussed in Sec. 5.

1.2. Relation between Coulomb Integral and Overlap Integral

The basis of the present derivation is the observation that the integral $C(\mathbf{R})$ of Eq. (1.1) can be regarded as the potential arising from a source distribution defined by the overlap integral

$$S(\mathbf{R}) = \int dV \, f^*(\mathbf{r}_A) g(\mathbf{r}_B). \tag{1.12}$$

In other words, the Poisson equation

$$\Delta_R C(\mathbf{R}) = -4\pi S(\mathbf{R}) \tag{1.13}$$

and its inverse

$$C(\mathbf{R}) = \int dV' \, S(\mathbf{R}')/|\mathbf{R} - \mathbf{R}'| \tag{1.14}$$

are valid between C and S.

Proof of Eq. (1.14). Figure 1 shows the two centers A and B, the points (1) and (2) at which the two volume elements dV_1 and dV_2 are taken, the vectors

[6] This definition differs from the one used by Roothaan in Ref. 4. The latter's charge distributions are obtained by multiplying the right-hand side of Eq. (1.3) with the factor $[(n+1)!\, 2^{n+l+1}/(n+l+1)!]$.

[7] A. E. Edmonds, *Angular Momentum in Quantum Mechanics* (Princeton University Press, Princeton, New Jersey, 1957), p. 24, Eq. (2.5.29).

$\mathbf{r}_{A1}$, $\mathbf{r}_{B2}$, $\mathbf{r}_{12}$ and moreover, a third center C placed so that the points $(BC12)$ form a parallelogram. For a fixed value of $\mathbf{r}_{A1}$, there exists a unique correspondence between $\mathbf{r}_{B2}$ and

$$\mathbf{R}' = \mathbf{r}_{A1} - \mathbf{r}_{C1} = \mathbf{r}_{A1} - \mathbf{r}_{B2}. \tag{1.15}$$

Hence, substituting

$$\mathbf{r}_{B2} = \mathbf{r}_{A1} - \mathbf{R}' \tag{1.16}$$

for $\mathbf{r}_{B2}$ in the integral (1.1), one can consider $\mathbf{R}'$ instead of $\mathbf{r}_{B2}$ as integration variable, whereby

$$C(\mathbf{R}) = \int dV_1 \int dV'\, f^*(\mathbf{r}_{A1}) g(\mathbf{r}_{A1} - \mathbf{R}')/|\mathbf{R} - \mathbf{R}'|, \tag{1.17}$$

with dV' the volume element corresponding to $\mathbf{R}'$. From this equation follows

$$C(\mathbf{R}) = \int dV'\, |\mathbf{R} - \mathbf{R}'|^{-1}\, S(\mathbf{R}') \tag{1.18}$$

with

$$S(\mathbf{R}') = \int dV_1\, f^*(\mathbf{r}_{A1}) g(\mathbf{r}_{C1}), \tag{1.19}$$

where

$$\mathbf{C} - \mathbf{A} = \mathbf{R}'. \qquad \text{q.e.d.} \tag{1.20}$$

Equation (1.13) can also be proven directly by means of Fourier transformation. If $\varphi(\mathbf{k})$ and $\gamma(\mathbf{k})$ are the Fourier transforms of $f(\mathbf{r})$ and $g(\mathbf{r})$, respectively, i.e.,

$$f(\mathbf{r}) = (2\pi)^{-\frac{3}{2}} \int d\mathbf{k}\, \varphi(\mathbf{k}) \exp(i\mathbf{kr}), \tag{1.21}$$

$$g(\mathbf{r}) = (2\pi)^{-\frac{3}{2}} \int d\mathbf{k}\, \gamma(\mathbf{k}) \exp(i\mathbf{kr}), \tag{1.21'}$$

then it is easily seen that $C(\mathbf{R})$ and $S(\mathbf{R})$ are given by[8]

$$S(\mathbf{R}) = \int d\mathbf{k}\, \varphi^*(\mathbf{k})\gamma(\mathbf{k}) \exp(i\mathbf{kR}), \tag{1.22}$$

$$C(\mathbf{R}) = 4\pi \int d\mathbf{k}\, \varphi^*(\mathbf{k})\gamma(\mathbf{k})\, |\mathbf{k}|^{-2} \exp(i\mathbf{kR}), \tag{1.23}$$

whence it is obvious that ΔC is equal $-4\pi S$.

Remark 1. The overlap integral $S(\mathbf{R})$ defined by Eq. (1.12) is taken between the charge distributions $[nlm]$ of Eq. (1.3). It therefore differs slightly, as indicated by Eq. (1.4), from the overlap integral $S_{nn'}^{ll'm}$ discussed in the preceding paper, which was taken between the normalized Slater orbitals (nlm).

[8] After the present results had been derived, F. P. Prosser and C. H. Blauchard [J. Chem. Phys. **36**, 1112 (1962)] as well as M. Geller [J. Chem. Phys. **39**, 853 (1963)] have independently suggested the use of Fourier transforms in molecular integrals.

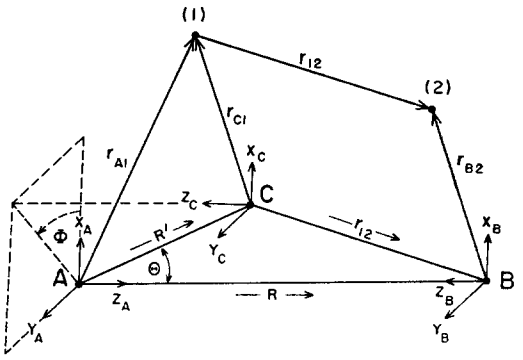

FIG. 1. Relative positions of the points A, B, C, (1), (2). Coordinate systems at points A, B, C. Definition of R', Θ, Φ.

Remark 2. In Eq. (1.14), $S(\mathbf{R}')$ has to be known as a function of $\mathbf{R}'$. This function results by moving the charge distribution $g(\mathbf{r}_B)$ while leaving the charge distribution $f(\mathbf{r}_A)$ fixed. This means that $g(\mathbf{r}_B)$ has to be moved by *displacing* its origin B *without rotating* it around this origin.

Remark 3. The result of Eq. (1.14) can be generalized to more general types of integrals. From the first proof given, it is obvious that the following identity is also valid:

$$\int dV_1 \int dV_2\, f^*(\mathbf{r}_{A1}) g(\mathbf{r}_{B2}) h(\mathbf{r}_{12}) = \int dV'\, h(\mathbf{R} - \mathbf{R}') S(\mathbf{R}'). \tag{1.24}$$

Such integrals may occur in connection with certain correlation functions.

1.3. Expression for Overlap Integral

According to the previous remarks, $S(\mathbf{R}')$ is the overlap integral between the charge distribution $[Anlm]$ at point A and the charge distribution $[Cn'l'm]$ at point C, where the angular dependence of $[Cn'l'm]$ is embodied in the spherical harmonics defined with respect to an axial system with origin at C, *but parallel to the original axis system at* B. The axial systems at A, B, and C are also shown in Fig. 1.

In order to obtain a formula for $S(\mathbf{R}')$, we express $[Anlm]$ in terms of spherical harmonics defined with respect to a coordinate system on A whose z axis points towards C. Similarly, we express $[Cn'l'm]$ in terms of spherical harmonics with respect to a coordinate system on C whose z axis points towards A. Moreover, the x and y axes on A and C are chosen parallel to each other. Now if an axial system is rotated from an old orientation (xyz) to a new

orientation $(x'y'z')$, then the corresponding spherical harmonics transform as follows:[9]

$$Y_{lm}(\theta\varphi) = \sum_{\nu=-l}^{l} Y_{l\nu}(\theta'\varphi')\, D^{l}_{\nu m}(\alpha, \beta, \gamma), \qquad (1.25)$$

where $D^{l}_{\nu m}(\alpha, \beta, \gamma)$ is the irreducible representation matrix of order $(2l + 1)$ for the rotation described by the Euler angles α, β, γ.[10] It is readily seen that, in the present case, one has $\alpha = \Phi$, $\beta = \Theta$, $\gamma = 0$ for the transformation on center A and $\alpha = \Phi$, $\beta = -\Theta$, $\gamma = 0$ for the transformation on center C, if R', Φ, Θ are the spherical coordinates locating the point C in terms of the *original* coordinate system (shown in Fig. 1) on center A. Inserting these transformations in the overlap integral, and noting that

$$[D^{l}_{\nu m}(\alpha\beta\gamma)]^* = (-1)^{\nu-m} D^{l}_{-\nu,-m}(\alpha\beta\gamma), \qquad (1.26)$$

$$D^{l}_{\nu m}(\alpha, -\beta, \gamma) = (-1)^{\nu-m} D^{l}_{\nu m}(\alpha\beta\gamma), \qquad (1.27)$$

one obtains

$$S(\mathbf{R}') = \sum_{\nu=-\min(l,l')}^{\min(l,l')} D^{l}_{-\nu-m}(\Phi\Theta 0)\, D^{l'}_{\nu m}(\Phi\Theta 0) \times S_{nl\nu,n'l'\nu}(R'), \qquad (1.28)$$

where

$$S_{nl\nu,n'l'\nu'} = \int dV\, [Anl\nu]^*\, [Cn'l'\nu'] = S_{nl\nu,n'l'\nu}(R')\, \delta_{\nu\nu'} \qquad (1.29)$$

is the overlap integral between the charge distributions on A and C aligned with respect to the *new*, rotated coordinate systems, which is the usual alignment in discussing overlap integrals. As indicated $S_{nl\nu,n'l'\nu'}$ vanishes for $\nu \neq \nu'$, a fact which has been used in arriving at Eq. (1.28), and it moreover depends on the absolute value $R' = |\mathbf{R}'|$ only.

1.4. Integration of Possion Equation

In order to obtain the Coulomb integral $C(\mathbf{R})$, we substitute in Eq. (1.14) the expression (1.28) for $S(\mathbf{R}')$, and the Laplace expansion[11]

$$|\mathbf{R} - \mathbf{R}'|^{-1} = \sum_{L=0}^{\infty} (R^{L}_{<}/R^{L+1}_{>}) P_L(\cos\Theta) \qquad (1.30)$$

for $|\mathbf{R} - \mathbf{R}'|^{-1}$. Here, $R_< = \min(R, R')$, $R_> = \max(R, R')$ and, as mentioned before, R', Φ, Θ are the spherical coordinates of point C in terms of the original coordinate system on A. The integration over the angles Θ, Φ can now be carried out explicitly since[12]

$$P_L(\cos\Theta) = D^{L}_{00}(\Phi, \Theta, 0),$$

and

$$\int_0^{2\pi} d\Phi \int_0^{\pi} d\Theta \sin\Theta\, D^{l}_{-\nu,-m}(\Phi, \Theta, 0)\, D^{l'}_{\nu',m'}(\Phi, \Theta, 0) \times D^{L}_{00}(\Phi, \Theta, 0) = 4\pi \begin{pmatrix} ll'L \\ m\ -m\ 0 \end{pmatrix} \begin{pmatrix} ll'L \\ \nu\ -\nu\ 0 \end{pmatrix} \delta_{\nu\nu'}\, \delta_{mm'}, \qquad (1.31)$$

where $\begin{pmatrix} j_1\ j_2\ j_3 \\ m_1 m_2 m_3 \end{pmatrix}$ are the Wigner $3j$-symbols.[13] Formulas for the calculation of the symbols were given in the preceding paper.[14] They vanish unless $|l - l'| \leq L \leq l + l'$. Hence we obtain for the Coulomb integral

$$C^{ll'm}_{nn'}(R) = 4\pi \sum_{L} \sum_{\nu} \begin{pmatrix} ll'L \\ m\ -m\ 0 \end{pmatrix} \begin{pmatrix} ll'L \\ \nu\ -\nu\ 0 \end{pmatrix} \times \left\{ R^{-L-1} \int_0^R dR' (R')^{L+2} S_{nl\nu,n'l'\nu}(R') + R^{L} \int_R^{\infty} dR' (R')^{-L+1} S_{nl\nu,n'l'\nu}(R') \right\}$$

$$C^{ll'm}_{nn'}(R) = \sum_{L=|l-l'|}^{l+l'} \begin{pmatrix} ll'L \\ m\ -m\ 0 \end{pmatrix} \left[\int_0^1 dt\, t^{L+2} \sigma^{ll'L}_{nn'}(t) + \int_1^{\infty} dt\, t^{-L+1} \sigma^{ll'L}_{nn'}(t) \right] \qquad (1.32)$$

with

$$\sigma^{ll'L}_{nn'}(t) = 4\pi R^2 \sum_{\nu=-\min(l,l')}^{\nu=\min(l,l')} \begin{pmatrix} ll'L \\ \nu\ -\nu\ 0 \end{pmatrix} S_{nl\nu,n'l'\nu}(Rt). \qquad (1.33)$$

In the summation, L is furthermore limited to values for which

$$L + l + l' = \text{even integer},$$

since the summation over ν cancels out the terms for which $(L + l + l')$ is odd.

Thus the Coulomb integral is reduced to a one-dimensional quadrature over several overlap integrals, and it is now possible to insert for the latter the expressions developed in the preceding paper.

[9] A. E. Edmonds, *Angular Momentum in Quantum Mechanics* (Princeton University Press, Princeton, New Jersey, 1957), p. 54, Eq. (4.1.4).
[10] Reference 9, page 7.
[11] H. Eyring, J. Walter, and G. E. Kimball, *Quantum Chemistry* (John Wiley & Sons, Inc., New York, 1947), p. 370, Eq. (V.6).
[12] M. Rotenberg, R. Bivis, N. Metropolis, and J. K. Wooten, Jr., *The 3-j and 6-j Symbols* (Technology Press, Cambridge, Massachusetts, 1959), p. 8, Eq. (1.42).
[13] A. E. Edmonds, *Angular Momentum in Quantum Mechanics* (Princeton University Press, Princeton, New Jersey, 1957), p. 46, Eq. (3.7.3).
[14] See Ref. 5, Sec. 3.2.

2. REDUCTION TO AUXILIARY FUNCTIONS

2.1. Expression for $\sigma_{nn'}^{ll'L}$

Using the charge distribution definitions of Eqs. (1.3, 3′), the elliptic coordinates

$$\xi = (r_A + r_C)/Rt, \qquad \eta = (r_A - r_C)/Rt, \tag{2.1}$$

and the parameter definitions

$$\bar{\zeta} = \tfrac{1}{2}(\zeta + \zeta'), \quad \rho = \bar{\zeta}R, \quad \tau = (\zeta - \zeta')/(\zeta + \zeta'), \tag{2.2}$$

one obtains for the combination of overlap integrals defined by Eq. (1.33), the expression

$$\begin{aligned}\sigma_{nn'}^{ll'L}(t) = {} & [(2l+1)(2l'+1)/2^{n+n'}(n+1)!\,(n'+1)!] \\ & \times \bar{\zeta}\rho^2(1+\tau)^{n+2}(1-\tau)^{n'+2}(\rho t)^{n+n'+1} \\ & \times \int_1^\infty d\xi \int_{-1}^1 d\eta(\xi+\eta)^n(\xi-\eta)^{n'}\Theta_{l,l'}^L(\theta_A, \theta_C) \\ & \times \exp(-2\rho t\xi - 2\tau\rho t\eta)\end{aligned} \tag{2.3}$$

with

$$\begin{aligned}\Theta_{l,l'}^L(\theta_A, \theta_C) = {} & 2(2l+1)^{-\frac{1}{2}}(2l'+1)^{-\frac{1}{2}} \\ & \times \sum_\nu \begin{pmatrix} ll'L \\ \nu\ -\nu 0 \end{pmatrix} \Theta_l^\nu(\cos\theta_A)\Theta_{l'}^\nu(\cos\theta_C).\end{aligned} \tag{2.4}$$

This function $\Theta_{ll'}^L$ can be transformed by inserting for the products $\Theta_l^\nu(\cos\theta_A)\,\Theta_{l'}^\nu(\cos\theta_B)$ the expansion given in Eq. (3.9) of the preceding paper, whence

$$\begin{aligned}&(\tfrac{1}{2}\xi + \tfrac{1}{2}\eta)^l\Theta_{ll'}^L \\ &= \sum_{k=0}^{l}\sum_{\Lambda=|l'-k|}^{l'+k} c_{\Lambda k}(2\Lambda+1)\begin{pmatrix} kl'\Lambda \\ 00\ 0 \end{pmatrix}(\tfrac{1}{2}\xi - \tfrac{1}{2}\eta)^k P_\Lambda(\cos\theta_C),\end{aligned} \tag{2.5}$$

where

$$\begin{aligned}c_{\Lambda k} = {} & \frac{(-1)^k}{(l-k)!}\sum_\nu \begin{pmatrix} ll'L \\ \nu\ -\nu 0 \end{pmatrix}\begin{pmatrix} kl'\Lambda \\ \nu\ -\nu 0 \end{pmatrix} \\ & \times \left[\frac{(l+\nu)!\,(l-\nu)!}{(k+\nu)!\,(k-\nu)!}\right]^{\frac{1}{2}}.\end{aligned} \tag{2.6}$$

The lower limit on k is the result of interchanging the summation over k with that over ν. The coefficients $c_{\Lambda k}$ can be expressed in terms of the 6–j symbols,[15] by virtue of the addition theorem[16]

$$\begin{aligned}c_{\Lambda k} = {} & \left[(2l+1)\begin{pmatrix} 2l \\ 2k \end{pmatrix}\right]^{\frac{1}{2}} \\ & \times \sum_\nu (-1)^{l+k+\nu}\begin{pmatrix} kl(l-k) \\ \nu\ -\nu 0 \end{pmatrix}\begin{pmatrix} kl'\Lambda \\ \nu\ -\nu 0 \end{pmatrix}\begin{pmatrix} ll'L \\ \nu\ -\nu 0 \end{pmatrix} \\ = {} & (-1)^{l'}\left[(2l+1)\begin{pmatrix} 2l \\ 2k \end{pmatrix}\right]^{\frac{1}{2}}\begin{pmatrix} (l-k)\Lambda L \\ 0\ \ 00 \end{pmatrix}\begin{Bmatrix} (l-k)\Lambda L \\ l'\ \ lk \end{Bmatrix},\end{aligned} \tag{2.7}$$

and insertion of these identities in Eq. (2.3) yields

$$\begin{aligned}\sigma_{nn'}^{ll'L}(t) = {} & \frac{(2l+1)(2l'+1)(n-l)!\,(n'-l')!}{2^{n+n'+2}(n+1)!\,(n'+1)!} \\ & \times \bar{\zeta}(1+\tau)^{n+2}(1-\tau)^{n'+2}(2\rho)^2\sum_{k=0}^{l}\sum_\Lambda (2\rho t)^{l-k} \\ & \times T_{n-l,n'+k-\Lambda}^\Lambda(2\rho t, \tau)(-1)^{l'}(2\Lambda+1)\frac{\Lambda!\,(n'+k-\Lambda)!}{(n'-l')!} \\ & \times \left[(2l+1)\begin{pmatrix} 2l \\ 2k \end{pmatrix}\right]^{\frac{1}{2}}\begin{pmatrix} kl'\Lambda \\ 00\ 0 \end{pmatrix}\begin{pmatrix} (l-k)\Lambda L \\ 0\ \ 00 \end{pmatrix}\begin{Bmatrix} (l-k)\Lambda L \\ l'\ \ lk \end{Bmatrix},\end{aligned} \tag{2.8}$$

where the functions $T_{ik}^\Lambda(x, y)$ are those defined and discussed in Eqs. (2.5)ff. of the preceding paper. The summation over Λ is subject to the properties of the 3-j symbols, namely[17]

$$\begin{aligned}&|l'-k| \le \Lambda \le (l'+k), \\ &\qquad l'+k+\Lambda = \text{even}, \\ &|l-k-L| \le \Lambda \le l-k+L, \\ &\qquad l-k+L+\Lambda = \text{even}.\end{aligned} \tag{2.9}$$

2.2. Expression for Coulomb-Integral

According to Eq. (1.32) the Coulomb integral results by integrating over the function σ, and hence, it will contain terms of the type

$$\begin{aligned}F = {} & \frac{\Lambda!\,(n'+k-\Lambda)!}{(n'-l')!}(2\rho)^2\left\{\int_0^1 dt\, t^{L+2}\right. \\ & \left. + \int_1^\infty dt\, t^{-L+1}\right\}(2\rho t)^{l-k}T_{n-l,n'+k-\Lambda}^\Lambda(2\rho t, \tau).\end{aligned} \tag{2.10}$$

Inserting for the functions T the expression given in Eq. (2.14) of the preceding paper, one obtains

$$\begin{aligned}F = {} & \frac{\Lambda!\,(n'+k-\Lambda)!}{(n'-l')!} \\ & \times \sum_{\lambda=0}^{\Lambda}\sum_{r=0}^{n-l}\sum_{s=0}^{n'+k-\Lambda}[k_{rs}^{\Lambda\lambda}/(\Lambda+r+s)!] \\ & \times J_{n-l-r,n'+k-\Lambda-s}^\lambda(\tau)(2\rho)^{l-k+\Lambda+r+s+2}\left\{\int_0^1 dt\, t^{L+2}\right. \\ & \left. + \int_1^\infty dt\, t^{-L+1}\right\}t^{l-k+\Lambda+r+s}e^{-2\rho t}I_{\lambda+r,\Lambda+s}(2\tau\rho t), \\ F = {} & (2\rho)_L\sum_\lambda\sum_r\sum_s b_{\lambda rs}J_{n-l-r,n'+k-\Lambda-s}^\lambda(\tau) \\ & \times \{H_{\lambda+r,\Lambda+s}^{-L+l-k+\Lambda+r+s+1}(2\rho, \tau) \\ & + (2\rho)_{-L+l-k+\Lambda+r+s+1}G_{\lambda+r,\Lambda+s}^{L-\lambda+l-k+1}(2\rho, \tau)\},\end{aligned} \tag{2.11}$$

[15] See Ref. 13, p. 94, Eq. (6.2.3).
[16] Reference 13, p. 95, Eq. (6.2.8).
[17] Reference 12, p. 2, Eq. (1.6)ff.

where

$$H^{\gamma}_{\alpha\beta}(\rho, \tau) = (\rho^{\gamma+1}/\gamma!) \int_1^{\infty} dt\, t^{\gamma} e^{-\rho t} I_{\alpha\beta}(\tau\rho t), \tag{2.12}$$

$$G^{\gamma}_{\alpha\beta}(\rho, \tau) = \rho \int_0^{1} dt\, t^{\alpha+\beta+\gamma+1} e^{-\rho t} I_{\alpha\beta}(\tau\rho t), \tag{2.13}$$

$$b_{\lambda r s} = \frac{L!\,\Lambda!\,(n' + k - \Lambda)!}{(n' - l')!} \quad \frac{(l - k - L + \Lambda + r + s + 1)!}{(\Lambda + r + s)!}\, k^{\Lambda\lambda}_{rs}, \tag{2.14}$$

and [similar to Eq. (2.16) of the preceding paper]

$$(2\rho)_{\alpha} = (2\rho)^{\alpha}/\alpha!. \tag{2.15}$$

The functions $I_{\alpha\beta}(x)$ and $J^{\lambda}_{\mu\nu}(\tau)$ were defined and discussed in Eqs. (1.9) to (1.32) and in Eqs. (2.17) to (2.25) of the preceding paper. Using for $k^{\Lambda\lambda}_{rs}$ the definition (2.15) of the preceding paper, we can write the coefficients $b_{\lambda rs}$ of Eq. (2.14) in the form

$$b_{\lambda rs} = \left[(l + l')! \binom{n' + l - L}{n' - l'} \Big/ \binom{l + l'}{L} \binom{2\Lambda}{\Lambda} \binom{n' + l - L}{n' + k - \Lambda}\right] \times \left[\binom{2\Lambda}{\Lambda + \lambda} \binom{\Lambda}{s} \binom{\lambda}{r} \binom{l - k - L + \Lambda + r + s + 1}{r + s + 1} \Big/ \binom{\lambda + \Lambda + r + s + 1}{r + s + 1}\right]. \tag{2.16}$$

The Coulomb integral is obtained by inserting the functions $\sigma^{ll'L}_{nn'}$ of Eq. (2.8) as integrand into Eq. (1.32), and it therefore is a sum of terms of the type defined in Eq. (2.10). Using for the latter the result derived in Eq. (2.11), we obtain for the Coulomb integral the formula

$$C^{ll'm}_{nn'} = \tilde{\zeta}(1 + \tau)^{n+2}(1 - \tau)^{n'+2} a_{nl} a_{n'l'} \times \sum_{L} \sum_{k=0}^{l} \sum_{\Lambda} A'_{Lk\Lambda}(n'll'm)(2\rho)_L + \sum_{r=0}^{n-l} \sum_{s=0}^{n'+k-\Lambda} \sum_{\lambda=0}^{\Lambda} D^{Lk\Lambda}_{rs\lambda}(l) J^{\lambda}_{n-l-r, n'+k-\Lambda-s}(\tau) \times \{H^{l-k-L+\Lambda+r+s+1}_{\lambda+r, \Lambda+s}(2\rho, \tau) + (2\rho)_{l-k-L+\Lambda+r+s+1} G^{l-k+L-\lambda+1}_{\lambda+r, \Lambda+s}(2\rho, \tau)\}. \tag{2.17}$$

Here, the coefficients a, A', D are defined as follows:

$$a_{nl} = (2l + 1)(n - l)!/2^{n+1}(n + 1)! \tag{2.18}$$

$$A'_{Lk\Lambda}(n'll'm) = \left[(-1)^{l'}(2\Lambda + 1)(l + l')! \binom{n' + l - L}{n' - l'} \Big/ \binom{l + l'}{L} \binom{2\Lambda}{\Lambda} \binom{n' + l - L}{n' + k - \Lambda}\right] \times \left[(2l + 1)\binom{2l}{2k}\right]^{\frac{1}{2}} \binom{kl'\Lambda}{00\ 0} \binom{(l-k)\Lambda L}{0 \quad 00} \begin{Bmatrix} (l-k)\Lambda L \\ l' \quad lk \end{Bmatrix} \binom{ll'L}{m\ -m0}, \tag{2.19}$$

$$D^{Lk\Lambda}_{rs\lambda}(l) = \binom{2\Lambda}{\Lambda + \lambda} \binom{\Lambda}{s} \binom{\lambda}{r} \binom{l - k - L + \Lambda + r + s + 1}{r + s + 1} \Big/ \binom{\lambda + \Lambda + r + s + 1}{r + s + 1}, \tag{2.20}$$

where $\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix}$ and $\begin{Bmatrix} j_1 & j_2 & j_3 \\ k_1 & k_2 & k_3 \end{Bmatrix}$ are the $3j$ and the $6j$ coefficients respectively.

According to the properties of the 6-j symbols, one has[18]

$$\begin{Bmatrix} (l-k)\Lambda L \\ l' \quad lk \end{Bmatrix} = \begin{Bmatrix} L(l-k)\Lambda \\ k \quad l' \quad l \end{Bmatrix} = \begin{Bmatrix} L \quad l' \quad l \\ k(l-k)\Lambda \end{Bmatrix} = \begin{Bmatrix} Ll' \quad l \\ kl''\Lambda \end{Bmatrix} \quad \text{with} \quad l = k + l'', \tag{2.21}$$

and in this special case, the 6-j symbol assumes a particularly simple form,[19] namely

$$\begin{Bmatrix} (l-k)\Lambda L \\ l' \quad lk \end{Bmatrix} = \left[\frac{(2g + 1)!\,(2g - 2l')!\,(2g - 2L)!\,(2g' - 2k)!\,(2g'' - 2l + 2k)!}{(2g' + 1)!\,(2g'' + 1)!\,(2g - 2l)!\,(2g' - 2l)!\,(2g' - 2\Lambda)!\,(2g'' - 2\Lambda)!\,(2g'' - 2L)!}\right]^{\frac{1}{2}} \times \left[(2l + 1)\binom{2l}{2k}\right]^{-\frac{1}{2}}, \tag{2.22}$$

[18] See Ref. 13, pp. 94–95, Eqs. (6.2.4), (6.2.5).
[19] See Ref. 13, p. 97, Eq. (6.3.1).

where

$$2g = l + l' + L, \qquad (2.23)$$
$$2g' = k + l' + L, \quad 2g'' = l - k + \Lambda + L.$$

Using this expression as well as those given for the 3-j symbols in the preceding paper, we can write the coefficients $A'_{Lk\Lambda}$ of Eq. (2.19) in the form

$$A'_{Lk\Lambda}(n'll'm) = A_L(n'll'm)B_{Lk\Lambda}(n'll') \qquad (2.24)$$

with

$$A_L(n'll'm) = \left[\binom{n'+l-L}{n'-l'}\binom{2g}{2l}\Big/\binom{l+l'}{L}\binom{2g}{L}\right] \times [(l+m)!\,(l-m)!\,(l'+m)!\,(l'-m)!]^{\frac{1}{2}} \times \sum_\alpha (-1)^{\alpha+c}\binom{2a}{l'-m-\alpha}\binom{2b}{l-m-\alpha}\binom{2c}{\alpha}, \qquad (2.25)$$

and

$$B_{Lk\Lambda}(n'll') = (-1)^k[(2\Lambda+1)/(2g'+1)(2g''+1)] \times \frac{\binom{2l}{2k}\binom{g'}{\Lambda}\binom{\Lambda}{g'-k}\binom{g''}{\Lambda}\binom{\Lambda}{g''-L}}{\binom{2\Lambda}{\Lambda}\binom{n'+l-L}{n'+k-\Lambda}\binom{2g'}{2k}\binom{2g''}{2l-2k}}, \qquad (2.26)$$

where

$$\begin{aligned} 2a &= -l + l' + L = 2g - 2l, \\ 2b &= l - l' + L \;\; = 2g - 2l', \\ 2c &= l + l' - L \;\; = 2g - 2L, \end{aligned} \qquad (2.27)$$

and the summation over α runs from max (0, $l-L-m, l'-L-m$) to min ($l-m, l'-m, l+l'-L$).

With this factorization of the A' coefficients, the Coulomb integral (2.17) finally assumes the form

$$C^{ll'm}_{nn'} = \tilde{\xi}(1+\tau)^{n+2}(1-\tau)^{n'+2}a_{nl}a_{n'l'} \times \sum_L A_L(n'll'm)(2\rho)_L \sum_{k=0}^{l}\sum_\Lambda B_{Lk\Lambda}(n'll') \times \sum_{r=0}^{n-l}\sum_{s=0}^{n'+k-\Lambda}\sum_{\lambda=0}^{\Lambda} D^{Lk\Lambda}_{rs\lambda}(l)J^\lambda_{n-l-r,n'+k-\Lambda-s}(\tau) \times \{H^{l-k-L+\Lambda+r+s+1}_{\lambda+r,\Lambda+s}(2\rho,\tau) + (2\rho)_{l-k-L+\Lambda+r+s+1}G^{l-k+L-\lambda+1}_{\lambda+r,\Lambda+s}(2\rho,\tau)\}. \qquad (2.28)$$

The summation over L is regulated by the conditions

$$|l-l'| \le L \le (l+l'), \qquad l+l'+L = \text{even},$$

and the summation over Λ by the conditions

$$|l'-k| \le \Lambda \le (l'+k), \qquad l'+k+\Lambda = \text{even},$$
$$|l-k-L| \le \Lambda \le (l-k+L), \qquad l-k+L+\Lambda = \text{even}.$$

The coefficients a, A, B, D are defined in Eqs. (2.18), (2.25), (2.26), and (2.20), respectively. The functions $H^\gamma_{\alpha\beta}$, $G^\gamma_{\alpha\beta}$, $(2\rho)_\alpha$ were defined in Eqs. (2.12), (2.13), and (2.15), respectively and are discussed in the next section. The functions $J^\lambda_{\mu\nu}$ were discussed in Eqs. (2.17)ff. of the preceding paper.

Equation (2.28) shows that, for fixed values of n, n', l, l', the expressions of the Coulomb integrals for different values of m differ only slightly, so that considerable computing can be saved by evaluating them all simultaneously.

It should be noted that those upper indices (γ) of $H^\gamma_{\alpha\beta}$ and $G^\gamma_{\alpha\beta}$ which occur in Eq. (2.28) are always positive, so that the definitions (2.12), (2.13) do not generate exponential integrals. Essential for this result is the fact that the overlap integral $S^{ll'm}_{nn'}$ contains the factor $R^{|l-l'|}$, as was pointed out after Eq. (3.11) of the preceding paper.

3. ANALYSIS OF AUXILIARY FUNCTIONS

3.1. Properties of $I_{\alpha\beta}$

According to the definitions (2.13) and (2.12), the properties of the functions G and H are determined by those of the functions $I_{\alpha\beta}$. From Eqs. (1.19), (1.21), (1.22), and (1.23) of the preceding paper we have

$$I_{\alpha\beta}(x) = [I_{\alpha-1,\beta}(x) - I_{\alpha,\beta-1}(x)](\alpha+\beta+1)/2x, \qquad (3.1)$$

$$I_{\alpha\beta}(x) = I_{\alpha+1,\beta}(x) + [2x(\beta+1)/(\alpha+\beta+2)(\alpha+\beta+3)]I_{\alpha+1,\beta+1}(x), \qquad (3.2)$$

$$I_{\alpha\beta}(x) = [(\alpha+1)I_{\alpha+1,\beta}(x) + (\beta+1)I_{\alpha,\beta+1}(x)]/(\alpha+\beta+2), \qquad (3.3)$$

$$I_{\alpha\beta}(x) = [1 - 2x/(\alpha+\beta+2)]I_{\alpha,\beta+1}(x) + [2x(\beta+2)/(\alpha+\beta+2)(\alpha+\beta+3)]I_{\alpha,\beta+2}(x). \qquad (3.4)$$

Next the following identities involving the derivatives of $I_{\alpha\beta}$ are readily deduced from the integral definitions given in Eq. (1.9) of the preceding paper:

$$(\alpha+\beta+2)(d/dx)e^xI_{\alpha\beta}(x) = 2(\beta+1)e^xI_{\alpha,\beta+1}(x), \qquad (3.5)$$

$$(\alpha+\beta+2)(d/dx)e^{-x}I_{\alpha\beta}(x) = -2(\alpha+1)e^{-x}I_{\alpha+1,\beta}(x), \qquad (3.6)$$

$$(d/dx)x^{\beta+1}e^xI_{\alpha\beta}(x) = (\beta+1)x^\beta e^xI_{\alpha-1,\beta+1}(x), \qquad (3.7)$$

$$(d/dx)x^{\alpha+1}e^{-x}I_{\alpha\beta}(x) = (\alpha+1)x^{\alpha}e^{-x}I_{\alpha+1,\beta-1}(x), \quad (3.8)$$

$$(d/dx)x^{\alpha+\beta+1}e^{x}I_{\alpha\beta}(x) = (\alpha+\beta+1)x^{\alpha+\beta}e^{x}I_{\alpha-1,\beta}(x), \quad (3.9)$$

$$(d/dx)x^{\alpha+\beta+1}e^{-x}I_{\alpha\beta}(x) = (\alpha+\beta+1)x^{\alpha+\beta}e^{-x}I_{\alpha,\beta-1}(x), \quad (3.10)$$

and furthermore

$$-(\alpha+\beta+2)(d/dt)e^{-\rho t}I_{\alpha\beta}(\tau\rho t) = \rho e^{-\rho t}[(1+\tau)(\alpha+1)I_{\alpha+1,\beta}(\tau\rho t) + (1-\tau)(\beta+1)I_{\alpha,\beta+1}(\tau\rho t)], \quad (3.11)$$

which in conjunction with Eqs. (1.20) and (1.21) of the preceding paper yields also

$$2\tau(d/dt)[t^{\alpha+\beta+1}e^{-\rho t}I_{\alpha\beta}(\tau\rho t)] = (\alpha+\beta+1)t^{\alpha+\beta}e^{-\rho t} \times [(1+\tau)I_{\alpha,\beta-1}(\tau\rho t) - (1-\tau)I_{\alpha-1,\beta}(\tau\rho t)]. \quad (3.12)$$

Finally, we recall the series representation

$$I_{\alpha\alpha}(x) = \sum_{k=0}^{\infty} b_{\alpha k}x^{2k}, \quad (3.13)$$

$$b_{\alpha k} = \binom{\alpha+k}{\alpha}\frac{(2\alpha+1)!}{(2\alpha+2k+1)!} \quad (3.14)$$

given in Eq. (1.15) of the preceding paper.

3.2. Properties of $G^{\gamma}_{\alpha\beta}$

If the relations (3.1), (3.2), (3.3), (3.4) are inserted in the definition (2.13) of the function $G^{\gamma}_{\alpha\beta}(\rho, \tau)$, the following recurrence formulas result for the latter

$$G^{\gamma}_{\alpha\beta} = G^{\gamma}_{\alpha+1,\beta-1} + [2\tau\rho/(\alpha+\beta+2)]G^{\gamma}_{\alpha+1,\beta}, \quad (3.15)$$

$$G^{\gamma+1}_{\alpha\beta} = G^{\gamma}_{\alpha+1,\beta} + [2\tau\rho(\beta+1)/(\alpha+\beta+2)(\alpha+\beta+3)]G^{\gamma}_{\alpha+1,\beta+1}, \quad (3.16)$$

$$G^{\gamma+1}_{\alpha\beta} = [(\alpha+1)G^{\gamma}_{\alpha+1,\beta} + (\beta+1)G^{\gamma}_{\alpha,\beta+1}]/(\alpha+\beta+2), \quad (3.17)$$

$$G^{\gamma+1}_{\alpha\beta} = G^{\gamma}_{\alpha,\beta+1} - [2\tau\rho/(\alpha+\beta+2)]G^{\gamma+1}_{\alpha,\beta+1} + [2\tau\rho(\beta+2)/(\alpha+\beta+2)(\alpha+\beta+3)]G^{\gamma}_{\alpha,\beta+2}. \quad (3.18)$$

In these and the following equations, $G^{\gamma}_{\alpha\beta}$ stands for $G^{\gamma}_{\alpha\beta}(\rho, \tau)$ as defined by Eq. (2.13).

Integrating by parts in Eq. (2.13), one obtains

$$(\alpha+\beta+\gamma+2)G^{\gamma}_{\alpha\beta} = \rho e^{-\rho}I_{\alpha\beta}(\tau\rho) - \rho\int_0^1 dt\, t^{\alpha+\beta+\gamma+2}\frac{d}{dt}e^{-\rho\tau}I_{\alpha\beta}(\tau\rho t), \quad (3.19)$$

whence, by virtue of Eq. (3.11),

$$(\alpha+\beta+\gamma+2)G^{\gamma}_{\alpha\beta} = \rho e^{-\rho}I_{\alpha\beta}(\tau\rho) + \rho[(1+\tau)(\alpha+1)G^{\gamma}_{\alpha+1,\beta} + (1-\tau)(\beta+1)G^{\gamma}_{\alpha,\beta+1}]/(\alpha+\beta+2). \quad (3.20)$$

Another integration by parts can be performed by using Eq. (3.5) in integrated form. Thereby

$$G^{\gamma}_{\alpha\beta} = \{\rho t^{\alpha+\beta+\gamma+1}e^{-(1+\tau)\rho t} \times [e^{\tau\rho t}I_{\alpha,\beta-1}(\tau\rho t)(\alpha+\beta+1)/\beta 2\tau\rho]\}_0^1 - \rho\int_0^1 dt[(d/dt)t^{\alpha+\beta+\gamma+1}e^{-(1+\tau)\rho t}] \times [(\alpha+\beta+1)/2\tau\rho\beta][e^{\tau\rho t}I_{\alpha,\beta-1}(\tau\rho t)]. \quad (3.21)$$

whence

$$[2\tau\rho\beta/(\alpha+\beta+1)]G^{\gamma}_{\alpha\beta} = \rho e^{-\rho}I_{\alpha,\beta-1} + \rho(1+\tau)G^{\gamma+1}_{\alpha,\beta-1} - (\alpha+\beta+\gamma+1)G^{\gamma}_{\alpha,\beta-1}. \quad (3.22)$$

Replacing β by $(\beta+2)$ in this equation and substituting the resulting expression for $G^{\gamma}_{\alpha,\beta+2}$ in the previous Eq. (3.18), one finally obtains

$$(\alpha+\beta+2)G^{\gamma}_{\alpha\beta} + \gamma G^{\gamma-1}_{\alpha,\beta+1} = \rho e^{-\rho}I_{\alpha,\beta+1} + \rho(1-\tau)G^{\gamma}_{\alpha,\beta+1}, \quad (3.23)$$

which, for $\gamma = 0$ becomes

$$(\alpha+\beta+2)G^{0}_{\alpha\beta} = \rho e^{-\rho}I_{\alpha,\beta+1} + \rho(1-\tau)G^{0}_{\alpha,\beta+1}. \quad (3.24)$$

The series (3.13) finally gives rise to the series representations

$$G^{\gamma}_{\alpha\alpha} = \sum_{k=0}^{\infty} b_{\alpha k}(\tau\rho)^{2k}g_{2\alpha+2k+\gamma+1}(\rho), \quad (3.25)$$

where

$$g_n(\rho) = \rho\int_0^1 dt\, t^n e^{-\rho t} \quad (3.26)$$

is essentially an incomplete gamma function.[20] Since $g_n(\rho)$ clearly decreases with increasing values of n, the series for $G^{\gamma}_{\alpha\alpha}$ converges faster than the Bessel function series (3.13). The $g_n(\rho)$ satisfy the recurrence relation

$$\rho g_n(\rho) = n g_{n-1} - \rho e^{-\rho}, \quad (3.27)$$

which is satisfactory in the upward direction for $n < \rho$, and always in the downward direction. The starting functions are given by

$$g_0(\rho) = 1 - e^{-\rho} \quad (3.28)$$

[20] Bateman Manuscript Project, *Higher Transcendental Functions*, edited by A. Erdelyi, W. Magnus, F. Oberhettinger, and F. G. Tricomi (McGraw-Hill Book Company, Inc., New York, 1953), Vol. 2, p. 133.

and

$$g_n(\rho) = e^{-\rho} \sum_{k=1}^{\infty} \rho^k/(n+1)(n+2)\cdots(n+k), \quad (3.29)$$

respectively.

3.3. Recurrence Scheme for Evaluating the Functions $G^{\gamma}_{\alpha\beta}$

For fixed values of ρ and τ, a complete table of $G^{\gamma}_{\alpha\beta}$ is required for the index values limited by

$$\begin{aligned} &0 \leq \alpha \leq n + l', \\ &0 \leq \beta \leq n' + l, \\ &0 \leq \gamma \leq 2l + l' + 1. \end{aligned} \quad (3.30)$$

These limits follow from the summation limits in Eq. (2.28). Moreover one recognizes that the index combinations $(\alpha + \gamma)$ and $(\beta + \gamma)$ are limited by

$$\alpha + \gamma \leq l + l' + n + 1, \quad (3.31)$$

$$\beta + \gamma \leq 2l + l' + n' + 1. \quad (3.32)$$

In the following discussion, the subtractions in the recurrence scheme are examined as regards their effect upon the numerical accuracy. It may be noted that

$$G^{\gamma}_{\alpha\beta} > 0 \quad (3.33)$$

for all parameter values. Furthermore, τ is assumed to be positive. The case of negative τ is disposed of in Sec. 3.6.

The evaluation of the G table starts with the evaluation of the last α row, viz. $G^{\gamma}_{A\beta}$, where γ and $A = (l + l' + n + 1 - \gamma)$ are fixed and $0 \leq \beta \leq n' + 2l + l' + 1 - \gamma$. The rest of the table $G^{\gamma}_{\alpha\beta}$ (for fixed γ) is then constructed by backwards recurrence on α and β, using alternately Eqs. (3.15) and (3.20). The initial row $G^{\gamma}_{A\beta}$ can be found by one of the following two paths.

First path for $G^{\gamma}_{A\beta}$ (A, γ fixed): One starts by using the series expansion (3.25) for the calculation of G^0_{NN} ($N = n + n' + l + l' + 1$), and then constructs the remainder of the row $G^0_{N\beta}$ ($0 \leq \beta \leq N$) by means of Eq. (3.24). Subsequently one constructs the rows $G^1_{N-1,\beta}$ ($0 \leq \beta \leq N - 1$), $G^2_{N-2,\beta}$ ($0 \leq \beta \leq N - 2$), $\cdots$, $G^{\gamma}_{N-\gamma,\beta}$ ($0 \leq \beta \leq N - \gamma$), for $\gamma = 1, 2, \cdots, (2l + l' + 1)$, by means of the relation (3.16).

Second path for $G^{\gamma}_{A\beta}$ (A, γ fixed): The series (3.25) is used to find G^{γ}_{MM} for $M = \max(n' + l, n + l')$ and the γ values $0 \leq \gamma \leq l + l' + 1$. The choice of M is obvious from Eq. (3.30). The upper limit of γ is seen to be sufficient if one solves the inequalities (3.31) and (3.32) for γ and substitutes the value of M for α and β, respectively; this yields four limits, all $\leq l + l' + 1$. It may be noted that

$$M \geq \gamma \quad (3.34)$$

follows for all γ. Next the row $G^0_{M\beta}$ ($0 \leq \beta \leq M$) is obtained by means of Eq. (3.24) and then the rows $G^{\gamma}_{M\beta}$ ($0 \leq \beta \leq M$) are constructed for $\gamma = 1, 2, \cdots, (l + l' + 1)$ using either (3.22), solved for $G^{\gamma+1}_{\alpha,\beta-1}$ or (3.23), solved for $G^{\gamma}_{\alpha\beta}$ so as to avoid loss of accuracy. If β is such that

$$\rho e^{-\rho} I_{M,\beta-1} \leq \tfrac{3}{4}(\beta + \gamma + M + 1) G^{\gamma}_{M,\beta-1}, \quad (3.35)$$

then the cancellation in Eq. (3.22) will not lose accuracy. On the other hand, if β is such that

$$\rho e^{-\rho} I_{M,\beta-1} \geq \tfrac{3}{4}(\beta + \gamma + M + 1) G^{\gamma}_{M,\beta-1}, \quad (3.36)$$

then by virtue of Eq. (3.34)

$$\rho e^{-\rho} I_{M,\beta-1} \geq \tfrac{3}{2}(\gamma + 1) G^{\gamma}_{M,\beta-1} \quad (3.37)$$

and hence Eq. (3.23) will not lose accuracy.[21]

3.4. Properties of $H^{\gamma}_{\alpha\beta}$

Substitution of Eqs. (3.1) to (3.3) into the definition (2.12) of $H^{\gamma}_{\alpha\beta}$, yields the recurrence formulas

$$H^{\alpha-1}_{\alpha-1,\beta} = H^{\gamma-1}_{\alpha,\beta-1} + 2\tau[\gamma/(\alpha + \beta + 1)] H^{\gamma}_{\alpha\beta}, \quad (3.38)$$

$$\begin{aligned} H^{\gamma}_{\alpha,\beta} = {} & H^{\gamma}_{\alpha+1,\beta} \\ & + 2\tau[(\gamma+1)(\beta+1)/(\alpha+\beta+2)(\alpha+\beta+3)] H^{\gamma+1}_{\alpha+1,\beta+1}, \end{aligned} \quad (3.39)$$

$$H^{\gamma}_{\alpha\beta} = [(\alpha+1)H^{\gamma}_{\alpha+1,\beta} + (\beta+1)H^{\gamma}_{\alpha,\beta+1}]/(\alpha+\beta+2). \quad (3.40)$$

Performing an integration by parts in Eq. (2.12), similar to that described by Eq. (3.21), and using again Eq. (3.5), we find

$$\begin{aligned} (1 + \tau) H^{\gamma}_{\alpha,\beta-1} = {} & H^{\gamma-1}_{\alpha,\beta-1} + 2\tau[\beta/(\alpha + \beta + 1)] H^{\gamma}_{\alpha\beta} \\ & + (\rho^{\gamma}/\gamma!) e^{-\rho} I_{\alpha,\beta-1}(\tau\rho), \end{aligned} \quad (3.41)$$

if $\gamma \neq 0$, and

$$\begin{aligned} (1 + \tau) H^{0}_{\alpha,\beta-1} = {} & 2\tau[\beta/(\alpha + \beta + 1)] H^{0}_{\alpha\beta} \\ & + e^{-\rho} I_{\alpha,\beta-1}(\tau\rho) \end{aligned} \quad (3.42)$$

for $\gamma = 0$. An analogous integration by parts, using however (3.6) in integrated form, i.e.,

$$\begin{aligned} &\int dt\, e^{-\tau\rho t} I_{\alpha\beta}(\tau\rho t) \\ &\quad = -[(\alpha + \beta + 1)/2\tau\rho\alpha] e^{-\tau\rho t} I_{\alpha-1,\beta}(\tau\rho t), \end{aligned} \quad (3.43)$$

[21] Here it is assumed that $G^{\gamma}_{M,\beta}$ is a smooth function of β.

yields the relations

$$(1-\tau)H^{\gamma}_{\alpha-1,\beta} = H^{\gamma-1}_{\alpha-1,\beta} - 2\tau[\alpha/(\alpha+\beta+1)]H^{\gamma}_{\alpha\beta} + (\rho^{\gamma}/\gamma!)e^{-\rho}I_{\alpha-1,\beta}(\tau\rho), \quad (3.44)$$

and

$$(1-\tau)H^{0}_{\alpha-1,\beta} = -2\tau[\alpha/(\alpha+\beta+1)]H^{0}_{\alpha\beta} + e^{-\rho}I_{\alpha,\beta-1}(\tau\rho). \quad (3.45)$$

Subtracting Eq. (3.38) from Eq. (3.41) one finds

$$(1+\tau)H^{\gamma}_{\alpha,\beta-1} = H^{\gamma-1}_{\alpha-1,\beta} + 2\tau[(\beta-\gamma)/(\alpha+\beta+1)]H^{\gamma}_{\alpha\beta} + (\rho^{\gamma}/\gamma!)e^{-\rho}I_{\alpha,\beta-1}(\tau\rho), \quad (3.46)$$

and, adding Eq. (3.38) to Eq. (3.44), one obtains

$$(1-\tau)H^{\gamma}_{\alpha-1,\beta} = H^{\gamma-1}_{\alpha,\beta-1} + 2\tau[(\gamma-\alpha)/(\alpha+\beta+1)]H^{\gamma}_{\alpha\beta} + (\rho^{\gamma}/\gamma!)e^{-\rho}I_{\alpha-1,\beta}(\tau\rho). \quad (3.47)$$

Finally one can carry out the following integration by parts in Eq. (2.12):

$$H^{\gamma}_{\alpha\beta} = \frac{\rho^{\gamma+1}}{\gamma!}\{t[t^{\gamma}e^{-\rho\tau}I_{\alpha\beta}(\tau\rho t)]\}_1^{\infty} - \frac{\rho^{\gamma+1}}{\gamma!}\int_1^{\infty} dt\, t\frac{d}{dt} \times \{t^{\gamma-\alpha-\beta-1}[t^{\alpha+\beta+1}e^{-\rho\tau}I_{\alpha\beta}(\tau\rho t)]\}.$$

Using the identity (3.12), this yields

$$(1-\tau)H^{\gamma}_{\alpha-1,\beta} = [2\tau/(\alpha+\beta+1)](\rho^{\gamma+1}/\gamma!)e^{-\rho}I_{\alpha\beta}(\tau\rho) + (1+\tau)H^{\gamma}_{\alpha,\beta-1} + 2\tau[-1+(\gamma+1)/(\alpha+\beta+1)]H^{\gamma}_{\alpha\beta}. \quad (3.48)$$

The series (3.13) gives rise to the expansion[22]

$$H^{\gamma}_{\alpha\alpha} = \sum_{k=0}^{\infty} b_{\alpha k}\frac{(\gamma+2k)!}{\gamma!}\tau^{2k}h_{\gamma+2k}(\rho), \quad (3.49)$$

where

$$h_n(\rho) = \frac{\rho^{n+1}}{n!}\int_1^{\infty} dt\, t^n e^{-\rho t} \quad (3.50)$$

is given by the recurrence scheme

$$h_0 = e^{-\rho}, \quad (3.51)$$

$$h_n = h_{n-1} + (\rho^n/n!)e^{-\rho}. \quad (3.52)$$

3.5. Recurrence Scheme for Evaluating the Functions $H^{\gamma}_{\alpha\beta}$

Required is, for fixed values of ρ and τ, a complete table $H^{\gamma}_{\alpha\beta}$ for the index values limited by

$$0 \le \alpha \le n + l', \quad 0 \le \beta \le n' + l, \quad 0 \le \gamma \le n + n' + 1 - |l - l'|, \quad (3.53)$$

as is seen from the summation limits in Eq. (2.28). Moreover one recognizes that

$$-(l+l') \le \gamma - (\alpha+\beta+1) \le l - |l-l'|. \quad (3.54)$$

In the following discussion, the subtractions in the recurrence scheme are examined as regards their effect upon the numerical accuracy. It may be noted that

$$H^{\gamma}_{\alpha\beta} > 0 \quad (3.54')$$

for all parameter values. Furthermore, τ is assumed to be positive. The case of negative τ is disposed of in Sec. 3.6.

As in the case of the G function, the table is constructed in two stages: First the rows with maximal $\alpha = M$ are constructed for all γ. Specifically

$$H^{\gamma}_{M\beta} \text{ for } M = \max(n+l', n'+l), \quad 0 \le \beta \le M$$

and

$$0 \le \gamma \le \max(n+n'+1-|l-l'|, M) = \gamma_{\max}. \quad (3.55)$$

Then the remainder of the table, for $\alpha = M-1, M-2, \cdots, 0$, is filled in by backward recurrence.

The first stage, i.e., construction of the rows $H^{\gamma}_{M\beta}$, starts by calculation of all diagonal elements H^{γ}_{MM} using the series (3.49) for all γ.[22] Then the row $H^{0}_{M\beta}$ $(0 \le \beta \le M)$ is constructed using Eq. (3.42), and the rows $H^{\gamma}_{M\beta}$ $(0 \le \beta \le M)$ for $\gamma = 1, 2, \cdots \gamma_{\max}$ are obtained using Eq. (3.41) raising the value of γ step by step.

The second stage, i.e., lowering the index α for all γ and β values, is executed by finding first all $H^{\gamma}_{M-1,\beta}$ $(0 \le \beta \le M,\ 0 \le \gamma \le \gamma_{\max})$, then $H^{\gamma}_{M-2,\beta}$ $(0 \le \beta \le M,\ 0 \le \gamma \le \gamma_{\max})$, and so forth for $\alpha = M-3$, $\alpha = M-4$, etc. The functions for $(\alpha-1)$ are obtained from those for α by using alternately the relations (3.38) and (3.40). Since the relation (3.38) simultaneously lowers the index γ, one has to use the relation (3.47) in conjunction with (3.40) in order to obtain the function $H^{\gamma}_{\alpha-1,\beta}$ for $\gamma = \gamma_{\max}$. According to the choice of $\gamma_{\max}$, as given by Eq. (3.55), the term $(\gamma_{\max}-\alpha)$ is guaranteed to be positive.

Actually, it is not necessary to construct this whole table, the condition given in Eq. (3.54) selects only a part of it. In particular if $l = l' = 0$, one needs only the elements for which $\gamma = \alpha + \beta + 1$.

[22] A more efficient way to calculate $H^{\gamma}_{\alpha\alpha}$ will be given elsewhere.

3.6. Recurrence Scheme for Negative Values of τ

The discussion given in Secs. 3.3 and 3.5 for the subtraction loss of accuracy applies to the case that τ is positive. It is always possible to identify the two atoms involved with A and B in such a way that this is the case. However, as in the case of the overlap integrals, it is also possible to choose A and B such that one has $\tau < 0$ in Eq. (2.28). In this case the calculation of the auxiliary functions G and H can be reduced to the calculation of auxiliary functions with positive argument values τ by virtue of the relations

$$G^{\gamma}_{\alpha\beta}(\rho, \tau) = G^{\gamma}_{\beta\alpha}(\rho, -\tau), \tag{3.56}$$

$$H^{\gamma}_{\alpha\beta}(\rho, \tau) = H^{\gamma}_{\beta\alpha}(\rho, -\tau), \tag{3.57}$$

which follow from Eq. (1.12) of the preceding paper.

Hence, the discussion in Secs. (3.3) and (3.5) still applies, except for the fact that the limits given for the indices α and β have to be altered correspondingly.

4. COULOMB INTEGRAL FOR EXTREME VALUES OF THE INTERNUCLEAR DISTANCE

4.1. Zero Internuclear Distance

If $A = B$, the Coulomb integral of Eq. (1.11) becomes a one-center integral and can be treated by standard one-center methods based on the Laplace expansion for r_{12}^{-1}.[11] This leads to the result[23]

$$C^{ll'm}_{nn'} = \delta_{ll'}\bar{\zeta}[F^{l}_{nn'}(x, y) + F^{l}_{n'n}(y, x)], \tag{4.1}$$

where

$$F^{l}_{\alpha\beta}(x, y) = x^{\alpha-l+2}y\left[\binom{\alpha+l+1}{l}\Big/\binom{\beta+1}{l+1}2^{\alpha+\beta+2l}(l+1)\right] \times \sum_{j=0}^{\beta-l}\binom{\alpha+l+1+j}{j}y^{j}, \tag{4.2}$$

$$x = \zeta/(\zeta+\zeta') = \tfrac{1}{2}(1+\tau),$$
$$y = \zeta'/(\zeta+\zeta') = \tfrac{1}{2}(1-\tau), \tag{4.2'}$$
$$\bar{\zeta} = \tfrac{1}{2}(\zeta+\zeta').$$

An alternate expression for this integral will be given elsewhere.[24]

4.2. Large Internuclear Distances

If we replace, in Eq. (1.32), the first quadrature by

$$\int_0^R dR' = \int_0^\infty dR' - \int_R^\infty dR', \tag{4.3}$$

then the expression (1.32) for the Coulomb integral can be written

$$C^{ll'm}_{nn'}(R) = M^{ll'm}_{nn'} + SR^{ll'm}_{nn'}. \tag{4.4}$$

The first term,

$$M^{ll'm}_{nn'} = \sum_{L=|l-l'|}^{l+l'}\begin{pmatrix} ll'L \\ m\ -m'0\end{pmatrix}\int_0^\infty dt\, t^{L+2}\sigma^{ll'L}_{nn'}(t), \tag{4.5}$$

has "long-range" character, in as much as it is proportional to a negative power of R, and is called the "multipole term." The second term,

$$SR^{ll'm}_{nn'} = \sum_{L=|l-l'|}^{l+l'}\begin{pmatrix} ll'L \\ m\ -m'0\end{pmatrix} \times \int_1^\infty dt(t^{-L+1} - t^{L+2})\sigma^{ll'L}_{nn'}(t), \tag{4.6}$$

is the "short-range term," in as much as it decreases exponentially for large values of R. Thus, there exists a distance R_M beyond which, to a specified accuracy, the Coulomb integral is given by the multipole term alone. This is of practical interest since the latter can be cast in a very simple form.

From considering the derivations in Eqs. (1.14) to (1.32) it is obvious that the multipole term can be written

$$M^{ll'm}_{nn'} = \sum_{L=0}^{\infty} R^{-L-1}\int dV'(R')^{L}P_L(\cos\Theta)S(\mathbf{R}'). \tag{4.7}$$

Now the following identity is justified below:

$$(R')^{L}P_L(\cos\Theta) = \sum_{l_1}\sum_{l_2}\sum_{\nu} A^{\nu}_{l_1l_2}r^{l_1}_{A1}Y_{l_1\nu}(\theta_{A1}, \varphi_{A1})r^{l_2}_{B2}Y^{*}_{l_2\nu}(\theta_{B2}, \varphi_{B2}), \tag{4.8}$$

where

$$A^{\nu}_{l_1l_2} = 4\pi(l_1+l_2)!/[(2l_1+1)(2l_2+1)(l_1+\nu)! \times (l_1-\nu)!\,(l_2+\nu)!\,(l_2-\nu)!]^{\frac{1}{2}} \tag{4.9}$$

and the summations are determined by

$$0 \le l_1 \le L, \quad 0 \le l_2 \le L, \quad l_1 + l_2 = L, \qquad -\min(l_1, l_2) \le \nu \le \min(l_1, l_2). \tag{4.10}$$

By virtue of this identity, the expression (4.7) for the multipole term can be transformed back to integrations over (dV_1) and (dV_2), respectively. The result is

$$M^{ll'm}_{nn'} = \sum_{L} R^{-L-1}\sum_{l_1}\sum_{l_2}\sum_{\nu} A^{\nu}_{l_1l_2} \times \int dV_1 f(r_{A1})^{*}r^{l_1}_{A1}Y_{l_1\nu}(\theta_{A1}, \varphi_{A1}) \times \int dV_2 g(r_{B2})r^{l_2}_{B2}Y^{*}_{l_2\nu}(\theta_{B2}, \varphi_{B2}). \tag{4.11}$$

[23] Similar formulas were given by S. Hagstrom, Ph.D. thesis, Iowa State University, Ames, Iowa, 1957, and by R. K. Nesbet, Rev. Mod. Phys. 35, 552 (1963).

[24] K. Miller and K. Ruedenberg, to be published.

Inserting for f and g the basic charge distributions of Eqs. (1.3), (1.3'), one finds that, of the fourfold series, only one term survives, namely,

$$M_{nn'}^{ll'm} = A_{ll'}^{m}\langle r_{A1}^{l}\rangle\langle r_{B2}^{l'}\rangle, \tag{4.12}$$

where

$$\langle r_{A1}^{l}\rangle = [(n+l+1)!/(n+1)!\,2^{n+l+1}] \times [(2l+1)/4\pi]^{\frac{1}{2}}\zeta^{-l} \tag{4.13}$$

and similar for $\langle r_{B1}^{l}\rangle$. Hence the multipole term is given by

$$M_{nn'}^{ll'm}(R) = M(nlm)M(n'l'm)(l+l')!/\zeta^{l}\zeta'^{l'}R^{l+l'+1},$$

where

$$M(nlm) = \left[\binom{n+l+1}{l}\Big/2^{n+l+1}\right]\left[\binom{2l}{l-m}\Big/\binom{2l}{l}\right]^{\frac{1}{2}}. \tag{4.14}$$

4.3. Derivation of Eq. (4.8)

According to Carlson and Rushbrooke,[25] the inverse distance between the two points 1 and 2 can be expressed in terms of the coordinates $(r_{A1}, \theta_{A1}, \varphi_{A1})$ and the coordinates $(r_{B2}, \theta_{B2}, \varphi_{B1})$ by the expansion

$$r_{12}^{-1} = R^{-1}\sum_{l_1=0}^{\infty}\sum_{l_2=0}^{\infty}\sum_{m=-l_<}^{l_<} A_{l_1l_2}^{m}\left(\frac{r_{A1}}{R}\right)^{l_1}\left(\frac{r_{B1}}{R}\right)^{l_2} \times Y_{l_1m}(\theta_{A1}, \varphi_{A1})Y_{l_2m}^{*}(\theta_{B2}, \varphi_{B2}) \tag{4.15}$$

provided that $r_{A1} + r_{B2} < R$. Here $A_{l_1l_2}^{m}$ is defined by Eq. (4.9) and $l_< = \min(l_1, l_2)$. Now it follows from the definition (1.15) of R' that in this case also

$$R' = [(\mathbf{r}_{A1} - \mathbf{r}_{B2})^2]^{\frac{1}{2}} \leq r_{A1} + r_{B2} < R, \tag{4.16}$$

so that Eq. (1.30) yields

$$r_{12}^{-1} = R^{-1}\sum_{L=0}^{\infty}\left(\frac{R'}{R}\right)^{L}P_L(\cos\Theta). \tag{4.17}$$

Comparison of the two series (4.15) and (4.17) yields directly the identity (4.8) provided that $r_{A1} + r_{B2} < R$. From Eq. (1.15) one finds, however,

$$(\mathbf{R}'\cdot\mathbf{R}) = RR'\cos\Theta = Rr_{A1}\cos\theta_{A1} + Rr_{B2}\cos\theta_{B2},$$

so that the left-hand side of the identity (4.8) can be expressed algebraically in terms of

$$R'\cos\Theta = r_{A1}\cos\theta_{A1} + r_{B2}\cos\theta_{B2} \tag{4.18}$$

and

$$(R')^2 = r_{A1}^2 + r_{B2}^2 - 2(\mathbf{r}_{A1}\cdot\mathbf{r}_{B2}). \tag{4.19}$$

Hence, Eq. (4.8) represents a polynomial identity in the variables x_{A1}, y_{A1}, z_{A1}, x_{B2}, y_{B2}, z_{B2}, and is therefore not restricted by any limitations on the variables. It can therefore be used in the integral $\int dV'$ pertaining over all space.

5. COULOMB INTEGRAL BETWEEN PRODUCTS OF ATOMIC ORBITALS

The Coulomb integrals commonly occurring in molecular quantum mechanics are obtained by substituting the orbital products

$$f = (An_1l_1m_1)(An_2l_2m_2), \tag{5.1}$$

$$g = (Bn_3l_3m_3)(Bn_4l_4m_4) \tag{5.1'}$$

for the charge distribution f and g in Eq. (1.1). Here

$$(Anlm) = (2\zeta)^{n+\frac{1}{2}}[(2n)!]^{-\frac{1}{2}}r_A^{n-1}e^{-\zeta r}\mathcal{Y}_{lm}(\theta_A, \varphi_A) \tag{5.2}$$

are normalized Slater-type orbitals, with $\mathcal{Y}_{lm}$ being the real spherical harmonics defined in Eqs. (1.8) to (1.9'').

Since the products of any two spherical harmonics can be expanded in terms of spherical harmonics,[26] the orbital product of Eq. (5.1) can be expanded in terms of the basic charge distributions [ANLM] of Eq. (1.3), (1.3') on the same center A, whence

$$(An_1l_1m_1)(An_2l_2m_2) = p_{n_1l_1}(\zeta_1/\zeta_a)p_{n_2l_2}(\zeta_2/\zeta_a) \times \binom{n_1+n_2}{n_1}\sum_{L,M} q_{LM}(l_1m_1l_2m_2)[ANLM] \tag{5.5}$$

where

$$p_{nl}(x) = 2^n n!\,[(2l+1)/(2n)!]^{\frac{1}{2}}x^{n+\frac{1}{2}}, \tag{5.5'}$$

$$\zeta_a = \tfrac{1}{2}(\zeta_1 + \zeta_2), \tag{5.6}$$

$$N = (n_1 + n_2 - 1), \tag{5.7}$$

and the summation over L is limited by

$$|l_1 - l_2| \leq L \leq (l_1 + l_2), \qquad L + l_1 + l_2 = \text{even}. \tag{5.8}$$

The summation over M is restricted to the two values M_+ and M_- given by

$$M_\pm = \text{sign}(m_1)\,\text{sign}(m_2)\,|(|m_1| \pm |m_2|)|, \tag{5.9}$$

where

$$\text{sign}(x) = x/|x| \quad \text{and} \quad \text{sign}(0) = +1. \tag{5.10}$$

The corresponding coefficients are

$$q_{LM+} = \epsilon_+(-1)^{m_1+m_2}[1 + \delta_{0,m_1m_2}]^{\frac{1}{2}}\,2^{-\frac{1}{2}} \times \binom{l_1l_2L}{000}\binom{l_1 \quad l_2 \quad L}{-|m_1| \; -|m_2| \; (|m_1|+|m_2|)}, \tag{5.11}$$

[25] B. C. Carlson and G. S. Rushbrooke, Proc. Cambridge Phil. Soc. **46**, 626 (1950). See also R. J. Buehler and Joseph O. Hirschfelder, Phys. Rev. **83**, 628 (1951).

[26] See Ref. 13, p. 63, Eq. (4.6.5).

$$q_{LM-} = \epsilon_-(-1)^{\max(|m_1|,|m_2|)}[1 + \delta_{0,m_1-m_2}]^{\frac{1}{2}}\, 2^{-\frac{1}{2}} \times \begin{pmatrix} l_1 l_2 L \\ 000 \end{pmatrix}\begin{pmatrix} l_1 & l_2 & L \\ -|m_1| & |m_2| & (|m_1|-|m_2|) \end{pmatrix}, \quad (5.11')$$

where ϵ_+ and ϵ_- depend upon the sign of (m_1m_2) and $(m_1 + m_2)$ according to the Table I, and the product of $3j$ coefficients is given by[27]

$$\begin{pmatrix} j_1 j_2 j_3 \\ 000 \end{pmatrix}\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 m_2 m_3 \end{pmatrix} = (-1)^{g+j_3+m_3} \times \left[\binom{2j_1}{j_1}\binom{2j_2}{j_2}\binom{2j_3}{j_3} \Big/ \binom{2j_1}{j_1-m_1}\binom{2j_2}{j_2-m_2}\binom{2j_3}{j_3-m_3}\right]^{\frac{1}{2}} \times \left[\binom{g}{j_1}\binom{j_1}{g-j_2} \Big/ \binom{2g}{j_1}\binom{2g-j_1}{j_2}(2g+1)\right] \times \sum_{\nu} (-1)^{\nu}\binom{2g-2j_1}{j_2+m_2-\nu}\binom{2g-2j_2}{j_1-m_1-\nu}\binom{2g-2j_3}{\nu}, \quad (5.12)$$

where

$$2g = j_1 + j_2 + j_3 \quad (5.13)$$

and the summation over ν goes from max $(0, j_2 - j_3 - m_1, j_1 - j_3 + m_2)$ to min $(2g - 2j_3, j_1 - m_1, j_2 + m_2)$.

[27] S. R. Polo, *Studies on Crystal Field Theory* (RCA Laboratories, Princeton, New Jersey, 1961), Vol. I, Eqs. (8.14) and Eq. (8.22).

TABLE I.

(m_1m_2)	$(m_1 + m_2)$	ϵ_+	ϵ_-
+	+	1	1
+	−	−1	1
−	+	1	−1
−	−	1	1
−	0	1	0
0	+, 0, −	1	0

Combination of the expansion (5.5) for orbital products with Eq. (1.11) for basic Coulomb integrals yields

$$[(An_1l_1m_1)(An_2l_2m_2) \mid (Bn_3l_3m_3)(Bn_4l_4m_4)] = p_{n_1l_1}(\zeta_1/\zeta_a)p_{n_2l_2}(\zeta_2/\zeta_a)p_{n_3l_3}(\zeta_3/\zeta_b)p_{n_4l_4}(\zeta_4/\zeta_b) \times \binom{n_1+n_2}{n_1}\binom{n_3+n_4}{n_3} \sum_{L,L'}\sum_{M,M'} q_{LM}(l_1m_1l_2m_2) \times q_{L'M'}(l_3m_3l_4m_4)\, \delta_{MM'} C_{NN'}^{LL'M}(R), \quad (5.14)$$

which reduces Coulomb integrals between products of real atomic orbitals to Coulomb integrals between basic charge distributions.

Solution of the Bloch–Nordsieck Model*

JAN TARSKI†
Courant Institute of Mathematical Sciences New York University, New York, New York
(Received 26 May 1965)

An operator solution is constructed for the Bloch–Nordsieck model, in the formulation of Bogolubov and Shirkov. The following are some of the features of the solution. (a) The method of solution depends on the following assumption: The renormalized electron current is identical with the free-field current. The resulting solution is consistent with this assumption. (b) The infraparticle structure of the electron is analogous to the infraparticle structure in Schroer's model. (c) The ordering of fields, i.e., the definition of their products, is considered in detail. A modification of Wick ordering (called F-ordering in the text) appears particularly satisfactory. (d) The solution allows a heuristic discussion of the infrared representations of the electromagnetic field. These depend, apparently, on the chosen velocity of the electron.

1. INTRODUCTION

THE model of Bloch and Nordsieck[1] gave us the first clue for understanding the peculiarities of the infrared in quantum electrodynamics. For this reason, and for others, the interest in the model remained. In particular, Bogolubov and Shirkov (BS)[2] presented an elegant formulation, and computed the electron Green's function. Their formulation also forms the basis of the present work.

In this paper we construct an explicit operator solution for the model, i.e., we express the Heisenberg electron field ψ and the photon field A in terms of the corresponding free fields. In this way we clarify the infraparticle structure of ψ. We find that this structure arises from regularization of integrals[3] at $k = 0$. Therefore, we have a close analogy with Schroer's model.[4,5]

The Bloch–Nordsieck model does not allow electron recoil or spin flip. Consequently, it is expected to be trivial, in the sense of having the S-matrix equal to unity. The main interest in the model lies therefore in the details of the solution, of which we have already noted one: the infraparticle structure.

Let us now describe the model. The essential feature, as is well known, is the assumption of a definite electron velocity $\mathbf{v}$. One then replaces the α_i-matrices by the respective components v_i, and similarly for the matrix β. For a four-dimensional formulation, we introduce the 4-vector

$$= (1, \mathbf{v})(1 - \mathbf{v}^2)^{-\frac{1}{2}}, \qquad u^2 = 1. \tag{1.1}$$

We assume the following field equations, which are implicit in BS (in, e.g., the Feynman gauge; $\Box = -\partial_t^2 + \nabla^2$):

$$(iu^\nu \partial_\nu - m)\psi = (eu^\nu A_\nu \psi)_{\text{ren}}, \tag{1.2a}$$

$$\Box A^\nu = -(eu^\nu \psi^* \psi)_{\text{ren}}. \tag{1.2b}$$

The subscript "ren" indicates renormalized interaction terms. We shall also consider the subsidiary condition on physical states Ψ:

$$\langle \partial_\nu A^\nu(x) \rangle_\Psi = 0. \tag{1.3}$$

One normally requires, in addition, renormalized canonical commutation (or anticommutation) relations. They will be examined in detail. However, we shall be referring to fields which satisfy Eqs. (1.2) as solutions of the model, even if they are not local. We construct, in fact, two solutions for ψ, which differ with respect to the ordering of fields in products. We shall see that the relation

$$[\psi(t, \mathbf{x}), u^0\, \psi^*(t, \mathbf{y})]_+ = Z_2^{-1}\, \delta(\mathbf{x} - \mathbf{y}) \tag{1.4}$$

is not fulfilled by one of our solutions, but the other solution can fulfill the usual relations for equal times.

The solution of the field equations hinges on the following assumption, where $\psi^{(0,\pm)}$ denote the creation (+) and the annihilation parts of the free-electron field:

$$(eu^\nu \psi^* \psi)_{\text{ren}} = eu^\nu \psi^{(0,+)} \psi^{(0,-)}. \tag{1.5}$$

This equation means, e.g., that there are no closed fermion loops in the theory, and this latter statement is in agreement with the conventional pertur-

* This work represents results obtained in part at the Courant Institute under the Ford Foundation grant for mathematical physics.
† Present address: Department of Physics, Columbia University, New York, New York.
[1] F. Bloch and A. Nordsieck, Phys. Rev. 52, 54 (1937).
[2] N. N. Bogolubov and D. V. Shirkov, *Introduction to the theory of quantized fields* (Interscience Publishers, Inc., New York, 1959), Sec. 41. This book is referred to as BS in the text. See also A. V. Svidzinskii, Zh. Eksperim. i Theor. Phys. 31, 324 (1956) [English transl.: Soviet Phys—JETP 4, 179].
[3] A. S. Wightman, lectures in Summer School in Cargèse, Corsica, 1964.
[4] B. Schroer, Forsch. Phys. 11, No. 1 (1963), Sec. III.
[5] J. Tarski, J. Math. Phys. 5, 1713 (1964).

bation-theoretic solution. In fact, each of the fields $\psi^{(0,\pm)}$ has only one sign of frequency, and consequently our solution does not allow antifermions.

We shall see that Eq. (1.5) is consistent with our solution for ψ, ψ^*, and with our renormalization prescriptions. Furthermore, these fields yield a Green's function, which is in agreement with the function obtained in BS by functional integration.

The solution of field equations is now straightforward, at least in principle. We find A from Maxwell's equations and Eq. (1.5), and when A is known, Dirac's equation can be integrated.

In the first part of our paper, in Secs. 2 and 3, the presentation is based on the Feynman gauge. This gauge is the most convenient, and allows a direct comparison with the results of BS. However, it seems that the subsidiary condition can be fulfilled only in a way which is rather unsatisfactory. One can avoid this difficulty by employing the Coulomb gauge, which we discuss in Sec. 4.

The operator solution allows us to examine the infrared representations of the electromagnetic field (i.e., of the corresponding canonical commutation relations). Unfortunately, we are able to give only a heuristic discussion. It appears that the infrared representations are characterized by translation, as in some other models,[1,5,6] and that these representations depend on the chosen velocity $\mathbf{v}$.

Another shortcoming of our work lies in the treatment of renormalization. We describe in Sec. 2 three separate prescriptions, two of which depend on particular ordering in products of fields, and the third, on the use of counter terms and singular multipliers. These prescriptions, however, are introduced in an *ad hoc* way. A search for a more systematic approach to renormalization should be attempted.

It may also be worthwhile to point out, that our analysis of the model is mathematically precise, even though the treatment of certain details remains to be improved or completed. In particular, the quantities which we introduce have essentially unambiguous meaning. (We largely ignore the questions of domain of field operators.) While our renormalization prescriptions appear somewhat arbitrary, each of them is mathematically precise, or can be readily made such. For contrast, the computation of the electron Green's function in BS proceeds through an infinite renormalization, and is therefore essentially formal.

[6] K. O. Friedrichs, *Mathematical aspects of the quantum theory of fields* (Interscience Publishers, Inc., New York, 1953), Sec. III and §19.

We now present an outline of the paper. In Sec. 2 we first describe the renormalization prescriptions. Then we make the restriction $u = (1, \mathbf{0})$ and construct the fields A, ψ, ψ^*. In Sec. 3 these fields are examined in detail. In Sec. 4 we construct the solution and discuss briefly its properties, for the case of a general vector u. Both the Feynman gauge and the Coulomb gauge are considered. Section 5 is devoted to the infrared representations. Finally, Sec. 6 contains a short discussion of our results.

2. SOLUTION CORRESPONDING TO A STATIC ELECTRON

A. Renormalization and Ordering

We have already stated that we shall express ψ and A in terms of the free fields $\psi^{(0,\pm)}$ and $A^{(0)}$. The former satisfy the equations

$$(iu^\nu \partial_\nu \pm m)\psi^{(0,\pm)} = 0,$$

and the canonical commutation relations

$$[\psi^{(0,\pm)}, \psi^{(0,\pm)}]_+ = 0,$$

$$[\psi^{(0,-)}(t, \mathbf{x}), u^0\psi^{(0,+)}(t, \mathbf{y})]_+ = \delta(\mathbf{x} - \mathbf{y}).$$

Therefore, one can try to renormalize the theory by Wick- or W-ordering of the fields and of the interaction terms.

However, we shall see that a modified ordering perscription appears to be preferable. This ordering resembles the E-ordering introduced by Friedrichs,[7] and we shall call it *F-ordering*. For this ordering we arrange the factors as follows. We put any $\psi^{(0,-)}$ standing alone to the right, any $\psi^{(0,+)}$ standing alone to the left, and leave any functional of the 4-current $u\psi^{(0,+)}\psi^{(0,-)}$ unordered. (The current commutes with itself.) For $A^{(0)}$, we use the usual Wick ordering. Examples will make this procedure clearer.

We shall derive the relevant formulas for each ordering perscription. The subscript "ord" will be used to indicate that either ordering can be used in a given product, provided this choice is consistently followed. Thus

$$(AB)_{\text{ord}} = (AB)_F \text{ or } = (AB)_W \equiv :AB:. \qquad (2.1)$$

These subscripts will be used sometimes for the electron fields as well. The two solutions ψ_F and ψ_W of course differ in some respects.

We introduce the notation

$$\psi^{[-1]} \equiv \psi, \qquad \psi^{[+1]} \equiv \psi^*, \qquad (2.2)$$

since these fields are proportional (in our solution) to $\psi^{(0,-)}$ and $\psi^{(0,+)}$, respectively. However, each of

[7] K. O. Friedrichs, Ref. 6, §24.

the fields $\psi^{[\pm]}$ may have contributions from all values of frequency. We shall refer to the two fields $\psi^{[\pm]}$ jointly as the *electron field*, and to $\psi^{(0,\pm)}$ as the *free-electron field* $\psi^{(0)}$.

Now, if we renormalize by ordering, we can write Eqs. (1.2) in the form

$$(iu^\nu \partial_\nu \pm m)\psi^{[\pm]} = \mp eu^\nu (A_\nu \psi^{[\pm]})_{\text{ord}}, \tag{2.3a}$$

$$\Box A^\nu = -eu^\nu (\psi^{[+]}\psi^{[-]})_{\text{ord}}, \tag{2.3b}$$

and our assumption about the current becomes

$$u^\nu (\psi^{[+]}\psi^{[-]})_{\text{ord}} = u^\nu \psi^{(0,+)}\psi^{(0,-)} \equiv j^{(0)\nu}. \tag{2.4}$$

Finally, let us note that an alternative approach to renormalization is to construct the interaction terms as, e.g., the limits[8]

$$\lim_{\epsilon\to 0} e\,\frac{u^\nu A_\nu(x+\epsilon)\psi^{[\pm]}(x) - C_1^{\pm}(\epsilon)\psi^{[\pm]}(x)}{1+\Gamma_1(\epsilon)}, \tag{2.5a}$$

$$\lim_{\epsilon\to 0} e\,\frac{u^\nu \psi^{[+]}(x+\epsilon)\psi^{[-]}(x) - \langle \psi^{[+]}u^\nu \psi^{[-]}\rangle_0 - C_2(\epsilon)A^\nu(x)}{1+\Gamma_2(\epsilon)}, \tag{2.5b}$$

where the functions C_i, Γ_i, and $\langle \psi^{[+]}\psi^{[-]}\rangle_0$ may be singular at $\epsilon = 0$. When we compare with our solutions, we shall see that the F-ordering is consistent with these expressions, if one sets

$$\Gamma_1 = C_2 = \langle \psi^{[+]}\psi^{[-]}\rangle_0 = 0. \tag{2.6}$$

On the other hand, W-ordering is inconsistent with the expressions (2.5). See Eq. (3.37) and the subsequent discussion.

In view of their consistency with F-ordering, the expressions (2.5) will not be considered as often as the ordering prescriptions.

B. The Free Fields

For the free electromagnetic field $A^{(0)}$ we shall use in Secs. 2 and 3 the Feynman gauge, so that the commutation rules are

$$[A_\mu^{(0)}(x), A_\nu^{(0)}(y)] = -i^{-1}g_{\mu\nu}D(x-y). \tag{2.7}$$

This gauge corresponds to the longitudinal coefficient $d_l = 1$. Naturally, we require

$$[A_\nu^{(0)}, \psi^{(0,\pm)}] = 0. \tag{2.8}$$

In the remainder of Sec. 2, and in Sec. 3, we shall specialize to the case

$$u = (1, 0). \tag{2.9}$$

The equations for the free electron field,

$$(i\,\partial_t \pm m)\psi^{(0,\pm)} = 0, \tag{2.10}$$

allow us to set

$$\psi^{(0,\pm)}(t, \mathbf{x}) = \chi^{(\pm)}(\mathbf{x})\, e^{\pm imt}. \tag{2.11}$$

The canonical anticommutation relations are

$$[\chi^{(\pm)}, \chi^{(\pm)}]_+ = 0, \qquad [\chi^{(-)}(\mathbf{x}), \chi^{(+)}(\mathbf{y})]_+ = \delta(\mathbf{x}-\mathbf{y}), \tag{2.12}$$

and if

$$\chi^{(\pm)}(\mathbf{x}) = \frac{1}{(2\pi)^{\frac{3}{2}}}\int d^3\mathbf{k}\, e^{\mp i\mathbf{k}\mathbf{x}}\tilde{\chi}^{(\pm)}(\mathbf{k}), \tag{2.13}$$

then

$$[\tilde{\chi}^{(-)}(\mathbf{k}), \tilde{\chi}^{(+)}(\mathbf{l})]_+ = \delta(\mathbf{k}-\mathbf{l}). \tag{2.14}$$

The free-electron current satisfies

$$j^{(0)} \equiv (\rho, \mathbf{0}), \quad \rho = \rho(\mathbf{x}) = \chi^{(+)}(\mathbf{x})\chi^{(-)}(\mathbf{x}), \tag{2.15}$$

$$[\rho(\mathbf{x}), \rho(\mathbf{y})] = 0, \tag{2.16}$$

$$[\rho(\mathbf{x}), \chi^{(\pm)}(\mathbf{y})] = \pm\delta(\mathbf{x}-\mathbf{y})\chi^{(\pm)}(\mathbf{y}). \tag{2.17}$$

We further set

$$\rho(\mathbf{x}) = \frac{1}{(2\pi)^{\frac{3}{2}}}\int d^3\mathbf{k}\, e^{\mp i\mathbf{k}\mathbf{x}}\tilde{\rho}^{\pm}(\mathbf{k}); \tag{2.18}$$

then

$$\tilde{\rho}^{+}(\mathbf{k}) = \tilde{\rho}^{-}(-\mathbf{k}) = \frac{1}{(2\pi)^{\frac{3}{2}}}\int d^3\mathbf{l}\,\tilde{\chi}^{(+)}(\mathbf{k}+\mathbf{l})\tilde{\chi}^{(-)}(\mathbf{l}). \tag{2.19}$$

C. The Interacting Fields

We consider first the electromagnetic field. The spatial components do not interact, so

$$A_s = A_s^{(0)}, \qquad s = 1, 2, 3. \tag{2.20}$$

For A_0, the solution is the sum of the free field and of the Coulomb potential,

$$A_0(t, \mathbf{x}) = A_0^{(0)}(t, \mathbf{x}) + e\Phi(\mathbf{x}), \tag{2.21}$$

$$\Phi(\mathbf{x}) = \frac{1}{4\pi}\int d^3\mathbf{y}\,\frac{\rho(\mathbf{y})}{|\mathbf{x}-\mathbf{y}|}. \tag{2.22}$$

As is well known, one can also derive the Coulomb interaction in the framework of covariant quantum electrodynamics.[9]

For the study of representations of canonical commutation relations, it will be useful to express the solution (2.21) in terms of translated canonical operators (or distributions). This is as in the usual

[8] W. Zimmermann (private communication).

[9] J. M. Jauch and F. Rohrlich, *The Theory of Photons and Electrons* (Addison-Wesley Publishing Company, Inc., Cambridge, Massachusetts, 1955), Sec. 6-1 and references given there.

treatment of fixed source scattering. We make the decomposition $A_0 = A_0^{(+)} + A_0^{(-)}$, and

$$A_0^{(\pm)}(t, \mathbf{x}) = \frac{1}{(2\pi)^{\frac{3}{2}}} \int \frac{d^3\mathbf{k}}{(2\omega)^{\frac{1}{2}}} a_0^{(\pm)}(t, \mathbf{k})\, e^{\mp i\mathbf{k}\mathbf{x}}, \quad (2.23a)$$

$$a_0^{(\pm)}(t, \mathbf{k}) = a_0^{(0,\pm)}(\mathbf{k})\, e^{\pm i\omega t} + e[2^{-\frac{1}{2}}(4\pi)^{-1}]\tilde{\rho}^{\pm}(\mathbf{k})\omega^{-\frac{3}{2}}, \quad (2.23b)$$

where $\omega = |\mathbf{k}|$. One has, in fact,

$$A_0^{(\pm)} = A_0^{(0,\pm)} + \tfrac{1}{2}e\Phi. \quad (2.24)$$

The distributions $a_0^{(0,\pm)}$ are associated with the plane-wave decomposition of $A_0^{(0)}$, and satisfy the *reversed commutation relations*,[10]

$$[a_0^{(0,-)}(\mathbf{k}), a_0^{(0,+)}(\mathbf{l})] = -\delta(\mathbf{k} - \mathbf{l}). \quad (2.25)$$

Finally, we come to the electron field $\psi^{[\pm]}$. A formal solution is

$$\psi^{[\pm]}(t, \mathbf{x}) = \left(\exp\left[\pm ie \int^t dt'\, A_0(t', \mathbf{x})\right]\chi^{(\pm)}(\mathbf{x})\right)_{\text{ord}} e^{\pm imt}.$$

We have to specify the indefinite integral. For the Coulomb potential part of A_0, we have no choice but to select an arbitrary lower limit T. This yields the expressions

$$(e^{\pm ie^2(t-T)\Phi(\mathbf{x})})_{\text{ord}} \equiv (e^{\pm i\beta\Phi(\mathbf{x})})_{\text{ord}} = e^{\pm i\beta\Phi(\mathbf{x})} \quad \text{or} \quad :e^{\pm i\beta\Phi(\mathbf{x})}:. \quad (2.26)$$

We return to these expressions in Sec. 3. In particular, we shall see that in case of F-ordering, the exponential becomes a unitary operator, and vacuum expectation values can be made independent of T.

However, for the free-field part $A_0^{(0)}$, we cannot simply take a lower limit T. Such a limit would strongly affect the vacuum expectation values, and one could hardly consider the electron field as an approximation, in some sense, to a local field. We may note that the resulting field would still be a solution of the field equations, and in fact would be analogous to the bilocal solution of Schroer's model,[5] but we shall not discuss this possibility further.

We can avoid these difficulties by using the plane wave decomposition of $A_0^{(0)}$, and by setting

$$\int^t dt'\, A_0^{(0)}(t', \mathbf{x}) = \frac{1}{(2\pi)^{\frac{3}{2}}} \int \frac{d^3\mathbf{k}}{(2\omega)^{\frac{1}{2}}} [a_0^{(0,+)}(\mathbf{k})e^{i\omega t - i\mathbf{k}\mathbf{x}}(i\omega)^{-1} + a_0^{(0,-)}(\mathbf{k})e^{-i\omega t + i\mathbf{k}\mathbf{x}}(-i\omega)^{-1}] \equiv B(t, \mathbf{x}). \quad (2.27)$$

[10] Throughout the paper, this phrase refers to Eq. (2.25) and its analogs, and never to the commutator–anticommutator interchange.

This definition is equivalent to setting, e.g.,

$$B(t, \mathbf{x}) = \int_{t+i\infty}^{t} d\tau\, A_0^{(0,+)}(\tau, \mathbf{x}) + \int_{t-i\infty}^{t} d\tau\, A_0^{(0,-)}(\tau, \mathbf{x}). \quad (2.28)$$

The commutation relations (2.25) give us, formally,

$$\langle B(x)B(y)\rangle_0 \sim \frac{-1}{(2\pi)^3} \int \frac{d^3\mathbf{k}}{2\omega^3} e^{-ik(x-y)}.$$

We see a divergence at $\mathbf{k} = 0$.

As in Schroer's model[3,5] one can restrict the test functions $\tilde{f}(k)$ for B by imposing the condition $\tilde{f}(0) = 0$, or else, regularize the integrand. We shall see in the next section that here, as in Schroer's model, regularization must be used if one wishes to exponentiate the field.

In the sequel we shall use the notation

$$:e^{ieB(x)}: \equiv \Upsilon(x) = \Upsilon^{(+)}(x)\Upsilon^{(-)}(x), \quad (2.29a)$$

$$:e^{-ieB(x)}: = \Upsilon^*(x). \quad (2.29b)$$

The electron field can now be expressed as

$$\psi^{[\pm]} = (e^{\pm ieB}e^{\pm i\beta\Phi}\chi^{(\pm)})_{\text{ord}}\, e^{\pm imt}, \quad (2.30)$$

or

$$\psi^{[+]}(t, \mathbf{x}) = \Upsilon(x)\chi^{(+)}(\mathbf{x})(e^{i\beta\Phi(\mathbf{x})})_{\text{ord}}\, e^{imt}, \quad (2.31a)$$

$$\psi^{[-]}(t, \mathbf{x}) = \Upsilon^*(x)(e^{-i\beta\Phi(\mathbf{x})})_{\text{ord}}\chi^{(-)}(\mathbf{x})\, e^{-imt}. \quad (2.31b)$$

One sees that our assumption about the current is fulfilled, for either ordering.

3. PROPERTIES OF THE FOREGOING SOLUTION

A. The Field B

We have already noted, that the field $B(x)$ can be defined by regularizing an integral with the help of distributions. We now elaborate on this, as follows. First, we define one regularization, and we compute the resulting two-point function $\langle BB\rangle_0$. Then we consider more general regularizations. Finally, we show that one cannot avoid distributions by restricting the test functions at $k = 0$.

We now consider the two-point function,[11]

$$\langle B(x)B(y)\rangle_0 \equiv -i^{-1}\mathfrak{D}^{(-)}(\xi) \quad (\text{where } \xi = x - y) \quad (3.1a)$$

$$= \frac{-1}{2(2\pi)^3} \int d^4k[\delta(\omega - |\mathbf{k}|)\omega^{-3}]_{\text{reg}}\, e^{-ik\xi}. \quad (3.1b)$$

The subscript "reg" indicates that the function in the

[11] The corresponding commutator function $\mathfrak{D}$ was used by J. Schwinger, Phys. Rev. 74, 1439 (1948), Sec. III.

bracket has to be regularized. The main point is to replace[3,5]

$$\theta(\omega)\omega^{-1} \to (\omega^{-1})_+ \equiv (d/d\omega)[\theta(\omega)\log\omega] = \theta(\omega)\omega^{-1} \quad \text{for} \quad \omega \neq 0. \tag{3.2}$$

We will adopt the following definition:

$$i^{-1}\mathfrak{D}^{(-)}(\xi) = \frac{1}{2(2\pi)^3}\int d\Omega_k \int_{-\epsilon}^{\infty} d\omega(\omega^{-1})_+ e^{-i\omega(\xi^0 - |\xi|\cos\theta)}. \tag{3.3}$$

We select $\epsilon > 0$, in order to avoid having $\omega = 0$ an end point of integration.

The function $\mathfrak{D}^{(-)}$ is now easily computed. One uses the relation[5]

$$\int_{-\infty}^{\infty} d\omega\, e^{-i\omega\tau}(\omega^{-1})_+ = -\log i\tau + \Gamma'(1) \tag{3.4}$$

and finds

$$i^{-1}\mathfrak{D}^{(-)}(\xi) = (4\pi^2)^{-1}[\Gamma'(1) + 1] - \frac{1}{8\pi^2}\log[-(\xi^0)^2 + \xi^2 + i\epsilon\xi^0] - \frac{1}{8\pi^2}\frac{\xi^0}{|\xi|}\log\frac{\xi^0 + |\xi|}{\xi^0 - |\xi| - i\epsilon}. \tag{3.5a}$$

We shall need the following special case:

$$H(t) \equiv -i^{-1}\mathfrak{D}^{(-)}(t, \mathbf{0}) = -(4\pi^2)^{-1}\Gamma'(1) + (8\pi^2)^{-1}\log(-t^2 + i\epsilon t). \tag{3.5b}$$

Let us now look at more general regularizations. An essential requirement for any two-point function $\langle BB\rangle_0 \equiv -i^{-1}\mathfrak{B}^{(-)}$ is that

$$-\partial_0^2\mathfrak{B}^{(-)}(\xi) = D^{(-)}(\xi). \tag{3.6}$$

The distribution $\mathfrak{D}^{(-)}$ satisfies this equation, and hence the most general solution is[12]

$$\mathfrak{B}^{(-)}(\xi) = \mathfrak{D}^{(-)}(\xi) + a\xi^0 + ib. \tag{3.7}$$

Equation (3.7) would be satisfied if a and b were distributions in ξ. However, they can be only polynomials in ξ^2 if one requires $\mathfrak{B}^{(-)}$ to be rotation-invariant, and if one imposes the usual spectral condition. (This condition most likely remains meaningful for indefinite metric.[13]) If one also requires (as in Ref. 11) that

$$\Box\mathfrak{B}^{(-)} = 0,$$

then a and b must be constants. For these reasons, and for simplicity, we confine ourselves to the case where a and b are real constants.

The constant b can be absorbed into $\mathfrak{D}^{(-)}$ if one alters the scale for the first logarithm in Eq. (3.5a), but at the moment there is no reason to choose $a = 0$. Later in Sec. 3 we shall see how a relates to the renormalization of the electron mass, and to the commutators $[A_0, \psi^{[\pm]}]$.

Of course, various other singular functions can be expressed in terms of $\mathfrak{D}^{(-)}$ or $\mathfrak{B}^{(-)}$. In particular, we later need the function

$$\langle A_0^{(0)}B\rangle_0 = -i^{-1}\,\partial_0\mathfrak{D}^{(-)}(\xi) - i^{-1}a. \tag{3.8}$$

Note that the function $\langle A_0^{(0)}B\rangle_0$ can be evaluated directly from the definition (2.27), without regularization, and this leads to $a = 0$. However, the possible presence of constant terms in such functions has been observed elsewhere.[14]

We shall also make use of the commutator functions,[11]

$$(-i)[B(x), B(y)] = \mathfrak{B}^{(-)}(\xi) - \mathfrak{B}^{(-)}(-\xi) \equiv \mathfrak{B}(\xi) \equiv \mathfrak{D}(\xi) + 2a\xi^0, \tag{3.9a}$$

$$\mathfrak{B}(0, \xi) = \partial_t^2\,\mathfrak{B}(0, \xi) = 0, \tag{3.9b}$$

$$\partial_t\,\mathfrak{B}(0, \xi) = (4\pi\,|\xi|)^{-1} + 2a. \tag{3.9c}$$

Finally, let us demonstrate the need for regularization. The following argument is in analogy with the analysis of Schroer's model; cf. Ref. 5, especially Sec. 3. We introduce a restricted field, B_r, which is the restriction of B to test functions $f \in \mathcal{S}$ satisfying

$$\int d^4x\, f(x) = 0 = \tilde{f}(0). \tag{3.10}$$

The field B_r leaves the Fock space of $A_0^{(0)}$ invariant, and allows one to replace $(\omega^{-1})_+$ by $\theta(\omega)\omega^{-1}$. We want to know if one can define restrictions of $:B^n:$, to be denoted by $:B_r^n:$, which would also have these two equivalent properties. We shall see that $:B_r^2:$ is defined only for test functions $\tilde{f}(k)$ which vanish for $k^2 = 0$, and that $:B_r^3:$ cannot be defined at all, if nonzero vacuum expectation values are desired. In this sense one can define the exponential only of B, not of B_r, and the distribution $(\omega^{-1})_+$ cannot be avoided.

The proof of the corresponding facts in Schroer's model depended on the decomposition $\varphi_P = \varphi_r + \varphi_l$, which does not apply immediately to the present case. Therefore we give a direct proof. If we pre-

[12] L. Schwartz, *Théorie des distributions* (Hermann & Cie, Paris, 1957–1959), 2nd ed., Vol. I, p. 53.

[13] I. S. Iohvidov and M. G. Krein, Trudy Moscov. Mat. Obsc. 5, 367 (1956) [translation: Am. Math. Soc. Transl. Ser. 2, 13, 105].

[14] O. Steinmann, J. Math. Phys. 4, 583 (1963), Sec. 5.

suppose the restriction to the Fock space of $A_0^{(0)}$, we can write

$$\frac{1}{(2\pi)^2}\int d^4\xi\, e^{ik\xi}\langle B_r(x)B_r(y)\rangle_0 = \frac{-1}{2\pi}(k^0)^{-2}\,\delta(k^2)\theta(k^0), \tag{3.11a}$$

$$\frac{1}{(2\pi)^2}\int d^4\xi\, e^{ik\xi}\langle :B_r^2(x): :B_r^2(y):\rangle_0$$
$$= \frac{2}{(2\pi)^4}\int d^4p(p^0)^{-2}\,\delta(p^2)\theta(p^0)$$
$$\times (k^0-p^0)^{-2}\,\delta((k-p)^2)\theta(k^0-p^0), \tag{3.11b}$$

$$\frac{1}{(2\pi)^2}\int d^4\xi\, e^{ik\xi}\langle :B_r^3(x): :B_r^3(y):\rangle_0$$
$$= \frac{-6}{(2\pi)^7}\int d^4p\, d^4q\cdots\delta((k-p-q)^2). \tag{3.11c}$$

In Eq. (3.11a) we can smear out the fields with a test function satisfying (3.10). For Eq. (3.11b), we must require that $\tilde{f}(k)\delta((k-p)^2) = 0$ for $p = 0$, and the previous assertion follows. However, for Eq. (3.11c) we would need

$$\tilde{f}(k)\,\delta((k-p-q)^2) = 0 \text{ for } q = 0, \quad \text{all } p, \quad \text{all } k.$$

This is possible only for $\tilde{f} = 0$, and the demonstration is complete.

B. The Coulomb Pontential Φ

We stated in Sec. 2C that the operator $e^{i\beta\Phi(x)}$ is unitary, if β is real. An equivalent assertion is that Φ is self-adjoint.[15] However, Φ has a real form, and so can be assumed self-adjoint. Alternately, one can use the Friedrichs extension theorem. Note that

$$\langle\chi^{(-)}(f^*)\Phi(\mathbf{x})\chi^{(+)}(f)\rangle_0 = \int d^3\mathbf{y}\,\frac{|f(\mathbf{y})|^2}{|\mathbf{x}-\mathbf{y}|} \geq 0. \tag{3.12}$$

Let us derive some commutation relations. Formally one has, if α is a scalar,

$$[A,B] = \alpha B \quad \text{implies} \quad e^A B = e^\alpha(Be^A). \tag{3.13}$$

We can apply this relation to Φ and $\chi^{(\pm)}$:

$$e^{i\beta\Phi(x)}\chi^{(\pm)}(\mathbf{y}) = e^{\pm i\beta/|\mathbf{x}-\mathbf{y}|}\chi^{(\pm)}(\mathbf{y})e^{i\beta\Phi(x)}. \tag{3.14}$$

The use of distributions $\chi^{(\pm)}(\mathbf{y})$ can be justified, since the exponential factor is bounded in the real region.

The observations just noted show, that the unordered exponential is a very convenient quantity to work with. Also, it can be brought easily to a normal form. We state the result, but we shall not use it:

$$e^{i\beta\Phi(x)} = :\exp\left[\int d^3\mathbf{y}\,\chi^{(+)}(\mathbf{y})\chi^{(-)}(\mathbf{y})(e^{i\beta/|\mathbf{x}-\mathbf{y}|}-1)\right]: \tag{3.15}$$

To obtain this equation, one can start with the substitution

$$\chi^{(+)}(\mathbf{y})\chi^{(-)}(\mathbf{y}) \to \chi^{(+)}(\mathbf{y})[\chi^{(-)}(\mathbf{y}) + \delta/\delta\chi^{(+)}(\mathbf{y})], \tag{3.16}$$

and proceed as in the reduction of time-ordered products.[16]

With regard to the Wick-ordered exponential $:e^{i\beta\Phi(x)}:$, the needed commutation rules can be likewise readily found[17]:

$$:e^{i\beta\Phi(x)}:\chi^{(\pm)}(\mathbf{y}) = (1+i\beta\,|\mathbf{x}-\mathbf{y}|^{-1})^{\pm 1}\chi^{(\pm)}(\mathbf{y})\;:e^{i\beta\Phi(x)}:. \tag{3.17}$$

This equation was derived formally, but it can be used to study the questions of domain of $:e^{i\beta\Phi}:$. We shall not study these questions, since the field $\psi_W^{[\pm]}$ turns out to be not altogether satisfactory. One can also see that

$$[:e^{i\beta\Phi(x)}:, \rho(\mathbf{y})] = [:e^{i\beta\Phi(x)}:, :e^{i\beta'\Phi(y)}:] = 0. \tag{3.18}$$

However, it seems difficult to relate $e^{i\beta\Phi(x)}$ to $:e^{i\beta\Phi(x)}:$.

C. Vacuum Expectation Values

We shall now consider various vacuum expectation values in turn. We first note that the electromagnetic field functions are just the free field ones,

$$\langle A_{\nu_1}(x_1)\cdots A_{\nu_n}(x_n)\rangle_0 = \langle A_{\nu_1}^{(0)}(x_1)\cdots A_{\nu_n}^{(0)}(x_n)\rangle_0. \tag{3.19}$$

The electron field functions are readily obtained from the preceding equations. We compute, in particular, the two-point time-ordered function, and compare it with the solution of BS. We will next consider briefly the arbitrary constants and renormalization. Lastly, we will make a remark about mixed electron–photon functions.

The electron two-point function is the same for the F-ordered as for the W-ordered solution. We set $a, b = 0$ in Eq. (3.8) and obtain, where $\xi^0 \equiv t$,

$$W(\xi) \equiv \langle\psi^{[-]}(x)\psi^{[+]}(y)\rangle_0 \tag{3.20}$$
$$= \langle\Upsilon^*\Upsilon\rangle_0\langle\chi^{(-)}\chi^{(+)}\rangle_0\, e^{-imt}$$
$$= C(-t^2+i\epsilon t)^{e^2/8\pi^2}\,\delta(\xi)\,e^{-imt}, \tag{3.21}$$
$$C = \exp[-e^2\Gamma'(1)/4\pi^2], \tag{3.22}$$

[15] F. Riesz and B. Sz.-Nagy, *Functional Analysis* (F. Ungar Publishing Company, New York, 1955), Secs. 123, 124, and 137.

[16] J. L. Anderson, Phys. Rev. 94, 703 (1954). See also BS, Sec. 39.

[17] J. Schwinger, Phys. Rev. 93, 615 (1954), Appendix B.

Since $\langle \Upsilon^* \Upsilon \rangle_0 = e^{e^2 H(t)}$, cf. Eq. (3.5b). In momentum space, $W(\xi)$ becomes the Riesz distribution,[18]

$$\frac{1}{(2\pi)^2} \int d^4\xi \, e^{i\xi k} W(\xi)$$

$$= \frac{C}{(2\pi)^2} \int_{-\infty}^{\infty} dt \, e^{it(\omega - m)} (-t^2 + i\epsilon t)^{e^2/8\pi^2} \tag{3.23}$$

$$= C(2\pi)^{-1} Z_l^{(1)}(\omega - m)$$

$$= C' \theta(\omega - m)(\omega - m)^{-(e^2/4\pi^2)-1}, \tag{3.24}$$

$$l = -e^2/4\pi^2,$$

$$C' = C(2\pi)^{-1} \pi^{\frac{1}{2}} 2^{1-l} [\Gamma(\tfrac{1}{2}l)\Gamma(\tfrac{1}{2}(l+1))]^{-1}.$$

In this model, the time-ordered function is

$$\langle (\psi^{[-1]}(x) \psi^{[+1]}(y))_+ \rangle_0 = W(\xi)\theta(t). \tag{3.25}$$

One sees that its Fourier transform is like the r.h.s. of Eq. (3.23), except that the lower limit is 0 rather than $-\infty$. Aside from an over-all constant factor [cf. Eq. (3.29a) below], this is in agreement with Eq. (41.25) of BS, for $u = (1, \mathbf{0})$ and $d_l = 1$. The convergence factor $e^{-\epsilon t}$ (as in BS) has to be supplied, in accordance with Feynman's rule $m \to m - i\epsilon$.

We now turn to the four-point functions, and let us assume the initial times T_x, etc., in the exponentials of Φ: e.g.,

$$(e^{\pm i\beta\Phi(\mathbf{x})})_{\text{ord}} = (\exp [\pm ie^2(x^0 - T_x)\Phi(\mathbf{x})])_{\text{ord}}.$$

We use Eqs. (3.14) and (3.17) to displace such exponentials toward the vacuum state. The results are, for the two respective orderings,

$$\langle \psi^{[-1]}(u)\psi^{[-1]}(v)\psi^{[+1]}(x)\psi^{[+1]}(y)\rangle_0$$

$$= R_{F,W} \langle \Upsilon^*(u)\Upsilon^*(v)\Upsilon(x)\Upsilon(y)\rangle_0 [\delta(\mathbf{u} - \mathbf{x})\delta(\mathbf{v} - \mathbf{y})$$

$$+ \delta(\mathbf{u} - \mathbf{y})\delta(\mathbf{v} - \mathbf{x})] e^{-im(u^0 + v^0 - x^0 - y^0)}, \tag{3.26}$$

$$R_F = \exp [-ie^2(v^0 - x^0 - T_v + T_x) |\mathbf{x} - \mathbf{y}|^{-1}], \tag{3.27a}$$

$$R_{W_-} = [1 - ie^2(v^0 - T_v) |\mathbf{x} - \mathbf{y}|^{-1}]$$

$$\times [1 + ie^2(x^0 - T_x) |\mathbf{x} - \mathbf{y}|^{-1}]. \tag{3.27b}$$

Note that we can set $|\mathbf{u} - \mathbf{v}| = |\mathbf{x} - \mathbf{y}|$, in view of the δ-functions.

One can see that these four-point functions already have the essential features of the general $2n$-point function.

In the case of F-ordering, one can eliminate the initial times T from the four-point function by choosing $T_v = T_x$. But the elimination is not possible for W-ordering. This is one reason for preferring the F-ordering.

We recall that $(e^{\pm ie^2 T\Phi})_F$ are unitary operators. Therefore, if we choose $T_v = T_x$, etc., we can eliminate the T-dependence of the fields by means of a unitary transformation of the Fock space of $\chi^{(\pm)}$. In fact, one can construct fields independent of T with the help of the reconstruction theorem,[19,20] and they would have just such an interpretation.

Let us now consider the effect of adding $at + ib$ to $\mathfrak{D}^{(-)}$. The two-point function then becomes

$$W(\xi) = C(-t^2 + i\epsilon t)^{e^2/8\pi^2} e^{ie^2 at} e^{-e^2 b} \delta(\xi) e^{-imt}. \tag{3.28}$$

We can interpret the new factors as renormalizations of field strength (of course, the unit of length, and hence an over-all factor of W, remain arbitrary), and of mass:

$$C \to C e^{-e^2 b}, \tag{3.29a}$$

$$m \to m - e^2 a. \tag{3.29b}$$

We shall return soon to the last relation.

Finally, we come to the mixed electron–photon functions. They could be readily obtained with the help of relations like

$$[A_0^{(-)}, \Upsilon] = ie\langle A_0 B\rangle_0 \Upsilon. \tag{3.30}$$

We only note here that these functions in general have terms proportional to a, cf. Eq. (3.8), and also depend on a through exponential factors, as in Eq. (3.28).

D. Canonical Commutation Relations and Counter Terms

The relations and the terms in question may well be considered part of the definition of the model. Like the vacuum expectation values, both the commutation relations and the counter terms show F-ordering as preferable to W-ordering. Both depend on the constant a in interesting ways. In particular, one counter term may reverse the mass renormalization (3.29b).

We first note that the photon field commutators reduce to the free-field function,

$$[A_\mu, A_\nu] = [A_\mu^{(0)}, A_\nu] = -ig_{\mu\nu} D(\xi). \tag{3.31}$$

Let us now assume a fixed time t. With either ordering,

$$[\psi^{[\pm 1]}(\mathbf{x}), \psi^{[\pm 1]}(\mathbf{y})]_+ = 0, \tag{3.32}$$

[18] L. Schwartz, Ref. 12, Vol. I, p. 49; Vol. II, p. 119.

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

[20] R. F. Streater and A. S. Wightman, *PCT, Spin and Statistics, and All That* (W. A. Benjamin, Inc., New York, 1964), Sec. 3-4.

$$[A_0(\mathbf{x}), \psi^{[\pm 1]}(\mathbf{y})] = \mp 2ae\psi^{[\pm 1]}(\mathbf{y}), \quad (3.33a)$$

$$[\dot{A}_0(\mathbf{x}), \psi^{[\pm 1]}(\mathbf{y})] = 0. \quad (3.33b)$$

We would expect both of these commutators to vanish. The term proportional to a comes from Eq. (3.9c) and is surprising, but we do not know what deeper meaning it may have.

For the remaining anticommutator, one has

$$[\psi_F^{[-1]}(\mathbf{x}), \psi_F^{[+1]}(\mathbf{y})]_+ = Z_2^{-1}\,\delta(\mathbf{x}-\mathbf{y}), \quad (3.34)$$

$$Z_2^{-1} = 0 = e^{e^2 H(0)} = \mathrm{T}^*(x)\mathrm{T}(x), \quad (3.35)$$

cf. Eq. (3.20) and subsequent lines. On the other hand, one can easily verify that

$$[\psi_W^{[-1]}(\mathbf{x}), \psi_W^{[+1]}(\mathbf{y})]_+ \neq 0 \quad \text{for} \quad \mathbf{x} \neq \mathbf{y}. \quad (3.36)$$

If the field $A_0^{(0)}$ had normal rather than reversed commutation relations, then we would have[21] $Z_2 = 0$ in Eq. (3.34). For example, we expect $Z_2 = 0$ for the Coulomb gauge. We may also note that for unequal times, the corresponding anticommutator is no longer a c-number. However, it would be finite also in cases where $Z_2 = 0$, and an equation like (3.34) would then express the singular limit of equal times.

Next let us turn to the counter terms. The interaction term for, e.g., $\psi_F^{[-1]}$ can be written in the form (2.5a) as follows:

$$e(A_0(x)\psi_F^{[-1]}(x))_F = \lim_{\epsilon\to 0} e[A_0(x+\epsilon)\psi_F^{[-1]}(x) + ie\langle A_0(x+\epsilon)B(x)\rangle_0\psi_F^{[-1]}(x)], \quad (3.37)$$

where we used Eq. (3.30). In order to write $(\mathrm{A}_0\psi_W^{[-1]})_W$ in such a form, we would need an equation

$$:\Phi\, e^{i\beta\Phi}: = \Phi\, :e^{i\beta\Phi}: + f\, :e^{i\beta\Phi}:,$$

where f is a scalar function. But this is impossible, since the vacuum expectation values of these operators would imply $f = 0$. Analysis of the form (2.5b) is similar.

Let us now return to the constant a and to the mass renormalization. Consider the interaction term (3.37). Equation (3.8) shows that, upon adding $a\xi^0$ to $\mathfrak{D}^{(-)}$, the counter term yields a contribution $-e^2a\psi_F^{[-1]}$ to the r.h.s. So far this is the same renormalization as in (3.29b). We may argue, however, that one should start with the original equation, add $-e^2a\psi_F^{[-1]}$ to both sides, and interpret these terms as follows: as mass renormalization on the l.h.s., and as a part of the counter term on the r.h.s. Then we would obtain a compensating displacement of the mass, $m \to m + e^2a$. This argument, of course, does not lead to any firm conclusion, but rather, it points to a situation that deserves to be clarified.

E. Subsidiary Condition and Positivity of the Metric

As is well known, the positive-definiteness condition requires that the Fourier transform of the two-point function (e.g. of $\psi^{[\pm 1]}$) be a positive measure.[19] In the present model, we had to introduce two different sources of indefinite metric: reversed commutation relations, and the extension of the electromagnetic field by a distribution-theoretic regularization.

For the static electron, the indefiniteness from both sources can be eliminated by the subsidiary condition, but in a way which is not very satisfactory. However, in Sec. 4 we are forced to consider the consequences of regularization in more detail.

In the Heisenberg picture, which we employ, the subsidiary condition is

$$\langle\Psi|\,\partial_\nu A^\nu(x)\,|\Psi\rangle = 0 \quad \text{or} \quad [\partial_\nu A^{(-)\nu}(x)]\Psi = 0. \quad (3.38)$$

In our model one has (also in the nonstatic case, as follows from covariance)

$$\partial_\nu A^{(\pm)\nu} = \partial_\nu A^{(0,\pm)\nu}. \quad (3.39)$$

The condition (3.38) is not fulfilled by, e.g., a one-electron state

$$\Psi_f \equiv \psi^{[+1]}(f)\,|0\rangle. \quad (3.40)$$

One could try to modify such a state as follows. We consider the longitudinal and transverse parts of $\mathbf{A}^{(0)}$, and we introduce

$$\Lambda = \nabla^{-2}\,\mathrm{div}\,\mathbf{A}^{(0)}, \quad (3.41)$$

so that

$$\nabla\Lambda = \mathbf{A}^{(0)L}, \qquad \mathbf{A}^{(0)\mathrm{tr}} = \mathbf{A}^{(0)} - \mathbf{A}^{(0)L}. \quad (3.42)$$

We further define

$$\psi_\Lambda^{[\pm]} = :e^{\pm ie\Lambda}:\,\psi^{[\pm]}, \quad (3.43)$$

and the subsidiary condition will be fulfilled by

$$\Psi_{\Lambda,h} = \psi_\Lambda^{[+]}(h)\,|0\rangle. \quad (3.44)$$

However, the two-point function of Λ needs to be regularized, and we can take

$$\langle\Lambda\Lambda\rangle_0 = -\langle BB\rangle_0 = i^{-1}\,\mathfrak{D}^{(-)}(\xi). \quad (3.45)$$

Then

$$\langle\psi_\Lambda^{[-]}\psi_\Lambda^{[+]}\rangle_0 = \langle\chi^{(-)}\chi^{(+)}\rangle_0 = \delta(\xi). \quad (3.46)$$

Since the field $\psi_\Lambda^{[\pm]}$ does not exhibit an infraparticle structure, it appears that the state $\Psi_{\Lambda,h}$ cannot be

[21] Cf. K. Symanzik, in *Lectures on Field Theory and the Many Body Problem*, edited by E. R. Caianiello (Academic Press Inc., New York, 1961), especially p. 74.

obtained by applying polynomials in $\psi^{[\pm]}$ and A, to the vacumm, and by taking reasonable limits. (The question of limits is delicate, in view of the indefinite metric.) If this is so, this state does not belong to what we would call the space of states of the theory. It also seems pausible, that such a space of states would have no state which satisfies the subsidiary condition, and which has nonzero charge.

If the subsidiary condition is fulfilled for a given state, then this state has a nonnegative norm. (The usual argument, as given for $A_0^{(0)}$ and $\mathbf{A}^{(0)L}$, can be readily adapted to B and Λ.) One therefore obtains a space with positive-definite metric by the standard construction of a quotient space.[20]

We may note, finally, the possibility of using the interaction picture. The electron field then fulfills the subsidiary condition in the following way[22]:

$$\left[\partial_\nu A^{(0,-)\nu}(x) - e\int d^3\mathbf{x}'\,\rho(\mathbf{x}')\,D^{(-)}(x^0 - y^0, \mathbf{x}-\mathbf{x}')\right] \times \psi^{[+]}(y)\,|0\rangle = 0, \qquad (3.47)$$

since $[\partial_0 A^{(0,-)0}, \Upsilon] = eD^{(-)}\Upsilon$. Furthermore, the strong singularity of $\langle\psi^{[-]}\psi^{[+]}\rangle_0$ in momentum space, cf. Eq. (3.24), can be made harmless if we restrict $\psi^{[\pm]}(f)$ to test functions satisfying

$$\int_{-\infty}^{\infty} dt\, e^{\pm imt}\, f(t, \mathbf{x}) = 0 = \tilde{f}(\omega = \mp m, \mathbf{k}).$$

One could therefore try to construct a suitable space of states for the interaction picture. But we do not consider this approach further.

4. FIELDS FOR THE CASE OF A MOVING ELECTRON

We now adapt the preceding results to other velocities, and also to the Coulomb gauge. The present analysis is of interest for several reasons, and we mention one in particular: The problem of proving positive-definiteness of the metric will be rather different from the corresponding discussion in Sec. 3E.

We now assume that the electron field has a velocity 4-vector u, and we consider a proper, orthochronous Lorentz transformation Λ such that

$$\Lambda(1, \mathbf{0}) = u = (u^0, u^1, u^2, u^3) = (u^0, \mathbf{u}) = (u_0, -u_1, -u_2, -u_3). \qquad (4.1)$$

Of course, Λ is not determined uniquely by u, and we suppose that a particular Λ has been chosen.

We continue using the notation of the previous sections, but the fields will now refer to the velocity u, i.e., to Eqs. (1.2), unless otherwise stated.

A. Free-Electron Field

The equations for the free-electron field,

$$(iu^0\,\partial_0 + iu\nabla \pm m)\psi^{(0,\pm)} = 0, \qquad (4.2)$$

have the solution

$$\psi^{(0,\pm)}(x) = e^{\pm im(u^0x^0 - ux)}\chi^{(\pm)}(x). \qquad (4.3)$$

The fields $\chi^{(\pm)}$ of course depend on only three variables,

$$\psi^{(0,\pm)}(x) = e^{\pm imt^0}\chi^{(\pm)}(\zeta), \text{ where } \zeta = \Lambda^{-1}x. \qquad (4.4)$$

The fields $\chi^{(\pm)}$ are therefore entirely analogous to the corresponding static fields, and we normalize them so that $\chi^{(+)} = \chi^{(-)*}$ and

$$[\chi^{(\pm)}, \chi^{(\pm)}]_+ = 0, \quad [\chi^{(-)}(\zeta), \chi^{(+)}(\zeta')]_+ = \delta(\zeta - \zeta'). \qquad (4.5)$$

On the other hand, the canonical anticommutation relations require adjoint fields,

$$\chi = \chi^{(-)}, \qquad \bar{\chi} = u^0\chi^{(+)}, \qquad (4.6)$$

$$[\chi(t, \mathbf{x}), \bar{\chi}(t, \mathbf{x}')]_+ = \delta(\mathbf{x} - \mathbf{x}'). \qquad (4.7)$$

In the momentum space, we have

$$\tilde{\chi}^{(\pm)}(t, \mathbf{k}) = (2\pi)^{-\frac{3}{2}}\int d^3\mathbf{x}\, e^{\pm ikx}\chi^{(\pm)}(t, \mathbf{x}) = e^{\pm ikut/u^0}\tilde{\chi}^{(\pm)}(0, \mathbf{k}), \qquad (4.8)$$

$$[\tilde{\chi}^{(-)}(t, \mathbf{k}), u^0\tilde{\chi}^{(+)}(t, \mathbf{k}')]_+ = \delta(\mathbf{k} - \mathbf{k}'). \qquad (4.9)$$

For convenience we also note

$$[\chi^{(-)}(t, \mathbf{x}), \chi^{(+)}(t', \mathbf{x}')]_+ = \delta[\mathbf{u}(t - t') - u^0(\mathbf{x} - \mathbf{x}')]. \qquad (4.10)$$

The free current now has up to four nonzero components,

$$j^{(0)\nu} = u^\nu\chi^{(+)}\chi^{(-)} \equiv u^\nu J = v^\nu\bar{\chi}\chi \quad \text{where} \quad v = (1, \mathbf{v}). \qquad (4.11)$$

The static case and an invariance argument show that

$$\partial_\nu j^{(0)\nu} = 0.$$

We also see that

$$[j^{(0)\mu}, j^{(0)\nu}] = 0, \qquad (4.12)$$

in contrast to the well-known result of Schwinger.[23] This difference in properties of the current can be

[22] J. M. Jauch and F. Rohrlich, Ref. 10, Sec. 6-3.

[23] J. Schwinger, Phys. Rev. Letters 3, 296 (1959).

related to the fact, that in a conventional relativistic theory,[24]

$$j(x)\,|0\rangle = 0 \quad \text{implies} \quad j(x) = 0. \tag{4.13}$$

This implication is not valid in the Bloch–Nordsieck model.

We shall make use of the Fourier transforms of J. We introduce

$$\kappa \equiv \Lambda^{-1}k, \qquad w \equiv \kappa^0 = u^0\omega - \mathbf{uk}, \tag{4.14}$$

$$J(\zeta) = \frac{1}{(2\pi)^{\frac{1}{2}}}\int d^3\kappa\, e^{\mp i\kappa\zeta}\mathfrak{J}^{\pm}(\kappa) = \frac{1}{(2\pi)^{\frac{1}{2}}}\int d^3\mathbf{k}\, e^{\mp ikx}\mathfrak{J}^{\pm}(t,\mathbf{k}). \tag{4.15}$$

Then, if the vectors k and κ are restricted to the future light cone,

$$\mathfrak{J}^{\pm}(\kappa) = (\omega/w)\, e^{\pm it(u^0w-\omega)}[(u^0\omega^{-1})uk]\mathfrak{J}^{\pm}(t, \mathbf{k} - w\mathbf{u}). \tag{4.16}$$

Here $(u^0\omega^{-1})uk$ is the Jacobian $\partial(\mathbf{k} - w\mathbf{u})/\partial(\mathbf{k})$.

B. Interacting fields in the Feynman gauge

Let us turn to Maxwell's equations. Their solution is

$$A^{\nu}(x) = A^{(0)\nu}(x) + eu^{\nu}\Phi(x), \tag{4.17}$$

$$\Phi(x) = \Phi(\zeta) = \frac{1}{4\pi}\int d^3\mathbf{n}\,\frac{J(\mathbf{n})}{|\mathbf{n} - \boldsymbol{\zeta}|}. \tag{4.18}$$

As in the static case, this solution can be split into two Hermitian-conjugate parts $A^{(\pm)}$. One can, e.g., use the fixed-source solution with reference to the moving frame, and then use Eq. (4.16):

$$A^{(\pm)\nu}(x) = \frac{1}{(2\pi)^{\frac{1}{2}}}\int \frac{d^3\mathbf{k}}{(2\omega)^{\frac{1}{2}}}\, a^{(\pm)\nu}(t,\mathbf{k})\, e^{\mp ikx}, \tag{4.19a}$$

$$a^{(\pm)\nu}(t,\mathbf{k}) = a^{(0,\pm)\nu}(\mathbf{k})e^{\pm i\omega t} + eu^{\nu}[2^{-\frac{1}{2}}(4\pi)^{-1}]\mathfrak{J}^{\pm}(t,\mathbf{k})E(\mathbf{k}), \tag{4.19b}$$

$$E(\mathbf{k}) = \omega^{\frac{1}{2}}w^{-2}(\mathbf{l}_k) = |\mathbf{k}|^{\frac{1}{2}}\,(u^0\,|\mathbf{l}_k| - \mathbf{ul}_k)^{-2}, \tag{4.20}$$

where $\mathbf{k} \to \mathbf{l}_k$ is the map inverse to $\mathbf{k} \to \mathbf{k} - w\mathbf{u}$.

For the electron field, we need integrals B^{ν} of all components $A^{(0)\nu}$. These integrals are readily evaluated if we use transformed coordinates, and proceed as in Eq. (2.27). The result is as in this equation, but with the following replacement of the denominator:

$$(2\omega)^{-\frac{1}{2}}\omega^{-1} \to (2\omega)^{-\frac{1}{2}}w^{-1}. \tag{4.21}$$

[24] P. G. Federbush and K. Johnson, Phys. Rev. **120**, 1926 (1960); K. Johnson, Nucl. Phys. **25**, 431 (1961). Also, Ref. 20, p. 163.

The electron field can now be expressed as follows:

$$\psi^{[\pm]} = :e^{\pm ieu^{\nu}B_{\nu}}: (e^{\pm i\beta\Phi}\psi^{(0,\pm)})_{\text{ord}}. \tag{4.22}$$

Here $\beta = e^1(\zeta^0 - Z^0)$, and Z^0 is an arbitrary value of ζ^0.

Let us consider the vacuum expectation values. For B^{ν}, let us adjust the constants so that

$$\langle B^{\mu}B^{\nu}\rangle_0 = -i^{-1}g^{\mu\nu}\mathfrak{D}^{(-)}(\Lambda^{-1}\xi). \tag{4.23}$$

This equation will result if we regularize w^{-1} rather than ω^{-1}. From this equation one obtains

$$\langle\psi^{[-]}\psi^{[+]}\rangle_0 = W(\Lambda^{-1}\xi), \tag{4.24}$$

and similarly for other functions. To express these functions in p-space, one only needs to take the expressions for the static case, and to make the replacement $\omega \to up$.

C. Interacting Fields in the Coulomb Gauge

The difficulties which we observed with the subsidiary condition can be eliminated by employing the Coulomb gauge.[11,25] We now present the relevant formulas. For the static case, this gauge yields $A_0 = e\Phi$, so that

$$\psi^{[\pm]} = (e^{\pm i\beta\Phi}\chi^{(\pm)})_{\text{ord}}\, e^{\pm imt}, \tag{4.25}$$

$$\langle\psi^{[-]}\psi^{[+]}\rangle_0 = \langle\chi^{(-)}\chi^{(+)}\rangle_0 = \delta(\xi), \tag{4.26}$$

as in Eq. (3.46).

For the case of general vector u, we select the transverse gauge, and write the field equations as

$$(iu^{\nu}\partial_{\nu} \pm m)\psi^{[\pm]} = \mp e((u^0A_0 - \mathbf{uA}^{\text{tr}})\psi^{[\pm]})_{\text{ord}}, \tag{4.27}$$

$$\nabla^2 A_0 = -ej_0^{(0)}, \qquad \Box\mathbf{A}^{\text{tr}} = -e\mathbf{j}^{(0)\text{tr}}. \tag{4.28}$$

We note the commutation relations for the spatial components,

$$[\dot{A}_r^{\text{tr}}(t,\mathbf{x}),\, A_s^{\text{tr}}(t,\mathbf{y})] = i^{-1}(\delta_{rs} - \partial_r\partial_s\nabla^{-2})\delta(\mathbf{x} - \mathbf{y}). \tag{4.29}$$

Maxwell's equations have the solution

$$A_0(x) = eu_0\Phi_C(x) \equiv \frac{eu_0}{4\pi}\int d^3\mathbf{y}\,\frac{J(x^0,\mathbf{y})}{|\mathbf{x} - \mathbf{y}|}, \tag{4.30}$$

$$\mathbf{A}^{\text{tr}} = \mathbf{A}^{(0)\text{tr}} + e(\mathbf{u}\Phi)^{\text{tr}}, \tag{4.31}$$

Note that $\Phi_C \neq \Phi$. The Fourier decomposition of the Hermitian conjugate parts can be obtained from Eqs. (4.19) by making the replacements (where $\boldsymbol{\varepsilon}_{k\lambda} \perp \mathbf{k}$)

$$\mathbf{a}^{(0,\pm)}(\mathbf{k}) \to \sum_{\lambda=1,2}\boldsymbol{\varepsilon}_{k\lambda}a_{\lambda}^{(0,\pm)}(\mathbf{k}), \qquad \mathbf{u} \to \sum_{\lambda=1,2}\boldsymbol{\varepsilon}_{k\lambda}(\boldsymbol{\varepsilon}_{k\lambda}\mathbf{u}). \tag{4.32}$$

[25] B. Zumino, J. Math. Phys. **1**, 1 (1960).

The electron field $\psi^{[\pm]}$ is given by

$$\psi^{[\pm]} = :\exp(\pm ie\mathbf{u}\mathbf{B}^{tr}):$$
$$\times (\exp\{\pm i\beta[u_0^2\Phi_C - \mathbf{u}(\mathbf{u}\Phi)^{tr}]\}\psi^{(0,\pm)})_{ord}. \quad (4.33)$$

The second exponent is unitary in case of F-ordering. As before, this follows from reality of the expression in the brackets.

We do not evaluate the electron two-point function. The relevant calculations are somewhat involved, and we only note that

$$\langle B_r^{tr}B_s^{tr}\rangle_0 = \frac{1}{2(2\pi)^3}\int \frac{d^3\mathbf{k}}{\omega}\frac{1}{w}(w^{-1})_+ e^{-ik\xi}\left(\delta_{rs} - \frac{k_rk_s}{\omega^2}\right). \quad (4.34)$$

Alternately, one could try to adapt functional methods [Ref. 25, especially Eq. (47)] to the present problem.

D. Positivity of the Metric

Most of the properties of the solution for a static electron, as considered in Sec. 3, carry over readily to the nonstatic case. We also indicated a few modifications earlier in this section. However, the question of positive-definiteness is more subtle than before, and is discussed presently. A fixed velocity vector u is assumed.

Consider a one-electron state $\psi^{[+]}(f)\,|0\rangle$ in the Feynman gauge, and suppose that the field B_0 has been omitted from the dressing exponential. Then it is easy to see, that this state will have a positive norm, for any nonnull function f. Moreover, such a conclusion appears very plausible for the case of the Coulomb gauge, where the longitudinal field is further omitted from the dressing exponential. We did not carry out the relevant calculations, however.

We shall now concentrate on the Coulomb gauge. For the investigation of positive-definiteness for the Feynman gauge, with the subsidiary condition assumed, one would in fact have to combine the present discussion with that of Sec. 3E.

Consider the applications of operators of the forms $\psi^{[\pm]}(f)$ and $\mathbf{n}\mathbf{A}^{tr}(g)$ to the vacuum. Linear combinations of such states and suitable limits define the space of physical states:

$$\mathfrak{K}^{phys} = \{\sum \mathbf{n}_1\mathbf{A}^{tr}(g_1)\cdots\mathbf{n}_m\mathbf{A}^{tr}(g_m) \times \psi^{[-]}(f_1)\cdots\psi^{[+]}(h_l)\,|0\rangle\}. \quad (4.35)$$

Our problem is to show that the states of this space have a positive norm (or at least, a nonnegative norm).

The sectors corresponding to different charges are, however, orthogonal, and in fact can be separated with a superselection rule. Therefore it suffices to establish that the norm is positive in each sector.

The corresponding problem for Schroer's model has been thoroughly discussed.[3] The technique there was to give the boson field a positive mass μ, and to let $\mu \to 0$ in a way which would manifestly preserve positive-definiteness. Such a technique will also work for the present case. However, we have to make one assumption:

$$\langle :\exp[ie\mathbf{u}\mathbf{B}^{tr}(x)]: :\exp[-ie\mathbf{u}\mathbf{B}^{tr}(y)]:\rangle_0 = \lim_{\mu\to 0}\mu^\alpha\langle :\exp[ie\mathbf{u}\mathbf{B}_\mu^{tr}(x)]: :\exp[-ie\mathbf{u}\mathbf{B}_\mu^{tr}(y)]:\rangle_0, \quad (4.36)$$

where B_μ refers to the time integral of the vector field $\mathbf{A}_\mu$ with mass μ, and α is a (real) constant. One could try to confirm this relation by explicit calculations, or perhaps by some ingenious device.[26] The analogous relation for Schroer's model is indeed fulfilled.

Let us now outline a proof which differs slightly from that of Ref. 3. We consider the following as a typical state having one unit of charge:

$$\Psi = [\psi^{[+]}(f_1) + \mathbf{n}\mathbf{A}^{tr}(g)\psi^{[+]}(f_2) + \psi^{[-]}(h_1)\psi^{[+]}(h_2)\psi^{[+]}(f_3)]\,|0\rangle. \quad (4.37)$$

We construct a new vector Ψ_μ, defined in terms of $\mathbf{A}_\mu$ and $\mathbf{B}_\mu$ and such that

$$||\Psi|| = \lim_{\mu\to 0}\mu^\alpha\,||\Psi_\mu||. \quad (4.38)$$

For each $\mu > 0$, one will have $||\Psi_\mu|| \geq 0$, on the basis of general theory of free fields, and $||\Psi|| \geq 0$ will follow.

To construct Ψ_μ, we make the replacements $\mathbf{A} \to \mathbf{A}_\mu$, $\mathbf{B} \to \mathbf{B}_\mu$ in (4.37), except that the operators $\psi^{[\pm]}(h_i)$ are treated in the following way. We write

$$\psi^{[-]}(h_1)\psi^{[+]}(h_2) = \int d^4x\,d^4y\,h_1(x)h_2(y)\psi^{(0,+)}(x)\psi^{(0,-)}(y)H(x,y),$$
$$H(x,y) = \exp\{e^2\langle[\mathbf{u}\mathbf{B}^{tr}(x)][\mathbf{u}\mathbf{B}^{tr}(y)]\rangle_0\} \times :\exp\{ie\mathbf{u}[\mathbf{B}^{tr}(x) - \mathbf{B}^{tr}(y)]\}:. \quad (4.39)$$

Now, the first exponential is a c-number function, and we leave it unaltered. In the second exponential we replace $\mathbf{B}^{tr}$ by $\mathbf{B}_\mu^{tr}$, as before.

One may check that this prescription indeed yields Ψ_μ satisfying Eq. (4.38), provided Eq. (4.36) holds. The motivation for this construction depends on the

[26] See, e.g., A. Jaffe, J. Math. Phys. **6**, 1172 (1965).

interpretation of the second exponential in (4.39) as a bilocal field[5] which leaves invariant (when smeared out) every charge sector of $\mathfrak{H}^{\text{phys}}$.

5. REMARKS ON INFRARED REPRESENTATIONS

A. Representations of the Spatial Components A^s

Our approach to the question of infrared representations follows the discussion of Schroer's model.[5] However, the present arguments are less complete. We assume a fixed velocity vector u.

For definiteness, we shall be referring to the Coulomb gauge, and to $\mathbf{A}^{\text{tr}}$. However, one may also wish to consider some framework where $\mathbf{A}^L$ is an independent field, whose effects are not cancelled by A^0, and where the metric is positive-definite. Then our arguments could be readily adapted to $\mathbf{A}^L$.

We now introduce an enlarged space of states $\mathfrak{H}^*$. For this space, the free electron field and the exponential are not assumed to be related in space–time. Thus

$$\mathfrak{H}^* = \sum_{n=0}^{\infty}{}^{\oplus} \mathfrak{H}^{[n]}(\psi^{(0)}) \otimes \mathfrak{H}_n(\mathbf{A}^{(0)\,\text{tr}}), \tag{5.1}$$

where $\mathfrak{H}_n(\mathbf{A}^{(0)\,\text{tr}})$ is generated by n exponentials of $ie\mathbf{u}\mathbf{B}^{\text{tr}}$ applied to the vacuum, and by $\mathbf{A}^{(0)\,\text{tr}}$. The space $\mathfrak{H}^{[n]}$ is the n-particle sector of $\mathfrak{H}_{\text{Fock}}(\psi^{(0)})$. In particular, for $n = 0$, the term is just $|0\rangle_{\psi^{(0)}} \otimes \mathfrak{H}_{\text{Fock}}(\mathbf{A}^{\text{tr}})$. We see that

$$\mathfrak{H}^{\text{phys}} \subset \mathfrak{H}^* \subset \mathfrak{H}^{\max} \equiv \mathfrak{H}_{\text{Fock}}(\mathbf{B}^{\text{tr}}) \otimes \mathfrak{H}_{\text{Fock}}(\psi^{(0)}). \tag{5.2}$$

A partial completion of $\mathfrak{H}_{\text{Fock}}(\mathbf{B}^{\text{tr}})$ is presupposed here.

Let us turn to the infrared representations of $\mathbf{A}^{\text{tr}}$ in sectors of nonzero charge. We start with some commutation relations:

$$[a_\lambda^{(\pm)}(t, \mathbf{k}), :\exp[ie\mathbf{u}\mathbf{B}^{\text{tr}}(t', \mathbf{y})]:] = K\mathbf{u}\boldsymbol{\varepsilon}_{\mathbf{k}\lambda} e\omega^{-\frac{1}{2}} w^{-1} e^{\pm i\mathbf{k}\mathbf{y}} e^{\pm i\omega(t-t')} \times :\exp[ie\mathbf{u}\mathbf{B}^{\text{tr}}(t', \mathbf{y})]:, \tag{5.3}$$

$$[a_\lambda^{(\pm)}(t, \mathbf{k}), \psi^{(0,+)}(t', \mathbf{y})] = K'\mathbf{u}\boldsymbol{\varepsilon}_{\mathbf{k}\lambda} eE(\mathbf{k})(u^0)^{-1} e^{\pm i\mathbf{k}\mathbf{y}} e^{\pm i\mathbf{k}u(t-t')/u^0} \psi^{(0,+)}(t', \mathbf{y}), \tag{5.4}$$

$$K = 2^{-\frac{1}{2}}(2\pi)^{-\frac{3}{2}}, \qquad K' = (4\pi)^{-1}2^{-\frac{1}{2}}(2\pi)^{-\frac{3}{2}}. \tag{5.5}$$

From these equations we obtain, for one unit of charge,

$$\{a_\lambda^{(-)}(t, \mathbf{k}) - \mathbf{u}\boldsymbol{\varepsilon}_{\mathbf{k}\lambda} e[K\omega^{-\frac{1}{2}} w^{-1} e^{-i\omega(t-t')} + K'E(\mathbf{k})(u^0)^{-1} e^{-i\mathbf{k}u(t-t')/u^0}] e^{-i\mathbf{k}\mathbf{y}}\} \psi^{[+1]}(t', \mathbf{y})\ |0\rangle = 0. \tag{5.6}$$

We see distributions $a_\lambda^{(-)}$ translated by a function, which has the expected singularity at $\mathbf{k} = \mathbf{0}$, i.e., $|\mathbf{k}|^{-\frac{1}{2}d}$ where d is the number of spatial dimensions. This translation has two separate origins: the structure of the electron field, and the current in Maxwell's equations.

There remains the problem of describing the action of the $a_\lambda^{(\pm)}$ on normalizable vectors, e.g., $\psi^{[+1]}(f)\ |0\rangle$. (We may observe that $\mathbf{A}^{\text{tr}}$ needs to be smeared out in space only, but $\psi^{[\pm]}$, also in time.) We conjecture that the sector $\mathfrak{H}$ of one unit of charge can be represented as a direct integral,

$$\mathfrak{H} = \int_0^{2\pi}{}^{\oplus} d\sigma\ \mathfrak{H}_\sigma \tag{5.7}$$

where in $\mathfrak{H}_\sigma$ the following distributions define a Fock representation:

$$a_\lambda^{(\pm)}(t, \mathbf{k}) - \mathbf{u}\boldsymbol{\varepsilon}_{k\lambda} e[K\omega^{-\frac{1}{2}} w^{-1} + e^{\pm i\sigma} K'E(\mathbf{k})(u^0)^{-1}] F(\mathbf{k}). \tag{5.8}$$

Here F is any real function, differentiable at $\mathbf{k} = \mathbf{0}$, and satisfying

$$F(\mathbf{0}) = 1, \qquad \int d^3\mathbf{k}\ |F|^2 < \infty. \tag{5.9}$$

To motivate this conjecture, in particular that the distributions (5.8) are Fock, we observe that the smearing of charges can be thought of as "providing an ultraviolet cutoff." Therefore, the function which defines the translation should be non-square-integrable only at the origin, and not at infinity.

This argument can be made precise for the effect of the dressing exponential. Indeed, we can replace t' by $t' + i\tau$, $\tau > 0$, in Eq. (5.3). Then we obtain infrared representations of $\mathbf{A}^{(0)\,\text{tr}}$ in $\mathfrak{H}_n(\mathbf{A}^{(0)\,\text{tr}})$, which are characterized by the distributions (5.8), but with the last term (proportional to K') omitted; i.e., the resulting distributions have the Fock property.

Unfortunately, we do not know if the effect of the current in Maxwell's equations can be analyzed in similar terms. Indeed, the following prototype problem is still unsolved, apparently: to determine the representations of φ, satisfying

$$\Box\varphi(t, \mathbf{x}) = g\rho(\mathbf{x}). \tag{5.10}$$

Here ρ is as in Sec. 2, and g is a constant.

We may note, finally, that Bloch and Nordsieck[1] assumed a c-number electron field, and therefore they did not encounter the effect of the current. Indeed, they solve the problem by a canonical transformation, which is a translation resembling (5.8), but with the last term again omitted. In the

approach of Bloch and Nordsieck, one can readily eliminate ultraviolet effects by introducing an electron form factor, but these authors worked with point electrons.

B. Representations of the temporal component A^0

If we work in, e.g., the Feynman gauge and ignore the subsidiary condition, then we encounter an infraparticle structure consisting of timelike photons. This structure seems more intricate than the corresponding effect of spacelike photons. Recall, e.g., the nonintegrable singularity of the electron two-point function, Eq. (3.24). In the Thirring model we can find another instance of infraparticle structure arising out of a field with reversed commutation relations.[3]

If one hopes to clarify the infraparticle structure in these cases, most likely it will be necessary to introduce the concept of equivalent representations of such fields. One could then hope, e.g., to relate the singularity of a vacuum expectation value to an equivalence class of representations.

In the study of spaces with indefinite metric, a standard technique[13] is to introduce a positive-definite form in addition to the indefinite one. This is familiar for the Fock space of $A_0^{(0)}$:

$$\langle \alpha, \beta\rangle_{\text{ind}} = \sum \langle \alpha_n, \beta_n\rangle, \quad \langle \alpha, \beta\rangle_{\text{pd}} = \sum (-1)^n \langle \alpha_n, \beta_n\rangle,$$

where we sum over contributions of n-particle sectors. The positive-definite form allows us, e.g., to study the expansion of the operator

$$\int d^4x\, d^4y\, f(x)\, g(y)\, :\exp\{ieu^0[B_0(x) - B_0(y)]\}:,$$

and one can show, as in Schroer's model, that this operator leaves $\mathfrak{C}_{\text{Fock}}(A_0^{(0)})$ invariant.

For the study of sectors of nonzero charge, we can proceed as with the spatial components. The commutation (5.3)–(5.4) remain valid, if we make the replacements

$$a_\lambda^{(\pm)} \to a_0^{(\pm)}, \quad \mathbf{u}\mathbf{B}^{\text{tr}} \to u^0 B_0, \quad K \to -K, \quad \mathbf{u}\varepsilon_{\mathbf{k}\lambda} \to u^0. \tag{5.11}$$

Again we may hope that the sector of, say, unit charge, can be decomposed as the direct integral (5.7) of Fock representations, which are defined by the following distributions:

$$a_0^{(\pm)}(t, \mathbf{k}) - e[-u^0 K\omega^{-\frac{1}{2}} w^{-1} + e^{\pm i\sigma} K' E(\mathbf{k})]F(\mathbf{k}). \tag{5.12}$$

The function F has the same restrictions as before.

In conclusion we may note, that such Fock representations would suggest the introduction of a positive-definite form (or forms) in the sector of given charge.

6. CONCLUSION

We should like to emphasize the following features of our solution as perhaps the most significant:

(1) The key to the solution is the identity $j_{\text{ren}} = j^{(0)}$.

(2) By selecting F-ordering and $a = 0$, we obtained a solution that has the properties which one normally expects of a quantum field theory, aside from relativistic invariance and crossing symmetry. On the other hand, the natural alternatives of W-ordering and of $a \neq 0$ exhibit some anomalies. For comparison we may note another model (two-dimensional electrodynamics) where a similar arbitrary constant destroys Lorentz covariance.[27]

(3) Another feature of the model is a striking resemblence to several two-dimensional models, e.g., Schroer's model, and the interaction of massless spinors with massive vector mesons.[28] This resemblance is not only with respect to the infraparticle structure, but also with respect to renormalization. Indeed, in all these models we have $Z_1 = Z_2 = 0$ (or $Z_1^{-1} = Z_2^{-1} = 0$, as a result of reversed commutation relations), $Z_3 = 1$, δ_m infinite (if $m \neq 0$), and δ_{μ^2} finite, or even zero. The assertions about mass renormalization are based on perturbation theory.

Of the foregoing features, one could readily conjecture the first, without knowledge of the solution. The remaining ones, however, do not seem so intuitive.

We now summarize some technical questions which arose but were not settled during this investigation.

(1) One could ask whether there may be other solutions to the field equations, e.g., solutions which violate our assumption $j_{\text{ren}} = j^{(0)}$.

(2) Perhaps a more systematic approach to renormalization can be found. In particular, most likely there is some deeper significance to the particular counter terms which occur in this model, and also to F-ordering.

(3) The constant a should be better understood, especially its role in the commutators $[A_0, \psi^{[\pm]}]$, the counter terms, and the electron mass renormalization.

(4) One could ask whether a more satisfactory

[27] L. S. Brown, Nuovo Cimento 29, 617 (1963).
[28] G. Wraith, thesis, Cambridge University, 1964.

way could be found to handle the subsidiary condition.

(5) We assumed Eq. (4.36), and this equation remains to be proved. Or perhaps an alternate way could be found to establish positivity of the metric, for the Coulomb gauge.

(6) One would like to determine the equivalence class of representations of fields for the case of scattering by a given operator source, like $\rho(\mathbf{x})$.

(7) It might be desirable to formulate the problem of equivalence of representations of fields, for the case of indefinite metric.

ACKNOWLEDGMENTS

I would like to thank the following for discussions: Drs. and Professors M. E. Arons, J. J. Loeffel, O. Steinmann, A. S. Wightman, W. Zimmermann, B. Zumino, and especially K. Symanzik. The initial stages of this work were spread out over visits to several institutions, namely Imperial College, London; Institut des Hautes Etudes Scientifiques, Bures-sur-Yvette; and Steklov Mathematical Institute, Moscow. This work was completed at the Courant Institute. I would like to express my gratitude for the hospitality which I received.

Indefinite Metric Resulting from Regularization in the Infrared Region*

DANIEL A. DUBIN
Physics Department, New York University, New York, New York
AND
JAN TARSKI†
Courant Institute of Mathematical Sciences, New York University, New York, New York
(Received 13 July 1965)

We consider the free scalar massless field in two dimensions, defined with the help of Wightman's regularization at $p = 0$. The one-particle sector of this field is completed, and is shown to be a space of type Π_1, studied by Iohvidov and Krein. (Alternately, one can also obtain a direct sum of two such spaces.)

1. INTRODUCTION

RECENT work on quantum field theory models has led to several instances in which peculiarities of the infrared were observed. A typical result would be the occurrence of power function branch points. Our current understanding of these peculiarities can be paraphrased as follows.

The free scalar massless field, ϕ, in two dimensions is anomalous at the origin[1] $p = 0$. In particular, the field gives rise to integrals divergent at $p = 0$. Such integrals can be regularized by making the replacement[2]

$$\theta(w)w^{-1} \to d/dw[\theta(w) \log w] \equiv (w^{-1})_+ = \theta(w)w^{-1} \quad \text{for} \quad w \neq 0. \tag{1.1}$$

This regularization is then the basis of the infrared behavior in various simple models in two and in four dimensions.[2-5]

We shall take the Fourier transform of $F(x)$ to be

$$\tilde{F}(p) = (2\pi)^{-1} \int d^2x\, e^{-ipx} F(x). \tag{1.2}$$

The scalar product for one-particle states of ϕ now has the form

$$\langle \phi(F^*)\phi(H)\rangle_0 = \pi \int d^2p\, \sigma(p)(F^*)^{\sim}(p)\tilde{H}(-p), \tag{1.3}$$

$$\sigma(p) = 2[\delta(u)(v^{-1})_+ + \delta(v)(u^{-1})_+], \tag{1.4}$$

where u and v are the light-cone coordinates,

$$u = p^0 - p^1, \qquad v = p^0 + p^1. \tag{1.5}$$

As a result of the foregoing regularization, one encounters two principal differences from the familiar properties of the free fields.[6] Firstly, the metric is indefinite, since the Fourier transform (1.4) of the two-point function is not a measure.[7] Secondly, Eq. (1.3) requires F and H to be test functions, e.g., in $\mathcal{S}$. It is not at all clear to what extent this equation can be made meaningful for arbitrary square-integrable functions. A suggestion was made[2] to restrict the underlying space of ϕ to suitably defined sequences of test functions, e.g., such as considered by Kristensen *et al.*[8]

The restriction of ϕ to very smooth functions, however, is not always convenient. We therefore consider the question of completion of spaces. We show how the completed one-particle sectors of ϕ, and of a similar field in four dimensions, can be related to the Π_κ -spaces which were investigated by Pontriagin[9] and by Iohvidov and Krein.[10] This completion allows us to consider simple examples of spectral decomposition. However, a brief examination of the sectors of two or more particles shows that there the question of completion is much more subtle.

We may observe, for contrast, that the more familiar examples of indefinite metric in quantum field theory are due to reversed commutation relations.[11] In such examples, the problem of completion

* This work was supported in part by the National Science Foundation, Grant NSF-GP-2416 (D.A.D.), and in part by a Ford Foundation grant for mathematical physics (J.T.).
† Present address: Department of Physics, Columbia University, New York, New York.
[1] This fact was noted, for instance, by K. Symanzik, in *Lectures on High Energy Physics* (at Hercegnovi, 1961), edited by B. Jakšić (Federal Nuclear Energy Commission of Yugoslavia, 1962), Vol. II, p. 500.
[2] A. S. Wightman, lectures at Summer School in Cargèse, Corsica, 1964.
[3] B. Schroer, Fortsch. Phys. 11, No. 1 (1963), Sec. III.
[4] J. Tarski, J. Math. Phys. 5, 1713 (1964). The present field ϕ corresponds to φ_k of this reference.
[5] J. Tarski, J. Math. Phys. 7, 560 (1966) (previous paper).
[6] E.g., A. S. Wightman and L. Gårding, Arkiv Fysik 28, 129 (1964).
[7] E.g., A. S. Wightman, Phys. Rev. 101, 860 (1956).
[8] P. Kristensen, L. Mejlbo, and E. Thue Poulsen, Commun. Math. Phys. 1, 175 (1965).
[9] L. S. Pontriagin, Izv. Akad. Nauk SSSR. Ser. Mat. 8, 243 (1944).
[10] I. S. Iohvidov and M. G. Krein, Trudy Moscov. Mat. Obsc. 5, 367 (1956) and 8, 413 (1959); [English transl.: Am Math Soc. Transl., Ser. 2, 13, 105; 34, 283].
[11] For instance, E. C. G. Sudarshan, Phys. Rev. 123, 2183 (1961); H. J. Schnitzer and E. C. G. Sudarshan, *ibid.*, 2193.

is trivial. We also refer the reader to a series of papers by Nevanlinna,[12] for investigations of indefinite metric from a more general point of view.

For reference we give a few more formulas. In coordinate space, Eq. (1.3) becomes, where C is a constant,

$$\langle\phi(F^*)\phi(H)\rangle_0 = -(4\pi)^{-1}\int d^2x\, d^2y\, F^*(x)H(y) \times [\log(-\xi^2 + i\epsilon\xi^0) + C], \quad (1.6)$$

$$\xi = x - y, \qquad \xi^2 = (\xi^0)^2 - (\xi^1)^2. \quad (1.7)$$

We observe that $\sigma(p)$ may have a contribution $C_1\delta(p)$, where C_1 may be made arbitrary, by changing the scale of the logarithm in (1.1). We choose $C_1 = 0$ since we may want to allow, later, functions discontinuous at $p = 0$. Then

$$C = (2\pi)^{-1}[\log 2 + \Gamma'(1)]. \quad (1.8)$$

We also note that $\sigma(p) = 2\omega(p)$ of Ref. 4. The factors are now adjusted so that

$$\sigma(p)\, d^2p \to d\Omega = dp^1/|p^1| \quad (1.9)$$

for functions vanishing at $p = 0$.

2. ONE-PARTICLE SECTOR OF ϕ

Let us first identify the one-particle sector of ϕ with a space of functions on the future light-cone. The inner product (1.3) is given by a sum of two terms, one of which can be written as

$$\pi\int_{-\infty}^{+\infty} du(u^{-1})_+(\tilde{F})^*(-\tfrac{1}{2}u, \tfrac{1}{2}u)\tilde{H}(-\tfrac{1}{2}u, \tfrac{1}{2}u), \quad (2.1)$$

since $(F^*)\tilde{}(p) = (\tilde{F})^*(-p)$. We now introduce the correspondence

$$\phi(H)\,|0\rangle \leftrightarrow \pi^{\frac{1}{2}}\tilde{H}(-\tfrac{1}{2}u, \tfrac{1}{2}u) \equiv h(u), \quad (2.2)$$

and $h(u)$ can be considered a function on the ray $u > 0, v = 0$. Similarly, for the other term of (1.3), functions on the ray $u = 0, v > 0$ are relevant.

We consider only one such ray now, i.e., $f(w)$ defined for $0 < w < \infty$. The following Hermitian inner product, implied by (2.1) and (2.2), is of interest to us:

$$\langle f, g\rangle = -\int_0^\infty dw \log w \frac{d}{dw}[f^*(w)g(w)]. \quad (2.3)$$

In order to study the consequences of this inner product, we introduce the space $\tilde{V}$ of functions f, having the following properties. (Here the tilde denotes an incomplete space.) The functions f must have a continuous derivative, which approaches a (finite) limit as $w \to 0$, i.e.,

$$\lim_{w\to 0+} f'(w) \equiv f'_r(0)$$

exists. Then $f_r(0)$, similarly defined, also exists. Moreover, f must vanish at infinity at least as fast as $w^{-\epsilon}$, for some $\epsilon > 0$.

Let $f, g \in \tilde{V}$. If $f_r(0) = 0$ or $g_r(0) = 0$, then Eq. (2.3) reduces to

$$\langle f, g\rangle = +\int_0^\infty \frac{dw}{w} f^*(w)g(w). \quad (2.4)$$

Next let us review the axioms for a space Π_κ, where κ is a positive integer.[10] We modify slightly the conventions of Ref. 10.

(1) Π_κ is a linear space over the complex number field.

(2) There is given on Π_κ a form which is linear in the second argument, and which satisfies $\langle\xi, \eta\rangle = \langle\eta, \xi\rangle^*$ (i.e., a Hermitian form).

(3) The form is nondegenerate, i.e., $\langle\xi, \eta\rangle = 0$ for all $\eta \in \Pi_\kappa$ implies $\xi = 0$.

(4) Π_κ contains at least one subspace of dimension κ, on which the form is negative-definite.

(5) There is no subspace of dimension $\kappa + 1$ on which the form is negative-definite.

(6) Consider all the resolutions of Π_κ into an orthogonal direct sum, $\Pi_\kappa = N \dotplus P$, where the form is negative-definite on N, and positive-definite on P. Then then there is at least one such resolution

$$\Pi_\kappa = \Pi_- \dotplus \Pi_+, \quad (2.5)$$

for which Π_+ is a (complete) Hilbert space.

For the discussion of topological properties of Π_κ, we introduce a positive-definite metric. Let $f, g \in \Pi_\kappa$, and let

$$f = f_+ + f_-, \qquad g = g_+ + g_-, \quad (2.6a)$$

$$f_\pm, g_\pm \in \Pi_\pm. \quad (2.6b)$$

Then we define

$$[f, g] = \langle f_+, g_+\rangle - \langle f_-, g_-\rangle, \quad (2.7a)$$

$$||f|| = [f, f]^{\frac{1}{2}}. \quad (2.7b)$$

The following two theorems of Iohvidov and Krein are relevant (cf. Ref. 10, Theorems 1.1–1.4):

Theorem 1. In an arbitrary decomposition considered in Axiom 6, P is a (complete) Hilbert space, and defines a norm $||f||_P$ in analogy with Eq. (2.7b). The two topologies, defined by $||f||$ and $||f||_P$ respectively, are equivalent.

[12] R. Nevanlinna, Ann. Acad. Sci. Fenn., Ser. A, Nos. 108, 113, 115, 163, 222 (1952–56).

Theorem 2. An incomplete space $\tilde{\Pi}_\kappa$, satisfying Axioms 1–5, can always be completed, so as to yield a space Π_κ, satisfying Axioms 1–6.

Let us verify that $\tilde{V}$ is a $\tilde{\Pi}_1$-space. Axioms 1 and 2 are, of course, valid. To verify Axiom 3, consider a function $c(w)$, having a continuous derivative, and satisfying

$$c(0) = 0; \quad 0 < c(w) \leq 1 \quad \text{for} \quad 0 < w < \infty. \tag{2.8}$$

Then for any nonzero $f \in \tilde{V}$, in view of Eq. (2.4),

$$\langle f, cf \rangle = \int_0^\infty \frac{dw}{w} c(w) \, |f(w)|^2 > 0. \tag{2.9}$$

For Axioms 4 and 5, consider

$$\begin{aligned} f(w) &= 1 - w \quad \text{for} \quad 0 < w \leq 1 - \epsilon \\ &= 0 \qquad \text{for} \qquad w \geq 1, \end{aligned} \tag{2.10}$$

and let f be suitably rounded for $1 - \epsilon < w < 1$. Then $\langle f, f \rangle < 0$. Let f and g be linearly independent functions, such that $\langle f, f \rangle < 0$ and $\langle g, g \rangle < 0$. Then for h defined by

$$h(w) = f(0)g(w) - g(0)f(w), \tag{2.11}$$

we have $\langle h, h \rangle > 0$. Thus, Axioms 4 and 5 hold, with $\kappa = 1$.

Axiom 6 will hold if we complete $\tilde{V}$, in accordance with Theorem 2. Let us now describe the details of the completion. We introduce two spaces,

$$\tilde{V}_0 = \{g : g \in \tilde{V} \quad \text{and} \quad g(0) = 0\}, \tag{2.12a}$$

$$V_0 = \left\{g : \int_0^\infty \frac{dw}{w} \, |g(w)|^2 < \infty\right\}, \tag{2.12b}$$

so that V_0 is the completion of $\tilde{V}_0$ in the relevant L_2-norm. Furthermore, let us select one vector $v^- \in \tilde{V}$ having a negative norm. Its scalar multiples form a space V^-, and

$$\tilde{V} = V^- \dotplus_{(\text{no})} \tilde{V}_0,$$

where (no) indicates that the direct sum is non-orthogonal.

We next define

$$\tilde{V}^\perp = \{g : g \in \tilde{V}_0 \quad \text{and} \quad \langle g, v^- \rangle = 0\}. \tag{2.13}$$

We will denote by $V^\perp$ the completion of $\tilde{V}^\perp$ to a subspace of V_0. Let V^B be the orthocomplement of $V^\perp$ in V_0, and let P^B and $P^\perp$ be the respective projections:

$$V_0 = V^\perp \dotplus V^B, \qquad P^\perp V_0 = V^\perp, \quad \text{etc.} \tag{2.14}$$

The space V^B is nonempty, since, e.g., $cv^- \notin V^\perp$ [cf. (2.8)]. We assert that V^B is one dimensional.

Consider vectors $f_i \in \tilde{V}_0$, $j = 1, 2$; note the decompositions $f_i = f_i^\perp + f_i^B$, and suppose that

$$0 \neq \langle v^-, f_i \rangle \equiv \langle v^-, f_i^B \rangle. \tag{2.15}$$

We do not know whether the f_i^B are differentiable, so we use this equation to define $\langle v^-, f_i^B \rangle$. If the f^B were linearly independent, we could form

$$g = \langle f_1^B, v^- \rangle f_2^B - \langle f_2^B, v^- \rangle f_1^B \neq 0$$

and g would belong to $V^\perp$. Thus, dim $(P^B \tilde{V}_0) = 1$, and, since a projection is a continuous operator, dim $(V^B) = 1$.

The spaces V^B and V^- are not orthogonal, and therefore we define

$$V^A = \{h : h \in V^B \dotplus_{(\text{no})} V^- \quad \text{and} \quad \langle h, v^- \rangle = 0\}. \tag{2.16}$$

This space is also one dimensional. We can now represent the completion V, of $\tilde{V}$, as follows:

$$V = V^- \dotplus V^A \dotplus V^\perp = V^- \dotplus_{(\text{no})} V_0. \tag{2.17}$$

It is an open question whether the functions which span V^A and V^B are differentiable. This question could be decided if we had, e.g., an explicit expression for $P^B(cv^-)$, for some function c as in (2.8). In view of Theorem 1, the completion is independent of the choice of v^- originally made, and this fact also follows from Eq. (2.17).

The preceding discussion can be clarified and partially summarized with the help of the following diagram:

$$V \equiv \Pi_1 \left\{ \begin{matrix} \supset \tilde{V} \supset \mathcal{S}_+ \supset \\ \supset V_0 \equiv L_2 \supset \end{matrix} \right\} \mathcal{S}_{+0}, \tag{2.18}$$

where $\mathcal{S}_+$ consists of functions in $\mathcal{S}$, but restricted to $(0, \infty)$, and $\mathcal{S}_{+0} = \mathcal{S}_+ \cap \tilde{V}_0$.

To complete the discussion of the one-particle sector, we need to specify how the functions on the two rays interrelate. If we consider the two respective functions as independent, then the completed one-particle sector will be the sum of two orthogonal Π_1-spaces. One can also say then, that this sector will be a Π_2-space.

However, suppose that we were to start with, e.g., functions $\tilde{f}(p) \in \mathcal{S}$, use the inner product (1.3) without reference to the two terms, and otherwise proceed as before. Then we would obtain a Π_1-space as the completion. This approach is closer to the customary one, but we see no clear reason for rejecting the former possibility. Indeed, the field ϕ is only auxiliary[4], it does not relate to any hypothetical measurements, and its definition is not restricted by the usual principles.

3. FURTHER REMARKS

We should now like to elaborate on the foregoing in three respects.

First: An analysis similar to that of Sec. 2 can be carried out for the case of four dimensions, e.g., for the field B of Ref. 5. Then the one-particle sector is again a Π_1-space.

Second: One can also try to construct tensor product spaces, with possible applications to more than one particle. Let, e.g., $f_i, g_i \in V$, and

$$f = (f_1, f_2),\ g = (g_1, g_2) \in V \otimes V, \tag{3.1}$$

$$\langle f, g\rangle \equiv \langle f_1, g_1\rangle\langle f_2, g_2\rangle,\ [f, g] = [f_1, g_1][f_2, g_2]. \tag{3.2}$$

Now, let us consider a typical element of the tensor product space, $f = \sum (f_{1i}, f_{2i})$, and expand $\langle f, f\rangle$ and $[f, f]$ in orthonormal series suggested by the decomposition of Sec. 2. Then in general, the two series will differ with respect to the signs of an infinite number of terms. Consequently $[f, f]$ may diverge, while $\langle f, f\rangle$ might converge if the terms are suitably arranged. Hence the problem of completion is more delicate than for the case of the one particle sector.

Third: Let us return to the completed one-particle sector $\mathfrak{H}^{[1]}$ of ϕ, in order to give two simple examples of spectral decomposition. The elements of this sector can be taken as functions on the p^1-axis, where now [in contrast to (2.2)]

$$H(x) \leftrightarrow \phi(H)\ |0\rangle \leftrightarrow \tilde{H}(-|p^1|, p^1) \equiv h(-p^1). \tag{3.3}$$

We do not try to give a precise description of the set of functions $H(x)$ corresponding to the completed sector. We confine ourselves to a formal discussion, as follows.

Let P^1 and U_a be the familiar operators

$$P^1F(x) = i^{-1}\,\partial_1 F(x), \tag{3.4a}$$

$$U_aF(x) = F(x^0, x^1 - a). \tag{3.4b}$$

Their spectral representation[13] is of course given in terms of the correspondence (3.3):

$$(P^1f)(p^1) = p^1 f(p^1), \tag{3.5a}$$

$$(U_af)(p^1) = e^{-iap^1} f(p^1). \tag{3.5b}$$

It is immediate that

$$\langle f, P^1g\rangle = \langle P^1f, g\rangle, \tag{3.6a}$$

$$\langle U_af, U_ag\rangle = \langle f, g\rangle. \tag{3.6b}$$

Of course, in Eqs. (3.4a)–(3.6a), the functions F, f, g have to be suitably restricted. We can also define spectral projections,

$$(E_af)(p^1) = f(p^1)\theta(q - p^1). \tag{3.7}$$

If $q \neq 0$, then clearly $E_af \in \mathfrak{H}^{[1]}$. (For $q = 0$, see below.) Equations (3.6) and (3.7) now suggest the standard relations,

$$(E_qF)(x) = \frac{1}{2\pi}\int d^2p\, e^{ipx}\tilde{F}(p)\theta(q - p^1), \tag{3.8}$$

$$P^1F = \int_{-\infty}^{\infty} q\, dE_qF, \tag{3.9a}$$

$$U_aF = \int_{-\infty}^{\infty} e^{-iaq}\, dE_qF. \tag{3.9b}$$

We now recall[9,10] some differences between the spectral theory of operators in Π_κ-spaces and in Hilbert spaces. In general, a self-adjoint or a unitary operator in a Π_κ-space may have up to 2κ eigenvalues which lie off the real axis or off the unit circle, respectively. Moreover, such an operator always leaves invariant one or more subspaces, each having a nonpositive metric, and whose total dimension is between κ and 2κ. In particular, there is at least one (proper) eigenvector.

It appears that in our example there is only one possibility for an eigenvalue of P^1 or of U_a, namely 0 or 1, respectively; furthermore, it appears likely that the common eigenspace is one-dimensional, and that its nonzero vectors have negative norms. If this conjecture is valid, then we interpret E_0 as E_{0-}, as usual, and Eqs. (3.8)–(3.9) are not affected.

This conjecture, if valid, may lead to a decomposition of V, which would be more concise than (2.17). Let us denote the analogous eigenspace in V by V_s. We now make the further conjecture

$$V_s \perp V_0, \qquad V = V_s \dotplus V_0. \tag{3.10}$$

ACKNOWLEDGMENTS

We have benefitted from discussions with Dr. J. J. Loeffel, Dr. O. Steinmann, and Professor K. Symanzik. One of us (J. T.) would like to express his appreciation to Professor E. Bromberg and to Professor B. Zumino for their hospitality at the Courant Institute.

[13] N. Dunford and J. T. Schwartz, *Linear Operators, Part II* (Interscience Publishers, Inc., New York, 1964), Secs. X.5 and XII.3.

Point Transformations and the Hard-Sphere Bose Gas*

F. M. EGER† AND E. P. GROSS
Brandeis University, Department of Physics, Waltham, Massachusetts
(Received 13 July 1965)

The method of point transformations is used to provide a Hamiltonian for strongly interacting particles that is equivalent to the usual Hamiltonian. The new Hamiltonian is Hermitian, Fourier analyzable, velocity dependent, and free of the difficulties of pseudopotentials. The present work is restricted to the case that two-body encounters are dominant. The theory is expressed in the quantized wave formalism so that number-violating approximations can be applied. The well known results for the spectrum and ground-state energy of the dilute hard-sphere gas are then derived as an elementary application.

I. INTRODUCTION

IN several articles[1-3] we have explored the use of point transformations in many-body theory. For two particles i and j of equal mass m, the method consists of introducing coordinates $\mathbf{y}_i$ and $\mathbf{y}_j$ in place of the original coordinates $\mathbf{x}_i$ and $\mathbf{x}_j$. The new coordinates are defined by

$$\begin{aligned}\mathbf{y}_i &= \mathbf{x}_i + (\partial/\partial x_i)F(r_{ij}),\\ \mathbf{y}_j &= \mathbf{x}_j + (-\partial/\partial x_j)F(r_{ij}),\end{aligned} \qquad f(r) = (2/r)\,dF/dr\,.$$

If $H_{ij}(\mathbf{p}_i, \mathbf{p}_j, \mathbf{x}_i, \mathbf{x}_j)$ is the original Hamiltonian, the transformed Hamiltonian expressed in terms of the original coordinates is

$$H'_{ij} = \frac{p_i^2 + p_j^2}{2m} + (p_i - p_j)_\mu \frac{\mathfrak{L}^{\mu\nu}(r_{ij})}{4m}(p_i - p_j)_\nu + \frac{\hbar^2}{m}U(r_{ij}) + V(r_{ij}(1+f)). \tag{1.1}$$

Here $V(r)$ is the original direct potential between the particles. It is modified by the transformation. In particular, if V contains a part representing hard-core interactions between spheres of diameter c, the class of functions $f = (c/r)\psi(r)$ with $\psi(r \to 0) = 1$ and $\psi(r \to \infty) \to 0$, removes the hard-core interaction. The effects of the strong short-range interaction are reflected in the velocity-dependent interaction as measured by $\mathfrak{L}^{\mu\nu}(r)$ and in a metric potential $(\hbar^2/m)U(r)$. These functions are defined in terms of the transformation $f(r)$ by

$$\begin{aligned}\mathfrak{L}^{\mu\nu}(r) &= -\mathfrak{F}(r)\,\delta_{\mu,\nu} + (x_\mu x_\nu/r^2)\mathfrak{K}(r),\\ \mathfrak{F}(r) &= 1 - (1+f)^{-2},\\ \mathfrak{K}(r) &= [1 + (rf)']^{-2} - (1+f)^{-2},\end{aligned} \tag{1.2}$$

$$U(r) = \frac{\sqrt{B}}{r^2}\frac{d}{dr}\frac{r^2}{[1+(rf)']^2}\frac{d}{dr}\frac{1}{\sqrt{B}}, \qquad B = \frac{1}{(1+f)^2}\frac{1}{1+(rf)'}. \tag{1.3}$$

We have then a family of Hermitian Hamiltonians that are fully equivalent to the original Hamiltonians. The chief gain is that the new Hamiltonians are regular, i.e., the functions $\mathfrak{L}^{\mu\nu}(r)$, $U(r)$ and $V(r(1+f))$ are Fourier-analyzable. There is an important advantage in the many-body problem since it opens up the possibility of applying standard methods, including variational approaches, without the special prescriptions and rules that are inherent in non-Hermitian pseudopotential approaches. Recent extensions of this type of approach have been made by Luban.[4] A Hermitian Fourier-analyzable Hamiltonian in terms of the original coordinates has been derived by Lieb.[5]

Before considering the application of point transforms in the many-body problem we recall the salient features of the two-body problem. Suppose, for concreteness, that the spatial extent of $f(r)$ is characterized by a range λ and we write $f(r) = (c/r)\psi(r/\lambda)$. In the two-body problem, as λ gets much larger than c, the effective interaction is spread out in space and perturbative techniques such as the Born approximation yield results which can be expressed as a series in c/λ. For small values of this ratio the motion occurs under the influence of a smooth slowly varying velocity-dependent potential. The theory can, however, also be used for $c/\lambda \sim 1$, since there is a unitary equivalence of all the Hamiltonians. We still have the advantage of a Fourier-analyzable Hamiltonian, but simple techniques such as the Born approximation will not be quantitatively reliable.

* This work was supported by AFOSR grant 1176-63.
† Present address: Lawrence Radiation Laboratory, University of California, Berkeley, California.
[1] F. M. Eger and E. P. Gross, Nuovo Cimento 34, 1225 (1964).
[2] F. M. Eger and E. P. Gross, Ann. Phys. (N.Y.) 24, 63 (1963).
[3] F. M. Eger and E. P. Gross, J. Math. Phys. 6, 891 (1965).
[4] M. Luban, Phys. Rev. 138 A1028, A1033 (1965).
[5] E. Lieb, Proc. Natl. Acad. Soc. U. S. 46, 1000 (1960).

In the many-body problem we would like to take as the Hamiltonian

$$H' = \sum_{i<j} H'_{ij}. \tag{1.4}$$

However, this is only valid to the extent that two-body encounters are dominant, as is the situation for dilute gases. In that case it is possible to choose λ so that it is both large compared to c and small compared to the interparticle separation $\rho^{-\frac{1}{3}}$. Of course, the correct application of a point transformation in the many-body problem involves using an inherently many-body generator.[2] As discussed in our earlier work, one obtains a transformed Hamiltonian which contains multibody interactions. Hopefully, the dominant terms will be of the above two-body form with the parameter λ fixed as a function of the density from many-body analysis.

Our concern here is, however, only with those situations where the transformed two-body Hamiltonian is itself a good approximation. In Sec. II we express the theory in the language of second quantization. In Sec. III we show how the well known results of Huang, Lee, and Yang[6] for the low-density Bose gas emerge in an elementary and unambiguous way. The dilute gas limit is reached by letting λ be a fixed small fraction of the interparticle separation $\rho^{-\frac{1}{3}}$. Then c/λ tends towards zero and the point transformation ensures convergence. The Bogolyubov approximation involves a corrected two-body scattering length and one obtains the exact result for the low-density limit only if the term in c/λ vanishes. This is shown to be the case in the point transformation approach.

II. THE MANY-BODY HAMILTONIAN

We now set down the Hamiltonian $H' = \sum_{i<j} H'_{ij}$ in the quantized field form. This is important because it gives us the freedom to use number-violating approximations. It is permissible to Fourier analyze functions of the space coordinates. The transform is defined by

$$\begin{aligned} U(\mathbf{r}) &= \frac{1}{\Omega} \sum_{\mathbf{k}} \tilde{U}(\mathbf{k}) \exp(i\mathbf{k}\cdot\mathbf{x}), \\ \tilde{U}(\mathbf{k}) &= \int_{\Omega} \exp(-i\mathbf{k}\cdot\mathbf{x}) U(\mathbf{x})\, d^3x \end{aligned} \tag{2.1}$$

so that when we go to the conventional limit $\Omega \to \infty$, $N \to \infty$, $\rho = N/\Omega$ finite, the transform $\tilde{U}(k)$ remains finite. Then, choosing units where $\hbar = 1$, we have

$$\begin{aligned} H' = \sum_i \frac{p_i^2}{2m} &+ \frac{1}{2\Omega} \sum_{\mathbf{k}} \sum_{i<j} \frac{\tilde{\mathcal{L}}^{\mu\nu}(\mathbf{k})}{2m} (p_i - p_j)_\mu \\ &\times \exp[i\mathbf{k}(\mathbf{x}_i - \mathbf{x}_j)](p_i - p_j)_\nu \\ &+ \frac{2}{\Omega} \sum_{\mathbf{k}} \sum_{i<j} \frac{\tilde{U}(k)}{2m} \exp[i\mathbf{k}(\mathbf{x}_i - \mathbf{x}_j)] \\ &+ \sum_{\mathbf{k}} \sum_{i<j} \frac{\tilde{V}(k)}{\Omega} \exp[i\mathbf{k}\cdot(\mathbf{x}_i - \mathbf{x}_j)]. \end{aligned} \tag{2.2}$$

Here, $\tilde{V}(k)$ is the Fourier transform of the *point-transformed potential.*

We introduce the bilinear operators

$$\begin{aligned} G(\mathbf{p} \mid \mathbf{k}) &= \sum_i \delta(\mathbf{p}_i - \mathbf{p}) \\ &\times \exp(-i\mathbf{k}_i\mathbf{x}_i) \leftrightarrow \frac{\Omega}{(2\pi)^3} a^+(\mathbf{p}) a(\mathbf{p} + \mathbf{k}), \end{aligned} \tag{2.3}$$

$$[a(\mathbf{p}), a^+(\mathbf{p}')]_\pm = \delta_{\mathbf{p},\mathbf{p}'}, \qquad \delta(\mathbf{p} - \mathbf{p}') = \frac{\Omega}{(2\pi)^3} \delta_{\mathbf{p},\mathbf{p}'},$$

where the $a^+(\mathbf{k})$ and $a(\mathbf{k})$ are creation and annihilation operators of plane-wave state. The commutation relation

$$\begin{aligned} [G(\mathbf{p}' \mid \mathbf{k}'), G(\mathbf{p} \mid \mathbf{k})] &= G(\mathbf{p}' \mid \mathbf{k} + \mathbf{k}')\, \delta(\mathbf{p}' + \mathbf{k}' - \mathbf{p}) \\ &- G(\mathbf{p} \mid \mathbf{k} + \mathbf{k}')\, \delta(\mathbf{p} + \mathbf{k} - \mathbf{p}') \end{aligned} \tag{2.4}$$

is valid for both Bose and Fermi statistics. From the definitions

$$\begin{aligned} \int G(\mathbf{p} \mid \mathbf{k})\, d^3p &= \tilde{\rho}(\mathbf{k}) = \sum_i \exp(-i\mathbf{k}\cdot\mathbf{x}_i) \\ \int f(\mathbf{p}) G(\mathbf{p} \mid \mathbf{k})\, d^3p &= \sum_i f(p_i) \exp(-i\mathbf{k}\cdot\mathbf{x}_i) \\ G(\mathbf{p} \mid 0) &= [\Omega/(2\pi)^3] n_{\mathbf{p}}. \end{aligned} \tag{2.5}$$

With the aid of these equations we may express the Hamiltonian (2.2) entirely in terms of the $G(\mathbf{p}/\mathbf{k})$. The first and third terms are standard. The second term is reduced as follows:

$$\begin{aligned} \sum_{\mathbf{k}} \tilde{\mathcal{L}}^{\mu\nu}(\mathbf{k}) \sum_{i<j} (p_i - p_j)^\mu \exp[i\mathbf{k}(\mathbf{x}_i - \mathbf{x}_j)](p_i - p_j)^\nu \\ = \sum_{\mathbf{k}} \tilde{\mathcal{L}}^{\mu\nu}(k) \{ \sum_{i,j} p_i^\mu \exp[i\mathbf{k}(\mathbf{x}_i - \mathbf{x}_j)] p_i^\nu \\ - \sum_{i,j} p_i^\mu \exp[i\mathbf{k}(\mathbf{x}_i - \mathbf{x}_j)] p_j^\nu \} \end{aligned}$$

where we have used the fact that $\tilde{\mathcal{L}}^{\mu\nu}(-\mathbf{k}) = \tilde{\mathcal{L}}^{\mu\nu}(\mathbf{k})$. Using the commutator relation

$$\exp(-i\mathbf{k}\cdot\mathbf{x}_i)\mathbf{p}_j = (\mathbf{p}_j + \mathbf{k}\, \delta_{i,j}) \exp(-i\mathbf{k}\cdot\mathbf{x}_i),$$

[6] K. Huang, T. D. Lee, and C. W. Yang, Phys. Rev. **106**, 1135 (1957).

one can write

$$\sum_k \tilde{\mathfrak{L}}^{\mu\nu}(\mathbf{k}) \sum_{i,j} p_i^\mu \exp(i\mathbf{k}\cdot\mathbf{x}_i) \exp(-i\mathbf{k}\cdot\mathbf{x}_j) p_j^\nu$$
$$= \sum_k \tilde{\mathfrak{L}}^{\mu\nu}(\mathbf{k}) \int p^\mu (p^\nu - k^\nu) G(\mathbf{p} \mid -\mathbf{k})\, d^3p \times \int G(\mathbf{q} \mid \mathbf{k})\, d^3q$$

and

$$\sum_k \tilde{\mathfrak{L}}^{\mu\nu}(\mathbf{k}) \sum_{i,j} p_i^\mu \exp(+ik\mathbf{x}_i) \exp(-i\mathbf{k}\cdot\mathbf{x}_j) p_j^\nu$$
$$= +\sum_k \tilde{\mathfrak{L}}^{\mu\nu}(\mathbf{k}) \int p^\mu \int (q^\nu + k^\nu) \times G(\mathbf{p} \mid -\mathbf{k}) G(\mathbf{q} \mid \mathbf{k})\, d^3q\, d^3p.$$

The full Hamiltonian is

$$H' = \int \frac{p^2}{2m} G(\mathbf{p} \mid 0)\, d^3p + \frac{1}{2\Omega} \sum_k \iint \frac{\tilde{\mathfrak{L}}^{\mu\nu}(k)}{2m} p^\mu (p^\nu - q^\nu - 2k^\nu) \times G(\mathbf{p} \mid -\mathbf{k}) G(\mathbf{q} \mid \mathbf{k})\, d^3p\, d^3q + \frac{1}{2\Omega} \sum_k \left[\frac{\tilde{U}(k)}{m} + \tilde{V}(k)\right] \times \left[\int G(\mathbf{p} \mid -\mathbf{k}) \int G(\mathbf{q} \mid \mathbf{k})\, d^3p\, d^3q - N\right]. \quad (2.6)$$

The transition to creation and annihilation operators for both types of statistics is made using

$$G(\mathbf{p} \mid -\mathbf{k}) G(\mathbf{q} \mid \mathbf{k}) = \frac{\Omega^2}{(2\pi)^6} [a^+(\mathbf{p}) a^+(\mathbf{q}) a(\mathbf{q} + \mathbf{k}) a(\mathbf{p} - \mathbf{k}) + a^+(\mathbf{p}) a^+(\mathbf{p})\, \delta(\mathbf{p} - \mathbf{k} - \mathbf{q})]. \quad (2.7)$$

We have, finally,

$$H' = \sum_p \frac{p^2}{2m} a^+(\mathbf{p}) a(\mathbf{p}) + \frac{1}{2\Omega} \sum_{p,q,k} M(\mathbf{p}, \mathbf{q}, \mathbf{k}) a^+(\mathbf{p}) a^+(\mathbf{q}) a(\mathbf{p} - \mathbf{k}) a(\mathbf{q} + \mathbf{k}), \quad (2.8)$$

with

$$M(\mathbf{p}, \mathbf{q}, \mathbf{k}) = \tfrac{1}{2} m^{-1} \times \{\tilde{\mathfrak{L}}^{\mu\nu}(\mathbf{k}) p^\mu [p^\nu - q^\nu - 2k^\nu] + 2\tilde{U}(\mathbf{k})\} + \tilde{V}(\mathbf{k}). \quad (2.9)$$

This is the usual form of the Hamiltonian with two-body interaction, the velocity dependences of the interaction finding expression in the dependence of M on $\mathbf{p}$ and $\mathbf{q}$ as well as $\mathbf{k}$. In the following section we consider only the hard-sphere gas. With the appropriate class of point transformations the term $\tilde{V}(k)$ then vanishes.

III. LOW DENSITY LIMIT

We now apply the elementary Bogolyubov method to our Hamiltonian to exhibit the low-density results of Huang, Lee, and Yang. In that limit one can neglect depletion and anharmonic effects, and assume that the zero-momentum state has an occupancy N. Replacing a_0 and a_0^+ by $N^{\frac{1}{2}}$ and going to units in which $2m = 1$, one picks out the reduced Hamiltonian

$$H_{eff} = N\rho\tilde{U}(0) + \sum \tilde{\epsilon}(k) a^+(\mathbf{k}) a(\mathbf{k}) + \sum_k \rho\tilde{U}(\mathbf{k}) [a^+(\mathbf{k}) a^+(-\mathbf{k}) + \text{H. c.}], \quad (3.1)$$

where

$$\tilde{\epsilon}(k) = k^2 + \rho[2\tilde{U}(k) + A - B], \quad (3.2)$$

$$B(k) - A(k) = \tfrac{1}{2} k^\mu k^\nu [\tilde{\mathfrak{L}}^{\mu\nu}(\mathbf{k}) - \tilde{\mathfrak{L}}^{\mu\nu}(0)]. \quad (3.3)$$

Use of the Bogolyubov linear transformation yields the excitation spectrum

$$\epsilon(k) = (\tilde{\epsilon}^2 - 4\rho^2\tilde{U}^2)^{\frac{1}{2}}, \quad (3.4)$$

and the ground-state energy

$$E_0 = N\rho\tilde{U}(0) + \tfrac{1}{2} \sum_k [\epsilon(k) - \tilde{\epsilon}(k)] = N\rho\tilde{U}(0) - 2 \sum_k \frac{\rho^2\tilde{U}^2}{\epsilon + \tilde{\epsilon}(k)}. \quad (3.5)$$

In our case, the convergence of the sum is automatic, arising from the properties of the Fourier transform of the metric potential $\tilde{U}(k)$. The Bogolyubov transform can, of course, be applied as a variational estimate of the ground-state energy. Within the domain of validity of our transformed Hamiltonian, this is perfectly legitimate provided one calculates depletion and anharmonic effects in the standard way using the whole Hamiltonian. It should be stressed that even the assumption of complete occupancy of the $\mathbf{k} = 0$ state for the transformed Hamiltonian may imply considerable depletion for the zero-momentum state that is conventionally discussed, since the wavefunctions of H' are connected to those of the original Hamiltonian by the Jacobian of the point transform. This effect is, however, unimportant in the dilute-gas limit.

We now proceed to extract the low-density results from the theory.

A. Spectrum

We pass to the limit of a dilute gas with λ a fixed small multiple of the interparticle separation $\rho^{-\frac{1}{3}}$. The transition from phonon to particle behavior in the spectrum occurs at $k^2 \sim 8\pi c\rho$. In this region $(k\lambda) \sim (8\pi c)^{\frac{1}{2}}\rho^{\frac{1}{6}}$ so that $k\lambda \ll 1$. We show, however, in the Appendix that for $k\lambda < 1$, the term $A - B$ is of order $(k\lambda)^4 c$, while the leading term in $\tilde{U}$ is of course $4\pi c$. Thus the standard form of the spectrum holds until k approaches the fixed fraction of $\rho^{\frac{1}{3}}$. The dilute gas spectrum thus includes the particle as well as phonon region of the excitation spectrum. These are the standard conditions for the validity of the form (3.4) of the spectrum in the low-density limit.

B. Ground-State Energy

Here the argument parallels that of Brueckner[7] or Landau and Lifshitz.[8] If one neglects the effect of the medium, i.e., puts $\rho = 0$ in the denominator, one finds

$$E_0 \to N\rho\tilde{U}(0) - \rho^2 \sum_k \frac{\tilde{U}^2}{k^2}.$$

This is just the energy due to $\frac{1}{2}[N(N-1)]$ hard-sphere interactions in a volume Ω. The second term is a correction to the scattering length $\tilde{U}(0)$. Now for general $\varphi(r/\lambda)$, $\tilde{U}(0)$ has the form[1]

$$\tilde{U}(0) = 4\pi c[1 + \alpha(c/\lambda) + \beta(c/\lambda)^2 \log c/\lambda + \cdots].$$

There are particular forms of $\varphi(r/\lambda)$ such that the coefficient α is zero, but this is not true in general. The correction c/λ however spoils the transition to the dilute gas limit, for it leads to an energy proportional to $\rho^{4/3}$ which is larger than the medium correction $\rho^{\frac{1}{2}}$ as $\rho \to 0$. Thus it is essential that the correction to the scattering length cancel the term in c/λ. The remaining error in the scattering length does not hinder the passage to the low-density limit. The cancellation is demonstrated in the Appendix.

We have then, including terms through $c \cdot c/\lambda$,

$$E_0 = N\rho 4\pi c - 2\sum_k \rho^2\tilde{U}^2\left[\frac{1}{\epsilon(k) + \tilde{\epsilon}(k)} - \frac{1}{2k^2}\right].$$

Now, however, the sum over k converges even with $\tilde{U}(k) = 4\pi c$, and the corrections due to $B - A$ as well as variations in $U(k)$ can be neglected in the low-density limit. Hence we find the well-known result in standard units

$$E_0 = N(\rho\hbar^2/m)2\pi c[1 + (128/15)(\rho c^3/\pi)^{\frac{1}{2}}].$$

CONCLUSION

It should be stressed that the two-body Hamiltonian equation (2.2) has a wider range of validity than the particular application treated here. It, of course, also applies to the Fermi case. In addition, by taking λ to be of the order of c, the Hamiltonian is applicable to systems with finite densities, perhaps within a factor one-fourth of physical densities of condensed systems. This allows one to study the effects of attractive interactions under conditions where they may play a qualitatively decisive role. Of particular interest is the weakening of the attractive potential which is directly exhibited by the transformed interaction $V(r(1 + f))$. The study of these questions, however, requires using theories which go beyond the elementary Bogolyubov analysis as applied to the transformed Hamiltonian.

APPENDIX: PROPERTIES OF THE MATRIX ELEMENTS

The Momentum-Dependent Interaction

We have $B - A = k^\mu k^\nu[\tilde{\mathfrak{U}}^{\mu\nu}(k) - \tilde{\mathfrak{U}}^{\mu\nu}(0)]$,

$$\tilde{\mathfrak{U}}^{\mu\nu}(k) = \lambda^3 \int \exp(-i\mathbf{k}\boldsymbol{\xi}\lambda)\left[-F\left(\frac{c}{\lambda}, \xi\right)\delta_{\mu,\nu} + \frac{\xi^\mu\xi^\nu}{\xi^2} H\left(\frac{c}{\lambda}, \xi\right)\right] d^3\xi,$$

$$F = 1 - \left(1 + \frac{c}{\lambda}\frac{\varphi}{\xi}\right)^{-2},$$

$$H = \left[1 + \frac{c}{\lambda}\varphi'\right]^{-2} - \left[1 + \frac{c}{\lambda}\frac{\varphi}{\xi}\right]^{-2}.$$

For $k\lambda \ll 1$, one can expand the exponential and

$$k^\mu k^\nu[\tilde{\mathfrak{U}}^{\mu\nu}(k) - \tilde{\mathfrak{U}}^{\mu\nu}(0)] = \lambda(k\lambda)^4 \int \left[-\tfrac{1}{3}F\left(\frac{c}{\lambda}, \xi\right) + \tfrac{1}{5}H\left(\frac{c}{\lambda}, \xi\right)\right]\xi^2 d^3\xi.$$

The leading term of $\int F[(c/\lambda), \xi^2]\xi^2 d^3\xi$ is proportional to c/λ so that the entire expression goes as $(k\lambda)^4 c$.

The Scattering Length

We wish to show that the two contributions to the scattering length, when combined, are free of a term $c \cdot c/\lambda$. In evaluation the term

$$\rho \sum \frac{\tilde{U}^2(k)}{k^2} = \frac{N\rho}{4\pi} \iint U(\mathbf{x}) \frac{1}{|x - y|} U(\mathbf{y})\, d^3x\, d^3y,$$

[7] K. Brueckner, *The Many Body Problem* (John Wiley & Sons, Inc., New York, 1959).

[8] L. D. Landau and E. M. Lifshitz, *Statistical Physics* (Addison Wesley Company, Inc., Reading, Massachusetts, 1958), p. 240.

it suffices to insert $U(x)$ to lowest order. This is

$$U(\mathbf{x}) = (c/\lambda^3)\ \nabla_\xi^2[(\varphi/\xi) + \tfrac{1}{2}\varphi'] \equiv (c/\lambda^3)\ \nabla^2 P(\xi).$$

Hence the contribution to the scattering length is

$$\frac{N\rho}{4\pi}\frac{c^2}{\lambda}\iint \nabla^2 P(\xi)\,\frac{1}{|\xi-\eta|}\,\nabla^2 P(\eta)\,d^3\xi\,d^3\eta$$
$$= -N\rho\left(c\cdot\frac{c}{\lambda}\right)\int P(\xi)\,\nabla^2 P(\xi)\,d^3\xi.$$

We consider next $\tilde{U}(0) = \int U(\mathbf{x})\,d^3x$,

$$\tilde{U}(0) = \lambda\int B^{\frac{1}{2}}\left[\frac{1}{\xi^2}\frac{d}{d\xi}\,\frac{\xi^2}{\left(1+\frac{c}{\lambda}\varphi'\right)^2}\,\frac{d}{d\xi}\frac{1}{B^{\frac{1}{2}}}\right]d^3\xi.$$

The term $[1 + (c/\lambda)\varphi']$ can be replaced by unity to order $c(c/\lambda)$. Then with $B^{\frac{1}{2}} \sim 1 + (c/\lambda)P(\xi)$, we have

$$\tilde{U}(0) = \lambda\int \sqrt{B}\,\nabla_\xi^2\frac{1}{B^{\frac{1}{2}}}\,d^3\xi$$
$$= 4\pi c - c\cdot\frac{c}{\lambda}\int P(\xi)\,\nabla^2 P(\xi)\,d^3\xi.$$

This establishes the cancellation of the terms in $c(c/\lambda)$, so that the scattering length is c to the order needed in the dilute-gas limit. These considerations deal cavalierly with the region $r < c/\lambda$. However, it has been shown[1] that the deviations give contributions of order $(c/\lambda)^2 \log (c/\lambda)$.

Functional Formulation of the Lee Model*

H. M. FRIED
Physics Department, Brown University, Providence, Rhode Island
(Received 26 April 1965)

The Lee model is defined as a particular member of a class of approximations to a relativistic, local, nontrivial field theory. Functional methods are applied to obtain generalizations of familiar properties, and to suggest variants of the model.

I. INTRODUCTION

THE Lee model has been extensively discussed[1] from the viewpoint of formulating a theory of a model field, the latter nonrelativistic and nonlocal. One can, however, describe the Lee model, together with appropriate generalizations, in terms of a model field theory; here the basic structure is that of a local, relativistic, nontrivial field theory, and the model is obtained by performing certain simple approximations—more properly, mutilations—on the bare V, N, and ϑ propagators which enter into the coupled Green's function equations. This latter strategy seems quite obvious and natural, but it does not seem to have been employed, at least in this context, in the voluminous literature on the subject. It is a rather useful way of defining the model, since compact functional techniques may then be brought to bear, and various generalizations suggested and explored.

The purpose of this brief note is to indicate how the Lee model may be treated from this point of view. No essentially new results are obtained, only a clear delineation of the class of such model approximations which may be termed the Lee model. The usual quantities are easily computed within this framework, and a general equation analogous to the state vector equation for one V and $n\vartheta$-particles, is displayed in functional terms. In particular, the corresponding equation obeyed by the $V\vartheta$ scattering amplitude is shown to be equivalent to one previously discussed and solved by Amado.[2]

Related functional methods are also applicable to the simple static model, generating that familiar[3] and completely soluble theory. Some generalizations of the ordinary Lee model are briefly noted here; in particular, that the limiting situation in which the N-particle mass becomes very large is completely soluble.

The objects of interest here are the set of time-ordered n-point Green's functions, off the mass shell, and these are most conveniently obtained from a generating functional. In any local field theory, in which the operators $A(x)\cdots$ are coupled with each other, one may define the generating functional of these Green's functions,

$$\mathfrak{z}\{j, \cdots\} = \left\langle \left(\exp\left\{ i \int [jA + \cdots] \right\} \right)_+ \right\rangle, \quad (1)$$

in terms of functional differentiation operations upon the free-field generating functional

$$\mathfrak{z}\{j, \cdots\} = \frac{1}{N_V} \exp\left[i \int dx\, \mathfrak{L}'\left\{ \frac{1}{i} \frac{\delta}{\delta j(x)}, \cdots \right\} \right] \times \mathfrak{z}^{(0)}\{j, \cdots\}. \quad (2)$$

Here, (1) represents the definition of $\mathfrak{z}$, in terms of the vacuum expectation value of time-ordered products of (here unrenormalized) operators $A(x)\cdots$ and associated c-number sources $j(x)\cdots$, while (2) gives the corresponding functional solution; in the latter, $\mathfrak{L}'\{A(x), \cdots\}$ denotes the interaction Lagrangian density, and N_V is a normalization constant so chosen that $\mathfrak{z}\{0, \cdots\} = 1$. For local theories, this solution follows directly from the action principle[4]; or, alternatively, it requires only the input information of field equations plus equal-time commutation relations.[5]

* Work supported by the U. S. Atomic Energy Commission.

[1] T. D. Lee, Phys. Rev. **95**, 1392 (1954) and other authors quoted in S. S. Schweber, *An Introduction to Relativistic Quantum Theory* (Row, Peterson and Company, New York, 1961). A calculation using the *LSZ* formalism has recently been given by M. Maxon and R. Curtis, Phys. Rev. **137**, B996 (1965); these authors claim to elucidate "the basic structure of the model in the most natural way . . .". It is believed that the functional methods employed in this paper provide a more general way.

[2] R. D. Amado, Phys. Rev. **122**, 696 (1961) and **132**, 485 (1963); R. P. Kenschaft and R. D. Amado, J. Math. Phys. **5**, 1340 (1964). Equation (36) of the present paper should be compared with Eq. (5) of the last-named reference.

[3] See, for example, S. S. Schweber, Ref. 1.

[4] J. Schwinger, "Summer Lectures on Quantum Field Theory," Stanford University, 1956 (unpublished).

[5] K. Symanzik, Z. Naturforsch. **9**, 809 (1954).

II. FUNCTIONAL SOLUTION AND MODEL APPROXIMATIONS

As originally formulated, the Lee model is nonlocal, and hence these functional considerations do not immediately apply. The method adopted here is to consider the completely relativistic theory involving the interactions of a pair of spin-zero boson fields ϕ_1, ϕ_2 written in the charge combinations ϕ, $\phi^\dagger$ according to the interaction

$$\mathcal{L}' = -g(\bar{\psi}_V \phi \psi_N + \bar{\psi}_N \phi^\dagger \psi_V), \tag{3}$$

where $\psi_{N,V}$ and $\bar{\psi}_{N,V}$ denote the N and V fermion fields. In this realistic theory, the quanta of these free fields are to appear in conjunction with the corresponding antiparticles; the quanta of the free ϕ and $\phi^\dagger$ fields are to have the same mass, M, corresponding to that of the ϑ (and $\bar{\vartheta}$)-particles. The free-field generating functional can then be written down immediately,

$$\mathfrak{z}^{(0)} = \exp\{i\bar{\eta}_V S_c^{(V)} \eta_V + i\bar{\eta}_N S_c^{(N)} \eta_N + ij^\dagger \Delta_c^{(\vartheta)} j\}, \tag{4}$$

where j and $j^\dagger$ are the sources of $\phi^\dagger$ and ϕ, while similar remarks hold for the anticommuting, c-number spinor sources η_N, η_V, $\bar{\eta}_N$, $\bar{\eta}_V$. The time-ordered Green's functions $S_c^{(N,V)}$ and $\Delta_c^{(\vartheta)}$ represent the relativistic, causal (Feynman) propagators of the N, V, and ϑ particles.

In accordance with (2), the generating functional is here given by

$$\mathfrak{z} = \frac{1}{N_V} \exp\left(-g\left\{\frac{\delta}{\delta\eta_V}\frac{\delta}{\delta j^\dagger}\frac{\delta}{\delta\bar{\eta}_N} + \frac{\delta}{\delta\eta_N}\frac{\delta}{\delta j}\frac{\delta}{\delta\bar{\eta}_V}\right\}\right)\mathfrak{z}^{(0)}, \tag{5}$$

and the functional translation operations of (5) may be performed in several equivalent ways. We choose to translate the η_N, $\bar{\eta}_N$ dependence,

$$\mathfrak{z} = \frac{1}{N_V} \exp\left\{i\left[\bar{\eta}_N - g\frac{\delta}{\delta\eta_V}\frac{\delta}{\delta j^\dagger}\right] \times S_c^{(N)}\left[\eta_N + g\frac{\delta}{\delta j}\frac{\delta}{\delta\bar{\eta}_V}\right]\right\} \times \exp(i\bar{\eta}_V S_c^{(V)}\eta_V + ij^\dagger\Delta_c^{(\vartheta)}j), \tag{6}$$

since, for simplicity, only those amplitudes having no external N-particle coordinates will be considered. Hence

$$\mathfrak{z}|_{\eta_N=\bar{\eta}_N=0} = \frac{1}{N_V} \exp\left(-ig^2\frac{\delta}{\delta\eta_V}\frac{\delta}{\delta j^\dagger} S_c^{(N)} \frac{\delta}{\delta j}\frac{\delta}{\delta\bar{\eta}_V}\right) \times \exp(i\bar{\eta}_V S_c^{(V)}\eta_V + ij^\dagger\Delta_c^{(\vartheta)}j). \tag{7}$$

Carrying through the η_V and $\bar{\eta}_V$ displacements,[6] this becomes

$$\mathfrak{z}|_{\eta_N=\bar{\eta}_N=0} = \frac{1}{N_V} \exp\left[i\bar{\eta}_V S_c^{(V)}\left(1 + g^2\frac{\delta}{\delta j^\dagger} S_c^{(N)}\frac{\delta}{\delta j} S_c^{(V)}\right)^{-1}\eta_V\right] \times \exp\left\{\mathrm{Tr}\ln\left(1 + g^2\frac{\delta}{\delta j^\dagger} S_c^{(N)}\frac{\delta}{\delta j} S_c^{(V)}\right)\right\} \times \exp(ij^\dagger\Delta_c^{(\vartheta)}j), \tag{8}$$

and, so far, all is exact.

We now make that approximation to this nontrivial field theory which may be termed the Lee model: the propagators $S_c^{(N,V)}$ and $\Delta_c^{(\vartheta)}$, in (8), are replaced by nonrelativistic propagators $S_{V,N}$ and Δ_ϑ, where each of the latter is, in configuration space, a retarded function. This essential point can be achieved in a variety of ways, any one of which may be described as resulting from the absence of antiparticles and the corresponding omission of a negative energy pole in the momentum representation of each function. The result of this approximation is dynamically (graph for graph) the same as the ordinary Lee model; this equivalence follows because the reduction to nonrelativistic ϑs means that the N, V symmetry of (3) is broken. Only a V can now emit a ϑ, and only an N can absorb one. The restriction to nonrelativistic N and V removes all processes involving their antiparticles; and all that remains is the desired model dynamics.

III. CALCULATIONS

Certain familiar computations can now be set up in functional terms, and we consider some of these in the order of increasing difficulty.

A. Vacuum Processes

The simplest example of solubility arises in the evaluation of the constant N_V, describing all vacuum-to-vacuum processes,

$$N_V = \exp\left[\mathrm{Tr}\ln\left(1 + g^2\frac{\delta}{\delta j^\dagger} S_N \frac{\delta}{\delta j} S_V\right)\right] \times \exp(ij^\dagger\Delta_\vartheta j)|_{j=j^\dagger=0} = 1, \tag{9}$$

which result follows from the retarded nature of the propagators combined with the trace operation of (9). For the same reason, such closed-loop quantities never appear in any n-point function.

[6] B. Zumino, "Lecture Notes on the Quantum Theory of Fields," New York University, 1958 (unpublished). See also C. Sommerfield, Ann Phys. (N.Y.) **26**, 1 (1963).

B. V-particle Propagator

The Green's function corresponding to the dressed V propagator is, similarly, easy to evaluate,

$$S_V'(x-y) = \langle x|\, S_V\Big(1 + g^2 \frac{\delta}{\delta j^\dagger} S_N \frac{\delta}{\delta j} S_V\Big)^{-1} |y\rangle \times \exp\,(ij^\dagger \Delta_\vartheta j)|_{j=j^\dagger=0}, \quad (10)$$

using an obvious notation. If the perturbation expansion of (10) is calculated, to any order, one finds that retardation requires that the combination

$$-g^2\langle z|\, \frac{\delta}{\delta j^\dagger} S_N \frac{\delta}{\delta j}\, |w\rangle,$$

in the expansion of (10), is always replaced by

$$-ig^2 S_N(z-w)\, \Delta_\vartheta(z-w),$$

which quantity we designate by $\Sigma\,(z-w) \equiv \langle z|\Sigma|w\rangle$. Thus, the sum of all such terms of (10) gives

$$S_V'(x-y) = \langle x|\, S_V(1 - \Sigma\, S_V)^{-1}\, |y\rangle, \quad (11)$$

which is, in momentum space, just the daisy chain solution corresponding to the iteration of the simplest V-particle self-energy graph. Lee's original result is reproduced (except for the cutoff factor, which can always, and later will, be inserted by hand) if one chooses the nonrelativistic fermion propagators corresponding to fixed V and N particles,

$$S_{V,N}(x) = S_{V,N}(\mathbf{r}, x_0) = i\delta(\mathbf{r})\theta(x_0)e^{-ix_0 m_{V,N}}, \quad (12)$$

and a retarded ϑ-propagator obtained by discarding the contribution from the negative energy pole in the Fourier representation of $\Delta_c^{(\vartheta)}$,

$$\Delta_\vartheta(\mathbf{r}, x_0) = i(2\pi)^{-3} \int \frac{d^3k}{2\omega} \exp\,(i\mathbf{k}\cdot\mathbf{r} - i\omega x_0)\vartheta(x_0), \qquad \omega^2 = \mathbf{k}^2 + M^2. \quad (13)$$

It is clear that a wide variety of such models, each differing only in the form of the retarded functions and each equally soluble, can be defined in the same way.

For the special choices (12), (13) one obtains the very well known solution,[1] briefly noted here for completeness, for the V propagator. It is not difficult to see that the spatial δ-function of (12) persists in the dressed $S_V'(\mathbf{r}, x_0)$, multiplying a function of x_0 only; hereafter, when the symbols $S_{N,V}(x)$ and $S_V'(x)$ are used, they represent this latter dependence only. Similarly, the quantity $\Delta_\vartheta(x)$ will refer to the x_0 dependence of $\Delta_\vartheta(\mathbf{0}, x_0)$ of (13). One easily finds that the Fourier transform of the free V propagator,

$$\tilde{S}_V(q) = \int dx\, e^{iqx} S_V(x) = [m_V - i\epsilon - q]^{-1}, \quad (14)$$

is replaced by the dressed function

$$\tilde{S}_V'(q) = [m_V - i\epsilon - q - \tilde{\Sigma}(q)]^{-1}, \quad (15)$$

from which one obtains the renormalized V propagator,

$$\begin{aligned}\tilde{S}_V^R(q) &= Z^{-1}\tilde{S}_V'(q) \\ &= [m_V^R - i\epsilon - q]^{-1} \cdot [1 - (m_V^R - i\epsilon - q)\beta(q)]^{-1}, \\ \beta(q) &= \frac{g_R^2}{(2\pi)^3} \int \frac{d^3k}{2\omega} |f(\omega)|\, (\omega + m_N - m_V^R)^{-2} \times (\omega' + m_N - i\epsilon - q)^{-1},\end{aligned} \quad (16)$$

defined in terms of the renormalized charge and mass and the renormalization constant. In (16), the cutoff factor $|f(\omega)|$ introduced into the Fourier representation of $\Delta_\vartheta(x)$ is displayed, although it is not necessary in this equation. The definition of m_V^R and Z are Lee's, and need not be reproduced here. It is also easy to see, from (6) and (8), that S_N' and Δ_ϑ' are given by their respective free-field values, when the replacement of causal propagators by retarded functions is made.

C. $V\vartheta$ Scattering

Higher n-point functions are not so easily constructed; the reason is that too many (more than two!) external coordinates will appear, and the simple retardation arguments cannot be applied. For example, elastic $V\vartheta$ scattering is governed by the four-point function

$$G(xy\,|\,z_1z_2) = -i\langle x|\, S_V\Big(1 + g^2 \frac{\delta}{\delta j^\dagger} S_N \frac{\delta}{\delta g} S_V\Big)^{-1} |y\rangle \times \frac{\delta}{\delta j^\dagger(z_1)} \frac{\delta}{\delta j(z_2)} e^{ij^\dagger \Delta_\vartheta j}\Big|_0, \quad (17)$$

and the extra indices $z_{1,2}$, corresponding to the emission and absorption of a ϑ-particle, make application of the previous arguments far too complicated to carry through. Nevertheless, the retardation properties do permit one to write an equation which couples this scattering amplitude to itself, in terms of S_N, Δ_ϑ, and S_V'; this property of the model might be termed recursive, and is a feature of all the higher $V + n\vartheta$ amplitudes.

It is somewhat simpler to rewrite (17) in the equivalent form

$$G(xy \mid z_1z_2) = i \exp\left(-i \frac{\delta}{\delta j^\dagger} \Delta_\vartheta \frac{\delta}{\delta j}\right) j^\dagger(z_1)j(z_2)$$
$$\times \langle x| S_V(1 - g^2 j^\dagger S_N j S_V)^{-1} |y\rangle|_0, \quad (18)$$

and for later convenience, we adopt the notation

$$G_V^\varrho(x, y) = \langle x| S_V(1 - g^2 j^\dagger S_N j S_V)^{-1} |y\rangle, \quad (19)$$

$$S_V^\varrho(x, y) = \exp\left(-i \frac{\delta}{\delta j^\dagger} \Delta_\vartheta \frac{\delta}{\delta j}\right) G_V^\varrho(x, y), \quad (20)$$

and then, $S_V'(x - y) = S_V^\varrho(x, y)\,|_0$. The integral equation satisfied by G_V^ϱ, which will be useful shortly,

$$G_V^\varrho(x, y) = S_V(x - y)$$
$$+ g^2 \int S_V(x - u) j^\dagger(u) S_N(u - v) j(v) G_V^\varrho(v, y), \quad (21)$$

indicates that G_V^ϱ is also a retarded function. Passing the exponential functional differentiation operator of (18) through the $j^\dagger(z_1)j(z_2)$ factors, one obtains

$$G(xy \mid z_1z_2) = \Delta_\vartheta(z_1 - z_2) S_V'(x - y)$$
$$- i \int \Delta_\vartheta(z_1 - u)\Delta_\vartheta(v - z_2) \frac{\delta}{\delta j(u)} \frac{\delta}{\delta j^\dagger(v)} S_V^\varrho(x, y)\Big|_0, \quad (22)$$

from which it is clear that the connected part of the configuration space scattering amplitude, amputated on the ϑ-coordinates, is given by

$$T(xy \mid \bar{z}_1\bar{z}_2) = -i \frac{\delta}{\delta j(z_1)} \frac{\delta}{\delta j^\dagger(z_2)} S_V^\varrho(x, y)\Big|_0. \quad (23)$$

Equation (23) would be exact if the proper causal functions had been used; with the approximation of substituting retarded functions, we can illustrate the recursive nature of the model solutions for (23), and for more general amplitudes, by studying the properties of S_V^ϱ.

Equations (20) and (21) show that S_V^ϱ can be written in the form

$$S_V^\varrho(x, y) = S_V(x - y) + g^2 \exp\left(-i \frac{\delta}{\delta j^\dagger} \Delta_\vartheta \frac{\delta}{\delta j}\right)$$
$$\times \int S_V(x - u) j^\dagger(u)$$
$$\times S_N(u - v) j(v) G_V^\varrho(v, y)$$
$$= S_V(x - y) + g^2 \int S_V(x - u)\Big[j^\dagger(u)$$
$$- i \int \Delta_\vartheta(u - w) \frac{\delta}{\delta j(w)}\Big] S_N(u - v)$$
$$\times \exp\left(-i \frac{\delta}{\delta j^\dagger} \Delta_\vartheta \frac{\delta}{\delta j}\right) j(v) G_V^\varrho(v, y). \quad (24)$$

Because of the retarded nature of G_V^ϱ and Δ_ϑ, $j(v)$ cannot factor pair with any $j^\dagger$ in $G_V^\varrho(v, y)$, and hence the combination

$$\exp\left(-i \frac{\delta}{\delta j^\dagger} \Delta_\vartheta \frac{\delta}{\delta j}\right) j(v) G_V^\varrho(v, y) \quad (25)$$

of (24) may be replaced by $j(v)S_V^\varrho(v, y)$. This step is indicative of the model nature of all $V + n\vartheta$ amplitudes: there is no term in (25) of form

$$-i \int \frac{\delta}{\delta j^\dagger(\xi)} \Delta_\vartheta(\xi - v) S_V^\varrho(v, y).$$

Such a term would couple any n-point function to an $(n + 2)$-point function, as in the causal theory, and the absence of this coupling makes the model recursive. Equation (24) may then be written in the form

$$S_V^\varrho(x, y) = S_V(x - y)$$
$$+ g^2 \int S_V(x - u) j^\dagger(u) S_N(u - v) j(v) S_V^\varrho(v, y)$$
$$+ \int S_V(x - u)\, \Sigma\, (u - v) S_V^\varrho(v, y)$$
$$- ig^2 \int S_V(x - u)\Delta_\vartheta(u - w)$$
$$\times S_N(u - v) j(v) \frac{\delta}{\delta j(w)} S_V^\varrho(v, y), \quad (26)$$

from which the defining integral equation for S_V' can be obtained by setting $j = j^\dagger = 0$. Here, we are interested in the coefficient of $j(z_1)j^\dagger(z_2)$ of the corresponding expansion of S_V^ϱ; that is, if

$$S_V^\varrho(x, y) = S_V'(x - y)$$
$$+ \int j^\dagger(u) j(v) S(xy \mid vu) + \cdots, \quad (27)$$

we can simply expand both sides of (26) and obtain

$$S(xy \mid z_1z_2) = g^2 S_V(x - z_1) S_N(z_1 - z_2) S_V'(z_2 - y)$$
$$+ \int S_V(x - u)\, \Sigma\, (u - v) S(vy \mid z_1z_2)$$
$$- ig^2 \int S_V(x - u)\Delta_\vartheta(u - w)$$
$$\times S_N(u - z_2) S(z_2y \mid z_1w). \quad (28)$$

The coefficients of higher powers of $(j^\dagger j)$ yield recursive equations for the higher $V + n\vartheta$ amplitudes, with the latter coupled to themselves [as in (28)] and to the amplitudes corresponding to processes of smaller n.

Equation (28) can be simplified by removing the

trivial one-particle structure which is still present. One writes

$$S(xy \mid z_1z_2) = \int S'_V(x-\xi)S(\bar{\xi}y \mid z_1z_2),$$

which permits the second line of (28) to be rewritten in the form

$$\begin{aligned}\int S_V(x-u)\,\Sigma\,(u-v)S'_V(v-\xi)S(\bar{\xi}y \mid z_1z_2)\\ = \int [S'_V(x-\xi) - S_V(x-\xi)]S(\bar{\xi}y \mid z_1z_2)\\ = S(xy \mid z_1z_2) - \int S_V(x-\xi)S(\bar{\xi}y \mid z_1z_2), \qquad (29)\end{aligned}$$

and hence (28) is equivalent to

$$\begin{aligned}S(\bar{x}y \mid z_1z_2) = g^2\,\delta(x-z_1)S_N(z_1-z_2)S'_V(z_2-y)\\ - ig^2S_N(x-z_2)\int \Delta_\vartheta(x-w)S'_V(z_2-u)S(\bar{u}y \mid z_1w). \qquad (30)\end{aligned}$$

A similar amputation performed on the y-coordinate,

$$S(\bar{x}y \mid z_1z_2) = \int S(\bar{x}\bar{\xi} \mid z_1z_2)S'_V(\xi-y),$$

leads to the equation

$$\begin{aligned}M(xy \mid z_1z_2) = -ig^2\,\delta(x-z_1)\\ \times S_N(z_1-z_2)\,\delta(z_2-y) - ig^2S_N(x-z_2)\\ \times \int \Delta_\vartheta(x-w)S'_V(z_2-u)M(uy \mid z_1w), \qquad (31)\end{aligned}$$

where we denote by $M(xy \mid z_1z_2)$ the fully amputated quantity

$$T(\bar{x}\bar{y} \mid \bar{z}_1\bar{z}_2) = -iS(\bar{x}\bar{y} \mid z_1z_2).$$

The renormalization of (31) may be accomplished by a renormalization of the original source function η_V, $\bar{\eta}_V$, and consequent multiplication of (22) by the factor $(Z^{-\frac{1}{2}})^2$. We are then calculating

$$\begin{aligned}Z^{-1}G(xy \mid z_1z_2) = S^R_V(x-y)\,\Delta_\vartheta(z_1-z_2)\\ + Z^{-1}\int S'_V(x-\xi)\,\Delta_\vartheta(z_1-u)\\ \times M(\xi\eta \mid uw)\,\Delta_\vartheta(v-z_2)S'_V(\eta-y), \qquad (32)\end{aligned}$$

and upon amputating on the ϑ-coordinates of the last term of (32), replacing S'_V by ZS^R_V, and amputating on the fermion coordinates with respect to S^R_V, one sees that the renormalized amplitude M_R is given by $M_R = ZM$; this, when substituted into (31) together with the definition of the renormalized charge, $g_R^2 = Zg^2$, produces the corresponding equation in terms of renormalized quantities,

$$\begin{aligned}M_R(xy \mid z_1z_2) = -ig_R^2\,\delta(x-z_1)S_N(z_1-z_2)\,\delta(z_2-y)\\ - ig_R^2S_N(x-z_2)\int \Delta_\vartheta(x-w)\\ \times S^R_V(z_2-u)M_R(uy \mid z_1\,w). \qquad (33)\end{aligned}$$

The mass-shell Fourier transform of M_R represents the renormalized scattering amplitude.

For the special choices (12), (13), the spatial δ-function dependence $\delta(\mathbf{x}-\mathbf{z}_1)\delta(\mathbf{z}_1-\mathbf{z}_2)\delta(\mathbf{z}_2-\mathbf{y})$ factors out of (33), and the remainder can then be considered as a function of the time dependence only, as described previously. If the transform of $M_R(x-z_2, z_2-z_1, z_1-y) \equiv M(a, b, c)$ is denoted by

$$\tilde{M}(q, p, k) = \int_{-\infty}^{+\infty} da\,db\,dc\,e^{i(qa+pb+kc)}M(a, b, c),$$

one obtains from (33) the integral equation

$$\begin{aligned}\tilde{M}(q, p, k) = -ig_R^2\tilde{S}_N(q+k-p)\\ - i\left(\frac{g_R^2}{2\pi}\right)\int_{-\infty}^{+\infty} dQ\;\tilde{\Delta}_\vartheta(p-Q)\\ \times \tilde{S}_N(Q+q-p)\tilde{S}^R_V(Q)\tilde{M}(Q, p, k). \qquad (34)\end{aligned}$$

Because of the retarded nature of the coordinate difference $x - z_2 = a \geq 0$, we may expect that $\tilde{M}(q, p, k)$ is analytic in the upper half q plane (the same remarks hold for the k dependence), and this is born out by examination of the functions $\tilde{S}_N$ and $\tilde{S}^R_V$ which enter into (34). Hence the only contribution to the integral of (34) comes from the cut of $\tilde{\Delta}_\vartheta(p-Q)$,

$$\begin{aligned}\tilde{\Delta}_\vartheta(p-Q) = (2\pi)^{-4}\int \frac{d^3k}{2\omega}\,|f(\omega)|\,(Q-[p-\omega+i\epsilon])^{-1}\\ \equiv (1/2\pi)\mathfrak{s}_\omega(Q+\omega-p-i\epsilon)^{-1}, \qquad (35)\end{aligned}$$

and one obtains

$$\begin{aligned}\tilde{M}(q, p, k) = -ig_R^2\tilde{S}_N(q+k-p)\\ + (g_R^2/2\pi)\mathfrak{s}_\omega\tilde{S}_N(q-\omega)\tilde{S}^R_V(p-\omega)\tilde{M}(p-\omega, p, k), \qquad (36)\end{aligned}$$

which is just Amado's equation, and has the solution constructed by him and Kenschaft in Ref. 2.

IV. THREE GENERALIZATIONS

There are several rather obvious generalizations of the Lee model, suggested by the above functional analysis, which may be worth mentioning.

(A) It has been noted that the conventional Lee model, expressed by (12) and (13), is only one possible choice of a class of retarded functions. We

remark here only that the model in which $\Delta_\vartheta(x)$ has the same form as the propagators of (12) [omitting the $\delta(\mathbf{x})$] is finite without the need of any cutoff, since the relationships between m_V, m_V^R, g, and g_V^R are algebraic, rather than integral.

(B) One can define classes of "anti-Lee models", by replacing all causal propagators by advanced functions; this corresponds to the propagation of antiparticles only.

(C) The limiting case in which $m_N \to \infty$ is completely soluble.[7] What is meant here is that situation in which $S_N(x)$ is replaced by $m_N^{-1}\delta(x)$, or equivalently, $\tilde{S}_N(q) = [m_N - q - i\epsilon]^{-1}$ is replaced by m_N^{-1}. The existence of this limit depends upon the type of integrals into which $\tilde{S}_N$ is folded, and it shall be assumed that a cutoff makes this limit well defined. The reason why this limit leads to a completely soluble theory is that the two essential functional operations can be performed:

(1) The V propagator, defined in the presence of the external sources $j^\dagger$, j can now be obtained exactly. If we use the S_V and Δ_ϑ of (12) and (13), but replace $S_N(x)$ by $m_N^{-1}\delta(x)$, (21) is replaced by

$$G_V^\vartheta(x, y) = S_V(x - y) + \frac{g^2}{m_N}\int S_V(x - u)j^\dagger(u)j(u)G_V^\vartheta(u, y), \qquad (37)$$

and (37) can be solved,

$$G_V^\vartheta(x, y) = i\vartheta(xy) \times \exp\left(-im_V(x - y) + i\frac{g^2}{m_N}\int_y^x dz\, j^\dagger(z)j(z)\right). \qquad (38)$$

[7] In a similar way, the case where the bare mass $m_V \to \infty$ is also completely soluble. It has been assumed everywhere in this paper that $m_N + M > m_V{}^R$.

Equation (38) is just the solution corresponding to that found in the simple static model,[3] when $(g/m_N)j^\dagger(z)j(z)$ is replaced by the boson source $J(z)$.

(2) The functional differentiation operations involved in the passage from G_V^ϑ to S_V^ϑ can be explicitly performed.[6] One obtains

$$S_V^\vartheta(x, y) = i\vartheta(xy)\exp[-im_V(x - y)] \times \exp[ij^\dagger B(1 - \Delta_\vartheta^T B)^{-1}j] \times \exp[\mathrm{Tr}\ln(1 - \Delta_\vartheta^T B)^{-1}], \qquad (39)$$

where $\langle z|B|w\rangle = (g^2/m_N)\vartheta(xz)\delta(z - w)\vartheta(wy)$. Similar but slightly more complicated forms are found when the functional operations are performed upon products of G_V^ϑ factors. It is not difficult to carry through the evaluation of the trace factor of (39) and show that, in the limit $m_N \to \infty$, it corresponds to taking the same limit in $\Sigma^\sim(q)$ of (15), namely $S_V' \to S_V$; this just corresponds to the suppression of every N line and is not particularly interesting. However, there is another option here which can make the limit interesting: if the cutoff is adjusted such that $g^2 m_N^{-1} = g_R^2(Zm_N)^{-1}$ is nonvanishing, and finite, as m_N becomes large, one finds

$$S_V'(x) = i\vartheta(x)e^{-im_V x + L(x)},$$

$$L(x) = \int\frac{dk}{2\pi}\int\frac{dp}{2\pi}\frac{e^{-i(p+k)x}}{(p + k + i\epsilon)^2}\tilde{\Delta}_\vartheta(k)\,\tilde{\Delta}_\vartheta(p)^{-1} \times \ln\left(1 - \frac{g_R^2}{Zm_N}\tilde{\Delta}_\vartheta(p)\right), \qquad (40)$$

which is an example of nontrivial structure. Similar results may be written down for all the other Green's functions.