Detection and "emission" processes of quantum particles in a "chaotic state"

Christine Benard and Odile Macchi

Laboratoire d'Etudes des Phénomènes Aléatoires, Université de Paris-Sud, Bâtiment 210 - 91 Orsay, France

(Received 29 June 1971)

We study the statistical properties of a *chaotic* beam of quantum particles (bosons or fermions) by using wavepacket formalism, which allows us to deal with bosons and fermions in a very symmetrical way. Our study is based on the definition and analysis of two different stochastic point processes consisting, respectively, of the "emission" time instants and detection time instants (or positions in a space) of the particles of the beam. The two processes are shown to be identical in the chaotic boson case and in the chaotic fermion case, respectively. The symmetry between the well-known results concerning the photon detection process and the new results concerning the fermion detection process is pointed out. And finally it is noted that in the *chaotic* case the bunching (or antibunching) effect is present already in the "emission" of the particles.

I. INTRODUCTION

The present paper is a part of our study of the statistical properties of quantum particle beams (bosons or fermions).[1] It is based on the use of wavepacket formalism which has the great advantage of dealing with bosons and fermions in a very symmetrical way. Wavepacket formalism consists of associating a stochastic wavepacket with every particle and of building the state of the whole set of particles by symmetrization or antisymmetrization.

The statistical properties we are specially interested in, in this paper, are the stochastic laws of two point processes which can be defined in a particle beam. One is the ideal detection[2] process, that is the point process consisting of the ideal-detection time instants (or positions in space) of the particles. The other one is defined here in connection with wavepacket formalism. It is called the "emission process" and actually corresponds to emission time instants of the wavepackets if certain conditions for the excitation of the particle source are fulfilled.

The beams considered here are "chaotic" beams.[2] In the photon case, this means that the moments of all orders of the electromagnetic field associated with the photons are identical to the moments of a stationary Gaussian stochastic function. From this definition, as shown by Mollow,[3] certain properties of the density matrix of a chaotic photon field follow. Similar properties are used by Glauber to define the chaotic fermion density matrix.[2] Thus, our starting point is the chaotic density matrix given by Glauber[2] in terms of wavepackets for bosons and fermions. We use it to define the chaotic "emission process" and establish several results for this process as well as for the detection process. After giving some definitions and notations in Sec. II, we deal first with the *photon* case in Sec. III. In this section, starting with the results of Glauber and from those of Macchi,[4] we establish the identity between "emission" and detection processes. At the same time, we give the expression for the exclusive probability densities of these processes in terms of the wavepacket covariance. This new result achieves the statistical description of these two processes. In Sec. IV, we deal with the chaotic *fermion* case and settle the probability distribution of both the detection and "emission" point processes for fermions. These two processes are shown to be identical. It is also pointed out how, in the photon case as well as in the fermion case, exclusive probability densities are related to the wavepacket covariance, while coincidence probability densities are related to a "field covariance."

The symmetry between the results obtained for chaotic bosons and fermions, respectively, is also pointed out.

II. DEFINITIONS AND NOTATIONS

The various processes formed with the particle positions in space consist of random events located in infinitesimal volumes $\{\mathbf{r}_j, \mathbf{dr}_j\}$ of $\mathbb{R}^3$. More simply, we may consider that a particle process consists of random points $\mathbf{r}_j$ of $\mathbb{R}^3$, and thus forms what is called a stochastic point process (pp). Obviously the number of particles within any finite volume U is finite with probability one. Moreover, the particles are assumed to be enclosed in a finite "parallelepipedic" volume V. Hence actually all our pp are restricted to V, which avoids mathematical difficulties linked with the unboundedness of $\mathbb{R}^3$.

We shall consider two different pp which are the "emission process" and the detection process. For the sake of simplicity, they are assumed to be purely spatial processes and in order to avoid confusion the events of these pp are denoted by $\{\mathbf{r}_j\}$ and $\{\mathbf{a}_j\}$, respectively.

A. General properties of point processes

In order to define the probability distribution of a pp in V, one must give all the values $H_n(A_1, j_1, \ldots, A_k, j_k)$ called "exclusive probabilities" which are the probabilities of exactly n events in V, with exactly j_1 events in $A_1, \ldots$, and exactly j_k events in A_k for all nonnegative integers $k, j_1, \ldots, j_k$. The sets $A_1, \ldots, A_k$ are any k Borel sets of V. Obviously, the set function H_n must be (α) nonnegative, (β) symmetrical, and (γ) σ-additive.[5] In addition, one must have the unit property (δ):

$$\sum_{n=0}^{\infty} H_n(V,n) = 1.$$

As intuition would lead us to expect, it can be shown that any system of set functions $H_n(A_1, j_1, \ldots, A_k, j_k)$ satisfying (α)–(δ) actually defines a pp.

Now almost all pp that exist in nature possess another important property, that of being "continuously distributed." Mathematically, this means that if $B_1, \ldots, B_n$ are n disjoint Borel sets of V, and if we call B the product set $B_1 \times \cdots \times B_n$, the set function defined in V^n by $H_n(B) = H_n(B_1, 1, \ldots, B_n, 1)$ is absolutely continuous with respect to the Lebesgue measure in V^n. In physics this means that (almost everywhere) all the exclusive probabilities H_n possess densities. Hence there exists a system of "point functions" $G_n^V[\{\mathbf{x}_j\}]$ such that for n infinitesimal intervals $\{\mathbf{x}_j, \mathbf{dx}_j\}$ of width $\mathbf{dx}_j$, each of them containing point $\mathbf{x}_j$, respectively, we have

$$H_n(\{\mathbf{x}_1, d\mathbf{x}_1\}, 1, \ldots, \{\mathbf{x}_n, d\mathbf{x}_n\}, 1) = G_n^V[\{\mathbf{x}_j\}]\, d\mathbf{x}_1 \cdots d\mathbf{x}_n.$$

Therefore, $G_n^V[\{\mathbf{x}_j\}]\, d\mathbf{x}_1 \cdots d\mathbf{x}_n$ is the probability that there is one event in $\{\mathbf{x}_1, d\mathbf{x}_1\}, \ldots,$ another one in $\{\mathbf{x}_n, d\mathbf{x}_n\}$, and none other in the whole volume V in which the particles are confined. The $G_n^V[\{\mathbf{x}_j\}]$ are called "exclusive probability densities" (epd). They are obviously symmetrical functions of their arguments because of (β), $G_n^V(\mathbf{x}_1, \ldots, \mathbf{x}_n) = G_n^V(\mathbf{x}_{\alpha_1}, \ldots, \mathbf{x}_{\alpha_n})$ for all permutations $\alpha_1, \ldots, \alpha_n$ of $1, \ldots, n$. Moreover, the $G_n^V\{\mathbf{x}_j\}]$ which are the densities of a positive set measure must be nonnegative

$$G_n^V[\{\mathbf{x}_j\}] \geq 0. \tag{P.0}$$

In addition the probability p_n of n events in V is

$$p_n = \frac{1}{n!} \int_{V^n} G_n^V[\{\mathbf{x}_j\}]\, d\mathbf{x}_1 \ldots d\mathbf{x}_n,$$

on account of the symmetry. Thus condition (δ) implies that

$$\sum_{n=0}^{\infty} \frac{1}{n!} \int_{V^n} G_n^V[\{\mathbf{x}_j\}]\, d\mathbf{x}_1 \cdots d\mathbf{x}_n = 1. \tag{P.1}$$

Conversely, a system of symmetrical, nonnegative functions $G_n^V[\{\mathbf{x}_j\}]$ verifying Eq. (P.1), defines the distribution of an actual pp, which is "continuously" distributed, with epd $G_n^V[\{\mathbf{x}_j\}]$. This system is therefore a very practical means to define a pp, and we use it herein.

Another system of probabilities appears useful for the study of pp. For disjoint Borel sets $A_1, \ldots, A_n (A_j \subset V)$ we define the coincidence probability $Q_n(A_1, \ldots, A_n)$ as the probability of exactly one event in $A_1, \ldots,$ exactly one event in A_n, the total number p of events in V being arbitrary. Let $B = V - \sum_{j=1}^{n} A_j$, then the probability of one event in each of the A_j and $p - n$ events in B is found to be

$$\frac{1}{(p-n)!} \int_{A_1} d\mathbf{x}_1 \cdots \int_{A_n} d\mathbf{x}_n \int_{B^{p-n}} G_p^V[\{\mathbf{x}_j\}]\, d\mathbf{x}_{n+1} \cdots d\mathbf{x}_p,$$

and thus

$$Q_n(A_1, \ldots, A_n) = \sum_{p=n}^{\infty} \frac{1}{(p-n)!} \times \int_{A_1} d\mathbf{x}_1 \cdots \int_{A_n} d\mathbf{x}_n \int_{B^{p-n}} G_p^V[\{\mathbf{x}_j\}]\, d\mathbf{x}_{n+1} \cdots d\mathbf{x}_p.$$

This equation shows that coincidence probabilities also admit densities, say $P_n[\{\mathbf{x}_j\}]$, called "coincidence probabilities densities" (cpd). Their physical definition is $P_n[\{\mathbf{x}_j\}]\, d\mathbf{x}_1 \cdots d\mathbf{x}_n = P_r$ [one event in each interval $\{\mathbf{x}_j, d\mathbf{x}_j\}$].

Moreover, it appears in the above series that the cpd are expressed in terms of the epd, by the series

$$P_n[\{\mathbf{x}_j\}] = \sum_{k=0}^{\infty} \frac{1}{k!} \int_{V^k} G_{n+k}^V(\mathbf{x}_1, \ldots, \mathbf{x}_{n+k})\, d\mathbf{x}_{n+1} \cdots d\mathbf{x}_{n+k}.$$

It can be shown (see Ref. 6) that this formula has an inverse,

$$G_n^V[\{\mathbf{x}_j\}] = \sum_{p=0}^{\infty} \frac{(-1)^P}{p!} \int_{V^p} P_{n+p}(\mathbf{x}_1, \ldots, \mathbf{x}_{n+p})\, dx_{n+1} \cdots dx_{n+p},$$

which is basic for Appendix D.

B. Modes in a cavity

We shall consider the modes in a finite cubic cavity of volume $V = L^3$. Each mode is labeled by a wavevector $\mathbf{k}$ with components

$$k_x = p_x/L, \quad k_y = p_y/L, \quad k_z = p_z/L,$$

where p_x, p_y, p_z are three arbitrary integers (but not all three zero). As is well known $\mathbb{Z}^3$ can be mapped into $\mathbb{Z}$ by a one-to-one correspondence. Hence there is a countable collection $\{\mathbf{k}\}$ of vectors $\mathbf{k}$. When necessary, they may be treated as a sequence $\mathbf{k}_1, \ldots, \mathbf{k}_p, \cdots$.

A product over all the modes of the cavity is symbolized by Π_k, and a sum by $\sum_k$. These quantities must be absolutely convergent in order to be significant, since no order is given in the collection $\{\mathbf{k}\}$.

The nonnegative integer n_k is the number of particles in mode $\mathbf{k}$. The mean value of n_k is denoted by $\langle n_k \rangle$.

The symbol $\{n_k\}$ indicates a sequence of nonnegative integers $\{n_1, \ldots, n_p, \cdots\}$, n_p being the number of particles in mode $\mathbf{k}_p$, and $\sum_{\{n_k\}}$ indicates the summation over all such sequences.

C. Other notations

We shall use the following notations: $\langle \cdots \rangle$, which means ensemble average; $|\rangle$, which is a vector in the Fock space. For short, we call it a "ket"; $\sum P_\alpha$, which indicates a sum over all the permutations $\alpha_1, \ldots, \alpha_n$ of $1, \ldots, n$; $\sum P_\alpha(-1)$, which indicates an algebraic sum over all the permutations $\alpha_1, \ldots, \alpha_n$, with coefficients $(-1)^{p(\alpha)}$, where $p(\alpha)$ is the order of the permutation $\alpha_1, \ldots, \alpha_n$ (these two latter notations can be summed up into one notation which is $\sum P_\alpha(\epsilon)$, with $\epsilon = \pm 1$); $\mathcal{L}^2(V^n)(\mathcal{L}^1(V^n))$, which is the space of functions $\Phi(\mathbf{r})$ verifying

$$\int_{V^n} |\Phi(\mathbf{r}_1, \ldots, \mathbf{r}_n)|^2\, d\mathbf{r}_1 \cdots d\mathbf{r}_n < \infty,$$

$$\left(\int_{V^n} |\Phi(\mathbf{r}_1, \ldots, \mathbf{r}_n)|\, d\mathbf{r}_1 \cdots d\mathbf{r}_n < \infty.\right)$$

III. PHOTON CASE

A. Expression of the "chaotic" density matrix in terms of wavepackets

The density matrix ρ of a set of photons in a "chaotic state" and localized in a finite volume V, can be written as[2,3]

$$\rho = \sum_{\{n_k\}} \prod_k \frac{\langle n_k \rangle^{n_k}}{(1 + \langle n_k \rangle)^{1+n_k}} |\{n_k\})(\{n_k\}|, \tag{1}$$

or, in an equivalent form, as

$$\rho = \int \prod_k \exp\left(\frac{-|\alpha_k|^2}{\langle n_k \rangle}\right) |\{\alpha_k\})(\{\alpha_k\}| \frac{d^2\alpha_k}{\langle n_k \rangle}. \tag{2}$$

In Eq. (2), α_k is any complex number and $|\{\alpha_k\})$ symbolizes a state in which, for each mode $\mathbf{k}$, the photons are in the coherent state $|\alpha_k)$.[7] [If a_k is the annihilation operator in mode $\mathbf{k}$, we have $a_k|\alpha_k) = \alpha_k|\alpha_k)$.]

By definition, a "chaotic state" is such that all the moments[7] of the electromagnetic field $\mathbf{E}(\mathbf{r}, t)$ associated with the photons are identical to the moments of a stationary Gaussian stochastic function. In Ref. 3, Mollow has shown that, if these moments are assumed to be bounded, the chaotic density matrix can be written in form (1) or (2). If $\langle n_k \rangle$ is chosen to be equal to $\exp \hbar\omega_k (1 - \exp \hbar\omega_k)^{-1}$ (where ω_k is the angular frequency of mode $\mathbf{k}$), the "chaotic state" describes thermal equilibrium.

Construction of the wavepackets $|n(\{r_j\})$

In Ref. 2 Glauber gives the chaotic density matric ρ in terms of the wavepackets $|n(\{\mathbf{r}_j\})$. The ket $|n(\{\mathbf{r}_j\})$ describes in the Heisenberg picture, a set of n one-photon wavepackets, in the neighborhood of points $\mathbf{r}_1, \ldots, \mathbf{r}_n$. If the excitation of the source is assumed to be "broadline" [8] (Condition H_1), wavepacket "emission times" t_j, or, more correctly, source "excitation times" can be introduced by setting $\mathbf{r}_j = -\mathbf{c}t_j$,[1] where $\mathbf{c}$ is the light velocity vector in direction $\mathbf{r}_j$. Then the ket $|n(\{\mathbf{r}_j\})$ describes a set of n one-photon wavepackets excited at time instants $t_1, \ldots, t_n$. Thus, if Condition H_1 is fulfilled, we shall speak of an "emission process" that is the point process built with points $\mathbf{r}_j$. In a general way, the point process built with points $\mathbf{r}_j$ will be called "emission process" even though no information about the source allows us to give a physical meaning to the time instants t_j. Let us now build the kets $|n(\{\mathbf{r}\})$ as done in Ref. 2. Let $|0)$ represent the vacuum state, $a_k^+(a_k)$ be the creation (annihilation) operator in mode $\mathbf{k}$ $[a_k^+|0) = |k)$, and $[a_k^+ a_{k'}] = \delta_{kk'}]$, and A^+ be the following operator

$$A^+ = \sum_k f(\mathbf{k})\, a_k^+,$$

where the complex function $f(\mathbf{k})$ is such that

$$\sum_k |f(\mathbf{k})|^2 = 1; \tag{3}$$

the ket $A^+|0)$ describes a one-photon wavepacket with spectral density $|f(\mathbf{k})|^2$.

The ket $A^+(\mathbf{r}_j)|0)$, deduced from $A^+|0)$ by a translation of $\mathbf{r}_j$ in the configuration space of the particles, is written

$$\begin{aligned} A^+(\mathbf{r}_j)|0) &= \exp - iP\mathbf{r}_j A^+ |0) \\ &= \sum_k f(\mathbf{k}) \exp(-2i\pi\, \mathbf{k}\cdot\mathbf{r}_j)\, a_k^+ |0). \end{aligned}$$

In this equation, P is the impulsion of the particle and the operator $\exp(-iP\mathbf{r}_j)$ is the displacement operator. The ket $A^+(\mathbf{r}_j)|0)$ describes a wavepacket "emitted" at point $\mathbf{r}_j$. Thus the ket $|n(\{\mathbf{r}_j\}))$ is written

$$|n(\{\mathbf{r}_j\})) = [W_n(\{\mathbf{r}_j\})]^{-1/2} \prod_{j=1}^{n} A^+(\mathbf{r}_j)\,|\,0). \tag{4}$$

Function $W_n(\{\mathbf{r}_j\})$, introduced to normalize vector $|n(\{\mathbf{r}_j\})$, is given by

$$\begin{aligned} W_n(\{\mathbf{r}_j\}) &= (0\,|\prod_{j=1}^{n} A(\mathbf{r}_j)\, A^+(\mathbf{r}_j)\,|\,0), \\ W_n(\{\mathbf{r}_j\}) &= \sum P_\alpha \prod_{j=1}^{n} \gamma(\mathbf{r}_j - \mathbf{r}_{\alpha_j}), \end{aligned} \tag{5}$$

where

$$\begin{aligned} \gamma(\mathbf{r}_j - \mathbf{r}_{\alpha_j}) &= (0\,|\,A(\mathbf{r}_j)\, A^+(\mathbf{r}_{\alpha_j})\,|\,0) \\ &= \sum_k |f(\mathbf{k})|^2 \exp[2i\pi\mathbf{k}\cdot(\mathbf{r}_j - \mathbf{r}_{\alpha_j})]. \end{aligned} \tag{6}$$

Let us emphasize that the latter equation can be written

$$\gamma(\mathbf{r}_j - \mathbf{r}_{\alpha_j}) = \sum_k |f(\mathbf{k})|^2\, V\psi_k(\mathbf{r}_j)\,\psi_k^*(\mathbf{r}_{\alpha_j}), \tag{6'}$$

where

$$\psi_k(\mathbf{r}) = (V)^{1/2} \exp(2i\pi\,\mathbf{k}\cdot\mathbf{r}). \tag{7}$$

The functions $\psi_k(\mathbf{r})$ form a complete orthonormal set (CON) for $\mathcal{L}^2(V)$, because, as $\mathbf{k}$ takes on the values given in Sec. I,

$$\int_V \psi_k(\mathbf{r})\,\psi_{k'}^*(\mathbf{r})\,d\mathbf{r} = \delta_{kk'}. \tag{8}$$

Hence the series (6') actually defines a good function $\gamma(\mathbf{r} - \mathbf{r}')$ because of conditions (3). The set of all kets $|n(\{\mathbf{r}_j\}))$ *uniquely generates by integration* the Fock space, as shown in Appendix A. By this we mean that any vector of the Fock space can be expressed by integrating the $|n(\{\mathbf{r}_j\}))$ vectors. The integration coefficients are unique.

Diagonal Expression for the density matrix in terms of the wavepackets

In Ref. 2, Glauber looks for a diagonal expression for ρ, on the generating system of the wavepackets

$$\rho = \sum_{n=0}^{\infty} \int_{V^n} F_n(\{\mathbf{r}_j\})\,|\,n(\{\mathbf{r}_j\}))(n(\{\mathbf{r}_j\})\,|\prod_{j=1}^{n} \frac{d\mathbf{r}_j}{V}. \tag{9}$$

Such a diagonal form of ρ is of great interest because it can be understood in terms of probabilities: If condition H_1 is fulfilled, $(n!/V^n)F_n(\{\mathbf{r}_j\})$ is the "emission" epd in $\{t_j = -|c|^{-1}|r_j|\}$, i.e., the probability density that any wavepacket will be "emitted" in t_1, any other in $t_2, \ldots$, any other in t_n, and that none other will be "emitted." In a general way, if Condition H_1 is not fulfilled, the quantity $(n!/V^n)F_n(\{\mathbf{r}_j\})$ is the epd that defines the "emission process" constituted by points $\mathbf{r}_j$.

After having given the principle of his demonstration [which is equating expressions (2) and (9)], Glauber obtains the following results:[2]

$$F_n(\{\mathbf{r}_j\}) = c_n W_n(\{\mathbf{r}_j\}), \tag{10}$$

with

$$c_n = (Z^n/n!) \prod_k (1 + \langle n_k\rangle)^{-1} \tag{11}$$

and

$$Z\,|f(\mathbf{k})|^2 = \langle n_k\rangle/(1 + \langle n_k\rangle). \tag{12}$$

From (3) and (12), it follows that

$$\sum_k \langle n_k\rangle/(1 + \langle n_k\rangle) < \infty, \quad \sum_k \langle n_k\rangle < \infty. \tag{3'}$$

Moreover,

$$Z = \sum_k \langle n_k\rangle/(1 + \langle n_k\rangle).$$

As the kets $|n(\{\mathbf{r}_j\}))$ is a uniquely generating system of the Fock space (cf. Appendix A), we know that this is the only solution. A demonstration of Eqs. (10)–(12) is given in Appendix B.

Thus the photon chaotic density matrix ρ is

$$\begin{aligned} \rho = \sum_{n=0}^{\infty} \frac{Z^n}{n!\,V^n} \prod_k (1 + \langle n_k\rangle)^{-1} \\ \times \int_{V^n} |\,0) \prod_{j=1}^{n} A(\mathbf{r}_j) A^+(\mathbf{r}_j)\,(0\,|\prod_{j=1}^{n} d\mathbf{r}_j. \end{aligned} \tag{13}$$

Exclusive probability densities of the "emission process"

From the results (10)–(12) of Glauber, we immediately deduce that the emission epd $G_n^V(\{\mathbf{r}_j\})$ is given by

$$\begin{aligned} G_n^V(\{\mathbf{r}_j\}) &= \frac{n!\,F_n(\{\mathbf{r}_j\})}{V^n} \\ &= \prod_k (1 + \langle n_k\rangle)^{-1} \sum P_\alpha \prod_{j=1}^{n} \frac{Z}{V} \gamma(\mathbf{r}_j - \mathbf{r}_{\alpha_j}), \end{aligned} \tag{14}$$

or, from Eqs. (6′) and (12), by

$$G_n^V(\{\mathbf{r}_j\}) = \prod_k (1+\langle n_k\rangle)^{-1} \sum P_\alpha \prod_{j=1}^n \sum_l \frac{\langle n_l\rangle}{1+\langle n_l\rangle}\, \psi_l(\mathbf{r}_j)\psi_l^*(\mathbf{r}_{\alpha_j}). \tag{15}$$

With Eq. (15) we have totally settled the statistical laws of the "emission process" of chaotic photons. It should be observed that $G_n^V(\{\mathbf{r}_j\})$ depends only on the differences $(\mathbf{r}_j - \mathbf{r}_{\alpha_j})$; thus the "emission process" is spatially stationary inside volume V.

B. Relation between exclusive probability densities and coincidence probability densities

Relation between epd and cpd for a certain type of compound poisson process

Let us consider a spatial compound Poisson process[9] whose events are denoted by $\{\mathbf{x}_j\}$. By definition, a pp is a compound Poisson process if its cpd are the various moments of a nonnegative stochastic function $\rho(\mathbf{x})$, called density.

If we assume that $\rho(\mathbf{x})$ is the square modulus of a stochastic, zero-mean, Gaussian complex signal $X(\mathbf{x})$,[10] which in addition is assumed to be analytic, that is

$$\langle X(\mathbf{x})X(\mathbf{x}')\rangle = 0 \quad \text{for all } x, x',$$

then the cpd are

$$P_n(\{\mathbf{x}_j\}) = \sum P_\alpha \prod_{j=1}^n C(\mathbf{x}_j, \mathbf{x}_{\alpha_j}), \tag{16}$$

where $C(x,x')$ is the covariance of $X(x)$. Such a pp will be denoted by Compound Poisson process of type C, standing for its chaotic character.[11] Equation (16) can be proved very easily. Indeed

$$P_n(\{\mathbf{x}_j\}) = \langle X(\mathbf{x}_1)\cdots X(\mathbf{x}_n)X^*(\mathbf{x}_1)\cdots X^*(\mathbf{x}_n)\rangle, \tag{16′}$$

where, as $X(\mathbf{x})$ is a complex Gaussian function, the $2n$ variables are jointly Gaussian. Hence we can use the well-known formula for the moments of Gaussian variables $Z_1, \ldots, Z_{2n}$,

$$\langle Z_1\cdots Z_{2n}\rangle = \sum_{\alpha\beta} \langle Z_{\alpha_1}Z_{\beta_1}\rangle\cdots\langle Z_{\alpha_n}Z_{\beta_n}\rangle, \tag{17}$$

where $\sum_{\alpha\beta}$ indicates a sum over all possible couplings of $Z_1, \ldots, Z_{2n}$. Because $X(\mathbf{x})$ is an analytic signal, formula (17) applied to Eq. (16′) gives nonzero terms only if $X(\mathbf{x}_i)$ is coupled to $X^*(\mathbf{x}_j)$: Thus we obtain Eq. (16).

It has been shown by Macchi[12] that, for such a process, the epd are given by

$$G_n^v(\{\mathbf{x}_j\}) = \prod_l (1+\lambda_l)^{-1} \sum P_\alpha \prod_{j=1}^n g(\mathbf{x}_j, \mathbf{x}_{\alpha_j}), \tag{18}$$

where

$$g(\mathbf{x},\mathbf{u}) = \sum_l [\lambda_l/(1+\lambda_l)]\, \varphi_l(\mathbf{x})\varphi_l^*(\mathbf{u}). \tag{19}$$

Index l scans the whole set of positive integers. Quantities λ_l and functions $\varphi_l(\mathbf{x})$ are, respectively, the eigenvalues and eigenfunctions of $C(\mathbf{x},\mathbf{x}')$, that is to say, they verify the following equation

$$\lambda_l\varphi_l(\mathbf{x}) = \int_V C(\mathbf{x},\mathbf{u})\,\varphi_l(\mathbf{u})\,d\mathbf{u}, \qquad \forall \mathbf{x} \in V. \tag{20}$$

Moreover the functions $\varphi_l(\mathbf{x})$ are chosen so as to form an orthonormal basis of $\mathcal{L}^2(V)$, as is possible, if the covariance $C(\mathbf{x},\mathbf{x}')$ is continuous.[13]

It should be recalled that, as shown in Sec. I, the whole set of epd defines completely the stochastic law of any point process and that it has been shown by Macchi[6] that the whole set of cpd defines completely the stochastic law of the process as well. Thus, what Macchi has shown in Ref. 4 is that the two definitions (16) and (18) are equivalent for a compound Poisson process of type (C).

Application of the foregoing results to the "emission" process

We now compare Eq. (15) and Eqs. (18) and (19). Let us call $g(\mathbf{r}-\mathbf{r}')$ the function

$$\frac{Z}{V}\,\gamma(\mathbf{r}-\mathbf{r}') = \sum_k \frac{\langle n_k\rangle}{1+\langle n_k\rangle}\,\psi_k(\mathbf{r})\psi_k^*(\mathbf{r}'),$$

which appears in Eq. (15). Since the set of functions $\psi_k(\mathbf{r})$ is a CON, $\sum_k \langle n_k\rangle/(1+\langle n_k\rangle)$ is finite, and $\langle n_k\rangle/(1+\langle n_k\rangle)$ is positive, this function $g(\mathbf{r}-\mathbf{r}')$ is a covariance. We can associate to $g(\mathbf{r}-\mathbf{r}')$ another covariance $C(\mathbf{r}-\mathbf{r}')$ in the same way as done in Eqs. (19) and (20):

$$C(\mathbf{r}-\mathbf{r}') = \sum_k \langle n_k\rangle\,\psi_k(\mathbf{r})\psi_k^*(\mathbf{r}'), \tag{21}$$

which is actually a covariance because the $\langle n_k\rangle$ are positive and of finite sum.

According to Eqs. (18)–(20), the compound Poisson process of type C associated with the covariance given in Eq. (21) has for epd,

$$G_n^V(\{\mathbf{r}_j\}) = \prod_k (1+\langle n_k\rangle)^{-1} \times \sum P_\alpha \prod_{j=1}^n \sum_k \frac{\langle n_k\rangle}{1+\langle n_k\rangle}\,\psi_k(\mathbf{r}_j)\psi_k^*(\mathbf{r}_{\alpha_j}),$$

which is identical to the "emission" process epd given by Eq. (15). Hence the "emission" process is a compound Poisson process of type C associated with the covariance given in Eq. (21), and, in particular, its cpd are given by Eq. (16):

$$P_n(\{\mathbf{r}_j\}) = \sum P_\alpha \prod_{j=1}^n \frac{1}{V}\sum_k \langle n_k\rangle \exp[2i\pi\mathbf{k}\cdot(\mathbf{r}_j - \mathbf{r}_{\alpha_j})]. \tag{22}$$

According to the Kahrunen–Loeve theorem[14] the stochastic function $X(\mathbf{r})$, the covariance of which is $C(\mathbf{r}-\mathbf{r}')$, given by Eq. (21), admits the following expansion

$$X(\mathbf{r}) = (V)^{-1/2} \sum_k \alpha_k \exp(2i\pi\mathbf{k}\cdot\mathbf{r}), \tag{23}$$

where α_k is a Gaussian, zero-mean, stochastic variable, the second order moments of which are

$$\langle \alpha_k \alpha_{k'}^*\rangle = \langle |\alpha_k|^2\rangle\,\delta_{kk'} = \langle n_k\rangle\,\delta_{kk'},$$
$$\langle \alpha_k \alpha_{k'}^*\rangle = 0 \text{ for all } \mathbf{k} \text{ and } \mathbf{k}'. \tag{24}$$

(For $\mathbf{k} \neq \mathbf{k}'$ the stochastic variables α_k and $\alpha_{k'}$, are independent.)

Let us compare $X(\mathbf{r})$ with the positive frequency part $\mathcal{E}(\mathbf{r})$ of the function associated, in the coherent state representation, with the electric field operator, at time $t = 0$, $E(\mathbf{r}, 0)$. We have[7]

$$E(\mathbf{r},0) = (V)^{-1/2} \sum_k (\tfrac{1}{2}\hbar\omega_k)^{1/2} \times [a_k \exp(2i\pi\mathbf{k}\cdot\mathbf{r}) - a_k^+ \exp(-2i\pi\mathbf{k}\cdot\mathbf{r})], \quad (25)$$

where ω_k is the angular frequency of mode $\mathbf{k}$, and

$$\mathcal{E}(\mathbf{r}) = (V)^{-1/2} \sum_k (\tfrac{1}{2}\hbar\omega_k)^{1/2} \alpha_k \exp(2i\pi\mathbf{k}\cdot\mathbf{r}). \quad (26)$$

As the field we are considering is a chaotic one, $\mathcal{E}(\mathbf{r})$ is Gaussian and so are the various α_k, which, moreover, satisfy Eq. (24). From Eqs. (23) and (26), it follows that $X(\mathbf{r})$ and $\mathcal{E}(\mathbf{r})$ differ only by the factors $\frac{1}{2}\hbar\omega_k$. In the optical frequency region, the spectral bandwidth of the field $\Delta\omega_0$ is always much smaller than the mean frequency ω_0 of the field, which allows us to write

$$\mathcal{E}(\mathbf{r}) = (V)^{-1/2} (\tfrac{1}{2}\hbar\omega_0)^{1/2} X(\mathbf{r}). \quad (27)$$

This shows, in particular, that the function $C(\mathbf{r}-\mathbf{r}')$ appearing in the cpd [Eq. (22)] of the "emission" process, is $2V/\hbar\omega_0$ times the covariance of the analytic electromagnetic field.

On the other hand, the function $g(\mathbf{r}-\mathbf{r}')$ appearing in the epd [Eq. (15)] of the "emission" process can be considered as the covariance of the wavepackets. In fact, if we rewrite the wavepacket $A^+(\mathbf{r}_j)\,|\,0)$ in the r representation

$$\Phi(\mathbf{r}-\mathbf{r}_j) = \sum_k f(\mathbf{k})(V)^{-1/2} \exp[2i\pi\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}_j)],$$

Eq. (6) can be expressed in the following way:

$$\gamma(\mathbf{r}_j - \mathbf{r}_{\alpha_j}) = \int_V \Phi^*(\mathbf{r}-\mathbf{r}_j)\Phi(\mathbf{r}-\mathbf{r}_{\alpha_j})\,d\mathbf{r} = \int_V \Phi^*(\mathbf{r}_{\alpha_j}-\mathbf{r})\Phi(\mathbf{r}_j-\mathbf{r})\,d\mathbf{r} = \langle \Phi^*(\mathbf{r}_{\alpha_j})\Phi(\mathbf{r}_j)\rangle V.$$

Thus, $g(\mathbf{r}-\mathbf{r}')$ is equal to $Z\langle\Phi^*(\mathbf{r}_{\alpha_j})\Phi(\mathbf{r}_j)\rangle$.

The chaotic detection and "emission" processes are identical

Let us now consider another pp which is the *detection* process, in the ideal case defined in Refs. 1 and 7. It is the pp consisting of the detection points of photons obtained with detectors of quantum efficiency one and very well localized in time and space. (Their dimensions in time and space are very small compared with the coherence time of the field $\tau_c = \Delta\omega_0^{-1}$, and the coherence length $l = c\tau_c$). It is well known that, under such conditions, the detection cpd are given by Eq. (22).[11] From this, it follows that the detection epd are given by Eq. (15), and that *the "emission process" and the ideal detection process are identical*. This means that, *for chaotic fields*, the bunching effect observed by detection is already present in the "emission".

It is possible to check up on this identity between the two pp, without using the well-known results concerning the ideal detection process. In fact we can obtain the detection epd and cpd directly by letting the coincidence operator[1,15] act on the density matrix given by Eq. (9). The virtue of such a calculation, given in Sec. IIIC, resides in the fact that a similar one must be performed to obtain the detection epd for fermions. Indeed the fermion epd are not yet known except for weak densities.[1]

C. Direct calculation of the detection coincidence probability densities in wavepacket formalism

The coincidence operator will be acting on the chaotic density matrix ρ given by Eq. (13):

$$\rho = \prod_k (1+\langle n_k\rangle)^{-1} \sum_{n=0}^{\infty} \frac{Z^n}{n!\,V^n} \times \int_{V^n} |0) \prod_{j=1}^{n} A(\mathbf{r}_j) A^+(\mathbf{r}_j)(0| \prod_{j=1}^{n} d\mathbf{r}_j. \quad (13)$$

Now the coincidence operator $\mathcal{P}_n(\{\mathbf{a}_j\})$, which is defined in such a way that

$$\mathrm{Tr}\rho\mathcal{P}_n(\{\mathbf{a}_j\}) = P_n(\{\mathbf{a}_j\}) \quad (28)$$

is given by

$$\mathcal{P}_n(\{\mathbf{a}_j\}) = \prod_{j=1}^{n} \mathcal{A}^{(+)}(\mathbf{a}_j)\mathcal{A}(\mathbf{a}_j), \quad (29)$$

where the operator $\mathcal{A}^{(+)}(\mathbf{a}_j)$ is

$$\mathcal{A}^{(+)}(\mathbf{a}_j) = (V)^{-1/2} \sum_k a_k^+ \exp(2i\pi\mathbf{k}\cdot\mathbf{a}_j).$$

Let us next imagine a measurement that would put the system in a state such that one photon is in each one of the n volumes $\{\mathbf{a}_j, d\mathbf{a}_j\}$ and no other photon is in volume V. The probability of this measurement being realized is the detection exclusive probability $G_n^V(\{\mathbf{a}_j\})\prod_{j=1} d\mathbf{a}_j$. Let N_j (N) be the numbers of particles in $\{\mathbf{a}_j, d\mathbf{a}_j\}(V)$. Then, according to the Bayes rule,

$$G_n^V(\{\mathbf{a}_j\}) \prod_{j=1}^{n} d\mathbf{a}_j = \Pr[(N_1=1, N_2=1, \ldots, N_n=1)/(N=n)]\, p_n, \quad (30)$$

where $p_n = \Pr[N=n]$. The *a posteriori* probability $\Pr[(N_1=1,\ldots,N_n=1)/(N=n)]$ may be considered as a coincidence probability, measured for the system in a n particle state.

If we call ρ_n the nth matrix coefficient of the density matrix ρ given by Eq. (13),

$$\rho_n = \frac{Z^n}{n!\,V^n} \prod_k (1+\langle n_k\rangle)^{-1} \int_{V^n} |0) \prod_{j=1}^{n} A(\mathbf{r}_j) A^+(\mathbf{r}_j)(0| \prod_{j=1}^{n} d\mathbf{r}_j, \quad (31)$$

then $\rho_n\cdot(\mathrm{Tr}\rho_n)^{-1}$ is the density matrix of the system in an n particle state. By applying Eq. (28) to such a system, we obtain

$$\Pr[(N_1=1,\ldots,N_n=1)/(N=n)] = \frac{\mathrm{Tr}\rho_n\mathcal{P}_n(\{\mathbf{a}_j\})}{\mathrm{Tr}\rho_n} \prod_{j=1}^{n} d\mathbf{a}_j. \quad (32)$$

Since $\mathrm{Tr}\rho_n = p_n$, we conclude from Eqs. (30) and (32) that

$$G_n^V(\{\mathbf{a}_j\}) = \mathrm{Tr}\rho_n\mathcal{P}_n(\{\mathbf{a}_j\}). \quad (33)$$

Thus from Eqs. (29), (31) and (33), it follows that

$$G_n^V(\{\mathbf{a}_j\}) = \prod_k (1+\langle n_k\rangle)^{-1} \sum P_\alpha \prod_{j=1}^{n} \frac{Z}{V}\gamma(\mathbf{a}_j - \mathbf{a}_{\alpha_j}), \quad (34)$$

which, according to Ref. 4, is equivalent to

$$P_n(\{\mathbf{a}_j\}) = \sum P_\alpha \prod_{j=1}^{n} C(\mathbf{a}_j - \mathbf{a}_{\alpha_j}). \quad (35)$$

With Eq. (35) we once again obtain the well-known detection cpd for a chaotic field and we show directly that with an "emission" epd of the form (14) in the density

matrix ρ, we obtain an identical detection epd [given by Eq. (34)].

IV. FERMION CASE

A. Expressing the "chaotic" density matrix in terms of wavepackets

Following Glauber,[2] we define a fermion "chaotic state" by the density matrix ρ,

$$\rho = \prod_k \rho_k ,$$

with

$$\rho_k = (1 - \langle n_k \rangle)\,|0)(0| + \langle n_k \rangle\,|k)(k|. \tag{36}$$

The ket $|k)$ describes a one-fermion state in mode $\mathbf{k}$. The mean number of fermions in mode $\mathbf{k}$, $\langle n_k \rangle$, verifies the relation $\langle n_k \rangle < 1$.

As he has already done for the photons, Glauber looks, in Ref. 2, for a diagonal expression of ρ on the uniquely generating system of the wavepackets $|n(\{\mathbf{r}_j\}))$,

$$\rho = \sum_{n=0}^{\infty} \int_{V^n} F_n(\{\mathbf{r}_j\})\,|n(\{\mathbf{r}_j\}))(n(\{\mathbf{r}_j\})|\prod_{j=1}^{n} \frac{d\mathbf{r}_j}{V} . \tag{9}$$

Construction of the wavepackets $|n(\{r_j\})$

The normalized wavepacket $|n(\{\mathbf{r}_j\}))$ is given by

$$|n(\{\mathbf{r}_j\})) = [W_n(\{\mathbf{r}_j\})]^{-1/2} \prod_{j=1}^{n} A^+(\mathbf{r}_j)\,|0). \tag{4}$$

The operator $A^+(\mathbf{r}_j)$ is

$$A^+(\mathbf{r}_j) = \sum_k f(\mathbf{k}) \exp(-2i\pi \mathbf{k}\cdot\mathbf{r}_j)\, a_k^+\,|0),$$

where $a_k^+(a_k)$ is now the fermion creation (annihilation) operator in mode $\mathbf{k}$ and verifies

$$[a_k^+ a_{k'}]_- = a_k^+ a_{k'} + a_{k'} a_k^+ = \delta_{kk'} .$$

The ket $A^+(\mathbf{r}_j)\,|0)$ describes a one-fermion wavepacket with spectral density $|f(\mathbf{k})|^2$, "emitted" at point $\mathbf{r}_j$. Function $|f(\mathbf{k})|^2$ is normalized according to Eq. (3). In a way similar to that for photons, we have

$$W_n(\{\mathbf{r}_j\}) = (0|\prod_{j=1}^{n} A(\mathbf{r}_j) A^+(\mathbf{r}_j)\,|0),$$

$$W_n(\{\mathbf{r}_j\}) = \sum P_\alpha(-1) \prod_{j=1}^{n} \gamma(\mathbf{r}_j - \mathbf{r}_{\alpha_j}), \tag{37}$$

with

$$\gamma(\mathbf{r} - \mathbf{r}') = (0|\,A(\mathbf{r}) A^+(\mathbf{r}')\,|0) = \sum_k |f(\mathbf{k})|^2 \exp[2i\pi \mathbf{k}\cdot(\mathbf{r} - \mathbf{r}')], \tag{6}$$

or also in terms of the $\psi_k(\mathbf{r})$ [Eq. (7)],

$$\gamma(\mathbf{r} - \mathbf{r}') = \sum_k |f(\mathbf{k})|^2 V \psi_k(\mathbf{r})\,\psi_k^*(\mathbf{r}'). \tag{6'}$$

The ket $|n(\{\mathbf{r}_j\}))$ describes, in the Heisenberg picture, a set of n one-fermion wavepackets in the neighborhood of points $\mathbf{r}_1, \ldots, \mathbf{r}_n$. As we did for photons, we call "emission process" the pp consisting of points $\mathbf{r}_j$ (even though no physical meaning can be given to the time instants t_j defined by $\mathbf{r}_j = -\mathbf{c}t_j$). As shown in Appendix A, the kets $|n(\{\mathbf{r}_j\}))$ uniquely generates the Fock space.

Diagonal expression for the density matrix on the wavepacket basis

By identifying Eqs. (9) and (36), we obtain with Glauber a possible solution for functions $F_n(\{\mathbf{r}_j\})$ (a demonstration of which is given in Appendix C), which is

$$F_n(\{\mathbf{r}_j\}) = c_n W_n(\{\mathbf{r}_j\}) \tag{10}$$

with

$$c_n = (Z^n/n!) \prod_k (1 - \langle n_k \rangle), \tag{38}$$

and

$$Z\,|f(\mathbf{k})|^2 = \langle n_k \rangle/(1 - \langle n_k \rangle). \tag{39}$$

From Eqs. (3) and (39) it follows that

$$\sum_k \langle n_k \rangle/(1 - \langle n_k \rangle) < \infty, \quad \sum_k \langle n_k \rangle < \infty, \tag{3''}$$

and

$$Z = \sum_k \langle n_k \rangle/(1 - \langle n_k \rangle). \tag{40}$$

Based on Appendix A, we can state that this is the only solution. Thus the fermion chaotic density matrix ρ is

$$\rho = \sum_{n=0}^{\infty} \frac{Z^n}{n!\,V^n} \prod_k (1 - \langle n_k \rangle) \times \int_{V^n} |0) \prod_{j=1}^{n} A(\mathbf{r}_j) A^+(\mathbf{r}_j)(0| \prod_{j=1}^{n} d\mathbf{r}_j. \tag{41}$$

Exclusive probability densities of the "emission process"

As already established for photons, the "emission" epd $G_n^V(\{\mathbf{r}_j\})$ is equal to $(n!/V^n) F_n(\{\mathbf{r}_j\})$. Thus we obtain the chaotic fermion "emission process", which is given by its epd $G_n^V(\{\mathbf{r}_j\})$,

$$G_n^V(\{\mathbf{r}_j\}) = \prod_k (1 - \langle n_k \rangle) \sum P_\alpha(-1) \prod_{j=1}^{n} \frac{Z}{V} \gamma(\mathbf{r}_j - \mathbf{r}_{\alpha_j}), \tag{42}$$

or, from Eqs. (6') and (39),

$$G_n^V(\{\mathbf{r}_j\}) = \prod_k (1 - \langle n_k \rangle) \sum P_\alpha(-1) \prod_{j=1}^{n} \times \sum_l \frac{\langle n_l \rangle}{1 - \langle n_l \rangle} \psi_l(\mathbf{r}_j)\,\psi_l^*(\mathbf{r}_{\alpha_j}).$$

B. Relation between exclusive probability densities and coincidence probability densities

Let $C(\mathbf{r}, \mathbf{r}')$ be a covariance function whose eigenvalues and eigenfunctions are, respectively, λ_l and φ_l on volume V: They are solutions to equation

$$\lambda_l \varphi_l(\mathbf{r}) = \int_V C(\mathbf{r}, \mathbf{r}')\,\varphi_l(\mathbf{r}')\,d\mathbf{r}', \quad \mathbf{r} \in V, \tag{20}$$

and furthermore, they can be selected to form a CON of $\mathcal{L}^2(V)$.

We show that if a pp is defined by its cpd (pp of type C)

$$P_n(\{\mathbf{r}_j\}) = \sum P_\alpha(\epsilon) \prod_{j=1}^{n} C(\mathbf{r}_j, \mathbf{r}_{\alpha_j}), \tag{43}$$

its epd are

$$G_n^V(\{\mathbf{r}_j\}) = \prod_{l=1}^{\infty} (1 + \epsilon\lambda_l)^{-\epsilon} \sum P_\alpha(\epsilon) \prod_{j=1}^{n} g(\mathbf{r}_j, \mathbf{r}_{\alpha_j}), \tag{44}$$

where $\epsilon = \pm 1$ according to the pp and where $g(\mathbf{r}, \mathbf{r}')$ is defined as the root of the integral equation

$$g(\mathbf{r}, \mathbf{r}') + \epsilon \int_V g(\mathbf{r}, \mathbf{r}'')\,C(\mathbf{r}'', \mathbf{r}')\,d\mathbf{r}'' = C(\mathbf{r}, \mathbf{r}'), \tag{45}$$

or, equivalently, as the sum of the series

$$g(\mathbf{r},\mathbf{r}') = \sum_{l=1}^{\infty} \frac{\lambda_l}{1+\epsilon\lambda_l}\,\varphi_l(\mathbf{r})\,\varphi_l^*(\mathbf{r}'). \tag{46}$$

The case $\epsilon = +1$ is the photon case, and has been dealt with in detail in Sec. III.

The proof of the case $\epsilon = -1$ is given in Appendix D.

Then in order that $P_n(\{\mathbf{r}_j\})$ and $G_n^V(\{\mathbf{r}_j\})$ actually be probability densities, it is necessary and sufficient that

$$0 \leq \lambda_l < 1 \quad \text{for all } l. \quad \text{(Condition F)}$$

Application of the foregoing results to the fermion "emission" process

Let us compare Eq. (42) with Eqs. (44) and (46). We call $g(\mathbf{r}-\mathbf{r}')$ the function

$$(Z/V)\gamma(\mathbf{r}-\mathbf{r}') = \sum_k \langle n_k\rangle/(1-\langle n_k\rangle)\,\psi_k(\mathbf{r})\,\psi_k^*(\mathbf{r}'),$$

appearing in Eq. (42). As the set of functions $\psi_k(\mathbf{r})$ constitutes a CON and $\langle n_k\rangle$ is positive, we can associate with $g(\mathbf{r}-\mathbf{r}')$ a covariance function $C(\mathbf{r}-\mathbf{r}')$, as done in Eq. (46),

$$C(\mathbf{r}-\mathbf{r}') = \sum_k \langle n_k\rangle\,\psi_k(\mathbf{r})\,\psi_k^*(\mathbf{r}'). \tag{21}$$

As $\sum_k \langle n_k\rangle < \infty$, this covariance actually exists and is identical to the one given in the photon case.

Let us also notice that the $\langle n_k\rangle$ fulfill condition F. Hence we may apply the preceding results of Sec. IVB. In other words, type C pp, whose cpd are

$$P_n(\{\mathbf{r}_j\}) = \sum P_\alpha(-1)\prod_{j=1}^{n}\frac{1}{V}\sum_k \langle n_k\rangle\,\exp[2i\pi\mathbf{k}\cdot(\mathbf{r}_j-\mathbf{r}_{\alpha_j})], \tag{47}$$

has its epd given by

$$G_n^V(\{\mathbf{r}_j\}) = \prod_k (1-\langle n_k\rangle)\sum P_\alpha(-1)\prod_{j=1}^{n} \times \sum_l \frac{\langle n_l\rangle}{1-\langle n_l\rangle}\,\exp[2i\pi\mathbf{l}\cdot(\mathbf{r}_j-\mathbf{r}_{\alpha_j})],$$

which are identical to the "emission" process epd given by Eq. (42). Hence the "emission" process can be defined as well by its cpd given in Eq. (47).

C. Fermion detection process

The coincidence operator $\mathcal{P}_n(\{\mathbf{a}_j\}^1$ is formally given by Eq. (29); but operator $a_k^+(a_k)$ is now the fermion creation (annihilation) operator in mode $\mathbf{k}$ defined earlier. The coincidence operator is defined[1] such that

$$\mathrm{Tr}\mathcal{P}_n(\{\mathbf{a}_j\}) = P_n(\{\mathbf{a}_j\}). \tag{28}$$

As we have done for photons, we now consider a fictive measurement giving the detection epd. Thus we can write

$$G_n^V(\{\mathbf{a}_j\}) = \mathrm{Tr}\rho_n\,\mathcal{P}_n(\{\mathbf{a}_j\}), \tag{33}$$

where, from Eq. (41),

$$\rho_n = \prod_k (1-\langle n_k\rangle)\,\frac{Z^n}{n!\,V^n}\int_{V^n}|0)\prod_{j=1}^{n}A(\mathbf{r}_j)\,A^+(\mathbf{r}_j)(0|\prod_{j=1}^{n}\mathbf{dr}_j. \tag{48}$$

From these two equations and from Eq. (29), we conclude that the detection epd is given by

$$G_n^V(\{\mathbf{a}_j\}) = \prod_k (1-\langle n_k\rangle)\sum P_\alpha(-1)\prod_{j=1}^{n}(Z/V)\gamma(\mathbf{a}_j-\mathbf{a}_{\alpha_j}), \tag{49}$$

where $\gamma(\mathbf{a}_j-\mathbf{a}_{\alpha_j})$ is given by Eq. (6) or (6′).

Since we know that the two definitions of a pp by its epd and its cpd are equivalent,[4] and more precisely, that Eqs. (43) and (44) are equivalent, we can define the fermion detection process by its cpd, which might be actually measured:

$$P_n(\{\mathbf{a}_j\}) = \sum P_\alpha(-1)\prod_{j=1}^{n}C(\mathbf{a}_j-\mathbf{a}_{\alpha_j}). \tag{50}$$

Therefore we have established with Eq. (50), the cpd of the ideal detection process of "chaotic" fermions. Such quantities were not known beforehand, with the exception of the case of an incoherent and weak fermion beam (see Ref. 1).[16] It need be emphasized that Eqs. (47) and (50) assert that for chaotic fermions as well as for chaotic photons, the "emission process" and the ideal detection process are identical.

It should also be noted that in the fermion case, as in the photon case, $g(\mathbf{r}-\mathbf{r}')$ is the wavepacket covariance. The physical meaning of the stochastic quantity which can be associated with the "field covariance" $C(\mathbf{r}-\mathbf{r}')$ in the fermion case is discussed elsewhere.[16]

V. CONCLUSION

Starting from the results of Glauber, we have defined an "emission" process for chaotic bosons or fermions and computed the detection process of such particles.

For bosons and fermions, respectively, these two processes have been shown to be *identical*. It has also been pointed out how, in both cases, epd are related only to the wavepacket covariance and cpd only to the "field covariance."

Moreover, the description of the statistics of chaotic fermions and bosons in terms of the cpd sets forth the symmetry between the two types of particles, as can be seen by comparing Eqs. (35) and (50).

Finally, we stress that the identity between "emission" and detection processes, in the photon as well as in the fermion case, has a physical interpretation, namely that the bunching (or antibunching) effect[1,11] observed by detection in the *chaotic* case is already present in the emission process.

ACKNOWLEDGMENTS

The authors are endebted to Dr. J. Tarsky for having suggested to them the possible use of the Wiener–Tauberian theorem in the demonstration of Appendix A. They are very grateful to Dr. T. Schuker who greatly improved the English of the present paper.

APPENDIX A

The purpose of this appendix is to demonstrate that the set of vectors

$$|n(\{\mathbf{r}_j\})) = (W_n[\{\mathbf{r}_j\}])^{-1/2}\prod_{j=1}^{n}A^+(\mathbf{r}_j)\,|0) \tag{4}$$

is a uniquely generating system (ugs) of the Fock space.

To do so, we first recall how a basis can be built in the

Fock space only making use of a basis in the one-particle state space H. Then, by applying the Wiener–Tauberian theorem, we prove that any vector of the Fock space is expressed in a unique way in terms of the vectors $|n(\{\mathbf{r}_j\})\rangle$.

Construction of a basis on the Fock space

Let $\mathcal{K}$ be the Hilbert space which is the state space of a particle. Its basis consists of vectors $|\psi_i\rangle$. The vectors of the basis of the dual $\mathcal{K}^*$ of $\mathcal{K}$ are written $\langle\psi_i^n|$. With the vector $|\mathbf{x}\rangle$ belonging to $\mathcal{K}$, the linear form $\langle\psi|\mathbf{x}\rangle$ is written $\psi(\mathbf{x})$.

Let us now consider the Hilbert space $\mathcal{K}^{\otimes n}$ which is the tensorial product $\underbrace{\mathcal{K}\otimes\mathcal{K}\otimes\cdots\otimes\mathcal{K}}_{n}$.[17] Its corresponding basis consists of vectors[18]

$$|\psi_{\{i\}}^n\rangle = |\psi_{i_1}\rangle\otimes|\psi_{i_2}\rangle\otimes\cdots\otimes|\psi_{i_n}\rangle.$$

The vectors of the basis of the dual $(\mathcal{K}^*)^{\otimes n}$ of $\mathcal{K}^{\otimes n}$ are written $\langle\psi_{\{i\}}^n|$.

If $|\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_n\rangle = |\mathbf{x}_1\rangle\otimes|\mathbf{x}_2\rangle\otimes\cdots\otimes|\mathbf{x}_n\rangle$, the form $\langle\psi_{\{i\}}^n|\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_n\rangle$ is given by

$$\psi(\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_n) = \psi_{i_1}(\mathbf{x}_1)\psi_{i_2}(\mathbf{x}_2)\cdots\psi_{i_n}(\mathbf{x}_n).$$

We define[19] the Hilbert space $S\mathcal{K}^{\otimes n}(A\mathcal{K}^{\otimes n})$ where

$$S = \frac{1}{n!}\sum P_\alpha\sigma_\alpha, \qquad A = \frac{1}{n!}\sum P_\alpha(-1)\sigma_\alpha.$$

The operator σ_α, on $\mathcal{K}^{\otimes n}$ is defined as

$$\langle\psi_{\{i\}}^n|\sigma_\alpha|\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_n\rangle = \psi_{i_{\alpha_1}}(\mathbf{x}_1)\psi_{i_{\alpha_2}}(\mathbf{x}_2)\cdots\psi_{i_{\alpha_n}}(\mathbf{x}_n),$$

where $\{\alpha_1,\ldots,\alpha_n\}$ is a permutation of $\{1,2,\ldots,n\}$. So we have

$$\langle\psi_{\{i\}}^n|S(A)|\mathbf{x}\rangle = \langle\psi_{\{i\}}^n|\sum P_\alpha(\epsilon)\sigma_\alpha|\mathbf{x}\rangle = \sum P_\alpha(\epsilon)\prod_{j=1}^n\psi_{i_{\alpha_j}}(\mathbf{x}_j). \tag{A1}$$

The set of all functions $S|\psi_{\{i\}}^n\rangle(A|\psi_{\{i\}}\rangle)$ is a basis of $S\mathcal{K}^{\otimes n}(A\mathcal{K}^{\otimes n})$.[20]

The space to which the vectors $|\{n_k\}\rangle, |\{\alpha_k\}\rangle$, and $|n(\{\mathbf{r}_j\})\rangle$ belong is the Fock space F_1 if we consider bosons and is F_2 if we consider fermions

$$F_1 = \bigoplus_{n=0}^{+\infty} S\mathcal{K}^{\otimes n}, \qquad F_2 = \bigoplus_{n=0}^{+\infty} A\mathcal{K}^{\otimes n},$$

where the symbol $\oplus$ means a tensorial sum. By definition, the space $H^{\otimes 0}$ is the set of complex numbers. A basis of $F_1(F_2)$ is the set of all vectors $S|\bar{\psi}_{\{i\}}\rangle(A|\psi_{\{i\}}\rangle)$ for all n. The basis of vectors $|\{n_k\}\rangle$ is built in this way, but the basis of vectors $|\{\alpha_k\}\rangle$ is not.

Let us show that $|n(\{\mathbf{r}_j\})\rangle$ can be written in a quite similar manner. Setting

$$A^+(\mathbf{r}_j)|0\rangle = |\Phi_j\rangle,$$

we know[21] that the creation and annihilation operator $A^+(\mathbf{r}_j), A(\mathbf{r}_j)$ in state $|\Phi_j\rangle$ satisfy the relations

$$[A^+(\mathbf{r}_j)A(\mathbf{r}_i)]_\pm = \langle\Phi_j|\Phi_i\rangle$$

and $[A(\mathbf{r}_j)A(\mathbf{r}_i)]_\pm = [A^+(\mathbf{r}_j)A^+(\mathbf{r}_i)] = 0, \quad \forall_i, \forall_j.$

Then, according to Ref. 21, it can easily be shown that

$$\prod_{j=1}^n A^+(\mathbf{r}_j)|0\rangle \equiv S(A)|\Phi_{\{j\}}^n\rangle,$$

where

$$|\Phi_{\{j\}}^n\rangle = |\Phi_{j1}\rangle\otimes|\Phi_{j2}\rangle\otimes\cdots\otimes|\Phi_{jn}\rangle.$$

Thus, according to Eq. (4), vector $|n(\{\mathbf{r}_j\})\rangle$ can be written as

$$|n(\{\mathbf{r}_j\})\rangle = (n!/W_n[\{\mathbf{r}_j\}])^{1/2}\ S(A)|\Phi_{\{j\}}^n\rangle. \tag{A2}$$

By using Eq. (A1) in conjunction with Eq. (A2) it can be shown that

$$\langle n(\{\mathbf{r}_j\})|\mathbf{x}_1,\ldots,\mathbf{x}_n\rangle = (n!/W_n[\{\mathbf{r}_j\}])^{1/2}\times\sum P_\alpha(\epsilon)\prod_{j=1}^n\Phi_{i_{\alpha_j}}(\mathbf{x}_j) = (W_n[\{\mathbf{r}_j\}])^{-1/2}\sum P_\alpha(\epsilon)\prod_{j=1}^n\Phi(\mathbf{x}_j-\mathbf{r}_{\alpha_j}). \tag{A3}$$

Expressing any vector of the Fock space in terms of vector $|n(\{\mathbf{r}_j\})\rangle$

Any state vector $|X\rangle$ of the Fock space is uniquely written in terms of the $|n(\{\mathbf{r}_j\})\rangle$, which is equivalent to showing, in the space $SH^{\otimes n}(AH^{\otimes n})$, for any nonnegative n, that a vector $|X\rangle$ can be written in a unique way:

$$|X\rangle = \int_{V^n} u(\mathbf{r}_1,\mathbf{r}_2,\ldots,\mathbf{r}_n)|n(\{\mathbf{r}_j\})\rangle d\mathbf{r}_1 d\mathbf{r}_2\cdots d\mathbf{r}_n. \tag{A4}$$

In this formula, V is arbitrary, $|n(\{\mathbf{r}_j\})\rangle$ is given by Eqs. (4) or (A3), and it is assumed that the spectral coefficients $f(\mathbf{k})$ of the wavepacket $|\Phi\rangle$,

$$|\Phi\rangle = \sum_k f(\mathbf{k})a_k^+|0\rangle$$

are all nonzero

$$f(\mathbf{k}) \neq 0, \qquad \forall\mathbf{k}. \tag{A5}$$

Moreover, $u(\mathbf{r}_1,\mathbf{r}_2,\ldots,\mathbf{r}_n)$ is (anti) symmetrical.

Instead of proving formula (A4) for any vector $|X\rangle$, we can equivalently prove it for every vector $|\psi_i^n\rangle$ of a basis of $SH^{\otimes n}(AH^{\otimes n})$. Moreover, property (A4) can be proved in the space $SH^{\otimes n}(AH^*)^{\otimes n}$. Thus we have to show that

$$\langle\psi_{\{i\}}^n|\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_n\rangle = \int_{V^n} u(\mathbf{r}_1,\mathbf{r}_2,\ldots,\mathbf{r}_n)\times\langle n(\{\mathbf{r}_j\})|\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_n\rangle d\mathbf{r}_1 d\mathbf{r}_2\cdots d\mathbf{r}_n \tag{A6}$$

for any vector $|\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_n\rangle$, where $(\mathbf{x}_1,\ldots,\mathbf{x}_n)$ is a point of V^n. The two inner products appearing in Eq. (A6) are functions of $\mathfrak{L}^1(V^n)$. According to formulae (A1) and (A3) and by choosing $\langle\psi_{i_p}|\mathbf{y}_p\rangle \equiv (V)^{-1/2}\exp 2i\pi\mathbf{k}_p\cdot\mathbf{y}_p$, it is necessary to show that a unique symmetrical function $u(\mathbf{r}_1,\mathbf{r}_2,\ldots,\mathbf{r}_n)$ exists which depends on $\mathbf{k}_1,\ldots,\mathbf{k}_n$, such that

$$V^{-n/2}\sum P_\alpha(\epsilon)\,e^{2i\pi\mathbf{k}_i\cdot\mathbf{x}_{\alpha_1}}\cdots e^{2i\pi\mathbf{k}_n\cdot\mathbf{x}_{\alpha_n}} = \int_{V^n} u(\mathbf{r}_1,\mathbf{r}_2,\ldots,\mathbf{r}_n)\times\frac{\sum P_\alpha(\epsilon)\Phi(\mathbf{x}_{\alpha_1}-\mathbf{r}_1)\cdots\Phi(\mathbf{x}_{\alpha_n}-\mathbf{r}_n)}{W_n[\{\mathbf{r}_j\}]^{1/2}}\,d\mathbf{r}_1\cdots d\mathbf{r}_n. \tag{A7}$$

This is a straightforward consequence of the relation

$$V^{-n/2}\, e^{2i\pi \mathbf{k}_1\cdot\mathbf{y}_1} \cdots e^{2i\pi \mathbf{k}_n\cdot\mathbf{y}_n} = \int_{V^n} u(\mathbf{r}_1, \mathbf{r}_2, \ldots \mathbf{r}_n) \times \frac{\Phi(\mathbf{y}_1 - \mathbf{r}_n)\cdots\Phi(\mathbf{y}_n - \mathbf{r}_n)}{W_n[\{\mathbf{r}_j\}]^{1/2}}\, \mathbf{dr}_1 \cdots \mathbf{dr}_n. \quad \text{(A8)}$$

From the Wiener–Tauberian theorem[22] we know that Eq. (A8) holds for at most one function $u(\mathbf{r}_1, \ldots, \mathbf{r}_n)$ in $\mathcal{L}^1(V^n)$ provided the Fourier coefficients $f(\mathbf{k}'_1)\cdots f(\mathbf{k}'_n)$ of the function $\Phi(\mathbf{r}_1)\cdots\Phi(\mathbf{r}_n)$ are all nonzero. This is satisfied given Eq. (A5).

Moreover, the uniqueness of $u(\mathbf{r}_1, \ldots, \mathbf{r}_n)$ can be shown as follows. Putting into Eq. (A8) the Fourier expansion of $u(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_n)\, W_n[\{\mathbf{r}_j\}]^{-1/2}$,

$$u(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_n)\, W_n[\{\mathbf{r}_j\}]^{-1/2} = \sum_{j_1,\ldots,j_n} V^{-n/2}\, t(\mathbf{j}_1, \ldots, \mathbf{j}_n)\, e^{2i\pi \mathbf{j}_1\cdot\mathbf{r}_1} \cdots e^{2i\pi \mathbf{j}_n\cdot\mathbf{r}_n}$$

and of $\Phi(\mathbf{y}_1 - \mathbf{r}_1)\cdots\Phi(\mathbf{y}_n - \mathbf{r}_n)$,

$$\Phi(\mathbf{y}_1 - \mathbf{r}_1)\cdots\Phi(\mathbf{y}_n - \mathbf{r}_n) = \sum_{l_1,\ldots,l_n} f(\mathbf{l}_1)\cdots f(\mathbf{l}_n) V^{-n/2} e^{2i\pi \mathbf{l}_1\cdot(\mathbf{y}_1-\mathbf{r}_1)} \cdots e^{2i\pi \mathbf{l}_n\cdot(\mathbf{y}_n-\mathbf{r}_n)},$$

we get

$$\begin{aligned} V^{-n/2}\, & e^{2i\pi \mathbf{k}_1\cdot\mathbf{y}_1} \cdots e^{2i\pi \mathbf{k}_n\cdot\mathbf{y}_n} \\ &= \sum_{\substack{j_1,\ldots,j_n \\ l_1,\ldots,l_n}} t(\mathbf{j}_1, \ldots, \mathbf{j}_n) f(\mathbf{l}_1)\cdots f(\mathbf{l}_n) \\ &\quad \times e^{2i\pi \mathbf{l}_1\cdot\mathbf{y}_1} \cdots e^{2i\pi \mathbf{l}_n\cdot\mathbf{y}_n} \\ &\quad \times \int_{V^n} V^{-n}\, e^{2i\pi \mathbf{r}_1\cdot(\mathbf{j}_1-\mathbf{l}_1)} \cdots e^{2i\pi \mathbf{r}_n\cdot(\mathbf{j}_n-\mathbf{l}_n)}\, \mathbf{dr}_1 \cdots \mathbf{dr}_n \\ &= \sum_{l_1,\ldots,l_n} t(\mathbf{l}_1, \ldots, \mathbf{l}_n) \\ &\quad \times f(\mathbf{l}_1)\cdots f(\mathbf{l}_n)\, e^{2i\pi \mathbf{l}_1\cdot\mathbf{y}_1} \cdots e^{2i\pi \mathbf{l}_1\cdot\mathbf{y}_n}. \end{aligned} \quad \text{(A9)}$$

Since the set of functions $V^{-n/2}\, e^{2i\pi \mathbf{l}_1\cdot\mathbf{y}_1} \cdots e^{2i\pi \mathbf{l}_n\cdot\mathbf{y}_n}$ is a basis of $\mathcal{L}^1(V^n)$, we have to set

$$\begin{aligned} & t(\mathbf{l}_1, \ldots, \mathbf{l}_n) = 0 \\ & \text{if } \mathbf{l}_1 \neq \mathbf{k}_1, \text{ or } \mathbf{l}_2 \neq \mathbf{k}_2, \cdots \text{or } \mathbf{l}_n \neq \mathbf{k}_n, \\ & t(\mathbf{k}_1, \ldots, \mathbf{k}_n) = V^{-n/2}[f(\mathbf{k}_1)\cdots f(\mathbf{k}_n)]^{-1}. \end{aligned} \quad \text{(A10)}$$

Thus $u(\mathbf{r}_1, \ldots, \mathbf{r}_n)\, W_n[\{\mathbf{r}_j\}]^{-1/2}$ is unique, and of course $u(\mathbf{r}_1, \ldots, \mathbf{r}_n)$ is unique.

APPENDIX B

The chaotic density matrix ρ can be written in the form

$$\rho = \rho_1 = \sum_{\{n_k\}} \prod_k \frac{\langle n_k\rangle^{n_k}}{(1 + \langle n_k\rangle)^{1+n_k}}\, |\{n_k\}\rangle\langle\{n_k\}| \quad (1)$$

on the basis of vectors $|\{n_k\}\rangle$.

We look for a diagonal expression ρ_2 of ρ on the uniquely generating system of vectors $|n(\{\mathbf{r}_j\})\rangle$,

$$\rho = \rho_2 = \sum_{n=0}^{+\infty} \int_{V^n} F_n(\{\mathbf{r}_j\})\, |n(\{\mathbf{r}_j\})\rangle\langle n(\{\mathbf{r}_j\})|\, V^{-n} \prod_{j=1}^{n} \mathbf{dr}_j. \quad (9)$$

In order to obtain function $F_n(\{\mathbf{r}_j\})$, assuming it exists, we identify $(\{n_k\}|\rho_1|\{n_{k'}\})$ and $(\{n_k\}|\rho_2|\{n_{k'}\})$ for any $|\{n_k\})$ and $|\{n_{k'}\})$.

From expression (4) for $|n(\{\mathbf{r}_j\}))$, we deduce that

$$\begin{aligned} |n(\{\mathbf{r}_j\})) = W_n[\{\mathbf{r}_j\}]^{-1/2} & \sum_{\{n_k\}/\sum n_k = n} C_n^{n_1}(j_1) f(\mathbf{k}_1)^{n_1} \\ & \times \prod_{j_1} \exp(-2i\pi \mathbf{k}_1\cdot\mathbf{r}_{j_1}) \cdots \\ & \times C_{n-n_1-\cdots-n_{p-1}}^{n_p}(j_p) f(\mathbf{k}_p)^{n_p} \\ & \times \prod_{j_p} \exp(-2i\pi \mathbf{k}_p\cdot\mathbf{r}_{j_p}) \prod_k \sqrt{n_k!}\, |\{n_k\}). \end{aligned} \quad \text{(B1)}$$

The symbol $\sum_{\{n_k\}/\sum n_k = n}$ indicates a double summation. First, for a given set $\{\mathbf{k}\} = \{\mathbf{k}_1, \ldots, \mathbf{k}_l, \ldots, \mathbf{k}_p\}$, $p \leq n$ of *different* wave vectors, we have to sum over all *different* ordered sets of positive integers $\{n_k\} = n_1, \ldots, n_l, \ldots, n_p$, fulfilling the equality $n_1 + n_2 + \cdots + n_p = n$. The integer n_1 is the number of photons in mode $\mathbf{k}_1, \ldots, n_p$ in mode $\mathbf{k}_p$. Secondly, we sum over all different sets $\{\mathbf{k}\}$.

The symbol $C_n^{n_l}(j_l)$ means that we take the sum over all possible sets of n_l nonordered elements $\{j_l\}$ taken in the N indexes not yet chosen among $1, 2, \ldots, n$.

Let us introduce all modes $\mathbf{k}$ in every term of the summation $\sum_{\{n_k\}/\sum n_k = n}$; since $\sum n_k = n$, all modes $\mathbf{k}$ can be obviously accepted in the product $\prod_k \sqrt{n_k!}$. Moreover, by introducing all the zero-particle modes ($n_p = 0$) for which there are $C_0^0 = 1$ terms and for which $\exp(-2i\pi \mathbf{k}_p\cdot\mathbf{r}_{j_p}$ is conventionally taken to be one, we may rewrite Eq. (B1) as

$$\begin{aligned} & |n(\{\mathbf{r}_j\})) \times (W_n[\{\mathbf{r}_j\}])^{1/2} \\ & \quad \times \sum_{\{n_k\}/\sum n_k = n} \prod_{p=1}^{\infty} [C_{n-n_1-\cdots-n_{p-1}}^{n_p}(j_p) f(\mathbf{k}_p)^{n_p} \\ & \quad \times \prod_{j_p} e^{-2\pi \mathbf{k}_p\cdot\mathbf{r}_{jp}} \sqrt{n_p!}\,]\{n_k\}). \end{aligned} \quad \text{(B1')}$$

The nondiagonal terms $(\{n_k\}|\rho_1|\{n_{k'}\})$ with $\{n_k\} \neq \{n_{k'}\}$ are zero. Thus, by using Eq. (B1'), the nondiagonal terms $(\{n_k\}|\rho_2|\{n_{k'}\})$ fulfill

$$\begin{aligned} \int_{V_n} \frac{F_n(\{r_j\})}{W_n[\{r_j\}]} & \prod_{p=1}^{\infty} [C_{n-n_1-\cdots-n_{p-1}}^{n_p}(j_p) f(\mathbf{k}_p)^{n_p} \\ & \times \prod_{j_p} \exp(-2i\pi \mathbf{k}_p\cdot\mathbf{r}_{jp}) \sqrt{n_p!}\,] \\ & \times \prod_{p=1}^{\infty} C_{n-n'_1-\cdots-n'_{p-1}}^{n'_p}(j'_p) f(\mathbf{k}_p)^{n'_p} \\ & \times \prod_{j'_p} \exp(-2i\pi \mathbf{k}_p\cdot\mathbf{r}_{j'_p}) \sqrt{n'_p!} \prod_{j=1}^{n} \mathbf{dr}_j / V^n = 0, \end{aligned} \quad \text{(B2)}$$

where at least one element of the set $\{n_k\} = n_1, n_2, \ldots, n_p, \cdots$ differs from one element of the set $\{n'_k\} = n'_1, n'_2, \ldots, n'_p, \cdots$. There is an obvious solution of Eq. (B2) which is

$$F_n(\{r_j\}) / (W_n[\{r_j\}]) = \text{const} = c_n. \quad \text{(B3)}$$

Equation (B3) is identical to Eq. (10) of Sec. III. In order to evaluate c_n let us now consider the diagonal terms. The density matrix ρ_2 must be such that

$$(\{n_k\} \mid \rho_2 \mid \{n_k\}) = \prod_k \left(\frac{\langle n_k\rangle}{1+\langle n_k\rangle}\right)^{n_k} \frac{1}{1+\langle n_k\rangle} \tag{B4}$$

By computing $(\{n_k\} \mid \rho_2 \mid \{n_k\})$ from Eqs. (9) and (B1'), we obtain

$$(\{n_k\} \mid \rho_2 \mid \{n_k\}) = \int \frac{F_n(\{r_j\})}{W_n[\{r_j\}]} \times \prod_{p=1}^{\infty} C_{n-n_1-\cdots-n_{p-1}}^{n_p}(j_p) \mid f(\mathbf{k}_p)\mid^{2n_p} n_p! \prod_{j=1}^{n} d\mathbf{r}_j / V^n, \tag{B5}$$

which gives, using Eq. (B3) and taking into account that

$$\prod_{p=1}^{\infty} C_{n-n_1-\cdots-n_{p-1}}^{n_p} n_p! = n!,$$

$$(\{n_k\} \mid \rho_2 \mid \{n_k\}) = c_n n! \prod_k \mid f(\mathbf{k})\mid^{2n_k}. \tag{B6}$$

From Eqs. (B4) and (B6) it follows that, if we set

$$Z \mid f(\mathbf{k})\mid^2 = \langle n_k\rangle / (1+\langle n_k\rangle), \tag{B7}$$

$$c_n = (Z^n/n!) \prod_k (1+\langle n_k\rangle^{-1}. \tag{B8}$$

Equations (B7) and (B8) are identical to Eqs. (12) and (11) of Sec. III, respectively.

Thus, we have shown that the solution given by Eqs. (10)–(12) in Sec. III is possible. Moreover, we know from Appendix A, that this solution is the only one possible.

APPENDIX C

The demonstration in this Appendix is very similar to that in Appendix B.

Let us call $\mid\{\mathbf{k}_j\})$ with $1 \leq j \leq n$ the vector describing the state where one fermion is in mode $\mathbf{k}_1, \ldots,$ one fermion in mode $\mathbf{k}_j, \ldots,$ and one fermion in mode $\mathbf{k}_n$.

On the basis of vectors $\mid\{\mathbf{k}_j\})$ (n varying from 1 to infinity), the chaotic density matrix of the fermions, expressed by Eq. (36), takes the form

$$\rho = \rho_1 = \prod_k (1-\langle n_k\rangle) \sum_{\{k_j\}} \prod_{j=1}^{n} \frac{\langle n_j\rangle}{1-\langle n_j\rangle} \mid\{\mathbf{k}_j\})(\{\mathbf{k}_j\}\mid. \tag{C1}$$

We look for an expression ρ_2 of ρ on the uniquely generating system of vectors $\mid n(\{\mathbf{r}_j\}))$,

$$\rho_2 = \sum_{n=0}^{+\infty} \int_{V^n} F_n(\{\mathbf{r}_j\}) \mid n(\{\mathbf{r}_j\}))(n(\{\mathbf{r}_j\}) \mid V^{-n} \prod_{j=1}^{n} d\mathbf{r}_j. \tag{9}$$

We obtain $F_n(\{\mathbf{r}_j\})$ by identifying $(\{\mathbf{k}_j\} \mid \rho_1 \mid \{\mathbf{k}'_j\})$ and $(\{\mathbf{k}_j\} \mid \rho_2 \mid \{\mathbf{k}'_j\})$ for any $\mid\{\mathbf{k}_j\})$ and $\mid\{\mathbf{k}'_j\})$. From Eq. (4), we deduce the following expression for $\mid n(\{\mathbf{r}_j\}))$ on the basis of the vectors $\mid\{\mathbf{k}_j\})$:

$$\mid n(\{\mathbf{r}_j\})) = (W_n[\{\mathbf{r}_j\}])^{-1/2} \sum_{\{k_j\}} \times \prod_{j=1}^{n} f(\mathbf{k}_j) \sum P_\alpha \exp(2i\pi \mathbf{k}_j \cdot \mathbf{r}_{\alpha_j}) \mid\{\mathbf{k}_j\}). \tag{C2}$$

Thus, the nondiagonal terms $(\{\mathbf{k}_j\} \mid \rho_2 \mid \{\mathbf{k}'_j\})$ with $\{\mathbf{k}_j\} \neq \{\mathbf{k}'_j\}$ verify

$$\int_{V^n} \frac{F_n(\{\mathbf{r}_j\})}{W^n[\{\mathbf{r}_j\}]} \prod_{j=1}^{n} f(\mathbf{k}_j) f^*(\mathbf{k}'_j) \sum P_\alpha \exp(2i\pi \mathbf{k}_j \cdot \mathbf{r}_{\alpha_j}) \times \sum P_\beta \exp(2i\pi \mathbf{k}'_j \cdot \mathbf{r}_{\beta_j}) d\mathbf{r}_j = 0. \tag{C3}$$

An obvious solution of Eq. (C3) is

$$F_n(\{\mathbf{r}_j\}) / (W_n[\{\mathbf{r}_j\}]) = \text{const} = c_n. \tag{C4}$$

Equation (C4) is identical to Eq. (10) of Sec. IV.

The diagonal terms $(\{\mathbf{k}_j\} \mid \rho_2 \mid \{\mathbf{k}_j\})$ satisfy

$$V^{-n} n! \int_{V_n} \frac{F_n(\{\mathbf{r}_j\})}{W_n[\{\mathbf{r}_j\})} \prod_{j=1}^{n} \mid f(\mathbf{k}_j)\mid^2 d\mathbf{r}_j = \prod_k (1-\langle n_k\rangle \prod_{j=1}^{n} \frac{\langle n_j\rangle}{1-\langle n_j\rangle}. \tag{C5}$$

By setting

$$Z \mid f(\mathbf{k})\mid^2 = \langle n_k\rangle/(1-\langle n_k\rangle), \tag{C6}$$

we obtain

$$c_n = (Z^n/n!) \prod_k (1-\langle n_k\rangle). \tag{C7}$$

Equations (C7) and (C6) are identical to Eqs. (38) and (39) of Sec. IV, respectively. Thus the solution given by Eqs. (10), (38), and (39) is a possible solution. According to Appendix A, it is only one.

APPENDIX D

It is shown in Ref. 13 that the probability distribution of a pp may be thoroughly stipulated either by the system of its cpd $P_n[\{\mathbf{r}_j\}]$ or by that of its epd in all volumes U, $G_n^U[\{\mathbf{r}_j\}]$.

In particular, it is shown that for any pp, both systems are related by

$$G_n^U[\{\mathbf{r}_j\}] = \sum_{p=0}^{+\infty} \frac{(-1)^p}{p!} \int_{U^p} P_{n+p}(\mathbf{r}_1, \ldots, \mathbf{r}_{n+p}) d\mathbf{r}_{n+1} \cdots d\mathbf{r}_{n+p}. \tag{D1}$$

Here we show that the cpd system

$$P_n[\{\mathbf{r}_j\}] = \sum P_\alpha (-1) \prod_{j=1}^{n} C(\mathbf{r}_j, \mathbf{r}_{\alpha j}), \tag{D2}$$

and the epd system

$$G_n^U[\{\mathbf{r}_j\}] = \prod_{l=1}^{\infty} (1-\lambda_l) \sum P_\alpha (-1) \prod_{j=1}^{n} g(\mathbf{r}_j, \mathbf{r}_{\alpha j}) \tag{D3}$$

define the same stochastic pp, provided that

$$g(\mathbf{r}, \mathbf{r}') - \int_U g(\mathbf{r}, \mathbf{r}'') C(\mathbf{r}'', \mathbf{r}') d\mathbf{r}'' = C(\mathbf{r}, \mathbf{r}'), \tag{D4}$$

the λ_l being the eigenvalues of the covariance $C(\mathbf{r}, \mathbf{r}')$ in volume U.

In order to prove this property, it suffices to show that Eq. (D2) and Eq. (D1) imply Eqs. (D3) and (D4). If this is true, let $\mathcal{P}_1$ and $\mathcal{P}_2$ be the stochastic pp whose distributions are defined by Eqs. (D2) and (D3), respectively. According to Ref. 6, $\mathcal{P}_1$ then has epd following Eq. (D3). Thus $\mathcal{P}_1$ has the same distribution as $\mathcal{P}_2$.

Proof of Eq. (D3)

We now proceed to prove Eq. (D3) from Eqs. (D1) and (D2).

This proof relies entirely on the covariance diagonalization by means of a complete orthonormal set (CON) over the volume U for the class $\mathcal{L}^2(U)$ (of functions with finite energy). It is well known[13] that, for continuous covariance $C(\mathbf{r}, \mathbf{r}')$, which we assume, a CON exists whose elements are eigenfunctions of $C(\mathbf{r}, \mathbf{r}')$, in other words solutions of Eq. (20). The diagonalization is

$$C(\mathbf{r}, \mathbf{r}') = \sum_{l=1}^{\infty} \lambda_l \varphi_l(\mathbf{r}) \varphi_l^*(\mathbf{r}'), \quad \lambda_l \geq 0 \tag{D5}$$

and the corresponding series is uniformly and absolutely convergent over $U \times U$.

More generally, regardless of any probability context, the proof of Eq. (D3) from Eqs. (D1) and (D2) given herein, is valid provided the function $C(\mathbf{r}, \mathbf{r}')$ of Eq. (D2) is known to be the sum of an absolutely convergent series $\sum_l \lambda_l \varphi_l(\mathbf{r}) \varphi_l^*(\mathbf{r}')$, where the various $\varphi_l(\mathbf{r})$ form a CON and where $\sum_l |\lambda_l| < \infty$. Then necessarily

$$\lambda_l \varphi_l(\mathbf{r}) = \int_U C(\mathbf{r}, \mathbf{r}') \varphi_l(\mathbf{r}') \, d\mathbf{r}' ;$$

but the eigenvalues λ_l are not necessarily positive.

By means of Eq. (D5), it is shown below that with Eq. (D2),

$$\begin{aligned} \int_{U^p} P_{n+p}(\mathbf{r}_1, \ldots, \mathbf{r}_{n+p}) \, d\mathbf{r}_{n+1} \cdots d\mathbf{r}_{n+p} \\ = \sum_{l_1, \ldots, l_n} \lambda_{l_1} \cdots \lambda_{l_n} \varphi_{l_1}(\mathbf{r}_1) \cdots \varphi_{l_n}(\mathbf{r}_n) \\ \times \sum P_\alpha (-1) \varphi_{l_{\alpha_1}}^*(\mathbf{r}_1) \cdots \varphi_{l_{\alpha_n}}(\mathbf{r}_n) \\ \times \sum (l_1, \ldots, l_n) \lambda_{l_{n+1}} \cdots \lambda_{l_{n+p}}, \end{aligned} \tag{D6}$$

where the symbol $\sum(l_1, \ldots, l_n)$ indicates a summation over all distinct integers $l_{n+1}, \ldots, l_{n+p}$ all of which differ from $l_1, \ldots, l_n$.

In order to shorten these equations, let us set

$$\begin{aligned} f_n(\{\mathbf{r}_j\}, \{l_j\}) = \varphi_{l_1}(\mathbf{r}_1) \cdots \varphi_{l_n}(\mathbf{r}_n) \lambda_{l_1} \cdots \lambda_{l_n} \\ \times \sum P_\alpha (-1) \varphi_{l_{\alpha_1}}^*(\mathbf{r}_1) \cdots \varphi_{l_{\alpha_n}}^*(\mathbf{r}_n). \end{aligned} \tag{D7}$$

Introducing Eqs. (D6) and (D7) into Eq. (D1) one arrives at

$$\begin{aligned} G_n^U[\{\mathbf{r}_j\}] = \sum_{p=0}^{\infty} \frac{(-1)^p}{p!} \\ \times \sum_{l_1, \ldots, l_n} f_n(\{\mathbf{r}_j\}, \{l_n\}) \sum (l_1, \ldots, l_n) \lambda_{l_{n+1}} \cdots \lambda_{l_{n+p}}. \end{aligned} \tag{D8}$$

Thus, by exchanging the order of the first two summations, one gets

$$\begin{aligned} G_n^U[\{\mathbf{r}_j\}] = \sum_{l_1, \ldots, l_n} f_n(\{\mathbf{r}_j\}, \{\mathbf{l}_j\}) \\ \times \sum_{p=0}^{\infty} \frac{(-1)^p}{p!} \sum (l_1, \ldots, l_n) \lambda_{l_{n+1}} \cdots \lambda_{l_{n+p}}. \end{aligned} \tag{D9}$$

This procedure is legitimate if the $(n+1)$-fold series of the absolute values, say Q,

$$\begin{aligned} Q = \sum_{p, l_1, \ldots, l_n} |f_n(\{\mathbf{r}_j\}, \{\mathbf{l}_j\})| \frac{1}{p!} \\ \times \sum (l_1, \ldots, l_n) \lambda_{l_{n+1}} \cdots \lambda_{l_{n+p}}, \end{aligned} \tag{D10}$$

is convergent. It is known[23] that the summation order does not matter for a positive series. Summing Eq. (D10) first with respect to p, we realize that

$$\begin{aligned} \sum_{p=0}^{\infty} \frac{1}{p!} \sum (l_1, \ldots, l_n) \lambda_{l_{n+1}} \cdots \lambda_{l_{n+p}} \\ = \prod_{j \neq l_1, \ldots, l_n} (1 + \lambda_j) \leq \prod_{j=1}^{\infty} (1 + \lambda_j). \end{aligned} \tag{D11}$$

The last infinite product converges because[13]

$$\sum_j \lambda_j = \int_U C(\mathbf{r}, \mathbf{r}) \, d\mathbf{r} < \infty, \tag{D12}$$

in turn coming from the covariance continuity.

Then, summing Eq. (D10) with respect to $l_1, \ldots, l_n$, and taking definition (D7) into account, one obtains

$$\begin{aligned} Q \leq \prod_{j=1}^{\infty} (1 + \lambda_j) \sum P_\alpha \\ \times \sum_{l_1, \ldots, l_n} \lambda_{l_1} |\varphi_{l_1}(\mathbf{r}_1) \varphi_{l_1}^*(\mathbf{r}_{\alpha_1})| \cdots \lambda_{l_n} |\varphi_{l_n}(\mathbf{r}_n) \varphi_{l_n}^*(\mathbf{r}_{\alpha_n})|, \\ Q \leq \prod_{j=1}^{\infty} (1 + \lambda_j) \sum P_\alpha \prod_{j=1}^{n} \Big(\sum_l \lambda_l |\varphi_l(\mathbf{r}_j) \varphi_l^*(\mathbf{r}_{\alpha_j})| \Big). \end{aligned} \tag{D13}$$

From Eq. (D12) and the absolute convergence of expression (D5), one deduces that Q is bounded. This establishes the validity of Eq. (D9).

Now, just as for expression (D11), it should be noted that

$$\begin{aligned} \sum_{p=0}^{\infty} \frac{(-1)^p}{p!} \sum (l_1, \ldots, l_n) \lambda_{l_{n+1}} \cdots \lambda_{l_{n+p}} \\ = \prod_{l \neq l_1, \ldots, l_n} (1 - \lambda_l), = \prod_{l=1}^{\infty} (1 - \lambda_l) \prod_{k=1}^{n} (1 - \lambda_{l_k})^{-1}. \end{aligned} \tag{D14}$$

Introducing Eqs. (D7) and (D14) into Eq. (D9), we get

$$\begin{aligned} G_n^U[\{\mathbf{r}_j\}] \\ = \prod_{l=1}^{\infty} (1 - \lambda_l) \sum_{l_1, \ldots, l_n} \frac{\lambda_{l_1} \cdots \lambda_{l_n}}{(1 - \lambda_{l_1}) \cdots (1 - \lambda_{l_n})} \\ \times \varphi_{l_1}(\mathbf{r}_1) \cdots \varphi_{l_n}(\mathbf{r}_n) \sum P_\alpha (-1) \varphi_{l_1}^*(\mathbf{r}_{\alpha_1}) \cdots \varphi_{l_n}^*(\mathbf{r}_{\alpha_n}) \\ = \prod_{l=1}^{\infty} (1 - \lambda_l) \sum P_\alpha (-1) \prod_{j=1}^{n} \Big(\sum_l \frac{\lambda_l}{1 - \lambda_l} \varphi_l(\mathbf{r}_j) \varphi_l^*(\mathbf{r}_{\alpha_j}) \Big). \end{aligned} \tag{D15}$$

Next, we suppose that no λ_l is one. So it can easily be seen that the root $g(\mathbf{r}, \mathbf{r}')$ of Eq. (D4) admits the following Mercer expansion, similar to expansion (D5), and converges absolutely and uniformly over $U \times U$:

$$g(\mathbf{r}, \mathbf{r}') = \sum_l [\lambda_l / (1 - \lambda_l)] \varphi_l(\mathbf{r}) \varphi_l^*(\mathbf{r}'). \tag{D16}$$

Now the result claimed in Eq. (D3) directly derives from Eqs. (D15) and (D16).

In order to complete the proof, we must justify Eq. (D6).

Introducing expansion (D5) into the expression (D2) of the cpd and integrating over U^p, we see that

$$\begin{aligned} \int_{U^p} P_{n+p}[\{\mathbf{r}_j\}] \, d\mathbf{r}_{n+1} \cdots d\mathbf{r}_{n+p} = \int_{U^p} \sum P_\alpha (-1) \\ \times \sum_{l_1, \ldots, l_{n+p}} \lambda_{l_1} \cdots \lambda_{l_{n+p}} \varphi_{l_1}(\mathbf{r}_1) \cdots \varphi_{l_{n+p}}(\mathbf{r}_{n+p}) \\ \times \varphi_{l_{\alpha_1}}^*(\mathbf{r}_1) \cdots \varphi_{l_{\alpha_{n+p}}}^*(\mathbf{r}_{n+p}) \, d\mathbf{r}_{n+1} \cdots d\mathbf{r}_{n+p}. \end{aligned} \tag{D17}$$

Notice that

$$\sum P_\alpha(-1)\,\varphi^*_{l_{\alpha_1}}(\mathbf{r}_1)\cdots\varphi^*_{l_{\alpha_{n+p}}}(\mathbf{r}_{n+p})$$
$$=\sum P_\alpha(-1)\,\varphi^*_{l_1}(\mathbf{r}_{\alpha_1})\cdots\varphi^*_{l_{n+p}}(\mathbf{r}_{\alpha_{n+p}})$$
$$=\begin{vmatrix}\varphi^*_{l_1}(\mathbf{r}_1), & \varphi^*_{l_1}(\mathbf{r}_2), & \cdots, & \varphi^*_{l_1}(\mathbf{r}_{n+p})\\ \varphi^*_{l_2}(\mathbf{r}_1), & \varphi^*_{l_2}(\mathbf{r}_2), & \cdots, & \varphi_{l_2}(\mathbf{r}_{n+p})\\ & & \cdots, & \\ \varphi^*_{l_{n+p}}(\mathbf{r}_1), & \varphi^*_{l_{n+p}}(\mathbf{r}_2), & \cdots, & \varphi^*_{l_{n+p}}(\mathbf{r}_{n+p})\end{vmatrix}. \quad \text{(D18)}$$

So if two l_i are equal, this determinant has two identical rows, and thus is zero. Therefore, in Eq. (D17), we may restrict the series to the sets of distinct $l_1,\ldots,l_{n+p}$.

Next we show that term by term integration is allowed in Eq. (D17); for any fixed permutation $\alpha_1,\ldots,\alpha_{n+p}$, the series $\sum_{l_1\neq\cdots\neq l_{n+p}}$ of the integrated modulus is bounded over by

$$R=\sum_{l_1,\ldots,l_{n+p}}\lambda_{l_1}\cdots\lambda_{l_{n+p}} \times|\,\varphi_{l_1}(\mathbf{r}_1)\cdots\varphi_{l_n}(\mathbf{r}_n)\,\varphi^*_{l_{\alpha_1}}(\mathbf{r}_1)\cdots\varphi^*_{l_{\alpha_n}}(\mathbf{r}_n)\,|, \quad \text{(D19)}$$

because of the Schwarz inequality

$$\int_U|\,\varphi_{l_{n+k}}(\mathbf{r}_{n+k})\,\varphi^*_{l_{\alpha_{n+k}}}(\mathbf{r}_{n+k})\,|\,d\mathbf{r}_{n+k} \leq\Big(\int_U|\,\varphi_{l_{n+k}}(\mathbf{r})\,|^2\,d\mathbf{r}\int_U|\,\varphi_{l_{\alpha_{n+k}}}(\mathbf{r})\,|^2\,d\mathbf{r}\Big)^{1/2}$$

and because all $\varphi_l(\mathbf{r})$ are normalized. Now it appears that the series in Eq. (D19) can be summed separately in all its arguments $l_1,\ldots,l_{n+p}$. It then becomes the product of $(n+p)$ factors some of which are of the form

$$a=\sum_l\lambda_l\,|\,\varphi_l(\mathbf{r}_i)\,\varphi^*_l(\mathbf{r}_j)\,|$$

and others of the form

$$b=\sum_l\lambda_l\,|\,\varphi_l(\mathbf{r}_k)\,|.$$

Expansion a is convergent since expansion (D5) is absolutely convergent. Expansion b converges because the $\varphi_l(r)$ are a CON and because of Eq. (D12). Hence R is bounded above. The term by term integration of expression (D17) then gives

$$\int_{U^p}P_{n+p}[\{\mathbf{r}_j\}]\,d\mathbf{r}_{n+1}\cdots d\mathbf{r}_{n+p} =\sum_{l_1\neq\cdots\neq l_{n+p}}\lambda_{l_1}\cdots\lambda_{l_{n+p}}\,\varphi_{l_1}(\mathbf{r}_1)\cdots\varphi_{l_n}(\mathbf{r}_n) \times\sum P_\alpha(-1)\,\varphi^*_{l_{\alpha_1}}(\mathbf{r}_1)\cdots\varphi^*_{l_{\alpha_n}}(\mathbf{r}_n) \times\prod_{k=1}^{p}\int_U\varphi_{l_{n+k}}(\mathbf{r}_{n+k})\,\varphi^*_{l_{\alpha_{n+k}}}(\mathbf{r}_{n+k})\,d\mathbf{r}_{n+k}. \quad \text{(D20)}$$

Suppose that some $n+k$ exists such that $\alpha_{n+k}\neq n+k$, then $l_{\alpha_{n+k}}\neq l_{n+k}$, since the l_i are distinct. Thus, with the CON $\varphi_l(r)$

$$\int_U\varphi_{l_{n+k}}(\mathbf{r})\,\varphi^*_{l_{\alpha_{n+k}}}(\mathbf{r})\,d\mathbf{r}=0.$$

Therefore in the summation $\sum P_\alpha(-1)$, we only have to consider the part $\sum P^n_\alpha(-1)$ of the permutations $\alpha_1,\ldots,\alpha_n,n+1,\ldots,n+p$ that leave $(n+1),\ldots,(n+p)$ unchanged. And for these, integration always gives one. The point now is that $\alpha_1,\ldots,\alpha_n$ is a permutation of $1,\ldots,n$, and that is why Eq. (D20) becomes

$$\int_{U^p}P_{n+p}[\{\mathbf{r}_j\}]\,d\mathbf{r}_{n+1}\cdots d\mathbf{r}_{n+p} =\sum_{l_1\neq\cdots\neq l_{n+p}}\lambda_{l_1}\cdots\lambda_{l_{n+p}}\,\varphi_{l_1}(\mathbf{r}_1)\cdots\varphi_{l_n}(\mathbf{r}_n) \times\sum P_\alpha(-1)\,\varphi^*_{l_1}(\mathbf{r}_{\alpha_1})\cdots\varphi^*_{l_n}(\mathbf{r}_{\alpha_n}). \quad \text{(D21)}$$

Now it is enough to observe, as for expression (D18), that the condition of distinct $l_1,\ldots,l_n$ does not actually change the value of the series, since it only eliminates null terms. Then Eq. (D21) provides result (D6) previously stated, and thus achieves the proof of Eq. (D3).

Probability meaning

We now turn to the probability meaning of these equations.

The system of functions $P_n[\{\mathbf{r}_j\}]$ expressed in Eq. (D2) may be viewed as probability densities of a pp only if two obvious conditions are fulfilled.

Condition 1: $P_n[\{\mathbf{r}_j\}]\geq 0, \quad \forall n,\forall\mathbf{r}_1,\ldots,\mathbf{r}_n.$ (D22)

Now expression (D2) is equivalent to

$$P_n[\{\mathbf{r}_j\}]=\begin{vmatrix}C(\mathbf{r}_1,\mathbf{r}_1), & \cdots, & C(\mathbf{r}_1,\mathbf{r}_n)\\ C(\mathbf{r}_2,\mathbf{r}_1), & \cdots, & C(\mathbf{r}_2,\mathbf{r}_n)\\ & \cdots & \\ & \cdots & \\ C(\mathbf{r}_n,\mathbf{r}_1), & \cdots, & C(\mathbf{r}_n,\mathbf{r}_n)\end{vmatrix}. \quad \text{(D23)}$$

In Ref. 24, it is shown that for such $P_n[\{\mathbf{r}_j\}]$, condition (D22) is fulfilled if and only if

$C(\mathbf{r},\mathbf{r}')$ is nonnegative definite. (D24)

Such is the case in our paper because $C(\mathbf{r},\mathbf{r}')$ is defined either as the covariance function of a stochastic process, or as the sum of an expansion of the kind (D5), with a CON of functions $\varphi_l(r)$ and nonnegative coefficients λ_l,

$$\lambda_l\geq 0 \quad \forall l,\forall U. \quad \text{(D25)}$$

Condition 2: The system of functions $G^U_n[\{\mathbf{r}_j\}]$ that derive from $P_n[\{\mathbf{r}_j\}]$ through relation (D1) actually represents exclusive probability densities only if

$$G^U_n[\{\mathbf{r}_j\}]\geq 0, \quad \forall n,\forall\mathbf{r}_1,\ldots\mathbf{r}_n,\forall U, \quad \text{(D26)}$$

or, in our case where $G^U_n[\{\mathbf{r}_j\}]$ is given by Eq. (D3), if

$$\prod_{l=1}^{\infty}(1-\lambda_l)\begin{vmatrix}g(\mathbf{r}_1,\mathbf{r}_1), & \cdots, & g(\mathbf{r}_1,\mathbf{r}_n)\\ g(\mathbf{r}_2,\mathbf{r}_1), & \cdots, & g(\mathbf{r}_2,\mathbf{r}_n)\\ & \cdots & \\ g(\mathbf{r}_n,\mathbf{r}_1), & \cdots, & g(\mathbf{r}_n,\mathbf{r}_n)\end{vmatrix}\geq 0, \quad \forall n,\mathbf{r}_1\cdots\mathbf{r}_{n1},\forall U. \quad \text{(D27)}$$

Condition (D27) implies that the n-order determinant built on $g(\mathbf{r},\mathbf{r}')$ is of constant sign. Therefore $g(\mathbf{r},\mathbf{r}')$ is either nonnegative definite or nonpositive definite. According to expansion (D16), and taking Eq. (D25) into account, this is equivalent either to

$\forall l,\ \lambda_l\leq 1, \quad g(\mathbf{r},\mathbf{r}')$ nonnegative definite (D28)

or to

$$\forall l,\ \lambda_l \geqslant 1, \quad g(\mathbf{r},\mathbf{r}') \text{ nonpositive definite.} \tag{D29}$$

But hypothesis (D29) is not acceptable because the eigenvalues λ_l decrease towards zero [see Eq. (D12)]. Moreover remember that $\lambda_l \neq 1$ appeared in the proof of relation (D3). Thus relations (D27) and (D25) can be grouped together into

$$0 \leqslant \lambda_l < 1, \quad \forall l, \forall U. \tag{D30}$$

Conversely it can be shown that if $G_n^U[\{\mathbf{r}_j\}]$ derives from $P_n[\{\mathbf{r}_j\}]$ by Eq. (D1), and if relations (D22) and (D26) hold, then there exists an actual pp whose distribution is stipulated either by the system $P_n[\{\mathbf{r}_j\}]$ or by that of $G_n^U[\{\mathbf{r}_j\}]$. Therefore relation (D30) is a necessary and sufficient condition for a probability interpretation of Eqs. (D2) and (D3).

Huang and Johnson[25] demonstrated the intuitive property of the eigenvalues that if the λ_l are numbered in decreasing order, then

$$\forall l \quad \lambda_l(U_n) \geqslant \lambda_l(U_l) \quad \text{if } U_1 \supset U_2. \tag{D31}$$

Now for the applications we have in mind, namely the distributions of the "emission" and the detection fermion-point processes, the particles are enclosed in a fixed volume V. The function $C(\mathbf{r},\mathbf{r}')$ is defined over $V \times V$ by the expansion

$$C(\mathbf{r},\mathbf{r}') = \sum_k \langle n_k \rangle\ e^{2i\pi\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')}/V, \tag{21}$$

where the vectors $\mathbf{k}$ are chosen in such a way that the diverse $e^{2i\pi\mathbf{k}\cdot\mathbf{r}}$ form a CON over V. The number $\langle n_k \rangle$ which is the mean number of fermions in mode $\mathbf{k}$, is positive. Hence $C(\mathbf{r},\mathbf{r}')$ is positive definite over $V \times V$, and thus relation (D25) holds for all volumes $U \subset V$. On the other hand, $\langle n_k \rangle$ is certainly less than one. With condition (D31), this shows that relation (D30) is valid inside volume V.

Therefore the functions $P_n[\{\mathbf{r}_j\}]$ expressed by Eq. (D2) actually represent the cpd of a point process. For that pp the epd are given by Eqs. (D3) and (D4).

[1] C. Benard, Phys. Rev. A **2**, 2140 (1970).
[2] R. J. Glauber, in *Quantum Optics*, edited by S. M. Kay and A. Maitland (Academic, New York, 1970).
[3] B. R. Mollow, Phys. Rev. **175**, 1555 (1968).
[4] O. Macchi, C.R. Acad. Sci. (Paris) **272**, 437 (1971).
[5] In particular, if the product set $A = (A_1 \times \ldots \times A_n)$, $A \subset V^n$ is the disjoint union $A = \Sigma_{j=1}^{\infty} A^j$ of certain product sets $A^j = A_1^j \times \ldots \times A_n^j$, then H_n satisfies $H_n(A_1, 1, \ldots, A_n, 1) = \Sigma_{j=1}^{\infty} H_n(A_1^j, 1, \ldots, A_n^j, 1)$.
[6] O. Macchi, IEEE Trans. Inf. Theory **17**, 2 (1971).
[7] R. J. Glauber, in *Optique et electronique quantique*, edited by C. de Witt et al. (Gordon and Breach, New York, 1964).
[8] C. Cohen Tannoudji, *Compléments de Mécanique Quantique* (published by the "Laboratoire de Physique de l'Ecole Normal Supérieure", Paris, 1967), Vol. 1, p. 265 and following.
[9] B. Picinbono, C. Bendjaballah, and J. Pouget, J. Math. Phys. **11**, 2166 (1970).
[10] M. J. Beran and G. B. Parrent, *Theory of Partial Coherence* (Prentice-Hall, New Jersey, 1964).
[11] L. Mandel and E. Wolf, Rev. Mod. Phys. **37**, 231 (1965).
[12] See Ref. 4 where the proof is given when the argument x of the random function $X(\mathbf{x})$ is an element of R. The whole proof extends word for word for $\mathbf{x}$ in $\mathbf{R}^n$. Here $\mathbf{x}$ is in $\mathbf{R}^3$.
[13] R. Courant and D. Hilbert, *Methods of Mathematical Physics* I, (Interscience, New York, 1953), p. 122.
[14] W. B. Davenport and W. L. Root, *Introduction to Random Signals and Noise* (McGraw-Hill, New York, 1958), p. 96.
[15] L. Mandel, Phys. Rev. **144**, 1071 (1966).
[16] C. Benard, Thesis (Université Paris-Sud, March 1972).
[17] D. Kastler, *Introduction à l'électrodynamique quantique* (Dunod, Paris, 1961), p. 74.
[18] See in Ref. 17, p. 51.
[19] See in Ref. 17, p. 57.
[20] See in Ref. 17, p. 59.
[21] See in Ref. 17, 62 and following.
[22] L. H. Loomis, *An Introduction to Abstract Harmonic Analysis* (Van Nostrand, New York, 1953), p. 148.
[23] L. Schwartz, *Méthodes mathématiques pour les sciences physiques* (Hermann, Paris, 1961), p. 19.
[24] E. F. Beckenbach and R. Bellan, *Inequalities* (Springer-Verlag, New York, 1965).
[25] R. Y. Huang and P. A. Johnson, IEEE Trans. Inf. Theory **19**, 84 (1963).

A spinor approach to the generalized singular electromagnetic field theory of Bel, Lapiedra, and Montserrat

Joseph Zund and William Maher*

Department of Mathematical Sciences, New Mexico State University, Las Cruces, New Mexico

(Received 26 April 1971; Revised 9 July 1971)

A spinor approach to the generalized singular electromagnetic field of Bel, Lapiedra, and Montserrat is presented. We include new proofs of the basis existence theorems of the field theory using the method of spin coefficients, and interpretations of these theorems in terms of the optical scalars associated with a null geodesic congruence. We also exhibit general classes of solutions of the generalized singular electromagnetic field equations and classify them according to the optical properties of their associated null geodesic congruences.

1. INTRODUCTION

Let $\mathfrak{M}$ be a four-dimensional differentiable manifold having a Riemannian metric g_{ab} of hyperbolic normal signature, and let Γ denote a congruence of curves in $\mathfrak{M}$ having null tangent vector l^a.

A source-free singular electromagnetic field associated with Γ in the sense that the propagation vector of the field l^a is defined by a skew-symmetric tensor F_{ab} such that

$$l^a F_{ab} = l^a \overset{*}{F}_{ab} = 0, \tag{1.1}$$

$$\nabla_a F^{ab} = 0, \tag{1.2}$$

$$\nabla_a \overset{*}{F}^{ab} = 0. \tag{1.3}$$

In these equations, we have denoted the covariant derivative by ∇_b and dual tensors are constructed using the Levi-Civita permutation tensor ϵ_{abcd}, e.g., $\overset{*}{F}_{ab} = \frac{1}{2}\epsilon_{abcd}F^{cd}$.

The existence theorems for this system are those of Mariot[1] and Robinson,[2] which state that there exist solutions if and only if the associated null congruence Γ is geodesic and shear free. Bel, Lapiedra, and Montserrat[3] contend that it is essential for the proper consideration of reflection problems that one be able to associate electromagnetic radiation with a shearing geodesic congruence, and they argue that a correct formulation of electromagnetic optics in general relativity must be phrased entirely in terms of singular electromagnetic fields. Such an analysis, together with a desire to generalize the theorem of Mariot and Robinson, led them to introduce the notion of a generalized singular electromagnetic field.

A generalized singular electromagnetic field associated with the null congruence Γ is defined by a pair of skew-symmetric tensors (F_{ab}, G_{ab}) such that

$$l^a F_{ab} = l^a \overset{*}{F}_{ab} = 0, \tag{1.4}$$

$$\nabla_a \overset{*}{F}^{ab} = 0, \tag{1.5}$$

$$\nabla_a G^{ab} = 0,$$
$$\text{where } G^{ab} \equiv \xi F^{ab} + \eta \overset{*}{F}^{ab} \text{ with } \xi \neq 0. \tag{1.6}$$

For brevity, we will refer to such a field as a BLM field and (1.4)-(1.6) as the BLM field equations.

The spinor and spin-coefficient formulations of the BLM theory is presented in Sec. 2 of this paper. In Sec. 3, we present a proof of the basic existence theorems for this field employing the spin-coefficient formalism, and we relate the conditions of the theorems to the optical scalars associated with Γ and the Bel–Petrov type of the Weyl tensor. In Sec. 4, we give a collection of general classes of solutions of the BLM system which is classified in terms of the optical scalars associated with Γ.

The spinor and spin-coefficient formalism of Newman and Penrose[4] will be assumed known throughout our discussion, and we follow the Battelle Rencontres spinor notation of Penrose.[5]

2. THE SPINOR FORMULATION OF THE BLM SYSTEM

The spinor formalism permits us to immediately translate Eqs. (1.5) and (1.6) into the form

$$\nabla_{AA'} {}^*F^{AA'BB'} = 0, \tag{2.1}$$

$$\nabla_{AA'} G^{AA'BB'} = 0, \tag{2.2}$$

where

$$G^{AA'BB'} = \xi F^{AA'BB'} + \eta {}^*F^{AA'BB'} \tag{2.3}$$

and $\xi \neq 0$.

By using the standard spinor expressions for F_{ab} and $\overset{*}{F}_{ab}$, we may rewrite (2.3) in terms of the symmetric spinor ϕ_{AB} as

$$G_{AA'BB'} = \zeta \phi_{AB} \epsilon_{A'B'} + \bar{\zeta}\bar{\phi}_{A'B'}\epsilon_{AB}, \tag{2.4}$$

where

$$\zeta \overset{\text{DEF}}{=} \zeta - i\eta \tag{2.5}$$

and

$$\text{Re}\,\zeta \neq 0. \tag{2.6}$$

Thus (2.1) and (2.2) may equivalently be written in the form

$$\nabla_B^{A'} \bar{\phi}_{A'B'} - \nabla_{B'}^{A} \phi_{AB} = 0, \tag{2.7}$$

$$\nabla_B^{A'}(\bar{\zeta}\bar{\phi}_{A'B'}) + \nabla_{B'}^{A}(\zeta \phi_{AB}) = 0. \tag{2.8}$$

Theorem 1: A BLM field associated with congruence Γ is determined by the complex functions $\{\phi, \zeta\}$ with $\text{Re}\,\zeta \neq 0$, such that

$$F_{AA'BB'} = \phi o_A o_B \epsilon_{A'B'} + \bar{\phi}\bar{o}_{A'}\bar{o}_{B'}\epsilon_{AB}, \tag{2.9}$$

$$G_{AA'BB'} = \zeta \phi o_A o_B \epsilon_{A'B'} + \bar{\phi}\bar{\zeta}\bar{o}_{A'}\bar{o}_{B'}\epsilon_{AB}, \tag{2.10}$$

and

$$\nabla_B^{A'}(\phi o_A o_B) = -\frac{1}{2\text{Re}\zeta}(\phi o_A o_B \nabla_B^{A'}\zeta + \bar{\phi}\bar{o}_{A'}\bar{o}_{B'}\nabla_B^{A'}\zeta). \tag{2.11}$$

Proof: The set of equations (2.6)-(2.8) is equivalent

to the single spinor equation

$$\nabla^A_{B'}\phi_{AB} = -1/2\mathrm{Re}\zeta(\phi_{AB}\nabla^A_{B'}\zeta + \bar{\phi}_{A'B'}\nabla^{A'}_{B}\bar{\zeta}), \tag{2.12}$$

where $\mathrm{Re}\zeta \neq 0$.

Clearly Eqs. (2.7) and (2.8) imply (2.12), for if we expand (2.8), we obtain

$$\bar{\phi}_{A'B'}\nabla^{A'}_{B}\bar{\zeta} + \phi_{AB}\nabla^A_{B'}\zeta + \bar{\zeta}\nabla^{A'}_{B}\bar{\phi}_{A'B'} + \zeta\nabla^A_{B'}\phi_{AB} = 0. \tag{2.8'}$$

Thus, the proof of (2.11) merely involves the observation that the difference of the product of $\bar{\zeta}$ and (2.7) and (2.8) is Eq. (2.12). Conversely Eq. (2.12) implies (2.7) and (2.8), since they are precisely the real and imaginary parts of (2.12). Equations (2.9) and (2.10) now follow since by virtue of (1.4) we may write $\phi_{AB} = \phi o_A o_B$, where l_a corresponds to $o_A\bar{o}_{A'}$.

Theorem 2: The spin coefficient formulation of the field equations reduces to

$$\kappa = 0, \tag{2.13}$$

$$D\phi = (\rho - 2\epsilon)\phi + \bar{\sigma}\bar{\phi}, \tag{2.14}$$

$$\delta\phi = (\tau - 2\beta)\phi + p, \tag{2.15}$$

$$D\zeta = -2\bar{\sigma}\bar{\phi}\phi^{-1}\mathrm{Re}\zeta, \tag{2.16}$$

$$\delta\zeta = -\phi^{-1}(p\mathrm{Re}\zeta + iq), \tag{2.17}$$

where p and q are real functions which depend on Γ and the initial conditions, and $\mathrm{Re}\zeta \neq 0$.

Proof: The spin coefficient resolution of Eq. (2.11) consists of the set of projections of this equation on the tetrad spinors $o^B\bar{o}^{B'}$, $o^B\bar{\iota}^{B'}$, $\iota^B\bar{o}^{B'}$, and $\iota^B\bar{\iota}^{B'}$. If one expands Eq. (2.11) to

$$\phi o_A o_B\nabla^A_{B'}\zeta + \bar{\phi}\bar{o}_{A'}o_{B'}\nabla^{A'}_{B}\bar{\zeta} + 2\mathrm{Re}\zeta(o_A o_B\nabla^A_{B'}\phi + \phi o_A\nabla^A_{B'}o_B + \phi o_B\nabla^A_{B'}o_A) = 0, \tag{2.18}$$

then the determination of the projections may be easily achieved through the following stages.

First, multiplying (2.18) by o^B, one obtains

$$\bar{\phi}\bar{o}_{A'}\bar{o}_{B'}o^B\nabla^{A'}_{B}\bar{\zeta} + 2\phi\mathrm{Re}\zeta\cdot o_A o^B\nabla^A_{B'}o_B = 0, \tag{2.19}$$

which may be expressed as

$$-\bar{\phi}\bar{o}_{B'}D\bar{\zeta} + 2\phi\mathrm{Re}\zeta\cdot o_A o^B\{[(-\kappa\iota_B + \epsilon o_B)\iota^A + (\rho\iota_B - \alpha o_B)o^A]\bar{\iota}_{Y'} + [(\sigma\iota_B - \beta o_B)\iota^A + (\gamma o_A - \tau\iota_A)o^A]\bar{o}_{B'}\} = 0. \tag{2.20}$$

This equation immediately reduces to

$$-\bar{\phi}\bar{o}_{B'}D\bar{\zeta} + 2\phi\mathrm{Re}\zeta\cdot\{\kappa\bar{\iota}_{B'} - \sigma\bar{o}_{B'}\} = 0. \tag{2.21}$$

Now, multiplying (2.21) by $\bar{o}^{B'}$, we obtain

$$\kappa = 0,$$

while multiplying (2.21) by $\bar{\iota}^{B'}$ yields

$$\bar{\phi}D\bar{\zeta} + \sigma\phi(\zeta + \bar{\zeta}) = 0. \tag{2.22}$$

Similarly, multiplying Eq. (2.18) by $\iota^B\bar{o}^{B'}$, one obtains

$$-\phi D\zeta + (\zeta + \bar{\zeta})(-D\phi - \phi\iota^B Do_B + \phi o^{B'}\nabla^A_{B'}o_A) = 0. \tag{2.23}$$

This equation can be reduced to

$$-\phi D\zeta + (\zeta + \bar{\zeta})[-D\phi + (\rho - 2\epsilon)\phi] = 0. \tag{2.24}$$

Finally, multiplying (2.18) by $\iota^B\bar{o}^{B'}$, one obtains

$$\phi\delta\zeta + \bar{\phi}\bar{\delta}\bar{\zeta} + [\delta\phi + (2\beta - \tau)\phi](\zeta + \bar{\zeta}) = 0. \tag{2.25}$$

Thus, the spin coefficient resolution of Eq. (2.10) is given by (2.13), (2.22), (2.24), and (2.25).

We complete the proof by showing that these four equations are equivalent to the set (2.13)-(2.17). Equation (2.25) with the initial assumption that $(\zeta + \bar{\zeta}) \neq 0$, implies that the combination $\delta\phi + (2\beta - \tau)\phi$ is real. We can, therefore, introduce the real function p and split (2.25) into the equivalent pair (2.15), and

$$\mathrm{Re}(\phi\delta\zeta) = -p\mathrm{Re}\zeta. \tag{2.26}$$

Since Eq. (2.26) imposes a condition on $\mathrm{Re}(\phi\delta\zeta)$ only, we can introduce the real function q and rewrite (2.26) as Eq. (2.17).

In conclusion, if one now considers the pair (2.16) and (2.24), together with the initial assumption $\mathrm{Re}\zeta \neq 0$, Eq. (2.14) immediately follows. Thus we have the previously stated equivalence. As a direct consequence of Eq. (2.22), we have:

Corollary: The shear of the congruence Γ and the function ζ are related by

$$\sigma\bar{\sigma} = \frac{D\zeta\cdot D\bar{\zeta}}{(\zeta + \bar{\zeta})^2}. \tag{2.27}$$

Equation (2.27) illustrates an interesting aspect of the BLM theorem. In particular if $\eta = 0$ and we choose $D = \partial/\partial r$, then (2.27) shows that

$$\xi = C\exp(2\int|\sigma|dr), \tag{2.28}$$

where C is a constant. Solutions of the BLM equations corresponding to this possibility are indicated in Sec. 4 under the cases I.1.1 and I.1.2.

3. EXISTENCE OF SOLUTIONS TO THE BLM FIELD EQUATIONS

The basic existence theorem for the BLM field is the following:

Theorem 3: The BLM equations (2.13)-(2.17) admit a solution if and only if the associated congruence Γ is geodesic.

Proof: The necessity is a direct consequence of Eq. (2.13).

To establish the sufficiency we assume $\kappa = 0$, and to prove that the remaining set of equations is integrable. Without loss of generality we may assume that our tetrad is parallelly propagated along Γ so that $\epsilon = \pi = 0$. Equations (2.14) and (2.15) are then integrable if and only if

$$(\delta D - D\delta)\phi = (\bar{\alpha} + \beta)D\phi - \sigma\bar{\delta}\phi + \bar{\rho}\delta\phi. \tag{3.1}$$

It will therefore be sufficient to show that the commutator constructed directly from the field equations assumes precisely the form

$$(\delta D - D\delta)\phi = (\bar{\alpha} + \beta)(\rho\phi + \bar{\sigma}\bar{\phi}) - \sigma\bar{\delta}\phi + \bar{\rho}[(\tau - 2\beta)\phi + p] \tag{3.2}$$

for appropriately determined functions p.

From Eqs. (2.14) and (2.15), respectively, it follows that $\delta D\phi$ and $D\delta\phi$ are

$$\delta D\phi = \phi\delta\rho + \rho[(\tau - 2\beta)\phi = p] + \bar{\phi}\delta\bar{\sigma} + \bar{\sigma}\delta\bar{\phi} \tag{3.3}$$

and

$$D\delta\phi = (D\tau - 2D\beta)\phi + (\tau - 2\beta)(\rho\phi + \bar{\sigma}\bar{\phi}) + Dp. \tag{3.4}$$

Hence

$$(\delta D - D\delta)\phi = \phi\delta\rho + \phi\delta\sigma + \rho p - (D\tau - 2D\beta)\phi - (\tau - 2\beta)\phi - Dp. \tag{3.5}$$

This expression may be conveniently expanded using the following Newman–Penrose equations[6]:

$$\delta\rho = [\bar{\delta}\sigma + \rho(\bar{\alpha} + \beta) - \sigma(3\alpha - \bar{\beta}) + (\rho - \bar{\rho})\tau - \Psi_1 + \phi_{01}],$$
$$-D\tau = -(\tau\rho + \bar{\tau}\sigma + \Psi_1 + \Phi_{01}),$$
$$2D\beta = 2(\alpha\sigma + \bar{\rho}\beta + \Psi_1).$$

Hence the commutator constructed directly from the field equations is

$$(\delta D - D\delta) = \phi\bar{\delta}\sigma + \bar{\phi}\delta\bar{\sigma} + \phi[\rho(\bar{\alpha} + \beta) - \alpha\sigma + \sigma\bar{\beta} + 2\bar{\rho}\beta - \tau\rho - \bar{\tau}\sigma] + \rho p + \bar{\sigma}\delta\bar{\phi} - (\tau - 2\beta)\bar{\sigma}\bar{\phi} - Dp. \tag{3.6}$$

Thus Eqs. (2.14) and (2.15) are integrable if

$$(\bar{\alpha} + \beta)(\rho\phi + \bar{\sigma}\bar{\phi}) - \sigma\bar{\delta}\phi + \bar{\rho}[(\tau - 2\beta)\phi + p] - \{\phi\bar{\delta}\sigma + \bar{\phi}\delta\bar{\sigma} + \phi[\rho(\bar{\alpha} + \beta) - \alpha\sigma + \sigma\bar{\beta} + 2\bar{\rho}\beta - \bar{\tau}\rho] + \rho p + \bar{\sigma}\delta\bar{\phi} - (\tau - 2\beta)\bar{\sigma}\bar{\phi} - Dp\} = 0. \tag{3.7}$$

Equation (3.7) obviously simplifies to

$$\phi\bar{\delta}\sigma + (\bar{\beta} - \alpha - \bar{\tau})\sigma\phi + \sigma\bar{\delta}\phi + \bar{\phi}\delta\bar{\sigma} + (\beta - \bar{\alpha} - \tau)\sigma\phi + \bar{\sigma}\delta\bar{\phi} + (\rho + \bar{\rho})p - Dp = 0. \tag{3.8}$$

Coordinates may be chosen so that $D = \partial/\partial r$. Thus one may rewrite and integrate Eq. (3.8) in the form

$$p = 2f^{-1}\int f\mathrm{Re}[\bar{\delta}(\sigma\phi) + (\bar{\beta} - \alpha - \bar{\tau})\sigma\phi]dr, \tag{3.9}$$

where

$$f \equiv \exp(-2\int \mathrm{Re}\rho dr). \tag{3.10}$$

Clearly the p given in Eq. (3.9) is real, and hence (2.14) and (2.15) are integrable for this choice of p.

Equations (2.16) and (2.17) for ζ may be handled in the same manner. The integrability conditions reduce to

$$(\zeta + \bar{\zeta})[i\Omega + \tfrac{1}{2}p(\bar{\sigma}\bar{\phi}\phi^{-1} - \sigma\phi\bar{\phi}^{-1})] + \sigma\phi\bar{\delta}\zeta - \bar{\sigma}\bar{\phi}\delta\bar{\zeta} - iDq + iq(\rho + \bar{\rho}) = 0, \tag{3.11}$$

where

$$\Omega \equiv \mathrm{Im}[\bar{\delta}(\sigma\phi) + (\bar{\beta} - \alpha - \bar{\tau})\sigma\phi]. \tag{3.12}$$

Thus as before, (3.11) may be integrated to yield

$$q = 2f^{-1}\int f\{\mathrm{Re}\zeta\cdot[\Omega - p\mathrm{Im}(\sigma\phi\bar{\phi}^{-1})] + \mathrm{Im}(\sigma\phi\bar{\delta}\zeta)\}dr \tag{3.13}$$

and the integrability conditions for (2.16) and (2.17) are satisfied for this choice of q. This completes our proof of Theorem 3.

Before attempting to determine classes of solutions of the BLM field equations, it would seem natural to consider the possibility that one might be able to proceed in the spirit of the Bel, Lapiedra, Montserrat generalization, using, however, a simpler modified set of field equations. In this direction they have established the following theorem showing that this is in general not possible.

Theorem 4: In order for a singular electromagnetic field to exist, in the generalized sense with $\eta = 0$, it is necessary and sufficient that Γ be a geodesic congruence and that

$$\tfrac{1}{2}\epsilon^{abcd}DL_{ce}\cdot L_d^e l_b = -2|\sigma|^2\omega l^a, \tag{3.14}$$

where

$$L_{ab} = \nabla_{(a}l_{b)} - \theta g_{ab},$$

$$DL_{ab} = l^c\nabla_c L_{ab}.$$

The spin coefficient formulation of Theorem 4 enables one to interpret clearly, in terms of the Bel–Petrov type of the Weyl tensor and the optical scalars associated with Γ, the restrictions imposed by the integrability conditions of the theorem. If we replace L_{ab} by its spinor equivalent,

$$L_{AA'BB'} = 2\mathrm{Re}\gamma\cdot o_A o_B\bar{o}_{A'}\bar{o}_{B'} - (\bar{\tau} + \bar{\beta} + \alpha)o_A o_B\bar{o}_{(A'}\bar{\iota}_{B')} - (\tau + \beta + \bar{\alpha})o_{(A}\iota_{B)}\bar{o}_{A'}\bar{o}_{B'} + \bar{\sigma}o_A o_B\bar{\iota}_{A'}\bar{\iota}_{B'} + \sigma\iota_A\iota_B\bar{o}_{A'}\bar{o}_{B'} + \mathrm{Re}\rho\cdot\{\iota_A o_B o_{A'}\bar{\iota}_{B'} + o_A\iota_B o_{A'}\bar{o}_{B'}\} - \theta\epsilon_{AB}\epsilon_{A'B'}$$

and evaluate (3.14), a lengthy but elementary calculation yields:

Theorem 4': The field equations (2.13)-(2.17) modified by the demand $\eta = 0$ and integrable if and only if

$$\kappa = 0 \tag{3.15}$$

and

$$4|\sigma|^2\omega = i[\sigma\bar{\Psi}_0 - \bar{\sigma}\Psi_0]. \tag{3.16}$$

Corollary 4: If the Weyl tensor is algebraically special in the sense of the Bel–Petrov classification, having repeated principal null direction l^a, then the modified system with $\eta = 0$ admits a solution.

If l^a is a principal null direction of the Weyl tensor which is of Bel–Petrov Type I, then the modified system admits a solution if and only if Γ is a twist-free geodesic congruence.

4. SOLUTIONS OF THE BLM SYSTEM

In this section, we present a collection of classes of solutions of the BLM field equations (2.13)-(2.17).

The mode of classification of the solutions which we employ, and indeed the determinants of the solutions themselves depend on the following fundamental theorem of Sachs.[7]

Theorem 5: If $R_{ab}l^a l^b = 0$ and $\Psi_0 = 0$, and Γ is a geodesic congruence of curves affinely parameterized by r; then the optical scalars σ, θ, and ω depend on the parameter r in one of the three ways:

Case 1: If $\sigma\bar{\sigma} \neq \theta^2 + \omega^2$, then

$$\sigma = \frac{b}{r^2 + a^2 - b^2}, \quad \theta = \frac{r}{r^2 + a^2 - b^2}, \quad \omega = \frac{a}{r^2 + a^2 - b^2};$$

Case 2: If $\sigma\bar{\sigma} = \theta^2 + \omega^2$ and $\theta \neq 0$, then

$$\sigma = 1 + ia/r, \quad \theta = 1/2r, \quad \text{and} \quad \omega = a/2r;$$

Case 3: If $\sigma\bar{\sigma} = \omega^2$ and $\theta = 0$, then

$$\sigma = ia, \quad \theta = 0, \quad \text{and} \quad \omega = a,$$

where a and b are real-valued functions which are independent of r.[8]

The symbol I.X.Y. will be used to index the solutions of the modified generalized singular electromagnetic field equations ($\eta = 0$), while II.X.Y. will be employed with solutions of the general system. The individual classes will be described according to whether Γ is a shearing or shear-free congruence: i.e., by writing $X = 1$ or $X = 0$, respectively; and according to the r dependence of the optical scalars given in Theorem 5: the values of Y indicating the appropriate case, e.g., II. 1. 1 is the class of solutions to the general system corresponding to the situation, where Γ is a shearing geodesic congruence with associated optical scalars of the form described by Theorem 5, Case 1.

The solutions of type I.X.Y. appearing in the following table are based on the second part of Corollary 4 and are the only solutions which can be determined without more detailed knowledge of the curvature of the underlying Lorentzian manifold.

Synopsis of solutions

In the following solutions a and b are functions introduced in Theorem 5, and A_0, A_1, B_0, B_1 are real-valued functions, independent of r, which are determined by conditions in the initial hypersurface whenever it exists.[9]

I. 0. 1 $$\phi = \frac{A_0 + A_1 r}{r^2 + a^2} + \frac{i}{a}\frac{A_1 a^2 - A_0 r}{r^2 + a^2} \text{ for } a \neq 0,$$
$$\zeta = B_0,$$
$$\phi = A_0/r + iA_1/r, \zeta = B_0 \text{ for } a = 0.$$

I. 0. 2 No solution exists by Theorem 5.

I. 0. 3 $$\phi = A_0 + iA_1,$$
$$\zeta = B_0.$$

I. 1. 1 $$\phi = A_0/(r + b),$$
$$\zeta = B_0(r + b/r - b),$$
or
$$\phi = iA_1/(r - b),$$
$$\zeta = B_0(r - b/r + b).$$

I. 1. 2 $$\phi = A_0, \ \phi = iA_1/r,$$
or
$$\zeta = B_0/r, \ \zeta = B_0 r.$$

I. 1. 3 No solution exists since $\sigma \neq 0, \Psi_0 = 0$ implies $\omega = 0$, which contradicts Case 3 of Theorem 5.

II. 0. 1, II. 0. 2, and II. 0. 3 are the same as I. 0. 1, I. 0. 2, and I. 0. 3 with the exception that in these cases we have

$$\zeta = B_0 + iB_1$$

II. 1. 1 $$\phi = \frac{A_0 + A_1 r}{r^2 + a^2 - b^2} + \frac{i}{a}\left[A_1 - \frac{(r + b)(A_0 + A_1 r)}{(r^2 + a^2 - b^2)}\right],$$
$$\zeta = \exp\left(-2\int \frac{b}{r^2 + a^2 - b^2}\frac{\phi_r^2 - \phi_i^2}{\phi_r^2 + \phi_i^2}\,dr\right) + 4i\int \frac{\zeta_r b}{r^2 + a^2 - b^2}\phi_i\phi_r dr.$$

II. 1. 2 $$\phi = A_0 + i(A_1/r - aA_0),$$
$$\zeta = \frac{B_0 r}{A_0^2(1 + a^2)r^2 - 2A_0A_1 r + A_1^2} + i\frac{(aB_0 r - A_1B_0/A_0)}{A_0^2(1 + a^2)r^2 - 2A_0A_1 r + A_1^2} + iB_1.$$

II. 1. 3 $$\phi = A_0(1 - 2iar),$$
$$\zeta = \frac{B_0}{1 + 4a^2r^2}(1 + 2iar) + iB_1.$$

The techniques employed in the determination of the solutions in each of these classes are essentially the same, and hence we will illustrate the methods in the following cases only.

We are assuming that $\eta = \kappa = \sigma = 0$ and that θ and ω assume the form

I. 0. 1 $$\theta = r/(r^2 + a^2), \ \omega = a/(r^2 + a^2).$$

If one writes $\phi \overset{\text{def}}{=} \phi_r + i\phi_i$, then Eq. (2.14) may be resolved into its real and imaginary parts yielding

$$D\phi_r = -1/(r^2 + a^2)(r\phi_r - a\phi_i), \tag{4.1}$$

$$D\phi_i = -1/(r^2 + a^2)(a\phi_r + r\phi_i). \tag{4.2}$$

One may rewrite these equations in a standard fashion to obtain

$$(r^2 + a^2)D\phi_r + 4rD\phi_r + 2\phi_r = 0, \tag{4.3}$$

$$\phi_i = 1/a[(r^2 + a^2)D\phi_r + r\phi_r].$$

Equation (4.3) may be integrated using elementary techniques to yield

$$\phi_r = \frac{A_0 + A_1 r}{r^2 + a^2}$$

and

$$\phi_i = \frac{1}{a}\frac{A_1 a^2 - A_0 r}{r^2 + a^2}.$$

Finally considering Eq. (2.16), we have

$$D\zeta = 0$$

and hence $\zeta = 0$.

We are assuming that $\phi = 0, \eta \neq 0$ and that the optical scalars σ, θ, ω assume the form

II.1.2 $\sigma = 1 + ia/2r,\ \theta = 1/2r,\ \omega = a/2r.$

If one writes $\phi = \phi_r + i\phi_i$, then Eq. (2.14) may be decomposed into

$$D\phi_r = 0, \tag{4.4}$$

$$D\phi_i = -1/r(\phi_i + a\phi_r). \tag{4.5}$$

And hence

$$\phi_r = A_0,$$

$$\phi_i = A_1/r - aA_0.$$

Similarly writing $\zeta \equiv \zeta_r + i\zeta_i$, one may decompose Eq. (2.16) into

$$\frac{D\zeta_r}{\zeta_r} = \frac{-1}{r}\ \frac{A_0^2(1 + a^2)r^2 - A_1^2}{A_0^2(1 + a^2)r^2 - 2aA_0A_1r + A_1^2} \tag{4.6}$$

and

$$D\zeta_i = \frac{-aB_0}{[A^2(1 + a^2)r^2 -]aA_0A_1r + A_1^2} + \frac{2A_0A_1B_0r}{\{[A^2(1 + a^2)r^2 -]aA_0A_1r + A_1^2\}^2}, \tag{4.7}$$

whereupon an elementary calculation yields

$$\zeta_r = \frac{B_0r}{[A^2(1 + a^2)r^2 -]A_0A_1r + A_1^2},$$

$$\zeta_i = \frac{aB_0r - A_1B_0/A_0}{[A^2(1 + a^2)r^2 -]A_0A_1r + A_1^2}.$$

*Present address: unknown.

[1]L. Mariot, Thèse, Fac. Sci. Paris (1957).
[2]I. Robinson, J. Math. Phys. **2**, 290 (1961).
[3]L. Bel, R. Lapiedra, and A. Montserrat, Cahiers Phys. **82**, 433 (1965).
[4]E. T. Newman and R. Penrose, J. Math. Phys. **3**, 566 (1962).
[5]R. Penrose in *Battelle Rencontres*, edited by C. M. DeWitt and J. A. Wheeler Benjamin (1968), P. 121.
[6]See Ref. 4, p. 569, Eqs. (4.2k, c, and e).
[7]R. Sachs, Proc. R. Soc. A **264**, 309 (1961). Theorem 5 is essentially Theorem 6.3 on p. 327 of this reference.
[8]These cases assume that $\epsilon = \kappa = \pi = 0$, *viz.* the tetrad $l^a, n^a, m^a, \bar{m}^a$ is parallelly propagated along Γ. The proof then essentially consists in integrating the Newman–Penrose equations (6.11a), (6.11b) of Ref. 4 for the indicated conditions.
[9]A discussion of solutions to the BLM equations (2.13)-(2.17), when Γ is shearfree has been given by W. Maher and J. Zund, C.R. Acad. Sci. A **268**, 1231 (1969) and C.R. Acad. Sci. A **268**, 1307 (1969). Other solutions have been given by L. Bel and A. Montserrat, C.R. Acad. Sci. (Paris) **258**, 4659 (1964), and R. Lapiedra, C.R. Acad. Sci. (Paris) **262**, 475 (1966), for plane monochromatic waves at spatial infinity in a Schwarzschild space-time. J. D. Zund, Nuovo Cimento Ser. X **55B**, 15 (1968), has constructed special solutions in conformally flat space-times.

A note on vector-meson dominance and the compactness of current algebra

J. A. de Azcárraga*
Department of Applied Mathematics and Theoretical Physics, University of Cambridge, England
(Received 8 July 1971)

It is shown that for any semisimple inner symmetry group (compact or not) the Lee–Weinberg–Zumino procedure applied to a Yang–Mills type Lagrangian leads to field algebra. The exclusion of the noncompact case is analyzed.

1. INTRODUCTION

The usual $SU(2) \otimes SU(2)$ current algebra commutation relations were introduced by Gell–Mann[1] who postulated for the hadronic axial charges the same commutation relations as for the leptonic ones, setting the real constant α in

$$[Q_i^A, Q_j^A] = \alpha \cdot i\epsilon_{ijk} Q_k^V \tag{1.1}$$

equal to 1 and fixing in this way the scale of the axial current. The choice $\alpha = -1$ would give, instead of the compact $SU_2 \otimes SU_2$ algebra, the Lie algebra of the Lorentz group.

Recently, the structure of a "relaxed" current algebra defined by the set of commutators

$$[Q_i^V, V_{j,\mu}(x)] = i\epsilon_{ijk} V_{k,\mu}(x), \quad [Q_i^V, A_{j,\mu}(x)] = i\epsilon_{ijk} A_{k,\mu}(x),$$
$$[Q_i^A, V_{j,\mu}(x)] = i\epsilon_{ijk} A_{k,\mu}(x), \quad [Q_i^A, A_{j,\mu}(x)] \equiv i X_{ij,\mu}(x) \tag{1.2}$$

has been considered[2] for the classical vertex $\rho A_1 \pi$ in a meson-pole dominance context. As a result, it has been found that the hypothesis of meson dominance is powerful enough to set

$$X_{ij,\mu} = \alpha \cdot i\epsilon_{ijk} V_{k,\mu} \tag{1.3}$$

with α *positive*; it is obvious that for $\alpha > 0$, an adequate definition of the scale of the axial current reproduces the $SU(2) \otimes SU(2)$ commutation relations, the Lorentz algebra being excluded.

The model used in Ref. 2 is equivalent to tree graphs in a Lagrangian framework. In it, the only theory which incorporates vector-meson dominance—through current-field identities—is the Yang–Mills theory; current-field identities lead in turn, through the Lee-Weinberg-Zumino[3,4] (LWZ) procedure, to the algebra of fields. The algebra of fields has been considered only in the compact case, and, because of the connection found between meson dominance and the usual $SU(2) \otimes SU(2)$ algebra in the $\rho A_1 \pi$ example, it was hoped in Ref. 2 that the analysis through the Yang–Mills formalism would lead to similar results. We consider the problem here closely analysing the LWZ method without making any assumption on the inner symmetry group and show that, for any semisimple group, the LWZ procedure gives field algebra. Its structure is the same as that of the symmetry group, and the (unintegrated) commutation relations exhibit constant Schwinger terms which involve the Killing form of the group. However, since the field algebra is necessarily associated with the corresponding inner symmetry group, the Lagrangian is fraught with physical difficulties if the algebra is not compact. Then if these difficulties are to be avoided, we conclude that the hypothesis of meson dominance implies the compactness of current algebra in the framework of the Yang–Mills theory.

The derivation of field algebra has its origin in the possibility of constructing Lagrangians with some definite properties under gauge transformations.[5,6] To show the results mentioned above, we consider first the construction of a Yang–Mills type Lagrangian, which is only possible when the symmetry group is semisimple; we derive then the equations of motion, and, once current-field identities have been established, we obtain the field algebra commutation relations.

In what follows, the Lie algebra of the symmetry group is defined by

$$[X_a, X_b] = C_a{}^c{}_b X_c, \quad a,b,c = 1,2,\ldots,r, \tag{1.4}$$

where the structure constants are real. Any matrix representation of the infinitesimal generators $X_a \to (T_a)^l{}_m$ satisfies

$$(T_a)^l{}_m (T_b)^m{}_n - (T_b)^l{}_m (T_a)^m{}_n = C_a{}^c{}_b (T_c)^l{}_n. \tag{1.5}$$

The *adjoint* representation is defined by

$$(\mathrm{ad}X_a)Y \equiv [X_a, Y] \tag{1.6a}$$

for any element Y of the algebra; the matrix corresponding to X_a is

$$(\mathrm{ad}X_a)^c{}_b = C_a{}^c{}_b. \tag{1.6b}$$

The *coadjoint* representation is the contragradient of $\mathrm{ad}X$; $(\mathrm{coad}X_i) = -(\mathrm{ad}X_i)^T$. In the vector space of the algebra there exists a metric defined by the *Killing form*[7,8]

$$\mathrm{Tr}(\mathrm{ad}X_a \mathrm{ad}X_b) \equiv -c(F)_{ab} = -c(F)_{ba} = C_a{}^l{}_m C_b{}^m{}_l, \tag{1.7}$$

where c is a positive constant which defines the "scale" of the Lie algebra. When the group is semisimple, Cartan's criterium states that F is nonsingular, and F and F^{-1} can be used to lower and raise indices. From the Jacobi identity it is easy to prove the equivalence

$$F(\mathrm{ad}X_a)F^{-1} = (\mathrm{coad}X_a) \tag{1.8a}$$

or

$$F_{sc} G_a{}^c{}_b = -G_a{}^c{}_s F_{cb}, \tag{1.8b}$$

which implies total antisymmetry for C_{abc}. For a compact group the Killing form is negative definite and in the adequate coordinate system, F can be taken as the unit matrix. Then, (1.8a) becomes an identity, the distinction between covariant and contravariant indices unnecessary and the structure constants $C_a{}^c{}_b$ totally antisymmetric.

The existence of fields undergoing gauge transformations in the Lagrangian requires the presence of the Yang–Mills universal field $\rho^{a,\mu}(\mu = 0,1,2,3; a = 1,\ldots,r;$ r being equal to the dimension of the algebra) to preserve its invariance.

It is immediate to check that, for a field which transforms under the symmetry group according to the representation defined by the matrix T,

$$\phi'^c = \phi^c + \delta\phi^c, \quad \delta\phi^c = (T^a)^c{}_b(-\alpha^a)\phi^b, \tag{1.9}$$

the *covariant derivative* ∇_μ,

$$\nabla_\mu\phi^c \equiv \partial_\mu\phi^c + \gamma(T^a)^c{}_b\rho^a{}_\mu\phi^b, \tag{1.10}$$

transforms in the same way

$$\delta(\nabla_\mu\phi^c) = (T_a)^c{}_b(-\alpha^a)\nabla_\mu\phi^b, \tag{1.11}$$

when

$$\delta\rho^c_\mu = C_a{}^c{}_b\alpha^b\rho^a_\mu + (1/\gamma)\partial_\mu\alpha^c, \tag{1.12}$$

i.e., when ρ^c_μ transforms according to the adjoint representation for constant gauge transformations. In an analogous fashion, (1.12) implies for the tensor

$$\rho^c_{\mu\nu} \equiv \partial_\mu\rho^c_\nu - \partial_\nu\rho^c_\mu + \gamma C_a{}^c{}_b\rho^a_\mu\rho^b_\nu \tag{1.13}$$

the adjoint representation transformation law,

$$\delta\rho^c_{\mu\nu} = C_a{}^c{}_b\alpha^b\rho^a_{\mu\nu}. \tag{1.14}$$

2. YANG-MILLS LAGRANGIAN AND FIELD ALGEBRA COMMUTATORS

It is a well-known fact[9] that the construction of Hermitian (or orthogonal) scalar products invariant under the action of a group G imposes severe restrictions on the representation of G according to which the factors involved transform. The most trivial example of this is provided by the Lorentz scalar product defined by $g_{\mu\nu}$, which is indeed invariant because the representation $D_{1/2,1/2}$ of the Lorentz group is equivalent to its contragradient, $gD_{1/2,1/2} = (D^{-1}_{1/2,1/2})^T g$. In the case of the Yang–Mills Lagrangian, the condition which allows one to contract the a indices of $\rho^a_{\mu\nu}$ in a G-invariant manner is the equivalence between the-real-adjoint and coadjoint representations, which is guaranteed for semisimple groups [(1.8)]. The metric is provided by the Killing form, and thus we write, distinguishing between covariant and contravariant a indices, the invariant (up to the mass term) Lagrangian[10]

$$\mathcal{L} = -\tfrac{1}{4}\rho^{a,\mu\nu}\rho_{a,\mu\nu} + \tfrac{1}{2}m^2\rho^{a,\mu}\rho_{a,\mu} + \mathcal{L}'. \tag{2.1}$$

$\mathcal{L}'$ includes G-invariant terms in which the only dependence on ρ is through the covariant derivative (1.10).

We have now

$$\frac{\partial\mathcal{L}}{\partial(\partial^\mu\rho^{a,\nu})} = -\rho_{a,\mu\nu}, \tag{2.2}$$

$$\frac{\partial\mathcal{L}}{\partial(\rho^{a,\nu})} = \gamma C_a{}^c{}_b\rho^{b,\mu}\rho_{c,\mu\nu} + m^2\rho_{a,\nu} + \frac{\partial\mathcal{L}'}{\partial(\rho^{a,\nu})}, \tag{2.3}$$

and the equations of motion read

$$-\partial^\mu\rho_{a,\mu\nu} = \gamma C_a{}^c{}_b\rho^{b,\mu}\rho_{c,\mu\nu} + m^2\rho_{a,\nu} + \frac{\partial\mathcal{L}'}{\partial(\rho^{a,\nu})}. \tag{2.4}$$

The momentun conjugate to $\rho^{a,\nu}$ is

$$\pi_{a,\nu} = \frac{\partial\mathcal{L}}{\partial(\partial^0\rho^{a,\nu})} = -\rho_{a,0\nu}, \quad \pi_{a,0} = 0, \ \pi_{a,i} = \rho_{a,i0} \tag{2.5}$$

and the canonical commutation relations are

$$[\rho^{a,i}(x), \pi_{b,j}(x')]_{x'_0=x_0} = i\delta^a_b\delta^i_j\delta(\mathbf{x}-\mathbf{x}'), \tag{2.6}$$

i.e.,

$$[\rho^i_a(x), \pi_{b,j}(x')]_{x_0=x'_0} = iF_{ab}\delta^i_j\delta(\mathbf{x}-\mathbf{x}'), \tag{2.7}$$

where F_{ab} is the matrix associated to the Killing form.

The only term which contributes to $\delta\mathcal{L}$ is the term in m^2, from which are obtained the current-field identities

$$j_{a,\nu}(x) = -\frac{\partial(\delta\mathcal{L})}{\partial(\partial^\nu\alpha^a(x))} = -\frac{m^2}{\gamma}\rho_{a,\nu}(x). \tag{2.8}$$

To calculate the equal time commutators between the currents, we express $\rho_{a,0}$ in terms of $\pi_{a,i}$. From the equations of motion we get

$$\rho_{a,0} = -\frac{1}{m^2}\partial^i\pi_{a,i} - \frac{\gamma}{m^2}C_a{}^c{}_b\rho^{b,i}\pi_{c,i} - \frac{1}{m^2}\frac{\partial\mathcal{L}'}{\partial(\rho^{a,0})}. \tag{2.9}$$

Because of the gauge invariance, the contributions to $\mathcal{L}'$ are of the form $\frac{1}{2}(\nabla_\mu\phi_a)^2$. Thus

$$\frac{\partial\mathcal{L}'}{\partial(\rho^{a,0})} = \gamma(T_a)^{b'}{}_{c'}\phi^{c'}\frac{\partial\mathcal{L}'}{\partial(\partial^0\phi^{b'})}, \tag{2.10}$$

i.e., the last term of (2.9) involves only the fields ϕ^a and their conjugate momenta, which commute with both ρ^i_a and π^i_a, and so does not need to be considered when calculating $[\rho^0_a(x), \rho^l_d(x')]_{x_0=x'_0}$. We obtain then,

$$\begin{aligned}[j^0_a(x), j^l_d(x')]_{x_0=x'_0} &= (m^2/\gamma)^2[-(1/m^2)\partial^i\pi_{a,i}(x) \\ &\quad - (\gamma/m^2)C_a{}^c{}_b\rho^{b,i}(x)\pi_{c,i}(x), \rho^l_d(x')]_{x_0=x'_0} \\ &= i(m^2/\gamma)F_{ad}\partial^l\delta(\mathbf{x}-\mathbf{x}') - iC_a{}^c{}_bF_{cd}j^{b,l}(x)\delta(\mathbf{x}-\mathbf{x}').\end{aligned} \tag{2.11}$$

From (1.8b) we find that

$$-C_a{}^c{}_bF_{cd} = C_a{}^c{}_dF_{cb}, \tag{2.12}$$

and, finally,

$$\begin{aligned}[j^0_a(x), j^l_d(x')]_{x_0=x'_0} &= i(m^2/\gamma)F_{ad}\partial^l\delta(\mathbf{x}-\mathbf{x}') \\ &\quad + iC_a{}^c{}_dj^l_c(x)\delta(\mathbf{x}-\mathbf{x}').\end{aligned} \tag{2.13a}$$

In the same way we obtain

$$[j^0_a(x), j^0_d(x')]_{x_0=x'_0} = iC_a{}^c{}_dj^0_c(x)\delta(\mathbf{x}-\mathbf{x}'), \tag{2.13b}$$

where use has been made of the Jacobi identity for the structure constants and of the commutation relation

$$\left[\frac{\partial\mathcal{L}'}{\partial(\rho^{a,0})}(x), \frac{\partial\mathcal{L}'}{\partial(\rho^{d,0})}(x')\right]_{x_0=x'_0} = i\gamma C_a{}^c{}_d\frac{\partial\mathcal{L}'}{\partial(\rho^{c,0})}(x) \times\delta(\mathbf{x}-\mathbf{x}'), \tag{2.14}$$

which is easily derived from (2.10) using (1.5).

Formulas (2.13a) and (2.13b) together with the evident relation

$$[j^i_a(x), j^l_d(x')]_{x_0=x'_0} = 0 \tag{2.13c}$$

are usually called the field algebra commutation relations. We recall that in its derivation no special properties of the structure constants (such as total antisym-

metry) have been assumed with the exception of the existence of a nonsingular F (semisimplicity).

3. DISCUSSION

The only external difference between Eqs. (2.13) and the usual field algebra commutators is the presence of the matrix F_{ad} instead of δ_{ad} in the Schwinger terms; F_{ad} also appears crucially in the canonical commutation relations (2.7). It is clear that, after double integration of (2.13), the charges reproduce the algebra of the symmetry group. If it is compact [$SU(2) \otimes SU(2)$, for instance], $F_{ad} \to \delta_{ad}$ and Eqs. (2.13) are the LWZ commutation relations. If the group is noncompact F is no longer definite, and this creates the physical difficulties that can be used to rule out noncompact algebras. For instance, to obtain the Lorentz field algebra [noncompact alternative to the $SU(2) \otimes SU(2)$ which appears when α is set equal to -1 in (1.2) and (1.3)], it is necessary to assume the Lorentz group as the group of inner symmetry. The Killing form is, in the appropriate reference, proportional to the diagonal matrix $(1, 1, 1, -1, -1, -1)$ (note that the Killing form gives the Casimir operator, which is $\mathbf{M}^2 - \mathbf{N}^2$ for the Lorentz group). Then, the introduction of F in the canonical commutation relations makes the space of states nonpositive definite,[8] and no subsidiary conditions—such as the Lorentz condition in the Maxwell case—can be used to eliminate negative norm states. Moreover, as is the general case for a noncompact group, no finite Hermitian matrices can represent the generators of algebra; to obtain Hermitian charges infinite matrices operating on infinite multiplets of states are required, which in turn prevents the identification of the universal field as a member of such multiplets. These physical difficulties, inherent to the noncompactness of the inner symmetry group, disappear when the group is compact. We are thus naturally led to the conclusion that the hypothesis of vector-meson dominance induces the compactness of current algebra, as was found in the hard pion calculation of Ref. 2.

ACKNOWLEDGMENTS

The author is in debt to B. Renner for several stimulating discussions and to the DAMTP for the kind hospitality extended to him.

---

*Supported by G.I.F.T. (Spain). Present address: Departamento de Física Teórica, Facultad de Ciencias, Diagonal s/n, Barcelona-14, Spain.

[1] M. Gell-Mann, Phys. Rev. **125**, 1067 (1962).
[2] J. A. de Azcárraga and B. Renner, Nucl. Phys. B **23**, 236 (1970).
[3] T. D. Lee, S. Weinberg, and B. Zumino, Phys. Rev. Lett. **18**, 1029 (1967).
[4] T. D. Lee and B. Zumino, Phys. Rev. **163**, 1667 (1967).
[5] C. N. Yang and R. L. Mills, Phys. Rev. **96**, 191 (1954).
[6] R. Utiyama, Phys. Rev. **101**, 1597 (1956).
[7] See, for instance, R. Hermann, *Lie groups for physicists* (Benjamin, New York, 1966) or *Group theory and its applications*, edited by E. M. Loebel (Academic, New York, 1968).
[8] B. S. DeWitt, *Dynamical theory of groups and fields* (Gordon and Breach, New York, 1965).
[9] See, for instance, Iu. M. Shirokov, Zh. Eksp. Teor. Fiz. **33**, 861 (1957) [Sov. Phys. JETP **6**, 664 (1958)], or M. Hamermesh, *Group theory and its applications to physical problems* (Addison-Wesley, Reading, Mass., 1962).
[10] This expression seems to contradict an earlier result of S. L. Glashow and M. Gell-Mann [Ann. Phys. (N.Y.) **15**, 437 (1961)] which limits the construction of $\mathfrak{L}$ to the case for which the structure constants are totally antisymmetric. However, they restrict themselves to Euclidean metric to contract inner symmetry indices, and then, as we have seen, the invariance can only be achieved with totally antisymmetric C's since F is the unit matrix.

On matrix superpropagators. I

J. Ashmore and R. Delbourgo
Physics Department, Imperial College, London SW7 2BZ, England
(Received 29 July 1971)

Given the free propagator of a matrix-valued field $\phi_{\alpha\beta}$ in the form $<\phi_{\alpha\beta}(x), \phi_{\gamma\delta}(0)> = (1/2)(\delta_{\alpha\gamma}\delta_{\beta\delta} + \delta_{\alpha\delta}\delta_{\beta\gamma} - 2c\delta_{\alpha\beta}\delta_{\gamma\delta})\Delta(x)$, we derive an integral representation for the matrix superpropagator $<\phi^N_{\alpha\beta}(x), \phi^N_{\gamma\delta}(0)>$ for arbitrary N, and apply this to find the exponentially parametrized gravity superpropagator $<|-g(x)|^\omega g_{\alpha\beta}(x), |-g(0)|^\omega g_{\gamma\delta}(0)>$ with $g_{\mu\nu}(x) \equiv [\exp\kappa\phi(x)]_{\mu\nu}$. Other applications are also mentioned.

1. INTRODUCTION

Nonpolynomial Lagrangians are finding increasing application in quantum field theories of elementary particle interactions. The techniques needed to cope with such complicated Lagrangians and the results that have appeared to date have been largely limited to the case when the field $\phi(x)$ is scalar[1] or a member of an $SU(2)$ multiplet,[2] though some progress has lately been achieved towards extending the methods[3] to $SU(3)$ fields. In the meantime one interesting problem which has arisen concerns the exponential parameterization of gravity[4] $g_{\mu\nu}(x) = [\exp\kappa\phi(x)]_{\mu\nu}$, and its associated superpropagator

$$\langle |-g(x)|^\omega g_{\alpha\beta}(x),\ |-g(0)|^\omega g_{\gamma\delta}(0)\rangle$$

for damping out ultraviolet infinities and providing a localizable field theory—a nontrivial generalization of the scalar field situation which is *not* amenable to any of the nonpolynomial techniques[2] presented heretofore. The main object of our paper will be to provide the answer in *closed form* in terms of the free propagator Δ and the gauge parameter c occurring in

$$\langle \phi_{\alpha\beta}(x), \phi_{\gamma\delta}(0)\rangle = \tfrac{1}{2}[\delta_{\alpha\gamma}\delta_{\beta\delta} + \delta_{\alpha\delta}\delta_{\beta\gamma} - 2c\,\delta_{\alpha\beta}\delta_{\gamma\delta}]\Delta(x), \tag{1}$$

although the new techniques which we develop find numerous other applications.

Basic to the success of the whole approach is to look for an integral representation in which an *invariant function* of the matrix field $\phi(x)$ appears to an *arbitrary power*. Only when this is found can one, by appropriate manipulation, derive an integral representation for $\langle \mathrm{Tr}[\phi^N(x)], \mathrm{Tr}[\phi^N(o)]\rangle$, with N *arbitrary*, and thence proceed to the matrix function

$$\langle \phi^N_{\alpha\beta}(x), \phi^N_{\gamma\delta}(0)\rangle \equiv K^{(N)}_{\alpha\beta\gamma\delta}\Delta^N(x)$$

as we shall show. Thus whereas the simple integral representation

$$|-g|^{-1/2} = \pi^{-2}\int d^4n e^{n_\alpha g_{\alpha\beta} n_\beta}$$

has been successfully used in the past[5] for dealing with the rational parameterization of gravity $g = \eta + \kappa\phi$, it is quite futile for our purpose since the (invariant) determinant $|-g|$ appears to be a fixed power of $-\frac{1}{2}$ and will not lend itself to determining matrix superpropagators

$$\langle F[\phi(x)]_{\alpha\beta}, F[\phi(0)]_{\gamma\delta}\rangle$$

for *general* matrix functions F. Combing through the literature, it seems that generalizations of integral representations of a single variable y to a $\nu \times \nu$ matrix Y involve the *determinant* $|Y|$ of the matrix. Indeed one of the simplest such generalizations is Siegel's integral[6] which provides an expression for $|Y|$ to an arbitrary power:

$$\int dX\,|X|^\mu e^{-\mathrm{Tr}(XY)} = \pi^{\nu(\nu-1)/4}\Gamma_\nu(\mu)|Y|^{-\mu-(1/2)(\nu+1)}, \tag{2}$$

where

$$\Gamma_\nu(\mu) \equiv \Gamma(\mu+1)\Gamma(\mu+\tfrac{3}{2})\cdots\Gamma[\mu+\tfrac{1}{2}(\nu+1)],$$
$$dX = d^{\nu(\nu+1)/2}x \equiv \prod_{\alpha\le\beta} dx_{\alpha\beta}, \tag{3}$$

and the integration is taken over all $\frac{1}{2}\nu(\nu+1)$ elements $x_{\alpha\beta}$ which maintain the $\nu\times\nu$ matrix X *symmetric* and *positive definite*. Initially representation (2) is defined for $\mathrm{Re}\mu > -1$; but we shall later analytically continue the formula to other μ values. When $\nu = 1$, Siegel's integral will be recognized as the conventional definition of the gamma function.

In Sec. 2 we shall show how (2) and (3) provide the key for finding integral representations of other invariants and specifically $\mathrm{Tr}(X^N)$. Used in conjunction with (1) we show how this enables us to arrive at integral representations of general matrix superpropagators in Sec. 3. In Sec. 4 we reduce these integrals to Pfaffians over double variable integrals following methods which have been extensively used in statistical mechanics.[7] The relevant Pfaffians are evaluated for gravity in Sec. 5 and the exponential superpropagator explicitly spelled out. We conclude the paper with the applicability of our new method to other cases such as chiral $SU(3)$.

2. AN INTEGRAL REPRESENTATION FOR $\langle \mathrm{Tr}(\phi^N), \mathrm{Tr}(\phi^N)\rangle$

Begin with Siegel's integral (2) and make the substitution $Y = 1 + \kappa\phi$. Then take the vacuum expectation value of the product of two such integrals, remembering from (1) that

$$\langle e^{[-\mathrm{Tr}(A\phi(x))]}, e^{[-\mathrm{Tr}(B\phi(0))]}\rangle = e^{\{[\mathrm{Tr}(AB) - c\,\mathrm{Tr}A\,\mathrm{Tr}B]\Delta(x)\}}. \tag{4}$$

Thus we get

$$\begin{aligned}
&\langle |1+\kappa\phi(x)|^{-\mu-(1/2)(\nu+1)}, |1+\kappa'\phi(0)|^{-\mu'-(1/2)(\nu+1)}\\
&= \int \frac{dX\,dX'}{\pi^{\nu(\nu-1)/2}} \frac{|X|^\mu |X'|^{\mu'}}{\Gamma_\nu(\mu)\Gamma_\nu(\mu')}\\
&\quad\times \langle e^{\{-\mathrm{Tr}(X[1+\kappa\phi(x)])\}}, e^{\{-\mathrm{Tr}(X'[1+\kappa\phi(0)])\}}\rangle\\
&= \int \frac{dX\,dX'}{\pi^{\nu(\nu-1)/2}} \frac{|X|^\mu |X'|^{\mu'}}{\Gamma_\nu(\mu)\Gamma_\nu(\mu')}\\
&\quad\times e^{-\mathrm{Tr}X - \mathrm{Tr}X' + \kappa\kappa'\Delta[\mathrm{Tr}(XX') - c\,\mathrm{Tr}X\,\mathrm{Tr}X']}\\
&= \int \frac{dX}{\pi^{\nu(\nu-1)/4}} \frac{|X|^\mu}{\Gamma_\nu(\mu)}\\
&\quad\times e^{-\mathrm{Tr}X}|1 - \kappa\kappa'\Delta(X - c\,\mathrm{Tr}X)|^{-\mu'-(1/2)(\nu+1)}.
\end{aligned} \tag{5}$$

Noting that

$$|1 + \kappa\phi|^{-n} = \exp[-n \operatorname{Tr} \log(1 + \kappa\phi)]$$
$$= 1 - n \operatorname{Tr} \log(1 + \kappa\phi) + o(n^2),$$

we can take μ derivatives of (5) at $-\frac{1}{2}(\nu + 1)$ to get

$$\langle \operatorname{Tr} \log[1 + \kappa\phi(x)], \operatorname{Tr} \log[1 + \kappa'\phi(0)] \rangle$$
$$= \frac{\partial}{\partial\mu}\Big|_{(\nu+1)/2} \int \frac{dX}{\pi^{\nu(\nu-1)/4}} \frac{|X|^\mu}{\Gamma_\nu(\mu)}$$
$$\times e^{-\operatorname{Tr}X} \operatorname{Tr} \log[1 - \kappa\kappa'\Delta(X - c \operatorname{Tr}X)]$$

and then pick terms of order $(\kappa\kappa')^N$ to reach our desired integral representation

$$\langle \operatorname{Tr}[\phi^N(x)], \operatorname{Tr}[\phi^N(0)] \rangle = N\Delta^N \frac{\partial}{\partial\mu} I_\nu^{(c)}(\mu, N)\Big|_{\mu=-(\nu+1)/2}, \quad (6)$$

where

$$I_\nu^{(c)}(\mu, N) \equiv \int \frac{dX}{\pi^{\nu(\nu-1)/4}} \frac{|X|^\mu}{\Gamma_\nu(\mu)} e^{-\operatorname{Tr}X} \operatorname{Tr}((X - c \operatorname{Tr}X)^N). \quad (7)$$

None of the nonpolynomial Lagrangian methods to be found in the literature have succeeded in deriving such a closed form expression, so in (6) and (7) we have a powerful new result. Of course as they stand the formulas are not useful because I is a $\frac{1}{2}\nu(\nu + 1)$—fold integral which has to be differentiated; but in Secs. 4 and 5 we show how to carry out such calculations. For the present, however, we simply note that the $c \neq 0$ case follows by straightforward binomial manipulation from the case $c = 0$. Thus

$$\int dX |X|^\mu e^{-\operatorname{Tr}X} \operatorname{Tr}((X - c \operatorname{Tr}X)^N)$$
$$= \sum_{n=0}^{N} \binom{N}{n} \left(c \frac{\partial}{\partial b}\right)^{N-n} \int dX |X|^\mu e^{-b \operatorname{Tr}X} \operatorname{Tr}(X^n)\Big|_{b=1}$$
$$= \sum_{n=0}^{N} \binom{N}{n} \frac{\Gamma(N + \mu\nu + \frac{1}{2}\nu(\nu + 1))}{\Gamma(n + \mu\nu + \frac{1}{2}\nu(\nu + 1))} (-c)^{N-n}$$
$$\times \int dX |X|^\mu e^{-\operatorname{Tr}X} \operatorname{Tr}(X^n)$$

upon rescaling X by the factor b. Hence

$$I_\nu^{(c)}(\mu, N) = \sum_{n=0}^{\infty} \frac{\Gamma(N + 1)\Gamma(N + \mu\nu + \frac{1}{2}\nu(\nu + 1))(-c)^{N-n}}{\Gamma(n + 1)\Gamma(n + \mu\nu + \frac{1}{2}\nu(\nu + 1))\Gamma(N - n + 1)} I_\nu^{(0)}(\mu, n). \quad (8)$$

3. MATRIX SUPERPROPAGATORS

We shall now prove that once we have calculated (6) so that the coefficients a_N in

$$\langle \operatorname{Tr}(\phi^N(x)), \operatorname{Tr}(\phi^N(0)) \rangle \equiv N!\nu a_N \Delta^N(x) \quad (9)$$

are known, this is sufficient to determine all matrix superpropagators

$$\langle F_{\alpha\beta}(\phi), F_{\gamma\delta}(\phi) \rangle, \qquad F(\phi) \equiv \sum_N \frac{F_N \phi^N(x)}{N!}. \quad (10)$$

To that end we first write the general form for the N'th power

$$\langle \phi^N_{\alpha\beta}(x), \phi^N_{\gamma\delta}(0) \rangle$$
$$\equiv N![\tfrac{1}{2}(\delta_{\alpha\gamma}\delta_{\beta\delta} + \delta_{\alpha\delta}\delta_{\beta\gamma})b_N - \delta_{\alpha\beta}\delta_{\gamma\delta}c_N]\Delta^N \quad (11)$$

and show that b_N and c_N are determined from the a_N. For if we trace over (11),

$$a_N = b_N - \nu c_N,$$

whereas a direct application of the Wick expansion on (9) gives

$$a_N = \tfrac{1}{2}(\nu + 1 - 2c)\, b_{N-1} - (1 - c\nu)c_{N-1}.$$

The pair of recurrence conditions demonstrate that

$$\tfrac{1}{2}(\nu + 2)(\nu - 1)b_N = \nu a_{N+1} - (1 - c\nu)a_N,$$
$$\tfrac{1}{2}(\nu + 2)(\nu - 1)c_N = a_{N+1} - \tfrac{1}{2}(\nu + 1 - 2c)a_N, \quad (12)$$

so the particular matrix superpropagator (11) reads

$$\langle \phi^N_{\alpha\beta}(x), \phi^N_{\gamma\delta}(0) \rangle = \frac{N!}{(\nu + 2)(\nu - 1)} \times \begin{pmatrix} \{\delta_{\alpha\gamma}\delta_{\beta\delta} + \delta_{\alpha\delta}\delta_{\beta\gamma} - 2\nu^{-1}\delta_{\alpha\beta}\delta_{\gamma\delta}\}\nu a_{N+1} \\ + \{(\delta_{\alpha\gamma}\delta_{\beta\delta} + \delta_{\alpha\delta}\delta_{\beta\gamma})(c\nu - 1) + \delta_{\alpha\beta}\delta_{\gamma\delta}(\nu + 1 - 2c)\}a_N \end{pmatrix} \quad (11')$$

Then for the most general matrix superpropagator (10) it follows that

$$\langle F_{\alpha\beta}(x), F_{\gamma\delta}(0) \rangle = \sum_N \frac{F_N^2 \Delta^N}{N!} \langle \phi^N_{\alpha\beta}(x), \phi^N_{\gamma\delta}(0) \rangle = \frac{1}{(\nu + 2)(\nu - 1)} \begin{pmatrix} \{\delta_{\alpha\gamma}\delta_{\beta\delta} + \delta_{\alpha\delta}\delta_{\beta\gamma} - 2\nu^{-1}\delta_{\alpha\beta}\delta_{\gamma\delta}\}\nu a'(\Delta) \\ + \{(\delta_{\alpha\gamma}\delta_{\beta\delta} + \delta_{\alpha\delta}\delta_{\beta\gamma})(c\nu - 1) + \delta_{\alpha\beta}\delta_{\gamma\delta}(\nu + 1 - 2c)\}a(\Delta) \end{pmatrix}, \quad (10')$$

where

$$a'(\Delta) = \sum_N a_{N+1} F_N^2 \Delta^N / N!,$$
$$a(\Delta) = \sum_N a_N F_N^2 \Delta^N / N! = \nu^{-1}\langle \operatorname{Tr}F, \operatorname{Tr}F \rangle. \quad (13)$$

In particular, for the interesting exponential matrix case where $F(\phi) = \exp\kappa\phi$, $F_N = \kappa^N$. Therefore

$$a(\Delta) = \sum_N a_N (\kappa^2\Delta)^N / N!,$$
$$a'(\Delta) = da(\Delta)/d(\kappa^2\Delta) \quad (13')$$

are supposedly determinate functions and have only to be substituted into (12) to provide us with the answer for

$$\langle [\exp\kappa\phi(x)]_{\alpha\beta}, [\exp\kappa\phi(0)]_{\gamma\delta} \rangle.$$

4. REDUCTION TO PFAFFIAN FORM

Our next task is to actually evaluate the integral

$$I_\nu^{(0)}(\mu, N) = \int \frac{dX}{\pi^{\nu(\nu-1)/4}} \frac{|X|^\mu}{\Gamma_\nu(\mu)} e^{-\operatorname{Tr}X} \operatorname{Tr}(X^N), \quad (7')$$

since every quantity we need stems from it. Observing that the integrand involves functions of X invariant under similarity transformations, it serves to parametrize the symmetric positive definite matrices X via the orthogonal group which diagonalizes them to $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_\nu)$. Thus we define

$$X(\theta, \lambda) = S(\theta)\Lambda \tilde{S}(\theta),$$

where θ are a set of $\frac{1}{2}\nu(\nu-1)$ angular parameters. The Jacobian of the transformation is

$$d^{\nu(\nu+1)/2} x = \prod_{j>i} |\lambda_j - \lambda_i| \prod_k d\lambda_k . J(\theta) d^{\nu(\nu-1)/2}\theta,$$

and the change of variables simplifies (7) to

$$I_\nu^{(0)}(\mu, N) = \frac{\gamma_\nu}{\Gamma_\nu(\mu)} \prod_k \int_0^\infty d\lambda_k e^{-\lambda_k} \lambda_k^\mu \prod_{j>i} |\lambda_j - \lambda_i| \sum_{l=1} \lambda_l^N, \tag{14}$$

where γ_ν is a normalization constant (coming from the angular integration) determined by the condition that $I_\nu^{(0)}(\mu, 0) = \nu$ according to Siegel's integral. Integrals of the type (14) involve the Vandermonde determinant

$$\prod_{j>i} |\lambda_j - \lambda_i| = \prod_{j>i} \epsilon(\lambda_j - \lambda_i) \cdot \begin{vmatrix} 1 & 1 & \ldots & 1 \\ \lambda_1 & \lambda_2 & \ldots & \lambda_\nu \\ \lambda_1^2 & \lambda_2^2 & \ldots & \lambda_\nu^2 \\ \cdot & \cdot & \ldots & \cdot \\ \lambda_1^{\nu-1} & \lambda_2^{\nu-1} & \ldots & \lambda_\nu^{\nu-1} \end{vmatrix} \tag{15}$$

and occur frequently in statistical mechanics where methods have been developed[7] for simplifying them which we shall adopt. From (14) and (15), using properties of determinants, the calculation reduces to finding

$$I_\nu^{(0)}(\mu, N) = \frac{\gamma_\nu \nu!}{\Gamma_\nu(\mu)} \sum_{l=1}^{\nu} \int_0^\infty d\lambda_1 \cdots d\lambda_\nu \prod_{j>i} \epsilon_{ji} E_1(\lambda_1) E_2(\lambda_2) \cdots E_{N+l}(\lambda_l) \cdots E_\nu(\lambda_\nu)$$

with

$$E_\mu(\lambda) \equiv \lambda^{\mu+n-1} e^{-\lambda}. \tag{16}$$

Hereafter we shall suppose that ν is even (the case ν odd is treated in the Appendix). Recall now that the Pfaffian of antisymmetric $\nu \times \nu$ matrices A is defined as

$$|A| \equiv \begin{vmatrix} a_{12} & a_{13} \cdots a_{1\nu} \\ & a_{23} \cdots a_{2\nu} \\ & \cdots \\ & a_{\nu-1\,\nu} \end{vmatrix} = \frac{2}{\nu} \sum_{j>i} a_{ij} A_{ij}$$

$$= [2^{\nu/2}(\tfrac{1}{2}\nu)!]^{-1} \sum_{j \text{ perms}} (-1)^P a_{j_1 j_2} a_{j_3 j_4} \cdots a_{j_{\nu-1} j_\nu}, \tag{17}$$

where A_{ij} is the cofactor of the element a_{ij}. It so happens that

$$|\epsilon| = \prod_{j>i} \epsilon_{ji} = \prod_{j>i} \epsilon(\lambda_j - \lambda_i). \tag{18}$$

Used in conjunction with (16) I becomes a sum of double integrals in a Pfaffian expansion

$$I_\nu^{(0)}(\mu, N) = \frac{\gamma_\nu \nu!}{\Gamma_\nu(\mu)} \sum_{l=1}^{\nu} \begin{vmatrix} a_{12} & a_{13} & \cdot & a_{1\,l+N} & \cdot & a_{1\nu} \\ & a_{23} & \cdot & a_{2\,l+N} & \cdot & a_{2\nu} \\ & & & & \cdots & \\ & & & & \cdot & \end{vmatrix}$$

with elements

$$a_{ij} \equiv \int_0^\infty d\lambda\, d\lambda' E_i(\lambda) E_j(\lambda') \epsilon(\lambda' - \lambda).$$

Furthermore, since Pfaffians are linewise additive,

$$I_\nu^{(0)}(\mu, N) = \frac{2\gamma_\nu \nu!}{\Gamma_\nu(\mu)} \sum_{j>i} a_{ij}^{(N)} A_{ij}, \tag{19}$$

where A_{ij} is the appropriate cofactor and

$$a_{ij}^{(N)} \equiv \tfrac{1}{2}[a_{i\,N+j} + a_{N+i\,j}], \qquad a_{ij}^{(0)} = a_{ij}. \tag{20}$$

Thus the normalization constant $\gamma_\nu = \Gamma_\nu(\mu)/(\nu-1)!\,|A|$.

It remains to work out the coefficients (20) which enter in the Pfaffian expansion (19). Changing variables to $\lambda + \lambda' = s$, $\lambda - \lambda' = st$, the integral representation of $a_{ij}^{(N)}$ boils down to

$$a_{ij}^{(N)} \equiv (\tfrac{1}{2})^{2\mu+i+j+N} \Gamma(2\mu + i + j + N) \alpha_{ij}^{(N)}$$

$$\alpha_{ij}^{(N)} \equiv \int_0^1 dt\,(1-t^2)^\mu [(1-t)^N + (t \longleftrightarrow -t)] \times [(1-t)^{i-1}(1+t)^{j-1} - (t \longleftrightarrow -t)]. \tag{21}$$

These α coefficients are embodied in the generating function

$$\left.\begin{aligned} \alpha_{ij}(z) &\equiv \sum_{N=0}^\infty \frac{\alpha_{ij}^{(N)} z^N}{N!} = 2\int_0^1 dt(1-t^2)^\mu e^z \cosh zt \begin{pmatrix} (1-t)^{i-1}(1+t)^{j-1} \\ -(i \longleftrightarrow j) \end{pmatrix} = e^{z\partial_{ij}} \alpha(z) \\ \partial_{ij} &\equiv (1-\partial)^{i-1}(1+\partial)^{j-1} - (i \longleftrightarrow j), \qquad \partial \equiv \frac{d}{dz} \\ \alpha(z) &= 2\int_0^1 dt(1-t^2)^\mu \sinh zt = \frac{\sqrt{\pi}\,\Gamma(\mu+1)}{(\frac{1}{2}z)^{\mu+(1/2)}} L_{\mu+(1/2)}(z) = \sqrt{\pi} \sum_{k=0}^\infty \frac{\Gamma(\mu+1)(\frac{1}{2}z)^{2k+1}}{\Gamma(k+\mu+2)\Gamma(k+\frac{3}{2})} \end{aligned}\right\}, \tag{22}$$

where L is the modified Struve function. Thus $\alpha_{ij}(z)$ is an entire function of z. Since the only cases of physical interest are for $\nu \leq 6$, when the i, j values are small integers, the most practical way of discovering the $\alpha_{ij}^{(N)}$ is to work out

$$\alpha_{ij}^{(N)} = \partial^N e^z \partial_{ij} \alpha(z)\Big|_{z=0} = (1+\partial)^N \partial_{ij} \alpha(z)\Big|_{z=0}, \tag{23}$$

a method which has another advantage as we shall find. In passing let us record the lowest few N values when $\nu = 4$ for future use:

$$\alpha_{12}^{(0)} = \alpha_{12}^{(1)} = \tfrac{1}{2}\alpha_{13}^{(0)} = \tfrac{1}{2}\alpha_{13}^{(1)} = -\tfrac{4}{3},$$
$$\alpha_{23}^{(0)} = \alpha_{23}^{(1)} = \tfrac{1}{2}\alpha_{24}^{(0)} = \tfrac{1}{2}\alpha_{24}^{(1)} = -4,$$

$$\alpha_{14}^{(0)} = \alpha_{14}^{(1)} = -\tfrac{4}{3}, \quad \alpha_{34}^{(0)} = \alpha_{34}^{(1)} = 4,$$
$$\alpha_{12}^{(2)} = \tfrac{1}{2}\alpha_{13}^{(2)} = \tfrac{4}{3}, \quad \alpha_{23}^{(2)} = \tfrac{1}{2}\alpha_{24}^{(2)} = -12,$$
$$\alpha_{14}^{(2)} = \tfrac{52}{3}, \quad \alpha_{34}^{(2)} = \tfrac{20}{3}, \text{etc.}$$

With $c \neq 0$, therefore, the final Pfaffian expansion reads

$$I_\nu^{(c)}(\mu, N) = \frac{\gamma_\nu \nu!}{\Gamma_\nu(\mu)} \sum_{j>i} \sum_{n=0}^{N} \times \frac{\Gamma(N+1)\Gamma(N+\mu\nu+\frac{1}{2}\nu(\nu+1))(-c)^{N-n}}{\Gamma(n+1)\Gamma(n+\mu\nu+\frac{1}{2}\nu(\nu+1))\Gamma(N-n+1)} a_{ij}^{(n)} A_{ij}$$

to be differentiated at $\mu = -\frac{1}{2}(\nu+1)$. Since

$$\Gamma_\nu(\mu) = \pi^{\nu/4} 2^{-\mu\nu+\nu^2/4} \Gamma(2\mu+2)\Gamma(2\mu+4)\cdots\Gamma(2\mu+\nu)$$

introduces multiple zeros which are cancelled by zeros in the Pfaffian coefficients, we finally write for the superpropagator

$$\langle \mathrm{Tr}\phi^N(x), \mathrm{Tr}\phi^N(o)\rangle = N!\Delta^N \nu!\, \gamma_\nu \sum_{n=0}^{N} \binom{N}{n}(-c)^{N-n} \times \frac{\partial}{\partial\mu}\Big|_{-(\nu+1)/2} \sum_{j>i} \frac{a_{ij}^{(n)} A_{ij}}{\Gamma_\nu(\mu)\Gamma(n+\mu\nu+\frac{1}{2}\nu(\nu+1))} \tag{24}$$

which is now a matter of straightforward computation as we will exhibit for the case of gravity ($\nu = 4$) below.

5. THE GRAVITY SUPERPROPAGATOR

In Lagrangian field theories which conform to the principles of general relativity, one meets interactions of the type

$$L = g^{\alpha\beta}(x)|-g(x)|^\omega T_{\alpha\beta},$$

where $g_{\alpha\beta}$ is the metric tensor, $T_{\alpha\beta}$ is a covariant tensor density of some fields ψ, and ω is the weight needed to render L a scalar density. The corresponding superpropagators of interest are defined by (the second order in L) expectation values

$$\langle g^{\alpha\beta}(x)|-g(x)|^\omega, g^{\gamma\delta}(0)|-g(0)|^\omega\rangle = K^{\alpha\beta\gamma\delta}(\Delta, c') \tag{25}$$

and are given functions of the flat space free graviton (h-field) propagator

$$\langle h^{\alpha\beta}(x), h^{\gamma\delta}(0)\rangle = \tfrac{1}{2}(\eta^{\alpha\gamma}\eta^{\beta\delta} + \eta^{\alpha\delta}\eta^{\beta\gamma} - 2c'\eta^{\alpha\beta}\eta^{\gamma\delta})\Delta(x) \tag{26}$$

in a particular gauge specified by c'. η is the Minkowski metric and $g = \eta + \kappa h + o(\kappa^2)$ where κ is related to the Newtonian gravitational constant.

Such a superpropagator was evaluated previously[5] for the rational parameterization $g = \eta + \kappa h$ when $l = \frac{1}{2}$ and 1. However, a localizable version of gravity[4] uses instead the exponential parameterization $g = \exp\kappa h$, and this is the example we shall concentrate on here as it is not amenable to any of the previous treatments.[2,5] Noting that

$$g^{\alpha\beta}(x)|-g(x)|^\omega = \{\exp\kappa[h(x) + \omega\eta\ \mathrm{Tr}h(x)]\}^{\alpha\beta},$$

we can interpret $(h + \omega\eta\ \mathrm{Tr}h)$ as a new field ϕ and by rotating to a Euclidean metric reestablish Eq. (1) with c given in terms of c' by

$$(1-4c) = (1-4c')(1+4\omega)^2.$$

All we have left to do is to evaluate the coefficients a_N appearing in (9),

$$a_N = 3!\,\gamma_4 \sum_{n=0}^{N} \binom{N}{n}(-c)^{N-n} \frac{\partial}{\partial\mu}\Big|_{-5/2} \times \sum_{j>i} \frac{a_{ij}^{(n)} A_{ij}}{\Gamma_4(\mu)\Gamma(n+4\mu+10)}, \tag{27}$$

to obtain the required form (12). Now in

$$\sum_{j>i}^{4} \frac{a_{ij}^{(n)} A_{ij}}{\Gamma_4(\mu)} = \frac{2^{4\mu+4}/\pi}{\Gamma(2\mu+2)\Gamma(2\mu+4)} \times \begin{pmatrix} a_{12}^{(n)}a_{34} - a_{13}^{(n)}a_{24} + a_{14}^{(n)}a_{23} \\ + a_{34}^{(n)}a_{12} - a_{24}^{(n)}a_{13} + a_{23}^{(n)}a_{14} \end{pmatrix}$$

we encounter terms

$$\frac{a_{ij}^{(n)} a_{\bar{i}\bar{j}}}{\Gamma(2\mu+2)\Gamma(2\mu+4)} = \frac{\Gamma(2\mu+i+j+n)\Gamma(2\mu+10-i-j)}{\Gamma(2\mu+4)\Gamma(2\mu+2)2^{4\mu+10+n}} \alpha_{ij}^{(n)} \alpha_{\bar{i}\bar{j}}^{(0)}$$

which vanish at $2\mu + 5 = 0$ and are easily differentiated there; a little work using definition (21) gives

$$\begin{aligned} 32\pi \frac{\partial}{\partial\mu}\Big|_{-5/2} \sum_{j>i} \frac{a_{ij}^{(n)} A_{ij}}{\Gamma_4(\mu)\Gamma(n+4\mu+10)} &= 12\delta_{n0}\alpha_{14}^{(0)}\alpha_{23}^{(0)} - 3\delta_{n1} \\ &\times (\alpha_{13}^{(1)}\alpha_{24}^{(0)} + \alpha_{12}^{(1)}\alpha_{34}^{(0)}) + \tfrac{3}{2}\delta_{n2}\alpha_{12}^{(2)}\alpha_{34}^{(0)} \\ &+ 2^{-n}3[n(n+1) \\ &\times \alpha_{34}^{(n)}\alpha_{12}^{(0)} + 2n\alpha_{24}^{(n)}\alpha_{13}^{(0)} + 2\alpha_{14}^{(n)}\alpha_{23}^{(0)} + 2\alpha_{23}^{(n)}\alpha_{14}^{(0)}]. \end{aligned}$$

Inserting this into (27), we finally arrive at

$$a_N = \sum_{n=0}^{N} \binom{N}{n}(-c)^{N-n} \times \begin{pmatrix} 2\delta_{n0} - \frac{3}{2}\delta_{n1} + \frac{1}{2}\delta_{n2} \\ -2^{-n-2}\left\{\begin{matrix} \frac{1}{2}n(n+1)\alpha_{34}^{(n)} + 2n\alpha_{24}^{(N)} \\ + 3\alpha_{14}^{(n)} + \alpha_{23}^{(n)} \end{matrix}\right\} \end{pmatrix} \tag{28}$$

with the α_{ij} provided by Eqs. (22) and (23). Thus in terms of the generating Struve function $\alpha(z)$ of (22) having $\mu = -\frac{5}{2}$,

$$a_N = 2(-c)^N - 3N(-c)^{N-1}/2 + N(N-1)(-c)^{N-2}/8 - \tfrac{1}{4}\sum_{n=0}^{N}\binom{N}{n}(-c)^{N-n}(\tfrac{1}{2}\partial)^n e^z \times \begin{pmatrix} n(n-1)(1-\partial^2)^2 \\ + 2n(1-\partial^2)(5-\partial^2) \\ + 4(5+\partial^2) \end{pmatrix} \partial\alpha(z)|_{z=0}, \tag{28'}$$

and the first few coefficients are

$$a_0 = 4, \quad a_1 = 1-4c, \quad a_2 = \tfrac{5}{2} - 2c + 4c^2,$$
$$a_3 = 4 - \tfrac{15}{2}c + 3c^2 - 4c^3, \text{etc.},$$

which may also be checked by the (more tedious) direct Wick expansion.

Lastly we form the function (13′):

$$a(\Delta)=\sum_N \frac{a_N(\kappa^2\Delta)^N}{N!}=\left(2-\frac{3(\kappa^2\Delta)}{2}+\frac{(\kappa^2\Delta)^2}{8}\right)e^{-c\kappa^2\Delta}$$

$$-\frac{1}{4}e^{-\kappa^2\Delta(c-1/2)}\begin{pmatrix}\frac{1}{4}(\kappa^2\Delta\partial)^2e^z(1-\partial^2)^2\\+(\kappa^2\Delta\partial)e^z(1-\partial^2)(5-\partial^2)\\+4e^z(5+\partial^2)\end{pmatrix}\partial\alpha(z)|_0$$

$$=(2-3z+\tfrac{1}{2}z^2)e^{-2cz}+e^{z(1-2c)}$$

$$\times\begin{pmatrix}2+3z-z^2\\+\frac{1}{2}z(z+\frac{1}{2})3\pi L_0(z)\\-\frac{1}{2}z^2\pi L_1(z)\end{pmatrix}\Bigg|_{z=\kappa^2\Delta/2} \tag{29}$$

(where L is the modified Struve function) which enters in

$$\langle g^{\alpha\beta}(x)|-g(x)|^\omega, g^{\gamma\delta}(0)|-g(0)|^\omega\rangle$$

$$=\langle[\exp\kappa\phi(x)]^{\alpha\beta},\exp[\kappa\phi(0)]^{\gamma\delta}\rangle$$

$$=\frac{1}{18}\begin{pmatrix}\{\eta^{\alpha\gamma}\eta^{\beta\delta}+\eta^{\alpha\delta}\eta^{\beta\gamma}-\frac{1}{2}\eta^{\alpha\beta}\eta^{\gamma\delta}\}\,4\dfrac{d}{d(\kappa^2\Delta)}+\\\{[\eta^{\alpha\gamma}\eta^{\beta\delta}+\eta^{\alpha\delta}\eta^{\beta\gamma}](4c-1)+\eta^{\alpha\beta}\eta^{\gamma\delta}(5-2c)\}\end{pmatrix}a(\Delta). \tag{30}$$

This is the gravity superpropagator in the gauge c' of (26) where $(1-4c)=(1-4c')(1+4\omega)^2$. Note that when $\omega=-\frac{1}{2}$, $c=c'$, and furthermore in the De Donder gauge $c=\frac{1}{2}$ makes for some simplification in (29) and (30). However, (29) and (30) cannot really be simplified very much further and by their nature we see how entirely nontrivial is the result for exponential gravity. It goes without saying, however, that the superpropagator is an entire function of Δ; indeed the leading behavior as $\Delta\to\infty$ is

$$\langle g^{\alpha\beta}|-g|^\omega, g^{\gamma\delta}|-g|^\omega\rangle$$

$$\sim[\eta^{\alpha\gamma}\eta^{\beta\delta}+\eta^{\alpha\delta}\eta^{\beta\gamma}+\eta^{\alpha\beta}\eta^{\gamma\delta}](\kappa^2\Delta)^{3/2}e^{\kappa^2\Delta(1-c)},$$

where as $\Delta\to 0$ we of course recover the perturbation series.

6. OTHER APPLICATIONS

We have demonstrated above how Siegel's integral (2) enables one to derive closed expressions for superpropagators of matrix fields as in (6) and how the calculation boils down to taking the derivative of a ν-dimensional Pfaffian as in (24). In practice this is not too difficult since the cases of practical interest involve $SU(\nu)$ or $SO(\nu)$ fields with $\nu\leq 6$, where the number of Pfaffian terms is small.[8] Thus for the example of $SU(3)$ there will be six terms in all to be differentiated, as we briefly discuss in the Appendix. For instance $SU(3)\otimes SU(3)$ with matrix interactions of the type $m\bar{\psi}^\alpha[\exp\gamma_5\kappa\phi]^\beta_\alpha\psi_\beta$, where ϕ is a nonet of pseudoscalar mesons propagating as

$$\langle\phi^\beta_\alpha(x)\phi^\delta_\gamma(o)\rangle=\delta^\beta_\gamma\delta^\delta_\alpha\Delta(x)$$

in the interaction picture, is perfectly amenable to our treatment. This is perhaps the single most important application of the new method besides the case of gravity covered in this paper, and will be the subject of a separate publication. What is not so obvious is how the new technique helps for calculating higher orders of perturbation theory which the previously used transform methods are at some pains to solve in any case.

APPENDIX

Assume ν is odd. All the manipulations done until Eq. (16) in Sec. 4 remain valid. However, as it stands (18) fails. In its place one can prove by induction that

$$\prod_{j>i}\epsilon(\lambda_j-\lambda_i)=Pf(\epsilon^+),$$

where the bordered $(\nu+1)\times(\nu+1)$ matrix ϵ^+ is defined by adding an extra column and row of ones:

$$\epsilon^+=\left(\begin{array}{ccc|c}&&&1\\&&&1\\&&&\cdot\\&&&\cdot\\\hline-1&-1&\cdots&0\end{array}\right)$$

with ϵ_{ji} defined as a $\nu\times\nu$ matrix element. Doing the same symmetrization steps as in the even-dimensional case, one obtains

$$I^{(0)}_\nu(\mu,N)=\frac{2\gamma_\nu\nu!}{\Gamma_\nu(\mu)}\sum_{j>i}^{\nu+1}a^{(N)}_{ij}A_{ij},$$

where a_{ij} are defined by (20) for $1\leq i<j\leq\nu$ while

$$a_{i\,\nu+1}=-a_{\nu+1\,i}=\int_0^\infty E_i(\lambda)\,d\lambda=\Gamma(\mu+i),$$

$$a^{(N)}_{i\,\nu+1}=\int_0^\infty\lambda^N E_i(\lambda)\,d\lambda=\Gamma(\mu+i+N),\qquad i=1,\ldots,\nu,$$

and trivially $a_{\nu+1\,\nu+1}=0$. [Again by construction we have $I^{(0)}_\nu(\mu,o)=\nu$.] In this way we can take over formula (24) with the summation over i,j values running from 1 to $\nu+1$.

It is quite straightforward to check that for $\nu=1$ one correctly reproduces the known result for the scalar case, viz., $a_N=(1-c)^N$. The case $\nu=3$ includes the superpropagator for chiral $SU(3)\otimes SU(3)$, and the relevant computation devolves upon differentiating at $\mu=-2$,

$$\frac{1}{\Gamma_3(\mu)}\sum_{1\leq i\leq j<4}a^{(N)}_{ij}A_{ij}=\frac{2^{2\mu+2}}{\sqrt{\pi}\,\Gamma(\mu+1)\Gamma(2\mu+3)}\begin{pmatrix}a^{(N)}_{12}a_{34}-a^{(N)}_{13}a_{24}\\+a^{(N)}_{14}a_{23}+a^{(N)}_{34}a_{12}\\-a^{(N)}_{24}a_{13}+a^{(N)}_{23}a_{14}\end{pmatrix}$$

$$=\frac{(\frac{1}{2}\partial)^N e^z}{\Gamma(2\mu+3)}\begin{pmatrix}\{\Gamma(\mu+3)\Gamma(2\mu+3+N)+\Gamma(2\mu+3)\Gamma(\mu+3+N)\}\,\partial_{12}\\-\{\Gamma(\mu+2)\Gamma(2\mu+4+N)+\Gamma(2\mu+4)\Gamma(\mu+2+N)\}\,\partial_{13}\\+\{\Gamma(\mu+1)\Gamma(2\mu+5+N)+\Gamma(2\mu+5)\Gamma(\mu+1+N)\}\,\partial_{23}\end{pmatrix}\frac{\alpha(z)}{\Gamma(\mu+1)}\Bigg|_0.$$

The details and results of this example will be given elsewhere because of its particular importance in strong interaction physics.

[1]G. V. Efimov, Nucl. Phys. **74,** 657 (1965); H. M. Fried, Phys. Rev. **174,** 1725 (1968).
[2]R. Delbourgo, J. Math. Phys. **13,** 464 (1972); A. Hunt, K. Koller and Q. Shafi, Phys. Rev. D 3, 1327 (1971); T. Martin and J. G. Taylor (Southampton preprint).
[3]J. Charap (private communication).
[4]C. Isham, A. Salam, and J. Strathdee, Phys. letters **35B**, 585 (1971).
[5]R. Delbourgo and A. Hunt. Nuovo Cimento Lett. **4,** 1010 (1970); C. Isham, A. Salam, and J. Strathdee, Phys. Rev. D **2**, 685 (1970).
[6]C. L. Siegel, Ann. Math. **36,** 527 (1935); R. Bellmann, Duke Math. J. **23,** 571 (1956); I. Olkin, Duke Math. J. **26,** 207 (1959).
[7]M. L. Mehta, Nucl. Phys. **18,** 395 (1960); N. G. de Bruijn, J. Ind. Math. Soc. **19,** 133 (1955).
[8]When $\nu=2$, there is only one such Pfaffian term and the results of Ref. 2 are easily recovered.

Exact bound states of some N-body systems with two- and three-body forces

F. Calogero
Istituto di Fisica dell'Università di Roma, Istituto Nazionale di Fisica Nucleare, Sezione di Roma, Rome, Italy

C. Marchioro*
Istituto Matematico dell'Università di Roma, Rome, Italy
(Received 1 July 1971)

The exact wavefunctions and the corresponding binding energies are exhibited, for some N-body systems with long-range two-body and three-body forces. These results may also be used to obtain upper and lower bounds for the binding energies of certain many-body systems with two-body forces only.

1. INTRODUCTION

Exact results are beautiful and may be useful to test approximation techniques or to develop (approximate) computational methods. No exactly solvable model is known for the nonrelativistic quantum-mechanical many-body problem with pair forces in three-dimensional space, besides the trivial case when all particles interact pairwise via harmonic oscillator forces (except at most one pair of them, that can interact through an arbitrary potential[1]). Here we present some many-body models, for whom some eigenstates, generally including the ground state, can be explicitly exhibited.[2] The main limitation of these models is the (unavoidable) presence of three-body forces, whose character and strength is determined by those of the two-body forces. Another limitation is the long-range nature of these forces. Most of the models involve N distinguishable particles or N bosons; for $N = 3$ and $N = 4$ solvable models involving (spinless) fermions are also introduced. Results are given for three-dimensional space; they could be easily extended to higher dimensional spaces.

2. OSCILLATOR AND CENTRIFUGAL PAIR FORCES, AND ADDITIONAL THREE-BODY FORCES

Consider the many-body problem characterized by the hamiltonian

$$H = -\frac{\hbar^2}{2m}\sum_{i=1}^{N}\Delta_i + \tfrac{1}{4}m\omega^2\sum_{i>j=1}^{N} r_{ij}^2 + g\sum_{i>j=1}^{N} r_{ij}^{-2} + G\sum_{k=1}^{N}\sum_{\substack{i>j=1\\ i\neq k, j\neq k}}^{N}\frac{(\mathbf{r}_{ki}\cdot\mathbf{r}_{kj})}{(r_{ki}^2 r_{kj}^2)}, \quad (1)$$

with $g > -\hbar^2/(4m)$.[3] Here Δ_i is the Laplace operator acting on the coordinate $\mathbf{r}_i$ of the ith particle, and $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$. The first term in the rhs is the usual nonrelativistic kinetic energy of N particles of mass m; the second term, is the usual harmonic oscillator pair potential; the third term is a "centrifugal" potential acting between every two particles; finally, the last term represents a three-body potential. It is easy to show that this last potential is always finite. The ratio of this three-body potential, acting between three particles, to the sum of the three "centrifugal" two-body potentials acting between the same three particles is expressed, in terms of the ratio α of the area of the triangle formed by the three particles, divided by the sum of the squares of the three sides of the same triangle, by the following formula:

$$[(\mathbf{r}_{12}\cdot\mathbf{r}_{13})/(r_{12}^2 r_{13}^2) + (\mathbf{r}_{23}\cdot\mathbf{r}_{21})/(r_{23}^2 r_{21}^2) + (\mathbf{r}_{31}\cdot\mathbf{r}_{32})/(r_{31}^2 r_{32}^2)]/(r_{12}^{-2} + r_{23}^{-2} + r_{31}^{-2}) = 32\alpha^2/(1 + 16\alpha^2). \quad (2)$$

This ratio is convenient to display the dependence of the three-body potential on the configuration of the triplet of interacting particles. Note that it is an increasing function of α, so that its maximal value $\frac{1}{2}$ obtains for the (equilateral) configuration for which α attains its maximal value $\alpha_{max} = (48)^{-1/2}$. This formula also implies that the three-body potential vanishes whenever the three particles in question are aligned (in fact in one-dimensional space the last term in the rhs of Eq. (1) would vanish identically; it is for this reason that the model analogous to that treated here, but with only two-body forces, is solvable in one dimension[4]).

If the coupling constant characterizing the three-body potential is related to that characterizing the two-body "centrifugal" potential by the equation

$$G = (\hbar^2/m)b^2, \quad (3)$$

with

$$b = \tfrac{1}{2}[(1 + 4mg\hbar^{-2})^{1/2} - 1], \quad (4)$$

a subset of the completely symmetrical eigenfunctions of the Hamiltonian (1) is given by the explicit formula (valid in the center-of-mass frame)

$$\psi_n = \text{const}\ z^{b/2}\exp[-\tfrac{1}{2}(m\omega/\hbar)(N/2)^{1/2}\rho^2] \times L_n^B[(m\omega/\hbar)(N/2)^{1/2}\rho^2],\ n = 0, 1, 2, \ldots\ldots \quad (5)$$

where L_n^B is a Laguerre polynomial[5] and

$$B = \tfrac{1}{2}[(3N - 5) + bN(N - 1)], \quad (6)$$

$$z = \prod_{i>j=1}^{N} r_{ij}^2, \quad (7)$$

$$\rho^2 = \frac{1}{N}\sum_{i>j=1}^{N} r_{ij}^2. \quad (8)$$

The corresponding eigenvalues of H are

$$E_n = \hbar\omega(N/2)^{1/2}\{2n + \tfrac{1}{2}[3N - 3 + bN(N - 1)]\}. \quad (9)$$

For $n = 0$, the wavefunction (5) has no nodes besides those implied by the singular nature of the centrifugal potential, that forces the wavefunction to vanish whenever the coordinates of two particles coincide. Thus it corresponds to the ground state of the system.

Note that, as $g \to 0$, also $G \to 0$, and the wavefunction (5) becomes (for $n = 0$) the ground-state wavefunction of the oscillator problem without centrifugal potential and with Bose statistics. Thus the situation is different from that of the one-dimensional problem of Ref. 4

where in the limit $g \to 0$ the wavefunction becomes that of the groundstate of the oscillator problem with Fermi statistics (because in the one-dimensional case the zeros of the wavefunction remain as $g \to 0$, while here as $g \to 0$ they disappear).

If instead

$$G = (\hbar^2/m)f^2, \tag{10}$$

with

$$f = \tfrac{3}{2}\{[1 + 4mg/(9\hbar^2)]^{1/2} - 1\}, \tag{11}$$

then for $N = 3$ and for $N = 4$ a subset of the completely antisymmetrical eigenfunctions of the Hamiltonian (1) and the corresponding eigenvalues can be given explicitly. Specifically for $N = 3$,

$$\psi_n = \text{const}\,(\mathbf{r}_{12} \wedge \mathbf{r}_{23})z^{f/2} \exp[-\tfrac{1}{2}(m\omega/\hbar)(\tfrac{3}{2})^{1/2}\rho^2] \times L_n^F[(m\omega/\hbar)(\tfrac{3}{2})^{1/2}\rho^2], \quad n = 0, 1, 2, \ldots, \tag{12}$$

with

$$F = 3f + 3 \tag{13}$$

and

$$E_n = \hbar\omega(\tfrac{3}{2})^{1/2}(2n + 3f + 4). \tag{14}$$

Note that each of the three cartesian components of the (axial) vector ψ_n is an eigenfunction of H, corresponding to the same eigenvalue E_n. Thus the wavefunctions (12) describe bound states of three identical (spinless) fermions, having unit total (orbital) spin. Of course the eigenvalues E_n are (at least) threefold degenerate. The fact that the wavefunctions (12) change sign if the coordinates of any two particles are exchanged is easily verified; indeed the product

$$\mathbf{r}_{12} \wedge \mathbf{r}_{23} \equiv \mathbf{r}_1 \wedge \mathbf{r}_2 + \mathbf{r}_2 \wedge \mathbf{r}_3 + \mathbf{r}_3 \wedge \mathbf{r}_1 \tag{15}$$

is clearly antisymmetrical, while the two coordinates z and ρ are clearly symmetrical.

For $N = 4$ a set of completely antisymmetrical eigenfunctions of H (with G always given by eqs. (10-11)) is

$$\psi_n = \text{const}[(\mathbf{r}_{12} \wedge \mathbf{r}_{23}) \cdot \mathbf{r}_{34}]z^{f/2} \exp[-\tfrac{1}{2}(m\omega/\hbar)(2)^{1/2}\rho^2] \times L_n^{F'}[(m\omega/\hbar)(2)^{1/2}\rho^2]; \; n = 0, 1, 2, \ldots, \tag{16}$$

with

$$F' = 6f + 5; \tag{17}$$

and the corresponding eigenvalues are

$$E_n = \hbar\omega 2^{1/2}(2n + 6f + 6). \tag{18}$$

The eigenfunctions ψ_n transform now as pseudoscalars; thus they represent bound states of four (spinless) fermions, with vanishing total (orbital) spin and negative parity.

3. COULOMB OR GRAVITATIONAL PAIR FORCES, AND ADDITIONAL TWO-BODY FORCES

Consider the N-body system described by the Hamiltonian

$$H = -\hbar^2 \sum_{i=1}^{N} (2m_i)^{-1} \Delta_i + \sum_{i>j=1}^{N} \left(\frac{q_i q_j}{r_{ij}}\right) + \hbar^{-2} \sum_{k=1}^{N} \sum_{\substack{i>j=1 \\ i\neq k, \\ j\neq k}}^{N} \frac{m_k m_i m_j}{(m_k + m_i)(m_k + m_j)} q_k^2 q_i q_j (\hat{\mathbf{r}}_{ki} \cdot \hat{\mathbf{r}}_{kj}), \tag{19}$$

where $\hat{\mathbf{r}}_{ki}$ indicates the unit vector in the direction $\mathbf{r}_{ki}$:

$$\hat{\mathbf{r}}_{ki} = \mathbf{r}_{ki}/r_{ki}. \tag{20}$$

This Hamiltonian describes N particles, of masses m_i and charges q_i, interacting via the usual Coulomb potential $q_i q_j/r_{ij}$, and in addition via a finite-valued three-body potential. The function

$$\psi = \text{const} \prod_{i>j=1}^{N} \exp\{[(m_i m_j)/(m_i + m_j)][(q_i q_j r_{ij}/\hbar^2]\} \tag{21}$$

is a solution of the stationary Schrödinger equation

$$H\psi = E\psi \tag{22}$$

with

$$E = -(2\hbar^2)^{-1} \sum_{i>j=1}^{N} \{[(m_i m_j/(m_i + m_j)]q_i^2 q_j^2\}. \tag{23}$$

Obviously the function ψ is nodeless. Thus, if it is normalizable, it corresponds to the ground state of the many-body system described by the Hamiltonian (19). A sufficient condition for normalizability is that

$$\text{Re}\; q_i m_i \left(\sum_{\substack{j=1 \\ j\neq i}}^{N} [(q_j m_j)/(m_i + m_j)]\right) < 0, \quad \text{for } i = 1, 2, \ldots, N. \tag{24}$$

Another case worth mentioning obtains setting

$$q_j = iKm_j, \tag{25}$$

so that the Hamiltonian becomes

$$H = -\hbar^2 \sum_{i=1}^{N} (2m_i)^{-1} \Delta_i - K^2 \sum_{i>j=1}^{N} \frac{m_i m_j}{r_{ij}} + K^4 \hbar^{-2} \sum_{k=1}^{N} \sum_{\substack{i>j=1 \\ i\neq k, \\ j\neq k}}^{N} \frac{m_k^3 m_i^2 m_j^2}{(m_k + m_i)(m_k + m_j)} \times (\hat{\mathbf{r}}_{ki} \cdot \hat{\mathbf{r}}_{kj}). \tag{26}$$

The two-body potential of this Hamiltonian describes now the usual gravitational interaction. In this case the wavefunction ψ of Eq. (22) is clearly always normalizable; and, if all the masses are equal, it is completely symmetrical.

4. LOWER AND UPPER BOUNDS TO THE GROUND-STATE ENERGY OF N-BODY SYSTEMS WITH COULOMB OR GRAVITATIONAL PAIR FORCES

Exploiting the fact that the three-body potential in Eqs. (19) and (26) is finite, and it possesses therefore finite upper and lower bounds, one can obtain from these results upper and lower bounds for the ground-state energies of some N-body models with the usual (two-body) Coulomb or gravitational forces. For instance from Eqs. (26) and (23) it follows that the ground-state energy E_N of the system of N bosons of mass m interacting via the gravitational potential $-K^2m^2/r$ is bounded as follows:

$$-(K^4m^5/\hbar^2)\tfrac{1}{8}N(N-1)[1 + \tfrac{1}{3}(N-2)\eta_{max}] \leq E_N \leq -(K^4m^5/\hbar^2)\tfrac{1}{8}N(N-1)[1 + \tfrac{1}{3}(N-2)\eta_{min}] \tag{27}$$

where $\eta_{\min}$, resp., $\eta_{\max}$ are the minimal and the maximal values for all possible spatial configurations of N particles, of the quantity

$$\eta = \sum_{k=1}^{N} \sum_{\substack{i>j=1 \\ i\neq k, \\ j\neq k}}^{N} (\hat{\mathbf{r}}_{ki} \cdot \hat{\mathbf{r}}_{kj})/[\tfrac{1}{6}N(N-1)(N-2)], \tag{28}$$

namely of the mean value, averaged over all the $\frac{1}{6}N(N-1)(N-2)$ different triplets of particles, of the sum of the three cosines of the three internal angles of the triangles formed by each triplet of particles. Note that, since clearly for any triplet the maximal and minimal values of this quantity are respectively $\frac{3}{2}$ and 1 (corresponding to an equilateral triangle and to a completely stretched triangle, respectively), obviously $\eta_{\max} \leq \frac{3}{2}$ and $\eta_{\min} = 1$. Thus a simpler upper and lower bound for E_N is

$$-(K^4 m^5/\hbar^2)\tfrac{1}{16}N^2(N-1) \leq E_N \leq -(K^4 m^5/\hbar^2)\tfrac{1}{24}N(N^2-1) \tag{29}$$

The upper bound is, however, less stringent than one previously known.[6] The lower bound, instead, coincides with a lower bound previously known[6]; moreover for $N > 4$, it should be possible to improve it, by simple geometrical considerations. This is an interesting possibility; indeed, together with the known upper bound,[6] the lower bound of Eq. (27) implies that at large N

$$E_N = -CN^3K^4m^5/(16\hbar^2), \tag{30}$$

with

$$[16/(6\pi)] \leq C \leq \tfrac{2}{3}\bar{\eta}, \tag{31}$$

where $\bar{\eta}$ is the value of $\eta_{\max}$ for $N \to \infty$, and it is certainly less than $\frac{3}{2}$.

5. GENERAL CASE

The solvable models of Sec. 3 obtain as special cases from the remark that the many-body wavefunction

$$\psi = \text{const} \prod_{i>j=1}^{N} F_{ij}(r_{ij}) \tag{32}$$

satisfies the equation

$$\left[-\sum_{i=1}^{N}(2m_i)^{-1}\Delta_i + \sum_{i>j=1}^{N}\left(\frac{(m_i+m_j)}{(2m_im_j)}\right) \times\left(\frac{2F'_{ij}(r_{ij})}{r_{ij}F_{ij}(r_{ij})} + \frac{F''(r_{ij})}{F(r_{ij})}\right) + \sum_{k=1}^{N}\sum_{\substack{i>j=1 \\ i\neq k, j\neq k}}^{N}(m_k)^{-1} \times(\hat{\mathbf{r}}_{ki}\cdot\hat{\mathbf{r}}_{kj})\left(\frac{F'_{ki}(r_{ki})}{F_{ki}(r_{ki})}\right)\left(\frac{F'_{kj}(r_{kj})}{F_{kj}(r_{kj})}\right)\right]\psi = 0. \tag{33}$$

Here the primes indicate differentiation relative to the argument of the function. Different choices of the functions $F_{ij}(r)$ allow the construction of a number of solvable many-body models, with two- and three-body potentials. The only requirement that the functions $F_{ij}(r)$ must satisfy in order that the many-body wavefunction ψ of Eq. (32) describe a many-body bound state, is that this function be normalizable (in the CM frame); a sufficient condition for this is that each function $F_{ij}(r)$ be square-integrable in three-dimensional space. As a consequence of this requirement the potentials of the solvable many-body models originating from Eq. (33) have generally a long range; specifically, at large r the two-body potential vanishes as r^{-2} or r^{-1}, and the three-body potential vanishes as r^{-2} or remains finite, depending whether the functions $F_{ij}(r)$ vanish asymptotically as a power or exponentially.

Clearly if all the functions $F_{ij}(r)$ coincide, $F_{ij}(r) = F(r)$, the wavefunction (32) is completely symmetrical, and in this case, the additional assumption that all the masses m_i coincide, $m_i = m$, makes Eq. (33) a convenient starting point for the introduction of solvable many-bosons models.

For $N = 3$ and $N = 4$ another convenient starting point for the introduction of solvable many-body models is the possibility to write in closed form solutions of the equation

$$\left[-\sum_{i=1}^{N}\Delta_i + 2\sum_{i>j=1}^{N}\left(\frac{4F'_{ij}(r_{ij})}{r_{ij}F'_{ij}(r_{ij})} + \frac{F''(r_{ij})}{F(r_{ij})}\right) + 2\sum_{k=1}^{N}\sum_{\substack{i>j=1 \\ i\neq k, j\neq k}}^{N}(\hat{\mathbf{r}}_{ki}\cdot\hat{\mathbf{r}}_{ki})\left(\frac{F'_{kj}(r_{kj})}{F_{kj}(r_{kj})}\right) \times[F'_{ki}(r_{ki})/F_{ki}(r_{ki})]\right]\psi = 0. \tag{34}$$

Indeed for $N = 3$, three solutions of this equation are provided by the three cartesian components of the vector

$$\psi = \text{const}(\mathbf{r}_{12} \wedge \mathbf{r}_{23}) \prod_{i>j=1}^{N} F_{ij}(r_{ij}), \tag{35}$$

while for $N = 4$ a solution is

$$\psi = \text{const}[(\mathbf{r}_{12} \wedge \mathbf{r}_{23})\cdot\mathbf{r}_{34}] \prod_{i>j=1}^{N} F_{ij}(r_{ij}). \tag{36}$$

Note that, if the functions $F_{ij}(r)$ are all equal, $F_{ij}(r) = F(r)$, these wavefunctions are antisymmetrical under the exchange of the coordinates of any two particles. Thus they provide a convenient starting point for the construction of solvable models of three or four interacting fermions.

ACKNOWLEDGMENTS

It is a pleasure to acknowledge conversations on the subject of this paper with Professor E. Salusti and Professor Yu. A. Simonov.

*Research partially supported by the Consiglio Nazionale delle Ricerche.

[1] J. Hurley, J. Math. Phys. **8**, 813 (1967).

[2] Such models are for simplicity referred to in the following as "solvable" models.

[3] This condition must be imposed to prevent two-body collapse [see, for instance, L. D. Landau and E. M. Lifshitz, *Quantum Mechanics* (Pergamon, New York, 1958), Sec. 35].

[4] F. Calogero, J. Math. Phys. **12**, 419 (1971).

[5] I. S. Gradshteyn and I. M. Ryzhik, *Tables of Integral, Series and Products*, (Academic, New York, 1965); *Higher Transcendental Functions*, edited by A. Erdélyi (McGraw-Hill, New York, 1965), Vol. II.

[6] F. Calogero and C. Marchioro, J. Math. Phys. **10**, 562 (1969).

The class sum operator approach for the point group O and D_4

J. Killingbeck
Physics Department, University of Hull, Hull, Yorkshire, England
(Received 10 May 1971; revised manuscript received 13 August 1971)

The class sum operator approach to the representation theory of the point groups O and D_4 is described and illustrated by means of several examples. Modified character tables are given for both groups, together with the class multiplication table for O. The construction of tensor operators within the group algebra of each group is discussed, using a modified version of traditional character analysis, and it is found that no E type tensor operator appears in the D_4 group algebra.

1. INTRODUCTION

In a previous paper[1] the principles of the class sum operator approach to group representation theory were outlined, and the specific group D_{3h} was treated as an example. D_{3h} is the point group used by physicists to describe the symmetry of the crystal field acting on the paramagnetic ion in various lanthanide salts (e.g., ethyl sulphates, double nitrates). Another point group of traditional interest in crystal field theory is the octahedral group O of proper rotations; the present paper applies the class sum operator approach to that group and to its tetragonal D_4 subgroup. A class sum operator $\mathcal{C}_j$ is constructed[1] by taking the "average" of the group operators R_k from within a given class K (of order h):

$$\mathcal{C}_j = (1/h)\sum R_k. \tag{1.1}$$

The operators $\mathcal{C}_j$ commute with one another, and thus fit naturally into the same formalism as employed in the Dirac commuting-operator approach to quantum mechanics. Each rep of the group is labelled by a set of eigenvalues of the $\mathcal{C}_j$, and various results of the traditional group representation theory can immediately be seen to be special cases of results of general linear space theory. Killingbeck[1] showed how this approach simplifies the decomposition of Kronecker products of reps; and also how it enables the projection operators of Löwdin[2] for the full rotation group to be directly related to the traditional projection operators for the finite point groups. Some of these points are illustrated further in the present paper by means of examples from the octahedral group. Section 2 gives the class multiplication table for O, and states an important symmetry property of the table. Section 3 gives the modified group character tables for O and D_4, briefly explaining their construction and use. Section 4 discusses the construction of tensor operators within the group algebras of O and D_4, and points out that both D_4 and D_{3h} have certain "forbidden" operator types, while O has no such restriction.

2. CLASS SUM MULTIPLICATION

Each class sum operator $\mathcal{C}_j$ commutes with every group operator, and, further, any product of two class sum operators can be written in the form[3]

$$\mathcal{C}_J\mathcal{C}_K = \sum h_N c_{JKN}\mathcal{C}_N = \sum g_{JKN}\mathcal{C}_N. \tag{2.1}$$

h_N is here the number of elements in the Nth class. Except for a few special cases, the coefficients c_{JKN} for the traditional points groups are invariant under interchange of any two subscripts, and the sum of the coefficients g_{JKN} is unity for fixed J, K, and variable N. The only nonzero coefficients in the case of the octahedral group O are as follows; the numbering of the classes from 1 to 5 being in accord with Table I of this paper:

$$g_{1NN} = 1 \quad \text{for all } N,$$
$$g_{222} = g_{244} = g_{245} = g_{255} = \tfrac{1}{2},$$
$$g_{333} = g_{345} = 2g_{344} = 2g_{355} = \tfrac{2}{3},$$
$$g_{232} = 1.$$

Table I. Modified character table for O.

	1	2	3	4	5
	E	$8C_3$	$3C_2$	$6C_2'$	$6C_4$
A_1	1	1	1	1	1
A_2	1	1	1	−1	−1
E	1	$-\frac{1}{2}$	1	0	0
	1	ω	1	1	1
	1	ω^2	1	−1	−1
T_1	1	0	$-\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$
	1	1	1	1	1
	1	ω	−1	−1	i
	1	ω^2	−1	−1	$-i$
T_2	1	0	$-\frac{1}{3}$	$\frac{1}{3}$	$-\frac{1}{3}$
	1	1	1	−1	−1
	1	ω	−1	1	$-i$
	1	ω^2	−1	1	i

The symmetry of the c_{JKN} is similar to that of the Kronecker products of reps. The existence of a duality between reps and classes is also suggested by the existence of orthogonality relations of first and second kinds in traditional representation theory; this duality has been stressed by Gamba,[4] and explored recently by Robinson.[5] However, the coefficients displayed above show that it is *not* possible to set up a direct one-to-one correspondence between the reps and classes of O so that the class multiplication structure and the Kronecker product structure become identical. (This can be seen directly by noting that the Kronecker product $T_1 \times T_2$ contains the four reps A_2, E, T_1, and T_2, while no product of two distinct class sums yields more than two class sums in its expansion). A similar conclusion also follows from the class multiplication rule for the group D_{3h}; although this feature was not specifically commented on previously.[1] The results given here for O may also be applied to the tetrahedral group T_d if the isomorphism between the groups is exploited. The classes 4 and 5 of Table I are then labeled σ_d and s_4 instead of C_2' and C_4. A similar comment applies to most of the results of this paper, after allowance for details such as the different conventional labeling of the T_1 and T_2 reps for O and T_d.

3. MODIFIED CHARACTER TABLES FOR O AND D_4

Tables I and II are modified character tables for O and D_4, constructed as suggested previously.[1] The principal numbers along each row are the eigenvalues (of the class sum operators $\mathcal{C}_J$), which are associated with the

Table II. Modified character table for D_4.

	1	2	3	4	5
	E	C_2	$2C_4$	$2C_2'$	$2C_2''$
A_1	1	1	1	1	1
A_2	1	1	1	−1	−1
B_1	1	1	−1	1	−1
B_2	1	1	−1	−1	1
E	1	−1	0	0	0
	1	−1	i	1	1
	1	−1	$-i$	−1	−1

given rep. For example, the rep E of O has the eigenvalue set $(1, -\frac{1}{2}, 1, 0, 0)$, and the eigenvalues $\lambda^{(\mu)}$ are related to the traditional characters $\chi^{(\mu)}$ by the equation

$$\chi^{(\nu)}(K) = n^{(\nu)}\lambda^{(\nu)}(K), \tag{3.1}$$

where K is the class label and $n^{(\nu)}$ the dimensionality of the νth rep. The bracketed sets of numbers in the table give the possible eigenvalues of the *single* named operator in the class name; when that operator acts within an irreducible subspace belonging to the given rep. These bracketed numbers provide useful information even though the named operators may not commute and therefore may not be simultaneously diagonizable. The presence of three ones in the $(E, 3)$ position of Table I shows at once that under the reduction $O \to D_2$ the rep E of O gives two A_1 type functions, for example, the two traditional d orbitals $(x^2 - y^2)$ and $(3z^2 - r^2)$ of crystal field theory. The descent in symmetry $O \to D_4$ may be investigated using Tables I and II if the axes of the groups are aligned; for example, by means of the class correspondence $(O \to D_4, E \to E, C_2 \to C_2, C_4 \to C_4, C_2' \to C_2', C_2 \to C_2'')$, which physically corresponds to expanding a cube along one fourfold axis. Thus, the rep E of O has eigenvalues (1, 1) for the single operator C_2, eigenvalues (1, 1) for the single operator C_2^1 and eigenvalues (1, 1) for the single operator C_4. This shows directly that E can only decompose into the reps A_1 and B_2 of D_4, and the validity of the argument is unaltered by the fact that the eigenvalue sets employed may refer to noncommuting operators. As another example, the Kronecker product $T_1 \times T_2$ has the eigenvalue set $(1, \omega, \omega^2)^2 = 3(1, \omega, \omega^2)$ for C_3 and the eigenvalue set $(1, i, -i)(-1, -i, i) = (1, i, -i) + (-1, -i, i) + (1, -1) + (-1)$ for C_4. This indicates the decomposition $T_1 \times T_2 = A_2 + E + T_1 + T_2$, and such a procedure may be applied to any Kronecker products of reps; the eigenvalues of other operators being used whenever necessary to resolve ambiguities. Table I shows that the operator $\mathcal{C}_3 + \frac{1}{3}$ will leave only the A_2 and E components when it operates on the nine basis functions arising from $T_1 \times T_2$. Further, the operator $(C_3 - \omega)(C_3 - \omega^2)$ will then eliminate any remaining E component. Thus the operator $(C_3^2 + C_3 + 1)(\mathcal{C}_3 + \frac{1}{3})$ will project out the A_2 part of the product functions. (Note that $\mathcal{C}_3$ here is a class sum operator, while C_3 is a single group operator.) The way in which the class sum approach permits economical formulation of projection operators was discussed previously,[1] and the present example provides another illustration.

For dihedral groups, the explicit form of the irreducible components of a Kronecker product can be obtained by inspection in simple cases. For example, the product $E \times E$ of D_4 contains the four functions $|i, i\rangle$, $|i, -i\rangle$, $|-i, i\rangle$, $|-i, -i\rangle$, if we use "C_4 quantization." The functions $|i, -i\rangle$ and $|-i, i\rangle$ both have eigenvalues 1 for C_4, C_2 and E. The operation C_2' permutes the two E basis functions, since it physically reverses the fourfold axis direction; thus, the linear combinations $(2)^{-1/2}[|i, -i\rangle \pm |-i, i\rangle]$ belong to eigenvalues 1 and −1 of C_2^1 and accordingly to reps A_1 and A_2, respectively. Note however that the requirement $C_2'|i\rangle = |-1\rangle$ is effectively a specification of which the two dihedral axes shall be called C_2', and any phase factors involved must be considered in detail when explicit functions are used to replace the symbolic vectors $|i, -i\rangle$, etc.

4. TENSOR OPERATORS IN THE GROUP ALGEBRA

The character analysis and projection operator methods of group representation theory are usually employed in conjunction with some operand set of functions ψ_j. If linear operators X_j are considered instead of functions, the traditional formalism still holds if $R\psi_j$ is replaced by RX_jR^{-1}, R being a group operator. This result is typical of quantum mechanical transformation theory. One way to establish it is to make the "obvious" convention that a set of operators T_M^K are to be termed tensor operators associated with rep type K if they act on an arbitrary A_1 type function ϕ to give a set of K rep functions ϕ_M^K. We then require that, for any group operator R,

$$R(T_M^K\phi) = \sum D_{MN}^K(R)(T_N^K\phi), \tag{4.1}$$

where the D^K are the rep matrices. This requirement may be rewritten

$$(RT_M^KR^{-1})(R\phi) = \left(\sum D_{MN}^K(R)T_N^K\right)\phi. \tag{4.2}$$

Since ϕ is of A_1 type we have $R\phi = \phi$ on the left-hand side; ϕ is otherwise arbitrary, and this implies that we must have

$$RT_M^KR^{-1} = \sum D_{MN}^K(R)T_N^K, \tag{4.3}$$

which accords with one traditional way of defining tensor operators for a group. This basic result then leads to the general rule stated at the beginning of this section. It is possible in some cases to take a class of a finite group and construct linear combinations of the individual group operators within it so that they obey (4.2), i.e., constitute tensor operators. Such a procedure uses the group elements both as operators and as members of a basis set (i.e., involves the group algebra), and is reminiscent of the process used in forming the regular representation of a group.[6] Each class clearly contains an A_1 type operator, namely the class sum operator: The use of traditional character analysis methods, modified as explained above, shows that the classes of O contain the following tensor operator types:

$$(1) \to A_1 \qquad (2) \to A_1 + A_2 + T_1 + T_2$$
$$(3) \to E + A_1, \qquad (4) \to A_1 + E + T_2$$
$$(5) \to A_1 + E + T_1.$$

A similar procedure for D_4 yields the results

$$(1) \to A_1, \qquad (2) \to A_1,$$
$$(3) \to A_1 + A_2, \qquad (4) \to A_1 + B_1,$$
$$(5) \to A_1 + B_2.$$

These operator types at first sight appear to be inconsistent with the class multiplication properties, but this

is not so. For example, in the group D_4, the Cartesian product of class 3 with itself gives classes 1 and 2, as indicated by the result $\mathcal{C}_2\mathcal{C}_2 = \frac{1}{2}\mathcal{C}_1 + \mathcal{C}_2$ for the group D_4. The Kronecker product of $A_1 + A_2$ with itself contains an A_2 part; but classes 1 and 2 contain no such operator, according to the result displayed above. As an example of the construction of the tensor operators, we deal here with the A_1 and A_2 operators of class 3 for D_4, and also show how to resolve the apparent contradiction just described above. It follows from the discussion at the beginning of this section that a tensor operator belonging to the μth rep can be obtained by taking some element X of a class and forming the sum

$$\sum_R \chi^{(\mu)*}(R)RXR^{-1},$$

where R runs over all group operators. The resulting operator is clearly a linear combination of group operators from the class of X. The A_1 rep of any group has all characters equal to unity and gives a multiple of the class sum operator as the A_1 tensor operator. For D_4 this is the sum $C_4 + C_4^3$ for class 3. The A_2 rep of D_4 has half the characters (those for which $RC_4R^{-1} = C_4$) equal to 1 and half of them (those for which $RC_4R^{-1} = C_4^3$) equal to -1. The resulting A_2 tensor operator for class 3 is a multiple of $C_4 - C_4^3$.

The Kronecker product of $A_1 + A_2$ with itself can now be expressed alternatively in terms of these explicit tensor operators:

$$[A_1; A_2]^2 \equiv [C_4 + C_4^3; C_4 - C_4^3]^2.$$

The A_2 terms in this product are of the form

$$(C_4 + C_4^3)(C_4 - C_4^3) = C_2 - E + E - C_2$$

and thus vanish identically within the group algebra. Thus the apparent contradictions between the results of this section and the class multiplication results of Sec. 2 are resolved by the explicit vanishing of some of the formally allowed operator types. For the group O, the results displayed above show that each type of tensor operator has a realization within the group algebra, whereas for D_4 no E type operator can be found in the D_4 group algebra. This can be seen by noting that each class of D_4 contains only one or two operators, and must contain the A_1 class sum operator. There is thus no possibility of forming the two extra E type operators within any class. In the case of the group D_{3h}, this dimensionality argument does not apply, and yet it is still found[1] that only certain tensor operator types can be realized in the group algebra. It is necessary that the "allowed" rep types should form a closed family under Kronecker product formalism, as the specific results of this paper illustrate. Whether this condition is sufficient is a problem to be investigated, together with several other group structural problems which have been suggested by the class sum operator approach.

[1] J. Killingbeck, J. Math. Phys. **11**, 2268 (1970).
[2] P-O. Löwdin, Rev. Mod. Phys. **34**, 520 (1962).
[3] G. G. Hall, *Applied group theory* (Longmans Green, London, 1967.
[4] A. Gamba, J. Math. Phys. **9**, 186 (1968).
[5] G. de B. Robinson, J. Math. Phys. **12**, 2212 (1971).
[6] J. Killingbeck and G. H. A. Cole, *Mathematical techniques and physical applications* (Academic, New York, 1971).

The mathematical problem of reflection solved by an extension of the WKB method

B. Chakraborty
Department of Mathematics, Jadavpur University, Calcutta-32, India
(Received 11 June 1971; revised manuscript received 3 August 1971)

Starting with two first order linear differential equations having slowly varying coefficients and mutually connected by the dependent variables, the reflection problem is solved approximately as in the cases of electromagnetic and acoustic waves. An extension of the WKB method is developed and applied to study this problem up to any higher order of accuracy one requires. The solution of the second order differential equation in normal form by the extended WKB method is used to find the characteristics of propagation at a point of discontinuity of higher order derivative of the parameter.

1. INTRODUCTION

Since Bremmer[1] initiated the study of reflection and transmission of electromagnetic waves through slowly varying dielectrics by the WKB method, much work has been done on various aspects of the theory of the WKB approximation[2-7] and its application to dielectrics and plasmas.[8-12] The necessity of finding a suitable extension of this method to the study of the very important problems of qualitative and quantitative estimation of coupling of different types of modes of propagation and scattering of radiation of one type of wave field by other types in slowly varying multicomponent plasma cannot be in any doubt.

The dielectric and plasma field equations are all of first order and the slowly varying parameters exist in them in one form or the other. It is therefore evident that all the peculiar tendencies found in the derived second and higher order wave equations should be hidden in the original first order field equations. Hence an analysis of a model set of first order equations containing the characteristic essential features and giving scope for an extended investigation is thought to be useful.

Our modest aim for this paper is the solution of the reflection problem as generally as possible with the help of two first order linear differential equations with slowly varying coefficients. An extension of the WKB method is developed and used to solve the problem up to any higher order of accuracy one requires. The results can be easily applied to the cases considered in Refs. 1, 8–10 with the gain of additional clarity. Moreover it would be evident that most of the properties previously associated exclusively with reflection and transmission of waves through optically transparent media can now be thought of as of general nature and are therefore properties of any pair of first order differential equations of the type considered in this note.

2. THE REFLECTION EQUATIONS AND THEIR WKB SOLUTIONS

We consider the two coupled first order differential equations

$$\alpha^2 u \pm i v' = 0, \tag{2.1}$$

$$\beta^2 v \pm i u' = 0, \tag{2.2}$$

where u and v are perturbation variables, α^2 and β^2 are slowly varying functions of the independent variable Σ and a primed quantity means derivative of the corresponding unprimed quantity with respect to Σ. For problems of propagation, α^2 and β^2 are both nonzero, their real parts should be of the same sign, and the magnitude of the real part must be much greater than that of the corresponding imaginary part if α^2 and β^2 are complex quantities. The second order equations for u and v are

$$u'' - (2\log\beta)'u' + \alpha^2\beta^2 u = 0, \tag{2.3}$$

$$v'' - (2\log\alpha)'v' + \alpha^2\beta^2 v = 0. \tag{2.4}$$

For real α^2 and β^2 the WKB solution for u is

$$u = [A\exp(i\Sigma_1) + B\exp(-i\Sigma_1)]/\mu, \tag{2.5}$$

where

$$\Sigma_1 = \int \alpha\beta d\Sigma, \tag{2.6}$$

$$\mu = (\alpha/\beta)^{1/2}. \tag{2.7}$$

Putting this solution in (2.2) and solving for v, we get

$$v = \pm\,\mu[A\exp(i\Sigma_1) - B\exp(-i\Sigma_1)] \pm \tfrac{1}{2}i\mu(\mu^2)', \tag{2.8}$$

where now a primed quantity means derivative with respect to Σ_1. Let

$$u_i = (1/\mu)A\exp(i\Sigma_1), \tag{2.9}$$

$$u_r = (1/\mu)B\exp(-i\Sigma_1), \tag{2.10}$$

then u_i and u_r can be regarded as solutions, respectively, for the forward-going (or incident) and backward-going (or reflected) wavefields. If v_i and v_r are the similar components of v, then isolating them from the solution (2.8), we get

$$v_i = \pm\,\tfrac{1}{2}u_i[i(\mu^2)' + 2\mu^2], \tag{2.11}$$

$$v_r = \pm\,\tfrac{1}{2}u_r[i(\mu^2)' - 2\mu^2]. \tag{2.12}$$

By using the upper sign in the right-hand side of (2.1), (2.8), (2.11), and (2.12) and the solutions (2.9) and (2.10) after some simplification, we obtain the equations

$$u_i' + u_r' = -(u_i + u_r)(\log\mu)' + i(u_i - u_r), \tag{2.13}$$

$$u_i' - u_r' = -(u_i - u_r)(\log\mu)' - i(u_i - u_r)[\mu(1/\mu)'' - 1]. \tag{2.14}$$

Elimination of u_r' and u_i' in turn yields the following coupled equations between the incident and the reflected waves

$$u_i' + [-i + \tfrac{1}{2}i\mu(1/\mu)'' + (\log\mu)']u_i = -\tfrac{1}{2}i\mu(1/\mu)''u_r, \tag{2.15}$$

$$u_r' + [i - \tfrac{1}{2}i\mu(1/\mu)'' + (\log\mu)']u_r = \tfrac{1}{2}i\mu(1/\mu)''u_i. \tag{2.16}$$

The coefficients of u_r in (2.15) and of u_i in (2.16) may be called the coupling coefficients for the incident and

reflected fields, respectively, and they depend on second-order derivatives of u.

Since the value of v in (2.8) was calculated by using Eq. (2.2), the residue of solutions (2.5) and (2.8) can be obtained by simplifying; after putting them on the left-hand side of (2.1), this gives

$$u + \frac{i}{\alpha^2}\frac{du}{d\Sigma} = -\frac{(u^2)''}{2u^2}u, \tag{2.17}$$

which is a quantity of the order neglected in the WKB solution (2.5). The coefficient $(u^2)''/2u^2$ for u on the right-hand side of (2.17) can be called the residue coefficient, and it should always be of the order of quantities neglected in the approximate solution for u.

If A and B of (2.5) are regarded as slowly varying functions of Σ_1, then substitution of (2.5) and (2.8) in (2.15) and (2.16) means that

$$\frac{dA}{d\Sigma} = \epsilon p[A + B\exp(-2i\Sigma_1)], \tag{2.18}$$

$$\frac{dB}{d\Sigma} = -\epsilon p[B + A\exp(2i\Sigma_1)], \tag{2.19}$$

where

$$\epsilon p = -\tfrac{1}{2}iu(1/u)''. \tag{2.20}$$

By following Bremmer,[1] the series solution

$$A = \sum_{r=1}^{\infty}\epsilon^r A_r, \qquad B = \sum_{r=1}^{\infty}\epsilon^r B_r \tag{2.21}$$

is sought. By equating coefficients of Σ^{m+1} from both sides, the recursion relations are

$$\frac{d}{d\Sigma_1}A^{m+1} = p[A_m + B_m\exp(-2i\Sigma_1)], \tag{2.22}$$

$$\frac{d}{d\Sigma_1}B^{m+1} = -p[B_m + A_m\exp(2i\Sigma_1)]. \tag{2.23}$$

The sequence of these solutions, being linear in A_0 and B_0, are superposable.

An exact solution of (2.18) and (2.19), maintaining some contribution from the nonuniformity of α^2 and β^2, exist if p is regarded as constant and $|B| \ll |A|$. Then ignoring the second term on the right-hand side of (2.8) the solutions are

$$A = A_0\exp(-ip\Sigma_1),$$

$$B = [\epsilon pA_0\exp(2i\Sigma_1 - i\epsilon p\Sigma_1)]/2(1-\epsilon p). \tag{2.24}$$

For $\epsilon = 1$, we get

$$u = A_0([\exp i(1+2p)^{1/2}\Sigma_1] + p\{\exp[-i(1+2p)^{1/2}\Sigma_1 - 2i\Sigma_1]\})\exp(-ip\Sigma_1)/[k(1+2p)]^{1/2}, \tag{2.25}$$

when $|2p| \ll 1$ this relation simplifies to

$$u = A\,(\exp i\Sigma_1)([\exp(-\tfrac{1}{2}ip^2\Sigma_1)] + p\{\exp[-i(2p-\tfrac{1}{2}p^2)\Sigma_1]\}/2(1-p))/[k(1+2p)]^{1/2}. \tag{2.26}$$

Thus the socalled secondary reflected wave obtained in the fashion of Bremmer[1] is only a higher order correction term of the primary incident wave. The fact that the power of the exponent of Σ_1 for the secondary field in (2.25) has sign opposite to that for the primary is deceptive as is shown by the expansion. To determine the amplitude of the reflected and transmitted fields, it becomes necessary to use boundary conditions to be satisfied at the point Σ_1 where α^2 and β^2 or any of their derivatives are discontinuous.

3. AN EXTENSION OF THE WKB METHOD

We consider the normal form equation

$$\frac{d^2}{d\Sigma^2}\psi(\Sigma) + K^2(\Sigma)\psi(\Sigma) = 0, \tag{3.1}$$

where K^2 is the slowly varying function of Σ.

Using the substitutions

$$\Sigma_1 = \int K d\Sigma, \qquad \psi_1 = (K)^{1/2}\psi \tag{3.2}$$

Eq. (3.1) reduces to

$$\frac{d^2}{d\Sigma_1^2}\psi_1(\Sigma_1) + K_1^2(\Sigma_1)\psi_1(\Sigma_1) = 0, \tag{3.3}$$

where

$$K_1^2 = 1 + \epsilon_2, \qquad \epsilon_2 = -\frac{1}{(K)^{1/2}}\frac{d^2}{d\Sigma_1^2}(K)^{1/2}. \tag{3.4}$$

The original equation (3.1) is thus reduced to (3.3) containing the coefficient K_1^2 which is the sum of unity and ϵ_2—a function proportional to second-order derivative of $(K)^{1/2}$ with respect to Σ_1. If therefore

$$|\epsilon_2| \ll 1, \tag{3.5}$$

the solution of (3.3) is the usual WKB solution

$$\psi = [A\exp(i\int K d\Sigma) + B\exp(-i\int K d\Sigma)]/(K)^{1/2}. \tag{3.6}$$

For a more general solution than (3.6) let

$$\Sigma_2 = \int K_1 d\Sigma = \int K(1+\epsilon_2)^{1/2}d\Sigma,$$
$$\psi_2 = (K_1)^{1/2}\psi_1 = (K)^{1/2}(1+\epsilon_2)^{1/4}\psi. \tag{3.7}$$

Then (3.3) reduces to

$$\frac{d^2}{d\Sigma_2^2}\psi_2(\Sigma_2) + K_2^2(\Sigma_2)\psi_2(\Sigma_2) = 0, \tag{3.8}$$

where

$$K_2^2 = 1 + \epsilon_4, \qquad \epsilon_4 = -\frac{1}{(1+\epsilon_2)^{1/4}}\frac{d^2}{d\Sigma_2^2}(1+\epsilon_2)^{1/4}. \tag{3.9}$$

If $|\epsilon_4| \ll 1$, a solution better than (3.6) can be constructed.

By further generalizing we can write

$$\frac{d^2}{d\Sigma_r^2}\psi_r + K_r^2\psi_r = 0, \tag{3.10}$$

where

$$\psi_r = (K_{r-1})^{1/2}\psi_{r-1} \tag{3.11}$$

$$= [K^2(1+\epsilon_2)(1+\epsilon_4)\cdots(1+\epsilon_{2r-2})]^{1/4}\psi, \tag{3.12}$$

$$d\Sigma_r = [K^2(1+\epsilon_2)(1+\epsilon_4)\cdots(1+\epsilon_{2r-2})]^{1/2}d\Sigma, \quad (3.13)$$

$$K_r^2 = (1+\epsilon_{2r}). \quad (3.14)$$

Then

$$\psi = (A\exp\{i\int[K^2(1+\epsilon_2)(1+\epsilon_4)\cdots(1+\epsilon_{2r})]^{1/2}d\Sigma + B\exp\{-i\int[K^2(1+\epsilon_2)(1+\epsilon_4)\cdots(1+\epsilon_{2r})]^{1/2}\}) /[K^2(1+\epsilon_2)(1+\epsilon_4)\cdots(1+\epsilon_{2r})]^{1/4}, \quad (3.15)$$

where

$$\epsilon_{2r+2} = -\frac{1}{(1+\epsilon_{2r})^{1/4}}\frac{d^2}{d\Sigma_r^2}(1+\epsilon_{2r})^{1/4} \quad (3.16)$$

and

$$|\epsilon_{2r+2}| \ll 1. \quad (3.17)$$

The solution (3.15) of (3.1) is correct up to the rth order of derivatives of the parameter K^2 with respect to Σ_r.

4. THE GENERAL REFLECTION PROBLEM

By using the extension of the WKB method developed above, the general solution of u of Eq. (2.3) correct up to rth order of derivatives of α^2 and β^2 is

$$u = [A\exp(i\Sigma_0) + B\exp(-i\Sigma_0)]/u_r, \quad (4.1)$$

where

$$u_r^2 = u^2K_1K_2\cdots K_r, \quad (4.2)$$

$$d\Sigma_0 = \alpha\beta K_1K_2\cdots K_r, \quad (4.3)$$

$$K_r^2 = 1 - \frac{1}{(K_{r-1})^{1/2}}\frac{d^2}{d\Sigma_r^2}(K_{r-1})^{1/2}, \quad (4.4)$$

with

$$K_1^2 = 1 - \frac{1}{u}\frac{d^2}{d\Sigma_1^2}u, \quad (4.5)$$

$$\Sigma_1 = \int\alpha\beta d\Sigma. \quad (4.6)$$

Then the first order equations connecting the forward and backward going solutions are

$$u_i^1 + \left(-i + \frac{i}{2}\frac{u_r''}{u_r} + \frac{u_r'}{u_r}\right)u_i = -\frac{i}{2}\frac{u_r''}{u_r}u_r, \quad (4.7)$$

$$u_r' + \left(i - \frac{i}{2}\frac{u_r''}{u_r} + \frac{u_r'}{u_r}\right)u_r = \frac{i}{2}\frac{u_r''}{u_r}u_i, \quad (4.8)$$

where now a primed quantity means the derivative of the corresponding unprimed quantity with respect to Σ_0.

The residue of the solution (4.1) for u and for v obtained from (2.2) is

$$u + \frac{i}{\alpha^2}\frac{du}{d\Sigma} = u\left[-\frac{u_r''}{u_r} - \left(1 + \frac{u_r''}{u_r}\right)\left(\sum_{i=1}^{r}\epsilon_{2i} + \sum_{i=1}^{r}\sum_{j=1}^{r}\epsilon_{2i}\epsilon_{2j} + \cdots\right)\right]. \quad (4.9)$$

The recursion relations more general than (2.18) and (2.19) are

$$A' = -\tfrac{1}{2}i(u_r''/u_r)[A + B\exp(-2i\Sigma_0)], \quad (4.10)$$

$$B' = \tfrac{1}{2}i(u_r''/u_r)[B + A\exp(2i\Sigma_0)]. \quad (4.11)$$

Since

$$\frac{u''}{u} - \frac{u_r''}{u_r} = \left(\frac{u'}{u} - \frac{u_r'}{u_r}\right)^2, \quad (4.12)$$

the extended WKB solution (4.1) is better than the usual solution.

5. REFLECTION AT A DISCONTINUITY

We consider the reflection at the layer $\Sigma_r = y$ for the solution (3.15) of ψ. Let the point $\Sigma_r = y$ be specified by the discontinuity of $K_r(\Sigma_r)$ and continuity of K_{r-n} $(n = 1, 2, \ldots, r-1)$, ψ_r and $(d/d\Sigma_r)\psi_r$, and zero or negligible value of $K_{r+n}(n = 1, 2, \ldots, \infty)$ on both sides of $\Sigma_r = y$. Let further $K_{r_1}, K_{r_2}, \epsilon_{2r_1}, \epsilon_{2r_2}$ be the values of K_r and ϵ_{2r}, respectively, on the two sides of $\Sigma_r = y$. For solution of the form

$$\psi_r = \exp\{iK_{r_1}(\Sigma_r - y)\} + R\exp\{-iK_{r_1}(\Sigma_r - y)\}, \quad -\infty < \Sigma_r < y, \quad (5.1)$$

$$= T\exp\{iK_{r_2}(\Sigma_r - y)\}, \quad \infty > \Sigma_r > y, \quad (5.2)$$

the conditions specified above applied at $\Sigma_r = y$ yield

$$T = \frac{2K_{r_1}(y)}{K_{r_1} + K_{r_2}}, \qquad R = \frac{K_{r_1} - K_{r_2}}{K_{r_1} + K_{r_2}}.$$

If ϵ_{2r} is small on both sides of $\Sigma_r = y$, then

$$R \approx \frac{[\epsilon_{2r}''/\epsilon_{2r}]_2^1}{K_{r_1} + K_{r_2}},$$

where $\epsilon_{2r}'' = (d^2/d\Sigma_r^2)\epsilon_{2r}$ and $[P]_2^1 = P_1 - P_2$.

The above expressions are more general than the usual results.[13]

ACKNOWLEDGMENT

The author expresses his thanks to the referee for some useful comments.

[1] H. Bremmer, Physica (Utr.) **15**, 593 (1949).
[2] F. W. Sluijter, J. Math. Anal. Appl. **27**, 2 (1969). It contains important earlier references.
[3] F. W. Sluijter, J. Opt. Soc. Am. **60**, 8 (1970).
[4] L. J. F. Broer and J. B. Van Vroonhoven, Physica **52**, 441 (1971).
[5] J. Heading, *An Introduction to Phase Integral Methods* (Methuen, London, 1962).
[6] N. Fröman and P. O. Fröman, *JWKB Approximation. Contribution to the Theory* (North-Holland, Amsterdam, 1965).
[7] L. J. F. Broer, Proc. Eleventh Intern. Congress Appl. Mech. (Germany), 59 (1964).
[8] J. R. Wait, *Electromagnetic Waves in Stratified Media* (Pergamon, New York, 1962).
[9] K. G. Budden, *Radio Waves in the Ionosphere* (Cambridge, New York, 1961).
[10] D. A. Tidman, Phys. Rev. **117**, 366 (1960).
[11] B. Chakraborty, J. Math. Phys. **11**, 2570 (1960).
[12] B. Chakraborty, J. Math. Phys. **12**, 529 (1971).
[13] R. Bellman, *Perturbation Techniques in Mathematics, Physics and Engineering* (Holt, Rinehart and Winston, New York, 1964), see p. 101, formula (13.10).

Mathieu function solutions to the radial Schrödinger equation for the $-f^2/r^4$ interaction

N. A. W. Holzwarth

The Department of Physics and The James Franck Institute, The University of Chicago, Chicago, Illinois 60637

(Received 21 April 1971; revised manuscript received 1 June 1971)

We describe the Mathieu function solutions to the radial Schrödinger equation for the $-f^2/r^4$ potential with reference to adiabatic elastic scattering of electrons from neutral atoms. By an appropriate choice of boundary conditions, the total scattering phase shift δ_l decomposes into two parts: $\delta_l = \gamma_l + \rho_l$. The "polarizability phase shift" ρ_l depends solely on parameters of the interaction and is easily calculated. The "polarizability extracted phase shift" γ_l contains information of the core interaction. We suggest that such a parametrization of this scattering problem is convenient for data analysis or for potential scattering calculations. We find that the Mathieu function solutions are characterized by a "polarizability range" $r_t = \sqrt{f}/k$. For $r > r_t$, the functions resemble spherical waves with phase shift ρ_l with respect to the free particle wavefunctions; while for $r < r_t$ they deviate strongly from spherical wave forms. The relationship of r_t with respect to the atomic radius grossly determines the behavior of the scattering phase shifts.

I. INTRODUCTION

Although the one-particle Schrödinger equation can be solved numerically for any given interaction, analytic representations of the solutions remain of considerable interest, especially for understanding the parametric dependences of the physical processes that the equation describes. In this paper we derive an analytic method for evaluating the contribution of the adiabatic charge-induced dipole interaction to the phase shifts for low energy elastic scattering. The system we have in mind is that of an electron incident on a gaseous target of neutral atoms; but this analysis is also appropriate for other charge-neutral systems.

For many electron-neutral atom elastic scattering experiments at low energy there is a region of space $r > d$ (outside the charge cloud of the atom) in which, to a good approximation, the effective interaction is represented by the electron-induced dipole potential

$$V(r) = -(\alpha e^2)/(2r^4) \text{ (cgs units)}, \tag{1}$$

where α is the (static) polarizability of the atom. This potential describes the Stark shift of the atomic energy levels by the dipole component of the electric field of the incident electron. According to Kleinman, Hahn, and Spruch[1] and Callaway, LaBahn, Pu, and Duxler,[2] contributions to the effective interaction from the quadrupole component of the electric field of the electron and from nonadiabatic interactions of the incident electron with the atomic electrons take the form

$$-(\alpha_q - 6a_0\beta)(e^2/2r^6) \text{ (cgs units)}.$$

These authors have shown that for He the atomic quadrupole polarizability α_q and the distortion parameter $6a_0\beta$ have roughly equal magnitude. Thus, at least for the e^--He system, Eq. (1) may describe the effective interaction with fair accuracy. For our purposes, since major corrections to (1) vanish as (const)$/r^6$, we assume that a distance d can be chosen such that (1) represents the interaction to arbitrary accuracy for all $r > d$. Therefore, in this region of space the relevant radial Schrödinger equation is

$$\left(\frac{1}{r^2}\frac{d}{dr}r^2\frac{d}{dr} - \frac{l(l+1)}{r^2} + \frac{f^2}{r^4} + k^2\right)\psi_l(r;k,f) \underset{r>d}{=} 0, \tag{1'}$$

where $f^2 = (-\alpha e^2/2)(2m/\hbar^2) = (\alpha/a_0)$ Bohr^2 and $k^2 = E(2m/\hbar^2)$ Bohr^{-2}. We seek two linearly independent solutions $\Phi_l^1(r;k,f)$ and $\Phi_l^2(r;k,f)$ of Eq. (1′) such that, without loss of generality, we can write the desired solution to a given electron-atom scattering problem, for $r > d$, in the form[3]

$$\psi_l(r;k,f) \underset{r>d}{=} \sqrt{(2k/\pi)}\,\{\cos\gamma_l\Phi_l^1(r;k,f) - \sin\gamma_l\Phi_l^2(r;k,f)\}. \tag{2}$$

However, it is clear from electron-neutral atom elastic cross section data that the form of an actual wavefunction at very large distances from the origin (at the target) is

$$\psi_l(r;k,f) \underset{kr\to\infty}{\sim} \sqrt{(2k/\pi)}\,[\sin(kr - l\pi/2 + \delta_l)/kr]. \tag{3}$$

The total scattering phase shift δ_l may be related to the parameter γ_l by equating the asymptotic form of (2) to Eq. (3). In particular, since the form of the differential equation (1′) with $k^2 > 0$ dictates that the functions $\Phi_l^{1,2}(r;k,f)$ become asymptotically spherical waves, we can conveniently require the normalization:

$$\begin{aligned}\Phi_l^1(r;k,f) &\underset{kr\to\infty}{\sim} \{\sin[kr - \tfrac{1}{2}l\pi + \rho_l(k,f)]/kr\},\\ \Phi_l^2(r;k,f) &\underset{kr\to\infty}{\sim} \{-\cos[kr - \tfrac{1}{2}l\pi + \rho_l(k,f)]/kr\},\end{aligned} \tag{4}$$

where the "polarization" phase shifts $\rho_l(k,f)$ are characteristic of the magnitudes of the parameters in the differential equation (1′). Hence, the parameters γ_l, which we term "polarization extracted phase shifts," are simply related to the experimental phase shifts by

$$\delta_l = \gamma_l + \rho_l. \tag{5}$$

It will be shown that in this formulation ρ_l can contribute up to a real magnitude of $\pi/4$ radians to the total scattering phase shift.

Although Eq. (1′) contains an irregular singularity at $r = 0$ for $f \neq 0$, the limit of (1′) as $f \to 0$ for finite k is well defined, so that its solutions may be chosen such that [in addition to normalization condition (4)]

$$\begin{aligned}\Phi_l^1(r;k,f=0) &= j_l(kr),\\ \Phi_l^2(r;k,f=0) &= y_l(kr).\end{aligned} \tag{6}$$

That is, $\Phi_l^{1,2}(r;k,f)$ may be constructed to go smoothly to the normal free particle solutions, spherical Bessel and Neumann functions, as the atom becomes nonpolarizable. Thus we can establish an absolute scale for the

"polarization extracted phase shift" from the requirement that $\gamma_l(k, f = 0) = \delta_l(k, f = 0)$. Fortunately, the solutions $\Phi_l^{1,2}(r; k, f)$ vary smoothly throughout the physical range of the parameters l, k, and f; that is, they tend toward the well-known radial Mathieu functions of integral order as the parameters tend to certain critical values, in addition to satisfying boundary conditions (4) and (6) above.

The scattering phase shifts δ_l may be determined by fitting experimental cross section data.[4] Recently there has been substantial progress in the accurate determination of "experimental" phase shifts δ_l. For example, Bransden and McDowell[5] have determined δ_0 and δ_1 for electrons on He for energies below the first excitation threshold with an accuracy of 5%. In addition, several experimental groups, e.g., Gibson and Dolder,[6] are now repeating the original work of Ramsauer and Kollath[7] with increased precision. In such cases, Eq. (5) can be utilized to determine γ_l "experimentally." Once γ_l is determined one can calculate the logarithmic derivative of the core wavefunction $R_l(d)$ at the point $r = d$:

$$R_l(d) = \frac{[(d/dr)\Phi_l^1(r; k, f)]_{r=d} - \tan\gamma_l[(d/dr)\Phi_l^2(r; k, f)]_{r=d}}{\Phi_l^1(d; k, f) - \tan\gamma_l\, \Phi_l^2(d; k, f)} \tag{7}$$

Alternatively, if $R_l(d)$ is known for a model core potential, Eqs. (7) and (5) may be utilized to determine the corresponding scattering phase shift δ_l. This procedure avoids some numerical integration in potential scattering calculations. In the limit of very low energy, the analysis reduces to the "modified effective range expansion" of O'Malley, Spruch, and Rosenberg.[8]

This formulation of the scattering problem suggests a speculation on the possibility of extrapolating, from a determined set of $\{\gamma_l(k, f)\}$, a new set of phase shifts $\{\gamma_l'(k, f')\}$ relevant to the same atomic target with reduced or zero polarizability. In the future we plan to apply the formalism reported to the comparison of excess electron–atom interactions in gaseous and solid insulators. In this case, to a fair approximation, the short range atomic core potential remains intact upon solidification; but the polarization interactions are strongly modified, as the density changes. In the muffin-tin representation of a crystal potential, in fact, all long range parts of the electron–atom interaction are discarded except in the sense that they slightly modify the effective core potential.

In Sec. II we review the properties of Mathieu functions as they apply to our problem. We derive from them the solutions $\Phi_l^{1,2}(r; k, f)$ and the "polarization phase shifts" $\rho_l(k, f)$. In Sec. III we discuss the parametric dependences of the solutions $\Phi_l^{1,2}(r; k, f)$ on the incident electron energy or on the target polarizability.

In Sec. IV we compare the results of Secs. II and III with the work of O'Malley, Spruch, and Rosenberg,[8] and Berger, Snodgrass, and Spruch,[9] and Hinckelmann and Spruch.[10]

II. MATHIEU FUNCTION SOLUTIONS TO THE $-f^2/f^4$ POTENTIAL SCATTERING PROBLEM[11-13]

To demonstrate the relationship between (1′) and Mathieu's equation, define

$$P_l^{1,2}(r) \equiv \sqrt{kr}\ \Phi_l^{1,2}(r; k, f). \tag{8}$$

Equation (1′), in terms of the functions $P_l^{1,2}(r)$, becomes

$$\left(r^2\frac{d^2}{dr^2} + r\frac{d}{dr} - (l + \tfrac{1}{2})^2 + \frac{f^2}{r^2} + k^2r^2\right) P_l^{1,2}(r) \underset{r>d}{=} 0. \tag{9}$$

Equation (9) is analogous to the cylindrical Bessel's equation just as Eq. (1′) is analogous to the spherical Bessel's equation. By following introduction of the scaled argument $x = \sqrt{(k/f)}r$, the equation then becomes

$$\left[x^2\frac{d^2}{dx^2} + x\frac{d}{dx} - (l + \tfrac{1}{2})^2 + kf\left(\frac{1}{x^2} + x^2\right)\right] P_l^{1,2}(x) = 0. \tag{9'a}$$

As expected, the properties of the solutions $P_l^{1,2}(x)$ are determined by the value of the angular momentum l and of the quantity kf, the square root of the product of energy and polarizability.

With the substitution $x = e^{\mu}$, Eq. (9′a) takes the form of the radial Mathieu equation

$$\left(\frac{d^2}{d\mu^2} - [(l + \tfrac{1}{2})^2 - 2kf\cosh 2\mu]\right) P_l^{1,2}(\mu) = 0. \tag{9'b}$$

Equations of the form (9b) have been extensively studied[11-13] in connection with solutions to the Helmholtz equation in elliptic coordinates. In this application μ is a radial variable $0 \leq \mu < \infty$, and the parameters corresponding to $(l + \frac{1}{2})^2$ and $2kf$ in Eq. (9′b) often have restricted values in order to satisfy the condition that the associated equation for the angular variable have periodic solutions, thereby defining the Mathieu functions of the second kind (and integral order). For our purposes, however,

(i) $-\infty < \mu < \infty$ since $0 \leq r < \infty$, and

(ii) periodicity of the solutions is not required so that, in general, solutions of nonintegral order are obtained according to the given magnitudes of the parameters l and kf.[14]

In the following, for the sake of clarity, we define a notation associated with each series representation of the Mathieu functions.

For most values of the parameters, two linearly independent solutions to Eq. (9′a) may be written

$$\mathfrak{M}_{\pm\tau}(x) = \sum_{n=-\infty}^{\infty} C_n(\tau)x^{\pm(\tau+2n)}. \tag{10}$$

This is demonstrated in Appendix A. The "characteristic exponent" $\tau = \tau_l(kf)$ is determined by the relation

$$\sin^2\tfrac{1}{2}\pi\tau = \tfrac{1}{2}\Delta^l(\tau = 0; kf), \tag{11}$$

where $\Delta^l(\tau = 0; kf)$ is an infinite determinant (called the "Hill determinant") which may be easily evaluated by use of recurrence relations. The coefficients $C_n(\tau)$ may be determined from the continued fractions (A11). For convenience, we choose the normalization $C_0(\tau) = 1$. From the above relations, the symmetry of the Mathieu functions about the point $x = 1$ $(r = \sqrt{f/k})$ is apparent[12]:

$$\mathfrak{M}_{+\tau}(1/x) = \mathfrak{M}_{-\tau}(x). \tag{12}$$

Solutions (10) cease to be linearly independent when τ achieves integral values. This special case will be

treated at the end of the section. Unless otherwise stated, in the following, τ is assumed to be nonintegral.

Although series (10) formally converges in N terms for N such that, at worst, $(x/2N)^2 \ll 1$ for $x > 1$ or $[(1/x)/2N]^2 \ll 1$ for $x < 1$, this representation becomes computationally inconvenient for relatively small or large values of x. For these values of the argument the Bessel product series or Bessel series representations may be used.[11-13]

The Bessel product series solution has the form

$$\mathcal{J}\mathcal{J}_{\pm\tau}(r) = \sum_{n=-\infty}^{\infty} (-1)^n C_n(\tau) J_{\pm(\tau+n)}(a) J_{\pm n}(b), \tag{13}$$

where a is the larger of $\{kr$ or $f/r\}$ and b is the smaller. The above restriction on the arguments a and b of Eq. (13) is imposed to ensure rapid convergence of the series for all values of r. There are two Bessel series representations termed "even" and "odd." The "even" Bessel series representation has the form

$$\mathcal{J}^e_{\pm\tau}(r) = \sum_{n=-\infty}^{\infty} (-1)^n C_n(\tau) J_{\pm(\tau+2n)}(kr + f/r). \tag{14}$$

The "odd" Bessel series representation has the form[15]

$$\mathcal{J}^o_{\pm\tau}(r) = \frac{|kr - f/r|}{kr + f/r} \sum_{n=-\infty}^{\infty} (-1)^n C_n(\tau)(\tau + 2n) \times J_{\pm(\tau+2n)}(kr + f/r). \tag{15}$$

The series (14) and (15) converge for all values of r, except $r = \sqrt{f/k}$ $(x = 1)$. The coefficients $C_n(\tau)$ and the characteristic exponent τ in series (13)–(15) are identical to those of the simple power series representation (10). All solutions defined in (13)–(15), in general, suffer some sort of discontinuity at the point $r = \sqrt{f/k}$ $(x = 1)$.

It can be shown that the above representations of Mathieu functions are interrelated in the following ways[13]:

$$\mathcal{J}^{\{e\}}_{\pm\tau}(r) = \mathfrak{N}^{\{e\}}_{\{o\}} \mathcal{J}\mathcal{J}_{\pm\tau}(r), \tag{16}$$

where $\mathfrak{N}^e = \mathfrak{N}^e(l; kf) \equiv \sum_{n=-\infty}^{\infty} C_n(\tau)$

and

$$\mathfrak{N}^o = \mathfrak{N}^o(l; kf) \equiv \sum_{n=-\infty}^{\infty} (\tau + 2n) C_n(\tau). \tag{16'a}$$

Further interrelations are[13]

$$\mathfrak{M}_{\pm\tau}(x) \underset{x\gg 1}{=} K^{\pm} \mathcal{J}\mathcal{J}_{\pm\tau}(r) \underset{x>1}{=} \frac{K^{\pm}}{\mathfrak{N}^{\{e\}}_{\{o\}}} \mathcal{J}^{\{e\}}_{\pm\tau}(r),$$
$$\mathfrak{M}_{\pm\tau}(x) \underset{x\ll 1}{=} K^{\mp} \mathcal{J}\mathcal{J}_{\mp\tau}(r) \underset{x<1}{=} \frac{K^{\mp}}{\mathfrak{N}^{\{e\}}_{\{o\}}} \mathcal{J}^{\{e\}}_{\mp\tau}(r) \tag{17}$$

provided neither $\mathfrak{N}^e = 0$ nor $\mathfrak{N}^o = 0$. The constants $K^{\pm}$ may be determined by calculating each series at any convenient matching point. In particular, since $\mathfrak{M}_{+\tau}(x=1) = \mathfrak{M}_{-\tau}(x = 1)$, the ratio of constants K^-/K^+ can be determined from the Bessel product representations alone, provided $\mathfrak{M}_{\tau}(x = 1) \neq 0$:

$$K^-/K^+ = \mathcal{J}\mathcal{J}_{+\tau}(r = \sqrt{f/k}) / \mathcal{J}\mathcal{J}_{-\tau}(r = \sqrt{f/k}). \tag{18}$$

In Table I, we list some representative values of the matching coefficients $\mathfrak{N}^{\{e\}}_{\{o\}}, K^{\pm}$. It is amusing to note that the $\mathfrak{N}^{\{e\}}_{\{o\}}$ constants of proportionality are related to the values and derivatives of the functions $\mathfrak{M}_{\pm\tau}(x)$ at $x = 1$:

$$\mathfrak{N}^e = \mathfrak{M}_{\pm\tau}(x = 1),$$
$$\mathfrak{N}^o = \pm \frac{d\mathfrak{M}_{\pm\tau}}{dx}\bigg|_{x=1}. \tag{16'b}$$

TABLE I. Some representative polarization phase shifts and matching coefficients.

kf	l	τ	ρ_l(rad)	$\mathfrak{N}^e$	$\mathfrak{N}^o$	K^+	K^-
0.25	0	0.5417	−0.0655	0.829	0.627	1.859	0.866
0.35	0	0.5824	−0.1294	0.752	0.709	1.7675	0.980
0.45	0	0.6387	−0.2179	0.662	0.817	1.728	1.0835
0.5	0	0.6743	−0.2738	0.607	0.884	1.724	1.146
0.6945	0	0.9985	−0.7673	0.0295	1.5185	1.985	1.958
0.25	1	1.4917	0.0130	1.092	1.361	10.511	−0.444
0.5	1	1.4671	0.0518	1.1675	1.235	6.107	−0.745
1.671	1	1.0124	0.7660	1.518	0.0386	2.201	−2.153
0.25	2	2.4988	0.0019	1.0265	2.436	106.156	0.074
0.5	2	2.4952	0.0076	1.058	2.366	44.463	0.177
3.245	2	2.0131	0.7648	2.373	0.0865	4.038	3.883
0.25	3	3.4996	0.00061	1.012	3.461	1487.995	−0.0074
0.5	3	3.4984	0.00249	1.025	3.421	441.830	−0.0249
0.25	4	4.4998	0.00028	1.007	4.471	26792.385	0.00053
0.5	4	4.4993	0.00114	1.014	4.443	5629.356	0.00251

Thus $\mathfrak{N}^e$ or $\mathfrak{N}^o$ [and consequently $\mathcal{J}^e_{\pm\tau}(r)$ and $\mathcal{J}^o_{\pm\tau}(r)$] vanish whenever the simple series solutions or their derivatives with respect to x, respectively, go to zero at $x = 1$. This, in general, occurs whenever τ attains integral values.

We have thus related three equivalent forms of Mathieu functions. Although there is some redundancy in the presentation, all three forms have been included for completeness and with the hope that these several forms have nonoverlapping domains which are useful for numerical computation.

Since we are now able to evaluate the functions $\mathfrak{M}_{\pm\tau}(x)$ over the entire range of r, we can form the appropriate linear combinations $P^{1,2}_l(r)$ which satisfy boundary conditions (4) and (6). The asymptotic form of a Bessel function of order $\pm p$ is[16]

$$J_{\pm p}(z) \underset{z\to\infty}{\sim} \sqrt{2/\pi z} \cos(z \mp \pi p/2 - \pi/4)[1 + O(1/z)]. \tag{19}$$

Consequently, from relations (17), and noting that $C_0(\tau) \equiv 1$, it is apparent that the asymptotic forms of $\mathfrak{M}_{\pm\tau}(x)$ are

$$\mathfrak{M}_{\pm\tau}(x) \underset{r\to\infty}{\sim} K^{\pm}\sqrt{2/\pi kr} \cos(kr \mp \pi\tau/2 - \pi/4)[1 + O(1/kr)], \tag{20a}$$

$$\mathfrak{M}_{\pm\tau}(x) \underset{r\to 0}{\sim} K^{\mp}\sqrt{2/\pi f/r} \cos(f/r \pm \pi\tau/2 - \pi/4) \times [1 + O(1/(f/r))]. \tag{20b}$$

Relations (20a) and (20b) remain valid for integral or complex τ. From (20a) we note that the value of the polarization phase shift must be closely linked to that of the characteristic exponent τ. By a comparison of (20a) and (4), we see that we must form two linear combinations of $\mathfrak{M}_{+\tau}(x)$ and $\mathfrak{M}_{-\tau}(x)$ such that their asymptotic $(r \to \infty)$ forms have equal amplitude but differ in phase by $\pi/2$. From the above relations (4) and (8), it immediately follows that we can choose

$$\Phi^1_l(r; k, f) = \sqrt{\pi/2kr}\,\{[\mathfrak{M}_{\tau}(\sqrt{k/f}\, r)]/K^+\} \tag{21}$$

with alternative representations

$$\Phi_l^1(r;k,f)=\begin{cases}\sqrt{\pi/2kr}\ \mathcal{JJ}_\tau(r;k,f) & \text{for } r\geq\sqrt{f/k}\\ \sqrt{\pi/2kr}\ (K^-/K^+)\ \mathcal{JJ}_{-\tau}(r;k,f) & \text{for } r\leq\sqrt{f/k}\end{cases} \tag{21'a}$$

or, provided $\mathfrak{N}^{\{e\}_o}\neq 0$,

$$\Phi_l^1(r;k,f)=\begin{cases}(1/\mathfrak{N}^{\{e\}_o})\sqrt{\pi/2kr}\ \mathcal{J}_\tau^{\{e\}_o}(r;k,f) & \text{for } r>\sqrt{f/k}\\ (1/\mathfrak{N}^{\{e\}_o})\sqrt{\pi/2kr}\ (K^-/K^+)\ \mathcal{J}_{-\tau}^{\{e\}_o}(r;k,f) & \text{for } r<\sqrt{f/k}\end{cases}. \tag{21b}$$

The second solution can be written

$$\Phi_l^2(r;k,f)=\sqrt{\frac{\pi}{2kr}}\left(\frac{\cos\pi\tau}{\sin\pi\tau}\ \frac{\mathfrak{M}_\tau(\sqrt{k/f}r)}{K^+}-\frac{1}{\sin\pi\tau}\times\frac{\mathfrak{M}_{-\tau}(\sqrt{k/f}r)}{K^-}\right). \tag{22}$$

For convenience we write alternative forms Φ_l^2 directly in Neumann product series form

$$\mathcal{YJ}_{\pm\tau}(r)=\sum_{n=-\infty}^{\infty}(-1)^nC_n(\tau)Y_{\pm(\tau+n)}(a)J_{\pm n}(b), \tag{23a}$$

or Neumann even or odd series form,

$$\mathcal{Y}_{\pm\tau}^{e}(r)=\sum_{n=-\infty}^{\infty}(-1)^nC_n(\tau)Y_{\pm(\tau+2n)}(kr+f/r), \tag{23b}$$

$$\mathcal{Y}_{\pm\tau}^{o}(r)=\frac{|kr-f/r|}{kr+f/r}\sum_{n=-\infty}^{\infty}(-1)^nC_n(\tau)(\tau+2n)\times Y_{\pm(\tau+2n)}(kr+f/r). \tag{23c}$$

Therefore

$$\Phi_l^2(r;k,f)=\begin{cases}\sqrt{\frac{\pi}{2kr}}\ \mathcal{YJ}_\tau(r;k,f) & \text{for } r\geq\sqrt{\frac{f}{k}}\\ -\sqrt{\frac{\pi}{2kr}}\ \frac{K^-}{K^+}\left(\mathcal{YJ}_{-\tau}(r;k,f)-\frac{1-(K^+/K^-)^2}{\sin\pi\tau}\times\mathcal{JJ}_\tau(r;k,f)\right) & \text{for } r\leq\sqrt{\frac{f}{k}}\end{cases} \tag{22'a}$$

or

$$\Phi_l^2(r;k,f)=\begin{cases}\frac{1}{\mathfrak{N}^{\{e\}_o}}\sqrt{\frac{\pi}{2kr}}\ \mathcal{Y}_\tau^{\{e\}_o}(r;k,f) & \text{for } r>\sqrt{\frac{f}{k}}\\ -\frac{1}{\mathfrak{N}^{\{e\}_o}}\sqrt{\frac{\pi}{2kr}}\ \frac{K^-}{K^+}\left(\mathcal{Y}_{-\tau}^{\{e\}_o}(r;k,f)-\frac{1-(K^+/K^-)^2}{\sin\pi\tau}\ \mathcal{J}_\tau^{\{e\}_o}(r;k,f)\right) & \text{for } r<\sqrt{\frac{f}{k}}.\end{cases} \tag{22'b}$$

When τ becomes an integer, $\mathfrak{M}_\tau(x)$ becomes an even or odd function with respect to x and $1/x$. As shown by Ince[17] there can be only one solution of that type for those critical values of the parameters l and kf. Thus, if $\Phi_l^1(r;k,f)$ is chosen as the first solution, Eq. (22) is no longer an independent solution. However, it has been shown[11–13] that the alternative forms (22a) and (22b), evaluated in the limit as $\tau\to$ integer L, remain valid, because they involve Neumann functions of integral order. Since Bessel or Neumann functions of integral order $Z_L(z)$ have the property that $Z_{-L}(z)=(-1)^LZ_L(z)$, it follows that

$$\frac{K^-(\tau=L)}{K^+(\tau=L)}=\pm(-1)^L\quad\text{for}\quad\begin{Bmatrix}\text{even}\\ \text{odd}\end{Bmatrix}\ \mathfrak{M}_L(x). \tag{24}$$

Consequently, the alternative forms of the solutions become

$$\Phi_l^1(r;k,f)=\begin{cases}\sqrt{\pi/2kr}\ \mathcal{JJ}_L(r;k,f), & r\geq\sqrt{f/k},\\ \pm\sqrt{\pi/2kr}\ \mathcal{JJ}_L(r;k,f), & r\leq\sqrt{f/k},\end{cases} \tag{21'a'}$$

or, for whichever solution $\{^e_o\}$ remains finite,

$$\Phi_l^1(r;k,f)=\begin{cases}(1/\mathfrak{N}^{\{e\}_o})\sqrt{\pi/2kr}\ \mathcal{J}_L^{\{e\}_o}(r;k,f), & r>\sqrt{f/k},\\ \pm(1/\mathfrak{N}^{\{e\}_o})\sqrt{\pi/2kr}\ \mathcal{J}_L^{\{e\}_o}(r;k,f), & r<\sqrt{f/k}.\end{cases} \tag{21'b'}$$

Also

$$\Phi_l^2(r;k,f)=\begin{cases}\sqrt{\pi/2kr}\ \mathcal{YJ}_L(r;k,f), & r\geq\sqrt{f/k},\\ -\sqrt{\pi/2kr}\ [\mathcal{YJ}_L(r;k,f)+A\mathcal{JJ}_L(r;k,f)], & r\leq\sqrt{f/k}.\end{cases} \tag{22'a'}$$

and, for whichever solution $\{^e_o\}$ remains finite,

$$\Phi_l^2(r;k,f)=\begin{cases}(1/\mathfrak{N}^{\{e\}_o})\sqrt{\pi/2kr}\ \mathcal{Y}_L^{\{e\}_o}(r;k,f), & r>\sqrt{f/k},\\ -(1/\mathfrak{N}^{\{e\}_o})\sqrt{\pi/2kr}\ [\mathcal{Y}_L^{\{e\}_o}(r;k,f)+A\mathcal{J}_L^{\{e\}_o}(r;k,f)], & r<\sqrt{f/k},\end{cases} \tag{22'b'}$$

where

$$A\equiv(-1)^{L+1}\left[\lim_{\tau\to L}\left(\frac{1-(K^+/K^-)^2}{\sin\pi\tau}\right)\right]. \tag{25}$$

The coefficient A may alternatively be calculated by matching procedure as described in the National Bureau of Standards publication, *Tables relating to Mathieu functions*.[18] If $\Phi_l^1(r;k,f)$ is even, $\mathcal{JJ}_L(r=\sqrt{f/k})$ is finite, and A may be calculated by requiring (22'a') to be continuous at $r=\sqrt{f/k}$:

$$A=-\frac{2\mathcal{YJ}_L(r=\sqrt{f/k})}{\mathfrak{N}^e/K^+}. \tag{26a}$$

If $\Phi_l^2(r;k,f)$ is odd, A may be calculated by requiring the derivative of (22'a') with respect to r to be continuous at $r=\sqrt{f/k}$:

$$A=\frac{2\sqrt{kf}\ \sum_{n=-\infty}^{\infty}(-1)^nC_n(\tau)[Y'_{n+L}(\sqrt{kf})J_n(\sqrt{kf})-Y_{n+L}(\sqrt{kf})J'_n(\sqrt{kf})]}{\mathfrak{N}^0/K^-} \tag{26b}$$

A table of such "joining coefficients" for representative values of the parameters is given in the NBS publication cited.[18]

From the asymptotic forms (20a) and definitions (21) and (22), we see that we have constructed solutions $\Phi_l^{1,2}(r;k,f)$ such that the "polarization phase shift" as defined in Eq. (4) is

$$\rho_l(k,f)=(l+\tfrac{1}{2})\pi/2-\tau\pi/2. \tag{27}$$

Thus, information about the parametric dependence of the polarization phase shift is carried in the Hill determinant $\Delta^l(\tau = 0; kf)$ via relation (11). In general, the solutions to (11) are complex numbers. Let $\tau = \lambda + i\mu$. In terms of these purely real variables Eq. (11) becomes

$$\sin^2(\pi\lambda/2)\cosh^2(\pi\mu/2) - \cos^2(\pi\lambda/2)\sinh^2(\pi\mu/2) + (i/2)\sin\pi\lambda\sinh\pi\mu = \tfrac{1}{2}\Delta^l(0). \quad (28)$$

Since $\Delta^l(\tau = 0; kf)$ is always real the left-hand side of (28) admits three types of solutions:

Case I: $\mu_l = 0$ for $0 \leq \Delta^l(0; kf) \leq 2$.

Case II: $\lambda_l = L_l^o$ (odd integer), $\mu_l \neq 0$ for $\Delta^l(0; kf) > 2$.

Case III: $\lambda_l = L_l^e$ (even integer), $\mu_l \neq 0$ for $\Delta^l(0; kf) < 0$

For physically reasonable values of l and kf, $\Delta^l(0; kf)$ attains values relevant to each of these three cases. (See Fig. 1 and Table II). Below we discuss the poliarization phase shift for each of these cases.

Case I: $\mu = 0; \tau = \lambda$

As seen from Fig. 1, $0 \leq \Delta^l(0; kf) \leq 2$ for small values of kf and relatively large angular momenta. From Eqs. (27) and (28) the polarization phase shift may be directly calculated from the relation

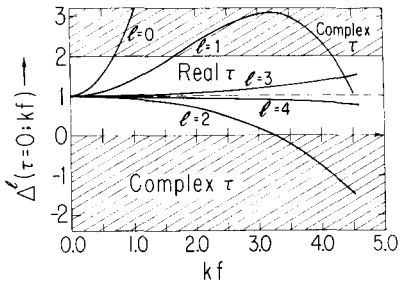

FIG. 1. Graph of Hill determinants $\Delta^l(\tau = 0; kf)$ vs. kf for angular momenta $l = 0, 1, 2, 3, 4$; illustrating regions of real and complex solutions $\Phi_l^{1,2}(r; kf)$.

$$\sin 2\rho_l(kf) = (-1)^{l+1}[\Delta^l(0; kf) - 1]. \quad (29)$$

In particular, it is apparent that $\Delta^l(0; kf \to 0) \to 1$, so that $\rho_l(kf \to 0) \to 0$. For fixed $k \neq 0$, this is the result which is required by boundary condition (7). It has been shown, by expanding the Hill determinant to order $(kf)^2$, that for small values of kf Eq. (29) reduces to[19]

$$\sin 2\rho_l \approx [2\pi(kf)^2/(2l + 3)(2l + 1)(2l - 1)] + O((kf)^4) \quad (30)$$

or

$$\rho_l \approx \pi(kf)^2/(2l + 3)(2l + 1)(2l - 1),$$

which is exactly the Born approximation result for $l \geq 1$.

TABLE II. Values of the parameters for rare gas targets.

			Energy of incident electron in eV at critical values of kf[a]			Values of transition point r_t (Bohr) at representative electron energies $\mathcal{E}$(eV)							
			$l = 0$	$l = 1$	$l = 2$								
	α(Bohr^3)[b]	f(Bohr)	$kf\rfloor_c = 0.695$	$kf\rfloor_c = 1.672$	$kf\rfloor_c = 3.246$	$\mathcal{E} = 0.001$	$\mathcal{E} = 0.005$	$\mathcal{E} = 0.01$	$\mathcal{E} = 0.05$	$\mathcal{E} = 0.1$	$\mathcal{E} = 0.5$	$\mathcal{E} = 1$	$\mathcal{E} = 5$
He	1.35	1.16	4.87	(28.17)	(106.17)	11.64	7.78	6.55	4.38	3.68	2.46	2.07	1.38
Ne	2.70	1.64	2.43	(14.08)	(53.09)	13.84	9.26	7.78	5.21	4.38	2.93	2.46	1.65
Ar	11.07	3.33	0.59	3.44	(12.95)	19.70	13.17	11.08	7.41	6.23	4.17	3.50	2.34
Kr	16.74	4.09	0.39	2.27	8.56	21.84	14.61	12.28	8.21	6.91	4.62	3.88	2.60
Xe	27.26	5.22	0.24	1.39	5.26	24.68	16.50	13.88	9.28	7.81	5.22	4.39	2.93

[a] Without correction for energy dependence of polarizability.
[b] From Ref. 20.

Equation (29) implies that $\rho_l(kf)$ attains a value of $\pm\pi/4$ and suffers a discontinuity in slope (kf) as $\Delta^l(0; kf) \to 2$ or 0, in accordance with τ attaining integral values. These "critical points" occur for the following values of the parameters in the "physical" region:

for $l = 0$, $kf\rfloor_c = 0.694\ 7316$,
$l = 1$, $kf\rfloor_c = 1.671\ 759$,
$l = 2$, $kf\rfloor_c = 3.246\ 0302$.

In Appendix B we discuss the nature of solutions $\Phi_l^{1,2}(r; k, f)$ for integral values of τ. For our purposes it suffices to know that they smoothly evolve from the solutions for nonintegral values of τ as the product kf increases to its "critical" value.

Cases II and III: $\tau = L + i\mu$ (L odd or even integer)

The range of parameters for which the characteristic exponent of the Mathieu functions is complex is often called the "unstable" region because Mathieu functions of the first kind having complex characteristic exponent diverge at one of the two limits of their argument. However, Mathieu functions of the second kind having complex characteristic exponent remain finite over all space, and are thus appropriate to our problem.

In this range, L_l is determined (mod 2) according to whether Case II or III applies. The imaginary part of τ may be calculated from the relation

$$\mu_l = (1/\pi)\ln(|\Delta^l(0) - 1| + \{[\Delta^l(0) - 1]^2 - 1\}^{1/2}). \quad (31)$$

Consequently, the real part of the polarization phase shift is fixed at $\pm\pi/4$ and only the imaginary part of the phase remains sensitive to the parameters of the problem. Since the experimental phase shift is, by definition, real for elastic potential scattering, it follows that

$$\mathrm{Im}\{\gamma_l(k, f)\} = -\mathrm{Im}\{\rho_l(k, f)\} = \mu_l\pi/2. \quad (32)$$

In Appendix B we also discuss the nature of solutions $\Phi_l^{1,2}(r; k, f)$ for complex values of τ. We are consequently also assured that the functions $\Phi_l^{1,2}(r; k, f)$ are well defined for Cases II and III and smoothly evolve from Case I as the value of kf increases, such as occurs in a sequence of scattering experiments with fixed target (fixed polarizability) and increasing energy.

In Table II we list representative values of the parameters for some real scattering systems, namely, electrons incident on rare gas targets. We see that, for example, the s-wave polarizability phase shifts are complex for electron energies $\gtrsim 5$ eV in the case of He, but for electron energies $\lesssim 1/4$ eV in the case of Xe. Although the values listed here have not been adjusted for the energy dependence of the polarizability, we must conclude that the "critical values" of kf and l are realized within the range of ordinary scattering experiments in which the adiabatic dipole response of the atom to the incident electron is a well-established approximation to the dominant long range interaction. In the sense of Eq. (5), then, we have shown that the "polarizability phase shift" can contribute up to the real magnitude of $\pi/4$ radians to the total observed phase shift.

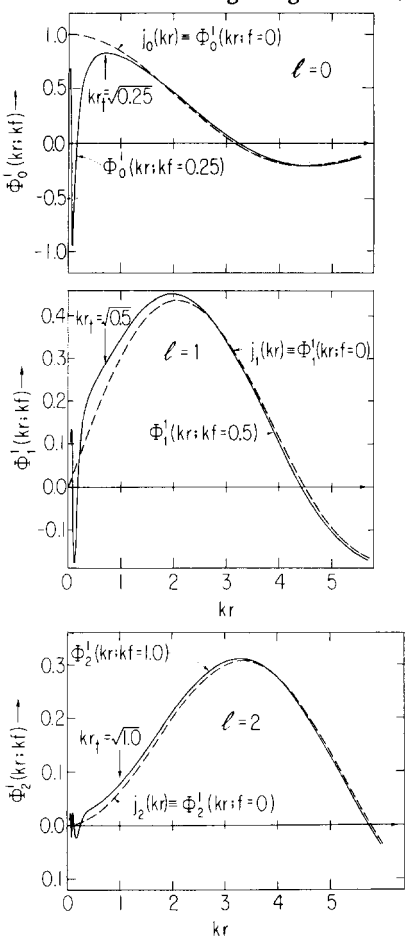

FIG. 2a. The functions $\Phi_l^1(kr; kf)$ for relatively small values of kf (solid line) compared with the corresponding spherical Bessel functions $j_l(kr) \equiv \Phi_l^1(kr; kf = 0)$(dashed line)

Having discussed the mathematical relations of the problem as posed, we now make a few statements about the qualitative nature of the solutions $\Phi_l^{1,2}(r; k, f)$. We have seen that for systems of finite polarizability, and for incident electrons at finite energy, the major properties of the polarization wavefunctions $\Phi_l^{1,2}(r; k, f)$ are determined by the magnitudes of the angular momentum l and of the product kf. The relative magnitudes of the polarizability and total energy figure only in the scale factor $r = \sqrt{k/f}\, x$.

In the sense of the symmetry relation (12) and the Bessel function representations (21'a), (21'b), (22'a), and (22'b), we see that $r_t\,(x = 1)$ is a transition point. For $r < r_t\,(x < 1)$ the functions $\Phi_l^{1,2}(r; k, f)$ begin to oscillate with respect to the argument f/r while for $r > r_t\,(x > 1)$ they achieve a spherical waveform. The separate magnitudes of k and f (of course) determine the "wavelengths" of the large r and small r oscillations, respectively. This is illustrated in Fig. 2 where we have plotted $\Phi_l^{1,2}(r; k, f)$ for small values of kf vs the argument $kr = \sqrt{kf}\, x$. For comparison, corresponding plots of $j_l(kr)$ and $y_l(kr)$ are also shown.

The point r_t has the following physical significance: At this distance the absolute value of the (classical) potential energy has the same magnitude as the total energy $k^2/2$ (in atomic units) so that the classical kinetic energy must be k^2. For $r < \sqrt{f/k}$, the kinetic energy is $> k^2$; for $r > \sqrt{f/k}$ the kinetic energy has an average value between k^2 and $k^2/2$. The kinetic energy is composed of the centripetal kinetic energy $-l(l+1)/r^2$ and the radial kinetic energy associated with the curvature of the radial wavefunction. In addition, we see that, in general, the polarization wavefunctions $\Phi_l^{1,2}(r; k, f)$ must have greater curvature than the free particle wavefunctions. The disparity is greatest when $l = 0$ and the centripetal kinetic energy is zero. We may then suppose that the "difficulty" of complex solutions $\Phi_l^{1,2}(r; k, f)$ for large values of kf is due to the fact that a real series expansion for these values of the parameters cannot assimilate enough curvature to be an eigenfunction of the problem. Instead, an expansion with leading term to the power $(L + i\mu)$, mathematically, has the effect of increasing the centripetal energy of the expression in the sense that it contributes a term of the same form and sign as the centripetal energy operator to the re-

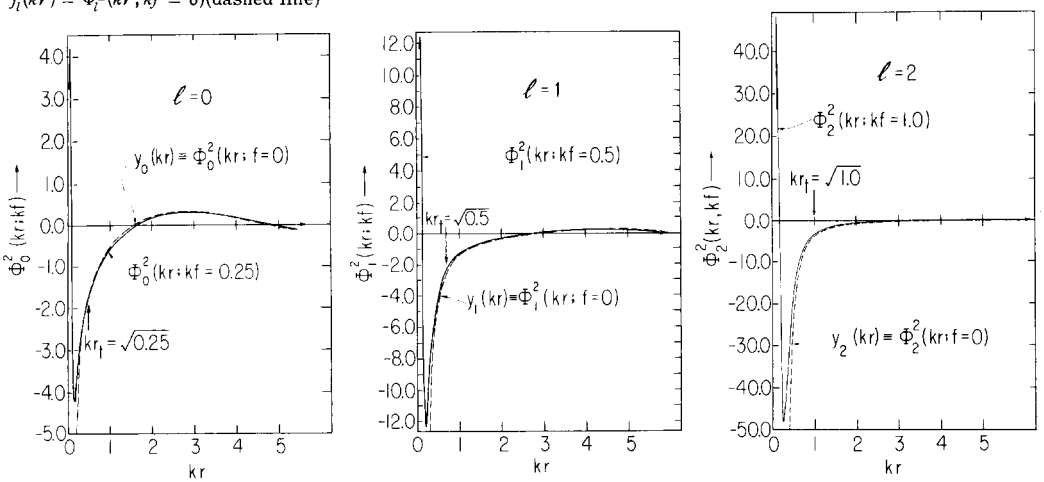

FIG. 2b. The functions $\Phi_l^2(kr; kf)$ for relatively small values of kf (solid line) compared with the corresponding spherical Neumann functions $y_l(kr) \equiv \Phi_l^2(kr; kf = 0)$ (dashed line).

currence relation. For example, relation (B5) becomes (for $\tau = L + i\mu$)

$$\frac{C_1(\tau)}{C_0(\tau)} + \frac{C_{-1}(\tau)}{C_0(\tau)} = \frac{(l+\frac{1}{2})^2 + \mu^2 + L^2 - 2i\mu L}{kf}. \tag{33}$$

In order to gain a qualitative understanding of the role of the polarization phase shift in determining the total phase shift of a given system, one should compare the location of the transition point $r_t = \sqrt{f/k}$ with distances which characterize the physical system. Two such distances are

(a) the mean free path D of the electron,

(b) the radius of the charge cloud of the atom, d, such that for $r > d$, Eq. (1′) adequately describes the system.

For example, we can make the following analysis of the importance of the mean free path of the electron for low energy scattering. In this case the polarization phase shift is given by Eq. (27) only if $D \gg \sqrt{f/k}$. The smallest value of the electron mean free path likely to be encountered is in a diffusion cross section measurement, where the appropriate value of D is approximately the mean free path of the gas atoms. We can thus estimate a conservative lower bound of incident electron energies for which Eq. (30) remains valid. This is

$$\mathcal{E}(\text{eV}) > 13.6 f^2/D^4 \tag{34}$$

where D is expressed in Bohr. For $D \approx 100$ Bohr and $f^2 \approx 30$ Bohr^2, the lowest energy is $\approx 4 \times 10^{-6}$ eV. It is not worthwhile to discuss the possibility of raising this lower bound by decreasing D, since for $D \lesssim 100$ Bohr the electron-induced dipole interaction begins to be screened by neighboring atoms, and this analysis breaks down anyway. We are therefore led to conclude that, for all ordinary experiments[21] to which this analysis applies, the detector "sees" the scattered electron as an outgoing spherical wave of wavelength $2\pi/k$ and phase shift $\delta_l = \gamma_l + \rho_l$. More precisely, it can be inferred that the asymptotic form of the electronic wavefunction (3) for this case is

$$\psi_l(r) \underset{kr\to\infty}{\sim} \sqrt{\frac{2k}{\pi}} \left(U \frac{\sin(kr - l\pi/2 + \delta_l)}{kr} - V \frac{\cos(kr - l\pi/2 + \delta_l)}{kr} \right), \tag{35}$$

where

$$U = 1 - \frac{l(l+1)[l(l+1)-2]}{2^2 2!\,(kr)^2} + \frac{l(l+1)[l(l+1)-2][l(l+1)-6][l(l+1)-12]}{2^4 4!\,(kr)^4} + \frac{(kf)^2}{(kr)^4}\left[\frac{l(l+1)}{2^4} + \frac{[l(l+1)-12]}{2\cdot 4!}\right] + O\left(\frac{1}{(kr)^6}\right)$$

and

$$V = -\frac{l(l+1)}{2kr} + \frac{l(l+1)[l(l+1)-2][l(l+1)-6]}{2^3 3!\,(kr)^3} - \frac{l(l+1)[l(l+1)-2][l(l+1)-6][l(l+1)-12][l(l+1)-20]}{2^5 5!\,(kr)^5} + \frac{(kf)^2}{3!\,(kr)^3} - \frac{(kf)^2}{(kr)^5}\left(\frac{l(l+1)[l(l+1)-2]}{2^4\cdot 5} + \frac{[l(l+1)-20]\{2^3[l(l+1)-12]+4!\,l(l+1)\}}{2^5\cdot 5!}\right) + O\left(\frac{1}{(kr)^7}\right).$$

Thus, as expected from the form of the differential equation (1′), the amplitudes of its asymptotic solutions, U and V, are very similar to those of spherical Bessel and Neuman function, the differences appearing only as the term $(kf)^2/[3!\,(kr)^3]$ + terms of $O(1/(kr)^n)$, $n = 4, 5$.

The importance of the distance d as it relates to this problem will be discussed in Sec. III below.

III. PARAMETRIC DEPENDENCE OF THE SCATTERING PHASE SHIFTS UPON THE ENERGY OF THE INCIDENT ELECTRON AND THE MAGNITUDE OF THE POLARIZABILITY OF THE TARGET ATOM

One may naively compare a set of experimentally determined phase shifts $\delta_l(E)$ with the appropriate polarization phase shifts $\rho_l(k, f)$, as we have done in Fig. 3, for the low energy range of the e^--He scattering data analyzed by Bransden and McDowell.[5] This comparison illustrates that ρ_l dominates the total scattering phase shift δ_l only for large values of angular momentum. In the case of He, we see that the s-wave experimental phase shift δ_0 seems to have a magnitude and variation with energy which is altogether different from that of ρ_0. For heavier rare gas target atoms Ne, Ar, Kr, and Xe,[22] the p-wave experimental phase shifts δ_1, also strongly deviate from the behavior of ρ_1.

These observations may be understood when it is noted that low angular momentum radial wavefunctions have relatively high amplitude in the core region. Consequently, it is reasonable to expect that low angular momentum eigenfunctions are dominated by the core potential, while higher angular momentum eigenfunctions are dominated by the long range potential. On the other hand, such "dominance" of the core potential in determining the low angular momentum scattering phase shift is different in the presence of the long range $-f^2/r^4$ potential tail in contrast to the case of the finite range core potential alone. That is, it is known from numerical computations[9] or asymptotic expansions[8,23] that the low angular momentum scattering phase shifts for finite range potentials (those which vanish at d at least exponentially) exhibit magnitudes and energy dependences which are qualitatively different from the corresponding phase shifts for long range potentials. For example, Thompson[24] calculated the low energy s-wave phase shift for e^- scattering on Ar. He found that the phase shift in the limit $k \to 0$ changed magnitude and sign when he used an Ar core potential with or without the $-f^2/r^4$ charge-induced dipole interaction. Roughly interpreted, this behavior occurs because a particle scattered from a potential which is zero for all $r \geq d$ must have a spherical waveform for $r > d$. However, the wavefunction of a particle scattered from a potential which has the asymptotic form $-f^2/r^4$ may achieve a spherical waveform only at distances $r > r_t = \sqrt{f/k}$ which, at low energy, can be significantly larger than d (see Table II). We may suppose that the behavior of the experimental e^--rare gas atom scattering phase shifts are also grossly determined by the relationship of the distances r_t and d (where $d \approx$ the atomic radius).

These statements may be made more quantitative by

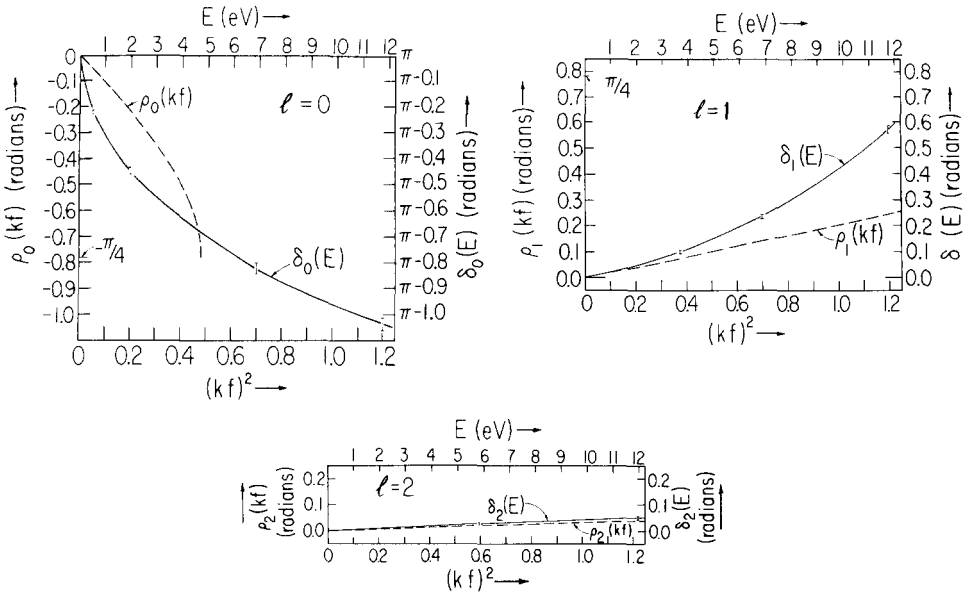

FIG. 3. Polarization phase shifts $\rho_l(kf)$ compared with experimentally inferred phase shifts for e^--He scattering, $\delta_l(E)$, given in Ref. 5.

examining the following model. The "core" of the atomic target may be roughly described as a deep potential well. In the Hartree approximation, the depth of the well goes as $-2Z/r$ Ryd as $r \to 0$. The "width" of the well is relatively small, as evidenced by the atomic radii listed in Table III. Elastic scattering at low energies from such a deep and narrow static potential may be described in terms of energy independent parameters.[8]

TABLE III. Parameters of rare gas atomic potentials.

	$2Z$	Atomic radius (Bohr)[a]
He	4	2.7
Ne	20	3.0
Ar	36	3.7
Kr	72	3.8
Xe	108	4.2

[a] See Ref. 25.

The logarithmic derivative, $R_l(d)$, of the wavefunction inside the deep potential, evaluated at atomic radius d, is virtually independent of energy for $k < k_{\max}$, where $k_{\max}$ in these cases is ≈ 0.1 Bohr^{-1}. For a given well depth, it is also only weakly dependent upon the exact form of the potential. In the following we assume that we can calculate $R_0(d)$ appropriate to a Hartree core potential. In the case of a nonpolarizable atom (such that $f = 0$, and the potential is zero for $r > d$) the s-wavefunction for $r > d$ would be

$$\psi_{l=0}(r) \underset{r>d}{=} \sqrt{2k/\pi}\,[\cos\delta_0 \sin(kr)/kr + \sin\delta_0 \cos(kr)/kr]. \tag{36}$$

In order to evaluate δ_0 for this model we match the logarithmic derivative $R_0(d)$ to that of Eq. (36) evaluated at $r = d$. If $k < k_{\max}$ and $kd \ll 1$, then $(\tan\delta_0)/k$ is independent of energy:

$$\lim_{k\to 0}[\tan(\delta_0)k] = -d^2R_0(d)/[dR_0(d) + 1]. \tag{37}$$

For the $f \neq 0$ case we must modify the above argument. $R_0(d)$ must be corrected to some approximation for the polarizability interaction in the core region. Equation (2) must be used instead of Eq. (36). For simplicity in the following analysis, we assume $kf \ll 0.695$, the "critical value" for $l = 0$. From Table II it is evident that, particularly for scattering from atomic targets of high atomic number, this requirement restricts the analysis to very low electron energies.

For such small values of the parameters we may approximate the functions $\Phi_l^{1,2}(r;k,f)$ analytically. We may use relation (30) to approximate:

$$\tau \approx l + \tfrac{1}{2} - [2(kf)^2/(2l+3)(2l+1)(2l-1)] + O((kf)^4). \tag{38}$$

From Table II and the list of atomic radii in Table III we see that $d < r_t$ for low electron energies, and consequently we must evaluate the functions in the regime $r < r_t$. From Eq. (18), the leading term for the amplitude K^-/K^+ is given by[8]

$$\frac{K^-}{K^+} \underset{kf\ll 1}{\approx} \left(\frac{\sqrt{kf}}{2}\right)^{2l+1} \frac{\Gamma(-l-\tfrac{1}{2}+1)}{\Gamma(l+\tfrac{1}{2}+1)}\,[1 + O((kf)^2 \ln kf) + O((kf)^2)] \tag{39}$$

since terms of $O(kf)$ cancel each other. If we assume $kf \lesssim 0.01$, the expressions will have approximately 1% accuracy. In this limit Eq. (21'a) and (22'a) become

$$\Phi_l^1(r;k,f) \underset{\substack{r<r_t\\ kf\ll 1}}{\approx} \sqrt{\pi/2kr}\,(K^-/K^+)J_{-(l+1/2)}(f/r), \tag{40}$$

$$\Phi_l^2(r;k,f) \underset{\substack{r<r_t\\ kf\ll 1}}{\approx} -\sqrt{\pi/2kr}\,(K^+/K^-)J_{(l+1/2)}(f/r); \tag{41}$$

for the $l = 0$ case these wavefunctions become

$$\Phi_{l=0}(r;k,f) \underset{\substack{r<r_t\\ kf\ll 1}}{\approx} \cos(f/r), \tag{42}$$

$$\Phi_{l=0}^2(r;k,f) \underset{\substack{r<r_t\\ kf\ll 1}}{\approx} -(1/kf)\sin(f/r). \tag{43}$$

Thus, in place of Eq. (36), the total s wavefunction becomes

$$\psi_{l=0}(r) \underset{\substack{d\le r<r_t\\ kf\ll 1}}{\approx} \sqrt{2k/\pi}\,\{\cos\gamma_0 \cos(f/r) + [\sin(\gamma_0)/kf]\sin(f/r)\}. \tag{44}$$

Using Eq. (5), we can calculate the "polarization extracted phase shift":

$$\lim_{k\to 0}\left(\frac{\tan\gamma_0}{k}\right) = \left(\frac{f\tan(f/d) - d^2R_0(d)}{d^2R_0(d)\tan(f/d) + f}\right)f. \tag{45}$$

Thus we see that if, for example, $0 < R_0(d) < [f\tan(f/d)]/d^2$, we can qualitatively understand Thompson's sign reversal result. In a sample rough calculation we have estimated that the Ar Hartree potential corrected for core polarizability has a zero energy s wave phase shift of roughly $\lim_{k\to 0}[(\tan\delta_0)/k] \cong -1.1$ Bohr. Using the above listed value of the atomic radius of Ar in relation (37), the s-wavefunction logarithmic derivative is approximately $R_0(d) \cong 0.11$ Bohr^{-1}. Consequently, we infer, for this case, $\lim_{k\to 0}[(\tan\gamma_0)/k] \cong 1.65$ Bohr. Since the appropriate "polarization phase shift" ρ_0 is negligible at the low energies considered, the above calculated value of $\lim_{k\to 0}[(\tan\gamma_0)/k]$ should correspond to the negative of the e^--Ar scattering length. A value of the e^--Ar scattering length of $a \cong -1.65$ Bohr is in agreement with the value $a \approx -1.7$ Bohr obtained from more careful estimates.[26] Note, however, that such good numerical agreement should not be taken seriously, since 10% errors in d and $\lim_{k\to 0}[(\tan\delta_0)/k]$ may correspond to roughly a 100% error in $\lim_{k\to 0}[(\tan\gamma_0)/k]$. However, our purpose here has only been to illustrate that the presence of the charge induced dipole interaction strongly modifies the wavefunction in the region outside the charge cloud of the atom, especially at low energies, from the usual spherical wave form, and that the Mathieu function representative may qualitatively predict this result and its consequences. The most interesting of these consequences, as noted previously,[8] is that the low energy e^--polarizable atom cross sections are much more sensitive functions of energy than finite range potential scattering cross sections. For example, in case of the Ar modified Hartree potential, the calculated total cross sections are energy independent (within 2%) up to electron energies of ≈ 0.5eV. However, when the long range induced dipole interaction is included the energy independent form (45) is only attained for $kf \lesssim 0.01$, which means for the case of Ar, incident electron energies less than $\approx 1.2 \times 10^{-4}$ eV.

This analysis can be carried further; by keeping terms of higher power in k in Eqs. (39)–(44) and by use of Eq. (5) we can straightforwardly obtain the "modified effective range" expansion given in Eq. (2.3a) of Ref. 26, retaining terms up to $O(k^3)$.

The above analysis can, in principle, be inverted to determine from a set of experimental e^--gaseous atom phase shifts, $\{\delta_l\}$, an analogous set of phase shifts appropriate to electron scattering from the same atomic core potential chopped off at, say, a distance r_s. In studying some properties of dielectric solids r_s could be the Wigner Seitz radius. For accuracy, one should use Eq. (7) to determine the logarithmic derivatives of the wavefunction at r_s rather than the approximate analytic forms given in this section. However, having demonstrated the sensitivity of the low energy scattering phase shifts to the long range $-f^2/r^4$ potential, we suggest that one be wary of unquestioned use of the muffin tin approximation, which sets all such long range parts of the interaction equal to zero. From the above considerations, the muffin tin approximation may be especially inaccurate for describing electron scattering in disordered systems.

It is important to emphasize the major limitation of the Mathieu function solutions relative to their analogous free particle solutions. All atomic wavefunctions are required to approach the origin as[27]

$$\psi_{\text{atomic}}(r) \underset{r\to 0}{\sim} C\{r^l - [Z/(l+1)]r^{l+1} + O(r^{l+2})\}. \quad (46)$$

This boundary condition is compatible with spherical Bessel function solutions. That is, for example, if one uses $\psi_{\text{atomic}}(r) \approx Cj_l(kr)$ as the initial step in iteration in an atomic scattering problem using the free particle Green's function (i.e., the Born approximation) satisfactory results are usually achieved. However, a Mathieu function Green's function has virtually no value for evaluating the wavefunction for $r < d$, since as $r \to 0$ the Mathieu function eventually oscillates sinusoidally with argument f/r, having no resemblance to (46). On the other hand, we recommend that numerical integration of the Schrödinger equation for a $-f^2/r^4$ potential be avoided and that, instead, one makes use of linear combinations of Mathieu functions to represent the desired function for the appropriate region of space.

IV. COMPARISON WITH PREVIOUS WORK

Since this work is closely related to that of O'Malley, Spruch, and Rosenberg,[8] Berger, Snodgrass, and Spruch,[9] and Hinckelmann and Spruch,[10] we feel it appropriate to compare the different approaches.

(a) "Modification of effective-range theory in the presence of a long range potential.[8]" The correspondence of our notation and that of O'Malley, Spruch and Rosenberg[8] is as follows:

Notation of O'Malley, *et al.*,	Our notation
β	f
m	K^-/K^+ (for nonintegral values of τ)
δ	$-\rho_l(k,f)$
$\eta(L)$	δ_l
A	$\left(\dfrac{d^2R_0(d) - f\tan(f/d)}{d^2R_0(d)\tan(f/d) + f}\right)f$ $= -\lim_{k\to 0}\left(\dfrac{\tan\gamma_0}{k}\right)$

These authors have solved a problem very similar to the one with which we are concerned here; however, they have imposed different boundary conditions. Roughly put, their boundary conditions are the reverse of ours, in the sense that they require simple sinusoidal behavior of the functions in the limit $r \to 0$, for the argument f/r, while we require simple sinusoidal behavior of the functions in the limit $r \to \infty$, for the argument kr. Their analysis of the low energy scattering phase shifts agrees with that given in Sec. III. We claim that such analytic expressions have fair accuracy for the range of parameters $kf \ll kf\rfloor_{\text{crit.}}$ and $d \ll \sqrt{f/k}$, as well as $kf \ll 1$ as stated by the authors.

(b) "Tables of coefficients to determine the long-range contributions to low-energy electron-atom scattering.[9]"

This paper deals with the numberical evaluation of the two linearly independent solutions $f(k,r)/kr$ and $g(k,r)/kr$ of the "cutoff" polarizability potential

$$\left(\frac{1}{r^2}\frac{d}{dr}r^2\frac{d}{dr} - \frac{l(l+1)}{r^2} + \frac{f^2}{r^4}\theta(r-d) + k^2\right)\psi(r) = 0, \quad \text{(Eq. (2.6))}$$

where $\theta(r-d)$ is the Heaviside unit step function. The boundary conditions imposed are the following:

$$f(k,r)/kr = \tilde{C}(k) j_l(kr), \quad r < d$$

$$f(k,r)/kr \underset{r\to\infty}{\sim} \sin(kr - l\pi/2 + \tilde{\rho})/kr$$

$$g(k,r)/kr = -Y_l(kr)/\tilde{C}(k) + \tilde{E}(k) j_l(kr), \quad r < d$$

$$g(k,r)/kr \underset{r\to\infty}{\sim} \cos(kr - l\pi/2 + \tilde{\rho})/kr.$$

(Eqs. (2.2) and (2.3))

We have denoted these authors' polarization phase shifts by $\tilde{\rho}$ to distinguish them from the $\rho_l(kf)$ discussed above.

We see that by matching $f(k,r)/kr$ and $g(k,r)/kr$ and their derivatives to linear combinations of Mathieu function solutions $\Phi_l^{1,2}(r;k,f)$ at $r=d$ we can determine parameters $\tilde{\rho}, \tilde{C}(k)$, and $\tilde{E}(k)$ without explicitly integrating the differential equation (Eq. (2.6)). This result may have some computational utility since the evaluation of series $\mathfrak{M}_{\pm\tau}(x)$ or $\mathfrak{J}_{\pm\tau}(r)$ is relatively straightforward.

We write

$$\frac{f(k,r)}{kr} \underset{r>d}{=} \cos\phi_l \Phi_l^1(r;k,f) - \sin\phi_l \Phi_l^2(r;k,f). \tag{47}$$

Evaluating (38) and its derivative at $r=d$, we can determine ϕ_l from a relation such as:

$$\tan\phi_l = \frac{\Phi_l^1(d;k,f)[(d/dr)\ln(j_l(kr))]_{r=d} - [(d/dr)\Phi_l^1(r;k,f)]_{r=d}}{\Phi_l^2(d;k,f)[(d/dr)\ln(j_l(kr))]_{r=d} - [(d/dr)\Phi_l^2(r;k,f)]_{r=d}}. \tag{48}$$

It follows that $\tilde{\rho}$ can be evaluated from

$$\tilde{\rho} = \rho_l(k,f) + \phi_l. \tag{49}$$

By inspection, if d is large enough that $\Phi_l^{1,2}(d;k,f)$ may be represented by their asymptotic forms $\{\sin[kr - l\pi/2 + \rho_l(k,f)]\}/kr$ or $-[\cos(kr - l\pi/2 + \rho_l)]/kr$, (48) becomes $\tan\phi_l = \tan(-\rho_l)$, so that $\tilde{\rho}=0$. $\tilde{C}(k)$ can be evaluated from

$$\tilde{C}(k) = [\cos(\phi_l)\Phi_l^1(d;k,f) - \sin(\phi_l)\Phi_l^2(d;k,f)]/j_l(kd). \tag{50}$$

By inspection, since

$$g(k,r)/kr \underset{r>d}{=} -[\cos(\phi_l)\Phi_l^2(r;k,f) - \sin\phi_l\Phi_l^1(r;k,f)], \tag{51}$$

$\tilde{E}(k)$ may be evaluated from

$$\tilde{E}(k) = [1/j_l(kd)]\{[y_l(kd)/\tilde{C}(k)] - [\cos(\phi_l)\Phi_l^2(d;k,f) - \sin(\phi_l)\Phi_l^1(d;k,f)]\}. \tag{52}$$

If d is such that the functions $\Phi_l^{1,2}(d;k,f)$ are well represented by their spherical wave asymptotic forms, then $\tilde{C}(k) \approx 1$ and $\tilde{E}(k) \approx 0$. These predictions are in agreement with the numerical values given by Berger, Snodgrass, and Spruch[9] for small values of f/d.

We also note that for $l=0$ these authors[9] show that $\tilde{\rho}$ jumps from an increasing function of kf to a decreasing function between the values of $d = f/1.4$ and $f/1.6$ and for $l=1$ between the values $d = f/2.8$ and $f/3.2$. Such behavior is predicted when the denominator of (48) passes through zero.

(c) "Low-energy scattering by long-range potentials.[10]" Included in this paper is a perturbation expansion for the low-energy scattering phase shifts of a $-f^2/r^4$ potential. The perturbation is carried to first order in the long range potential for all $r > d$. Since this model is very similar to ours, the two results should be compatible as long as the perturbing potential is small. In particular, their analysis requires that the zero-order wavefunction for $r > d$ have the form

$$\psi_l(r) \underset{r>d}{\approx} j_l(kr) - \tan(\eta_{ls}) y_l(kr). \tag{53}$$

We have shown above that such a condition may be satisfied for all values of angular momentum if

$$1 > (kf)^2 > f^4/d^4. \tag{54}$$

For example, in order for there to be an energy range in which the perturbation analysis holds, the potential parameters must satisfy $d > f$. Such a restriction is satisfied by the values of d listed in Table III for He and Ne, but not for Ar, Kr, and Xe.

ACKNOWLEDGMENTS

This research was supported by the Directorate of Chemical Sciences, Air Force Office of Scientific Research. We have also benefited from the use of facilities provided by the Advanced Research Projects Agency for materials research at the University of Chicago. The author would like to thank Dr. Stuart A. Rice for highly appreciated advice, encouragement, and support during this work. Valuable advice from Dr. Norman Lebovitz, Dr. Morrel Cohen, Dr. John Light, and Dr. A.R.P. Rau is also gratefully acknowledged.

APPENDIX A: DERIVATION OF VARIOUS REPRESENTATIONS OF MATHIEU FUNCTIONS

A solution to Eq. (9'a) may be generally represented by an infinite series:

$$\mathfrak{M}_\tau(x) \equiv \sum_{n=-\infty}^{\infty} C_n(\tau) x^{(\tau+2n)}, \tag{A1}$$

where τ is the "characteristic exponent" of the solutions and where the coefficients $C_n(\tau)$ satisfy the recurrence relation

$$[(\tau+2n)^2 - (l+\tfrac{1}{2})^2]C_n(\tau) + kf[C_{n+1}(\tau) + C_{n-1}(\tau)] = 0. \tag{A2}$$

In order for there to be a nontrivial solution for the constants $C_n(\tau)$, the determinant of their coefficients in Eq. (A2) must vanish. Equivalently we require the following infinite (but convergent) determinant to vanish:

$$\Delta^l(\tau) = \begin{vmatrix} \frac{[2(n-1+\tau]^2-(l+\frac{1}{2})^2}{4(n-1)^2-(l+\frac{1}{2})^2} & \frac{kf}{4(n-1)^2-(l+\frac{1}{2})^2} & 0 & \cdots \\ \frac{kf}{4n^2-(l+\frac{1}{2})^2} & \frac{(2n+\tau)^2-(l+\frac{1}{2})^2}{4n^2-(l+\frac{1}{2})^2} & \frac{kf}{4n^2-(l+\frac{1}{2})^2} & \cdots \\ 0 & \frac{kf}{4(n+1)^2-(l+\frac{1}{2})^2} & \frac{[2(n+1)+\tau]^2-(l+\frac{1}{2})^2}{4(n+1)^2-(l+\frac{1}{2})^2} & \cdots \\ 0 & 0 & \frac{kf}{4(n+2)^2-(l+\frac{1}{2})^2} & \cdots \\ \vdots & \vdots & \vdots & \end{vmatrix}. \tag{A3}$$

This is one of the many forms of the Hill determinant. The solutions of are given, for example, in Morse and Feshbach,[11] as

$$\sin^2\tfrac{1}{2}\pi\tau = [\sin^2\tfrac{1}{2}\pi(l+\tfrac{1}{2})]\Delta^l(\tau=0), \tag{A4}$$

which is equivalent to Eq. (11) in the text. The determinant $\Delta^l(\tau=0;kf)$ is a "continuant" whose Nth approximation has the symmetric form

$$\Delta_N \equiv \begin{vmatrix} 1 & a_N & & & & & & \\ a_{N-1} & 1 & a_{N-1} & & & & & \\ 0 & a_{N-2} & 1 & a_{N-2} & & & & \\ & & a_1 & 1 & a_1 & & & \\ & & & a_0 & 1 & a_0 & & \\ & & & & a_1 & 1 & a_1 & \\ & & & & & a_{N-1} & 1 & a_{N-1} \\ & & & & & & a_N & 1 \end{vmatrix}, \tag{A5}$$

where

$$a_n \equiv kf/[4n^2-(l+\tfrac{1}{2})^2].$$

By expanding (A5) in minors of its center row, Δ_N can be written

$$\Delta_N = (D_N^1)^2 - 2a_0a_1(D_N^2)(D_N^1), \tag{A6}$$

where the subdeterminants D_N^k are defined by

$$D_N^k = \begin{vmatrix} 1 & a_k & & & & \\ a_{k+1} & 1 & a_{k+1} & & & \\ & & \ddots & & & \\ & & & a_{N-1} & 1 & a_{N-1} \\ & & & & a_N & 1 \end{vmatrix}. \tag{A7}$$

These determinants may be augmented from the "top" or the "bottom" according to the recurrence relations

$$\begin{aligned} D_N^k &= D_N^{k+1} - a_k a_{k+1} D_N^{k+2}, \\ D_N^k &= D_{N-1}^k - a_N a_{N-1} D_{N-2}^k. \end{aligned} \tag{A8}$$

For our calculations we have used the starting relations

$$\begin{aligned} D_1^0 &= 1 - a_1 a_0, \\ D_2^0 &= D_1^0 - a_2 a_1, \\ D_1^1 &= 1, \\ D_2^1 &= D_1^1 - a_2 a_1. \end{aligned} \tag{A9}$$

and the algorithm

$$\begin{aligned} D_N^0 &= D_{N-1}^0 - a_N a_{N-1} D_{N-2}^0, \\ D_N^1 &= D_{N-1}^1 - a_N a_{N-1} D_{N-2}^1, \\ \Delta_N &= D_N^1(2D_N^0 - D_N^1) \end{aligned} \tag{A10}$$

to generate successive approximations to $\Delta^l(0;kf)$.

Once τ has been determined from Eq. (A4), the coefficients $C_n(\tau)$ may be determined (up to a normalization constant) from the convergent continued fractions:

$$\begin{aligned} C_n(\tau)/C_{n-1}(\tau) &= -kf/[(\tau+2n)^2 - (l+\tfrac{1}{2})^2 + kfC_{n+1}(\tau)/C_n(\tau)], \\ C_n(\tau)/C_{n+1}(\tau) &= -kf/[(\tau+2n)^2 - (l+\tfrac{1}{2})^2 + kfC_{n-1}(\tau)/C_n(\tau)]. \end{aligned} \tag{A11}$$

These continued fractions can be evaluated by iterating from large values of n. By using the starting values

$$\begin{aligned} C_N/C_{N-1} &= -kf/[(2N+\tau)^2 - (l+\tfrac{1}{2})^2], \\ C_{-N}/C_{-(N-1)} &= -kf/[(-2N+\tau)^2 - (l+\tfrac{1}{2})^2]. \end{aligned} \tag{A12}$$

and the algorithms

$$\begin{aligned} C_n/C_{n-1} &= -kf/[(2n+\tau)^2 - (l+\tfrac{1}{2})^2 + kfC_{n+1}/C_n], \\ C_{-n}/C_{-(n-1)} &= -kf/[(-2n+\tau)^2 - (l+\tfrac{1}{2})^2 + kfC_{-(n+1)}/C_{-n}], \end{aligned} \tag{A13}$$

the desired ratios of coefficients could be calculated up through C_1/C_0 and C_{-1}/C_0.

It is apparent that, for nonintegral values of τ, two independent solutions of (9'a) are thus given by

$$\mathfrak{M}_{\pm\tau}(x) = \sum_{n=-\infty}^{\infty} C_n(\pm\tau)x^{(\pm\tau+2n)}. \tag{A14}$$

We are allowed the further simplification $C_n(-\tau) = C_{-n}(+\tau)$, so that the solutions become Eq. (10).

APPENDIX B: PROPERTIES OF MATHIEU FUNCTIONS HAVING INTEGER OR COMPLEX CHARACTERISTIC EXPONENTS

The nature of the solutions $\Phi_l^{1,2}(r;k,f)$ for integral values of the characteristic exponent is primarily determined by some consequent interrelationships between the coefficients $C_n(\tau=L)$

$$C_n/C_{n-1} = C_{-(n+L)}/C_{-(n+L-1)},$$
$$C_n/C_{n+1} = C_{-(n+L)}/C_{-(n+L+1)}. \tag{B1}$$

These relationships ensure that, for example, terms in series (10) $C_n x^{(2n+L)}$ and $C_{-(n+L)}x^{-(2n+L)}$ will be paired symmetrically or antisymmetrically. For illustration, consider $L > 0$. For odd $\tau = L^o$, relations (B1) determine that

$$(C_{-[(L^o+1)/2]}/C_{-[(L^o-1)/2]})^2 = 1 \tag{B2}$$

and consequently that

$$C_{-L^o}/C_o = \pm 1, \quad C_{-(n+L^o)}/C_o = \pm C_n/C_o, \tag{B3}$$

leading to $\{^{\text{even}}_{\text{odd}}\}$ Mathieu functions of odd integral order.

For even $\tau = L^e$, relations (B1) determine that

$$C_{-L^e}/C_o = 1. \tag{B4}$$

Consequently, one can form even Mathieu functions of even integral order by setting $C_o = 1$. Odd solutions can be formed by setting $C_o = -1$ if the coefficient of the neutral term, $C_{-(L^e/2)} = 0$. (Similar arguments hold for $L < 0$).

In our problem for angular momenta $l = 0$ or $l = 1$, $\tau \to L^o$, and thus relations (B2), (B3) are applicable. From consideration of the general equality

$$C_1/C_0 + C_{-1}/C_0 = [(l + \tfrac{1}{2})^2 - \tau^2]/kf, \tag{B5}$$

we conclude that, for $l = 0$, at the critical value of kf, $C_{-1}/C_0 = -1$ and consequently $C_n(\tau_0 = 1) = -C_{-n-1}(\tau_0 = 1)$. Thus for this case $\Phi^1_{l=0}(r;k,f)$ takes the familiar form which is an "odd" function with respect to $(\sqrt{k/f}\,r)$ and $(\sqrt{k/f}\,r)^{-1}$:

$$\Phi^1_{l=0}(r;k,f) = \frac{1}{K^+}\sqrt{\frac{\pi}{2kr}}\sum_{n=0}^{\infty} C_n(\tau_0 = 1) \times \left[\left(\sqrt{\frac{k}{f}}\,r\right)^{2n+1} - \left(\sqrt{\frac{k}{f}}\,r\right)^{-(2n+1)}\right]$$
$$= \sqrt{\frac{\pi}{2kr}}\sum_{n=0}^{\infty}(-1)^n C_n(\tau_0 = 1)\left[J_{n+1}(kr)J_n\left(\frac{f}{r}\right) - J_n(kr)J_{n+1}\left(\frac{f}{r}\right)\right] = \frac{2}{\mathfrak{N}^o}\sqrt{\frac{\pi}{2kr}}\left(\frac{kr - f/r}{kr + f/r}\right) \times \sum_{n=0}^{\infty}(-1)^n(2n+1)C_n(\tau_0 = 1)J_{2n+1}\left(kr + \frac{f}{r}\right), \quad \text{for } r \neq \sqrt{\frac{f}{k}}. \tag{B6}$$

The second solution may be written

$$\Phi^2_{l=0}(r;k,f) = \begin{cases} \sqrt{\frac{\pi}{2kr}}\sum_{n=0}^{\infty}(-1)^n C_n(\tau_0 = 1)\left[Y_{n+1}(kr)J_n\left(\frac{f}{r}\right) - Y_n(kr)J_{n+1}\left(\frac{f}{r}\right)\right], & r \geq \sqrt{\frac{f}{k}}, \\ A\Phi^1_{l=0}(r;k,f) + \sqrt{\frac{\pi}{2kr}}\sum_{n=0}^{\infty}(-1)^n C_n(\tau_0 = 1) \times \left[Y_{n+1}\left(\frac{f}{r}\right)J_n(kr) - Y_n\left(\frac{f}{r}\right)J_{n+1}(kr)\right], & r \leq \sqrt{\frac{f}{k}}, \end{cases}$$
$$= \begin{cases} \frac{2}{\mathfrak{N}^o}\sqrt{\frac{\pi}{2kr}}\left(\frac{kr - f/r}{kr + f/r}\right)\sum_{n=0}^{\infty}(-1)^n C_n(\tau_0 = 1) \times (2n+1)Y_{2n+1}\left(kr + \frac{f}{r}\right), & r > \sqrt{\frac{f}{k}}, \\ A\Phi^1_{l=0}(r;k,f) + \frac{2}{\mathfrak{N}^o}\sqrt{\frac{\pi}{2kr}}\left(\frac{f/r - kr}{f/r + kr}\right) \times \sum_{n=0}^{\infty}(-1)^n C_n(\tau_0 = 1)(2n+1) \times Y_{2n+1}\left(kr + \frac{f}{r}\right), & r < \sqrt{\frac{f}{k}}, \end{cases} \tag{B7}$$

where $A = 0.744$.

Similarly, for the $l = 1$ case, at its critical value of kf, $C_{-1}/C_0 = +1$, and consequently $C_n(\tau_1 = 1) = C_{-n-1}(\tau_1 = 1)$. $\Phi^1_{l=1}(r;k,f)$ then becomes the "even" function

$$\Phi^1_{l=1}(r;k,f) = \frac{1}{K^+}\sqrt{\frac{\pi}{2kr}}\sum_{n=0}^{\infty} C_n(\tau_1 = 1) \times \left[\left(\sqrt{\frac{k}{f}}\,r\right)^{2n+1} + \left(\sqrt{\frac{k}{f}}\,r\right)^{-(2n+1)}\right]$$
$$= \sqrt{\frac{\pi}{2kr}}\sum_{n=0}^{\infty}(-1)^n C_n(\tau_1 = 1)\left[J_{n+1}(kr)J_n\left(\frac{f}{r}\right) + J_n(kr)J_{n+1}\left(\frac{f}{r}\right)\right] = \frac{2}{\mathfrak{N}^e}\sqrt{\frac{\pi}{2kr}}\sum_{n=0}^{\infty}(-1)^n C_n(\tau_1 = 1) \times J_{2n+1}\left(kr + \frac{f}{r}\right), \quad \text{for } r \neq \sqrt{\frac{f}{k}}. \tag{B8}$$

The second solution may be written

$$\Phi^2_{l=1}(r;k,f) = \begin{cases} \sqrt{\frac{\pi}{2kr}}\sum_{n=0}^{\infty}(-1)^n C_n(\tau_1 = 1) \times \left[Y_{n+1}(kr)J_n\left(\frac{f}{r}\right) + Y_n(kr)J_{n+1}\left(\frac{f}{r}\right)\right], & r \geq \sqrt{\frac{f}{k}}, \\ -A\Phi^1_{l=1}(r;k,f) - \sqrt{\frac{\pi}{2kr}}\sum_{n=0}^{\infty}(-1)^n C_n(\tau_1 = 1) \times \left[Y_{n+1}\left(\frac{f}{r}\right)J_n(kr) + Y_n\left(\frac{f}{r}\right)J_{n+1}\left(\frac{f}{r}\right)\right], & r \leq \sqrt{\frac{f}{k}}, \end{cases}$$
$$= \begin{cases} \frac{2}{\mathfrak{N}^e}\sqrt{\frac{\pi}{2kr}}\sum_{n=0}^{\infty}(-1)^n C_n(\tau_1 = 1) \times Y_{2n+1}\left(kr + \frac{f}{r}\right), & r > \sqrt{\frac{f}{k}}, \\ -A\Phi^1_{l=1}(r;k,f) - \frac{2}{\mathfrak{N}^e}\sqrt{\frac{\pi}{2kr}} \times \sum_{n=0}^{\infty}(-1)^n C_n(\tau_1 = 1)Y_{2n+1}\left(kr + \frac{f}{r}\right), & r < \sqrt{\frac{f}{k}}, \end{cases} \tag{B9}$$

where $A = 1.102$.

For the case $l = 2$, at the critical value of kf, τ approaches the even integer $L^e = 2$. $\Phi^1_{l=2}(r;k,f)$ then becomes the completely even function

$$\Phi^1_{l=2}(r;k,f) = \frac{1}{K^+}\sqrt{\frac{\pi}{2kr}}\left\{C_{-1}(\tau_2=2) + \sum_{n=0}^{\infty} C_n(\tau_2=2) \times \left[\left(\sqrt{\frac{k}{f}}\,r\right)^{2n+2} + \left(\sqrt{\frac{k}{f}}\,r\right)^{-(2n+2)}\right]\right\}$$

$$= \sqrt{\frac{\pi}{2kr}}\left\{C_{-1}(\tau_2=2)J_1(kr)J_1\left(\frac{f}{r}\right) + \sum_{n=0}^{\infty}(-1)^n \times C_n(\tau_2=2)\left[J_{n+2}(kr)J_n\left(\frac{f}{r}\right) + J_n(kr)J_{n+2}\left(\frac{f}{r}\right)\right]\right\}$$

$$= \frac{1}{\mathfrak{N}^e}\sqrt{\frac{\pi}{2kr}}\left\{-C_{-1}(\tau_2=2)J_0\left(kr+\frac{f}{r}\right) + 2\sum_{n=0}^{\infty}(-1)^n C_n(\tau_2=2)J_{2n+2}\left(kr+\frac{f}{r}\right)\right\},$$

$$\text{for } r \neq \sqrt{\frac{f}{k}}, \qquad \text{(B10)}$$

with the second solution

$$\Phi^2_{l=2}(r;k,f) = \begin{cases} \sqrt{\frac{\pi}{2kr}}\left\{C_{-1}(\tau_2=2)Y_1(kr)J_1\left(\frac{f}{r}\right) + \sum_{n=0}^{\infty}(-1)^n C_n(\tau_2=2)\left[Y_{n+2}(kr)J_n\left(\frac{f}{r}\right) + Y_n(kr)J_{n+2}\left(\frac{f}{r}\right)\right]\right\}, & r \geq \sqrt{\frac{f}{k}}, \\ -A\Phi^1_{l=2}(r;k,f) - \sqrt{\frac{\pi}{2kr}} \times \left\{C_{-1}(\tau_2=2)Y_1\left(\frac{f}{r}\right)J_1(kr) + \sum_{n=0}^{\infty}(-1)^n C_n(\tau_2=2) \times \left[Y_{n+2}\left(\frac{f}{r}\right)J_n(kr) + Y_n\left(\frac{f}{r}\right)J_{n+2}(kr)\right]\right\}, & r \leq \sqrt{\frac{f}{k}}, \end{cases}$$

$$= \begin{cases} \frac{1}{\mathfrak{N}^e}\sqrt{\frac{\pi}{2kr}}\left\{-C_{-1}(\tau_2=2)Y_0\left(kr+\frac{f}{r}\right) + 2\sum_{n=0}^{\infty}(-1)^n C_n(\tau_2=2) \times Y_{2n+2}\left(kr+\frac{f}{r}\right)\right\}, & r > \sqrt{\frac{f}{k}}, \\ -A\Phi^1_{l=2}(r;k,f) - \frac{1}{\mathfrak{N}^e}\sqrt{\frac{\pi}{2kr}} \times \left[-C_{-1}(\tau_2=2)Y_0\left(kr+\frac{f}{r}\right) + 2\sum_{n=0}^{\infty}(-1)^n C_n(\tau_2=2) \times Y_{2n+2}\left(kr+\frac{f}{r}\right)\right], & r < \sqrt{\frac{f}{k}}, \end{cases} \qquad \text{(B11)}$$

where $A = 1.900$.

In the case of a complex characteristic exponent $\tau = L + i\mu$ the expansion coefficients $C_n(L + i\mu)$ are again subject to certain interrelationships as a consequence of the fact that $Re\{\tau\} =$ integer. For this case relations (B1) become the more general equalities

$$C_n(\tau)/C_{n-1}(\tau) = (C_{-(n+L)}(\tau)/C_{-(n+L-1)}(\tau))^*,$$

$$C_n(\tau)/C_{n+1}(\tau) = (C_{-(n+L)}(\tau)/C_{-(n+L+1)}(\tau))^* \quad \text{(complex conjugate)} \qquad \text{(B12)}$$

Using the notation $C_{-n}(\tau)/C_{-n+1}(\tau) \equiv |C_{-n}/C_{-n+1}|\,e^{i\upsilon_{-n}}$, relations (B3) and (B4) become

$$\frac{C_{-L^o}}{C_0} = \exp\left[i\left(\theta_{-[(L^o+1)/2]} + 2\sum_{k=1}^{(L^o-1)/2}\theta_{-k}\right)\right] \equiv e^{i\sigma^o},$$

$$\frac{C_{-n+L^o}}{C_0} = \exp(i\theta_{-[(L^o+1)/2]})\left(\frac{C_n}{C_0}\right)^*, \qquad \text{(B13)}$$

$$\frac{C_{-L^e}}{C_0} = \exp\left[2i\left(\sum_{k=1}^{L^e/2}\theta_{-k}\right)\right] \equiv e^{i\sigma^e}. \qquad \text{(B14)}$$

It is amusing to note that we can write (10) in the following form for these cases:

$$\mathfrak{M}_{(L+i\mu)}(x) = e^{i\sigma/2}x^{i\mu}\left[\sum_{n=0}^{\infty} a_n(x^{2n+L} + x^{-(2n+L)}) + i\sum_{n=0}^{\infty} b_n(x^{2n+L} - x^{-(2n+L)})\right]$$

Since $x^{i\mu} \equiv \cos(\mu \ln x) + i\sin(\mu \ln x)$, this means that

$\mathrm{Re}[e^{-i\sigma/2}\mathfrak{M}_{(L+i\mu)}(x)]$ is even with respect to x and $1/x$ while

$\mathrm{Im}[e^{-i\sigma/2}\mathfrak{M}_{(L+i\mu)}(x)]$ is odd with respect to x and $1/x$.

---

[1] C. J. Kleinman, Y. Hahn, and L. Spruch, Phys. Rev. **165**, 53 (1968).

[2] Joseph Callaway, R. W. LaBahn, R. T. Pu, and W. M. D. Duxler, Phys. Rev. **168**, 12 (1968).

[3] The wavefunctions are "energy" normalized according to the definition given by H. A. Bethe and E. E. Salpeter, *Quantum mechanics of one- and two-electron atoms* (Springer-Verlag, Berlin, 1957), p. 22, $\int_0^\infty r^2\,dr\,\psi_l(r;k,f)\int_{k-\Delta k}^{k+\Delta k} k'^2\,dk'\,\psi_l(r;k',f) = 1$.

[4] For example, experimental cross sections may be fitted to the following relations: for the total elastic cross section, $k^2\sigma_t/4\pi = \Sigma_{l=0}^{\infty}(2l+1)\sin^2\delta_l$ (Bohr^2); for the differential cross section, $[k^2(d\sigma(\theta)/d\Omega) = |\Sigma_{l=0}^{\infty}(2l+1)e^{i\delta_l}\sin\delta_l P_l(\cos\theta)|^2 = \Sigma_{l,l'=0}^{\infty}{}_{l\neq l'}\{(2l+1)(2l'+1)(\sin^2\delta_l\sin^2\delta_{l'} + \tfrac{1}{4}\sin 2\delta_l\sin 2\delta_{l'})P_l(\cos\theta)P_{l'}(\cos\theta)\}$ (Bohr^2/Steradian); or for the momentum transfer cross section, $k^2\sigma_p/4\pi \equiv \tfrac{1}{2}\int_0^\pi d\sigma(\theta)/d\Omega\,(1-\cos\theta)d\theta = \Sigma_{l=0}^{\infty}(l+1)\sin^2(\delta_l - \delta_{l+1})$ (Bohr^2).

[5] B. H. Bransden and M. R. C. McDowell, J. Phys. B 2, 1187 (1969), and M. R. C. McDowell, *Atomic physics II* (Plenum, New York, 1970) (in press).

[6] J. R. Gibson and K. T. Dolder, J. Phys. B **2**, 1180 (1969).

[7] C. Ramsauer and R. Kollath, Ann. Phys. (Leipz.) **9**, 756 (1931); Ann. Phys. (Leipz.) 12, 529 (1932), Ann. Phys. (Leipz.) **12**, 836 (1932).

[8] T. F. O'Malley, L. Spruch, and L. Rosenberg, J. Math. Phys. **2**, 491 (1961).

[9] R. O. Berger, H. B. Snodgrass, and L. Spruch, Phys. Rev. **185**, 113 (1969).

[10] Otto Hinckelmann and Larry Spruch, Phys. Rev. A **3**, 642 (1971).

[11]P. M. Morse and H. Feshbach, *Methods of theoretical physics* (McGraw-Hill, New York, 1953), Vol. 1, pp. 556ff., Vol. II, pp. 1407ff., pp. 1568ff.

[12]N. W. McLachlan, *Theory and application of Mathieu functions* (Dover, New York, 1964).

[13]J. Meixner and F. W. Schäfke, *Mathieusche Funktionen und Spharoidfunktionen* (Springer-Verlag, Berlin, 1954).

[14]There are some misprints in the paper by N. W. McLachlan, J. Math. Phys. **26**, 29 (1947). For example, using his notation, Eq. (10) in Sec. 3 of his paper should read, for $0 \leqslant z < \infty, \beta \neq 0$, $K \Sigma_{r=-\infty}^{\infty}(-1)^r A_{2r} J_{2r+\beta}(2k \cosh z) = Ce_{2n+\beta}(z, q) + Se_{2n+\beta}(z, q) = 2\Sigma_{r=-\infty}^{\infty} A_{2r} e^{(2r+\beta)z}$. Equation (6) in Sec. 4 should read, for $0 \leqslant z < \infty, \beta \neq 0$, $K' \tanh z \, \Sigma_{r=-\infty}^{\infty}(-1)^r (2r+\beta) A_{2r} J_{2r+\beta}(2k \cosh z) = Ce_{2n+\beta}(z, q) + Se_{2n+\beta}(z, q) = 2\Sigma_{r=-\infty}^{\infty} A_{2r} e^{(2r+\beta)z}$, where K and K' are constants. In general, one should exercise caution in using any of the connecting relations in McLachlan's paper. The correct connecting relations are given in the later work of Meixner and Schäfke.[13] We stress this point here in order to prevent further confusion.

[15]The absolute value of $(kr - f/r)$ is written for convenience in notation. Even without this absolute value sign, the functions $\mathcal{J}^0_{t\tau}(r)$ are not rigorously odd with respect to kr and f/r, since $\lim_{kr \to f/r} (\mathcal{J}^0_{t\tau}(r)) \neq 0$ except for some special values of the parameters.

[16]G. N. Watson, *A treatiseon the theory of Bessel functions* (Cambridge U.P., Cambridge, 1958), Chap. 7, p. 199.

[17]E. L. Ince, *Ordinary differential equations* (Dover, New York, 1956), p. 175, Sec. 7.4.

[18]National Bureau of Standards, *Tables relating to Mathieu functions* (U. S. Government Printing Office, Washington, D. C., 1967). For convenience, we list

their notation	our notation
s	$4kf$
b	$(l + \frac{1}{2})^2 + 2kf$
r	L

[19]M. J. O. Strutt, Ergeb. Math. 1, 1932, Sec. III. 1. b., Eq. (3).

[20]H. Dalgarno, Adv. Phys. **11**, 281 (1962).

[21]Possibly smaller values of D and $\mathcal{E}$ may be achieved in a pressure broadening experiment. Here the presence of foreign gases causes a frequency shift of the absorption lines corresponding to transitions to highly excited states of alkali atoms (for example). The magnitude of the frequency shift is related to the elastic cross section of a very low energy electron with the foreign gas atom. However, the interaction in this case is perturbed by the Coulomb field of the far off alkali positive ion. In addition, cross sections measured in such experiments have much lower precision than direct cross section measurements.

[22]P. S. Hoeper, W. Franzen, and R. Gupta, Phys. Rev. **168**, 50 (1968); our unpublished work.

[23]B. R. Levy and J. B. Keller, J. Math. Phys. **4**, 54 (1963).

[24]D. G. Thompson, Proc. R. Soc. A **294**, 160 (1966).

[25]For an estimate of an atomic radius we have used one-half the distance of the Lennard-Jones minimum of interatomic energy. J. O. Hirschfelder, C. F. Curtiss, and R. B. Bird, *Molecular theory of gases and liquids* (Wiley, New York, 1954), Table 1-A, p. 1110; G. L. Pollack, Rev. Mod. Phys. **36**, 748 (1964).

[26]T. F. O'Malley, Phys. Rev. **130**, 1020 (1963).

[27]D. R. Hartree, *The calculation of atomic structures* (Wiley, New York, 1957), p. 81.

Off-energy-shell t matrix for local potentials with nonlocal square well core interaction*

O. Zohni†
International Center for Theoretical Physics, Trieste, Italy
(Received 27 March 1972)

An analytic expression is obtained for the s-wave t matrix of a nonlocal square well core interaction combined with an outside local square well. The result is compared with the t matrix of the hard-core square well. In particular, the t matrix obtained is found to remain finite for large values of the energy parameter in contrast to the hard-core square well t matrix.

1. INTRODUCTION

A mathematically interesting interaction proposed recently by Razavy[1] is that of the nonlocal square well. This interaction hasthe characteristic feature that its kernel is related to the Green's function of the second-order differential equation. In the present paper, analytic expressions are derived for the off-shell two-body t matrix for a combination of an outside local square well with a core interaction of the nonlocal square well type.

Actually, our calculation presents an example where exact analytic expressions of the off-shell t matrix could be obtained.[2-4] Using a wavefunction approach, which is generalized here to the case of nonlocal potentials, Van Leeuven and Reiner[2] obtained explicit expressions for the t matrix in case of a local square well combined with a hard-core interaction. We also aim at comparing the present nonlocal-core t matrix with that of the hard core, which allows to demonstrate the more generalized features of the former.

It should be remarked that the interest in nonlocal potentials has recently evolved in connection with nuclear structure investigations,[5] being suggested as possible replacements for the singular potential in perturbation expansions. Using our t matrix, it should be possible to make simple and rather indicative study of such nonlocality effects in a number of situations. For example, it is the main input to the Faddeev equations which provide exact solutions to the three-body problem,[6] and it is also closely related to the Brueckner G matrix of nuclear matter.[7]

In Sec. 2 the analytic expressions are given and their properties are discussed. The comparison is also made with the case of the hard-core square well potential. In particular, the nonlocal-core t matrix is found to remain finite as the energy parameter goes to infinity in contrast to the hard-core t matrix. Further, the t matrix obtained is shown to contain the features of the hard-core t matrix as a limiting case.

2. THE OFF-SHELL t MATRIX

The two-body t matrix satisfies the Lippman–Schwinger equation:

$$t(q) = v + v\,\frac{1}{(\hbar^2q^2/2\mu) - (P_{\rm op}^2/2\mu)}\,t(q), \tag{2.1}$$

where $q = (2\mu E/\hbar^2)^{1/2}$ with E the (complex) energy parameter and $P_{\rm op}$ is the relative momentum operator having the normalized plane waves

$$\langle \mathbf{r}\,|\mathbf{k}\rangle = (2\pi)^{-3/2}\,\exp(i\mathbf{k}\cdot\mathbf{r}) \tag{2.2}$$

as state eigenfunctions.

The potential v in (2.1) will be assumed to be central and nonlocal. One then writes

$$v(\mathbf{r},\mathbf{r}') = \sum_l \frac{(2l+1)}{4\pi}\,\frac{v_l(r,r')}{rr'}\,P_l(\hat{\mathbf{r}},\hat{\mathbf{r}}'). \tag{2.3}$$

The approach of Van Leeuven and Reiner can be readily generalized to the case of nonlocal potentials. The off-shell matrix elements of the t operator in (2.1) may be expressed as

$$\langle \mathbf{k}'\,|t(q)\,|\mathbf{k}\rangle = \langle \mathbf{k}'\,|\psi_{q,\mathbf{k}}\rangle, \tag{2.4}$$

where

$$|\psi_{q,\mathbf{k}}\rangle = |\mathbf{k}\rangle + \frac{1}{(\hbar^2q^2/2\mu) - (P_{\rm op}^2/2\mu)}\,v\,|\psi_{q,\mathbf{k}}\rangle. \tag{2.5}$$

It is useful here to consider the coordinate representation of the above equation whose radial part may then be expressed as

$$u_l(q,k;r) = rj_l(kr) - \frac{2\mu}{\hbar^2}\int_0^\infty dr'\int_0^\infty dr''G_{l,q}(r,r') \times v_l(r',r'')u_l(q,k;r'') \tag{2.6}$$

where

$$\langle \mathbf{r}\,|\mathbf{k}\rangle = (2\pi)^{-3/2}\sum_l(2l+1)i^lj_l(kr)P_l(\hat{\mathbf{k}},\hat{\mathbf{r}}), \tag{2.7}$$

$$\langle \mathbf{r}\,|\psi_{q,\mathbf{k}}\rangle = (2\pi)^{-3/2}\sum_l(2l+1)i^l\,\frac{u_l(q,k;r)}{r}\,P_l(\hat{\mathbf{k}},\hat{\mathbf{r}}), \tag{2.8}$$

$j_l(kr)$ is the usual spherical Bessel function, and $P_l(\hat{\mathbf{k}},\hat{\mathbf{r}})$ is a Legendre polynomial. Also, $G_{l,q}(r,r')$ in Eq. (2.6) is a radial Green's function satisfying

$$\left(\frac{d^2}{dr^2} + q^2 - \frac{l(l+1)}{r^2}\right)G_{l,q}(r,r') = -\,\delta(r-r'), \tag{2.9}$$

which on introducing into (2.6) gives the following integro-differential equation for $u_l(q,k;r)$:

$$\left(\frac{d^2}{dr^2} + q^2 - \frac{l(l+1)}{r^2}\right)u_l(q,k;r) = (q^2-k^2)rj_l'(kr) + \frac{2\mu}{\hbar^2}\int_0^\infty dr'v_l(r,r')u_l(q,k;r'). \tag{2.10}$$

By substituting the expansion

$$\langle \mathbf{k}'\,|t(q)\,|\mathbf{k}\rangle = \frac{\hbar^2}{2\mu}\,\frac{1}{2\pi^2}\sum_l(2l+1)t_l(k',k;q)P_l(\hat{\mathbf{k}},\hat{\mathbf{k}}') \tag{2.11}$$

into (2.4), it is straightforward to write the lth matrix element for a nonlocal potential as

$$t_l(k',k;q) = -\,\frac{2\mu}{\hbar^2}\int_0^\infty dr\int_0^\infty dr'\,r'j_l(k'r')v_l(r,r')u_l(q,k;r). \tag{2.12}$$

From now on we consider only $l = 0$, since the wavefunction $u_l(q,k;r)$ cannot be obtained analytically for $l \neq 0$ in case of the potential chosen for the core region. In particular, this is the nonlocal square well interaction[1] which is derived from a Green's function-type equation of the form

$$\left(\frac{d^2}{dr^2} - \beta^2\right) V_0(r,r') = -\frac{\hbar^2}{2\mu} U_0\beta^2\delta(r-r') \tag{2.13}$$

having the solution

$$v_0(r,r') = \frac{\hbar^2}{2\mu}\frac{U_0\beta}{\sinh\beta r_0} \times \begin{cases} \sinh\beta r \sinh\beta(r'-r_0) & \text{for } 0 \leq r < r' \leq r_0, \\ \sinh\beta(r-r_0)\sinh\beta r' & \text{for } 0 \leq r' < r \leq r_0. \end{cases} \tag{2.14}$$

Here, r_0 is the range of the interaction and $1/\beta$ defines the range of nonlocality. It is of interest to note from (2.13) that in the limit of $\beta \to \infty$ the interaction becomes strictly local, whereas in the limit of $\beta \to 0$ it turns out to be purely separable.[1] Further, the parameter U_0 in (2.14) is defined such that it is positive for a repulsive interaction and negative for an attractive one.[1]

To study the features of the solution, we first consider the t matrix for the nonlocal square well only and then proceed to get the result for its combination with an outside local square well.

A. The nonlocal square well

The potential $v_0(r,r')$ in this case is given by (2.14) for $0 \leq r < r' \leq r_0$ and vanishes outside the range r_0, i.e.,

$$v_0(r,r') = (\hbar^2/2\mu)\delta(r-r')U_1 = 0 \tag{2.15}$$

with U_1 being assumed the strength of a local square well outside r_0.

Using Eq. (2.9) into (2.6), one easily gets the following integro-differential equation for the wavefunction $u_0(q,k;r)$:

$$\left(\frac{d^2}{dr^2} + q^2\right) u_0(q,k;r) = (q^2-k^2)rj_0(kr) + \frac{2\mu}{\hbar^2}\int_0^\infty dr'\, v_0(r,r')u_0(q,k;r'). \tag{2.16}$$

Taking advantage of Eq. (2.13) defining the nonlocal interaction, one then obtains, on applying the operator $L_0 = (d^2/dr^2) - \beta^2$ to both sides of (2.16), the following differential equation for $u_0(q,k;r)$:

$$\frac{d^4u_0}{dr^4} - (\beta^2+q^2)\frac{d^2u_0}{dr^2} - \beta^2(q^2-U_0)u_0 = (k^2-q^2)(k^2+\beta^2)rj_0(kr) \tag{2.17}$$

with the appropriate boundary conditions and where use has been made of

$$L_0 f_0(r) = (k^2-q^2)(k^2+\beta^2)rj_0(kr) - U_0\beta^2 u_0(q,k;r), \tag{2.18}$$

where $f_0(r)$ stands for the right-hand side of (2.16).

One notes here the mathematical flexibility of the nonlocal interation used which allows the transformation from an integro-differential equation to a differential equation for the wavefunction. This, of course, is made possible through its relation to the Green's function of the second-order differential equation as seen from Eq. (2.13).

The solution of Eq. (2.17) is then easily obtained:

$$u_0(q,k;r) = A_0 rj_0(kr) + B_0\sinh\nu^- r + \tilde{B}_0\sinh\nu^+ r \quad \text{for } 0 \leq r \leq r_0, \tag{2.19}$$

where

$$A_0 = \frac{(k^2-q^2)(k^2+\beta^2)}{[k^4 + k^2(\beta^2-q^2) - \beta^2(q^2-U_0)]}, \tag{2.20}$$

$$\nu^\pm = (1/\sqrt{2})\{(\beta^2-q^2) \pm [(\beta^2-q^2) + 4\beta^2(q^2-U_0]^{1/2}\}^{1/2}. \tag{2.21}$$

Further, the wavefunction for $r \geq r_0$ is obtained by substituting (2.15) into (2.16) which yields the solution

$$u_0(q,k;r) = rj_0(kr) + B_1^+ rh_0^+(\alpha_1 r) \quad \text{for } r \geq r_0, \tag{2.22}$$

where $h_0^+(\alpha_1 r)$ is a spherical Hankel function of the first kind and $\alpha_1 = (q^2 - U_1)^{1/2}$. It is also assumed that $\mathrm{Im}\alpha_1 > 0$ so that the correct boundary conditions are satisfied.

Adopting the procedure given by Van Leeuven and Reiner,[2] the (complex) coefficients B_0, $\tilde{B}_0$, and B_1^+ can be determined from the vanishing of the wavefunction at $r = 0$ and the continuity of its value and derivative at $r = r_0$. This leads to the results $B_0 = -\tilde{B}_0$ and

$$B_0 = -(A_0 - 1)X_0^{-1}(\beta\alpha_1^+|r_0)X_0(k\alpha_1^+|r_0), \tag{2.23}$$

where

$$X_0(\beta\alpha_1^+|r_0) = \begin{vmatrix} u_0(\beta;r) & r_0h_0^+(\alpha_1 r_0) \\ \partial_{r_0}u_0(\beta;r) & \partial_{r_0}rh_0^+(\alpha,r \end{vmatrix} \tag{2.24}$$

with $u_0(\beta;r) = (\sinh\nu^+ r + \sinh\nu^- r)$. $X_0(k\alpha_1^+|r_0)$ has a similar expression with k standing for $rj_0(kr)$. The symbol ∂_{r_0} denotes the differential $d/dr|_{r=r_0}$.

Inserting (2.19) and (2.23) into (2.12), one finally gets the following expression for the s-wave t matrix:

$$t_0(k';k;q) = \frac{U_0\beta^2}{\sinh\beta r_0(k^2+\beta^2)}X_0^{-1}(\beta\alpha_1^+|r_0) \begin{vmatrix} e & d_1 \\ (A_0-1)X_0(k\alpha_1^+|r_0) & X_0(\beta\alpha_1^+|r_0) \end{vmatrix}, \tag{2.25}$$

where

$$e = \frac{A_0}{kk'}\left[\frac{\sin k'r_0}{(\beta^2+k^2)}[(\beta\cosh\beta r_0\sin kr_0) - (k\sinh\beta r_0\cos kr_0)] - \frac{\sinh\beta r_0}{2}\left(\frac{\sin(k'-k)r_0}{(k'-k)} - \frac{\sin(k'+k)r_0}{(k'+k)}\right)\right], \tag{2.26a}$$

$$d_1 = \frac{\sin k'r_0}{k'(\beta^2-\nu^{-2})}(\beta\cosh\beta r_0\sinh\nu^- r_0 - \nu^-\sinh\beta r_0\cosh\nu^- r_0) - \frac{1}{2}\left(\frac{\sinh(\beta-\nu^-)r_0}{(\beta-\nu^-)} - \frac{\sinh(\beta+\nu^-)r_0}{(\beta+\nu^-)}\right) - \frac{\sinh\beta r_0}{(k^2+\nu^{+2}}\left(\frac{(\nu^+\cosh\nu^+ r_0\sin kr_0)}{k-(\sinh\nu^+ r_0\cos kr_0)}\right) + \frac{1}{2}\left(\frac{\sin(k-\nu^-)r_0}{(k-\nu^-)} - \frac{\sin(k+\nu^-)r_0}{(k+\nu^-)}\right). \tag{2.26b}$$

One notes here that the off-shell t matrix in (2.25) not only depends on the range r_0 of the core region as it would for a hard-core potential,[8] but also on the parameters U_0 and β. Such dependence should manifest itself in multiparticle calculations whose sensitivity to off-shell effects connected with the "inner-core" nuclear interaction has recently aroused much interest.[7,9]

The limits of the present t matrix for large k', k, and q^2 are of further interest. One finds

$$\lim_{k'\to\infty} t_0(k', k; q) = 0, \tag{2.27}$$

as $1/k'$ and

$$\lim_{k\to\infty} t_0(k', k; q) = 0, \tag{2.28}$$

as $1/k$ which also follows from the symmetry character of $t_0(k', k; q)$ under the interchange of k, k' as may be inspected from (2.25) and (2.26). The above two limits are similar to what is found in case of the hard-core t matrix.[8]

As regards the limit for large q^2, one notes from (2.25) that the main dependence of $t_0(k', k; q)$ on q^2 is contained in the factor A_0. According to (2.20), this factor obviously exists for all values of q^2. Hence, one concludes that the present t matrix remains finite as $|q^2| \to \infty$. This is in contrast to the hard-core t matrix which goes to ∞ as $|q^2| \to \infty$.[8]

One may remark that the above features persist when an interaction outside the core is allowed.

B. Combination with an outside local square well

The results of the preceding subsection are readily extended to the case of a nonlocal square well combined with an outside local square well. For the region $0 \leq r < r' \leq r_0$, the interaction is again given by (2.14). For the outer region, we have

$$\begin{aligned} v_0(r, r') &= (\hbar^2/2\mu)\delta(r - r')U_1 < 0 \quad \text{for } r_0 \leq r \leq r_1 \\ v_0(r, r') &= (\hbar^2/2\mu)\delta(r - r')U_2 = 0 \quad \text{for } r \geq r_1, \end{aligned} \tag{2.29}$$

where r_0 is the range of the nonlocal core interaction, r_1 is the range of the outer local square well, and the $U_i (i = 1, 2)$ are the potential strengths.

For the potential so defined, the integro-differential equation in (2.10) has the following solutions:

$$\begin{aligned} u_0(q, k; r) &= A_0 r j_0(kr) + B_0 \sinh\nu^- r + \tilde{B}_0 \sinh\nu^+ r, \\ &\qquad \text{for } 0 \leq r \leq r_0 \\ u_0(q, k; r) &= A_1 r j_0(kr) + B_1^+ r h_0^+(\alpha, r) + B_1^- r h_0^-(\alpha_1 r), \\ &\qquad \text{for } r_0 \leq r \leq r_1, \\ u_0(q, k; r) &= r j_0(kr) + B_2^+ h_0^+(\alpha_2 r), \quad \text{for } r \geq r_1, \end{aligned} \tag{2.30}$$

where A_0 is given by (2.20), $A_1 = (q^2 - k^2)/(q^2 - k^2 - U_1)$ and $\alpha_1 = (q^2 - U_1)^{1/2}$ with $i = 1, 2$. The (complex) coefficients $B_0, \tilde{B}_0, B_1^+$, ord B_2^+ are again determined from the matching of the wavefunctions.

By proceeding in a similar way as in the previous subsection, the t matrix may then be expressed as

$$t_0(k', k; q) = \frac{U_0\beta^2}{(k^2 + \beta^2)\sinh\beta r_0} D_0^{-1}(r_0, r_1; \alpha_1, \alpha_2; \beta) \begin{vmatrix} e & d_1 & d_2 & d_3 \\ (A_0 - A_1) r_0 j_0(kr_0) & u_0(\beta; r_0) & -r_0 h_0^+(\alpha_1 r_0) & -r_0 h_0^-(\alpha_1 r_0) \\ (A_0 - A_1)\partial_{r_0} r j_0(kr) & \partial_{r_0} u_0(\beta; r) & -\partial_{r_0} r h_0^+(\alpha_1 r) & -\partial_{r_0} r h_0^-(d_1 r) \\ (A_1 - 1) X_0(k\alpha_2^+ | r_1) & 0 & X_0(\alpha_1^+\alpha_2^+ | r_1) & X_0(\alpha_1^-\alpha_2^+ | r_1) \end{vmatrix} \tag{2.31}$$

where

$$\begin{aligned} e &= -A_0 \int_0^{r_0} dr\, \chi(k', \beta; r) r j_0(kr) - A_1 U_1 \int_{r_0}^{r_1} dr\, r j_0(k'r) r j_0(kr), \\ d_1 &= -\int_0^{r_0} dr\, \chi(k', \beta; r) u_0(\beta; r), \\ d_2 &= -U_1 \int_{r_0}^{r_1} dr\, \chi(k', \beta; r) r h_0^+(\alpha, r), \\ d_3 &= -U_1 \int_{r_0}^{r_1} dr \chi(k', \beta; r) r h_0^-(\alpha, r), \end{aligned} \tag{2.32}$$

$$\begin{aligned} D_0 &= u_0(\beta; r_0)[-\partial_{r_0} r h_0^+(\alpha, r) X_0(\alpha_1^-\alpha_2^+ | r_1) \\ &\quad + \partial_{r_0} r h_0^-(\alpha_1 r) X_0(\alpha_1^+\alpha_2^+ | r)] - \partial_{r_0} u_0(\beta; r) \\ &\times [-r_0 h_0^+(\alpha_1 r_0) X_0(\alpha_1^-\alpha_2^+ | r_1) + r_0 h_0^-(\alpha_1 r_0) X_0(\alpha_1^+\alpha_2^+ | r_1)], \end{aligned}$$

with $\chi(k', \beta; r)$ given by

$$\begin{aligned} \chi(k', \beta; r) = \{&\sin k'r[\sinh\beta(r - r_0)\sinh\beta r \\ &- \cosh\beta(r - r_0)\sinh\beta r] + \sinh\beta r \sin k r_0\}. \end{aligned} \tag{2.33}$$

For the purpose of the comparison with the t matrix of a hard-core combined with a local square well,[2,10] it is useful to consider the poles of the t matrix in (2.31) which arise from negative real q^2 values defining bound states for which $D_0(r_0, r_1; \alpha_1, \alpha_2; \beta)$ vanishes. For the case such that $U_1 < q^2 < 0$ which develops zeros for D_0, one gets the following condition on the binding energies:

$$\begin{aligned} &\{(|q^2|)^{1/2} \sin[(q^2 - U_1)^{1/2}(r_1 - r_0)] - (q^2 - U_1)^{1/2} \\ &\quad \times \cos[(q^2 - U_1)^{1/2}(r_1 - r_0)]\} \\ &= -\frac{u_0(\beta; r_0)}{\partial_{r_0} u(\beta; r)} (q^2 - U_1)^{1/2} \{(|q|^2)^{1/2} \\ &\quad \times \cos[(q^2 - U_1)^{1/2}(r_1 - r_0)] \\ &\quad - (q^2 - U_1)^{1/2} \sin[(q^2 - U_1)^{1/2}(r_1 - r_0)]\}. \end{aligned} \tag{2.34}$$

Noting that the function $u_0(\beta; r)$, which represents a solution for the nonlocal square well alone, should vanish in the limit of the hard core,[11] one gets immediately the condition

$$\begin{aligned} &\{(|q^2|)^{1/2} \sin[(q^2 - U_1)^{1/2}(r_1 - r_0)] + (q^2 - U_1)^{1/2} \\ &\quad \times \cos[(q^2 - U_1)^{1/2}(r - r_0)]\} = 0, \end{aligned} \tag{2.35}$$

which is exactly the condition found for the hard-core square well t matrix.[10]

A further check is to find the parameters U_0, U_1, r_0, r_1,

and β which give zero binding energy. Setting $q^2 = 0$ in (2.34), one gets the condition

$$\cot[(-U_1)^{1/2}(r_1 - r_0)] = (-U_1)^{1/2}\left(\frac{u_0(\beta; r_0)}{\partial_{r_0} u_0(\beta, r)}\right), \quad (2.36)$$

which gives in the limit of the hard core

$$\cos[(-U_1)^{1/2}(r_1 - r_0)] = 0. \quad (2.37)$$

Again, this is the same as that found in case of the hard-core square well.[10]

It is aimed in future work to study the features of the present t matrix as applied to multiparticle problems. Of particular interest in this respect is to compare the off-shell effects associated with the nonlocal-core t matrix against those associated with the hard-core t matrix. This would throw some light on the sensitivity of multiparticle calculations to effects connected with the inner-core nuclear interaction.

ACKNOWLEDGMENTS

The author is thankful to Professor J. Dabrowski for a careful reading of the manuscript and for useful comments. He wishes also to thank Professor A. Salam, Professor P. Budini, the International Atomic Energy Agency and UNESCO for hospitality offered to him at the International Centre for Theoretical Physics.

*ICTP Preprint IC/71/90.

†Present address: Institute of Theoretical Physics, Warsaw University, Warsaw, Poland.

[1] M. Razavy, Nucl. Phys. **78**, 256 (1966); V. De La Cruz, B. A. Orman, and M. Razavy, Can. J. Phys. **44**, 629 (1966).
[2] J. M. J. Van Leeuven and A. S. Reiner, Physica (Utr.) **27**, 99 (1961).
[3] M. G. Fuda, J. Math. Phys. **12**, 1163 (1971).
[4] Y. Yamaguchi, Phys. Rev. **95**, 1628 (1954).
[5] A. K. Kerman, in *Cargese lectures in physics* (Gordon and Breach, New York, 1969), Vol. 3, p. 395.
[6] See, e.g., Y. E. Kim and A. Tubis, Phys. Rev. C **1**, 1628 (1970) and references cited there.
[7] See, e.g., M. I. Haftel and F. Tabakin, Phys. Rev. C **3**, 921 (1971) and references cited there; M. E. Grypeos, J. Math. Phys. **12**, 1655 (1971).
[8] R. Laughlin and B. L. Scott, Phys. Rev. **171**, 1196 (1968).
[9] V. F. Kharchenko, S. A. Shadchin, and S. A. Storozhenko, Phys. Lett. B **37**, 131 (1971).
[10] Y. E. Kim and Tubis, Phys. Rev. C **1**, 414 (1970).
[11] R. G. Newton, *Scattering theory of waves and particles* (McGraw-Hill, New York, 1966), p. 419.

Space-times with a future projective infinity

D. Eardley*
Department of Physics, University of California, Berkeley, California 94720

R. K. Sachs†
Department of Mathematics and Department of Physics, University of California, Berkeley, California 94720
(Received 31 July 1972)

We use a projective structure to make precise the concept of "future timelike infinity" for certain space-times. We apply our definitions to many exact and approximate solutions of the Einstein equations. Some physical restrictions are necessary.

I. FUTURE PROJECTIVE INFINITY

Penrose's elegant concept of lightlike conformal infinity[1] is hard to apply to rest mass nonzero fields or to timelike curves on a space-time. We shall give a definition intended to supplement Penrose's ideas. The long-range goal is to find ways of analyzing all the "information" that "leaks" into or out of a space-time through infinity or singularities. The basic ideas of our treatment are due to Hawking and Geroch.[2]

We will use the conventions of Misner, Thorne, and Wheeler[3] and of Bishop and Goldberg.[4] M will be a connected, Hausdorff, real, four-dimensional C^k manifold ($k \geq 2$) throughout. Greek indices will run over $\{0, 1, 2, 3\}$ and Latin indices over $\{1, 2, 3\}$; the summation convention is used throughout.

We review some facts about projective structures.[5,6] Let Γ and $\overline{\Gamma}$ be C^{k-2} symmetric linear connections on M. Γ and $\overline{\Gamma}$ are called *projectively equivalent* if every geodesic of Γ can be reparametrized so that it becomes a geodesic of $\overline{\Gamma}$. In local coordinates Γ and $\overline{\Gamma}$ are projectively equivalent iff

$$\Gamma^{\lambda}_{\mu\nu} - \tfrac{1}{5}(\Gamma^{\rho}_{\mu\rho}\delta^{\lambda}_{\nu} + \Gamma^{\rho}_{\nu\rho}\delta^{\lambda}_{\mu}) = \overline{\Gamma}^{\lambda}_{\mu\nu} - \tfrac{1}{5}(\overline{\Gamma}^{\rho}_{\mu\rho}\delta^{\lambda}_{\nu} + \overline{\Gamma}^{\rho}_{\nu\rho}\delta^{\lambda}_{\mu}). \qquad (1)$$

We shall now use projective structure to define "timelike future infinity" as a 3-manifold for certain space-times. Let ds^2 be a C^{k-1} Lorentzian metric on M; let Γ be the Levi-Civita connection. Then (M, ds^2, Γ) is a space-time, assumed time-orientable and time-oriented henceforth. We shall say (M, ds^2, Γ) has a *boring future* iff every future-inextendible timelike geodesic has infinite length. Unless explicitly stated otherwise, each space-time considered has a boring future.

One can group the timelike geodesics into equivalence classes as follows. Let TM be the tangent bundle with projection π; $p \in TM$ will be denoted by $p = (m, v)$ with $m = \pi p \in M$ and $v \in M_m$. The *future* $F \subset TM$ is $F = \{p \mid p \in TM, v \text{ timelike future-pointing}\}$. Let $U \subseteq F$ be an open neighborhood of $p \in F$. Let $V \subseteq F$ be the open set swept out by moving U indefinitely toward the future along the geodesic flow. Then $\pi V \subseteq M$ is an open set, called a *future thickening* of the geodesic with initial point πp and initial tangent v. Two future-pointing timelike geodesics $\gamma, \bar{\gamma}$ will be called *eventually parallel* if every future thickening of either contains a future thickening of the other. Eventual parallelism is an equivalence relation. We shall be interested in the case that a 3-manifold structure can be assigned to the set of equivalence classes, as follows.

We shall say (M, ds^2, Γ) has a C^k *regular future projective infinity* τ iff there exists a C^k Hausdorff manifold-with-boundary $\overline{M}$ and a C^{k-2} symmetric connection $\overline{\Gamma}$ on $\overline{M}$ such that (a) τ is the boundary of $\overline{M}$; (b) $\overline{M} = M \cup \tau$; (c) Γ and $\overline{\Gamma}|_M$ are projectively equivalent; (d) for each $p \in F \subset TM$ the corresponding geodesic of $\overline{\Gamma}$ on $\overline{M}$ can be extended to intersect τ; then at least one geodesic hits each point of τ. It can be shown that when $(\overline{M}, \overline{\Gamma})$ exists the geodesics of $\overline{\Gamma}$ which hit any one point of τ correspond to the geodesics of precisely one equivalence class of eventually parallel geodesics: Two future-pointing timelike geodesics intersect at future timelike infinity iff they are eventually parallel.

Minkowski space-time has a regular future projective infinity. The following example does not. Take Minkowski space-time with the usual coordinates and metrics $ds^2 = dt^2 - dx^2 - dy^2 - dz^2$. Identify points $(t, x, y, z) \equiv (t, x + k, y + l, z + m)$, where k, l, and m are arbitrary integers, to form M. M is thus $R^1 \times S^1 \times S^1 \times S^1$. Time orient (M, ds^2); then it has a boring future. It has no future projective infinity because every future thickening contains $(A, \infty) \times S^1 \times S^1 \times S^1$ for some A, so a 3-manifold τ cannot exist.

We now give examples which clarify somewhat the physical restrictions needed to insure that (M, ds^2, Γ) has a regular future projective infinity. We take cosmological constant $\Lambda = 0$ unless stated otherwise. The strategy in each example is the following. Choose local coordinates $\{x^{\mu}\}$ such that $x^0 = \text{const}$ is a plausible candidate for all or part of τ. Define a $\overline{\Gamma}$ by

$$\overline{\Gamma}^{\lambda}_{\mu\nu} = \Gamma^{\lambda}_{\mu\nu} - \tfrac{1}{5}(\Gamma^{\rho}_{\mu\rho}\delta^{\lambda}_{\nu} + \Gamma^{\rho}_{\nu\rho}\delta^{\lambda}_{\mu}) \qquad (2)$$

on the coordinate patch. Then try to extend $\overline{\Gamma}$ to τ by continuity. $\overline{\Gamma}$ depends on the choice of coordinate patch, but when a $(\overline{M}, \overline{\Gamma})$ exists the projective structure is unique.

II. HOMOGENEOUS COSMOLOGICAL MODELS

We shall show that $k \leq 0$ Friedmann models have a regular future projective infinity. We shall state some results and conjectures on other homogeneous cosmological models.

The Friedmann models[7] begin with a curvature singularity (big bang) in the past. Each $k = +1$ (closed) model ends with a bang and does not have a boring future. By estimating the behavior of timelike geodesics one finds that each $k \leq 0$ model has a boring future; in fact, a $k = -1$ model approaches Minkowski space-time in the future. We now exhibit a regular future projective infinity when $k \leq 0$.

Let $(L, d\sigma^2)$ be a C^{∞} 3-dimensional Riemannian manifold and (M, ds^2) a space-time with $M = (0, \infty) \times L$ and $ds^2 = R^2(\eta)(d\eta^2 - d\sigma^2)$, where $0 < \eta < \infty$. For the $k = 0$ Friedmann model $(L, d\sigma^2)$ is flat and $R(\eta) = \eta^2$; for $k = -1$, $(L, d\sigma^2)$ is of constant negative curvature and

$R(\eta) = a(\cosh\eta - 1)$, where $a > 0$ is a constant. We work in a patch $\eta \in (1, \infty)$ to avoid the big bang. Let

$$x^0 \equiv \chi = \int_\eta^\infty \frac{d\eta'}{R(\eta')}. \tag{3}$$

The integral exists and is finite for our examples; for $k = 0$, $\chi = 1/\eta$; for $k = -1$, $\chi = 2/[a(\exp\eta - 1)]$. As $\eta \to \infty$, $\chi \to 0^+$. For $\chi(\eta = 1) > \chi > 0$, corresponding to $1 < \eta < \infty$,

$$ds^2 = R^2[R^2 d\chi^2 - d\sigma^2]. \tag{4}$$

If $\{x^i\}$ are local coordinates on L, then $\{\chi, x^i\} \equiv \{x^\mu\}$ are local coordinates on M for $\chi(\eta = 1) > \chi > 0$, and we simply attach $\chi = 0$ as the boundary τ to M.

Now define $\overline{\Gamma}$ by Eq. (2). The only coefficients of Γ and $\overline{\Gamma}$ which depend on χ are

$$\Gamma^0_{00} = 2R'/R, \quad \Gamma^i_{j0} = \delta^i_j R'/R, \quad \Gamma^0_{ij} = (d\sigma^2)_{ij} R'/R^3,$$
$$\overline{\Gamma}^0_{00} = 0, \quad \overline{\Gamma}^i_{j0} = 0, \quad \overline{\Gamma}^0_{ij} = (d\sigma^2)_{ij} R'/R^3, \tag{5}$$

where $R' \equiv dR/d\chi$. In the $k \leq 0$ Friedmann models, $R'/R \to \infty$ as $\chi \to 0$ but $R'/R^3 \to 0$; in fact, the latter is a polynomial in χ and hence is C^∞ at $\chi = 0$. So $\overline{\Gamma}$ is C^∞ on $\overline{M}$, and each of these models has a C^∞ regular future projective infinity.

Some further results about cosmological models are: (a) Any $p \leq \rho/3$, $k = -1$ Robertson–Walker[8] model has a regular future projective infinity by the construction just given; so does any $p < \rho/3$, $k = 0$ Robertson–Walker model. (b) Any Heckmann–Schücking[7] dust model has a regular future projective infinity. (c) To mention a case with positive cosmological constant, a de Sitter[7] space-time has a regular future projective infinity.

Some unproved conjectures about homogeneous cosmological models are: (a) A dust model which is "approximately Robertson–Walker" in the far future in the sense of MacCallum[9] has a boring future and a regular future projective infinity. (b) Kasner vacuum[10] and dust Kantowski–Sachs[11] models do not have a regular future projective infinity.

III. BONDI-TOLMAN MODELS

We next consider a broad subclass of Bondi–Tolman models, which are spherically symmetric dust solutions.[12,13] We shall be able to construct a regular future projective infinity only if the dust is asymptotically homogeneous in the future in a sense that will be made precise below.

In the usual comoving coordinates ds^2 for these models has the form

$$ds^2 = dt^2 - X^2(r,t)dr^2 - Y^2(r,t)d\Omega^2. \tag{6}$$

Transformations $r \to f(r)$ and time translations $t \to t +$ const maintain this form. One of the Einstein equations implies that $(\partial Y/\partial r)/X \equiv W(r)$ is independent of t. We shall be concerned only with solutions for which $W(0) = 1$, $W'(r) > 0$ if $r > 0$, and $W(r) \to \infty$ as $r \to \infty$. These conditions insure that each mass shell $r =$ const will start at a curvature singularity $Y = 0$ and expand to infinite area in the future instead of recollapsing to another singularity. Now fix the radial coordinate as $v = [W^2(r) - 1]^{1/2}$, $0 \leq v < \infty$. Then each solution that we shall consider is specified by two functions $a(v)$ and $t_0(v)$, assumed C^∞ for $0 \leq v < \infty$. The metric and density ρ are given in terms of a parameter η, $0 < \eta < \infty$:

$$t = a(v)(\sinh\eta - \eta) + t_0(v),$$
$$Y(v,t) = va(v)(\cosh\eta - 1),$$
$$X(v,t) = \left[\left(\frac{a'}{a} + \frac{1}{v}\right) Y - v\left(1 + \frac{2av}{Y}\right)^{1/2} \left(\frac{a'}{a}(t - t_0) + t_0'\right)\right] \Big/ (v^2 + 1)^{1/2}, \tag{7}$$
$$\rho(v,t) = 3(av^3)'/(4\pi Y^3).$$

Here and below the prime $'$ denotes $\partial/\partial v$ and $c = G = 1$. We require $t_0' \leq 0$ and $a(v) > 0$, $(av^3)' \geq 0$. Then each solution (7) has everywhere nonnegative density and has a boring future.

The $k = -1$ Friedmann models appear as a special case $a =$ const, $t_0 =$ const.

First, consider the subcase $a =$ const. We shall construct a regular future projective infinity. We work in the region $\eta > 1$ to avoid the initial singularity at $t = t_0$. Define coordinates χ, u by

$$\chi = 2/[a(\exp\eta - 1)],$$
$$u = v + (v^2 + 1)t_0'\chi. \tag{8}$$

Then $\chi \to 0$ as $t \to \infty$. Future projective infinity will be the surface $\chi = 0$. Note that $u = v + O(\chi)$ near $\chi = 0$. Here and below $O(f(\chi))$ signifies a quantity vanishing at least as fast as $f(\chi)$ as $\chi \to 0$ for fixed u.

The metric in terms of χ and u is

$$ds^2 = \chi^{-4}Ad\chi^2 + 2\chi^{-1}Bd\chi du - \chi^{-2}(Cdv^2 + Dd\Omega^2), \tag{9}$$

where A, B, C, and D are C^∞ functions of χ and u near $\chi = 0$ satisfying

$$A = 1 - 2a\chi + O(\chi^2),$$
$$C = (v^2 + 1)^{-1} + O(\chi), \tag{10}$$
$$D = v^2 + O(\chi).$$

Now define $\overline{\Gamma}$ through (2). One calculates that the $\overline{\Gamma}^\lambda_{\mu\nu}$ are C^∞ near $\chi = 0$. So the models with $a =$ const have a regular future projective infinity τ at $\chi = 0$.

Now consider the case $a \neq$ const. The above construction fails. One can find coordinates such that (9) holds with A, B, C, and D at least C^0 at $\chi = 0$, and with $A = 1 - 2\chi f(u) + O(\chi^2(\ln\chi)^2)$ for some C^∞ function $f(v)$ determined by $a(v)$. But now $\overline{\Gamma}^u_{00} = -(u^2 + 1)f'/\chi + O(1)$, while the rest of the coefficients of $\overline{\Gamma}$ are $O(1)$. When $a =$ const, then $f = a$ and the singular term is absent. On the basis of further computations we conclude that it is very unlikely that the models with $a \neq$ const have a regular future projective infinity. The boundary τ still exists, but $\overline{\Gamma}$ is not regular there.

We interpret the restriction $a =$ const as follows. In a model with $a =$ const one finds that the boundary τ is totally geodesic, and so inherits a unique projective structure $\overline{\Gamma}|_\tau$ from the projective structure $\overline{\Gamma}$ of space-time. From (2), (9), and (10), $\overline{\Gamma}|_\tau$ is determined by the nonvanishing part of the spatial metric in the surfaces

$\chi = \text{const}$ at $\chi = 0$, which is

$$d\sigma^2 = \frac{du^2}{(u^2 + 1)} + u^2 d\Omega^2. \tag{11}$$

Because $(\tau, d\sigma^2)$ is a space of constant curvature, $(\tau, \overline{\Gamma}|_\tau)$ is projectively flat.[5] In turn, $d\sigma^2$ is determined by $\overline{\Gamma}|_\tau$ up to projective equivalence. We can then fix $d\sigma^2$ uniquely by a construction employing the metric structure of space-time. Thus τ has an invariant volume element

$$(u^2 + 1)^{-1/2} u^2 du\, d\Omega. \tag{12}$$

From (7) we find that the total rest mass of a shell dv in space-time is

$$dM = (3a/4\pi)(v^2 + 1)^{-1/2} v^2 dv\, d\Omega \tag{13}$$

if $a = \text{const}$, and is of course independent of t. Compare (12) and (13), and recall that u and v agree at τ. We may say that if $a = \text{const}$, the density of rest mass is constant on τ. Further, any attempt to put an inhomogeneous rest-mass density on τ by taking $a \neq \text{const}$ also destroys the regular projective structure there. So a Bondi–Tolman space-time has a regular future projective infinity only if the rest mass is asymptotically homogeneous in the future in this sense.

Further calculations show that if the cosmological constant $\Lambda > 0$, a Bondi–Tolman space-time that expands to vanishing density and has a boring future always has a C^2 regular future projective infinity.

A coordinate patch covering a part of the Schwarzschild space-time can be obtained from (6) and (7) by taking $t_0 = 0$, $av^3 = \text{const}$. This space-time does not have a boring future. However, from those timelike geodesics which escape to infinity with nonvanishing velocity we can fashion a partial boundary manifold τ. We can discuss the behavior of the space-time projective structure at τ.

Since $a \neq \text{const}$, the space-time does not have a partial regular future projective infinity. However, there are coordinates χ, u so that τ lies at $\chi = 0$ and

$$\overline{\Gamma}^u_{\chi\chi} = g(u)/\chi + G(\chi, u) \tag{14}$$

and $G(\chi, u)$ and all the rest of the $\overline{\Gamma}^\lambda_{\mu\nu}$ are C^∞ near $\chi = 0$.

In fact in a well-defined sense $(\tau, \overline{\Gamma}|_\tau)$ is still totally geodesic and projectively flat.

IV. DUST PERTURBATIONS OF k ≤ 0 FRIEDMANN MODELS

We now argue that each member of a rather wide class of solutions of the dust Einstein equations has a regular future projective infinity. We shall use dust perturbations of the $k \leq 0$ Friedmann models in the linear approximation. These perturbation models incorporate the Einstein equations into the analysis in a physically reasonable way. They are also rather general, exhibiting arbitrary small density, rotational, and gravitational-wave perturbations. We shall be able to construct a regular future projective infinity only if the rest mass is asymptotically homogeneous in the future.

First we shall recall some results about these models.[14,15] The cases $k = 0$ and $k = -1$ are similar in most respects, and we shall work with the former. The outstanding difference between the two cases concerns one of the modes of density perturbation, which we shall call the "eighth mode." When $k = 0$ this mode is unstable at large times and the perturbation approximation fails. Physically, this mode may lead to gravitational collapse, destroying the boring-future property. So we will not consider the eighth mode for $k = 0$. When $k = -1$, this mode is stable.

In the notation of Sec. II the perturbed metric for $k = 0$ is

$$ds^2 = R^2(\eta)[d\eta^2 - d\sigma^2 + h_{\mu\nu} dx^\mu dx^\nu], \tag{15}$$

where $h_{\mu\nu}$ is the small perturbation. The coordinates are $x^0 \equiv \eta$ and x^i. All tensor calculus below will be done with respect to the metric $d\sigma^2 = dx^i dx^i$. We take the usual gauge condition[14] and give initial conditions on an initial hypersurface $\eta = 1$. When the perturbed Einstein equations are solved for $h_{\mu\nu}$, eight functions of integration arise, as follows:

(1) "Density perturbations" are governed by two scalar functions $A(x^i), B(x^i)$.

(2) "Dust rotations" are governed by a vector field $C_i(x^j)$ subject to $C_{i,i} = 0$, determined by two functions of the x^j.

(3) "Gravitational waves" are governed by a tensor field $D_{ij}(x^k, \eta)$ subject to $D_{ij} = D_{ji}$, $D_{ij,j} = 0 = D_{ii}$, and $(\partial^2/\partial\eta^2 - \nabla^2) D_{ij} = 0$, determined by four functions of the x^k given at $\eta = 1$; call these functions $d_p(x^i)$, $p = 1, 2, 3, 4$.

Here and below $_{,i} \equiv \partial/\partial x^i$ and $\nabla^2 \equiv \partial^2/\partial x^i \partial x^i$. Then $h_{\mu\nu}$ is

$$\begin{aligned} h_{00} &= 0, \\ h_{0i} &= -2\eta^{-2} \nabla^2 C_i, \\ h_{ij} &= \eta^{-3} A_{,ij} - (\delta_{ij} B + \eta^2 B_{,ij}/10) \\ &\quad - (16\eta^{-3} - 2\eta^{-1}\nabla^2)(C_{i,j} + C_{j,i}) \\ &\quad + \eta^{-1}(\partial/\partial\eta)(\eta^{-1} D_{ij}). \end{aligned} \tag{16}$$

As $\eta \to \infty$ all terms remain duly small except for that part containing B, which grows without bound. B is the eighth mode mentioned above, and we shall set $B = 0$.

When $k = -1$ the results are similar,[15] except that the eighth mode is stable.

One finds from the linearly perturbed geodesic equations that these models have a boring future. Now define $\chi \equiv x^0$ by (3) in order to adjoin $\chi = 0$ as the boundary τ and calculate $\overline{\Gamma}$ through (2), exactly as in Sec. II. For the first seven modes the analysis is similar for $k = 0$ and $k = -1$, and we will discuss only the former case. Exactly as before the zeroth-order part of $\overline{\Gamma}$ is C^∞ at $\chi = 0$. The first-order terms containing A, C_i, or D_{ij} are all proportional to χ^n for integer $n \geq 0$, so the contribution of A and C_i is C^∞.

D_{ij} itself depends on χ and it is necessary to put some restrictions on its initial conditions d_p. Roughly, as one travels along an unperturbed dust line toward $\chi = 0$ $(\eta = \infty)$ one sees "news" determining D_{ij} coming in along the past light cone from increasingly remote reaches of the initial hypersurface. The news influences the connection; the leading contribution is $\overline{\Gamma}^i_{\chi j} = \frac{1}{2}\nabla^2 D_{ij} + \cdots$. Even if the d_p and all their derivatives are bounded in the initial hypersurface $\eta = 1$, $\overline{\Gamma}$ may

oscillate infinitely often as $\chi \to 0$ and fail to have a limit. We, therefore, restrict the d_p to have compact support. Then $\overline{\Gamma}$ will be C^∞ at $\chi = 0$. It is likely that this restriction can be weakened.

When $k = -1$ and the eighth mode is present, the above construction fails. The boundary manifold τ still exists, but $\overline{\Gamma}$ is singular there, with $\overline{\Gamma}^i_{\chi\chi} = O(\chi^{-1})$. A detailed analysis indicates that a regular future projective infinity does not exist in this case.

Thus each of these models has a regular future projective infinity, as long as the gravitational radiation falls off sufficiently rapidly in the initial hypersurface, and as long as the eighth mode is absent. When regular future projective infinity exists, it is always totally geodesic and projectively flat. The dust is always homogeneously distributed on τ, judged by the invariant volume element of Sec. III.

These results are consistent with those of Sec. III, since a perturbation $a \neq$ const to a homogeneous Bondi–Tolman model is an eighth-mode perturbation.

V. DISCUSSION

Judging from the examples, the requirement that a dust or vacuum model with a boring future have a regular future projective infinity seems to impose two general restrictions: (1) In some sense, only a finite amount of gravitational radiation may be present. (2) The distribution of rest mass must be asymptotically homogeneous in the future. If either of these restrictions is violated, the boundary manifold τ still generally exists, but the projective structure is not regular there. However, the nature of the singularity in the projective structure is quite different in the two cases. Roughly, if restriction (1) is violated, $\overline{\Gamma}$ can exhibit an arbitrarily ugly essential singularity. But if (2) is violated, $\overline{\Gamma}$ has only a single pole term as in (14) and is otherwise smooth. The requirement that the projective structure have only this mild form of singularity is nearly as useful and tractible as the requirement of complete regularity that we have discussed here. We are at present systematically investigating this case.

ACKNOWLEDGMENTS

The authors are grateful to M. MacCallum for suggesting the investigation of cosmological models at large times and to R. Petti for clarifying many aspects of geodesic "thickenings" for us. We thank R. Budic and R. Hermann for comments.

*National Science Foundation Predoctoral Fellow. Present address: Kellogg Laboratory, California Institute of Technology, Pasadena, California 91109.

†Partially supported by the National Science Foundation under Grant No. 21495. Temporary address: Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge, England.

[1]R. Penrose, "Structure of Space–Time" in *Battelle Rencontres. 1967 Lectures in Mathematics and Physics*, edited by C. M. DeWitt and J. A. Wheeler (Benjamin, New York, 1968).

[2]R. Geroch, J. Math. Phys. **9**, 450 (1968).

[3]C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation* (Freeman, San Francisco, 1972).

[4]R. L. Bishop and S. I. Goldberg, *Tensor Analysis on Manifolds* (Macmillan, New York, 1968).

[5]L. P. Eisenhart, *Non-Riemannian Geometry* (Amer. Math. Soc., New York, 1927), Chap. III.

[6]J. A. Schouten, *Ricci-Calculus* (Springer-Verlag, Berlin, 1954), Chap. VI.

[7]G. Heckmann and E. Schücking, "Relativistic Cosmology," in *Gravitation: An Introduction to Current Research*, edited by L. Witten (Wiley, New York, 1962).

[8]R. Adler, M. Bazin, and M. Schiffer, *Introduction to General Relativity* (McGraw-Hill, New York 1965), p. 356.

[9]M. A. H. MacCallum, Commun. Math. Phys. **20**, 57 (1971).

[10]E. Kasner, Am. J. Math. **43**, 217 (1921).

[11]R. Kantowski and R. K. Sachs, J. Math. Phys. **7**, 443 (1966).

[12]R. C. Tolman, Proc. Natl. Acad. Sci. USA **20**, 169 (1934).

[13]H. Bondi, Mon. Not. R. Astron. Soc. **107**, 410 (1947).

[14]R. K. Sachs and A. M. Wolfe, Astrophys. J. **147**, 73 (1967).

[15]E. M. Lifshitz and I. M. Khalatnikov, Usp. Fiz. Nauk **80**, 391 (1963) [Sov. Phys.-Usp. **6**, 495 (1964)]; see also E. M. Lifshitz and I. M. Khalatnikov, Adv. Phys. **12**, 185 (1963).

Closed-form Glauber cross sections in p–H and e^-–H collisions. I*

Brian K. Thomas and E. Gerjuoy
Department of Physics, University of Pittsburgh, Pittsburgh, Pennsylvania 15213
(Received 25 May 1972)

The Glauber predicted integrated (over scattering angle) cross section for the elastic scattering of charged particles by ground state atomic hydrogen is evaluated in closed form. The asymptotic form of the Glauber cross section for large incident particle momenta is obtained and compared with the exact first Born approximation result. The present analytic results, together with previous numerical results, indicate that (neglecting exchange) the Glauber integrated elastic scattering cross section is well represented, *for both incident electrons and incident protons,* by $\sigma_{\text{elastic}}=(7/3)v_i^{-2}\pi a_0^{\,2}$ at essentially *all* incident particle speeds v_i.

I. INTRODUCTION

The Glauber approximation[1] for scattering amplitudes has been applied, with considerable success, to the elastic and inelastic scattering of electrons[2–4] and of protons[5] by ground state atomic hydrogen. For these collisions, the Glauber approximation has been shown to be more useful than Born approximation for estimating differential and integrated (over scattering angle) cross sections. Recently, Thomas and Gerjuoy[6] obtained closed form expressions for these Glauber amplitudes (always neglecting exchange); in particular, for the transitions $1s \to 1s, 2s, 2p$ they found that the Glauber amplitudes reduce to simple sums of hypergeometric functions, which sums are relatively easy to compute. Although these analytic results considerably increase the practical utility of the Glauber approximation in charged particle–hydrogen atom collisions, a further integration over the allowed range of momentum transfers still is required in order to obtain the integrated cross sections from the amplitudes. Heretofore, this integration has been done numerically; it would be a useful reduction of effort if these integrated cross sections could be expressed in closed form. Furthermore, in p + H($1s$) collisions, Franco and Thomas[5] found that the Born and Glauber predicted integrated elastic scattering cross sections are essentially indistinguishalbe, even at moderately low incident proton speeds, despite the fact that the differential cross sections are quite dissimilar.

Therefore, still neglecting exchange, we have reexamined the Glauber predicted cross section integral for the elastic scattering of arbitrary structureless charged particles by ground state atomic hydrogen. By exploiting the known relations between the generalized Legendre functions P_ν and the hypergeometric functions appearing in the elastic scattering amplitude expression obtained by Thomas and Gerjuoy,[6] we have been able to re-express the amplitude as a simple sum of Legendre functions. The ensuing cross section integral then reduces to a sum of known definite integrals. In particular, the elastic scattering cross section reduces to a comparatively simple sum of products of Legendre functions (or hypergeometric functions) which are as easy to compute as the amplitude itself.

The contents of this paper now can be summarized as follows. In Sec. II we describe the reduction of the Glauber integrated elastic ($1s \to 1s$) scattering cross section, for arbitrary incident particle speeds. In Sec. III we obtain the asymptotic form of the cross section for large incident particle momenta and compare this result with the first Born approximation result. These latter results clearly show, for the incident proton energies considered by Franco and Thomas,[5] that the Glauber and Born predicted elastic scattering cross sections should be essentially indistinguishable.

One well might expect that the techniques employed in this present paper also should be usable to evaluate the Glauber predicted cross sections for the excitation of atomic hydrogen. This is indeed the case: The closed form evaluation of the Glauber predicted cross sections for excitation of the $n = 2$ level will be presented in a future paper.

II. THE GLAUBER 1s-1s ELASTIC SCATTERING CROSS SECTION

A. The amplitude and cross section formulas

We are concerned with the scattering of an arbitrary, structureless, spinless particle of charge Z by ground state atomic hydrogen. Let $\hbar\mathbf{K}_i, \hbar\mathbf{K}_f \equiv \mu\mathbf{v}_i, \mu\mathbf{v}_f$ define the initial and final momenta of the incident particle in the center of mass system, where μ is the reduced mass of the incident particle–hydrogen atom pair, and $\mathbf{v}_i$ and $\mathbf{v}_f$ are, respectively, the initial and final relative velocities of the colliding particles. Furthermore, define the momentum transfer vector $\mathbf{q}$ by

$$\mathbf{q} \equiv \mathbf{K}_i - \mathbf{K}_f.$$

The Glauber approximation to the scattering amplitude, in the center of mass system, corresponding to elastic scattering from atomic hydrogen without exchange has been obtained in closed form by Thomas and Gerjuoy.[6,7] They found, ignoring the effects of particle indistinguishability and spin, that the amplitude is given by

$$F(1s \to 1s; \mathbf{q}) = -iK_i\left(\frac{2}{a_0}\right)^3 \frac{1}{4}\left(\frac{\partial}{\partial\lambda} I_0(\lambda, q)\right)\Big|_{\lambda=2/a_0}, \quad (1)$$

where

$$I_0(\lambda, q) = -4i\eta\,\Gamma(1+i\eta)\Gamma(1-i\eta)\lambda^{-2-2i\eta}\,q^{-2+2i\eta} \times {}_2F_1(-i\eta+1, -i\eta+1; 1; (-\lambda^2/q^2)). \quad (2)$$

In Eqs. (1) and (2), $\eta \equiv -Ze^2/\hbar v_i$, where v_i is the speed of the incident particle, a_0 is the Bohr radius, and the ${}_2F_1$ is the usual hypergeometric function. We also remark that the amplitude was evaluated by taking the Glauber path integral along the direction $\hat{\xi}$ which is coplanar with $\mathbf{K}_i$ and $\mathbf{K}_f$, but is perpendicular to the momentum transfer $\mathbf{q}$ at each $\mathbf{K}_f$ for given $\mathbf{K}_i$.

The center of mass system differential cross section for the elastic scattering of the incident particle by ground state atomic hydrogen into the solid angle $d\Omega$ about the direction $\mathbf{n}_f = (K_f^{-1})\mathbf{K}_f$ is obtained from the scattering amplitude in the usual way, namely,

$$\frac{d\sigma(1s \to 1s; K_i)}{d\Omega} = \frac{K_f}{K_i}\,|F(1s \to 1s; \mathbf{q})|^2. \quad (3)$$

Of course, for elastic scattering $K_f = K_i$. In general, the Glauber amplitude for scattering from initial state i to final state f depends upon the choice of quantization axis. However, since the initial and final states are s states in the case of present interest, at any given $\mathbf{q}$ (i.e., at any given scattering angle) the quantity $|F(1s \to 1s; \mathbf{q})|^2$ in (3) will be independent of the choice of quantization axis. Therefore, the integrated cross section now may be written

$$\sigma(1s \to 1s; K_i) = \int |F(1s \to 1s; q)|^2 d\Omega, \tag{4}$$

where $F(1s \to 1s; q)$ is as given by Eqs. (1) and (2), and where the integration in Eq. (4) is over all directions of the outgoing particle. Taking advantage of the fact that

$$q^2 = 2K_i^2(1 - \cos\theta),$$

where θ is the scattering angle in the center of mass system, the integrated cross section becomes

$$\sigma(1s \to 1s; K_i) = \frac{1}{K_i^2} \int_0^{2K_i} q\, dq \int_0^{2\pi} d\phi_q |F(1s \to 1s; \mathbf{q})|^2. \tag{5}$$

Since $|F(1s \to 1s; \mathbf{q})|^2$ is azimuthally symmetric, Eq. (5) reduces to

$$\sigma(1s \to 1s; K_i) = \frac{2\pi}{K_i^2} \int_0^{2K_i} q\, dq\, |F(1s \to 1s; \mathbf{q})|^2. \tag{6}$$

B. Simplification of the amplitude

When Eq. (2) is used in Eq. (1) to generate the elastic scattering amplitude, one obtains an expression for $F(1s \to 1s; q)$ which involves a sum of two hypergeometric functions.[6] Using this result for the amplitude in Eq. (6) leads to an expression for the integrated cross section containing definite integrals of products of hypergeometric functions; of a sort which usually are not reduced easily to closed form. However, for the integrals arising in the present paper, it turns out that reduction to closed form is possible by exploiting known relations between the hypergeometric function appearing in Eq. (2) and the Legendre functions.

In particular we first note the relation[8]

$$\begin{aligned} &{}_2F_1(a, b; a - b + 1; x) \\ &\quad = \Gamma(a - b + 1)(1 - x)^{-b}(-x)^{(b-a)/2} P^{b-a}_{-b}\left(\frac{1 + x}{1 - x}\right) \end{aligned} \tag{7}$$

which is valid provided x is real and $-\infty < x < 0$.

Via Eq. (7), the hypergeometric function appearing in Eq. (2) can be written as

$$\begin{aligned} &{}_2F_1\left(1 - i\eta, 1 - i\eta; 1; -\frac{\lambda^2}{q^2}\right) \\ &\quad = \left(1 + \frac{\lambda^2}{q^2}\right)^{-1+i\eta} P_{-1+i\eta}\left(\frac{1 - (\lambda^2/q^2)}{1 + (\lambda^2/q^2)}\right), \end{aligned} \tag{8}$$

where $y \equiv [1 - (\lambda^2/q^2)]/[1 + (\lambda^2/q^2)]$ satisfies $-1 \le y \le 1$ for all physical q^2. The generating function $I_0(\lambda, q)$ then becomes

$$\begin{aligned} I_0(\lambda, q) &= -4i\eta\, \Gamma(1 + i\eta)\Gamma(1 - i\eta) q^{-4} (\lambda^2/q^2)^{-1-i\eta} \\ &\quad \times (1 + \lambda^2/q^2)^{-1+i\eta} P_{-1+i\eta}\left(\frac{1 - (\lambda^2/q^2)}{1 + (\lambda^2/q^2)}\right) \qquad (9a) \\ &\equiv -4i\eta\Gamma(1 + i\eta)\Gamma(1 - i\eta) q^{-4} M_0(\eta; \lambda^2/q^2). \qquad (9b) \end{aligned}$$

Equations (9a) and (9b) together define the quantity $M_0(\eta; (\lambda^2/q^2))$. Then, from Eq. (1),

$$\begin{aligned} F(1s \to 1s; q) &= iK_i \left(\frac{2}{a_0}\right)^3 (i\eta)\Gamma(1 + i\eta)\Gamma(1 - i\eta) q^{-4} \\ &\quad \times \left[\frac{\partial}{\partial\lambda} M_0\left(\eta; \frac{\lambda^2}{q^2}\right)\right]\Bigg|_{\lambda = 2/a_0} \qquad (10a) \\ &= iK_i \left(\frac{2}{a_0}\right)^3 (i\eta)\Gamma(1 + i\eta)\Gamma(1 - i\eta) q^{-5} \\ &\quad \times 2\left(x^{1/2} \frac{\partial}{\partial x} M_0(\eta; x)\right)\Bigg|_{x = 4/(a_0^2 q^2)} \qquad (10b) \end{aligned}$$

where $x \equiv \lambda^2/q^2$ and q^2 is fixed. But from Eqs. (9),

$$\begin{aligned} \frac{\partial}{\partial x} M_0(\eta; x) &= \frac{\partial}{\partial x}\left[x^{-1-i\eta}(1 + x)^{-1+i\eta} P_{-1+i\eta}\left(\frac{1 - x}{1 + x}\right)\right] \\ &= x^{-1-i\eta}(1 + x)^{-1+i\eta} \frac{\partial}{\partial x} P_{-1+i\eta}\left(\frac{1 - x}{1 + x}\right) \\ &\quad - x^{-2-i\eta}(1 + x)^{-2+i\eta}[(1 + i\eta) + 2x] P_{-1+i\eta}\left(\frac{1 - x}{1 + x}\right). \end{aligned} \tag{11}$$

If we now let $y \equiv (1 - x)/(1 + x)$, then $1 - y^2 = 4x/(1 + x)^2$ and

$$\begin{aligned} \frac{\partial}{\partial x} P_\nu\left(\frac{1 - x}{1 + x}\right) &= -\frac{2}{(1 + x)^2} \frac{\partial}{\partial y} P_\nu(y) \\ &= -\frac{1}{2x}(1 - y^2) \frac{\partial}{\partial y} P_\nu(y). \end{aligned}$$

However,[9]

$$(1 - y^2) \frac{d}{dy} P^\mu_\nu(y) = (\nu + \mu) P^\mu_{\nu-1}(y) - \nu y P^\mu_\nu(y).$$

Consequently,

$$\frac{\partial}{\partial x} P_{-1+i\eta}\left(\frac{1 - x}{1 + x}\right) = -\frac{1}{2x}(-1 + i\eta)[P_{-2+i\eta}(y) - y P_{-1+i\eta}(y)];$$

and therefore

$$\begin{aligned} \frac{\partial}{\partial x} M_0(\eta; x) &= -x^{-2-i\eta}(1 + x)^{-2+i\eta}\{(1 + x)\tfrac{1}{2}(-1 + i\eta) \\ &\quad \times [P_{-2+i\eta}(y) - y P_{-1+i\eta}(y)] + [1 + i\eta + 2x] P_{-1+i\eta}(y)\} \\ &= -\tfrac{1}{2} x^{-2-i\eta}(1 + x)^{-1+i\eta}[(-1 + i\eta) P_{-2+i\eta}(y) \\ &\quad + (3 + i\eta) P_{-1+i\eta}(y)]. \end{aligned} \tag{12}$$

Inserting Eq. (12) into (10b), and using the relation[10]

$$P^\mu_{-\nu-1}(y) = P^\mu_\nu(y),$$

we finally obtain the following expression for the elastic scattering amplitude:

$$\begin{aligned} &F(1s \to 1s; q) = -iK_i(i\eta)\Gamma(1 + i\eta)\Gamma(1 - i\eta)\tfrac{1}{4}(a_0^2) \\ &\times \{x^{1-i\eta}(1 + x)^{-1+i\eta}[(-1 + i\eta) P_{1-i\eta}(y) + (3 + i\eta) P_{-i\eta}(y)]\}. \end{aligned} \tag{13}$$

where $x \equiv 4/(a_0^2 q^2)$ and $y \equiv (1 - x)/(1 + x)$. It is the structure of Eq. (13) which permits us to evaluate the cross section integral [Eq. (6)] in closed form.

C. Evaluation of the cross section integral

We now use the result of the preceding subsection to evaluate the elastic scattering cross section in closed

form. Since $[\Gamma(1+i\eta)\Gamma(1-i\eta)]$ is real, using Eq. (13) in Eq. (6) leads to

$$\sigma(1s\to 1s;K_i)=(2\pi)\eta^2[\Gamma(1+i\eta)\Gamma(1-i\eta)]^2\left(\frac{a_0}{2}\right)^4 \times\int_0^{2K_i} q\,dq\left(\frac{x}{1+x}\right)^2\left|[(-1+i\eta)P_{1-i\eta}(y)+(3+i\eta)P_{-i\eta}(y)]\right|^2, \tag{14}$$

where x and y are as previously defined. However,

$$y=\frac{1-x}{1+x}=\frac{a_0^2q^2-4}{a_0^2q^2+4}$$

so that $[a_0^2x^2/(1+x)^2]\,qdq=dy$. Hence

$$\sigma(1s\to 1s;K_i)=\pi a_0^2\eta^2[\Gamma(1+i\eta)\Gamma(1-i\eta)]^2 2^{-3} \times\int_{-1}^{\beta} dy\left|[(-1+i\eta)P_{1-i\eta}(y)+(3+i\eta)P_{-i\eta}(y)]\right|^2 \tag{15}$$

with the upper limit $\beta\equiv(a_0^2K_i^2-1)/(a_0^2K_i^2+1)$.

In order to expand the integrand of Eq. (15) we require the quantity $[P_\nu(y)]^*$. We first note when $-1<y<1$ and m is a nonnegative integer, that[11]

$$P_\nu^{-m}(y)=\left(\frac{1-y}{1+y}\right)^{m/2}\frac{1}{\Gamma(1+m)}\,{}_2F_1\left(-\nu,\nu+1;1+m;\frac{1-y}{2}\right); \tag{16a}$$

moreover[10]

$$P_\nu^m(y)=(-1)^m[\Gamma(\nu+1+m)/\Gamma(\nu+1-m)]P_\nu^{-m}(y), \tag{16b}$$

so that

$$P_\nu^m(y)=(-1)^m\frac{\Gamma(\nu+1+m)}{\Gamma(\nu+1-m)\Gamma(1+m)}\left(\frac{1-y}{1+y}\right)^{m/2} \times{}_2F_1\left(-\nu,\nu+1;1+m;\frac{1-y}{2}\right). \tag{16c}$$

Therefore, since y is real it follows from (16a) and (16c) that

$$[P_\nu^m(y)]^*=P_{\nu^*}^m(y) \tag{17}$$

for arbitrary integers m. Therefore, Eq. (15) becomes

$$\begin{aligned}\sigma(1s\to 1s;K_i)&=2^{-3}\eta^2[\Gamma(1+i\eta)\Gamma(1-i\eta)]^2\pi a_0^2\\ &\times\int_{-1}^{\beta}dy[-(1-i\eta)P_{1-i\eta}(y)+(3+i\eta)P_{-i\eta}(y)]\\ &\times[-(1+i\eta)P_{1+i\eta}(y)+(3-i\eta)P_{i\eta}(y)] \qquad (18)\\ &=2^{-3}\eta^2[\Gamma(1+i\eta)\Gamma(1-i\eta)]^2\pi a_0^2\\ &\times\int_{-1}^{\beta}dy[(1+\eta^2)P_{1-i\eta}P_{1+i\eta}-(3+i\eta)(1+i\eta)P_{-i\eta}P_{1+i\eta}\\ &-(1-i\eta)(3-i\eta)P_{1-i\eta}P_{i\eta}+(9+\eta^2)P_{-i\eta}P_{i\eta}]. \qquad (19)\end{aligned}$$

Now, however, definite integrals of products of Legendre functions $P_\nu P_\sigma$ can be evaluated in closed form.[12] In particular when $-1<a\le y\le b<1$,

$$\int_a^b P_\nu(y)P_\sigma(y)dy=[(\nu-\sigma)(\nu+\sigma+1)]^{-1}(1-y^2) \times\mathrm{Wr}[P_\nu(y),P_\sigma(y)]\Big|_a^b \tag{20a}$$

$$=[(\nu-\sigma)(\nu+\sigma+1)]^{-1}(1-y^2)^{1/2}[P_\sigma P_\nu{}^1-P_\nu P_\sigma{}^1]\Big|_a^b, \tag{20b}$$

where $\mathrm{Wr}[f,g]$ is the usual Wronskian determinant of f and g. For convenience we also define the quantity

$$\mathfrak{W}(\nu,\sigma;y)\equiv(1-y^2)\,\mathrm{Wr}[P_\nu,P_\sigma] \tag{21a}$$

$$=(1-y^2)^{1/2}[P_\sigma P_\nu{}^1-P_\nu P_\sigma{}^1]. \tag{21b}$$

Therefore, via Eqs. (20) and (21), Eq. (19) becomes

$$\begin{aligned}\sigma(1s\to 1s;K_i)&=2^{-3}\pi a_0{}^2\eta^2[\Gamma(1+i\eta)\Gamma(1-i\eta)]^2\\ &\times\left(-\frac{(1+\eta^2)}{6i\eta}\mathfrak{W}(1-i\eta,1+i\eta;y)-\frac{(9+\eta^2)}{2i\eta}\mathfrak{W}(-i\eta,+i\eta;y)\right.\\ &+\frac{(3+i\eta)(1+i\eta)}{2(1+2i\eta)}\mathfrak{W}(-i\eta,1+i\eta;y)\\ &\left.-\frac{(1-i\eta)(3-i\eta)}{2(1-2i\eta)}\mathfrak{W}(1-i\eta,i\eta;y)\right)\Big|_{y=-1}^{\beta}. \qquad (22)\end{aligned}$$

Equation (22) can be simplified immediately. Since Eq. (17) holds, it is clear from Eq. (21b) that

$$\begin{aligned}\mathfrak{W}(\nu,\nu^*;y)&=(1-y^2)^{1/2}[P_{\nu^*}P_\nu{}^1-\{P_{\nu^*}P_\nu{}^1\}^*]\\ &=2i(1-y^2)^{1/2}\,\mathrm{Im}[P_{\nu^*}P_\nu{}^1]. \qquad (23a)\end{aligned}$$

Moreover, it is also true from Eq. (17) and (21b) that

$$\begin{aligned}\mathfrak{W}(1-i\eta,i\eta;y)&=-\mathfrak{W}(i\eta,1-i\eta;y)\\ &=-\mathfrak{W}^*(-i\eta,1+i\eta;y). \qquad (23b)\end{aligned}$$

Hence Eq. (22) becomes

$$\begin{aligned}\sigma(1s\to 1s;K_i)&=2^{-3}\pi a_0^2\eta^2[\Gamma(1+i\eta)\Gamma(1-i\eta)]^2\\ &\times\left[-\frac{(1+\eta^2)}{3\eta}(1-y^2)^{1/2}\,\mathrm{Im}[P_{1+i\eta}P_{1-i\eta}^1]\right.\\ &-\frac{(9+\eta^2)}{\eta}(1-y^2)^{1/2}\,\mathrm{Im}[P_{i\eta}P_{-i\eta}^1]\\ &\left.+\mathrm{Re}\left(\frac{(3+i\eta)(1+i\eta)}{1+2i\eta}\mathfrak{W}(-i\eta,1+i\eta;y)\right)\right]\Big|_{-1}^{\beta}. \qquad (24)\end{aligned}$$

The quantity $\mathfrak{W}(-i\eta,1+i\eta;y)$ can be further simplified. From Eq. (21a)

$$\mathfrak{W}(-i\eta,1-i\eta;y)=(1-y^2)\,\mathrm{Wr}[P_{-i\eta},P_{1+i\eta}].$$

But, as we noted previously, $P_{1+i\eta}=P_{-2-i\eta}$. However,[9]

$$(2\nu+1)yP_\nu^\mu(y)=(\nu-\mu+1)P_{\nu+1}^\mu(y)+(\nu+\mu)P_{\nu-1}^\mu(y),$$

so that

$$\begin{aligned}&P_{-2-i\eta}(y)\\ &\quad=(-1-i\eta)^{-1}[(-2i\eta-1)yP_{-1-i\eta}(y)-(-i\eta)P_{-i\eta}(y)].\end{aligned}$$

Therefore

$$\begin{aligned}\mathfrak{W}(-i\eta,1+i\eta;y)&=(1-y^2)\\ &\times\mathrm{Wr}\left[P_{-i\eta},\frac{1+2i\eta}{1+i\eta}yP_{-1-i\eta}-\frac{i\eta}{1+i\eta}P_{-i\eta}\right]\\ &=(1-y^2)\frac{1+2i\eta}{1+i\eta}\,\mathrm{Wr}[P_{-i\eta}]\\ &=(1-y^2)\frac{1+2i\eta}{1+i\eta}\{P_{-i\eta}P_{-1-i\eta}+y\,\mathrm{Wr}[P_{-i\eta},P_{-1-i\eta}]\}.\end{aligned}$$

However $P_{-1-i\eta} = P_{i\eta}$, so that, again using Eq. (21a) we have

$$\mathcal{W}(-i\eta, 1+i\eta; y) = \frac{1+2i\eta}{1+i\eta}\{(1-y^2)P_{-i\eta}P_{+i\eta} + y\mathcal{W}(-i\eta, +i\eta; y)\}$$
$$= \frac{1+2i\eta}{1+i\eta}\{(1-y^2)|P_{+i\eta}|^2 + 2iy(1-y^2)^{1/2}\,\mathrm{Im}[P_{i\eta}P^1_{-i\eta}]\}. \quad (25)$$

We therefore find, since y is real, that

$$\mathrm{Re}\left(\frac{(3+i\eta)(1+i\eta)}{1+2i\eta}\mathcal{W}(-i\eta, 1+i\eta; y)\right) = 3(1-y^2)|P_{i\eta}|^2 - 2\eta\, y(1-y^2)^{1/2}\,\mathrm{Im}[P_{i\eta}P^1_{-i\eta}],$$

and Eq. (24) reduces to

$$\sigma(1s\to 1s; K_i) = 2^{-3}\pi a_0^2\eta^2[\Gamma(1+i\eta)\Gamma(1-i\eta)]^2 \times \left\{-\frac{(1+\eta^2)}{3\eta}(1-y^2)^{1/2}\,\mathrm{Im}[P_{1+i\eta}P^1_{1-i\eta}] - \frac{[9+\eta^2(1+2y)]}{\eta}(1-y^2)^{1/2}\,\mathrm{Im}[P_{i\eta}P^1_{-i\eta}] + 3(1-y^2)|P_{i\eta}|^2\right\}\Big|_{-1}^{\beta}. \quad (26)$$

The term in braces in Eq. (26) can be evaluated without any real difficulty at the lower limit $y = -1$ (for fixed η), although it is necessary to be careful. Using Eq. (16c) the product $(1-y^2)^{1/2}\,P_\nu P^1_{\nu^*}$ becomes

$$(1-y^2)^{1/2}P_\nu P^1_{\nu^*} = -(1-y^2)^{1/2}\,{}_2F_1\left(-\nu, 1+\nu; 1; \frac{1-y}{2}\right) \times \frac{\Gamma(\nu^*+2)}{\Gamma(\nu^*)}\left(\frac{1-y}{1+y}\right)^{1/2}{}_2F_1\left(-\nu^*, 1+\nu^*; 2; \frac{1-y}{2}\right). \quad (27)$$

Near $y = -1$ the hypergeometric functions appearing in Eq. (27) are given by their standard analytic continuations. In particular,[13]

$${}_2F_1(-\nu, 1+\nu; 1; \tfrac{1}{2}(1-y) \sim [\Gamma(-\nu)\Gamma(1+\nu)]^{-1}[2\psi(1) - \psi(-\nu) - \psi(1+\nu) - \ln[\tfrac{1}{2}(1+y)] + O\{\tfrac{1}{2}(1+y)\ln[\tfrac{1}{2}(1+y)]\}] \quad (28a)$$

where the psi function has its usual definition[14] $\psi(z) = (d/dz)\ln\Gamma(z)$; and[13]

$${}_2F_1(-\nu^*, 1+\nu^*; 2; \tfrac{1}{2}(1-y) \sim [\Gamma(-\nu^*+1)\Gamma(\nu^*+2)]^{-1} \times O\{(\tfrac{1}{2}1+y)\ln[\tfrac{1}{2}(1+y)]\}. \quad (28b)$$

Combining Eqs. (28) with (27) we find that for $y \sim -1$

$$(1-y^2)^{1/2}P_\nu P^1_{\nu^*} \sim \frac{-(1-y)}{\Gamma(-\nu)\Gamma(1+\nu)\Gamma(\nu^*)\Gamma(1-\nu^*)} \times \left[2\psi(1) - \psi(-\nu) - \psi(1+\nu) - \ln\left(\frac{1+y}{2}\right)\right]$$
$$= \frac{(1-y)}{\Gamma(\nu)\Gamma(1-\nu)\Gamma(\nu^*)\Gamma(1-\nu^*)} \times \left[2\psi(1) - \psi(-\nu) - \psi(1+\nu) - \ln\left(\frac{1+y}{2}\right)\right]. \quad (29)$$

Although Eq. (29) contains a logarithmic divergence at $y = -1$, this poses no problem since in Eq. (26) we only require the imaginary part of Eq. (29) at $y = -1$. In particular, since $\Gamma(\nu)\Gamma(1-\nu)\Gamma(\nu^*)\Gamma(1-\nu^*)$ is real,

$$\lim_{y\to -1}\mathrm{Im}[(1-y^2)^{1/2}P_\nu P^1_{\nu^*}] = \frac{-2}{|\Gamma(\nu)\Gamma(1-\nu)|^2} \times \mathrm{Im}[\psi(-\nu) + \psi(1+\nu)]. \quad (30)$$

In Eq. (26) we are required to evaluate Eq. (30) for $\nu = i\eta$ and $\nu = 1 + i\eta$; we use the relation[14]

$$\psi(z+1) = \psi(z) + (1/z)$$

and its corollary

$$\psi(z-1) = \psi(z) - [1/(z-1)].$$

In particular, then, when $\nu = i\eta$

$$\mathrm{Im}[(1-y^2)^{1/2}P_{i\eta}P^1_{-i\eta}]\Big|_{y=-1} = \frac{-2}{|\Gamma(i\eta)\Gamma(1-i\eta)|^2} \times \mathrm{Im}[\psi(-i\eta) + \psi(1+i\eta)].$$

But

$$\psi(-i\eta) + \psi(1+i\eta) = \psi(-i\eta) + \psi(i\eta) + (1/i\eta).$$

Since[14]

$$\psi^*(z) = \psi(z^*), \qquad \psi(-i\eta) + \psi(i\eta) = 2\,\mathrm{Re}\psi(i\eta).$$

Therefore

$$\mathrm{Im}[(1-y^2)^{1/2}P_{i\eta}P^1_{-i\eta}]\Big|_{y=-1} = \frac{2}{|\Gamma(i\eta)\Gamma(1-i\eta)|^2}\left(\frac{1}{\eta}\right). \quad (31)$$

When $\nu = 1 + i\eta$, Eq. (31) becomes

$$\mathrm{Im}[(1-y^2)^{1/2}P_{1+i\eta}P^1_{1-i\eta}]\Big|_{y=-1} = \frac{-2}{|\Gamma(1+i\eta)\Gamma(-i\eta)|^2} \times \mathrm{Im}[\psi(-1-i\eta) + \psi(2+i\eta)]$$
$$= \frac{-2}{|\Gamma(1-i\eta)\Gamma(i\eta)|^2}\,\mathrm{Im}\left(\psi(-i\eta) + \frac{1}{1+i\eta} + \psi(i\eta) + \frac{1}{1+i\eta} + \frac{1}{i\eta}\right),$$

and finally, after some manipulation,

$$\mathrm{Im}[(1-y^2)^{1/2}P_{1+i\eta}P^1_{1-i\eta}]\Big|_{y=-1} = \frac{2}{|\Gamma(i\eta)\Gamma(1-i\eta)|^2}\,\frac{(1+3\eta^2)}{(1+\eta^2)\eta}. \quad (32)$$

The quantity $(1-y^2)|P_{i\eta}|^2$ appearing in Eq. (26) can be evaluated at $y = -1$ in a similar fashion. Again using Eq. (16c),

$$(1-y^2)|P_{i\eta}|^2 = (1-y^2)|{}_2F_1(-i\eta, 1+i\eta; 1; \tfrac{1}{2}(1-y))|^2.$$

When $y \sim -1$, we again apply Eq. (28a) and find that

$$(1-y^2)|P_{i\eta}|^2 \sim (1+y)(1-y)|[\Gamma(-i\eta)\Gamma(1+i\eta)]^{-1} \times \{2\psi(1) - \psi(-i\eta) - \psi(1+i\eta) - \ln[\tfrac{1}{2}(1+y)]\}|^2;$$

therefore

$$\lim_{y\to -1}(1-y^2)|P_{i\eta}|^2 = 0. \quad (33)$$

We now are in a position to evaluate the term in braces

in Eq. (26) at the lower limit $y = -1$; we denote this term by $\{\}|_{-1}$. Using Eq. (31), (32), and (33) we find that at $y = -1$

$$\{\}|_{-1} = \frac{2}{\eta^2 |\Gamma(i\eta)\Gamma(1-i\eta)|^2}\left(-\frac{1+3\eta^2}{3} - 9 + \eta^2\right)$$
$$= -\frac{2}{\eta^2 |\Gamma(i\eta)\Gamma(1-i\eta)|^2}\left(\frac{28}{3}\right). \quad (34)$$

However, $\eta^2 |\Gamma(i\eta)\Gamma(1-i\eta)|^2 = |\Gamma(1+i\eta)\Gamma(1-i\eta)|^2$ and $\{\}|^{\beta}_{-1} = \{\}|_{\beta} - \{\}|_{-1}$.

Therefore inserting Eq. (34) into Eq. (26) and collecting terms we find that (neglecting exchange) the Glauber elastic scattering cross section for arbitrary structureless charged particles incident upon ground state atomic hydrogen is given by

$$\sigma(1s \to 1s; K_i) = \pi a_0^2 \eta^2 \Big[\tfrac{7}{3} + 2^{-3}[\Gamma(1+i\eta)\Gamma(1-i\eta)]^2$$
$$\times \Big(3(1-\beta^2)|P_{i\eta}(\beta)|^2 - \frac{(1-\beta^2)^{1/2}}{\eta}$$
$$\times \frac{(1+\eta^2)}{3}\,\mathrm{Im}[P_{1+i\eta}(\beta)P^1_{1-i\eta}(\beta)]$$
$$- \frac{(1-\beta^2)^{1/2}}{\eta}[9 + \eta^2(1+2\beta)]\mathrm{Im}[P_{i\eta}(\beta)P^1_{-i\eta}(\beta)]\Big)\Big], \quad (35)$$

where $\beta \equiv (a_0^2 K_i^2 - 1)/(a_0^2 K_i^2 + 1)$.

III. THE ASYMPTOTIC FORM OF $\sigma(1s \to 1s; K_i)$ AT LARGE K_i

Equation (35) above is our desired closed form result for any K_i. However, it is desirable to have the asymptotic form of (35) for large K_i, the situation so often encountered in practice. For the purpose of obtaining this asymptotic form, it is convenient to introduce atomic units. We also stress at this time that $K_i = \mu_i v$ *large does not mean* that $\eta \equiv -Z/v_i$ is necessarily small.

The Legendre functions $P^m_\nu(\beta)$ appearing in Eq. (35) are evaluated via Eq. (16c) using their equivalent representation in terms of the hypergeometric functions. Since $\beta = (K_i^2 - 1)/(K_i^2 + 1)$, when $y = \beta$ in Eq. (16c), the argument of the hypergeometric function becomes $(1-\beta)/2 = (1+K_i^2)^{-1}$. When K_i is large, therefore, to obtain the asymptotic expansion of $[(1-\beta^2)^{m/2}P^m_\nu(\beta)]$ we simply expand the hypergeometric functions in (16c) in the usual series, retaining only the first few terms. In particular,[15] if $K_i \gg 1$ and $\chi \equiv (1+K_i^2)^{-1}$,

$$P_\nu(\beta) = {}_2F_1(-\nu, 1+\nu; 1; \chi) \sim 1 + (-\nu)(\nu+1)\chi$$
$$+ \tfrac{1}{4}(-\nu)(-\nu+1)(\nu+1)(\nu+2)\chi^2 \cdots, \quad (36a)$$

and

$$(1-\beta^2)^{1/2}P^1_\nu(\beta) = -\nu(\nu+1)(1-\beta)\,{}_2F_1(-\nu, \nu+1; 2; \chi)$$
$$\sim -\nu(\nu+1)2\chi\{1 + \tfrac{1}{2}(-\nu)(\nu+1)\chi$$
$$+ \tfrac{1}{12}(-\nu)(-\nu+1)(\nu+1)(\nu+2)\chi^2 + \cdots\}. \quad (36b)$$

Therefore, using Eqs. (36),

$$(1-\beta^2)^{1/2}P_\nu(\beta)P^1_{\nu^*}(\beta) \sim -\nu^*(1+\nu^*)2\chi\{1 - \nu(\nu+1)\chi$$
$$- \tfrac{1}{4}\nu(1-\nu)(\nu+1)(\nu+2)\chi^2 + \cdots\}$$
$$\times \{1 - \tfrac{1}{2}\nu^*(1+\nu^*)\chi - \tfrac{1}{12}\nu^*(1+\nu^*)(1-\nu^*)$$
$$\times (2+\nu^*)\chi^2 + \cdots\}$$
$$= -\nu^*(1+\nu^*)2\chi\{1 - \nu(\nu+1)\chi - \tfrac{1}{2}\nu^*(1+\nu^*)\chi$$
$$+ \tfrac{1}{2}|\nu(1+\nu)|^2\chi^2 - \tfrac{1}{4}\nu(1+\nu)(1-\nu)(\nu+2)\chi^2$$
$$- \tfrac{1}{12}\nu^*(1+\nu^*)(1-\nu^*)(2+\nu^*)\chi^2 + O(\chi^3)\}. \quad (37)$$

In Eq. (35) we require the imaginary part of Eq. (37). Since $(1-\nu)(\nu+2) = 2 - \nu(\nu+1)$, and χ is real we find that

$$\mathrm{Im}(1-\beta^2)^{1/2}P_\nu P^1_{\nu^*} \sim 2\chi\,\mathrm{Im}\{-\nu^*(1+\nu^*) + \tfrac{1}{2}[\nu^*(1+\nu^*)]^2\chi$$
$$- \tfrac{1}{2}|\nu(1+\nu)|^2\nu^*(1+\nu^*)\chi^2 - \tfrac{1}{4}|\nu(1+\nu)|^2\nu(\nu+1)\chi^2$$
$$+ \tfrac{1}{12}[\nu^*(1+\nu^*)]^2[2 - \nu^*(1+\nu^*)]\chi^2\}. \quad (38)$$

We now explicitly construct the imaginary part of the term in braces in Eq. (38) via $\mathrm{Im}A = -\tfrac{1}{2}i(A - A^*)$. Equation (38) then becomes, after some algebraic manipulation,

$$\mathrm{Im}(1-\beta^2)^{1/2}P_\nu P^1_{\nu^*} \sim -i\chi[\nu(1+\nu) - \nu^*(1+\nu^*)]$$
$$\times \{1 - \tfrac{1}{2}[\nu(1+\nu) + \nu^*(1+\nu^*)]\chi$$
$$+ \tfrac{1}{2}|\nu(1+\nu)|^2\chi^2 - \tfrac{1}{6}[\nu(1+\nu) + \nu^*(1+\nu^*)]\chi^2$$
$$+ \tfrac{1}{12}[\nu(1+\nu) - \nu^*(1+\nu^*)]^2\chi^2 + O(\chi^3)\}. \quad (39)$$

And therefore, to third order in χ,

$$\mathrm{Im}(1-\beta^2)^{1/2}P_\nu P^1_{\nu^*} \sim 2\chi\,\mathrm{Im}[\nu(1+\nu)]\{1 - \mathrm{Re}[\nu(1+\nu)]\chi$$
$$+ \tfrac{1}{2}|\nu(1+\nu)|^2\chi^2 - \tfrac{1}{3}[\mathrm{Re}[\nu(1+\nu)] + (\mathrm{Im}[\nu(1+\nu)])^2]\chi^2\}. \quad (40)$$

Hence, when $\nu = i\eta$,

$$\mathrm{Im}(1-\beta^2)^{1/2}P_{i\eta}P^1_{-i\eta} \sim 2\chi\eta\{1 + \eta^2\chi + \tfrac{1}{2}\eta^2(1+\eta^2)\chi^2\}; \quad (41a)$$

and, when $\nu = 1 + i\eta$,

$$\mathrm{Im}(1-\beta^2)^{1/2}P_{1+i\eta}P^1_{1-i\eta} \sim 2\chi(3\eta)\{1 - (2-\eta^2)\chi$$
$$+ \tfrac{1}{2}\chi^2[\eta^4 - \tfrac{1}{3}\eta^2 + \tfrac{8}{3}]\}. \quad (41b)$$

We still require, however, the asymptotic form of $|P_\nu|^2$ when K_i is large. We apply Eq. (36a), to obtain

$$(1-\beta)|P_\nu(\beta)|^2 = 2\chi\,|\,{}_2F_1(-\nu, 1+\nu; 1; \chi)|^2$$
$$\sim 2\chi[1 - \nu(1+\nu)\chi - \tfrac{1}{4}\nu(1+\nu)(1-\nu)(2+\nu)\chi^2 + \cdots]$$
$$\times [1 - \nu^*(1+\nu^*)\chi - \tfrac{1}{4}\nu^*(1+\nu^*)(1-\nu^*)(2+\nu^*)\chi^2 + \cdots].$$

Again, after some algebraic manipulation, we find that to order χ^3

$$(1-\beta)|P_\nu(\beta)|^2 \sim 2\chi\{1 - 2\,\mathrm{Re}[\nu(1+\nu)]\chi + |\nu(1+\nu)|^2\chi^2$$
$$- \mathrm{Re}[\nu(1+\nu)]\chi^2 + \tfrac{1}{2}\,\mathrm{Re}(\nu(1+\nu))^2\chi^2\}. \quad (42)$$

Hence, when $\nu = i\eta$,

$$(1-\beta)|P_{i\eta}|^2 \sim 2\chi\{1 + 2\eta^2\chi + \tfrac{3}{2}\chi^2\eta^2(\eta^2+1)\}. \quad (43)$$

The asymptotic form for $\sigma(1s \to 1s; K_i)$, when K_i is large now can be obtained directly from Eq. (35) using Eqs. (41) and (43) together with the fact that $\beta = 1 - 2\chi$. We ultimately find that, to third order in $\chi = (1+K_i^2)^{-1}$,

$\sigma(1s \to 1s; K_i)$ is given, in atomic units, by

$$\sigma(1s \to 1s; K_i) \sim \pi\eta^2\left\{\frac{7}{3} - |\Gamma(1+i\eta)\Gamma(1-i\eta)|^2\left[\frac{(1+\eta^2)}{1+K_i^2} + \frac{(1-\eta^2)^2}{(1+K_i^2)^2} + \frac{1}{(1+K_i^2)^3}\left(\frac{1}{3} + \frac{13}{6}\eta^2 - \frac{5}{3}\eta^4 + \frac{1}{2}\eta^6\right)\right]\right\}, \tag{44}$$

where $K_i = \mu v_i$ and $\eta \equiv -Z/v_i$ in (a.u.). Equation (44) is to be compared with the exact first Born result (again in a.u.)

$$\sigma^B(1s \to 1s; K_i) = \eta^2\pi\left(\frac{7}{3} - \frac{1}{(1+K_i^2)} - \frac{1}{(1+K_i^2)^2} - \frac{1}{3(1+K_i^2)^3}\right). \tag{45}$$

If $|\eta|$ is sufficiently small so that in Eq. (44) we may expand $|\Gamma(1+i\eta)\Gamma(1-i\eta)|^2$ in powers of η^2, then[16]

$$\Gamma(1+i\eta)\Gamma(1-i\eta) = \frac{\pi\eta}{\sinh\pi\eta} \sim 1 - \tfrac{1}{6}\pi^2\eta^2 + \cdots.$$

We see immediately that for small $|\eta|$ the asymptotic form of the Glauber elastic scattering cross section contains the exact Born cross section, plus terms which vanish as $\eta \to 0$. This, of course, is not unexpected since the first term in the expansion of the Glauber amplitude in powers of η is[2,17] just the first Born amplitude.

In fact, if v_i is so large that $K_i \gg 1$ and $\eta \ll 1$, then we can obtain from Eq. (44) the asymptotic expansion of $\sigma(1s \to 1s; K_i)$ in powers of v_i^{-2}. We find that when v_i is large in this sense, the elastic scattering cross section is given to order v_i^{-6} by

$$\sigma(1s \to 1s; K_i) \sim \pi\eta^2\left\{\frac{7}{3} - \frac{1}{K_i^2} + \frac{\eta^2}{K_i^2}\left(\frac{\pi^3}{3} - 1\right) + O(v_i^{-6})\right\}. \tag{46}$$

The equivalent Born asymptotic expansion is obtained by letting $K_i \to \infty$ in Eq. (45). We see immediately that the Born and Glauber elastic cross sections are equal to order v_i^{-4}. However, the Born cross section contains no term proportional to v_i^{-6}; indeed, to order v_i^{-6} the Glauber cross section is strictly greater than the Born.

We also observe that as long as $K_i \gg 1$ and $\eta \lesssim 1$, the terms involving $|\Gamma(1+i\eta)\Gamma(1-i\eta)|^2$ in Eq. (44) always are $\ll 1$, because $|\Gamma(1+i\eta)\Gamma(1-i\eta)|^2 \leq 1$ for all η. Therefore, for incident protons (where $\mu = M_p/2 \gg 1$ in atomic units) $\sigma(1s \to 1s)$ predicted by Eq. (44) will be essentially identical to $\sigma^B(1s \to 1s)$ from Eq. (45), for all $v_i \gtrsim 1$ in atomic units. Moreover, in this $v_i \gtrsim 1$ range for protons, it follows from Eqs. (44) or (45) that

$$\sigma(1s \to 1s; K_i) \cong \tfrac{7}{3}\pi\eta^2. \tag{47}$$

Equation (47) and the foregoing remarks—all inferred from our exact analytic treatment—now are seen to confirm the results obtained by Franco and Thomas[5] via numeral integration, and plotted as Fig. 1 of their paper. That the equality of the Born and Glauber elastic cross sections at $v_i > \sim$ is not *a priori* obvious is made manifest by the fact that the corresponding Born and Glauber angular distributions are markedly different; the equality is obtained only after integrating the differential cross sections over all angles.

Indeed, at $\eta \sim v_i \sim 1$, the Glauber and Born amplitudes are not expected to be the same, because in this range dropping the higher-order terms (in powers of η) in the expansion of the Glauber amplitude is not justified. Therefore, it is even more surprising that for incident electrons the exact Glauber formula (35) reduces to the simple Born high energy limit (47) at all incident electron energies ≥ 1 eV, as can be seen by comparing Fig. 1 of Franco[2] (again obtained by numerical integration) with Eq. (47). That in e^-–H(1s) elastic scattering Eq. (35) is well approximated by Eq. (47) at energies ≥ 1 eV has not been shown by us analytically at all energies; but it has been shown analytically by one of us (Thomas) that for electrons Eq. (47) is within a percent of Eq. (35) in the limit $v_i \to 0$, and at the value of v_i corresponding to $K_i = 1$. We add that for incident protons the analysis shows that Eqs. (35) and (47) are identical to extremely high accuracy at $v_i \to 0$ and $K_i = 1$. Thus, in view of the foregoing results, we are led to the remarkable conclusion that there is good reason to believe the Glauber predicted $\sigma(1s \to 1s; K_i)$ is well represented by the very simple Eq. (47) for all incident electron and proton energies.

*Supported in part by the Advanced Research Projects Agency under Contract No. DA-31-124-ARO-D-440 and by the National Aeronautics and Space Administration under Contract No. NGL-39-011-035.

[1]R. J. Glauber, in *Lectures in Theoretical Physics*, edited by W. E. Brittin *et al.* (Interscience, New York, 1959) Vol. I, p. 315.
[2]V. Franco, Phys. Rev. Lett. **20**, 209 (1968).
[3]H. Tai, P. J. Teubner, and R. H. Bassel, Phys. Rev. Lett. **22**, 1415 (1969); Phys. Rev. Lett. **23**, 453E (1969).
[4]H. Tai, R. H. Bassel, E. Gerjuoy and V. Franco, Phys. Rev. A **1**, 1819 (1970).
[5]V. Franco and B. K. Thomas, Phys. Rev. A **4**, 945 (1971).
[6]B. K. Thomas and E. Gerjuoy, J. Math. Phys. **12**, 1567 (1971).
[7]We take this opportunity to remark that Eq. (2) of Thomas and Gerjuoy (Ref. 6) contains a misprint; the right-hand side of that equation should be multiplied by a factor of -1. Their results, however, are consistent with the proper sign in this equation.
[8]M. Abramovits and I. A. Stegun, *Handbook of Mathematical Functions* (Natl. Bur. Stds., Washington, D. C., 1964), p. 562, Eq. (15.4.15).
[9]W. Magnus, F. Oberhettinger, and R. P. Soni, *Formulas and Theorems for the Special Functions of Mathematical Physics* (Springer-Verlag, New York, 1966) p. 171.
[10]Reference 9, p. 170.
[11]Reference 9, p. 166.
[12]Reference 9, p. 191.
[13]Reference 8, p. 559.
[14]Reference 8, p. 258.
[15]Reference 8, p. 556.
[16]Reference 8, p. 256.
[17]E. Gerjuoy, in *The Physics of Electronic and Atomic Collisions, VIIth International Conference, Amsterdam, July 1971, Invited Papers and Progress Reports*, edited by T. R. Govers and F. J. de Heer (North-Holland, Amsterdam, 1972), p. 247.

Quantum mechanics of the anharmonic oscillator*

Francis R. Halpern
Department of Physics, University of California, San Diego, La Jolla, California 92037
(Received 23 June 1972; revised manuscript received 15 September 1972)

The dependence of the energy levels and the value of the matrix elements of the momentum and position operators of a quantum anharmonic oscillator are studied as functions of the quantum number and the strength of the anharmonic term. The principal technique employed is the construction of a canonical transformation.

I. INTRODUCTION

There is no procedure for dealing systematically with even the simplest problems in one-dimensional quantum mechanics, e.g., an anharmonic oscillator. The textbook techniques, although elegant when they are applied to the specific problems for which they were designed, are inappropriate when they are applied in different situations.[1] There is nontheless a structure and logic to the way in which the eigenvalue problems of quantum mechanics resolve themselves. It is my purpose in treating the anharmonic oscillator to make a preliminary step in the development of a general theory of quantum mechanical problems. Although much of what I do is intuitive and ad hoc in character, I believe that the ideas presented here can be developed into a comprehensive set of procedures. In the present paper I will alternate between detailed calculations for the anharmonic oscillator and a discussion of an arbitrary potential.

Before beginning the exposition, it is worth commenting on the very old WKB method[2] and the very new summation techniques[3] for treating similar problems. In many ways the ideas that I develop are close to the spirit of the WKB method. I shall emphasize the importance of determining the wavefunctions correctly for large quantum numbers. The advantage that my approach has is that it is easier to relate to the usual procedures of quantum theory than are the WKB methods and that I can more readily generalize to problems in several dimensions. I treat the same class of problems that is currently being discussed by Padé series and Borel summation. The numerical results to be obtained by these methods is probably superior to that which I have attained. These techniques are, however, limited to dealing with a single level at a time, usually the ground state energy. They do not naturally yield the matrix elements of operators nor do they give insight into the dependence of an energy level on the coupling parameter.

The problem to be treated is specified by the Hamiltonian H

$$H = \tfrac{1}{2}(p^2 + x^2) + \lambda x^4. \tag{1}$$

It is natural at first to consider this problem as a perturbation λx^4 to the harmonic oscillator Hamiltonian H_0,

$$H_0 = \tfrac{1}{2}(p^2 + x^2).$$

If λ is small, the expansion of the eigenvalues and eigenfunctions of H in terms of those of H_0 with coefficients that are power series in λ is a standard procedure. It will be worthwhile to indicate why this attempt must fail. From the simplest point of view the perturbation is never uniformly small. If λx^2 is greater than one, then the perturbation is large, and there are always values of x for which λx^2 is larger than one. Alternatively the matrix elements of the perturbation are approximately equal to λn^2 while those of H_0 are approximately equal to n, where n is either the row or column index of the energy matrix. Again for a sufficiently large value of n, λn is greater than one and the perturbation will be larger than the unperturbed energy. Since in high orders of perturbation theory any state will be connected to states with large values of n, there is a strong suggestion that the perturbation series will diverge. One may also consider the behavior of the perturbation series if the sign of λ is reversed to again indicate that the series in λ is divergent. There is a final qualitative behavior that we will note when we study the asymptotic form of the wave functions.

The relation between the behavior of an operator for large x and the behavior of its matrix elements for large values of n is determined. An operator that behaves like x^η for large values of x will have matrix elements that behave like $n^{\eta/2}$ for large values of n. More generally if $A(x,p)$ is a function that behaves like $x^\eta p^\xi$ for large values of x and p, then its matrix elements will behave like $n^{(\eta+\xi)/2}$ for large values of n.[4]

These considerations suggest that, in order to separate a Hamiltonian H into an unperturbed portion H_0 and a perturbation V, it is essential that H and H_0 behave identically for large values of x and p or equivalently that the matrix elements of H and H_0 be identical for large values of n. The effect of such a procedure if it could be carried through would be to reduce the matrix that is to be diagonalized when H is expanded in the eigenfunctions of H_0 to a matrix whose only nonvanishing off-diagonal elements occur for small finite values of n. The diagonalization of this residual finite matrix can be regarded as a technical problem rather than a problem in principle.

The result suggested above is probably too ambitious. As a practicable first step I have been able to choose H_0 so that $(H/H_0) = O(1)$ for large values of x and p. This will be seen to imply that the off-diagonal matrix elements are not large compared to the diagonal ones for large values of n. The actual result is better than this.

It is useful and simple to consider the asymptotic behavior of the wavefunction $\Psi(x)$ for large values of x. This is a standard technique for mathematicians and physicists.[5] It is most easily carried through by writing $\Psi(x) = \exp(-ax^s)$. If this expression is substituted into the Schrödinger equation, the result is

$$-\tfrac{1}{2}s^2a^2x^{2s-2} + s(s-1)ax^{s-2} + \tfrac{1}{2}x^2 + \lambda x^4 = E.$$

In order that the leading terms cancel out, it is necessary that $s = 3$ and $9a^2 = 2\lambda$. Since the wavefunction

must vanish for $x = \pm\infty$, we see that the required asymptotic form is

$$\Psi(x) = \exp[-(2\lambda/a)^{1/2}|x|^3].$$

In a general situation it is no more difficult to determine the asymptotic behavior of a wavefunction; the difficult point is to construct a complete orthornormal set having the specified behavior. In the present case there is a slight additional complication because of the appearance of the "unnatural" function $|x|^3$ in the exponent. This makes the asymptotic form a poor choice even for so elementary a use as a trial ground state for a variational calculation. The dubious character of the perturbation procedure is again indicated by the difference in the asymptotic character of the wavefunctions. The harmonic oscillator wavefunctions are quadratic in x and the anharmonic oscillator functions are cubic in x. It is unreasonable to try to synthesize a completely different kind of function $\exp(-x^3)$ from $\exp(-x^2)$ rather than to start *ab initio* with functions having the correct asymptotic behavior.

One method for producing a new set of functions from s given set is to simply consider $\Psi[g(x)]$ instead of $\Psi(x)$. If $\Psi_n(x)$ is a complete orthonormal set of functions,

$$\int \Psi_m^*(x)\Psi_n(x)dx = \delta_{mn},$$

then $\Psi_m[g(x)][g'(x)]^{1/2}$ is also a complete orthonormal set:

$$\int \{\Psi_m^*[g(x)][g'(x)]^{1/2}\}\{\Psi_n[g(x)][g'(x)]^{1/2}\}dx = \delta_{mn}$$

as can easily be checked by the change of variables $y = g(x)$. If the $\Psi_n(x)$ are the harmonic oscillator wavefunctions and $g(x) = O(x^{3/2})$, we will have produced a set of functions that behaves appropriately for large x.

In the succeeding sections this procedure is shown to be a canonical or unitary transformation, and the transformed Hamiltonian is calculated. It is shown that to a certain extent the Hamiltonian is made diagonal in the leading terms in n. The type of canonical transformation required to carry the diagonalization further is considered.

II. THE CANONICAL TRANSFORMATION

The observation that the set of functions $\Psi_n(g)(g')^{1/2}$ is a complete orthonormal set of functions implies the existence of a unitary or canonical transformation connecting this set of functions with the set Ψ_n. The matrix element U_{ab} of this transformation is given by the integral

$$U_{ab} = \int \Psi_a^*(x)\Psi_b[g(x)][g'(x)]^{1/2}dx. \tag{2}$$

For most choices of g and Ψ this integral will be difficult to evaluate, while for many applications knowledge of the transformed operators $x' = U^+xU$ and $p' = U^+pU$ will be adequate and turns out to be quite simple to obtain.

The calculation of the transformed operators is made from the formulas

$$x'_{ab} = \int \Psi_a^*[g(x)][g'(x)]^{1/2}x\Psi_b[g(x)][g'(x)]^{1/2}dx$$

and

$$p'_{ab} = \int \Psi_a^*[g(x)][g'(x)]^{1/2}p\Psi_b[g(x)][g'(x)]^{1/2}dx.$$

The change of variable $y = g(x)$, where $x = f(y)$, gives the result that

$$x'_{ab} = \int \Psi_a^*(y)f(y)\Psi_b(y)dy,$$
$$p'_{ab} = \int \Psi_a^*(y)[f'(y)]^{-1/2}p[f'(y)]^{-1/2}\Psi_b(y)dy.$$

From the form of these formulas one may read off directly the transformation formulas

$$x' = f(x), \quad p' = [f'(x)]^{-1/2}p[f'(x)]^{-1/2}. \tag{3}$$

By direct calculation of the commutator of p' and x' it may be seen that $[p',x'] = -i$ so that the transformation is canonical. If the symbols in (3) are interpreted as ordinary variables rather than as operators, the Poisson bracket of the variables p' and x' may be calculated with respect to the variables p and x. This Poisson bracket is unity so that the transformation (3) is also a classical canonical transformation. The generating function F of this classical canonical transformation is $F(p',x) = p'f(x)$. In general classical canonical transformations are not quantum canonical transformations and vice versa.[6] The unitary operator U whose matrix elements are defined by (2) has the property

$$U^+xU = x' = f(x), \quad U^+pU = p' = [f'(x)]^{-1/2}p[f'(x)]^{-1/2}.$$

The formulas (2) and (3) permit the transformation of the Hamiltonian $H(x,p)$ by the unitary operator U to be carried out, and this gives the result

$$U^+H(x,p)U = H(x',p') = H\{f(x),[f'(x)]^{-1/2}p[f'(x)]^{-1/2}\}. \tag{4}$$

In general the Hamiltonian has the form $H = \frac{1}{2}p^2 + V(x)$. The transformation of the potential energy is $V[f(x)]$ while the transformation of the kinetic energy is given by

$$\tfrac{1}{2}p'^2 = \tfrac{1}{2}[f'(x)]^{-2}p^2 + i[f'(x)]^{-3}f''(x)p + \tfrac{1}{4}[f'(x)]^{-3}f'''(x) - \tfrac{5}{8}[f'(x)]^{-4}[f''(x)]^2. \tag{5}$$

The next step is the determination of the complete orthonormal set Ψ and the function g that are appropriate for the anharmonic oscillator problem.

III. THE CHOICE OF THE CANONICAL TRANSFORMATION

The only complete orthonormal set appropriate to the problem of the anharmonic oscillator are the harmonic oscillator wave functions. In Sec. I it was noted that asymptotically the wavefunctions of the anharmonic oscillator behave like $\exp(-\alpha|x|^3)$. In order for this to be true, if we use the harmonic oscillator functions for the set Ψ_n, the function $g(x)$ must behave like $x^{3/2}$ for large x while the inverse function $f(x)$ must behave like $x^{2/3}$. Both f and g should be even functions of x. I originally tried to use $f(x) = x^{2/3}$. This is inadequate because the derivatives of f introduced singularities at the origin. For very small values of x the harmonic part of the potential is dominant, and so one may try to construct a function that behaves like x at the origin and $x^{2/3}$ for large values of x. The combinations that I tried were of the form

$$f(x) = x(\alpha|x|^{e-1/3} + \beta)/(\gamma|x|^e + \delta)$$

with $e > \frac{1}{3}$, and

$$f(x) = \alpha x e^{-\beta x^2} + \gamma x^{2/3}(1 - e^{-\delta x^2}).$$

These functions lead to quite complicated expressions for the Hamiltonian and to intractable integrals.

I then speculated on letting the Hamiltonian itself generate a suitable transformation function f. Here the technique approaches the WKB method quite closely. By using $f(x) = x^{2/3}$ I determined that the terms in the kinetic energy that arise from the noncommutability of p and x give rise to terms of lower order in the quantum number than the term $[f'(x)]^{-2}p^2$. This is easy to see if one takes $f'(x) \sim x^{-1/3}$, $f''(x) \sim x^{-4/3}$, and $f'''(x) \sim x^{-7/3}$. The terms that arise from the commutator are $[f'(x)]^{-3} \times f''(x)p$, which is asymptotically like $x^{-1/3}p$, $[f'(x)]^{-4} \times [f''(x)]^2$, and $[f'(x)]^{-3}f'''(x)$ which both behave like $x^{-4/3}$. Since the dependence on the quantum number n of an operator is related to the operators dependence on x and p for large values of x and p by the remarks in Sec. I and the detailed calculations in Appendix A, I decided to neglect these terms temporarily. In the same spirit I dropped the harmonic term compared to the quartic term. I was then left with the operator

$$\tfrac{1}{2}[f'(x)]^{-2}p^2 + \lambda f^4(x).$$

My final arguments were that p^2 and x^2 were asymptotically equal in magnitude but opposite in sign so that I could replace the expression above by

$$-\tfrac{1}{2}[f'(x)]^{-2}x^2 + \lambda f^4(x).$$

In order to obtain cancellation between the kinetic and potential energy portions of the Hamiltonian I should set this expression equal to zero as a differential equation. The solution for the equation is

$$f(x) = \alpha x^{2/3}.$$

Since I had already seen that $x^{2/3}$ is an inappropriate function, I included the harmonic part of the potential, the next most significant term in the differential equation. The differential equation becomes

$$-\tfrac{1}{2}[f'(x)]^{-2}x^2 + \tfrac{1}{2}f^2(x) + \lambda f^4(x) = 0.$$

The solution of this equation is readily obtained and is

$$f(x) = A[(1 + Bx^2)^{2/3} - 1]^{1/2}.$$

This function proves satisfactory both at zero and infinity and it is reasonably simple to use. The transformation generated by this function will be used to study the anharmonic oscillator. The abbreviation

$$u = (1 + Bx^2)^{1/3}$$

will prove to be useful.

The derivatives of f that are necessary for the calculation of the transformed Hamiltonian are given by

$$f'(x) = \tfrac{2}{3}ABxu^{-1}(u^2 - 1)^{-1/2},$$

$$f''(x) = -\tfrac{2}{9}AB(u^2 - 1)^{-3/2}(u + u^{-1} - 4u^{-2} + 2u^{-4}),$$

$$f'''(x) = \tfrac{2}{27}AB^2x(u + 1)^{-5/2}(u - 1)^{-1/2} \times (4u^{-2} + 8u^{-3} + 22u^{-4} - 4u^{-5} - 32u^{-6} - 16u^{-7}).$$

The transformed Hamiltonian is found by using these expressions for the derivatives of f and the expressions (4) and (5) for H. The result is

$$H = \frac{9}{8A^2B^2x^2}(u^4 - u^2)p^2 + \frac{3}{4iA^2B^2x^3}(u^4 + u^2 - 4u + 2u^{-1})p + \frac{3u^3 + 9u^2 + 28u + 12 - 48u^{-1} - 64u^{-2} - 36u^{-3} - 12u^{-4}}{32A(u^5 + 3u^4 + 5u^3 + 5u^2 + 3u + 1)} + \tfrac{1}{2}A^2(u^2 - 1) + \lambda A^4(u^4 - 2u^2 + 1). \quad (6)$$

In the subsequent sections we choose the constants A and B in the transformation and evaluate the diagonal matrix elements of H.

IV. THE DIAGONAL MATRIX ELEMENTS OF THE HAMILTONIAN

The diagonal matrix elements of H are evaluated with the help of the generating function $h(A,B,\lambda,z)$ which is defined by:

$$h(A,B,\lambda,z) = \sum z^n\langle n|H|n\rangle.$$

The operator H has the form $H = h_1(x)p^2 + ih_2(x)p + h_3(x)$. The treatment of each of the three terms is slightly different and they will be considered separately. The sums are evaluated with the help of Mehler's formula[7]:

$$\sum \frac{z^nH_n(x)H_n(y)}{2^n n!} = (1 - z^2)^{-1/2}\exp\left(\frac{[2xyz - (x^2 + y^2)z^2]}{(1 - z^2)}\right), \quad (7)$$

which is used with x set equal to y to give

$$\sum \frac{z^nH_n^2(x)}{2^n n!} = (1 - z^2)^{-1/2}\exp\left(\frac{2x^2z}{(1 + z)}\right),$$

and differentiated once with respect to y and with y set equal to x to give

$$\sum \frac{z^nH_n(x)H_{n+1}(x)}{2^n n!} = (1 - z^2)^{-1/2}\frac{2x}{1 + z}\exp\left(\frac{2x^2z}{(1 + z)}\right).$$

The evaluation of the contribution of the term $h_1(x)p^2$ to the generating function proceeds by replacing p^2 by $(p^2 + x^2) - x^2$. Since the states $|n\rangle$ are eigenstates of the harmonic oscillator, $(p^2 + x^2)|n\rangle = (2n + 1)|n\rangle$. Inside the sum that defines the generating function n may be replaced by $z\,d/dz$. With the use of these two substitutions the calculation of the first part of the generating function may be reduced to an integral in a routine way. The result is

$$\sum z^n\langle n|h_1(x)p^2|n\rangle = \frac{\pi^{-1/2}}{2(1 + z)(1 - z)^2}\int dy\; e^{-y^2}h_1[(1 + z)^{1/2}y/(1 - z)^{1/2}] \times [(1 + z^2) - (1 - z)^2y^2].$$

The variable of integration y is related to x by $(1 - z)^{1/2}x/(1 + z)^{1/2} = y$.

The contribution of the term $ih_2(x)p$ to the generating function is calculated by using $p|n\rangle = -i(n/2)^{1/2}|n-1\rangle +$

$i[(n+1)/2]^{1/2}|n+1\rangle$. The use of the differentiated version of the Mehler formula gives

$$\sum z^n\langle n|ih_2(x)p|n\rangle = [\pi(1-z^2)]^{-1/2}\int dy\ e^{-y^2}yh_2[(1+z)^{1/2}y/(1-z)^{1/2}].$$

The final term $h_3(x)$ gives the result

$$\sum z^n\langle n|h_3(x)|n\rangle = \pi^{-1/2}(1-z)^{-1}\int dy\ e^{-y^2}h_3[(1+z)^{1/2}y/(1-z)^{1/2}].$$

In some of the integrals that occur in the definitions of the generating function there appear to be singularities at the origin caused by terms like x^{-2} and x^{-4}. A brief inspection will show that all of these terms are cancelled by zeros of equivalent order in the numerator. The apparently singular factors are to be removed from the integrals through integration by parts. The portion of the integrand of the form $dx\ x^{-4}$ and $dx\ x^{-2}$ are to be replaced by $-\frac{1}{3}d(x^{-3})$ and $-d(x^{-1})$ respectively. When these partial integrations are carried out all of the integrals may be expressed in terms of the function $G(s,\theta)$ defined by

$$G(s,\theta) = \int_{-\infty}^{\infty} dx\ e^{-x^2}(1+\theta x^2)^s.$$

The evaluation and the properties of G are studied in Appendix D. The function G is closely related to the confluent hypergeometric function.

In terms of $G(s,\theta)$ the generating function $h(A,B,\lambda,z)$ is given by

$$\begin{aligned}\pi^{1/2}h(A,B,\lambda,z) &= \{9[(1-z)^2+2(1+z^2)]/8A^2B^2(1-z)(1+z)^2\}\\ &\times[G(\tfrac{2}{3},\theta)-G(\tfrac{4}{3},\theta)]\\ &+[3(1+z^2)/2A^2B(1-z)^2(1+z)]\\ &\times[2G(\tfrac{1}{3},\theta)-G(-\tfrac{1}{3},\theta)]\\ &+[3(1-z)/2A^2B^2(1+z)^2]\\ &\times[G(\tfrac{4}{3},\theta)+G(\tfrac{2}{3},\theta)-4G(\tfrac{1}{3},\theta)+2G(-\tfrac{1}{3},\theta)]\\ &-[1/A^2B(1+z)]\\ &\times[2G(\tfrac{1}{3},\theta)+G(-\tfrac{1}{3},\theta)-2G(-\tfrac{2}{3},\theta)-G(-\tfrac{4}{3},\theta)]\\ &+[1/32A^2(1-z)]\quad(3G(-\tfrac{2}{3},\theta)+13G(-\tfrac{4}{3},\theta)\\ &+\sum_{t=0}^{\infty}[-(42+90t)G((-5-6t)/3,\theta)\\ &+4G((-6-6t)/3,\theta)\\ &+(66+90t)G((-7-6t)/3,\theta)\\ &-(83+90t)G((-8-6t)/3,\theta)+(115+90t)\\ &\times G((-10-6t)/3,\theta)])\\ &+[A^2/2(1-z)][G(2/3,\theta)-G(0,\theta)]\\ &+[\lambda A^4/(1-z)][G(\tfrac{4}{3},\theta)-2G(\tfrac{2}{3},\theta)+G(0,\theta)].\end{aligned}$$

The results of Appendix D are used to expand the generating function h in a power series in z. The leading terms for large values of n are

$$\langle n|H|n\rangle = [2^{1/3}\Gamma(\tfrac{5}{6})/24\pi^{1/2}\Gamma(\tfrac{7}{3})]\times[27A^{-2}B^{-2/3}+40\lambda A^4B^{4/3}]n^{4/3}+O(n^{2/3}).$$

This part of the matrix element that behaves like $n^{4/3}$ is only a function of $AB^{1/3}$. If it is varied with respect to this parameter the stationary value occurs for $AB^{1/3} = (27/80\lambda)^{1/6}$. This is the first relation used to determine the constants A and B.

There is another significance to this relation. To the leading order in n the matrix element $\langle n|H|n+2\rangle$ vanishes if this value for $AB^{1/3}$ is chosen. The easiest way to show this result is to observe that the leading powers in n come from the leading terms in x^ap^b, that is, for the largest values of $a+b$. In the Hamiltonian these terms are

$$H = (\tfrac{9}{8})(AB^{1/3})^{-2}x^{2/3}p^2+\lambda(AB^{1/3})^4x^{8/3}. \tag{8}$$

The matrix element $\langle n|H|n+2\rangle$ is given in this approximation by

$$\langle n|H|n+2\rangle = \tfrac{9}{4}(AB^{1/3})^{-2}\langle n|x^{2/3}|n+2\rangle + \lambda[(AB^{1/3})^4-\tfrac{9}{8}(AB^{1/3})^{-2}]\langle n|x^{8/3}|n+2\rangle, \tag{9}$$

where p^2 has been written as $p^2+x^2-x^2$. If the results of Appendix A are used to evaluate this expression it is easily shown to be zero. The detailed calculation is carried out in the next section.

Since it is reasonable to expect that the matrix elements of H, $\langle a|H|b\rangle$ decrease with increasing values of $|a-b|$, it is reassuring to see that choosing the parameter $AB^{1/3}$ in the transformation in the manner specified above to make the diagonal element $\langle n|H|n\rangle$ stationary also has the effect of making the immediate and presumably the largest off-diagonal vanish.

Another relation between A and B is required to complete the specification of the transformation. Since the form of the transformation was partially dictated by the behavior for large values of x and partially by the behavior near the origin, it is reasonable to choose the parameters A and B in the same fashion. Thus we find the remaining relation between A and B to control the transformation near the origin.

Again there are two methods of doing this that give substantially the same result. If we regard the transformed ground state Ψ_0

$$\Psi_0 = \pi^{-1/4}[g'(x)]^{1/2}\exp[-g^2(x)],$$

where g is the function inverse to f; as the ground state of some potential V we can solve the Schrödinger equation for this potential:

$$V-E = \Psi_0''/2\Psi_0.$$

A relation between the parameters A and B can be chosen so that V behaves like $\frac{1}{2}x^2$ near the origin. This relation is

$$A^{-4}(9/4B^2-27/8B-1) = 1. \tag{10}$$

Alternatively one may require that the function f behave like x for small values of x. This gives the relation:

$$9/4A^4B^2 = 1,$$

which agrees with the preceding result for small values of λ. It would be useful to minimize the ground state energy but this has proved to be too difficult.

The choices of the parameters is then given by $AB^{1/3} = (27/80\lambda)^{1/6}$ and (10). This completely specifies the transformation.

V. THE OFF-DIAGONAL MATRIX ELEMENTS OF H

In this section the off-diagonal matrix elements of H are calculated to leading order. We start from the formulas (8) and (9), by substituting $2s$ for 2. This gives the result

$$\langle n|H|n+2s\rangle = \tfrac{9}{4}(AB^{1/3})^{-2}(n+2s)\langle n|x^{2/3}|n+2s\rangle + \lambda(AB^{1/3})^4 - \tfrac{9}{8}(AB^{1/3})^{-2}\langle n|x^{8/3}|n+2s\rangle.$$

If the results of Appendix A are used to estimate this result, it follows that

$$\langle n|H|n+2s\rangle = \{9n^{4/3}2^{1/3}\Gamma(\tfrac{4}{3})\Gamma(\tfrac{5}{6})/4\Gamma[(\tfrac{7}{3})-s]\Gamma[(\tfrac{7}{3})+s]\} \times (AB^{1/3})^{-2}(1-s^2).$$

The matrix element decreases like $s^{-5/3}$ as s increases. This rate of decrease is not fast enough to prevent the sum of the absolute values of the off-diagonal matrix elements from being approximately equal to the value of the diagonal element. The results of this calculation are in this sense less satisfying than the heuristic calculation presented in Appendix C.

VI. A SUBGROUP OF THE GROUP OF CANONICAL TRANSFORMATIONS

In this section a subgroup of the group of all possible canonical or unitary transformations is described. The transformation that brings the Hamiltonian (6) to a diagonal form belongs to this subgroup. Since our operators x and p have the correct asymptotic behavior for large values of x and p, any canonical transformation that tends to make the Hamiltonian more diagonal must belong to the subgroup which leaves the asymptotic behavior unaltered. That is, if $P = P(x,p)$ and $X = X(x,p)$ are a new set of operators in the sense of Appendix II, then

$$\lim_{\substack{x\to\infty\\ p\to\infty}} \frac{P(x,p)}{p} = \lim_{\substack{x\to\infty\\ p\to\infty}} \frac{X(x,p)}{x} = 1.$$

It is easy to give an example of a P or an X that has this behavior, but it is difficult to exhibit a nontrivial canonical pair. For example,

$$P = \alpha p + \beta(1+x^2)^{-1}p^3$$

satisfies the above relation. I do not know whether or not there is an X that is conjugate to this P. A conjecture that has some reasonable features is that

$$P = \sum \alpha_r (1+x^2)^{-r}p^{2r+1}, \qquad X = \sum \beta_r (1+p^2)^{-r}x^{2r+1}$$

are appropriate forms for the transformations of this type with relations between α and β specified by (B1).

VII. CONCLUSIONS

In the anharmonic oscillator problem there are two parameters: the coupling constant λ and the quantum number n. The present treatment determines with some success the leading dependence of the dynamical variables x and p and the energy levels on these parameters and the relations between them. In this way the structure of the solution of this problem are illustrated.

The operators p and x satisfy the canonical commutation relationship $[p,x] = -i\hbar$. Hence, if x behaves like $\lambda^{-\alpha}$, p must behave like $\lambda^{+\alpha}$. Since p^2 and λx^4 must have the same asymptotic behavior in λ, it follows that 2α and $1-4\alpha$ are equal so that $\alpha = \frac{1}{6}$. Since the energy behaves like p^2 or λx^4, it depends on $\lambda^{1/3}$. Finally since $p = m\dot{x} = m[H,x]$ the dependence of p like $\lambda^{1/6}$, H like $\lambda^{1/3}$, and x like $\lambda^{-1/6}$ are compatible.

In the same way the n dependence of p^2 and x^4 must be the same from the form of the Hamiltonian. The commutator implies that the product px is of order n so that $p = n^{2/3}$ and $x = n^{1/3}$, and $H = n^{4/3}$. Since $p_{ij} = (E_i - E_j)x_{ij}$, the n dependence of these quantities is verified.

Finally the asymptotic dependence of the wavefunction Ψ on the coordinate x follows from the n dependence of these operators. The solution of a one-dimensional problem is of value only in so far as it can suggest the solution of problems involving coupled degrees of freedom. The only direct generalization of the present result would be to a set of identical oscillators coupled in an SOn invariant fashion so that the problem could be reduced to a one-dimensional one. If the potential is not invariant under SOn, it still appears that a "radial" coordinate should be introduced and the attempt made to relate the energy primarily to a principal quantum number.

APPENDIX A: THE MATRIX ELEMENTS OF THE POWERS OF THE COORDINATE OPERATOR

In this appendix the matrix elements $\langle n|x^\eta|n+s\rangle$ will be calculated for an arbitrary $\eta > -1$. The approach taken is to evaluate the generating function

$$g(\eta,s,z) = \sum z^n\langle n|x^\eta|n+s\rangle[(n+s)!\,2^s/n!\,]^{1/2}.$$

If the matrix element is written as an integral and the order of summation and integration are exchanged the generating function is given by:

$$\pi^{1/2}g = \int dx\; e^{-x^2}x^\eta \sum z^n H_n(x)H_{n+s}(x)/2^n n!\,.$$

From Mehler's formula it is possible to develop an expression for the sum. Both sides of (7) are to be differentiated s times with respect to y. If the formula $H_m'(y) = 2mH_{m-1}(y)$ is taken into account, the left-hand side becomes

$$\sum z^{n+s}H_{n+s}(x)H_n(y)/2^n n!\,.$$

In order to deal with the right-hand side, we observe that it has the form $(1-z^2)^{-1/2}\exp(A+By+Cy^2)$. It is easy to prove by induction the lemma

$$\left(\frac{d}{dy}\right)^s \exp(A+By+Cy^2) = \sum \frac{s![2(B+Cy)]^{s-2r}(2C)^r}{(s-2r)!\,r!\,2^r}.$$

If the appropriate expressions for $A, B,$ and C are entered, it follows that

$$(d/dy)^s \exp[(-x^2z^2 + 2xzy - y^2z^2)/(1-z^2)] = \sum [(-1)^r s!\,2^s z^s (x-zy)^{s-2r}/(s-2r)! \times r!\,2^r(1-z^2)^{s-r}].$$

If we set $y = x$ in the differentiated version of Mehler's formula, it becomes

$$\sum \frac{z^n H_n(x)H_{n+s}(x)}{2^n n!}$$

$$= (1-z^2)^{-1/2}\sum \frac{(-1)^r s!\, 2^s x^{s-2r}}{(s-2r)!\, r!\, 2^{2r}(1+z)^{s-r}(1-z)^r} \exp\left(\frac{2x^2 z}{1+z}\right).$$

With this result the expression for g becomes

$$\pi^{1/2} g = (1-z^2)^{-1/2}\sum \frac{(-1)^r s!\, 2^s}{(s-2r)!\, 2^{2r}(1+z)^{s-r}(1-z)^r} \times \int dx\, x^{\eta+s-2r} \exp\left[-x^2\left(\frac{1-z}{1+z}\right)\right].$$

The integration can easily be carried out to give

$$\pi^{1/2} g = (1+z)^{(\eta-s)/2}(1-z)^{-(\eta+s+2)/2} 2^s \times \sum \frac{(-1)^r s!\, \Gamma[(\eta+s-2r+1)/2]}{(s-2r)!\, r!\, 2^r}.$$

In carrying out the integration it was assumed that x^η was an even or odd function of x depending on whether s is an even or odd integer, that is, $x^{\eta+s-2r}$ is an even function of x.

In evaluating the sum in the above expression it is convenient to factor out a term $\Gamma[(\eta+1)/2]\{\Gamma[(\eta+2)/2]\}$ from the sum depending on whether s is even (odd). The remaining sum is a polynomial in η of degree $(\frac{1}{2}s - \frac{1}{2})$. It can easily be shown that this polynomial vanishes for $\eta = 0, 2, 4, \ldots, s-2$ ($\eta = 1, 3, 5, \ldots, s-2$) so that the polynomial is of the form $A_s n(\eta-2)(\eta-4)\cdots(\eta-s+2)$ ($B_s(\eta-1)(\eta-3)\cdots(\eta-s+2)$). The constants are determined by evaluating the polynomial at $\eta = s$. The result is that $A_s = 2^{-s/2}$, $B_s = 2^{\frac{1}{2}-\frac{1}{2}s}$. If the results are now combined, it follows that the sum is given by

$$\sum \frac{(-1)^r s!\, \Gamma[(\eta+s-2r+1)/2]}{(s-2r)!\, r!\, 2^{2r}} = \frac{\Gamma(\frac{1}{2}\eta+1)\Gamma(\frac{1}{2}\eta+\frac{1}{2})}{\Gamma(\frac{1}{2}\eta-\frac{1}{2}s+1)}.$$

With this result the generating function becomes

$$\pi^{1/2} g = (1+z)^{\frac{1}{2}\eta-\frac{1}{2}s}(1-z)^{-\frac{1}{2}\eta-\frac{1}{2}s-1} 2^s \Gamma(\tfrac{1}{2}\eta+1)\Gamma(\tfrac{1}{2}\eta+\tfrac{1}{2})/ \times \Gamma(\tfrac{1}{2}\eta-\tfrac{1}{2}s+1).$$

The matrix elements are found by expanding this function in powers of z. For our purposes this is most easily carried through as follows:

$$(1+z)^{\frac{1}{2}\eta-\frac{1}{2}s}(1-z)^{-\frac{1}{2}\eta-\frac{1}{2}s-1} = (1+z)^{1+\eta}(1-z^2)^{-\frac{1}{2}\eta-\frac{1}{2}s-1}.$$

The exponent $1+\eta$ will always be a number of the order of one for our purposes while the exponent $-\frac{1}{2}\eta - \frac{1}{2}s - 1$ will be an arbitrary negative number. The expansion is given by

$$(1+z)^{1+\eta}(1-z^2)^{-\frac{1}{2}\eta-\frac{1}{2}s-1} = \sum z^n \sum \times \frac{\Gamma(\eta+2)\Gamma(q+\frac{1}{2}\eta+\frac{1}{2}s+1)}{\Gamma(\eta+2-n+2q)\Gamma(\frac{1}{2}\eta+\frac{1}{2}s+1)q!\,(n-2q)!}.$$

For values of n that are large compared to s the sum over q may be approximated, and this gives

$$(1+z)^{1+\eta}(1-z^2)^{-\frac{1}{2}\eta-\frac{1}{2}s-1} = \sum \frac{\Gamma(n+1+\frac{1}{2}\eta+\frac{1}{2}s)}{n!\,\Gamma(1+\frac{1}{2}\eta+\frac{1}{2}s)}(z^{2n}+z^{2n+1})2^\eta.$$

The generating function g has the expansion

$$g(\eta, s, z) = 2^{s+\eta}\frac{\Gamma(\frac{1}{2}\eta+1)\Gamma(\frac{1}{2}\eta+\frac{1}{2})}{\Gamma(\frac{1}{2}\eta-\frac{1}{2}s+1)} \times \sum \frac{\Gamma(n+1+\frac{1}{2}\eta+\frac{1}{2}s)}{n!\,\Gamma(1+\frac{1}{2}\eta+\frac{1}{2}s)}(z^{2n}+z^{2n+1}).$$

The matrix elements can be read off from this expression:

$$\langle n|x^\eta|n+s\rangle = \frac{2^s\Gamma(\frac{1}{2}\eta+1)\Gamma(\frac{1}{2}\eta+\frac{1}{2})}{\Gamma(\frac{1}{2}\eta-\frac{1}{2}s+1)}\left(\frac{n!}{(n+s)!\,2^s}\right)^{1/2} \times \frac{\Gamma([\frac{1}{2}n]+1+\frac{1}{2}\eta+\frac{1}{2}s)2^\eta}{[\frac{1}{2}n]!\,\Gamma(1+\frac{1}{2}\eta+\frac{1}{2}s)}.$$

For $n \gg s$ we can use Stirling's formula to yield

$$\langle n|x^\eta|n+s\rangle = (2n)^{\eta/2}\Gamma(\tfrac{1}{2}\eta+1)\Gamma(\tfrac{1}{2}\eta+\tfrac{1}{2})/ \Gamma(\tfrac{1}{2}\eta-\tfrac{1}{2}s+1)\Gamma(1+\tfrac{1}{2}\eta+\tfrac{1}{2}s).$$

For values of s that are large compared to one but small compared to n, this matrix element decreases like $s^{-1-\eta}$.

APPENDIX B: CANONICAL TRANSFORMATIONS

The commutator $[p^m, x^n]$ is given by

$$[p^m, x^n] = \sum_{t=1} \frac{[p,x]^t m!\, n!}{t!\,(m-t)!\,(n-t)!}\, x^{n-t} p^{m-t}.$$

This relation can easily be established by induction. From this it follows that

$$[x^n p^m, x^s p^r] = x^n[p^m, x^s]p^r - x^s[p^r, x^n]p^m = \sum \frac{[p,x]^t}{t!} : \frac{\partial^t(x^s p^r)}{\partial x^t}\frac{\partial^t(x^n p^m)}{\partial p^t} - \frac{\partial^t(x^s p^r)}{\partial p^t}\frac{\partial^t(x^n p^m)}{\partial x^t} : .$$

In this formula the operators x and p must be rearranged in a normal order with all of the x's moved to the left and all of the p's moved to the right without regard to their noncommutability. Hence the notation : : .

If one considers the transformation $p \to P = P(x,p)$ and $x \to X = X(x,p)$, then the commutator

$$[P, X] = \sum \frac{[p,x]^t}{t!} : \frac{\partial^t X}{\partial x^t}\frac{\partial^t P}{\partial p^t} - \frac{\partial^t X}{\partial p^t}\frac{\partial^t P}{\partial x^t} : . \qquad \text{(B1)}$$

To derive this formula, it is imagined that X and P are expanded in power series in x and p and that the previous result is used to compute the commutator. The classical analog is the first term in the series which is the zero order approximation in Planck's constant. This equation shows that many transformations that are canonical classically are not in quantum mechanically and vice versa.

APPENDIX C: SOME HEURISTIC CONSIDERATIONS

I have made some attempts to directly construct operators x and p that satisfy the canonical commutation rules and diagonalize H the anharmonic oscillator Hamiltonian (1). The attempts are not successful, but they are instructive. The simplest attempt is

$$\sqrt{2}x = \alpha \begin{pmatrix} 0 & 1^{1/3} & 0 & 0 & 0\cdots \\ 1^{1/3} & 0 & 2^{1/3} & 0 & 0\cdots \\ 0 & 2^{1/3} & 0 & 3^{1/3} & 0\cdots \\ 0 & 0 & 3^{1/3} & 0 & 4^{1/3}\cdots \\ 0 & 0 & 0 & 4^{1/3} & 0\cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix}$$

and

$$\sqrt{2}\,ip = \beta \begin{bmatrix} 0 & 1^{2/3} & 0 & 0 & 0\cdots \\ -1^{2/3} & 0 & 2^{2/3} & 0 & 0\cdots \\ 0 & -2^{2/3} & 0 & 3^{2/3} & 0\cdots \\ 0 & 0 & -3^{2/3} & 0 & 4^{2/3}\cdots \\ 0 & 0 & 0 & -4^{2/3} & 0\cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix}$$

The commutator is

$$i[p,x] = \alpha\beta \begin{bmatrix} 1 & 0 & -\tfrac{1}{2}(2^{1/3}1^{2/3} - 1^{1/3}2^{2/3}) & 0 & \cdots \\ 0 & 1 & 0 & -\tfrac{1}{2}(3^{1/3}2^{2/3} - 2^{1/3}3^{2/3})\cdots & \\ \tfrac{1}{2}(2^{1/3}1^{2/3} - 1^{1/3}2^{2/3}) & & 1 & 0 & \cdots \\ 0 & \tfrac{1}{2}(3^{1/3}2^{2/3} - 2^{1/3}3^{2/3}) & 0 & 1 & \cdots \\ 0 & 0 & \tfrac{1}{2}(4^{1/3}3^{2/3} - 3^{1/3}4^{2/3}) & 0 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix}.$$

The off-diagonal element is asymptotically $\frac{1}{6}\alpha\beta$. The anharmonic oscillator Hamiltonian calculated with these approximations for p and x is given by

$$\begin{aligned} H_{mn} = {} & (\alpha^4\lambda/4)[(m-3)(m-2)(m-1)m]^{1/3}\delta_{m-4,n} \\ & + \{(\alpha^4\lambda/4)(m-1)m^{1/3}[(m-2)^{2/3} + (m-1)^{2/3} \\ & + m^{2/3} + (m+1)^{2/3}] + (\alpha^2/4)[(m-1)m]^{1/3} \\ & - (\beta^2/4)[(m-1)m]^{2/3}\}\delta_{m-2,n} \\ & + \{(\lambda\alpha^4/4)[m^{2/3}(m-1)^{2/3} \\ & + m^{4/3} + 2m^{2/3}(m+1)^{2/3} \\ & + (m+1)^{4/3} + (m+1)^{2/3}(m+2)^{2/3}] \\ & + (\alpha^2/4)[m^{2/3} + (m+1)^{2/3}] + (\beta^2/4)[m^{4/3} \\ & + (m+1)^{4/3}]\}\delta_{m+4,n} \\ & + \text{terms in } \delta_{m+2,n} \text{ and } \delta_{m+4,n} \text{ necessary to make } H \\ & \text{Hermitian.} \end{aligned}$$

The choice $\alpha\beta = 1$ is obvious to make the commutator as nearly correct as possible. In order to make H as diagonal as possible for large values of the quantum number, the portion of H that is proportional to $n^{4/3}\,\delta_{m-2,n}$ should be set equal to zero. This gives

$$2\lambda\alpha^4 = \beta^2,$$

or

$$\alpha = (1/2\lambda)^{1/6}.$$

The leading term in the remaining off-diagonal element which is proportional to $n^{4/3}\delta_{m-4,n}$ is then

$$\lambda\alpha^4 = \tfrac{1}{4}(\lambda/4)^{1/3}.$$

The leading diagonal element is:

$$\tfrac{1}{2}\beta^2 + \tfrac{3}{2}\alpha^4\lambda = \tfrac{1}{2}\alpha^4\lambda(1/\alpha^6\lambda + 3) = \tfrac{1}{2}(\lambda/4)^{1/3}(2+3) = \tfrac{5}{2}(\lambda/4)^{1/3}.$$

Since there are two off-diagonal elements in any row or column, the ratio of the off-diagonal elements to the diagonal element is $\frac{1}{5}$. The numerical value of the diagonal matrix element is $1.6\lambda^{1/3}$.

Very marked improvements can be made in this solution by making the matrices p and x more complicated. The next simplest form would be to consider the matrix p to have terms proportional to $\delta_{m\pm3,n}$. This example gives the energy proportional to $2.22\lambda^{1/3}$ essentially the WKB result. If the nonvanishing terms proportional to $\delta_{m\pm3,n}$ are included in the matrix for x, the energy becomes $2.19\lambda^{1/3}$, exactly the WKB value, while the ratio of the off-diagonal to the diagonal element falls to 0.05.

There are a number of difficulties to be overcome before a satisfactory method can emerge from these considerations. The problem being treated in all these instances is really $\frac{1}{2}p^2 + \lambda x^4$. In the large quantum number approximation which is all that has been treated there is no contribution from the harmonic part of the potential. In the earlier work this part of the potential was included through the form of the function f. To do this in the present approach without destroying the simplicity of the method may be difficult. Further, two approximations must be made simultaneously as the matrices for x and p are developed. On the one hand, there are terms of order lower than $n^{4/3}$ with coefficients that may be large and on the other there are terms proportional to $n^{4/3}$ with small coefficients. How these two types of terms can be eliminated simultaneously and compatibly is the obstacle to giving a rigorous account of the anharmonic oscillator by this technique.

APPENDIX D: THE FUNCTION $G(s, \theta)$

The function $G(s,\theta)$ is defined by the integral

$$G(s,\theta) = \int e^{-x^2}(1 + \theta x^2)^s dx.$$

By differentiating under the integral sign with respect to θ it can be seen that G satisfies the differential equation

$$2\theta^2 G'' + [2 + (3 - 2s)\theta]G' - sG = 0. \qquad \text{(D1)}$$

The equation can be solved in a series of descending powers of θ. If s is not half of an integer, there are two separate series that solve the equation, and the general solution is

$$G(s,\theta) = \alpha(s)\theta^s \sum \frac{\Gamma(n-s)\theta^{-n}}{n!\,\Gamma(n-s+\frac{1}{2})} + \beta(s)\theta^{-1/2} \sum \frac{\Gamma(n+\frac{1}{2})\theta^{-n}}{n!\,\Gamma(n+s+\frac{3}{2})}. \quad \text{(D2)}$$

The coefficients α and β may be evaluated by considering the integral for large values of θ. If $s > -\frac{1}{2}$, the dominant term is $\alpha(s)\theta^s\Gamma(-s)/\Gamma(-s-\frac{1}{2})$, while, if $s < \frac{1}{2}$, the dominant term is $\beta(s)\theta^{-1/2}\Gamma(\frac{1}{2})/\Gamma(s+\frac{3}{2})$. In each of these regions it is easy to work out the asymptotic form of the integral. For $s < -\frac{1}{2}$ make the change of variable $\theta x^2 = t/(1-t)$. The integral becomes

$$\theta^{-1/2}\int t^{(1/2)-1}(1-t)^{-(1/2)-s-1}\exp[-t/\theta(1-t)]dt.$$

For large values of θ the exponential is approximately one and it may be neglected. The asymptotic value of G when $s < -\frac{1}{2}$ is then given by

$$\lim_{\theta\to\infty} G(s,\theta) = \theta^{-1/2}B(\tfrac{1}{2},-\tfrac{1}{2}-s) = \theta^{-1/2}\Gamma(\tfrac{1}{2})\,\Gamma(-\tfrac{1}{2}-s)/\Gamma)-s), \quad s < -\tfrac{1}{2}.$$

For $s > \frac{1}{2}$ and large values of θ we may simply replace $1+\theta x^2$ by θx^2, and G becomes

$$\lim_{\theta\to\infty} G(s,\theta) = \theta^s\Gamma(s+\tfrac{1}{2}), \quad s > -\tfrac{1}{2}.$$

From these results it follows that

$$\alpha(s) = -\beta(s) = \Gamma(s+\tfrac{1}{2})\,\Gamma(-s+\tfrac{1}{2})/\Gamma(-s). \quad \text{(D3)}$$

For $\alpha(s)$ the result is true for $s > -\frac{1}{2}$ while for $\beta(s)$ the result is true for $s < -\frac{1}{2}$. We shall now show that G satisfies a mixed difference-differential equation in s and θ and a difference equation in s and that these values of α and β are true for all values of s.

By differentiating under the integral sign in the definition of G the two mixed differential-difference equations

$$2\theta^2G'(s) + (\theta+2)G(s) - 2G(s+1) = 0 \quad \text{(D4)}$$

and

$$\theta G'(s+1) + (s+1)G(s) - (s+1)G(s+1) = 0$$

can be deduced.

By substituting the series solution (13) for G in these equations it follows that

$$(s+1)\alpha(s) = \alpha(s+1), \quad (s+1)\beta(s) = \beta(s+1).$$

The expressions (D3) for $\alpha(s)$ and $\beta(s)$ satisfy these equations so the values for $\alpha(s)$ and $\beta(s)$ are correct for all values of s and the integral is represented by the sum (D2). The series are uniformly and absolutely convergent for all real positive values of θ. The function G is singular at $\theta = 0$ although the value at $\theta = 0$ is just $\pi^{1/2}$.

The differential equation (D1) may be deduced by eliminating the terms dependent on $G(s+1)$ between the pair of equations (D4). Alternatively the derivatives may be eliminated to give the difference equation

$$G(s+1,\theta) - [(s+\tfrac{1}{2})\theta + 1]G(s,\theta) + s\theta G(s-1,\theta) = 0.$$

In the present work the function G appears as a generating function with the argument $\theta = B(1+z)/(1-z)$. The expansion of G in power of z is required to determine matrix elements. The series (D2) for G exhibit this function in terms of powers of $1+z$ and $1-z$. It has singularities at $z = \pm 1$.

If $G(s,z)$ is expanded in a power series in z,

$$G(s,z) = \sum A_N z^N,$$

the behavior of the coefficients A_N for large values of N will be related to the behavior of G near $z = \pm 1$. The singularity at $z = 1$ is a branch point where $(1-z)^eG$ vanishes for a finite value of e. The singularity at $z = -1$ is much more complicated since $(1+z)^eG$ is infinite for all values of e. This singularity comes from the essential singularity in the exponential function in the integrand. It makes minor contributions to the asymptotic behavior of A_N, but it is difficult to treat because it is necessary to treat all the powers of $(1+z)$ simultaneously rather than term by term.

The function $(1-z)^sG(s,z)$ is easy to expand in powers of z. Direct differentiation of the integral defining G gives the result

$$(1-z)^sG(s,z) = \sum [(-1)^N\Gamma(N-s)z^N/\Gamma(-s)N!] \times \int e^{-x^2}(1+Bx^2)^{s-N}(Bx^2-1)^N dx.$$

In order to obtain the expansion of $G(s,z)$, the factor $(1-z)^s$ must be brought to the right-hand side of the equation and expanded. This is done differently for positive and negative values of s. If $s > 0$, the expansion of $(1-z)^{-s}$ is given by

$$(1-z)^{-s} = \sum [\Gamma(q+s)/\Gamma(-s)q!]z^q.$$

The expansion of G becomes:

$$G(s,z) = \sum [\Gamma(N+s)/\Gamma(s)N!]z^N \times \int F(-s,-N,-N-s+1, \times (1-Bx^2)/(1+Bx^2)) \times e^{-x^2}(1+Bx^2)^s dx,$$

where F is the Gauss hypergeometric series. For large values of N the hypergeometric function F becomes $[1+(Bx^2-1)/(Bx^2+1)]^s$ so that G is approximately given by:

$$G(s,z) = \sum [\Gamma(N+s)/\Gamma(s)N!]z^N \int e^{-x^2}(2Bx^2)^s dx = (2B)^s\Gamma(s+\tfrac{1}{2}) \sum [\Gamma(N+s)/\Gamma(s)N!]z^N.$$

The final simplification is to use Stirling's formula to approximate the factorials. The result for G is

$$G(s,z) = [\Gamma(s+\tfrac{1}{2})/\Gamma(s)](2B)^s \sum N^{s-1}z^N.$$

For $s < -\frac{1}{2}$ the following approach shows that for large values of N the coefficient A_N^s of z^N in the power series expansion of $G(s,z)$ is bounded by a constant independent of N. In order to show this, G is written as

$$G(s,z) = B^{-1/2} \sum (z^N/2^N N!)$$

$$\times \int dy(1+y^2)^s(1-z)e^{-y^2/B}[H_N^2(yB^{-1/2}) - 2NH_{N-1}^2(yB^{-1/2})].$$

The inequality $H_N(x) < ke^{x^2/2}N!\theta 2^{N/2}$ can be used to bound the coefficients A_N^s by

$$A_N^s \leq B^{-1/2}2k^2 \int dy(1+y^2)^s$$

which is finite N-independent constant for $s < -\frac{1}{2}$.

*This work is supported in part by the United States Atomic Energy Commission.

[1]P. A. M. Dirac, *The principles of quantum mechanics* (Oxford at the Clarendon Press, Oxford, 1947), p. 136.
[2]N. Fröman and O. Fröman, *JWKB approximation* (North-Holland, Amsterdam, 1965).
[3]S. Graffi, V. Grecchi, and G. Turchetti, Nuovo Cimento B **11,** 313 (1971).
[4]Appendix A.
[5]David Bohm *Quantum theory* (Prentice Hall, New York, 1951), p. 297.
[6]Appendix B.
[7]A. Erdélyi, et al., *Higher transcendental functions* (McGraw-Hill, New York, 1953), Vol. 2, p. 194.

Spaces admitting gravitational fields

Kishore B. Marathe
Physics Department, St. John Fisher College, Rochester, New York 14618
(Received 27 September 1971; revised manuscript received 5 November 1971)

We start with a modern version of Einstein's definition of a gravitational field. Tensors of curvature type and the curvature product of symmetric tensors are defined. The interaction tensor is defined as the curvature product of the fundamental tensor and the energy-momentum tensor. The tensor W obtained by coupling the Riemann tensor and the interaction tensor is used to obtain a characterization of gravitational fields. The linear transformation of the space of second-order differential forms, induced by W, is used to give a new definition of a gravitational field. The field equations are expressed in terms of the gravitational sectional curvature function f. Thorpe's theorem characterizing Einstein spaces is obtained as a corollary. New formulations of the field equations are used to solve the problem of classification of gravitational fields. The mathematical foundations of the theory of classification are examined and a geometric interpretation of classification is obtained by using the critical point theory on a suitable manifold.

1. INTRODUCTION

In recent years, global analysis and the theory of manifolds have become increasingly important tools for investigations into the mathematical theory of relativity. It is interesting to note that work in this direction had already been done by Elie Cartan many years ago.[1] Some of these techniques were also used by H. Weyl and are found, for example, in his classic "Raum, Zeit, Materie." However, it was not until Ehresmann[2] formulated the concept of connection in a general setting that theory of manifolds got its momentum.

In the theory of gravitation and relativity, attention has been concentrated on the analysis of Einstein's field equations and their solutions—both exact and approximate —and on the possible alternatives to, and generalizations of, the theory. That the theory of manifolds is suitable for diverse applications has been shown by the well-known work of Lichnerowicz, Choquet-Bruhat, and more recently by Ehlers.[3]

The setting of manifold theory is used here to obtain a characterization of gravitational fields. This result is stated as Theorem 1, and its proof is the main objective of this paper, as it enables us to give alternative formultations of Einsteins' field equations.

2. DEFINITION OF GRAVITATIONAL FIELD

By a differential four-dimensional manifold M of class C^r, we mean a Hausdorff, connected, locally Euclidean topological space with a fixed four-dimensional C^r atlas, r need not be infinite, but in what follows we assume it to be large enough to ensure smoothness of the operations involved. We denote by $L(M)$ the bundle of linear frames on M. By the pseudo-Riemannian structure Γ_g on M corresponding to the symmetric fundamental tensor g, we shall understand the unique torsion-free Levi–Civita connection Γ_g on $L(M)$ such that g is parallel with respect to Γ_g. We will say that a tensor is of type (r, s) if it is contravariant of degree r and covariant of degree s. We use the same letter g for the fundamental tensor of type (0, 2) or its dual of type (2, 0) and a similar usage is followed for other tensors.

Definition 1 is essentially Einstein's definition of gravitational fields. It is stated here in a form acceptable to present day mathematics.

Definition 1: Let T be a symmetric tensor of type (2, 0) on M. We define the gravitational field F, with source (energy–momentum tensor) T, as the triple (M, g, T), which satisfies the following conditions:

F–1: g has signature $(-,-,-,+)$;
F–2: T is divergence-free;
F–3: g satisfies Einstein's field equations with source T.

Using a local coordinate chart and the induced basis of the tensor algebra of M, we can write conditions F–2 and F–3 in the familiar form

$$\nabla_i T^{ij} = 0 \tag{1}$$

and

$$R^{ij} - \tfrac{1}{2}Rg^{ij} = -T^{ij}, \tag{2}$$

where ∇_i is the covariant derivative with respect to the vector $\partial/\partial x^i$, and the units are chosen so that the coefficient of T^{ij} in (2) is -1.

Condition (1) shows that an arbitrary tensor cannot act as a source of the gravitational field. In view of the Bianchi identities, Condition (1) acts as a consistency condition for (2).

We say that M is the carrier of the field F or that the space M admits the gravitational field F. In what follows we assume M to be a pseudo-Riemannian manifold with fundamental tensor g satisfying F–1.

3. TENSORS OF CURVATURE TYPE

We now introduce tensors of curvature type. They are defined so as to have the algebraic properties of the Riemann–Christoffel curvature tensor and play a vital role in the analysis of gravitational fields.

Definition 2: Let S be a tensor of type (4, 0) on M. We can regard S as a quadrilinear mapping (pointwise) so that for each $x \in M$, S_x can be identified as the map

$$S_x : T_x^*(M) \times T_x^*(M) \times T_x^*(M) \times T_x^*(M) \to R,$$

where R is the real field and $T_x^*(M)$ is the space of first order differential forms at $x \in M$. It is dual to the tangent space $T_x(M)$ at $x \in M$. We say that S is of curvature type if S_x satisfies the following conditions for each $x \in M$:

C–1: $S_x(e^1, e^2, e^3, e^4) = -S_x(e^2, e^1, e^3, e^4)$;
C–2: $S_x(e^1, e^2, e^3, e^4) = -S_x(e^1, e^2, e^4, e^3)$;
C–3: $S_x(e^1, e^2, e^3, e^4) + S_x(e^1, e^3, e^4, e^2) + S_x(e^1, e^4, e^2, e^3) = 0$;

where $e^i \in T_x^*(M)$ for $i = 1, 2, 3, 4$.

From the definition, it follows that S also satisfies the following condition,

C-4: $S_x(e^1, e^2, e^3, e^4) = S_x(e^3, e^4, e^1, e^2)$.

Example 1: The Riemann–Christoffel tensor is of curvature type. In fact, the definition given above is modeled after this example.

Example 2: The tensor G defined by

$$G_x(e^1, e^2, e^3, e^4) = g_x(e^1, e^3)g_x(e^2, e^4) - g_x(e^2, e^3)g_x(e^4, e^1),$$

where g is the fundamental tensor of M, is of curvature type.

A useful characterization of tensors of curvature type is given by the following lemma.

Lemma 1: Let S be a tensor of curvature type. Then S_x induces a linear transformation of $\Lambda_x^2(M)$ for each $x \in M$, where $\Lambda_x^2(M)$ is the space of differential forms of degree two at x and conversely the induced tranformation determines S_x completely.

Proof: Let (e^1, e^2, e^3, e^4) be a basis of $T_x^*(M)$. Then an induced basis of $\Lambda_x^2(M)$ is given by the elements b^i, $i = 1, 2, \ldots, 6$ as follows:

$$b^1 = e^2 \wedge e^3, \quad b^2 = e^3 \wedge e^1,$$

$$b^3 = e^1 \wedge e^2, \quad b^4 = e^1 \wedge e^4,$$

$$b^5 = e^2 \wedge e^4, \quad b^6 = e^3 \wedge e^4.$$

We now introduce a pseudo-inner product G_x on $\Lambda_x^2(M)$ defined by

$$G_x(e^i \wedge e^j, e^k \wedge e^l) = g_x(e^i, e^k)g_x(e^j, e^l) - g_x(e^j, e^k)g_x(e^l, e^i) \quad (3)$$

If the basis (e^1, e^2, e^3, e^4) is g orthonormal, then the basis $(b^1, b^2, b^3, b^4, b^5, b^6)$ is G-orthonormal. The signature of G is $(+, +, +, -, -, -)$. Henceforth, we will work with this G-orthonormal basis of $\Lambda_x^2(M)$.

To define the linear transformation S_x, it is enough to give its action on the basis elements. We define $S_x(e^i \wedge e^j)$ to be the element whose G product with $e^k \wedge e^l$ is S^{ijkl}; i.e.,

$$G(S_x(e^i \wedge e^j)\, e^k \wedge e^l) = S^{ijkl}. \quad (4)$$

It is easy to verify that this defines a linear transformation of $\Lambda_x^2(M)$. The formula (4) also shows that the transformation gives all the components of S.

To obtain the matrix of S_x relative to the basis $(b^1, b^2, b^3, b^4, b^5, b^6)$, we introduce the quantities $s^{\alpha\beta}$, $\alpha, \beta = 1, 2, \ldots, 6$, obtained by replacing pairs of suffixes of S by single suffixes, by the correspondence $23 \leftrightarrow 1$, $31 \leftrightarrow 2$, $12 \leftrightarrow 3$, $14 \leftrightarrow 4$, $24 \leftrightarrow 5$, $34 \leftrightarrow 6$. For example,

$$s^{11} = S^{2323} \quad \text{and} \quad s^{56} = S^{2434}.$$

Using the quantities $s^{\alpha\beta}$, the matrix of S can be written as

$$\begin{pmatrix} s^{11} & s^{12} & s^{13} & -s^{14} & -s^{15} & -s^{16} \\ s^{21} & s^{22} & s^{23} & -s^{24} & -s^{25} & -s^{26} \\ s^{31} & s^{32} & s^{33} & -s^{34} & -s^{35} & -s^{36} \\ s^{41} & s^{42} & s^{43} & -s^{44} & -s^{45} & -s^{46} \\ s^{51} & s^{52} & s^{53} & -s^{54} & -s^{55} & -s^{56} \\ s^{61} & s^{62} & s^{63} & -s^{64} & -s^{65} & -s^{66} \end{pmatrix}. \quad (5)$$

We now introduce the concept of curvature product of two symmetric tensors of type (2, 0).

Definition 3: Let g and T be two symmetric tensors of type (2, 0) on M. The curvature product of g and T denoted by $g \times_C T$ is a tensor of type (4,0) defined by (all tensors acting pointwise):

$$g \times_C T(X^1, X^2, X^3, X^4) = \tfrac{1}{2}[g(X^1, X^3)\cdot T(X^2, X^4) + g(X^2, X^4)\cdot T(X^1, X^3) - g(X^1, X^4)\cdot T(X^2, X^3) - g(X^2, X^3)\cdot T(X^1, X^4)], \quad (6)$$

where $X^i \in T_x^*(M)$, the dual of the tangent space at $x \in M$, for $i = 1, 2, 3, 4$. The following properties of the curvature product follow immediately from the definition.

P-1: $g \times_C T = T \times_C g$.

P-2: $g \times_C T$ is a tensor of curvature type.

P-3: $g \times_C g = G$, where G is the tensor defined in example 2.

Remark: When g is the fundamental tensor of M and T is the energy–momentum tensor, we call $g \times_C T$ the interaction tensor between the field and the source. It plays an essential role in the characterization of the gravitational fields given here.

Definition 4: A complex structure on a real vector space V is a linear transformation J of V such that $J^2 = -I$, where I is the identity transformation of V.

Lemma 2: The Hodge star operator J on $\Lambda_x^2(M)$ defines a complex structure on $\Lambda_x^2(M)$.

For a definition and description of the properties of the Hodge star operator and proof of Lemma 2, see, e.g., Thrope.[4] The set of complex transformations of $\Lambda_x^2(M)$ can be regarded as a subset of the set of real transformations of $\Lambda_x^2(M)$. We now state without proof, the following Lemma. For details see Ref. 5.

Lemma 3: The following are equivalent:

(1) e is a complex transformation of $\Lambda_x^2(M)$;

(2) e commutes with J;

(3) The matrix of e is of the form

$$\begin{pmatrix} A & B \\ -B & A \end{pmatrix},$$

where A, B are real 3×3 matrices.

4. A CHARACTERIZATION OF GRAVITATIONAL FIELDS

We are now in a position to state the main result as the following:

Theorem 1: Let M be a pseudo-Riemannian manifold with fundamental tensor g of signature $(-,-,-,+)$. Let T by a symmetric, divergence-free tensor of type $(2,0)$ whose trace is equal to the scalar curvature. Let R be the Riemann–Christoffel curvature tensor of type $(4,0)$. Define W by the equation

$$W = R + g \times_C T. \tag{7}$$

Then (M,g,T) is a gravitational field if and only if W, regarded as a linear transformation of $\Lambda_x^2(M)$, commutes with the Hodge star operator J.

Proof: Since R and $g \times_C T$ are of curvature type, clearly W is also of curvature type, and therefore, by Lemma 1 induces a linear transformation of $\Lambda_x^2(M)$. The matrix of W is obtained by using the construction and notation of Lemma 1.

Now, using Lemma 3, we see that W commutes with J if an only if the following conditions are satisfied:

$$\begin{pmatrix} w^{11} & w^{12} & w^{13} \\ w^{21} & w^{22} & w^{23} \\ w^{31} & w^{32} & w^{33} \end{pmatrix} = -\begin{pmatrix} w^{44} & w^{45} & w^{46} \\ w^{54} & w^{55} & w^{56} \\ w^{64} & w^{65} & w^{66} \end{pmatrix} \tag{8}$$

and

$$\begin{pmatrix} w^{14} & w^{15} & w^{16} \\ w^{24} & w^{25} & w^{26} \\ w^{34} & w^{35} & w^{36} \end{pmatrix} = \begin{pmatrix} w^{41} & w^{42} & w^{43} \\ w^{51} & w^{52} & w^{53} \\ w^{61} & w^{62} & w^{63} \end{pmatrix}. \tag{9}$$

We now introduce a symmetric tensor of type $(2,0)$ and a scalar and use them to express Conditions (8) and (9) in a simpler form. They are defined by using local coordinates as follows:

$$W^{ij} = g_{lk} W^{lijk}, \tag{10}$$

$$W = g_{ij} W^{ij}. \tag{11}$$

We now assert that Conditions (8) and (9) are equivalent to the following condition:

$$W^{ij} = \tfrac{1}{4} W \cdot g^{ij}. \tag{12}$$

Proof of the assertion: It is enough to prove the result in a suitable coordinate chart. We use the chart such that g = diagonal $(-1,-1,-1,+1)$. The components W^{ij} and W are then given in terms of $w^{\alpha\beta}$ by the following:

$$W^{11} = -w^{22} - w^{33} + w^{44}, \tag{13a}$$

$$W^{12} = w^{21} + w^{45}, \tag{13b}$$

$$W^{13} = w^{31} + w^{46}, \tag{13c}$$

$$W^{14} = w^{35} - w^{26}, \tag{13d}$$

$$W^{22} = -w^{11} - w^{33} + w^{55}, \tag{13e}$$

$$W^{23} = w^{32} + w^{56}, \tag{13f}$$

$$W^{24} = -w^{34} + w^{16}, \tag{13g}$$

$$W^{33} = -w^{11} - w^{22} + w^{66}, \tag{13h}$$

$$W^{34} = w^{24} - w^{15}, \tag{13i}$$

$$W^{44} = -w^{44} - w^{55} - w^{66}, \tag{13j}$$

where $W^{ij} = W^{ji}$ and

$$W = 2[w^{11} + w^{22} + w^{33} - w^{44} - w^{55} - w^{66}]. \tag{13k}$$

Using formula (13) it is easy to verify that Conditions (8) and (9) imply Condition (12). Conversely, if Condition (12) is given, then from (13d), (13g), and (13i), we get Condition (9).

Equation (13a), (13e), (13h), and (13j) lead to

$$\begin{aligned} w^{11} - w^{22} - w^{33} &= -w^{44} + w^{55} + w^{66}, \\ -w^{11} + w^{22} - w^{33} &= w^{44} - w^{55} + w^{66}, \\ -w^{11} - w^{22} + w^{33} &= w^{44} + w^{55} - w^{66}, \\ w^{11} + w^{22} + w^{33} &= -w^{44} - w^{55} - w^{66}. \end{aligned}$$

These imply

$$w^{11} = -w^{44}, \quad w^{22} = -w^{55}, \quad w^{33} = -w^{66}.$$

These results and Eqs. (13b), (13c), and (13f) lead to Condition (8). This proves the assertion.

By using (7), (10), (11), Condition (12) can be written as

$$\begin{aligned} &R^{ij} + T^{ij} + \tfrac{1}{2} T g^{ij} = \tfrac{1}{4}(R + 3T) g^{ij}, \text{ i.e.,} \\ &R^{ij} - \tfrac{1}{4} R g^{ij} = -T^{ij} + \tfrac{1}{4} T g^{ij}. \end{aligned} \tag{14}$$

We can rewrite (14) as

$$R^{ij} - \tfrac{1}{2} R g^{ij} = -T^{ij} + \tfrac{1}{4}(T - R) g^{ij}. \tag{15}$$

Thus Eqs. (14) and (15) are equivalent to

Condition M: $WJ = JW$.

Now the condition that the trace of the source tensor equals the scalar curvature can be written as

$$T - R = 0. \tag{16}$$

Thus under the conditions of the theorem, Condition M is equivalent to the Einstein field equations (2).

We observe that if Condition (16) is dropped from the statement of the theorem and Eq. (15) is satisfied, then the Bianchi identities and the fact that the source tensor is divergence-free, imply

$$T - R = \text{const.} \tag{16'}$$

Thus Condition (16) fixes the arbitrary constant that occurs in Eq. (16'). This completes the proof of the theorem.

5. A NEW DEFINITION OF GRAVITATIONAL FIELD

Theorem 1, proved in the preceding section enables us to give a new definition of a gravitational field. The most striking feature of the new definition is that it makes use of a linear transformation of the space of second-order differential forms on the carrier manifold. The field equations appear as a commutation condition. The linearity introduced here is not merely a formal result. The new definition leads directly to the study of an important feature of gravitational fields, namely their classification. This problem is treated in detail in Secs 7, 8, and 9. We now state the new definition of a gravitational field.

Definition 6: A gravitational field F is a triple (M, g, T) satisfying the following conditions:

GF-1: M is a pseudo-Riemannian manifold with fundamental tensor g of signature $(-,-,-,+)$ and the Levi-Civita connection Γ_g.

GF-2: T is a symmetric, divergence-free tensor of type $(2,0)$ on M.

GF-3: W commutes with J; i.e.,

$$[W, J] = 0, \tag{17}$$

where $W = R + g \times_C T$ and J is the Hodge star operator.

The last condition may be stated alternatively as

GF-3′: W induces a complex linear transformation of $\Lambda_x^2(M)$, regarded as a complex vector space by using the complex structure defined by J, for each $x \in M$.

We observe that the new definition leads to Einstein's field equations with the cosmological constant λ. This follows by writing the constant in (16′) as -4λ and rearranging Eq. (15) in the form

$$R^{ij} - \tfrac{1}{2}Rg^{ij} + \lambda g^{ij} = -T^{ij}. \tag{18}$$

Equation (18) is the Einstein field equation with the cosmological constant λ.

To obtain the field equation without the cosmological constant, we need only modify Condition GF-2, by requiring that the trace of the source tensor T be equal to the scalar curvature. This requirement which is a consequence of Einstein's field equation (2) is necessary here since it does not follow from Eq. (17).

We conclude this section by giving yet another condition equivalent to GF-3 and GF-3′. We define the submanifolds D_+ and D_- of $\Lambda_x^2(M)$ as follows:

$$D_+ = \{\xi \in \Lambda_x^2(M) \mid G(\xi, \xi) = +1 \text{ and } \xi \wedge \xi = 0\}.$$

D_- is defined similarly. We define the real-valued function f on $D_+ \cup D_-$ by

$$f(\xi) = \sigma(\xi) G(W\xi, \xi), \tag{19}$$

where $\sigma(\xi) = +1$ (if $\xi \in D_+$) or -1 (if $\xi \in D_-$) and $W = R + g \times_C T$.

The real valued function f^* on $D_+ \cup D_-$ is defined by

$$f^*(\xi) = f(J\xi), \tag{20}$$

where J is the Hodge star operator on $\Lambda_x^2(M)$. This definition makes sense in view of the fact that

$$J(D_+ \cup D_-) = D_+ \cup D_-. \tag{21}$$

We now assert that Condition GF-3 (or GF-3′) is equivalent to the following condition:

$$\text{GF-3}'': f = f^*. \tag{22}$$

Proof of the assertion: Using Eqs. (19) and (20), we can write Eq. (22) as

$$\sigma(\xi) G(W\xi, \xi) = \sigma(J\xi) G(WJ\xi, J\xi). \tag{23}$$

We note that J has the following properties:

$$\sigma(\xi) = -\sigma(J\xi)$$

and

$$G(\xi, J\eta) = G(J\xi, \eta).$$

Therefore, Eq. (23) is equivalent to

$$-G(W\xi, \xi) = G(JWJ\xi, \xi). \tag{24}$$

It is easy to see that (24) and the fact that JWJ and $-W$ are both of curvature type, leads to their equality following a standard argument, see, for example, Kobyashi and Nomizu, Chap. V.[6]

The condition $JWJ = -W$, is equivalent to

$$J^2WJ = -JW$$

or

$$WJ = JW, \quad \text{using } J^2 = -I.$$

The function f introduced here seems to have deep singificance so far as the gravitational fields are concerned. The last condition tells us that by studying f, we can determine whether the manifold M is a carrier of a gravitational field; if it is a carrier of a gravitational field, then f also tells us what type of field it carries.

6. EINSTEIN SPACES

Starting from Einstein's field equations and setting T (the energy-momentum tensor) equal to zero, we are led to the so-called vacuum field equation

$$R^{ij} = 0. \tag{25}$$

The spaces satisfying Eq. (25) belong to a class of spaces called Einstein spaces. A pseudo-Riemannian manifold M is called an Einstein space if its Ricci tensor and the scalar curvature satisfy

$$R^{ij} = \tfrac{1}{4}Rg^{ij}. \tag{26}$$

The scalar curvature R is then necessarily a constant.

Arbitrary Einstein spaces do not occur in the usual discussion of Einstein's field equations. However, an examination of the proof of Theorem 1 shows that the methods used there lead directly to the study of arbitrary Einstein spaces. In fact, putting the source tensor T equal to zero in Eq. (12) or Eq. (14), we obtain Eq. (26). But Condition GF-3 is equivalent to Eq. (12); therefore, writing T equal to zero in this condition we obtain the following characterization of Einstein spaces:

M is an Einstein space if and only if $[R, J] = 0$, where R is the Riemann-Christoffel curvature tensor.

Treating Conditions GF-3′ and GF-3″ similarly, we obtain other characterizations of Einstein spaces. We collect these results together to obtain the following theorem due to Thorpe.[4]

Theorem 2: The following conditions are equivalent:

(1) M is an Einstein space;
(2) the Riemann-Christoffel tensor R satisfies $[R, J] = 0$;

(3) The Riemann–Christoffel tensor R induces a complex linear transformation of $\Lambda_x^2(M)$—regarded as a complex vector space by using the complex structure defined by J;
(4) the functions f and f^* defined by Eqs. (19) and (20), respectively, with W replaced by R, satisfy $f = f^*$.

7. THE PROBLEM OF CLASSIFICATION

The theory of classification of gravitational fields is easily the most significant development in the theory of gravitation in recent years. Foundations of the theory of classification of gravitational fields were laid down by Petrov more than twenty years ago and his excellent account of the theory is now available in English.[7] The literature on the theory of classification is quite extensive and the problem of classification has been studied from different viewpoints by Debever,[8] Penrose,[9] Synge,[10] and others. The importance of the theory of classification for gravitational radiation was first discussed by Pirani[11] and has since been investigated by several people (see, e.g., Pirani's article in Ref. 12 and Sachs[13]).

We use the new formulation of the field equations of gravitation to study the problem of classification of gravitational fields.

8. ALGEBRAIC THEORY OF CLASSIFICATION

The problem of classification was initially discussed for Einstein spaces and in particular for the case of vacuum fields by Petrov.[7] He used the algebraic properties of the Riemann–Christoffel curvature tensor and introduced the notion of replacing pairs of suffixes by single suffixes with a different range. This formal procedure led him to introduce a 6-dimensional space in which the Riemann–Christoffel curvature tensor is mapped as a symmetric tensor of order two. The classification then proceeds by obtaining distinct canonical reductions of the 6×6 symmetric matrix of the components of this symmetric tensor. In the case of general gravitational fields, the source tensor (energy-momentum tensor) must be taken into account in carrying out the classification. This is done by replacing the Riemann–Christoffel curvature tensor by another tensor constructed by using the source tensor in addition to the Riemann–Christoffel curvature tensor. For example, the tensor W defined by Eq. (7) is such a tensor. The Weyl tensor is another example, differing from the tensor W by a multiple of $g \times_C g$, when the field equations are satisfied. For the purpose of classification it is immaterial whether we use the tensor W or the Weyl tensor.

Our new formulations of the field equations of gravitation make use of the linear transformation of $\Lambda_x^2(M)$ induced by W. We, therefore, use this result as the basis for the classification of gravitational fields. In the standard discussion of classification, use is made of the quadratic form Q which is related to the linear transformation W by

$$Q(v) = G(Wv, v) \text{ for } v \in \Lambda_x^2(M), \tag{27}$$

where W stands for either the Weyl tensor or the tensor defined by Eq. (7). We observe that the matrix of W with respect to the basis b^i of $\Lambda_x^2(M)$ is not symmetric when (M, g, T) is a gravitational field, but that the matrix of Q is symmetric in this case.

Now let (M, g, T) be a gravitational field. Then Conditions GF-3 and GF-3′ are satisfied. Condition GF-3 allows us to write the matrix of W with respect to the basis b^i of $\Lambda_x^2(M)$ in the form

$$\begin{pmatrix} A & B \\ -B & A \end{pmatrix}, \tag{28}$$

where A, B are real 3×3 matrices (see Lemma 3).

Then it is easy to verify that there exists a basis of the complex vector space $\Lambda_x^2(M)$, such that the matrix K of the complex linear transformation of $\Lambda_x^2(M)$ induced by W is given by

$$K = A + iB. \tag{29}$$

If x is the eigenvalue of K corresponding to an eigenvector v, then

$$Kv = xv. \tag{30}$$

The characteristic equation for the determination of the complex eigenvalues x of K is given by

$$\det(K - xI) = 0, \tag{31}$$

where I is the unit 3×3 matrix. Equation (31) is a cubic equation and we denote its roots by x_1, x_2, and x_3. The classification is now carried out according to the Penrose diagram. (See for example, Marathe.[14])

9. GEOMETRIC THEORY OF CLASSIFICATION

The field equations of gravitation are formulated in Condition GF-3″ of Sec. 5 by using the function f defined on $D_+ \cup D_-$ by Eq. (19). In this section we show that the general problem of classification can be solved by studying the critical points of this function f.

We recall that using the complex structure defined by J, $\Lambda_x^2(M)$ can be made into a complex, three-dimensional vector space, with multiplication by complex numbers defined by

$$(a + ib)v = av + bJv, \quad v \in \Lambda_x^2(M), \quad a, b \text{ real numbers.} \tag{32}$$

If (M, g, T) is a gravitational field, then condition GF-3″ of Sec. 5 is satisfied. Using Eq. (32) we have, therefore, the following result:

v is a eigenvector of the complex transformation induced by W with eigenvalue $a + ib$ if and only if v satisfies the equation

$$Wv = av + bJv, \tag{33}$$

where W is regarded as a real transformation of the real vector space $\Lambda_x^2(M)$.

The relation between the critical points of the function f and the eigenvectors of the transformation W is given by the following theorem. Its proof is similar to a lemma due to Thorpe[4] and is, therefore, omitted.

Theorem 3: Let (M, g, T) be a gravitational field. Then the following conditions are equivalent:

C-1: v is a critical point of the function f;

C-2: Jv is a critical point of the function f;

C-3: v is an eigenvector of the complex transformation W; i.e.,

$$Wv = xv,$$

where x is the eigenvalue corresponding to the eigenvector v.

Theorem 3 establishes the correspondence between nonnull eigenvectors of W and the critical points of the function f, since any nonnull eigenvector of W can be suitably normalized to give an eigenvector of W belonging to $D_+ \cup D_-$.

The Petrov class of the gravitational field (M, g, T) is determined by the number of critical points of the function f. The relation of the Petrov classes to the various types of the Penrose diagram and to the number of critical points of the function f is given below. Critical points are counted by equivalence classes in D_+ (either v or Jv is in D_+), where v is equivalent to w if v and w determine the same 2-plane.

Petrov class one corresponds to type I on the Penrose diagram, and the gravitational field (M, g, T) belongs to this class if the number of critical points of the function f is three. Petrov class two corresponds to the types II and D on the Penrose diagram and the gravitational field (M, g, T) belongs to this class if the function f has only one critical point. Petrov class three corresponds to the types III, N and O on the Penrose diagram and the gravitational field belongs to this class if the function f has infinitely many critical points (type O) or if it has no critical points, or in the terminology of Morse and Cairns[15] if the function f has only ordinary points (types III and N).

10. CONCLUSION

The modern version of Einstein's definition of a gravitational field has been used here. Tensors of curvature type and the curvature product of symmetric tensors have been defined. The interaction tensor was defined as the curvature product of the fundamental tensor and the energy-momentum tensor. W, the tensor obtained by coupling the Riemann and interaction tensors, was used to obtain a characterization of gravitational fields. The linear transformation of the space of second order differential forms, induced by W, was used to give a new definition of a gravitational field. Thorpe's theorem characterizing Einstein spaces was then obtained as a corollary.

We have examined the mathematical foundations of the theory of classification of gravitational fields. The real six-dimensional space used by Petrov[7] has been shown to arise naturally as the space of second-order differential forms at each point of the manifold M acting as the carrier manifold of the gravitational field (M, g, T). A geometrical interpretation of classification is obtained by applying critical point theory to a function f defined on a suitable manifold.

We remark that the function f may be regarded as defined over the Grassmann manifold of nondegenerate tangent two planes. It reduces to the sectional curvature function in the case when the source tensor is zero. We, therefore, call f the gravitational sectional curvature function. It seems to have deep significance so far as gravitational fields are concenred. The gravitational sectional curvature function f gives all the information regarding the gravitational field to which it corresponds. In terms of f, the field equations can be written as $f = fJ$, and the Petrov class of the field is determined by the number of critical points of f.

Conditions GF-3, GF-3′, and GF-3″ are seen to be fundamental for the study of classification. These conditions remain equivalent if certain conditions of Theorem 1 are relaxed, but they are then not equivalent to Einstein's field equations. It has been found that the study of these conditions leads to generalized field equations of gravitation. These results are reported in another paper.

ACKNOWLEDGMENTS

A part of this paper was presented at the annual meeting of the American Mathematical Society in January, 1971 at Atlantic City, New Jersey. The author's thanks are due to Professor Thorpe for discussions at this meeting which led to the inclusion of Sec. 6. The author is also grateful to his research advisor, Professor William F. Eberlein of the University of Rochester, for his assistance during the research for this paper. The author wishes to thank the Consiglio Nazionale Ricerche, for offering him a Visiting Professorship in Mathematics at the Istituto Matematico "Ulisse Dini", Firenze, during the summer of 1971, when this paper was written. Thanks are also due to Rettore dell'Universita di Firenze, Professor Giorgio Sestini, Direttore dell'Istituto Matematico, Professor Roberto Conti, and Professor Marco Modugno for their hospitality and kindness in making my stay very pleasant. Thanks are due to the referee for his suggestions and comments which have strengthened this paper.

[1] E. Cartan, Articles 56, 57, 59, and 80 in *Oeuvres completes* (Gauthier-Villars, Paris, 1955), Partie II, Vols. 1 and 2.
[2] C. Ehresmann, "Les connexions infinitesimales dans un espace fibre differentiable," Colloque de topologie, Bruxelles (1950).
[3] J. Ehlers, Proc. C.I.M.E. 1970, Italy, Relativistic Fluid Dynamics. 301-88, Edizioni Cremonese, Roma 1971.
[4] J. A. Thorpe, J. Math. Phys. **10**, 1 (1969).
[5] S. Kobayashi and D. Nomizu, *Foundations of Differential Geometry* (Interscience, New York, 1969), Vol. II, Chap. IX.
[6] Reference 5, Vol. I (second printing, 1964).
[7] A. Z. Petrov, *Einstein Spaces* (Pergamon, New York, 1969).
[8] R. Debever, C.R. Acad. Sci. (Paris) **249**, 1324 (1959).
[9] R. Penrose, Ann. Phys. (N.Y.) **10**, 171 (1960).
[10] J. L. Synge, Commun. Dublin Inst. Adv. Stud. A (15), ,(1964).
[11] F. A. E. Pirani, Phys. Rev. **105**, 1089 (1957).
[12] F. A. E. Pirani, in *Gravitation: An Introduction to Current Research* (Wiley, New York, 1962).
[13] R. K. Sachs, "Gravitational Radiation," in *Relativity Groups and Topology* (Gordon and Breach, New York, 1964).
[14] K. B. Marathe, Proc. Natl. Inst. Sci. India A **34**, 6 (1968).
[15] M. Morse and S. Cairns, *Critical Point Theory in Global Analysis and Differential Topology—An Introduction* (Academic, New York, 1969).

Measure theoretic convergences of observables and operators

Stanley P. Gudder and Harry C. Mullikin
Department of Mathematics, University of Denver, Denver, Colorado 80210
Department of Mathematics, Pomona College, Claremont, California 91711
(Received 12 July 1972)

Definitions of different types of measure theoretic convergence for observables and operators are given. In particular we define convergence in measure, almost everywhere, everywhere, almost uniformly, and uniformly. These types of convergence are compared and characterized. Furthermore, our theory is compared to that of Segal–Stinespring. Convergence theorems such as a bounded convergence theorem, Fatou's lemma, and a special case of Egoroff's theorem are proved. We show that the general Egoroff's theorem does not hold.

1. INTRODUCTION

In Ref. 1 Varadarajan has pointed out that Mackey's formulation of quantum mechanics[2] gives a direct generalization of the probability theory of Kolmogorov. In this generalized theory the σ algebra of subsets of a set, basic to the Kolmogorov formulation of probability theory, is replaced by a less richly endowed algebraic structure called a logic. The random variables or observables on the logic structure are the σ homomorphisms from the Borel subsets of the real line into the logic.

In addition to the probabilistic aspects of this theory, logics also give an abstract generalization of Hilbert space theory. Among the most distinguished examples of logics are those consisting of all the closed subspaces of a Hilbert space. For this case the observables on the logic correspond to the self-adjoint operators on the Hilbert space.

A considerable amount of work has been done toward the development of the probabilistic aspects of this thory.[3–8] In Refs. 5 and 9 Segal and Stinespring have developed a theory of measure theoretic convergences of operators on a Hilbert space. However, in their work the "measures" used differ radically from those of the Mackey formulation. Since measure theoretic convergences of random variables play a significant role in the conventional probability theory, an extension of these concepts to the generalized theory is quite desirable. In addition to being of interest in a purely mathematical sense, measure theoretic convergences of observables could serve as a useful tool in quantum mechanics.

2. DEFINITIONS AND PRELIMINARY RESULTS

In this paper L will denote a *sum logic*. That is, L is a quite full orthocomplemented σ lattice in which every pair of bounded observables has a unique sum. We refer the reader to Refs. 4 and 10 for these definitions and any others that are omitted in the sequel. We denote the set of all states on L by M and the set of bounded observables on L by X. X is a normed linear space under the spectral norm $\|x\| = \sup\{|\lambda| : \lambda \in \sigma(x)\}, x \in X$.

Let (Ω, F, m) be a measure space with $m(\Omega) = 1$, and let $f, f_1, f_2, \cdots$ be F measurable real-valued functions on Ω. Observe that $f_n \to f$ in measure $[m]$ if and only if $\lim m[(f_n - f)^{-1}([-\epsilon, \epsilon])] = 1$ for every $\epsilon > 0$ as $n \to \infty$. We can similarly characterize everywhere and almost everywhere convergence in terms of $(f_n - f)^{-1}$. For $f_n \to f$ everywhere if and only if for every $\epsilon > 0$, $\lim_{n\to\infty} \inf (f_n - f)^{-1}([-\epsilon, \epsilon]) = \cup_{k=1}^{\infty} \cap_{n=k}^{\infty} (f_n - f)^{-1}([-\epsilon, \epsilon]) = \Omega$, and $f_n \to f$ a.e. $[m]$ if and only if $m[\lim_{n\to\infty} \inf(f_n - f)^{-1}([-\epsilon, \epsilon])] = 1$ for every $\epsilon > 0$. Now for $\Lambda \in F, f_n \to f$ uniformly on Λ if and only if for every $\epsilon > 0$ there exists N such that $n \geq N$ implies that $(f_n - f)^{-1}([-\epsilon, \epsilon]) \supseteq \Lambda$. Finally $f_n \to f$ almost uniformly $[m]$ if and only if for $\epsilon > 0$ there exists $\Lambda \in F$ such that $m(\Lambda') \leq \epsilon$ and $f_n \to f$ uniformly on Λ. Since observables correspond to inverses of measurable functions, we are led to the following definitions for measure theoretic convergences of observables.

In the sequel $x, x_1, x_2, \cdots$ will denote elements of X; $m, m_1, m_2, \cdots$ elements of M; and $a, a_1, a_2, \cdots$ elements of L. We define $\limsup a_n = \wedge_{k=1}^{\infty} \vee_{n=k}^{\infty} a_n$ and $\liminf a_n = \vee_{k=1}^{\infty} \wedge_{n=k}^{\infty} a_n$. We say that x_n converges to x *in measure* $[m]$ if for every $\epsilon > 0$, $\lim_{n\to\infty} m((x_n - x)[-\epsilon, \epsilon]) = 1$ as $n \to \infty$; x_n converges to x *almost everywhere* $[m]$ (a.e. $[m]$) if for every $\epsilon > 0$ $m(\liminf (x_n - x)[-\epsilon, \epsilon]) = 1$; x_n converges to x *everywhere* if $\liminf (x_n - x)[-\epsilon, \epsilon] = 1$ for every $\epsilon > 0$; x_n converges *uniformly on* a if for every $\epsilon > 0$ there exists N such that $n \geq N$ implies $(x_n - x)[-\epsilon, \epsilon] \geq a$; x_n converges to x *uniformly if* x_n converges to x *uniformly* on L; x_n converges to x *almost uniformly* $[m]$ (a.u. $[m]$) if for every $\epsilon > 0$ there exists $a \in L$ such that $m(a') \leq \epsilon$ and x_n converges to x uniformly on a.

For an observable $x \in X$ and $1 \leq p < \infty$ we let $|x|^p$ denote the observable $|x|^p = h_p(x)$, where $h_p(\lambda) = |\lambda|^p$ for $\lambda \in \mathrm{R}$. Then $|x|^p(E) = x(h_1^{-1}(E)) = x(\{\lambda : |\lambda|^p \in E\})$, for example. For $x \in X$ and $m \in M$, we let m_x denote the measure $m_x(E) = m(x(E))$ for $E \in B(R)$. We define $\|x\|_p^{(m)} = m(|x|^p)^{1/p} = [\int_R |\lambda|^p m_x(d\lambda)]^{1/p}$ for $1 \leq p < \infty$ and $\|x\|_\infty^{(m)} = \inf\{r \geq 0 : m_x([-r, r]) = 1\}$. When there is no chance for confusion we simple write $\|\cdot\|_p$ for $\|\cdot\|_p^{(m)}$. We say that x_n converges to x *in mean* p $[m]$ if $\|x_n - x\|_p^{(m)} \to 0$ as $n \to \infty$.

Although we are using the conventional p norm notation, $m(|x|^p)^{1/p}$ is not, in general, a norm. In order to see this, let us consider $L = P(R^2)$, the logic of closed subspaces of real Euclidean 2-space.

Let $\phi = (1, 0)$ and $m(P) = \langle \phi, P\phi \rangle$ for $P \in L$. Let

$$A = \begin{pmatrix} 1 & 0 \\ 0 & -2 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

A straightforward computation shows that $m(|A + B|) = 8/\sqrt{13} > 2 = m(|A|) + m(|B|)$. One can easily verify that $\|A + B\|_\infty = \frac{1}{2}(1 + \sqrt{13}) > \frac{3}{2} = \|A\|_\infty + \|B\|_\infty$.

Let $x_1, x_2, \cdots$ be mutually comptaible observables. By a well-known representation theorem (Ref. 1, Theorem 3.4), there exists a measurable space (Ω, F), a σ homomorphism h of F onto L, and F measurable real-valued functions $f_1, f_2, \cdots$ such that $x_i(E) = h(f_i^{-1}(E))$ for every $E \in B(R)$. Using this representation theorem one can

use standard measure theoretic techniques to prove results concerning single observables and sequences of compatible observables. For example, using this method it is straightforward to show that for compatible observables the convergences we have defined do, in fact, generalize those of function theory.

Theorem 2.1: Let $x, x_1, x_2, \cdots$ be mutually compatible. Let (Ω, F) be a measurable space, h a σ homomorphism of F onto L, and $f, f_1, f_2, \cdots$ real-valued functions such that $x_i(E) = h(f_i^{-1}(E))$, $x(E) = h(f^{-1}(E))$ for every $E \in B(R)$. For $m \in M$ let m_h denote the measure $m_h(\Lambda) = m(h(\Lambda))$ on (Ω, F). Then $x_n \to x$ in measure $[m]$, a.e. $[m]$, a.u. $[m]$, in mean $p[m]$, respectively, if and only if $f_n \to f$ in measure $[m_h]$, a.e. $[m_h]$, a.u. $[m_h]$, in mean $p[m_h]$, respectively. Also $x_n \to x$ everywhere (uniformly) if and only if there is a $\Lambda \in F$ such that $h(\Lambda') = 0$ and $f_n \to f$ everywhere (uniformly) on Λ. Finally $x_n \to x$ uniformly on $a \in L$ if and only if there is a $\Lambda \in F$ such that $h(\Lambda) \geq a$ and $f_n \to f$ uniformly on Λ. If h is a σ isomorphism then $x_n \to x$ everywhere (uniformly) if and only if $f_n \to f$ everywhere (uniformly).

Although $\|\cdot\|_p$ is not in general a norm, using the above representation theorem, the following result is easily obtained.

Lemma 2.1: (i) For $1 \leq p \leq \infty$ and $x \in X$, $\|x\| = \sup\|x\|_p^{(m)}; m \in M$. (ii) For $x \in X$ and $m \in M$, $\|x\|_\infty = \lim\|x\|_p$ as $p \to \infty$.

If $0 < \lambda_i < 1$ and $\sum \lambda_i = 1$, the state $m = \sum \lambda_i m_i$ defined by $m(a) = \sum \lambda_i m_i(a)$ is called a *mixture* of the m_i. A state which is not a mixture is *pure*. We now investigate how convergence with respect to m is related to convergence with respect to m_i.

Lemma 2.2: If $m = \sum \lambda_i m_i$ is a mixture, then $x_n \to x$ (i) in measure $[m]$, (ii) a.e. $[m]$, (iii) a.u. $[m]$, respectively, if and only if $x_n \to x$ (i) in measure $[m_i]$, (ii) a.e. $[m_i]$, (iii) a.u. $[m_i]$, respectively, for every i.

Proof: It is trivial to show that for any of the convergences $x_n \to x$ if and only if $x_n - x \to 0$. Therefore, without loss of generality, let $x = 0$. Observe that for $a \in L$

$$m_i(a) \leq (1/\lambda_i)\sum \lambda_j m_j(a) = (1/\lambda_i)m(a). \tag{1}$$

(i) Suppose $x_n \to 0$ in measure $[m]$. Let $\epsilon > 0$ be given. Then $\lim_{n\to\infty} m(x_n[-\epsilon, \epsilon]') = \lim_{n\to\infty} 1 - m(x_n[-\epsilon, \epsilon]) = 0$. Now by (1) for every i $0 \leq m_i(x_n[-\epsilon, \epsilon]') \leq (1/\lambda_i)$ $m(x_n[-\epsilon, \epsilon]') \to 0$ as $n \to \infty$. Therefore for every i, $\lim m_i(x_n[-\epsilon, \epsilon]') = 0$ as $n \to \infty$. Hence

$$\lim_{n\to\infty} 1 - m_i(x_n[-\epsilon, \epsilon]) = \lim_{n\to\infty} m_i(x_n[-\epsilon, \epsilon]') = 0,$$

so that $x_n \to 0$ in measure $[m_i]$ for every i.

Suppose $x_n \to 0$ in measure $[m_i]$ for every i. Then for $\epsilon > 0$, $\lim m_i(x_n[-\epsilon, \epsilon]) = 1$ as $n \to \infty$. Now for each i and each n, $|\lambda_i m_i(x_n[-\epsilon, \epsilon])| \leq \lambda_i$ so that since $\sum \lambda_i < \infty$:

$$\lim_{n\to\infty} \sum_i \lambda_i m_i(x_n[-\epsilon, \epsilon]) = \sum_i \lambda_i \lim_{n\to\infty} m_i(x_n[-\epsilon, \epsilon]) = 1.$$

Thus for $\epsilon > 0$, $\lim_{n\to\infty} m(x_n[-\epsilon, \epsilon]) = 1$ and hence $x_n \to 0$ in measure $[m]$.

(ii) Trivially $m(a) = 1$ if and only if $m_i(a) = 1$ for each i. (ii) is then immediate.

(iii) Suppose $x_n \to 0$ a.u. $[m]$. Let $\epsilon > 0$ be given and fix i. Now there exists $a \in L$ such that $m(a') \leq \epsilon\lambda_i$ and $x_n \to 0$ uniformly on a. Then by (1) $m_i(a') \leq (1/\lambda_i)$ $m(a') \leq \epsilon$. Therefore, $x_n \to 0$ a.u. $[m_i]$.

Suppose $x_n \to 0$ a.u. $[m_i]$ for every i. Let $\epsilon > 0$ be given. Choose K such that $\sum_{K+1} \lambda_i \leq \epsilon/2$. For $i = 1, 2, \ldots, K$ there exists $a_i \in L$ such that $m_i(a_i') \leq \epsilon/2K\lambda_i$ and $x_n \to 0$ uniformly on a_i. Let $a = a_i \vee a_2 \vee \cdots \vee a_K$. Then for $i = 1, 2, \ldots, K$, $a' = a_1' \wedge a_2' \wedge \cdots \wedge a_K' \leq a_i'$ so we have $m_i(a') \leq m_i(a_i') \leq \epsilon/2K\lambda_i$. Therefore $m(a') = \sum_1^K \lambda_i m_i(a') + \sum_{K+1} \lambda_i m_i(a') \leq \epsilon$. Now for $\delta > 0$, since $x_n \to 0$ uniformly on a_i, there exists N_i such that $n \geq N_i$ implies that $x_n([-\delta, \delta]) \geq a_i$. Let $N = \max\{N_1, N_2, \ldots, N_K\}$. Then if $n \geq N$, for $i = 1, 2, \ldots, K$, $x_n([-\delta, \delta]) \geq a_i$ so that $x_n([-\delta, \delta]) \geq a_1 \vee a_2 \vee \cdots \vee a_K = a$. Thus $m(a') \leq \epsilon$ and $x_n \to 0$ uniformly on a. Therefore $x_n \to 0$ a.u. $[m]$.

The following example shows that if $m = \sum \lambda_i m_i$ is a mixture and $x_n \to x$ in mean $p[m_i]$ for each i, then it does not necessarily follow that $x_n \to x$ in mean $p[m]$. Let L be the collection of all subsets of the positive integers with $\leq$ as set theoretic inclusion. For $i = 1, 2, \ldots$ and $E \in L$, let

$$m_i(E) = \begin{cases} 1 & \text{if } i \in E \\ 0 & \text{if } i \notin E. \end{cases}$$

Then each m_i is a state on L. Let $m = \sum_1^\infty (1/2^i)m_i$. For $n = 1, 2, \cdots$ define observables x_n on L by $\sigma(x_n) = \{0, 2^n\}$ and $x_n(\{2^n\}) = \{n\}$. Then for each i

$$m_i(x_n) = \int \lambda\, m_i(x_n(d\lambda)) = 0 \cdot m_i(x_n(\{0\})) + 2^n m_i(x_n(\{2^n\})) = 2^n m_i(\{n\}).$$

Then as $m_i(\{n\}) = 0$ if $i \neq n$ and $m_n(\{n\}) = 1$, we have

$$m_i(x_n) = \begin{cases} 2^n & \text{if } n = i \\ 0 & \text{if } n \neq i. \end{cases}$$

Therefore $m_i(|x_n|) = m_i(x_n) \to 0$ as $n \to \infty$, and hence $x_n \to 0$ in mean $1[m_i]$ for each i. Now observe that $m(\{n\}) = \sum_i (1/2i)m_i(\{n\}) = (1/2^n)m_n(\{n\}) = (1/2^n)$ for every n. Then

$$m(|x_n|) = m(x_n) = \int \lambda\, m(x_n(d\lambda)) = 0 \cdot m(x_n(\{0\})) + 2^n m(x_n(\{2^n\})) = 2^n m(\{n\}) = 2^n \cdot (1/2^n) = 1.$$

Thus $x_n \not\to 0$ in mean $1[m]$. If we let y_n be defined by $\sigma(y_n) = \{0, 1\}$ and $y_n(\{1\}) = \{n\}$, one can verify that $\|y_n\| \leq 1$ and $y_n \to 0$ in mean $\infty[m_i]$ for every i, but $y_n \not\to 0$ in mean $\infty[m]$. However, for $1 \leq p < \infty$ and $m = \sum \lambda_i m_i$ any mixture, convergence in mean $p[m]$ of a uniformly bounded sequence is equivalent to convergence in mean $p[m_i]$ for every i. In order to prove this we will need the following result.

Lemma 2.3: Let $y \in X$ and let $m = \sum \lambda_i m_i$ be a mixture. Then

(i) $m(y) = \sum \lambda_i m_i(y)$,

(ii) $\|y\|_\infty^{(m)} = \sup_i \|y\|_\infty^{(m_i)}$

Proof: (i) The measures m_{iy} are all supported on

$I = [-\|y\|, \|y\|]$ and $m_y = \sum \lambda_i m_{iy}$. Hence $\int f dm_y = \sum \lambda_i \int f dm_{iy}$ for any continuous f on I and, in particular, $m(y) = \sum \lambda_i m_i(y)$. (ii) $\{M; m_y([-M, M]') = 0\} = \{M: m_{iy}([-M, M]') = 0 \text{ for all } i\}$.

Lemma 2.4: Let $m = \sum \lambda_i m_i$ be a mixture, and let $x, x_1, x_2, \cdots \in X$. Then

(i) For $1 \leq p \leq \infty$, if $x_n \to x$ in mean $p[m]$, then $x_n \to x$ in mean $p[m_i]$ for every i.

(ii) If there exists $K < \infty$ such that $\|x_n\| \leq K$ for every n, then for $1 \leq p < \infty$, $x_n \to x$ in mean $p[m]$ if $x_n \to x$ in mean $p[m_i]$ for every i.

Proof: Without loss of generality, let $x = 0$.

(i) First let $1 \leq p < \infty$. By (i) of Lemma 2.3, $m(|x_n|^p) = \sum \lambda_i m_i(|x_n|^p)$. Then

$$0 \leq m_i(|x_n|^p) \leq (1/\lambda_i) \sum_j \lambda_j m_j(|x_n|^p) = (1/\lambda_i) m(|x_n|^p) \to 0$$

as $n \to \infty$. Therefore $x_n \to 0$ in mean $p[m_i]$ for every i. By (ii) of Lemma 2.3, $\|x_n\|_\infty^{(m)} = \sup_i \|x_n\|_\infty^{(m_i)}$. Then for every i, $0 \leq \|x_n\|_\infty^{(m_i)} \leq \|x_n\|_\infty^{(m)} \to 0$ as $n \to \infty$, so that $x_n \to 0$ in mean $\infty[m_i]$ for every i.

(ii) By Lemma 2.1, $\|x_n\|_p^{(m_i)} \leq \|x_n\| \leq K$ so that $m_i(|x_n|^p) \leq K^p$. Then for every n and i, $|\lambda_i m_i(|x_n|^p)| \leq \lambda_i K^p$. Therefore since $\lim m_i(|x_n|^p) = 0$ as $n \to \infty$ and $\sum_i \lambda_i K^p < \infty$, we have $\lim_{n\to\infty} m(|x_n|^p) = \lim_{n\to\infty} \sum_i \lambda_i m_i(|x_n|^p) = \sum \lambda_i \lim_{n\to\infty} m_i(|x_n|^p) = 0$. Thus $x_n \to 0$ in mean $p[m]$.

3. CHARACTERIZATIONS OF THE CONVERGENCES

In function theory measure theoretic convergences can be characterized in terms of characteristic functions. For example, a sequence $\{f_n\}$ of functions converges to a function f in measure $[m]$ if and only if for every $\epsilon > 0$ there is a sequence of characteristic functions χ_{E_n} such that $\lim_{n\to\infty} m(E_n) = 1$ and $E_n \subset (f_n - f)^{-1}[-\epsilon, \epsilon]$ where this last statement is equivalent to the fact that the supremum norm of $(f_n - f) \cdot \chi_{E_n}$ is less than ϵ for every n. In this section we generalize these results to convergences of observables. In the observable theory characteristic functions correspond to the so-called proposition observables which are defined to be those observables whose spectra lie in the set $\{0, 1\}$. In the logic of all closed subspaces of a Hilbert space, the proposition observables are the orthogonal projections. When our results are applied to this special case in the next section, the measure theoretic convergences of operators are seen to be closely related to the theory developed by Segal and Stinespring.

Theorem 3.1: $x_n \to x$ (i) in measure $[m]$, (ii) a.e. $[m]$, (iii) everywhere, respectively, if and only if for every $\epsilon > 0$ there is a sequence of propositions a_n such that $a_n \leq (x_n - x)[-\epsilon, \epsilon]$ and (i) $\lim_{n\to\infty} m(a_n) = 1$, (ii) $a_n \leq a_{n+1}$, $n = 1, 2, \ldots$, and $\lim_{n\to\infty} m(a_n) = 1$, (iii) $a_n \leq a_{n+1}$, $n = 1, 2, \ldots$, and $\lim m(a_n) = 1$ as $n \to \infty$ for every $m \in M$, respectively.

Proof: Without loss of generality we assume $x = 0$. (i) For sufficiency we have $m(a_n) \leq m(x_n[-\epsilon, \epsilon])$. Thus $\lim m(x_n[-\epsilon, \epsilon]) = 1$ as $n \to \infty$ and $x_n \to x$ in measure $[m]$. For necessity let $a_n = x_n[-\epsilon, \epsilon]$. Then $\lim_{n\to\infty} m(a_n) = 1$. (ii) For sufficiency we have $m(\liminf x_n[-\epsilon, \epsilon]) \geq m(\liminf a_n) = m(\vee a_i) = \lim m(a_i) = 1$. For necessity, let $a_n = \wedge_{i=n}^\infty x_i[-\epsilon, \epsilon]$. Then $a_n \leq a_{n+1}$, $a_n \leq x_n[-\epsilon, \epsilon]$ and $\lim_{n\to\infty} m(a_n) = \lim_{n\to\infty} m(\wedge_{i=n}^\infty x_i[-\epsilon, \epsilon]) = 1$. (iii) follows from applying (ii) for all $m \in M$.

We write x_a for the proposition observable x that satisfies $x(\{1\}) = a$. Recall that x is a proposition observable if and only if $x^2 = x$ and $m(x_a) = m(a)$.

Lemma 3.1: Let $y \in X$ and let x_a be a proposition observable such that $x_a \leftrightarrow y$. Then $\|yx_a\| \leq \epsilon$ if and only if $a \leq y[-\epsilon, \epsilon]$.

This is proved by a straightforward application of the representation theorem of the previous section.

Theorem 3.2: (i) $x_n \to x$ in measure $[m]$ if and only if for every $\epsilon > 0$ there are proposition observables y_n such that $\lim_{n\to\infty} m(y_n) = 1$, $y_n \leftrightarrow x_n - x$ and $\|(x_n - x)y_n\| \leq \epsilon$. (ii) If for every $\epsilon > 0$ there are proposition observables y_n such that $y_n \leftrightarrow x_n - x$, $\|(x_n - x)y_n\| \leq \epsilon$, $y_n \leq y_{n+1}$, $n = 1, 2, \cdots$ and $m(y_n) \to 1$, then $x_n \to x$ a.e. $[m]$. If $x_n - x \leftrightarrow x_k - x$, $n, k = 1, 2, \ldots$, then the converse holds. (iii) If for every $\epsilon > 0$ there are proposition observables y_n such that $y_n \leftrightarrow x_n - x$, $\|(x_n - x)y_n\| \leq \epsilon$, $y_n \leq y_{n+1}$, $n = 1, 2, \cdots$ and $m(y_n) \to 1$ for every $m \in M$, then $x_n \to x$ everywhere. If $x_n - x \leftrightarrow x_k - x$, $n, k = 1, 2, \ldots$, then the converse holds.

Proof: This theorem follows from Lemma 3.1 and the proof of Theorem 3.1.

Theorem 3.3: $x_n \to x$ uniformly if and only if $\lim \|x_n - x\| = 0$ as $n \to \infty$.

Proof: Assume $x = 0$. Let $\epsilon > 0$. Then $\|x_n\| \leq \epsilon$ for $n \geq N$ if and only if $\sigma(x_n) \subset [-\epsilon, \epsilon]$ which holds if and only if $x_n[-\epsilon, \epsilon] = 1$ for $n \geq N$.

Theorem 3.4: (i) If for every $\epsilon > 0$ there is a proposition observable $y(\epsilon)$ such that $y \leftrightarrow x_n - x$ for every n, $m(I - y) \leq \epsilon$, and $\lim_{n\to\infty} \|(x_n - x)y\| = 0$, then $x_n \to x$ a.u. $[m]$. (ii) if $x_n - x \leftrightarrow x_k - x$ for all n, k then the converse of (i) holds.

Proof: (i) follows from Lemma 3.1. (ii) Let $\epsilon > 0$ be given. Then there is an $a \in L$ such that $m(a') \leq \epsilon$ and $x_n \to 0$ uniformly on a (we are assuming, as usual, that $x = 0$). Now for $k = 1, 2, \ldots$, there is an N_k such that $n \geq N_k$ implies $x_n[-1/k, 1/k] \geq a$. Let $b = \wedge_{k=1}^\infty \wedge_{n=N_k}^\infty x_n[-1/k, 1/k]$.

Now $b \geq a$ so $m(b') \leq m(a') \leq \epsilon$. For $\delta > 0$ choose l such that $1/l \leq \delta$. Then if $n \geq N_l$,

$$x_n[-\delta, \delta] \geq x_n[-1/l, 1/l] \geq \bigwedge_{j=N_l}^\infty x_j[-1/l, 1/l] \geq \bigwedge_{k=1}^\infty \bigwedge_{j=N_k}^\infty x_j[-1/k, 1/k] = b.$$

Therefore $m(b') \leq \epsilon$ and $x_n \to 0$ uniformly on b. Furthermore since $x_n \leftrightarrow x_k$ for every n, k, $b \leftrightarrow x_n(E)$ for all n and $E \in B(R)$. If we let $y = x_b$, then $y \leftrightarrow x_n$ for all n and $m(I - y) = m(b') \leq \epsilon$. If $\delta > 0$ there is an N such that $n \geq N$ implies $x_n[-\delta, \delta] \geq b$ and by Lemma 3.1 $\|x_n y\| \leq \delta$ for $n \geq N$. Hence $\lim \|x_n y\| = 0$ as $n \to \infty$.

In function theory a sequence f_n converges in measure $[m]$ to a function f if and only if

$$\int \frac{|f_n - f|}{1 + |f_n - f|}\, dm \to 0 \text{ as } n \to \infty.$$

We now generalize this result. We denote the nonnegative reals by R^+.

Theorem 3.5: Let $g: R^+ \to R^+$ be a strictly monotonic bounded Borel function which is continuous at 0 and $g(0) = 0$. For $x, y \in X$, let $\rho(x, y) = m(g(|x - y|))$. Then $x_n \to x$ in measure $[m]$ if and only if $\rho(x_n, y) \to 0$ as $n \to \infty$.

Proof: Since $\rho(x_n, x) = \rho(x_n - x, 0)$, without loss of generality let $x = 0$. Suppose $x_n \to 0$ in measure $[m]$. Let $\epsilon > 0$ be given. As g is continuous at 0 with $g(0) = 0$, there exists $\eta > 0$ such that $|\lambda| \leq \eta$ implies that $g(|\lambda|) \leq \epsilon/2$. If we let K be a bound for the function g, then

$$|\rho(x_n, 0)| = \int g(|\lambda|)\, m_{x_n}(d\lambda) = \int_{-\eta}^{\eta} g(|\lambda|)\, m_{x_n}(d\lambda)$$

$$+ \int_{[-\eta,\eta]'} g(|\lambda|)\, m_{x_n}(d\lambda) \leq \tfrac{1}{2}\epsilon + K \cdot m_{x_n}([-\eta, \eta]').$$

Now $x_n \to 0$ in measure $[m]$, so that $m_{x_n}([-\eta, \eta]') \to 0$ as $n \to \infty$. Therefore, there exists N such that $n \geq N$ implies that $m_{x_n}([-\eta, \eta]') \leq \epsilon/2K$. Then for $n \geq N$, $|\rho(x_n, 0)| \leq \frac{1}{2}\epsilon + K \cdot m_{x_n}([-\eta, \eta]') \leq \epsilon$. Hence, $\rho(x_n, 0) \to 0$ as $n \to \infty$.

Suppose now that $\rho(x_n, 0) \to 0$ as $n \to \infty$. Let $\epsilon > 0$ be given. Then using the fact that g is monotone increasing

$$\rho(x_n, 0) = \int g(|\lambda|)\, m_{x_n}(d\lambda) \geq \int_{[-\epsilon,\epsilon]'} g(|\lambda|)\, m_{x_n}(d\lambda)$$

$$\leq g(\epsilon)\, m_{x_n}([-\epsilon, \epsilon]').$$

Since $g(0) = 0$, $\epsilon > 0$, and g is strictly monotone, $g(\epsilon) > 0$. Therefore, $0 \leq m_{x_n}([-\epsilon, \epsilon]') \leq (1/g(\epsilon)]$ $\rho(x_n, 0) \to 0$ as $n \to \infty$. Consequently, $\lim_{n\to\infty} m(x_n([-\epsilon, \epsilon])) = 1 - \lim_{n\to\infty} m_{x_n}([-\epsilon, \epsilon]') = 1$ so that $x_n \to 0$ in measure $[m]$.

Corollary 3.1: Let $1 \leq p < \infty$. Then $x_n \to x$ in measure $[m]$ if and only if $m(|x - x_n|^p (I + |x - x_n|^p)^{-1}) \to 0$ as $n \to \infty$.

Proof: This result is immediate from Theorem 3.5 with $g(\lambda) = |\lambda|^p (1 + |\lambda|^p)^{-1}$.

4. CONVERGENCE OF OPERATORS

Let us now consider the special case in which L is the logic $P(H)$ of all closed subspaces of a separable Hilbert space H. In this case the bounded observables on L can be identified with the bounded self-adjoint operators on H. For a self-adjoint operator A on H, we let $A(\cdot)$ denote its resolution of identity. We notice that the proposition observables correspond to orthogonal projections and that two bounded self-adjoint operators A and B are compatible if and only if $AB = BA$.

Segal and Stinespring[5,9] have developed a measure and integration theory of operators on a Hilbert space. Although this theory is also motivated by certain investigations in quantum mechanics, the "measures" used are quite different from the states in the present theory. Nevertheless, we shall formally compare the two theories. First, we give the definition of a gage space which is basic to the work of Segal and Stinespring.

Let H be a separable Hilbert space. Let Λ be a ring of operators on H, i.e., Λ is an algebra of bounded everywhere defined linear operators which is closed under adjunction, is closed in the weak operator topology, and contains the identity I. A *gage* m is a nonnegative map on the projections in Λ which satisfies the following.

(i) If $\{P_n\}$ is a sequence of mutually orthogonal projections in Λ, then $m(\vee P_n) = \sum m(P_n)$.

(ii) If U is a unitary operator in Λ and P is a projection in Λ, then $m(U^* P U) = m(P)$.

(iii) If P is a projection in Λ, then there exist projections $\{P_n\} \subset \Lambda$ such that $P = \vee P_n$ and $m(P_n) < \infty$ for every n.

The system (H, Λ, m) is called a *gage space*. If $m(I) = 1$, then (H, Λ, m) is called a *probability gage space*.

We shall only consider the case when (H, Λ, m) is a probability gage space. In what follows we think of Λ as being the ring generated by $P(H)$. We emphasize that although gages and states are both "measures" on projections, they have essential differences. For example, gages are subadditive, while states are, in general, not subadditive.

Let $A, A_1, A_2, \cdots$ be operators on H. The Stinespring definition for convergence in measure is that $A_n \to A$ in measure $[m]$ if for every $\epsilon > 0$ there exist projections $P_n \in \Lambda$ such that $\|(A_n - A)P_n\| \leq \epsilon$ and $m(P_n) \to 1$. Theorem 3.2 gives a similar result for our convergence in measure. However in that theorem we require that P_n commute with $A_n - A$. This is because there is no satisfactory definition for the multiplication of two noncompatible observables. However, in the Hilbert space case the usual multiplication can be used and we can, therefore, remove this restriction.

Theorem 4.1: Let m be a state on $P(H)$. Then $A_n \to A$ in measure $[m]$ if and only if for every $\epsilon > 0$ there exist projections P_n such that $\|(A_n - A)P_n\| \leq \epsilon$ and $m(P_n) \to 1$ as $n \to \infty$.

Proof: Assume $A = 0$. By Ref. 11 there exists an orthonormal set $\{\phi_n\} \subset H$ and a set $\{\lambda_n\} \subset R$ with $0 \leq \lambda_n \leq 1$, $\sum \lambda_n = 1$ such that $m(P) = \sum \lambda_n \langle \phi_n, P\phi_n \rangle$. By (i) of Lemma 2.1, without loss of generality we can take $m(P) = \langle \phi, P\phi \rangle$ for some $\phi \in H$ with $\|\phi\| = 1$. The necessity is done by Theorem 3.2. For sufficiency, let $\epsilon > 0$ be given. Then by hypothesis there exist projections P_n such that $\|A_n P_n\| \leq \epsilon/2$ and $m(P_n) \to 1$. Now

$$|m(|A_n|(I + A_n|)^{-1})| = |\langle \phi, |A_n|(I + |A_n|)^{-1}\phi \rangle|$$

$$= |\langle \phi, |A_n|(I + |A_n|)^{-1}[P_n + (I - P_n)]\phi \rangle|$$

$$\leq |\langle \phi, |A_n|(I + |A_n|)^{-1} P_n \phi \rangle| + |\langle \phi, |A_n|(I + |A_n|)^{-1} P_n^{\perp} \phi \rangle|.$$

Now by Schwarz's inequality,

$$|\langle \phi, |A_n|(I + |A_n|)^{-1} P_n \phi \rangle| \leq \|\phi\| \cdot \|(I + |A_n|)^{-1} |A_n| P_n \phi\|$$

$$\leq \|(I + |A_n|)^{-1}\| \cdot \| |A_n| P_n \phi\| \leq \| |A_n| P_n\| \cdot \|\phi\|$$

$$= \|A_n P_n\| \leq \epsilon/2.$$

Again using Schwarz's inequality,

$$|\langle \phi, |A_n|(I+|A_n|)^{-1}P_n^{\perp}\phi\rangle| \leq \|\phi\| \cdot \| |A_n|(I+|A_n|)^{-1}P_n^{\perp}\phi\|$$

$$\leq \| |A_n|(I+|A_n|)^{-1}\| \cdot \|P_n^{\perp}\phi\| \leq \|P_n^{\perp}\phi\| = \langle P_n^{\perp}\phi, |P_n^{\perp}\phi\rangle^{1/2}$$

$$= (m(P_n^{\perp}))^{1/2}.$$

Now $m(P_n^{\perp}) = 1 - m(P_n) \to 0$ as $n \to \infty$. Therefore there exists N such that $n \geq N$ implies that $(m(P_n^{\perp}))^{1/2} \leq \frac{1}{2}\epsilon$. Then for $n \geq N$,

$$|m(|A_n|(I+|A_n|)^{-1})| \leq \|A_nP_n\| + (m(P_n^{\perp}))^{1/2} \leq \epsilon.$$

Consequently $m(|A_n|(I+|A_n|)^{-1}) \to 0$ as $n \to \infty$. Then by Corollary 3.1 $A_n \to 0$ in measure $[m]$.

Lemma 4.1: Let A be a bounded self-adjoint operator and P a projection on H. (i) If $AP = PA$, then $\|AP\| \leq \epsilon$ implies $P \leq A([-\epsilon, \epsilon])$. (ii) If $P \leq A([-\epsilon, \epsilon])$ then $\|AP\| \leq \epsilon$.

The proof of this lemma is the same as the proof of Lemma 3.1. The crucial point is that (ii) holds in the above lemma even if P does not commute with A. In the Segal–Stinespring definition of nearly everywhere convergence, $A_n \to A$ n.e. $[m]$, if for every $\epsilon > 0$ there exist projections $P_n \in \Lambda$ such that $\|(A_n - A)P_n\| \leq \epsilon$, $P_n \leq P_{n+1}$, $n = 1, 2, \ldots$, and P_n converges strongly to the identity I. Also $A_n \to A$ a.u. $[m]$ if for any $\epsilon > 0$ there is a projection $P \in \Lambda$ such that $m(P^{\perp}) < \epsilon$ and $\lim \|(A_n - A)P\| = 0$ as $n \to \infty$.

Theorem 4.2: Let m be a state on $P(H)$. (i) If $A_n \to A$ a.e. $[m]$ then for every $\epsilon > 0$ there are projections P_n such that $\|(A_n - A)P_n\| \leq \epsilon, P_n \leq P_{n+1}$, $n = 1, 2, \ldots$, and $m(P_n) \to 1$. If we require that P_n commute with $A_n - A$ then the converse holds. (ii) If $A_n \to A$ everywhere there exist projections P_n such that $\|(A_n - A)P_n\| \leq \epsilon$, $P_n \leq P_{n+1}$, $n = 1, 2, \ldots$, and $P_n \to I$ strongly. If we require that P_n commute with $A_n - A$ then the converse holds. (iii) If $A_n \to A$ a.u. $[m]$ then for any $\epsilon > 0$ there is a projection P such that $m(P^{\perp}) \leq \epsilon$ and $\lim \|(A_n - A)P\| = 0$ as $n \to \infty$.

Proof: (i) This follows from Lemma 4.1 and Theorem 3.1. The converse follows from Theorem 3.2 (ii). (ii) By Lemma 4.1 and Theorem 3.1 there are projections P_n such that $\|(A_n - A)P_n\| \leq \epsilon$, $P_n \leq P_{n+1}$, $n = 1, 2, \ldots$, and $m(P_n) = 1$ for every state m. Let $\phi \in H$ with $\|\phi\| = 1$ and let $m(P) = \langle \phi, P\phi\rangle$. Then

$$\|(I - P_n)\phi\|^2 = \langle P_n^{\perp}\phi, P_n^{\perp}\phi\rangle = \langle \phi, P_n^{\perp}\phi\rangle = m(P_n^{\perp})$$

$$= 1 - m(P_n) \to 0$$

as $n \to \infty$. Thus $P_n\phi \to \phi$ for every $\phi \in H$ and hence $P_n \to I$ strongly. The converse follows from Theorem 3.2 (iii). (iii) This follows from the definition of convergence a.u. $[m]$ and Lemma 4.1.

We now give an example to show that the converse in Theorem 4.2 (i) fails in the noncommutative case.

Let $H = R^2$, $m(P) = \langle \phi, P\phi\rangle$, where $\phi = (1, 0)$. For $n = 2, 3, \cdots$ define self-adjoint operators A_n on H by $\sigma(A_n) = \{1/n, 1\}$ and

$$A_n(\{1/n\}) = \frac{1}{1+n^2}\begin{pmatrix} n^2 & n \\ n & 1 \end{pmatrix}, \qquad A_n(\{1\}) = \frac{1}{1+n^2}\begin{pmatrix} 1 & -n \\ -n & n^2 \end{pmatrix}.$$

Then

$$A_n = \frac{1}{1+n^2}\begin{pmatrix} n+1 & 1-n \\ 1-n & (1/n)+n^2 \end{pmatrix}.$$

Let $\epsilon > 0$ be given. Let $P = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$. Then

$$A_nP = \begin{pmatrix} \frac{n+1}{1+n^2} & 0 \\ \frac{1-n}{1+n^2} & 0 \end{pmatrix} \to \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

uniformly. Therefore, there exists N such that $n \geq N$ implies that $\|A_nP\| \leq \epsilon$. For $n < N$, let $P_n = 0$ and for $n \geq N$ let $P_n = P$. Then $\|A_nP_n\| \leq \epsilon$, $P_n \leq P_{n+1}$ for every n, and $m(P_n) \to 1$.

However, we now show that $A_n \not\to 0$ a.e. $[m]$. It is easy to verify that $A_n(\{1/n\}) = LH(\{(1, 1/n)\})$, the linear subspace spanned by $(1, 1/n)$. It is then clear that for every n, $\wedge_{k=n}^{\infty} A_k[-\frac{1}{2}, \frac{1}{2}] = \wedge_{k=n}^{\infty} A_k(\{1/k\}) = 0$. Therefore, $\liminf A_n([-\frac{1}{2}, \frac{1}{2}]) = 0$ so that $A_n \not\to 0$ a.e. $[m]$.

This example also shows that the converse of Theorem 4.2 (iii) does not hold. One can give examples which show that the converse in Theorem 4.2 (ii) fails in the noncommutative case.

5. CONVERGENCE THEOREMS

In this section we investigate how the various types of convergence are interrelated and prove a generalized bounded convergence theorem and Fatou's lemma. We also consider Egoroff's theorem. The following lemma is easily proved.

Lemma 5.1: Uniform convergence implies any of the other types of convergence.

Our next result shows that if $L = P(H)$ with H a finite-dimensional Hilbert space, then everywhere convergence is equivalent to uniform convergence. We enlarge the class of logics for which this result holds through the following definition.

A logic L is said to have *finite chain condition* (f.c.c.) if $\{a_n\} \subset L$ with $a_1 \geq a_2 \geq \cdots$ implies that there exists N such that $a_n = a_N$ for $n \geq N$.

Lemma 5.2: If L has f.c.c., then everywhere convergence implies uniform convergence.

Proof: Suppose $x_n \to 0$ everywhere. Let $\epsilon > 0$ be given, and let $a_n = \wedge_{k=n}^{\infty} x_k([-\epsilon, \epsilon])$. Then $a_1' \geq a_2' \geq \cdots$. Then since L has f.c.c. there exists N such that $a_n' = a_N'$ for $n \geq N$. Then $a_n = a_N$ for $n \geq N$ so that $1 = \liminf x_n([-\epsilon, \epsilon]) = \vee_{n=1}^{\infty} a_n = a_N$. Hence if $n \geq N$, then $1 \geq x_n([-\epsilon, \epsilon]) \geq a_N = 1$. Thus $x_n([-\epsilon, \epsilon]) = 1$ for $n \geq N$, and consequently $x_n \to 0$ uniformly.

Lemma 5.3: (i) implies (ii), and (ii) implies (iii), where

(i) $x_n \to x$ everywhere;

(ii) $x_n \to x$ a.e. $[m]$;

(iii) $x_n \to x$ in measure $[m]$.

Proof: That (i) implies (ii) is obvious. That (ii) implies (iii) follows from Theorem 3.1.

In function theory if a sequence converges in measure then there exists a subsequence which converges a.e. We now show that this does not, in general, hold for observables. Let H, m, and $\{A_n\}$ be as in the example at the end of the last section. In the discussion of that example, we showed that for $\epsilon > 0$ there exist projections P_n such that $\|A_n P_n\| \leq \epsilon$ and $m(P_n) \to 1$. Then by Theorem 4.1, $A_n \to 0$ in measure $[m]$. Let $\{A_{n_k}\}$ be any subsequence. As noted earlier, $A_{n_k}(\{1/n_k\}) = LH(\{(1, 1/n_k)\})$. It is then obvious that for every l, $\wedge_{k=l}^{\infty} A_{n_k}([-\frac{1}{2}, \frac{1}{2}]) = \wedge_{k=l}^{\infty} A_{n_k}(\{1/n_k\}) = 0$. Therefore

$$m(\liminf_{k\to\infty} A_{n_k}([-\tfrac{1}{2}, \tfrac{1}{2}])) = 0$$

so that $A_{n_k} \not\to 0$ a.e. $[m]$.

Lemma 5.4: (i) For $1 \leq p \leq \infty$, if $x_n \to x$ in mean $p[m]$, then $x_n \to x$ in measure $[m]$. (ii) If $x_n \to x$ in measure $[m]$ and there exists K such that $\|x_n\| \leq K$ for every n, then $x_n \to x$ in mean $p[m]$ for $1 \leq p < \infty$.

Proof: Without loss of generality, let $x = 0$.

(i) Let $1 \leq p < \infty$. Then

$$0 \leq m(|x_n|^p(I + |x_n|^p)^{-1}) = \int \frac{|\lambda|^p}{1 + |\lambda|^p} m_{x_n}(d\lambda)$$
$$\leq \int |\lambda|^p m_{x_n}(d\lambda) = (\|x_n\|_p)^p \to 0 \quad \text{as } n \to \infty.$$

Therefore, $m(|x_n|^p(I + |x_n|^p)^{-1}) \to 0$ as $n \to \infty$, so that by Corollary 3.1 $x_n \to 0$ in measure $[m]$.

For $p = \infty$, let $\epsilon > 0$ be given. Then there exists N such that $n \geq N$ implies that $\|x_n\|_\infty = \inf\{r > 0 : m(x_n([-r, r])) = 1\} < \epsilon$. But then for $n \geq N$, $m(x_n([-\epsilon, \epsilon])) = 1$. Hence, $x_n \to 0$ in measure $[m]$.

(ii) Now $\|x_n\| \leq K$ implies that $m_{x_n}([-K, K]') = 0$ for every n. Then $(\|x_n\|_p)^p = \int_{-K}^{K} |\lambda|^p m_{x_n}(d\lambda)$. For

$$|\lambda| \leq K, \quad \frac{|\lambda|^p}{1 + K^p} \leq \frac{|\lambda|^p}{1 + |\lambda|^p}$$

so that $|\lambda|^p \leq (1 + K^p)[|\lambda|^p/(1 + |\lambda|^p)]$. Then using Corollary 3.1 we have

$$(\|x_n\|_p)^p \leq (1 + K^p)\int_{-K}^{K} \frac{|\lambda|^p}{1 + |\lambda|^p} m_{x_n}(d\lambda) = (1 + K^p)$$
$$m(|x_n|^p(I + |x_n|^p)^{-1}) \to 0$$

as $n \to \infty$. Therefore, $\|x_n\|_p \to 0$ as $n \to \infty$, and hence $x_n \to 0$ in mean $p[m]$.

Corollary 5.1 (Bounded Convergence Theorem): If $x_n \to x$ in measure $[m]$ and there exists K such that $\|x_n\| \leq K$ for every n, then $m(x_n) \to m(x)$ as $n \to \infty$.

Proof: Taking $p = 1$ in Lemma 5.4, $m(|x_n - x|) \to 0$ as $n \to \infty$. Then

$$|m(x_n - x)| = |\int \lambda m_{x_n - x}(d\lambda)| \leq \int |\lambda| m_{x_n - x}(d\lambda)$$
$$= m(|x_n - x|) \to 0$$

as $n \to \infty$. Hence $m(x_n) \to m(x)$ as $n \to \infty$.

We wish now to prove a Fatou lemma. Toward this end, we prove the following sequence of lemmas.

Lemma 5.5: Let $x, y \in X$ with $x \leftrightarrow y$. Then for $\epsilon > 0$, $m((x + y)[-\epsilon, \epsilon]') \leq m(x[-\epsilon/2, \epsilon/2]') + m(y[-\epsilon/2, \epsilon/2]')$.

Proof: This is clear once x and y are represented by functions.

Lemma 5.6: (i) Let $\{x_n\}, \{y_n\} \subset X$ and suppose $x_n \to 0, y_n \to 0$ in measure $[m]$. Suppose further that $x_n \leftrightarrow y_n$ for every n. Then $x_n + y_n \to 0$ in measure $[m]$. (ii) If $x_n \to 0$ in measure $[m]$, then $|x_n| \to 0$ in measure $[m]$.

Proof: (i) is a corollary of Lemma 5.5 and (ii) is trivial.

Lemma 5.7: Let $x_1, x_2 \in X$, and let $y = \frac{1}{2}(x_1 + x_2 - |x_1 - x_2|)$. Then

(i) $y \leq x_1$ and $y \leq x_2$;

(ii) if $x_1 \leftrightarrow x_2$ and $x_1 \geq 0$, then $\|y\| \leq \|x_2\|$.

Proof: (i) Observe that for $m \in M$,

$$m(|x_1 - x_2|) = \int |\lambda| m_{x_1 - x_2}(d\lambda) \geq |\int \lambda m_{x_1 - x_2}(d\lambda)|$$
$$= |m(x_1 - x_2)| \geq m(x_2) - m(x_1).$$

Then $-m(|x_1 - x_2|) \leq m(x_1) - m(x_2)$. Then using the linearity of m,

$$m(y) = \tfrac{1}{2}[m(x_1) + m(x_2) - m(|x_1 - x_2|)] \leq \tfrac{1}{2}[m(x_1)$$
$$+ m(x_2) + m(x_1) - m(x_2)] = m(x_1).$$

Thus $m(y) \leq m(x_1)$ for every m, so that $y \leq x_1$. Similarly, $y \leq x_2$.

(ii) Represent x_1 and x_2 by functions f_1, f_2 and use the fact that $f_1 \geq 0$ implies $-|f_2| \leq \min(f_1, f_2) \leq |f_2|$.

We now prove a Fatou Lemma. In order to get the result, we shall have to assume that each observable in the sequence is compatible with the limit observable.

Theorem 5.1 (Fatou's Lemma): Let $\{x_n\}$ be a sequence of nonnegative observables in X (i.e., $x_n \geq 0$ for every n), and suppose $x \in X$ with $x \leftrightarrow x_n$ for every n. If $x_n \to x$ in measure $[m]$, then $0 \leq m(x) \leq \liminf m(x_n)$ as $n \to \infty$.

Proof: We first show that $m(x) \geq 0$. Since $x_n \geq 0$ and $x_n \leftrightarrow x$, it is easy to show that $x(-\infty, -\epsilon) \leq (x - x_n)(-\infty, -\epsilon)$ for every n and every $\epsilon > 0$. Then for $k = 1, 2, \cdots$

$$0 \leq m(x(-\infty, -1/k)) \leq m(x - x_n)(-\infty, -1/k)$$
$$\leq m((x - x_n)[-1/k, 1/k]') \to 0$$

as $n \to \infty$. Hence $m(x(-\infty, -1/k)) = 0$ for $k = 1, 2, \cdots$. Then

$$m(x(-\infty, 0)) = m\left(x\left[\bigcup_{k=1}^{\infty}\left(-\infty, -\frac{1}{k}\right)\right]\right) = m\left(\bigvee_{k=1}^{\infty} x\left(-\infty, -\frac{1}{k}\right)\right)$$
$$= \lim_{k\to\infty} m(x(-\infty, -1/k)) = 0.$$

Thus $m(x(-\infty, 0)) = 0$ so that $m(x) = \int_0^{\infty} \lambda m_x(d\lambda) \geq 0$.

For $n = 1, 2, \cdots$ let $y_n = \frac{1}{2}(x_n + x - |x_n - x|)$. Then by Lemma 5.7, $\|y_n\| \leq \|x\|$ for every n. Observe that $y_n - x = \frac{1}{2}[(x_n - x) - |x_n - x|]$. Now by hypothesis $x_n - x \to 0$ in measure $[m]$. Then by Lemma 5.6 (ii), $-|x_n - x| \to 0$ in measure $[m]$. Then since $x_n - x \leftrightarrow -|x_n - x|$ by Lemma 5.6 (i), $(x_n - x) - |x_n - x| \to 0$ in measure $[m]$. Then for $\epsilon > 0$,

$$m((y_n - x)([-\epsilon, \epsilon])) = m([\tfrac{1}{2}\{(x_n - x) - |x_n - x|\}]$$
$$([-\epsilon, \epsilon])) = m([(x_n - x) - |x_n - x|]([-2\epsilon, 2\epsilon])) \to 1$$

Thus $y_n \to x$ in measure $[m]$ and $\|y_n\| \leq \|x\|$ for every n. Therefore by Corollary 5.1, $m(y_n) \to m(x)$. Now by Lemma 5.7, $m(y_n) \leq m(x_n)$ for every n. Hence, $m(x) = \lim_{n\to\infty} m(y_n) \leq \lim_{n\to\infty} \inf m(x_n)$.

We now prove the converse of Egoroff's Theorem.

Theorem 5.2: If $x_n \to x$ a.u. $[m]$, then $x_n \to x$ a.e. $[m]$.

Proof: Without loss of generality, let $x = 0$. Let $\epsilon > 0$ be given. Now for $\eta > 0$, there exists $a \in L$ such that $m(a') \leq \eta$ and $x_n \to 0$ uniformly on a. Then there exists N such that $n \geq N$ implies that $x_n[-\epsilon, \epsilon] \geq a$. Therefore for $n \geq N$, $\wedge_{k=n}^{\infty} x_k([-\epsilon, \epsilon]) \geq a$ and hence $m(\wedge_{k=n}^{\infty} x_k([-\epsilon, \epsilon])) \geq m(a) = 1 - m(a') \geq 1 - \eta$. Thus for $\eta > 0$, there exists N such that if $n \geq N$ then $0 \leq 1 - m(\wedge_{k=n}^{\infty} x_k([-\epsilon, \epsilon])) \leq \eta$. Hence,

$$m(\liminf x_n([-\epsilon, \epsilon])) = \lim_{n\to\infty} m\left(\bigwedge_{k=n}^{\infty} x_k([-\epsilon, \epsilon])\right) = 1.$$

We next prove a strong form of Egoroff's theorem for logics having f.c.c.

Theorem 5.3: Let L have f.c.c. If $x_n \to x$ a.e. $[m]$, then $x_n \to x$ a.u. $[m]$. In fact, there exists $a \in L$ such that $m(a') = 0$ and $x_n \to x$ uniformly on a.

Proof: Without loss of generality, let $x = 0$. Let p be a positive integer. Then $\wedge_{k=1}^{\infty} x_k[-1/p, 1/p] \leq \wedge_{k=2}^{\infty} x_k[-1/p, 1/p] \leq \cdots$ so that since L has f.c.c., there exists N_p such that $\wedge_{k=n}^{\infty} x_k[-1/p, 1/p] = \wedge_{k=N}^{\infty} x_k[-1/p, 1/p]$ for $n \geq N_p$. Without loss of generality, take $N_1 \leq N_2 \leq \ldots$ and let $a_p = \wedge_{k=N}^{\infty} x_k([-1/p, 1/p])$. Then $m(a_p) = m(\liminf x_n([-1/p, 1/p])) = 1$, and for $n \geq N_p$

$$x_n([-1/p, 1/p]) \geq \bigwedge_{k=n}^{\infty} x_k([-1/p, 1/p]) = a_p. \tag{2}$$

Now for every k and every p, $x_k[-1/p, 1/p] \geq x_k[-1/p + 1, 1/p + 1]$. Then as $N_{p+1} \geq N_p$,

$$a_p = \bigwedge_{k=N_p}^{\infty} x_k[-1/p, 1/p] = \bigwedge_{k=N_{p+1}}^{\infty} x_k[-1/p, 1/p]$$
$$\geq \bigwedge_{k=N_{p+1}}^{\infty} x_k[-1/p + 1, 1/p + 1] = a_{p+1}.$$

Therefore $a_1 \geq a_2 \geq \cdots$ so that there exists K such that $a_p = a_K$ for $p \geq K$. Let $a = a_K$. Then $m(a') = 1 - m(a_K) = 0$. For $\eta > 0$ choose p such that $p \geq K$ and $1/p \leq \eta$. Then using (2) for $n \geq N_p$

$$x_n[-\eta, \eta] \geq x_n[-1/p, 1/p] \geq a_p = a_k = a.$$

Thus $m(a') = 0$ and $x_n \to 0$ uniformly on a.

Corollary 5.4: Let H be a finite-dimensional Hilbert space. If $A_n \to A$ a.e. $[m]$ then there exists a projection P such that $m(P^{\perp}) = 0$ and $\lim\|(A_n - A)P\| = 0$ as $n \to \infty$.

Corollary 5.5: If H is a finite-dimensional Hilbert space and $A_n \to A$ a.e. $[m]$, then $m(A_n) \to m(A)$.

We now obtain a characterization for almost uniform convergence of operators. This result will be useful in the next section where we show Egoroff's theorem need not hold. Let H be a separable Hilbert space. For $U \subset H$, we let $\bar{U}$ denote the closure of U in H. In what follows when we refer to S as a linear subspace of H we are not assuming that S is closed. It is easy to verify that the following lemma holds.

Lemma 5.8: Let $\{S_{p,k} : p, k = 1, 2, \cdots\}$ be linear subspaces of H. Then $\bigcap_{k=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{p=n}^{\infty} S_{p,k}$ is also a linear subspace of H.

Theorem 5.4: Let $m(P) = \langle \phi, P\phi \rangle$ where $\phi \in H$ with $\|\phi\| = 1$. Then the following are equivalent:

(i) $A_n \to A$ a.u. $[m]$;

(ii) $\phi \in \overline{\bigcap_{k=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{p=n}^{\infty} (A_p - A)\left(\left[-\frac{1}{k}, \frac{1}{k}\right]\right)}$.

Proof: Without loss of generality, let $A = 0$. Let $S = \bigcap_{k=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{p=n}^{\infty} A_p([-1/k, 1/k])$. By Lemma 5.8, S is a linear subspace of H.

(i) implies (ii): Let $\epsilon > 0$ be given. Then there exists a projection P such that $m(P^{\perp}) \leq \epsilon^2$ and $A_n \to 0$ uniformly on P. Let $\psi = P\phi$. Then $\|\phi - \psi\|^2 = \|P^{\perp}\phi\|^2 = \langle \phi, P^{\perp}\phi \rangle \leq \epsilon^2$ so that $\|\phi - \psi\| \leq \epsilon$. Now for every k, there exists N_k such that $p \geq N_k$ implies that $A_p([-1/k, 1/k]) \geq P$. Letting P also denote the range of P then $\bigcap_{p=N_k}^{\infty} A_p([-1/k, 1/k]) \supseteq P$ for every k. Then trivially $S \supseteq P$. Then $\psi = P\phi \in P \subseteq S$. Thus for $\epsilon > 0$ there exists $\psi \in S$ such that $\|\phi - \psi\| \leq \epsilon$. Therefore $\phi \in \bar{S}$.

(ii) implies (i): Let $\epsilon > 0$ be given. Then there exists $\psi \in S$ such that $\|\psi - \phi\| \leq \sqrt{\epsilon}$. Let $P = LH(\{\psi\})$. Then P is a one-dimensional linear subspace of H and, hence, is closed. Therefore, P is a projection on H. By a standard result in Hilbert space theory,

$$\|P^{\perp}\phi\| = \|\phi - P\phi\|$$
$$= \inf\{\|\phi - \xi\| : \xi \in P\} \leq \|\phi - \psi\| \leq \sqrt{\epsilon}.$$

Then $m(P^{\perp}) = \langle \phi, P^{\perp}\phi \rangle = \|P^{\perp}\phi\|^2 \leq \epsilon$.

Now $\psi \in S$ so that for every k, $\psi \in \bigcup_{n=1}^{\infty} \bigcap_{p=n}^{\infty} A_p([-1/k, 1/k])$. Then, for $k = 1, 2, \cdots$ there exists N_k such that $\psi \in \bigcap_{p=N_k}^{\infty} A_p([-1/k, 1/k])$. Thus for $p \geq N_k$, $\psi \in A_p([-1/k, 1/k])$. But then for $p \geq N_k$, $A_p([-1/k, 1/k]) \geq LH(\{\psi\}) = P$. Therefore, $A_n \to 0$ uniformly on P completing the proof.

Corollary 5.4: Let $m(P) = \sum \lambda_i \langle \phi_i, P\phi_i \rangle$ where $\{\phi_i\}$ is an orthonormal set in H and $0 < \lambda_i \leq 1$ with $\sum \lambda_i = 1$. Let

$$S = \bigcap_{k=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{p=n}^{\infty} (A_p - A)\left(\left[-\frac{1}{k}, \frac{1}{k}\right]\right).$$

Then $A_n \to A$ a.u. $[m]$ if and only if $\phi_i \in \bar{S}$ for every i.

Proof: Let $m_i(P) = \langle \phi_i, P\phi_i \rangle$. Then by Lemma 2.1, $A_n \to A$ a.u.$[m]$ if and only if $A_n \to A$ a.u.$[m_i]$ for every i. The result is then immediate from Theorem 5.2.

6. COUNTEREXAMPLE FOR EGOROFF'S THEOREM

Theorem 5.3 shows that Egoroff's theorem holds on a finite-dimensional Hilbert space. We now wish to show that this does not hold for infinite-dimensional Hilbert spaces, and consequently Egoroff's theorem is not valid for observables.

In preparation for the construction of a counterexample to Egoroff's theorem, we consider now the collection of all polynomials on the unit interval $[0, 1]$. Let $C[0, 1]$ denote the continuous functions on $[0, 1]$, and let Π be the collection of polynomials in $C[0, 1]$. Let $|\cdot|$ denote the supremum norm on $C[0, 1]$.

Lemma 6.1: Let $\{r_k\}_{k=1}^n \subset [0, 1]$. Let $Z = \{p \in \Pi : p(r_k) = 0 \text{ for } 1 \leq k \leq n\}$, and let Z^C denote the closure of Z in $C[0, 1]$. Then

$$Z^C = \{f \in C[0, 1] : f(r_k) = 0 \text{ for } 1 \leq k \leq n\}.$$

Proof: We show that Z^C is a closed ideal in $C[0, 1]$. Clearly Z is a linear subspace of $C[0, 1]$, and hence Z^C is a closed linear subspace. Let $f \in Z^C$ and $g \in C[0, 1]$ with $|f| = |g| = 1$. Let $\epsilon > 0$ be given. Now Π is dense in $C[0, 1]$ so that there exists $q \in \Pi$ such that $|g - q| < \min\{\epsilon/2, 1\}$. Then $|q| > 1 - |g| = 0$. Now $f \in Z^C$ so that there exists $s \in Z$ such that $|f - s| < \epsilon/2|q|$. Let $p = sq$. Then since $s \in Z$ and $q \in \Pi$, $p \in Z$. Furthermore

$$|fg - p| = |fg - sq| \leq |fg - fq| + |fq - sq| \leq |f| \cdot |g - q| + |q| \cdot |f - s| < \epsilon.$$

Therefore $fg \in Z^C$ for every $f \in Z^C$ and $g \in C[0, 1]$ with $|f| = |g| = 1$. Then as Z^C is a linear space, $fg \in Z^C$ for every $f \in Z^C$ and every $g \in C[0, 1]$. Thus Z^C is a closed ideal in $C[0, 1]$. Then by Ref. 12, (Problem 1, p. 879) there exists a closed subset F of $[0, 1]$ such that $Z^C = \{f \in C[0, 1] : f(F) = 0\}$. Clearly, $\{r_k\} \subset F$. Furthermore since $p(\lambda) = (\lambda - r_1)(\lambda - r_2)\cdots(\lambda - r_n) \in Z^C$, we must have $F = \{r_k\}$. Therefore, $Z^C = \{f \in C[0, 1] : f(r_k) = 0 \text{ for } 1 \leq k \leq n\}$.

Lemma 6.2: Let Z be as in Lemma 6.1, and let $L^2[0, 1]$ be the square integrable functions with respect to Lebesgue measure μ on $[0, 1]$. Then Z is dense in $L^2[0, 1]$.

Proof: Let $f \in L^2[0, 1]$, and let $\epsilon > 0$ be given. Now $C[0, 1]$ is dense in $L^2[0, 1]$ so that there exists $g \in C[0, 1]$ such that $\|f - g\| \leq \epsilon/3$. Let K be a bound for g, and let $\eta = \min\{|r_k - r_i| : i \leq i, k \leq n\}$. Choose δ such that $0 < \delta < \min\{\epsilon^2/9K^2, \eta\}$ and let

$$E = \bigcup_{k=1}^{n} \left[r_k - \frac{\delta}{2n}, r_k + \frac{\delta}{2n}\right].$$

Let

$$h(\lambda) = g(\lambda) \cdot [1 - 2n/\delta) \inf\{|\lambda - \mu| : \mu \in E'\}].$$

Then $h = g$ on E' and $|h - g| \leq |g| \leq K$. Therefore $\int |h(\lambda) - g(\lambda)|^2 d\lambda = \int_E |h(\lambda) - g(\lambda)|^2 d\lambda \leq K^2\mu(E) \leq K^2\delta < \epsilon^2/9$. Therefore $\|g - h\| < \epsilon/3$. Now $h \in C[0, 1]$ and $h(r_k) = 0$ for $1 \leq k \leq n$. Then by Lemma 6.1, there exists $p \in Z$ such that $|h - p| < \epsilon/3$. Then $\|h - p\| < \epsilon/3$ so that

$$\|f - p\| \leq \|f - g\| + \|g - h\| + \|h - p\| < \epsilon.$$

Hence $f \in \overline{Z}$ so that $\overline{Z} = L^2[0, 1]$.

We need one further result before constructing the counterexample. This lemma holds on arbitrary logics L.

Lemma 6.3: Let $a_1 \geq a_2 \geq \cdots \geq a_n$ be in L. Then there exists $x \in X$ such that $x([-1/k, 1/k]) = a_k$ for $1 \leq k \leq n$.

Proof: Define x by $\sigma(x) \subset \{1/n, 1/(n-1), \ldots, \frac{1}{2}, 1, 2\}$, $x(\{1/k\}) = a_k \wedge a'_{k+1}$ for $1 \leq k < n$, $x(\{1/n\}) = a_n$ and $x(\{2\}) = a'_1$. It is straightforward to verify that $x \in X$ and $x([-1/k, 1/k]) = a_k$ for $1 \leq k \leq n$.

Let $H = L^2[0, 1]$ and let $\{r_k\}_{k=1}^{\infty}$ be an infinite subset of $[0, 1]$. For $p \in \Pi$ let $d(p)$ denote the degree of p. Let n be a positive integer and keep n fixed. For $k = 1, 2, \ldots, n$ let

$$P_k = \{p \in \Pi : d(p) \leq 2n - k \text{ and } p(r_j) = 0 \text{ for } 1 \leq j \leq k\}.$$

Suppose $1 \leq k \leq i \leq n$. Then for $p \in P_i$, $d(p) \leq 2n - i \leq 2n - k$ and $p(r_j) = 0$ for $1 \leq j \leq i$ and, hence, for $1 \leq j \leq k$. Thus $p \in P_k$, and consequently $P_1 \supseteq P_2 \supseteq \cdots \supseteq P_n$. Now for $1 \leq k \leq n$, $P_k \subset \{p \subset \Pi : d(p) \leq 2n - k\}$ which is a $(2n - k)$-dimensional subspace of H. Therefore, P_k is finite-dimensional and consequently closed. Thus $P_1 \geq P_2 \geq \cdots \geq P_n$ are closed subspaces of H. Then by Lemma 6.3, there exists a self-adjoint operator A_n such that $A_n([-1/k, 1/k]) = P_k$ for $1 \leq k \leq n$.

Thus we have a sequence A_n of self-adjoint operators such that $A_n([-1/k, 1/k]) = \{p \subset \Pi : d(p) \leq 2n - k \text{ and } p(r_j) = 0 \text{ for } 1 \leq j \leq k\}$ for $1 \leq k \leq n$. We show now that this sequence gives a counterexample to Egoroff's theorem.

Lemma 6.4: $A_n \to 0$ everywhere.

Proof: Let k be a positive integer. Then for $i \geq k$, $A_i([-1/k, 1/k]) = \{p \in \Pi : d(p) \leq 2i - k \text{ and } p(r_j) = 0 \text{ for } 1 \leq j \leq k\}$. Then for $n \geq k$,

$$\bigcap_{i=n}^{\infty} A_i([-1/k, 1/k]) = \bigcap_{i=n}^{\infty} \{p \in \Pi : d(p) \leq 2i - k \text{ and } p(r_j) = 0 \text{ for } 1 \leq j \leq k\}$$

$$= \{p \in \Pi : d(p) \leq 2n - k \text{ and } p(r_j) = 0 \text{ for } 1 \leq j \leq k\}.$$

Then as $\cap_{i=1}^{\infty} A_i([-1/k, 1/k]) \subseteq \cap_{i=2}^{\infty} A_i([-1/k, 1/k]) \subseteq \cdots$,

$$\bigcup_{n=1}^{\infty} \bigcap_{i=n}^{\infty} A_i([-1/k, 1/k]) = \bigcup_{n=k}^{\infty} \bigcap_{i=n}^{\infty} A_i([-1/k, 1/k])$$

$$= \bigcup_{n=k}^{\infty} \{p \in \Pi : d(p) \leq 2n - k \text{ and } p(r_j) = 0 \text{ for } 1 \leq j \leq k\}$$

$$= \{p \in \Pi : p(r_j) = 0 \text{ for } 1 \leq j \leq k\}.$$

By Lemma 6.2, this set is dense in H. Therefore, for $k = 1, 2, \ldots,$

$$\liminf A_n \left(\left[-\frac{1}{k}, \frac{1}{k}\right]\right) = \overline{\bigcup_{n=1}^{\infty} \bigcap_{i=n}^{\infty} A_i([-1/k, 1/k])} = H.$$

Therefore, $A_n \to 0$ everywhere.

Lemma 6.5: For every state m on $L(H)$, $A_n \not\to A$ a.u. $[m]$.

Proof: From the proof of Lemma 6.4, for $k = 1, 2, \ldots$,

$$\bigcup_{n=1}^{\infty} \bigcap_{i=n}^{\infty} A_i \left(\left[-\frac{1}{k}, \frac{1}{k}\right]\right) = \{p \in \Pi : p(r_j) = 0 \text{ for } 1 \leq j \leq k\}.$$

Then

$$S = \bigcap_{k=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{i=n}^{\infty} A_i \left(\left[-\frac{1}{k}, \frac{1}{k}\right]\right) = \bigcap_{k=1}^{\infty} \{p \in \Pi : p(r_j) = 0 \text{ for } 1 \leq j \leq k\}$$

Then $\bar{S} = \{0\}$ and the result is immediate from Corollary 5.4.

Thus we have a counterexample to Egoroff's theorem. In fact, we have a sequence which converges a.e. $[m]$ for every m, but fails to converge a.u. $[m]$ for *any* m.

We close this section with a few remarks on the comparison of the theory of Segal and Stinespring and that of this paper. Although we saw certain similarities in Secs. 3 and 4, the developments of the two theories lead to quite different results. In Ref. 13 Theorem 3.1 Padmanabhan proves that Egoroff's theorem holds in the Segal–Strinespring theory. However, as we have just seen, this is not the case in our theory. Padmanabhan's proof relies heavily on the subadditivity of gages, while the theorem fails here essentially because of the lack of subaddivity for states. It would seem that the difficulties incurred in attempting to generalize function theoretic results to observables arise from the inherent nature of states, rather than from the convergence definitions we have taken.

[1]V. Varadarajan, Commun. Pure Appl. Math. **15,** 189 (1962).
[2]G. Mackey, *The mathematical foundations of quantum mechanics* (Benjamin, New York, 1963).
[3]G. Bodiou, *Theorie dialectique probabilities* (Gauthier-Villar, Paris, 1964).
[4]S. Gudder, J. Math. Anal. Appl. **20,** 48 (1967).
[5]I. E. Segal, Ann. Math. **57,** 401 (1953).
[6]I. E. Segal, Am. J. Math. **76,** 721 (1954).
[7]H. Umegaki, Kodai Math. Sem. Rep. 14, 59 (1962).
[8]H. Umegaki, Proc. Jap. Acad. **37,** 459 (1961).
[9]W. Stinespring, Trans. Am. Math. Soc. **90,** 15 (1959).
[10]S. Gudder, Pac. J. Math. **19,** 81 (1966); Pac. J. Math. **19,** 588 (1966).
[11]A. Gleason, J. Ration. Mech. Anal. 6, 885 (1957).
[12]N. Dunford, and J. Schwarz, *Linear operators, Part II* (Interscience, New York, 1963).
[13]A. Padmanabhan, Trans. Am. Math. Soc. **128,** 359 (1967).

Anomalous dimensions in one-dimensional quantum field theory*

G. Parisi
Laboratori Nazionali del CNEN, Frascati, Italy

F. Zirilli
Scuola Normale Superiore, Pisa, Italy
(Received 23 June 1972)

We study the short distance behavior of the Green's functions of two operators in a soluble one-dimensional model of quantum field theory with dimensionless coupling constant. Integer power behavior does not occur. The leading terms of the Wilson expansion of two operators at short distances are determined.

I. INTRODUCTION

The behavior of Green's functions at short distances is one of the most interesting and controversial problems in quantum field theory.[1]

The first fundamental step towards a deeper understanding of the problem was made by Wilson.[2] He suggested that the product of two operators satisfies the following asymptotic expansion at short distances:

$$\phi(x)\phi(0) \simeq \sum_n \xi_n(x) O_n(0), \tag{1}$$

where the functions $\xi_n(x)$ become singular when x goes to zero.

In free field theory these functions are integer powers; in perturbation theory this simple behavior is destroyed by the appearance of terms of the form $g \log(x^2)$.

Wilson suggested that the logarithms may sum to a power which may not be integral: This happens in the soluble two-dimensional Thirring model.[3] The exponent of the power is dependent on the renormalized coupling constant.

In this work we prove that similar pathologies arise also in a very simple one-dimensional model of quantum field theory.

We study the Lagrangian

$$\mathcal{L} = \int dt \{\tfrac{1}{2}[\dot\phi(t)]^2 - \tfrac{1}{2}m^2[\phi(t)]^2 - g[\phi(t)]^{-2}\}. \tag{2}$$

This is the only possible one with dimensionless coupling constant: the Lagrangian density must have dimension 1, the field ϕ, therefore, has dimension $-\frac{1}{2}$, so that the only possible scale invariant interaction is $g\phi^{-2}$.

This model can be solved using the equivalence between one-dimensional quantum field theory and nonrelativistic quantum mechanics. The analogous quantum mechanical problem is the quantal oscillator: a harmonic oscillator with centrifugal potential g/x^2; its solution is known from the early days of quantum mechanics.[4]

We find that for $g > -\frac{1}{8}$ the Wilson expansion for the product of two fields ϕ is

$$\phi(t)\phi(0) \simeq \phi^2(0) + t\dot\phi(0)\phi(0) + \tfrac{1}{2}t^2\ddot\phi(0)\phi(0) + Dt^{a+2}\phi(0)^{-2a-1}\delta[\phi(0)], \tag{3}$$

where $a = \frac{1}{2}(1+8g)^{1/2}$ and all the neglected terms in the Wilson expansion are of the type $t^{\beta_i}O_i(0)$, where β_i is equal to n or to $n + a + 2$ with integer n.

We note that $\phi^{-2a-1}\delta[\phi]$ is not an operator but a nonbounded sesquilinear form, which is defined on any finite linear combination of the energy eigenvectors.

This result is interesting because it shows that anomalous dimensions in the short distance behavior are a very common phenomenon, which is present not only in relativistic quantum field theory, but also in the old nonrelativistic quantum mechanics.

In Sec. II we rederive the analogy between one-dimensional quantum field theory and nonrelativistic quantum mechanics. We write the solution of the quantal oscillator and use it to compute the Wightman functions for the quantum field problem.

In Sec. III we study the behavior of the two-point Wightman function at short distances and find anomalous dimensions. We extend our study to the two-point correlation function between two arbitrary energy eigenstates and finally arrive at the Wilson expansion (3).

In Sec. IV we briefly discuss our results and make an interesting but unproven conjecture.

II. SOLUTION OF THE MODEL

The Lagrangian of our one-dimensional problem is (2).

One can easily find the associated Euler–Lagrange equation

$$\ddot\phi(t) = m^2\phi(t) - 2g[\phi(t)]^{-3} \tag{4}$$

and the Hamiltonian

$$H = \tfrac{1}{2}[\pi(t)]^{-2} + \tfrac{1}{2}m^2[\phi(t)]^2 + g[\phi(t)]^{-2}. \tag{5}$$

ϕ and π satisfy the canonical commutation relations

$$[\phi(t), \pi(t)] = -i. \tag{6}$$

In order to find the eigenvectors of the Hamiltonian we use the standard representation

$$\phi(0) \to x, \qquad \pi(0) \to i\frac{d}{dx}. \tag{7}$$

This trick reduces the problem to finding the eigensolutions of the following differential equation:

$$\left(-\frac{d^2}{2dx^2} + \tfrac{1}{2}m^2x^2 + \frac{g}{x^2}\right)\psi(x) = E\psi(x); \tag{8}$$

Equation (8) is the Schrödinger equation for the quantal oscillator. The eigenfunctions and eigenvectors of Eq. (8) are

$$E_n = m[2n + a + 1];$$

$$\psi_n(x) = (4m)^{1/4}\left(\frac{\Gamma(n+1)}{\Gamma(a+n+1)}\right)^{1/2}[mx^2]^{(2a+1)/4} \times \exp\left(-\frac{mx^2}{2}\right) L_n^a[mx^2], \tag{9}$$

where $a = \frac{1}{2}(1 + 8g)^{1/2}$, n is a nonnegative integer, and $L_n^a(x)$ are the Laguerre polynomials.[5]

If $g \lesssim -\frac{1}{8}$ the spectrum of the Hamiltonian is no longer bounded below; there exists no ground state and the physical meaning of the problem is lost.

The Wightman functions of the theory are

$$\langle 0|\phi(t_1)\cdots\phi(t_n)|0\rangle = \langle\psi_0|x(t_1)\cdots x(t_n)|\psi_0\rangle, \tag{10}$$

where ψ_0 is the ground state of (8) and $x(t)$ is the position operator at time t in the Heisenberg representation. We note that $x(t)$ satisfies the same equation of motion (4) as $\phi(t)$.

If we define

$$x_{nm} = \langle\psi_n|x|\psi_m\rangle, \tag{11}$$

we have from (10)

$$\langle 0|\phi(t)\phi(0)|0\rangle = \sum_n e^{itE_n}|x_{0n}|^2. \tag{12}$$

Similar expression can be easily derived for general N-point Wightman functions.

III. THE WILSON EXPANSION

In this section we compute the two-point Wightman function and study its behavior at small t.

The formula for x_{0n} is

$$x_{0n} = -[2(\pi m)^{-1/2}][\Gamma(a+n+1)\Gamma(a+1)\Gamma(n+1)]^{-1/2} \times \Gamma(a+\tfrac{3}{2})\Gamma(n-\tfrac{1}{2}). \tag{13}$$

In the limit $n \to \infty$ we find

$$x_{0n} \simeq [-\tfrac{1}{2}(\pi m)^{-1/2}]\,\Gamma(a+\tfrac{3}{2})\Gamma^{-1/2}(a+1)n^{-[(a/2)+(3/2)]}. \tag{14}$$

This asymptotic behavior of x_{0n} implies that in the small t region

$$\langle 0|\phi(t)\phi(0)|0\rangle = (1/m)[C_0 + C_1\cdot(mt) + C_2\cdot(mt)^2 + C_3\cdot(mt)^{a+2} + \cdots], \tag{15}$$

where C_0, C_1, C_2, C_3 are g-dependent constants and the neglected terms have higher power in t. It is interesting to observe that the power of the fourth term is not integral and is a continuous function of the coupling constant.

We now look for the two-point correlation function between two arbitrary energy eigenstates s and r.

We need only to compute the asymptotic behavior for large n of x_{sn}.

If we decompose

$$L_s^a(x^2m) = \sum_0^s {}_k\, b_k^s (x^2m)^k \tag{16}$$

we find, using Eq. (9), that

$$x_{sn} = \frac{1}{m^{1/2}}\left(\frac{\Gamma(s+1)\Gamma(n+1)}{\Gamma(a+s+1)\Gamma(n+1+a)}\right)^{1/2} \times \sum_0^s {}_k \left(b_k^s \frac{\Gamma(a+k+\frac{3}{2})\Gamma(n-k-\frac{1}{2})}{\Gamma(n+1)\Gamma(-k-\frac{1}{2})}\right)$$

$$\simeq \frac{1}{m^{1/2}}\Gamma^{1/2}(s+1)\Gamma^{-1/2}(a+s+1)\sum_0^s {}_k \times [\Gamma(s+k+\tfrac{3}{2})b_k^s n^{-(a/2)-(3/2)-k}]. \tag{17}$$

The terms proportional to b_k^s with $k \neq 0$ go faster to zero. The final result for small t is

$$\langle r|x(t)x(0)|s\rangle = \frac{1}{m}[D_0^{r,s} + D_1^{r,s}\cdot(mt) + D_2^{rs}\cdot(mt)^2 + D_3\left(\frac{\Gamma(s+1)\Gamma(r+1)}{\Gamma(a+s+1)\Gamma(a+r+1)}\right)^{1/2} b_0^s b_0^r (mt)^{2+a}] \tag{18}$$

where $D_0^{r,s}, D_1^{r,s}, D_2^{r,s}$ are constants dependent on g, r and s, but D_3 is a function of only g. b_0^s can also be defined as

$$b_0^s = \lim_{x\to 0}\frac{1}{(4m)^{1/4}}\frac{\Gamma(a+s+1)}{\Gamma(s+1)}(mx^2)^{-(2a+1)/4}\psi_s(x), \tag{19}$$

so that (18) is equivalent to

$$\langle r|x(t)x(0)|s\rangle \simeq \langle r|x^2(0)|s\rangle + t\langle r|\dot{x}(0)x(0)|s\rangle + \tfrac{1}{2}t^2\langle r|\ddot{x}(0)x(0)|s\rangle + t^{a+2}\tfrac{1}{2}D_3\int\psi_r^+(x)\psi_s(x)x^{-2a-1} \times \delta(x) + \cdots. \tag{20}$$

The coefficient of t^{a+2} can also be interpreted as the mean value of the sesquilinear form $x^{-2a-1}\delta(x)$ between the states.

Equation (20) can be rewritten in operational form, and in this way we find the Wilson expansion (3) for the product of two fields.

We note that the operational form of the short distance singularities is mass independent, and the index of the power depends only on the dimensionless coupling constant.

One can investigate the general form of the neglected terms, computing the exact two-point correlation function. This can be done by inserting in (18) the exact expression (17) for x_{rn}, and not its asymptotic expansion.

One finds

$$\langle r|x(t)x(0)|s\rangle = \frac{e^{-ir\cdot tm}}{m}\left(\frac{\Gamma(s+1)\Gamma(r+1)}{\Gamma(a+s+1)\Gamma(a+r+1)}\right)^{1/2} \times \frac{1}{\Gamma(1+a)}\sum_0^r {}_k \sum_0^s {}_{k'}\, b_k^s b_k^r\,\Gamma(a+k+\tfrac{3}{2})\Gamma(a+k'+\tfrac{3}{2}) \times F(-k-\tfrac{1}{2},-k'-\tfrac{1}{2};1+a;e^{imt}). \tag{21}$$

Using the well-known decomposition of the hypergeometric function[5], we arrive at

$$\langle r|x(t)x(0)|s\rangle = \frac{e^{-irtm}}{m}\left(\frac{\Gamma(s+1)\Gamma(r+1)}{\Gamma(a+s+1)\Gamma(a+r+1)}\right) \times \sum_0^r {}_k \sum_0^s {}_{k'}\, b_k^s b_k^r \Big[\Gamma(a+k+k'+2) \times F(-k-\tfrac{1}{2},k'-\tfrac{1}{2};-a-k-k'-1;1-e^{imt}) + [1-e^{imt}]^{2+a+k+k'} \times \frac{\Gamma(-a-k-k'-2)\Gamma(a+k+\frac{3}{2})\Gamma(a+k'+\frac{3}{2})}{\Gamma(-k-\frac{1}{2})\Gamma(-k'-\frac{1}{2})} \times F(a+k+\tfrac{3}{2},a+k'+\tfrac{3}{2};a+k+k'+3;1-e^{imt})\Big]. \tag{22}$$

The first term generates short-distance singularities with integer powers, and the second term contains only powers of the form $n + a + 2$. (The hypergeometric function is regular at the origin.)

IV. CONCLUSIONS

The results of our study show that in a one-dimensional model of quantum field theory with dimensionless coupling constant, the fundamental field does not change dimension; but operators with anomalous dimension appear in the Wilson expansion of the product of two fields. The dimension of these operators is coupling-constant dependent.

We also find a very simple expression for the leading anomalous term of the Wilson expansion.

The next step along this line of work is to study other models with singular potential and with dimensional-coupling constant. Our feeling is that in this model too there are nonintegral powers in the short-distance Wilson expansion; but the anomalous dimensions should not be coupling-constant dependent. It would be very interesting to verify this conjecture.

ACKNOWLEDGMENTS

We thank Professor F. Calogero, Professor G. F. Dell'-Antonio, and Professor M. Marchioro for interesting discussions, and we are also grateful to S. Graffi for much illuminating advice and discussion. One of the authors (G.P.) thanks Professor W. Zimmermann for the kind hospitality extended to him at New York University, where part of this work was done.

---

*Work partially supported by GNSM-CNR.

[1] H. Lehmann, Nuovo Cimento **11**, 342 (1954).
[2] K. Wilson, Phys. Rev. **179**, 1499 (1969).
[3] K. Wilson, Phys. Rev. D **2**, 1473 (1970).
[4] F. Calogero, J. Math. Phys. **10**, 2191 (1969) and references therein.
[5] I. S. Gradshteyn and I. M. Ryzhik, *Table of integrals, Series and products* (Academic, New York, 1965).

Nonlinear Padé approximants for Legendre series*

J. Fleischer†
University of California, Los Alamos Scientific Laboratory, Los Alamos, New Mexico 87544
(Received 7 July 1972)

We introduce a new kind of Padé approximants for Legendre series based on the solution of a nonlinear system of equations. These Padé approximants have many properties in common with the well-known Padé approximants for Taylor series. In the examples studied here, the poles and zeroes of the Padé approximants lie on the cut and separate each other—a property which one expects to hold in general for Padé approximants. A proof of convergence follows the same lines as for Taylor series. Moreover, it turned out that these nonlinear approximants converge more rapidly than the linear approximants introduced in an earlier work. They may become a powerful tool in the summation of Legendre series.

1. INTRODUCTION

In a recent paper we have introduced Padé approximants (PA's) for Legendre series, which were obtained by solving a linear system of equations—as it is the case for PA's for Taylor series. However, "linear" Legendre PA's do not have the property that their first expansion coefficients agree with the first coefficients of the original series. If one wants to define PA's for Legendre series which have this property, one has to solve nonlinear equations and we call these PA's "nonlinear" Padé approximants.

These nonlinear PA's have some interesting properties, which we demonstrate by numerical calculations. First of all, they can improve the rate of convergence of a convergent series quite remarkably, and also they can give very good results even when the series diverges. Expanding the PA in a Legendre series, one can recover higher coefficients of the original series to a high precision. Furthermore, in the cases under consideration all the poles lie on the cut and are separated by zeroes—a property which one expects to hold in general for PA's. The residues of the poles give us the imaginary part on the cut with reasonable accuracy.

A proof of convergence is valid for these PA's, and it is hoped that the nonlinear PA's will have many other properties of PA's for Taylor series, such as, e.g., the properties proven for Stieltjes functions.

There should be many physical applications of these nonlinear PA's as all problems of scattering theory involve Legendre series or related series. Especially we have in mind application to phase shift analysis. Fitting an experimentally known scattering amplitude with a PA-ansatz with a limited number of parameters, one may obtain information about higher coefficients by expanding the PA.

2. THE NONLINEAR APPROXIMANTS

The definition of Padé approximants (PA's) for Taylor series can be made as

$$f(z)Q_M(z) - P_N(z) = Az^{M+N+1} + \cdots \tag{1}$$

or equivalently

$$f(z) - P_N(z)/Q_M(z) = A'z^{M+N+1} + \cdots, \tag{2}$$

where P_N and Q_M are polynomials of degree N and M, respectively. By introducing PA's for Legendre series, analogous definitions lead to different approximations with completely different properties. The reason is that in the expansion of a product of Legendre polynomials more than one term contributes:

$$P_i(z)P_k(z) = \sum_{l=|i-k|}^{i+k} \alpha_l^{(i,k)} P_l(z). \tag{3}$$

In analogy to Eq. (1) we have introduced PA's for Legendre series in a recent paper[1] by the definition

$$\begin{aligned}(a_0P_0 + a_1P_1 + \cdots + a_LP_L)(d_0P_0 + \cdots + d_MP_M) \\ = n_0P_0 + n_1P_1 + \cdots + n_NP_N + 0 + \cdots + 0 \\ + \bar{n}_{M+N+1}P_{M+N+1} + \cdots\end{aligned} \tag{4}$$

($L = 2M + N$) and discussed applications to the crossing of Feynman graphs.

In this paper we are concerned with the definition analogous to Eq. (2), which now reads

$$f(z) - R_N(z)/S_M(z) = \bar{a}_{M+N+1}P_{M+N+1} + \cdots, \tag{5}$$

or explicitly

$$\begin{aligned}(n_0P_0 + n_1P_1 + \cdots + n_NP_N)/(d_0P_0 + d_1P_1 + \cdots + d_MP_M) \\ = a_0P_0 + a_1P_1 + \cdots + a_{M+N}P_{M+N} + \cdots,\end{aligned} \tag{6}$$

where the a_i are the coefficients of the expansion of $f(z)$ into a Legendre series. We will show that the determination of the n_i and d_i leads to a linear system of equations for the n_i and a nonlinear one for the d_i.

Putting $d_0 = 1$, we expand $1/S_M$ in a Legendre series:

$$\begin{aligned}\frac{1}{S_M} &= \frac{1}{d_0P_0 + d_1P_1 + \cdots + d_MP_M} \\ &\equiv \frac{s_0}{(z - s_1)(z - s_2)\cdots(z - s_M)} \\ &= D_0P_0 + D_1P_1 + \cdots + D_{M-1}P_{M-1} + \cdots.\end{aligned}$$

Performing the projection into partial waves, we have

$$\begin{aligned}D_l &= \frac{2l+1}{2} s_0 \int_{-1}^{+1} \frac{P_l(x)}{(x - s_1)(x - s_2)\cdots(x - s_m)} dx \\ &= -(2l+1)\sum_{\nu=1}^{M} \frac{P_l(s_\nu)}{S_M'(s_\nu)} Q_0(s_\nu), \qquad l = 0, \ldots, M-1,\end{aligned}$$

where we assumed that the roots s_ν of the polynomial S_M are not degenerate, Q_0 being the Legendre function of the second kind and index 0.

We obtain the higher coefficients recursively:

$$D_M = \frac{2M+1}{d_M}\left(1 - \sum_{\lambda=0}^{M-1} \frac{d_\lambda}{2\lambda+1} D_\lambda\right)$$

and

$$D_{M+\nu} = -\frac{1}{d_M \alpha_\nu^{(M,M+\nu)}} \sum_{\lambda=\lambda_0}^{M+\nu-1} \left(\sum_{\mu=0}^{M} d_\mu \alpha_\nu^{(\mu,\lambda)}\right) D_\lambda,$$

$$\nu = 1, 2, 3, \cdots,$$

with

$$\lambda_0 = \max(0, \nu - M),$$

$\alpha_\nu^{(\mu,\lambda)}$ being the coefficients in Eq. (3).

Equation (6) now reads

$$(n_0P_0 + n_1P_1 + \cdots + n_NP_N)(D_0P_0 + D_1P_1 + \cdots + D_{M-1}P_{M-1} + \cdots) = a_0P_0 + a_1P_1 + \cdots + a_{M+N}P_{M+N} + \cdots.$$

Using (3) and equating the coefficients of Legendre polynomials with equal index on both sides of this equation, we finally obtain the following system of equations for the coefficients $n_i (i = 0, \ldots, N)$, and $d_i (i, \ldots, M)$:

$$\sum_{\nu=0}^{N} n_\nu \sum_{\mu=|\lambda-\nu|}^{\lambda+\nu} D_\mu \alpha_\lambda^{(\mu,\nu)} = a_\lambda, \quad \lambda = 0, \ldots, N+M. \tag{7}$$

This is a linear system of equations for the n_ν ($\lambda = 0, 1, \ldots, N$) and a nonlinear system for the d_μ ($\lambda = N+1, \ldots, N+M$).

Solutions of the nonlinear system have been found using the fitting routine FIT_4, developed by W. Anderson and T. Doyle. If the nonlinear equations are written in the form

$$f_1(d_1, d_2, \ldots, d_M) = 0,$$
$$f_2(d_1, d_2, \ldots, d_M) = 0,$$
$$\vdots$$
$$f_M(d_1, d_2, \ldots, d_M) = 0,$$

FIT_4 searches for a minimum of

$$\varphi = f_1^2 + f_2^2 + \ldots + f_M^2,$$

which in the case of a solution should be equal to 0. It turned out that the system of equations (7) has more than one solution in general, but only one solution with all poles s_ν outside the region $-1 \le z \le +1$ has been found in the cases under consideration.

We finally remark that the PA defined by Eq. (4) does not have the property that its expansion into a Legendre series has its first coefficients equal to the first coefficients of the original function. As a consequence, the proof of convergence given in Ref. 1 is invalid for those PA's, but is valid for the nonlinear PA's introduced here. For completeness we recall the following theorem:

Theorem: Let $Q_k(z)$ be any infinite sequence of $[N,M]$ PA's to a formal Legendre series where $M+N$ tends to infinity with k. If the absolute value of the Q_k is uniformly bounded in the ellipse (its boundary included) with foci at $+1$ and -1 and semimajor axis A, then the Q_k converge uniformly in the ellipse with semimajor axis a, $a < A$ to an analytic function $f(z)$, the Legendre series of which has a semimajor axis of at least A.

In the proof given in Ref. 1 we only have to replace $2M+N$ by $M+N$ wherever it occurs.

3. MATHEMATICAL EXAMPLES

In order to test the accuracy of the nonlinear PA's, we calculate them for some Legendre series for which the series is exactly summable. We are particularly interested in the improvement of convergence, the expansion of the PA's in a Legendre series, and whether or not any information can be obtained about the imaginary part on the cut.

A. Generating function of Legendre polynomials

We test the PA's for

$$f(z) = \frac{1}{\sqrt{1-2az+a^2}} = \sum_{l=0}^{\infty} a^l P_l(z).$$

Choosing $a = 0.3$, the cut starts at $x_C = (1+a^2)/2a|_{0.3} = 1.817$. In Table I some numerical results are presented. Quite generally we observe that the nonlinear PA's are a better approximation than the linear PA's and that their improvement of the convergence is quite remarkable. Even when the Legendre series diverges ($z = -3.5$), they still give an extremely good result.

TABLE I. Shows the improvement of the convergence for the generating function. PS is the partial sum ($l_{\max} = 6$), LP the linear [2, 2]-PA, NP the nonlinear [3, 3]-PA–both involving the same number of coefficients–and the exact function.

z	PS	LP	NP	exact
− 3.5	–	0.561	0.5600	0.5599
− 1.75	0.77	0.6838	0.683 591	0.683 586
− 1.0	0.769 40	0.769 236	0.769 230 81	0.769 230 77
− 0.5	0.848 24	0.848 188 4	0.848 188 95	0.848 188 93
− 0.0	0.957 81	0.957 829	0.957 826 26	0.957 826 29
+ 0.5	1.125 05	1.125 080	1.125 087 92	1.125 087 90
+ 1.0	1.4283	1.428 49	1.428 570 9	1.428 571 4
+ 1.5	2.13	2.26	2.290	2.294
+ 1.75	2.82	3.19	4.52	5.00

In view of the good agreement of the nonlinear PA's and the exact function in the region $-1 \le z \le +1$, it is apparent that the higher coefficients of the expansion of the PA in a Legendre series must agree to a high precision with the expansion coefficients of the exact function. If we introduce $r_l = a_l^{\mathrm{PA}}/a_l^{\mathrm{exact}}$ ($l = 0, 1, \cdots$), we have, e.g., in the case under consideration

$$r_7 = 0.9994, \quad r_8 = 0.9975, \quad r_9 = 0.9937.$$

This property might be of interest if an amplitude is known experimentally and one wants to decompose it into its partial waves. Fitting the amplitude with a PA-ansatz with a limited number of partial waves, one may also obtain some information about higher ones.

Finally, we check whether or not the residues of the poles can give us any information about the imaginary part on the cut. If (a, b) is an interval on the cut containing a pole p of the PA, we expect that

$$\int_a^b \mathrm{Im} f_+(z)dz = -\pi \,\mathrm{Res} f(z)\big|_{z=p} \tag{8}$$

would be a good approximation (f_+ refers to the upper side of the cut). The range (a, b) of the integration is

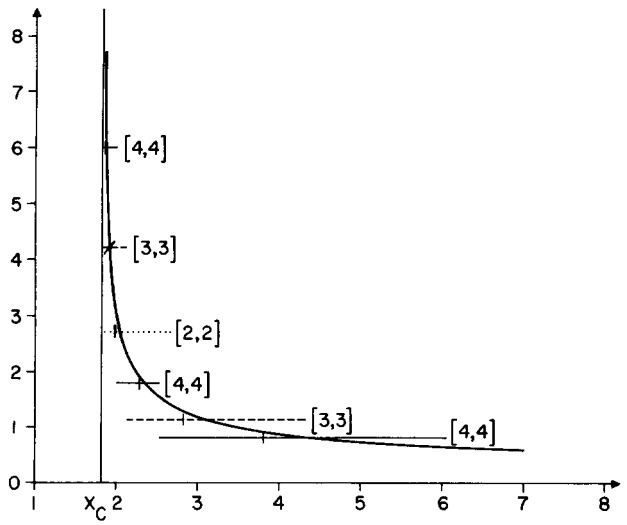

FIG. 1. For the generating function of the Legendre polynomials the mean values for the imaginary part on the cut are presented, calculated from the residues of the poles s_ν for the [2, 2] through [4, 4]-PA's. The poles themselves are indicated by small vertical lines.

certainly ambiguous, but the following choice seems to be a natural one:

(1) from the beginning of the cut to the first zero,

(2) two zeroes enclosing the respective pole.

Our results are represented in Fig. 1 for the [2, 2]–[4, 4] approximants, where the horizontal lines represent the mean values

$$(-\pi \operatorname{Res} f(z)|_{z=s_\nu})/(b-a)$$

in the respective intervals (a, b) containing the poles s_ν, which are also indicated by small vertical lines. We observe that actually the mean values at the poles themselves are a very good approximation. This property might be useful whenever one wants to calculate an imaginary part of a scattering amplitude in an unphysical kinematical region, which is not accessible by other means.

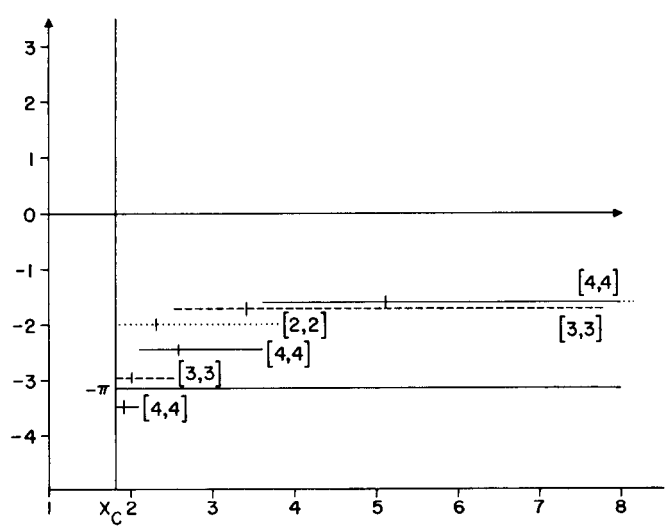

FIG. 2. Same as Fig. 1 for the logarithm.

B. The logarithm

Our second example is

$$f(z) = \ln\frac{1-az}{1-a} = \left[Q_0\left(\frac{1}{a}\right) + Q_1\left(\frac{1}{a}\right)\right]P_0(z) + \sum_{l=1}^{\infty}\left[Q_{l+1}\left(\frac{1}{a}\right) - Q_{l-1}\left(\frac{1}{a}\right)\right]P_l(z).$$

In order to have a reasonable comparison with the former case, we chose the cut starting at the same point x_C, which means $1/a = 1.817$. Our numerical results concerning the improvement of convergence are presented in Table II and they are again very good. Our main interest is, however, the approximation of the imaginary part on the cut, which is constant in this case. Figure 2 shows that the poles and zeroes are further distanced, which was to be expected. The accuracy at the beginning of the cut is reasonably good and one seems to observe a convergence with higher orders.

TABLE II. Same as Table I for the logarithm.

z	PS	LP	NP	exact
− 3.5	—	1.865	1.8727	1.8734
− 1.75	1.42	1.4735	1.474 13	1.474 16
− 1.0	1.237 97	1.238 06	1.238 0783	1.238 0784
− 0.5	1.042 62	1.042 655	1.042 653 57	1.042 653 64
− 0.0	0.799 537	0.799 521	0.799 527 67	0.799 527 58
+ 0.5	0.477 65	0.477 65	0.477 627 52	0.477 627 55
+ 1.0	0.000 20	0.000 18	0.000 0013	0.0
+ 1.5	− 0.854	− 0.907	− 0.942	− 0.947
+ 1.75	− 1.57	− 1.90	− 2.24	− 2.51

4. CONCLUSION

We have shown in this paper that it is possible to solve the nonlinear equations for the "nonlinear" Legendre PA's. It turned out, however, that for the fitting routine to be able to find the solutions, a double precision version for our equations was necessary (28 digits on the CDC 6600). For higher order PA's one loses significance again and one has to develop a reasonable way of evaluating the recursion relations for the D_λ with sufficient precision. Having this in mind, our procedure may become a powerful tool for the summation of Legendre series.

ACKNOWLEDGMENTS

The author is grateful to Professor G. Baker for a discussion on the Padé approximants introduced in this work. He appreciates many useful discussions with William Anderson about the FIT_4 routine. A correspondence with Dr. J. Engels, Bielefeld, Germany has been very useful. Financial support from NATO and the U.S. Atomic Energy Commission is gratefully acknowledged.

*Work partially supported by the U. S. Atomic Energy Commission.
†NATO Fellow.

[1] J. Fleischer, Nucl. Phys. B 37, 59 (1972). An earlier work on the "linear" approximants has been done by J. T. Holdeman Jr., Math. Comput. 23, 275 (1969). This has come to our knowledge after publication of the above work.

Vector space models of abstract quantum logics

Platon C. Deliyannis
Department of Mathematics, Illinois Institute of Technology, Chicago, Illinois
(Received 8 September 1972; revised manuscript received 3 October 1972)

A representation theorem for a class of logics is established in which the elements of the logic are represented as subspaces of some real vector space. A partial converse is also proved, which allows the systematic construction of various kinds of logics.

INTRODUCTION

One of the outstanding problems in the theory of quantum logics is the characterization of the logic of all (closed) subspaces of a Hilbert space in purely internal terms. By the use of methods in projective geometry significant results have been obtained which guarantee that an abstract logic is isomorphic to the lattice of all "closed" subspaces of some vector space (see, e.g., Refs. 1-4). One of the unsatisfactory features in this approach is that the field of scalars can be almost anything, and so no connection is established with the classical working model of physics. It is possible, by making further assumptions, to obtain that the field must be the reals, complexes, or quaternions, and also to rule out the noncommutative case.[1] It is unfortunate, however, that these hypotheses although verified in the Hilbert space case are of a rather nonphysical nature, not bearing any immediate physical interpretation. It should be stressed, though, that the Piron result[2,4] is quite sharp in case the scalars turn out to be the reals, complex, or quaternions, in that the vector space is then complete and the representing subspaces closed (in the topology generated by the norm).

In this paper we shall follow a different method to represent the elements of our logic as subspaces of a *real* vector space (equally well we can do the same over the complex field, should we wish to); the conditions we shall impose on the logic are quite plausible physically, and are satisfied in the classical cases. We shall also obtain by reversing the arguments a method of constructing logics of a certain kind, starting with a real vector space, a bilinear functional, and a convex cone. This class of logics is actually wider than the one we obtained the representation for and could be useful for various purposes; it does not appear unreasonable to hope even for some classification theorems.

Our method is elementary and in principle simple, while the arguments, although on occasion long, are basically straightforward.

We shall now describe the class of logics $\mathfrak{L}$ we shall be studying: the partial order $\leq$ (implication) and the complementation (negation) will satisfy, besides the usual conditions (i) to (iv), three more involving properties of states, as follows:

(i) For all $A, B \in \mathfrak{L}$ we have $A \leq B$ implies $B' \leq A'$ and $(A')' = A$.

(ii) With $\wedge, \vee$ denoting infimum and supremum respectively we have $A \wedge A' = 0, A \vee A' = I$ for some fixed $0, I \in \mathfrak{L}$ and all $A \in \mathfrak{L}$.

(iii) The orthomodular law: If $A \leq B$, then $A' \wedge B$ exists and $B = A \vee (A' \wedge B)$.

(iv) Existence of disjoint suprema: If $A_i \leq A_j'$ for $i \neq j$ then the supremum $\sum A_i$ exists.

(v) Quite fullness[5]: If for all states m, $mA = 1$ implies $mB = 1$ then $A \leq B$.

(vi) A weak form of the Jauch–Piron–Zierler axiom: if for some pure state m we have $m(A_i) = 1$ for all i, then the infimum A of the $\{A_i\}$ exists and $mA = 1$ also.

(vii) The pure states of the system generate all states: there exists a measurable space X such that to each m there corresponds a probability measure μ on X and a map $x \to m_x$ from X to the pure states with $mA = \int_X m_x(A)\, d\mu(x)$. In other words, every state is a mixture of pure states.

To obtain our representation theorem, we shall make two more assumptions on the behavior of the pure states. We shall state and make some preliminary use of them in the next section. The representation theorem follows, then the converse, and we conclude with some remarks and examples. We shall write $\mathfrak{M}$ for the set of all states and $\mathfrak{M}_p$ for the set of pure states.

ADDITIONAL HYPOTHESES AND PRELIMINARY RESULTS

We first exploit property (vii).

Proposition 1: For any $A \in \mathfrak{L}, A \neq 0$, there exists a pure state m with $mA = 1$. Further, the set $\{m \in \mathfrak{M}_p \mid mA = 1\}$ determines A.

Proof: Take any $m \in \mathfrak{M}$ with $mA = 1$ (see Ref. 5). Since we have $mA = \int_X m_x(A)\, d\mu(x)$ for some probability measure μ on X and $0 \leq m_x(A) \leq 1$, we see that $m_x(A) = 1$ a.e. with respect to μ and so $\{m \in \mathfrak{M}_p \mid mA = 1\} \neq \phi$. In fact, $mA = 1$ iff $m_x(A) = 1$ a.e., so that if $\{m \in \mathfrak{M}_p \mid mA = 1\} = \{m \in \mathfrak{M}_p \mid mB = 1\}$, then $m_x(A) = m_x(B)$ a.e. and hence $mA = 1$ iff $mB = 1$ for any $m \in \mathfrak{M}$. By (v) we then obtain $A = B$.

We now consider, for a fixed $m \in \mathfrak{M}_p$, the set of all $A \in \mathfrak{L}$ for which $mA = 1$; by (vi) the infimum of this set exists and occurs with certainty in m.

Notation: Write L_m for the element $\inf\{A \mid mA = 1\}$.

Definition: For $n, m \in \mathfrak{M}_p$, the number $n(L_m)$ will be called the *probability of transition* from the state n to the state m; we shall write it as $\langle n \mid m \rangle$.

We are introducing this terminology on the following grounds: since our relation $\leq$ on events is being interpreted as implication, the event L_m is the "ultimate cause" of anything that happens in m; thus the probability of L_m occurring in the state n is the probability essentially of n switching over to m.

Hypothesis I: The transition probability is symmetric in its two arguments: $\langle n \mid m \rangle = \langle m \mid n \rangle$.

It is easy to see that this holds in the two classical models, the Hilbert space model and the ordinary probability theory model.

Now consider two pure states n, m and suppose that $\langle n \mid m \rangle = 1$. The interpretation of $\langle n \mid m \rangle$ as a transition probability forces us to the conclusion that n and m cannot be distinguished; it also means that all events occurring with certainty in one state also occur with certainty in the other. So we shall formulate our last hypothesis as follows:

Hypothesis II: The set of all events occurring with certainty in some pure state determines this state completely.

We shall use this in the form $\langle n \mid m \rangle = 1$ implies $n = m$, or the obviously equivalent form, $L_n = L_m$ implies $n = m$.

It is now quite clear that we have the following:

Proposition 2: If $\langle m_1 \mid m \rangle = \langle m_2 \mid m \rangle$ for all $m \in \mathfrak{M}_p$, then $m_1 = m_2$.

Proposition 3: For every $m \in \mathfrak{M}_p$, L_m is an atom of $\mathfrak{L}$. Conversely, for each atom A of $\mathfrak{L}$ there is a unique $m \in \mathfrak{M}_p$ with $A = L_m$.

Proof: Suppose that $B < L_m$, so that $L_m = B + C$ with $C \neq 0$. By Proposition 1 there is a pure state n for which $nC = 1$, so that $n(L_m) = 1$ also. This implies $n = m$ and therefore $mC = 1$ hence $C \geq L_m$ and $B = 0$. The converse is clear from Proposition 1 and Hypothesis II.

Proposition 4: Every $A \neq 0$ is the sum of (disjoint) atoms.

Proof: Consider a maximal pairwise disjoint family of L_m contained in A, which exists by Proposition 1. Their sum, if distinct from A, admits a complement in A which will be not 0, and thus contains a new L_n disjoint from all the others; this is impossible by maximality of the chosen set.

Remark: We have L_n disjoint from L_m iff $\langle n \mid m \rangle = 0$. Because if they are disjoint, then $p(L_n) + p(L_m) \leq 1$ for all states p, and taking $p = n$ we find $n(L_m) = 0$. Conversely, if $n(L_m) = 0$ we have $n(L'_m) = 1$, i.e., $L_n \leq L'_m$.

THE REPRESENTATION SPACE

Consider the vector space F spanned over the reals by the set $\mathfrak{M}_p$, i.e., the space of all real valued functions on $\mathfrak{M}_p$ with finitely many nonzero values. Since $\mathfrak{M}_p$ forms a basis, we can define uniquely a bilinear form $\langle \ | \ \rangle$ on F by $\langle \sum a_i m_i \mid \sum b_j n_j \rangle = \sum a_i b_j \langle m_i \mid n_j \rangle$; this is, of course, symmetric, but may be degenerate. So, instead of F, we shall work with its quotient by $K = \{ \sum a_i m_i \mid \sum a_i \langle m_i \mid m \rangle = 0, \text{for all } m \in \mathfrak{M}_p \}$, and call it $\mathfrak{K}$. It is clear that $\mathfrak{K}$ inherits the form $\langle \ | \ \rangle$ which is now nondegenerate as well as symmetric.

Each element of $\mathfrak{K}$ will contain at most one pure state because of Proposition 2, and in such a case we shall use the name of the state to denote the corresponding element of $\mathfrak{K}$. In this way $\mathfrak{M}_p$ is imbedded intact in $\mathfrak{K}$ and evidently spans $\mathfrak{K}$.

Notation: We shall write M_A for the set of all $m \in \mathfrak{M}_p$ for which $mA = 1$ and H_A for the subspace of $\mathfrak{K}$ spanned by M_A. Also, for $S \subseteq \mathfrak{K}$ we shall write $S^\perp$ for the set $\{x \mid \langle y \mid x \rangle = 0 \text{ for all } y \in S\}$.

Proposition 5: For any $A, B \in \mathfrak{L}$ we have $A \leq B$ iff $M_A \subseteq M_B$.

Proof: The first half is trivial, so suppose that $M_A \subseteq M_B$. Take any state m with $mA = 1$ and analyze into pure states: $m = \int_X m_x d\mu(x)$; as before we have $m_x(A) = 1$ for almost all x and since m_x is pure, we also obtain $m_x(B) = 1$ for almost all x. This implies that $mB = 1$ and by quite fullness we have $A \leq B$.

Proposition 6: For any $A, B \in \mathfrak{L}$ we have $A \leq B$ iff $H_A \subseteq H_B$.

Proof: Again half is obvious, so assume that $H_A \subseteq H_B$ and $m \in M_A$. Since then $m \in H_B$ we shall have $m = \sum a_i m_i$ with $m_i \in M_B$, i.e., $\langle m \mid n \rangle = \sum a_i \langle m_i \mid n \rangle$ for all $n \in \mathfrak{M}_p$. By symmetry we obtain $n(L_m) = \sum a_i n(L_{m_i})$; call b_j the positive a_i and $-c_k$ the negative a_i to obtain $n(L_m) \leq n(L_m) + \sum c_k n(L_{m_k}) = \sum b_j n(L_{m_j})$. But $m_j \in M_B$ so $L_{m_j} \leq B$, hence $n(L_m) \leq (\sum b_j) n(B)$ for any $n \in \mathfrak{M}_p$. Using (vii) again, we have this last relation valid for all $n \in \mathfrak{M}$, and by the lemma below we get $L_m \leq B$ or $mB = 1$. Thus $M_A \subseteq M_B$ or $A \leq B$.

Lemma: If for some $k > 0$ and all states m we have $mA \leq kmB$, then $A \leq B$.

Proof: Because then $mB = 0$ implies $mA = 0$, i.e., $m(B') = 1$ implies $m(A') = 1$, or $B' \leq A'$.

Remark: The above argument shows that $m \in \mathfrak{M}_p$, $m \in H_A$ imply $mA = 1$, i.e., $\mathfrak{M}_p \cap H_A = M_A$.

Proposition 7: For any $A \in \mathfrak{L}$ we have $H_{A'} \subseteq (H_A)^\perp$; also, if $H_B \subseteq (H_A)^\perp$, then $B \leq A'$.

Proof: To have $\sum a_i m_i \in H_{A'}$ means $m_i(A') = 1$, or $m_i A = 0$. Now if $m \in M_A$, then $L_m \leq A$ and hence $m_i(L_m) = 0$, or $\langle m_i \mid m \rangle = 0$. Thus $\langle \sum a_i m_i \mid m \rangle = 0$ for any $m \in M_A$ and hence for any element of H_A, i.e., $H_{A'} \subseteq (H_A)^\perp$. Now let $H_B \subseteq (H_A)^\perp$ and take $m \in M_B$. For each $n \in M_A$ we then have $\langle m \mid n \rangle = 0$, or $m(L_n) = 0$ for every $L_n \leq A$. But by proposition 4 we have A the disjoint sum of such $L'_n s$, and thus $mA = 0$. We have shown that $mB = 1$ implies $mA = 0$, or $B \leq A'$ by quite fullness.

Corollary: For $A, B \in \mathfrak{L}$ we have A, B disjoint iff the subspaces H_A, H_B are orthogonal. Hence $H_{\sum A_i}$ is the smallest space H_B containing the direct sum of the spaces H_{A_i}.

Proposition 8: Let $\{m_i\}$ be a maximal orthogonal set of pure states in H_A (hence in M_A). For any pure state m we have $mA = \sum \langle m \mid m_i \rangle$.

Proof: Our hypothesis means that $\{L_{m_i}\}$ is a maximal disjoint family of atoms contained in A, and thus implies that A is their supremum. So $mA = \sum m(L_{m_i}) = \sum \langle m \mid m_i \rangle$.

The converse also holds:

Proposition 9: Let $\{m_i\}$ be any orthogonal family of pure states. Then there exists a unique $A \in \mathfrak{L}$ such that $\{m_i\}$ is maximal orthogonal in H_A.

Proof: Let A be the sum of the corresponding L_{m_i}, so that each m_i is in H_A. Now, if $\langle m|m_i\rangle = 0$ for all i, we cannot have $m \in M_A$, because then $mA = 1$, while $mA = \sum m(L_{m_i}) = 0$. Therefore $\{m_i\}$ is maximal orthogonal in H_A. Uniqueness now follows from Proposition 8.

We shall summarize now what the above propositions, put together, have established. Later we shall state this result formally and in some detail as a theorem: There exists a one-to-one correspondence between the elements of $\mathfrak{L}$ and certain subspaces of $\mathfrak{K}$, so that all algebraic and order structure of $\mathfrak{L}$ is described in terms of inclusion and orthocomplements with respect to some bilinear nondegenerate form $\langle\ |\ \rangle$: further the pure states of $\mathfrak{L}$ correspond to vectors in $\mathfrak{K}$ and their values on the elements of $\mathfrak{L}$ again obtain via the form $\langle\ |\ \rangle$.

Now we come to the details. First note that $\mathfrak{M}_p$ generates a convex cone $\mathfrak{C}$ in $\mathfrak{K}$ having the elements of $\mathfrak{M}_p$ as extreme rays. Because $\mathfrak{C}$ will consist by definition of all $\sum a_i m_i$ where $m_i \in \mathfrak{M}_p$ and $a_i > 0$ so that if $m = \sum a_i m_i$, we shall have $\sum a_i = 1$ [since $m(I) = 1$] and therefore m would be a mixture of states, which is not possible. On the other hand no sum $\sum a_i m_i$ with $a_i > 0$ can produce an extreme ray without consisting of a single term. Also we see that for any $u, v \in \mathfrak{C}$ we have $\langle u|v\rangle \geq 0$.

Theorem 1: There exists a vector space $\mathfrak{K}$ over the reals carrying a bilinear symmetric nondegenerate form $\langle\ |\ \rangle$ and containing a convex cone $\mathfrak{C}$ so that:

(i) For any $u, v \in \mathfrak{C}$ we have $\langle u|v\rangle \geq 0$.

(ii) The set $\mathcal{E}$ of extreme rays of $\mathfrak{C}$ generates $\mathfrak{K}$, and no such ray is self-orthogonal.

(iii) If $\mathfrak{N} = \{u/\langle u|u\rangle^{1/2} \mid u \in \mathcal{E}\}$, then for any $u, v \in \mathfrak{N}$ we have $\langle u|v\rangle \leq 1$ with equality iff $u = v$.

(iv) For any orthogonal set of vectors $\{u_i\}$ in $\mathfrak{N}$ and any $u \in \mathfrak{N}$ we have $\sum \langle u|u_i\rangle \leq 1$ with equality iff $\{u_i\}$ is maximal.

(v) For any orthogonal set $\{u_i\}$ in $\mathfrak{N}$, let N be the set of all $u \in \mathfrak{N}$ for which $\sum \langle u|u_i\rangle = 1$. Then

  (a) no linear combination of vectors in N is orthogonal to all u_i, and

  (b) if $\{v_j\}$ is a maximal orthogonal family in N then $\sum \langle u|u_i\rangle = \sum \langle u|v_j\rangle$ for all $u \in \mathfrak{N}$.

Further, there exists a one-to-one correspondence $A \to H_A$ between $\mathfrak{L}$ and certain subspaces of $\mathfrak{K}$ such that:

(1) $A \leq B$ iff $H_A \subseteq H_B$.

(2) $H_{A'}$ is the largest subspace H_B contained on the orthogonal complement of H_A.

(3) A subspace H of $\mathfrak{K}$ is in the range of this map iff:

  (a) $\mathfrak{N} \cap H$ spans H.

  (b) for any maximal orthogonal family $\{u_i\}$ in $\mathfrak{N} \cap H$ and any $u \in \mathfrak{N}$, the number $\sum \langle u|u_i\rangle$ depends only on u; call it $u(H)$.

  (c) $u \in \mathfrak{N} \cap H$ iff $u(H) = 1$.

(4) There exists a one-to-one correspondence $m \to u_m$ between the pure states of $\mathfrak{L}$ and $\mathfrak{N}$ such that $m(A) = u_m(H_A)$.

Proof: It is clear that (i), (ii), (iii) are valid by the very definition of the objects involved. Parts (iv), (v) are a restatement of Propositions 8, 9. Parts (1) and (2) are Propositions 6 and 7, while (4) follows again from Propositions 8, 9. It is part (3) that requires some elaboration. We have already noted that $M_A = \mathfrak{M}_p \cap H_A$, so that (a) follows. Since (b), (c) are consequences of Propositions 8, 9 all we have left is the converse. So consider any subspace H of $\mathfrak{K}$ satisfying (3) and take a maximal orthogonal family $\{m_i\}$ in $\mathfrak{N} \cap H$; let A be the sum of the corresponding L_{m_i} so that the family $\{m_i\}$ is also maximal orthogonal in H_A, according to Proposition 9. For any $m \in \mathfrak{M}_p$ we have $m(A) = \sum \langle m|m_i\rangle$ which by (b) is the same as $m(H)$; hence, using (c), we have $m \in M_A$ iff $m \in \mathfrak{N} \cap H$ so that H and H_A are spanned by the same set of vectors.

THE CONSTRUCTION

We shall now turn the argument around and show that under certain conditions a suitable family of subspaces in some real vector space constructed as described below will form a quantum logic $\mathfrak{L}$.

So we consider a real vector space $\mathfrak{K}$ with a bilinear symmetric nondegenerate form $\langle\ |\ \rangle$ and select a convex cone $\mathfrak{C}$ such that properties (i), (ii), (iii), (iv) of Theorem 1 hold. It is convenient to select our family $\mathfrak{L}$ in two stages.

The family $\mathfrak{K}$ consists of all subspaces K of $\mathfrak{K}$ such that:

(i) $\mathfrak{N} \cap K$ spans K.

(ii) for any maximal orthogonal family $\{u_i\}$ in $\mathfrak{N} \cap K$ and for any $u \in \mathfrak{N}$ the number $\sum \langle u|u_i\rangle$ depends only on u; call it $u(K)$, and note that $0 \leq u(K) \leq 1$.

The family $\mathfrak{L}$ consists of all $K \in \mathfrak{K}$ for which $u \in \mathfrak{N}$ and $u(K) = 1$ imply $u \in K$. Partial order on $\mathfrak{L}$ is just inclusion; the orthocomplement A' of an element $A \in \mathfrak{L}$ is defined to be the subspace of $\mathfrak{K}$ spanned by $\mathfrak{N} \cap A^\perp$. To show that $A' \in \mathfrak{L}$, we shall need the following lemma:

Lemma: If $\langle u|u_i\rangle = 0$ where $u, u_i \in \mathfrak{N}$ and $\{u_i\}$ is maximal orthogonal in $A \in \mathfrak{L}$, then $u \in A^\perp$.

Proof: Take any $v \in \mathfrak{N} \cap A$ and augment it to a maximal set $\{v, v_j\}$ in A; by (ii) we shall have $\sum \langle u|u_i\rangle = u|v + \sum \langle u|v_j\rangle$, and, as $\langle u|u_i\rangle$ is 0, we have $\langle u|v\rangle = 0$ also.

Proposition 10: For each $A \in \mathfrak{L}$ we also have $A' \in \mathfrak{L}$, and $u(A') = 1 - u(A)$ for all $u \in \mathfrak{N}$.

Proof: We first show that $A' \in \mathfrak{K}$. Since $\mathfrak{N} \cap A' \supseteq \mathfrak{N} \cap A^\perp$ and $\mathfrak{N} \cap A^\perp$ spans A', we have that (i) is satisfied. Now consider a maximal orthogonal set $\{u_j\}$ in $\mathfrak{N} \cap A'$ and a maximal orthogonal set $\{v_i\}$ in $\mathfrak{N} \cap A$; we shall show that $\{u_j, v_i\}$ is maximal orthogonal in $\mathfrak{N}$. For if u is orthogonal to all these, then by the lemma it is in $\mathfrak{N} \cap A^\perp$ and we should be able to adjoin it to the family $\{u_j\}$ which cannot happen by hypothesis. Thus we have $\sum \langle u|u_j\rangle + \sum \langle u|v_i\rangle = 1$ for any $u \in \mathfrak{N}$, and thus $\sum \langle u|u_i\rangle = 1 - u(A)$, which means that (ii) holds. Finally we need $u(A') = 1$ to imply $u \in A'$; but as we saw, $u(A') = 1 - u(A)$, and so we have $u(A) = 0$. This means $\langle u|v\rangle = 0$ for any $v \in \mathfrak{N} \cap A$, since any such v forms part of a maximal orthogonal set in $\mathfrak{N} \cap A$. So $u(A') = 1$ means $u \in A^\perp$, i.e., $u \in A'$.

Theorem 2: The map $A \to A'$ defined above is an orthocomplementation on $\mathfrak{L}$:

(i) $(A')' = A$.

(ii) $A \leq B$ implies $B' \leq A'$.

(iii) $A \wedge A' = 0, A \vee A' = I$, where $0, I$ are the subspaces $\{0\}, \mathfrak{K}$ (evidently members of $\mathfrak{L}$).

Proof: (i) By Proposition 10 we have $u((A')') = u(A)$ for any $u \in \mathfrak{N}$. On the other hand we see that given a subspace $B \in \mathfrak{L}$ we have $u \in B$ iff $u(B) = 1$ (provided $u \in \mathfrak{N}$): half of this is the definition of $\mathfrak{L}$, and the other half is obtained by imbedding u in a maximal orthogonal family to obtain $u(B) = \langle u | u \rangle = 1$. Thus we have for any $u \in \mathfrak{N}$ that $u \in (A')'$ iff $u \in A$.

(ii) Same principle: Let $A \leq B$, or $\mathfrak{N} \cap A \subseteq \mathfrak{N} \cap B$; then $u \in B'$ iff $u(B') = 1$ iff $u(B) = 0$ which implies $u(A) = 0$ or $u(A') = 1$ or $u \in A'$.

(iii) Suppose $B \leq A, B \leq A'$; we must show that $B = 0$, i.e., $\mathfrak{N} \cap B = \emptyset$. Taking $u \in \mathfrak{N} \cap B$ we obtain $u \in \mathfrak{N} A$ and $u \in \mathfrak{N} \cap A'$ so that u must be orthogonal to $\mathfrak{N} \cap A$; but this implies $\langle u | u \rangle = 0$ which contradicts condition (ii) of Theorem 1. Now suppose $B \geq A, B \geq A'$ and choose maximal orthogonal sets $\{u_i\}, \{v_j\}$ in $\mathfrak{N} \cap A$ and $\mathfrak{N} \cap A'$. As before we have $\{u_i, v_j\}$ maximal orthogonal in $\mathfrak{N}$, and thus we obtain $\sum \langle u | u_i \rangle + \sum \langle u | v_j \rangle = 1$ for all $u \in \mathfrak{N}$. But since all u_i and v_j are in B this means $u \in B$ for all $u \in \mathfrak{N}$, and as $\mathfrak{N}$ spans $\mathfrak{K}$ we have $B = \mathfrak{K}$.

It does not seem possible to proceed much further in establishing properties of $\mathfrak{L}$ such as orthomodularity without further hypotheses on $\mathfrak{K}$. Observe that we have used only four of the five properties listed in Theorem 1. This fifth is precisely what we need to prove existence of suitable elements of $\mathfrak{L}$. We shall therefore assume it from now on.

Proposition 11: Let $\{u_i\}$ be any orthogonal set in $\mathfrak{N}$. Then there exists a unique $A \in \mathfrak{L}$ with $\{u_i\}$ maximal orthogonal in $\mathfrak{N} \cap A$.

Proof: Consider the set $\mathfrak{N}_0$ of all $u \in \mathfrak{N}$ such that $\sum \langle u | u_i \rangle = 1$ and let A be the subspace it spans; clearly all $u_i \in A$. Also, any $u \in \mathfrak{N}_0$ is obviously in A, so $\mathfrak{N} \cap A \supseteq \mathfrak{N}_0$ and hence $\mathfrak{N} \cap A$ spans A. Part (b) of property (v) completes the proof of $A \in \mathfrak{K}$. Finally, let $u(A) = 1$ with $u \in \mathfrak{N}$; since this means $\sum \langle u | u_i \rangle = 1$, we have $u \in \mathfrak{N}_0$, or $u \in A$.

Theorem 3: $\mathfrak{L}$ is orthomodular and any family of pairwise disjoint elements admits a supremum.

Proof: Assuming $A \leq B$ choose a maximal orthogonal set $\{u_i\}$ in $\mathfrak{N} \cap A$ and adjoin $\{v_j\}$ to obtain a maximal orthogonal set $\{u_i, v_j\}$ in $\mathfrak{N} \cap B$. Let C be the element of $\mathfrak{L}$ determined by the $\{v_j\}$ according to Proposition 11. Since u_i is orthogonal to v_j for all i, j we have the subspaces A, C orthogonal and hence disjoint as elements of $\mathfrak{L}$; clearly $C \leq B$. We must show that if D is disjoint from $A, D \leq B$, then $D \leq C$. So consider $u \in \mathfrak{N} \cap D$; then $u \in \mathfrak{N} \cap B$ and so $\sum \langle u | u_i \rangle + \sum \langle u | v_j \rangle = 1$. But D is disjoint from A and so $\langle u | u_i \rangle = 0$ which implies $\sum \langle u | v_j \rangle = 1$, i.e., $u \in C$.

A similar argument shows immediately that if the A_i are pairwise disjoint, by choosing maximal orthogonal families $\{u_{i,k}\}$ in the A_i and putting them together, the element obtained via Proposition 11 will be the supremum of the A_i.

We have thus established that the five conditions of Theorem 1 allow us to construct a quantum logic. We shall now investigate the existence of arbitrary infima (or dually suprema).

Definition: We shall call two orthogonal sets $\{u_i\}$, $\{v_j\}$ in $\mathfrak{N}$ *equivalent if* for all $u \in \mathfrak{N}$ we have $\sum \langle u | u_i \rangle = \sum \langle u | v_j \rangle$.

Lemma: Two sets $\{u_i\}, \{v_j\}$ are equivalent iff they generate the same element of $\mathfrak{L}$.

Proof: Letting A, B be these two subspaces note that $A = B$ implies equivalence since each will be maximal orthogonal in $\mathfrak{N} \cap A$. So assume equivalence; but then we have $u(A) = u(B)$ for all $u \in \mathfrak{N}$, which implies $A = B$.

Proposition 12: Let $\{A_i\}$ be a family of elements of $\mathfrak{L}$ and $\mathfrak{N}_0$ be the set $\mathfrak{N} \cap (\bigcap_i A_i)$. The infimum of $\{A_i\}$ exists if any two maximal orthogonal sets in $\mathfrak{N}_0$ are equivalent, in which case the infimum is generated by any one of them.

Proof: Suppose the infimum exists, and call it A. For any orthogonal set $\{u_i\}$ in $\mathfrak{N}_0$ we see that the corresponding element of $\mathfrak{L}$ will be contained in each A_i, hence $\leq A$; so $\{u_i\}$ is in A, and since each vector in $\mathfrak{N}_0$ is part of some orthogonal set we have $\mathfrak{N}_0 \subseteq A$ or $\mathfrak{N} \cap A = \mathfrak{N}_0 \cap A$. Further, any maximal orthogonal set in $\mathfrak{N}_0$ will generate A, and thus any two such will be equivalent. Conversely, assume that any two maximal orthogonal sets in $\mathfrak{N}_0$ are equivalent and let A be the subspace generated by any one of them; evidently $A \leq A_i$ for all i. But if $B \leq A_i$ for all i, then $\mathfrak{N} \cap B \subseteq \mathfrak{N}_0$ and hence any maximal orthogonal set in $\mathfrak{N} \cap B$ extends to such a set in $\mathfrak{N}_0$; this generates A and thus $B \leq A$.

Proposition 13: Every vector $u \in \mathfrak{N}$ gives rise to a state of $\mathfrak{L}$ via the map $A \to u(A)$.

Proof: Recall that the supremum of a family of disjoint elements is generated by the union of maximal orthogonal families of vectors, one for each element; this clearly implies additivity.

Proposition 14: $\mathfrak{L}$ is quite full.

Proof: Supposing that $mB = 1$ for all states m for which $mA = 1$, take a vector $u \in \mathfrak{N} \cap A$; then $u(A) = 1$ and so $u(B) = 1$, or $u \in \mathfrak{N} \cap B$. But this means $A \leq B$.

EXAMPLES AND FINAL REMARKS

Example 1: The simplest case to consider is that of a commutative $\mathfrak{L}$. Our basic hypotheses imply that $\mathfrak{L}$ is a complete atomic Boolean algebra and hence isomorphic to all subsets of some set Ω; the points of Ω are in a one-to-one correspondence with the pure states of $\mathfrak{L}$ via the map $\omega \to \delta_\omega$, the Dirac measure at the point ω. Thus distinct pure states are orthogonal and $\mathfrak{K}$ is a pre-Hilbert space with respect to $\langle \ | \ \rangle$.

Example 2: Let $\mathfrak{L}$ now be the logic of all closed subspaces of a Hilbert space H with inner product $(\ | \ )$. The pure states being rays in H we identify them in the usual way to unit vectors in H. For any such state ψ we

see that L_ψ is just the ray of ψ and we identify it to the projection P_ψ: $\varphi \to (\varphi|\psi)\psi$, so that $\langle \varphi|\psi\rangle = \varphi(L_\psi) = (P_\psi\varphi|\varphi) = |(\varphi|\psi)|^2$. Since the vectors $\sum a_i\psi_i$ of $\mathcal{H}$ can be identified to the maps $\psi \to \sum a_i\langle\psi_i|\psi\rangle$, we see that the vector $u = \sum a_i\psi_i$ of $\mathcal{H}$ gives rise to the map $\psi \to \sum a_i|(\psi_i|\psi)|^2 = (T_u\psi|\psi)$ where T_u is the operator $\sum a_i\psi_i \otimes \psi_i$: $T_u\psi = \sum a_i(\psi|\psi_i)\psi_i$. Since a map of the form $\psi \to (A\psi|\psi)$ determines the operator A completely, we have defined a map $u \to T_u$ from $\mathcal{H}$ to operators on H, which is evidently linear. Each T_u is self-adjoint since the a_i are real, and of finite rank. To evaluate $\langle u|v\rangle$ in terms of T_u, T_v, note that we can always assume that in the form $T_u = \sum a_i\psi_i \otimes \psi_i$ the vectors ψ_i are a complete set of orthogonal eigenvectors of T_u with a_i the corresponding nonzero eigenvalues; this means that we can write $u = \sum a_i\psi_i$ for the same a_i, ψ_i. So we write $v = \sum b_j\varphi_j$ with the same conditions imposed and calculate:

$$\begin{aligned}\langle u|v\rangle &= \sum_{i,j} a_ib_j\langle\psi_i|\varphi_j\rangle \sum_{i,j} a_ib_j|(\psi_i|\varphi_j)|^2\\ &= \sum_{i,j} a_ib_j(\varphi_i|\psi_j)(\psi_j|\varphi_i)\\ &= \sum_i\Big(\sum_j b_j(\psi_i|\varphi_j)\varphi_j\Big|\sum_k a_k(\psi_i|\psi_k)\psi_k\Big)\\ &= \sum_i (T_u\psi_i|T_v\psi_i) = \mathrm{Tr}(T_uT_v).\end{aligned}$$

This shows also that if $T_u = 0$, then $u = 0$, so that we find $\mathcal{H}$ imbedded in the space of Hilbert–Schmidt operators of H isometrically.

Remark 1: It is interesting to observe that even if $\mathfrak{L}$ is given as a class of subspaces of a Hilbert space, our construction does not recapture the space except perhaps in an indirect way. So Theorems 1 and 2 do not form a dual pair.

Remark 2: It should also be noticed that it is by no means clear that the vectors in $\mathfrak{N}$ exhaust the pure states of $\mathfrak{L}$; in fact we do not know whether they are pure states to begin with. It is therefore impossible without further deeper study to decide whether $\mathfrak{L}$, constructed according to the above satisfies all hypotheses of Theorem 1. To be more precise, we should rephrase this as follows: What geometric conditions on $\mathcal{H}$ should we impose in order to have the hypotheses of Theorem 1 satisfied? We shall see next an example of some $\mathcal{E}$ for which some of these are violated. This question involves the calculation of the set $\mathfrak{M}$ of all states of $\mathfrak{L}$, which appears to be a complicated problem as exemplified by Gleason's theorem.

Example 3: Consider a pre-Hilbert space $\mathcal{H}$ (over the reals) with inner product $\langle\ |\ \rangle$, choose a fixed vector $a \in \mathcal{H}$ of unit length, and let $\mathcal{C} = \{u \in \mathcal{H} | \langle a|u\rangle \geq \frac{1}{2}\sqrt{2}\,\|u\|\}$. It is straightforward to verify that $\mathcal{C}$ is a convex cone; we shall show that $\mathcal{E} = \{u \in \mathcal{H} | \langle a|u\rangle = \frac{1}{2}\sqrt{2}\,\|u\|\}$. If this holds and $u = v + w$, with $v, w \in \mathcal{C}$, then $\frac{1}{2}\sqrt{2}\,\|u\| = \langle a|u\rangle = \langle a|v\rangle + \langle a|w\rangle \geq \frac{1}{2}\sqrt{2}\,(\|v\| + \|w\|)$ so that u, v, w are on a line. Conversely, if $\langle a|u\rangle > \frac{1}{2}\sqrt{2}\,\|u\|$ and u is not on a line with a, we can find a scalar k such that $ku = a + v$ for some $v \in \mathcal{C}$, which means that u is not extreme (e.g., we can take

$$k = [\|u\|\langle a|u\rangle + (\|u\|^2 - \langle a|u\rangle^2)^{1/2}]/(2\langle a|u\rangle^2 - \|u\|^2)$$

the case a, u on a line can be treated similarly.

To verify the five conditions of Theorem 1, we let $Pu = \langle a|u\rangle a$ and $Qu = u - Pu$, so that $PQ = QP = 0$ and $P + Q$ is the identity. We have $\|Qu\|^2 = \|u\|^2 - \|Pu\|^2 = \|u\|^2 - \langle a|u\rangle^2$, so that if $u \in \mathcal{C}$ we obtain $\|Qu\| \leq \frac{1}{2}\sqrt{2}\,\|u\|$. Therefore, for $u, v \in \mathcal{C}$ we get $\langle u|v\rangle = \langle Pu + Qu|Pv + Qv\rangle = \langle Pu|Pv\rangle + \langle Qu|Qv\rangle = \langle a|u\rangle\langle a|v\rangle + \langle Qu|Qv\rangle \geq \frac{1}{2}\|u\|\,\|v\| - \frac{1}{2}\|u\|\,\|v\| = 0$.

Since $\langle\ |\ \rangle$ is an inner product no ray is self-orthogonal. To see that $\mathcal{E}$ generates $\mathcal{H}$ we take $w \in \mathcal{H}$ and write $w = Pw + Qw$. Since either Pw or $-Pw$ is in $\mathcal{E}$ we consider $Qw = z$, which is orthogonal to a; then for any scalar k we have $\langle a|a + kz\rangle = 1$ and so it suffices to find a k for which $\frac{1}{2}\|a + kz\|^2 \leq 1$, which is evidently possible. Property (iii) follows from the condition that $\langle\ |\ \rangle$ is an inner product and all vectors in $\mathfrak{N}$ are of unit length lying on the same side of the hyperplane $\langle a|u\rangle = 0$. Now we determine the orthogonal sets in $\mathfrak{N}$. Observe that $\langle u|v\rangle = 0$ means $\frac{1}{2} + \langle Qu|Qv\rangle = 0$ or $\langle Qu|Qv\rangle = -\frac{1}{2}$; since $\|Qu\| = \|Qv\| = \frac{1}{2}$, we obtain the orthogonality condition in the form $Qu = -Qv$ or $u + v = \sqrt{2}\,a$. This means that the nonempty orthogonal sets in $\mathfrak{N}$ are either singletons or contain exactly two members related by the above equation. Property (iv) now follows trivially. Finally we verify (v) for the case of an orthogonal set with two elements (the other being obvious) by noting that in this case N consists of all of $\mathfrak{N}$.

The structure of $\mathfrak{L}$ is now transparent. Except 0 and I the only other subspaces are one-dimensional and are just the various rays of $\mathfrak{N}$. It is worth noticing that $A' \neq A^\perp$. We can describe $\mathfrak{L}$ in an abstract way by considering a family of 4-element Boolean algebras $\{\mathfrak{B}_\lambda\}_{\lambda\in\Lambda}$ identifying all 0_λ's to 0, all I_λ's to I and taking their set union. The states of $\mathfrak{L}$ are in a one-to-one correspondence with the points of a Hilbert cube $\mathsf{X}_{\lambda\in\Lambda} J_\lambda$, where all the J_λ are copies of the interval $[0, 1]$. If $\mathfrak{B}_\lambda = \{0, A_\lambda, A'_\lambda, I\}$ and m is a state, then m is completely determined by the numbers $p_\lambda = m(A_\lambda)$; and vice versa, any such family of numbers produces a state of $\mathfrak{L}$ when assigned as values of the various A_λ. Thus a pure state of $\mathfrak{L}$ must be a corner point of the cube, since the convexity structure is preserved in this correspondence. This means that a state is pure iff its values are either 0 or 1; this cannot hold for any element of $\mathfrak{N}$, i.e., $\mathfrak{N} \cap \mathfrak{M}_p = \emptyset$. Further, this $\mathfrak{L}$ does not satisfy our *JPZ* property, because for A, B in distinct $\mathfrak{B}_\lambda$ we have $A \wedge B = 0$, while there exists a large number of pure states m with $mA = mB = 1$.

ACKNOWLEDGEMENT

The author wishes to thank the Graduate Office and the Mathematics Department of I.I.T. for the award of a Faculty Fellowship which greatly facilitated the research involved in the present paper.

[1] N. Zierler, Pac. J. Math. **11**, 1151 (1961).
[2] C. Piron, Helv. Phys. Acta **37**, 439 (1964).
[3] N. D. MacLaren, Pac. J. Math. **14**, 597 (1964).
[4] V. S. Varadarajan, *Geometry of Quantum Theory* (Van Nostrand, Princeton, N. J., 1968).
[5] S. P. Gudder, Pac. J. Math. **19**, 81 (1968).

New forms for the representations of the three-dimensional Lorentz group

N. Mukunda and B. Radhakrishnan
Tata Institute of Fundamental Research, Bombay, India
(Received 17 May 1971; revised manuscript received 8 July 1971)

We present a uniform construction of all the principal series representations of the three-dimensional Lorentz group with generators constructed in terms of oscillator operators. In all cases, the Hilbert space and a hyperbolic generator have a simple appearance, while the other two generators give rise to nonlocal transformations involving Bessel functions.

INTRODUCTION

The purpose of this paper is to present the unitary irreducible representations (UIR's) of the group $SU(1, 1)$ in a new form. This group is closely related to the group of Lorentz transformations $O(2, 1)$ in three-dimensional space-time, being in fact the "spinor group" of the latter; as such, the properties of its representations, and the ways in which these representations may be constructed, are of relevance in various contexts in elementary particle physics.

We study here some properties of the principal series of UIR's of the group $SU(1, 1)$; as is well known, this series is made up of a family of UIR's called the continuous series, and another family called the discrete series.[1] We will obtain a uniform description of the UIR's of both these series; this description is characterized by the fact that the transformations generated by a particular hyperbolic ("noncompact") generator have a specially simple form.[2] Being able to construct the two families of UIR's in very similar forms is interesting for the following reasons. A common technique for the construction of UIR's of a noncompact group is this: One considers (complex-valued) functions on a space on which the group acts transitively as a group of point transformations, and then defines the action of an element of the group on a given function by making the former act on the argument of the latter as a point transformation, that is, one essentially has a change of argument. In addition one includes a "multiplier" in this action, subject to the associative law for group multiplication being valid. In this way, the action of group elements on functions is simple. To make the representation unitary, one must next impose a suitable positive-definite scalar product on the functions considered, and ensure that this product is preserved under the group transformations. If this can be achieved by means of a local expression for the scalar product, then both the group action and the Hilbert space have simple appearances; but in some cases, the scalar product becomes nonlocal, and then only the group action remains simple. In a common method of construction of the UIR's of $SU(1, 1)$, the continuous series of UIR's belongs to the former type, but the discrete series to the latter (we do not speak here of the exceptional series).[3] We have imposed the requirement from the beginning that the Hilbert space have a simple form, with a local scalar product, for both kinds of representations, and that both representations be similarly realized. We then find that the action of group elements is in general nonlocal, though the forms are rather similar in the two types of representations. This nonlocality is the consequence of demanding uniformity in the description. Of course, the similarity in our constructions of the two types of representations must be limited by the existence of basic differences between discrete and continuous representations. For example, the eigenvalues of a hyperbolic generator possess multiplicity one in the former and two in the latter. To be precise, then, the uniformity lies in having as similar forms as possible for the Hilbert spaces and for the action of one chosen hyperbolic generator.

In practical applications of the unitary representations of noncompact groups, many authors have utilized the technique of constructing the generators out of a set of harmonic oscillator creation and annihilation operators.[4] Generally such constructions have led to a subset of all UIR's of the relevant group. [In Ref.4(a), however, all the UIR's of the universal covering group of $SU(1, 1)$ have been constructed, using the method of complex rank tensors.] We use this oscillator operator method for all the principal series UIR's of $SU(1, 1)$. We describe the construction of the discrete series UIR's in Sec. 1, following brief comments on $SU(1, 1)$ and its UIR's, and consider the continuous series in Sec. 2.

1. REPRESENTATIONS OF THE DISCRETE SERIES

We begin with a few remarks concerning the group $SU(1, 1)$. This is the group of all two-dimensional complex unimodular matrices of the form

$$\begin{pmatrix} \alpha & \beta \\ \bar{\beta} & \bar{\alpha} \end{pmatrix}, \qquad \alpha\bar{\alpha} - \beta\bar{\beta} = 1; \tag{1.1}$$

that is, the group of pseudo unitary 2×2 matrices. It has three generators which may be written J_0, J_1, J_2; in a unitary representation, they are Hermitian, and their commutation rules (C R 's) are

$$[J_0, J_1] = iJ_2, \qquad [J_0, J_2] = -iJ_1, \qquad [J_1, J_2] = -iJ_0. \tag{1.2}$$

In the defining nonunitary matrix representation, we may identify the J's in terms of Pauli matrices as $J_0 \to \frac{1}{2}\sigma_3$, $J_1 \to \frac{1}{2}i\sigma_2, J_2 \to \frac{1}{2}i\sigma_1$. J_0 generates a maximal compact subgroup of $SU(1, 1)$; J_1 and J_2 are hyperbolic generators giving rise to noncompact $O(1, 1)$ subgroups. There are two kinds of discrete UIR's of $SU(1, 1)$, the positive type D_k^+ and the negative type D_k^-, $k = \frac{1}{2}, 1, \frac{3}{2}, \ldots$. In the former, J_0 has the eigenvalues $k, k + 1, \ldots, \infty$, in the latter $-k, -k - 1, \ldots, -\infty$; in both D_k^+ and D_k^- the Casimir operator $Q \equiv (J_1)^2 + (J_2)^2 - (J_0)^2$ has the value $k(1 - k)$.

Let us now introduce two sets of oscillator creation and annihilation operators a_j and a_j^+, $j = 1, 2$, obeying the CR's:

$$[a_j, a_j^+] = \delta_{jk}, \qquad [a_j, a_k] = [a_j^+, a_k^+] = 0. \tag{1.3}$$

Then one can obtain a solution to the CR's equation (1.2) by the choice

$$\begin{aligned} J_0 &= \tfrac{1}{2}(a_1^+ a_1 + a_2^+ a_2 + 1), \\ J_1 &= \tfrac{1}{4}[(a_1^+)^2 + (a_1)^2 + (a_2^+)^2 + (a_2)^2], \\ J_2 &= -\tfrac{1}{4}i\,[(a_1^+)^2 - (a_1)^2 + (a_2^+)^2 - (a_2)^2]. \end{aligned} \tag{1.4}$$

These are Hermitian, and since J_0 is positive definite this construction will lead to the D_k^+ UIR's.[5] It is evident that if one introduces the Hermitian operator K by

$$K = \tfrac{1}{2} - \tfrac{1}{2}(a_1^+ a_2 - a_2^+ a_1), \tag{1.5}$$

then all the generators commute with K, and the Casimir operator is a function of K (as a short calculation will show):

$$Q \equiv (J_1)^2 + (J_2)^2 - (J_0)^2 = K(1-K) \tag{1.6}$$

The above construction using oscillator operators allows us to deal in a simple way with the eigenstates of the compact generator J_0 which is just the sum of oscillator Hamiltonians, so that these normalized eigenstates are generated from the common vacuum state $|0\rangle$ in the following familiar way:

$$|m,n\rangle = (m!n!)^{-1/2}(a_1^+)^m (a_2^+)^n |0\rangle,$$
$$J_0|m,n\rangle = \tfrac{1}{2}(m+n+1)|m,n\rangle. \tag{1.7}$$

Diagonalization of the Hermitian operator K simultaneously with J_0 involves taking suitable linear combinations of these oscillator states keeping the total occupation number $(m+n)$ fixed, and this leads to irreducible representations in the familiar form in which the compact generator is diagonal. We are, however, interested in exhibiting the generator J_2 in a simple form and so proceed somewhat differently. The Hilbert space of the entire representation continues however to be spanned by the states $|m,n\rangle$.

The representation of $SU(1,1)$ generated by J_0, J_1, and J_2 of Eq. (1.4) is reducible, and is a direct sum of all the UIR's D_k^+, once each. To effect the reduction, we use configuration space variables. In terms of Cartesian coordinates x_1, x_2 varying from $-\infty$ to $+\infty$, we have

$$a_j = -\frac{i}{\sqrt{2}}\left(x_j + \frac{\partial}{\partial x_j}\right), \quad a_j^+ = \frac{i}{\sqrt{2}}\left(x_j - \frac{\partial}{\partial x_j}\right) \tag{1.8}$$

In the next step, we introduce radial and polar coordinates $x_1 = r\cos\theta$ and $x_2 = r\sin\theta$ in terms of which the Hilbert space consists of functions $f(r,\theta)$ for which

$$\|f(r,\theta)\|^2 = \int_0^{2\pi} d\theta \int_0^\infty r dr\, |f(r,\theta)|^2 < \infty. \tag{1.9}$$

Then the operator K and the generators J_0, J_1, J_2 become

$$K = \tfrac{1}{2}\left(1 - i\frac{\partial}{\partial\theta}\right),$$
$$J_0 = \frac{1}{4}\left(r^2 - \frac{\partial^2}{\partial r^2} - \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}(2K-1)^2\right),$$
$$J_1 = \frac{-1}{4}\left(r^2 + \frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} - \frac{1}{r^2}(2K-1)^2\right), \tag{1.10}$$
$$J_2 = -\frac{i}{2}\left(r\frac{\partial}{\partial r} + 1\right).$$

It is clear that by considering the action of these generators on functions f whose dependence on θ is of the form $e^{i(2k-1)\theta}$, where k can take on one of the values $\frac{1}{2}, 1, \frac{3}{2}, \ldots$, we get the UIR D_k^+. Therefore, restricting ourselves to one such UIR, the structure of the scalar product in the Hilbert space, and the generators, are

$$(f,g) = \int_0^\infty r dr \bar{f}(r) g(r),$$
$$J_0 = \tfrac{1}{4}\left(r^2 - \frac{d^2}{dr^2} - \frac{1}{r}\frac{d}{dr} + \frac{1}{r^2}(2k-1)^2\right),$$
$$J_1 = -\tfrac{1}{4}\left(r^2 + \frac{d^2}{dr^2} + \frac{1}{r}\frac{d}{dr} - \frac{1}{r^2}(2k-1)^2\right), \tag{1.11}$$
$$J_2 = -\frac{i}{2}\left(r\frac{d}{dr} + 1\right),$$

The scalar product is the same for all k, while the generators J_0 and J_1 depend on k.[6]

We now compute the form of the finite transformations generated by each J. Since J_0 and J_1 are second-order differential operators, the corresponding finite transformations will be nonlocal, while J_2 does give rise to a local transformation. We have

$$[e^{i\zeta J_2} f](r) = e^{\zeta/2} f(re^{\zeta/2}). \tag{1.12}$$

For J_0 and J_1 we write

$$[e^{i\mu J_0} f](r) = \int_0^\infty r' dr' L^{(k,+)}(r,r';\mu) f(r'),$$
$$[e^{i\nu J_1} f](r) = \int_0^\infty r' dr' M^{(k,+)}(r,r';\nu) f(r'). \tag{1.13}$$

The kernels $L^{(k,+)}$ and $M^{(k,+)}$ must be solutions of the following boundary value problems:

$$-i\frac{\partial}{\partial\mu} L^{(k,+)}(r,r';\mu)$$
$$= \frac{1}{4}\left(r^2 - \frac{\partial^2}{\partial r^2} - \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}(2k-1)^2\right) L^{(k,+)}(r,r';\mu);$$
$$L^{(k,+)}(r,r';0) = (1/r)\delta(r-r');$$
$$-i\frac{\partial}{\partial\nu} M^{(k,+)}(r,r';\nu)$$
$$= -\frac{1}{4}\left(r^2 + \frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} - \frac{1}{r^2}(2k-1)^2\right) M^{(k,+)}(r,r';\nu);$$
$$M^{(k,+)}(r,r';0) = (1/r)\delta(r-r'); \tag{1.14}$$

The kernel $L^{(k,+)}$ is easily obtained from the knowledge of the Green's function for a simple harmonic oscillator, since the operator J_0 given in Eq. (1.4) is just a sum of two such oscillator Hamiltonians.[7] That is, the solution to the boundary value problem

$$\frac{1}{4}\left(r^2 - \frac{\partial^2}{\partial r^2} - \frac{1}{r}\frac{\partial}{\partial r} - \frac{1}{r^2}\frac{\partial^2}{\partial\theta^2}\right) \mathcal{G}(r,r';\theta,\theta';\mu)$$
$$= -i\frac{\partial}{\partial\mu}\mathcal{G}(r,r';\theta,\theta';\mu),$$
$$\mathcal{G}(r,r';\theta,\theta';0) = (1/r)\delta(r-r')\delta(\theta-\theta') \tag{1.15}$$

is known to be given by

$$\mathcal{G}(r,r';\theta,\theta';\mu) = (i/2\pi\sin\tfrac{1}{2}\mu)$$
$$\times \exp\{-i[(r^2+r'^2)\cos\tfrac{1}{2}\mu - 2rr'\cos(\theta-\theta')]/2\sin\tfrac{1}{2}\mu\}. \tag{1.16}$$

(The dependence of $\mathcal{G}$ only on $\theta-\theta'$ reflects the fact that the operator K commutes with J_0.) It follows that

we obtain the kernel $L^{(k,+)}$ obeying Eq. (1.14) by just making a Fourier decomposition of $\mathcal{G}$ in the variable $\theta - \theta'$, and picking up the term with the behavior $e^{i(2k-1)(\theta-\theta')}$

$$L^{(k,+)}(r,r';\mu) = (i/2\pi \sin\tfrac{1}{2}\mu)\exp[-i(r^2+r'^2)\cot\tfrac{1}{2}\mu]$$
$$\times \int_0^{2\pi} d\theta\, e^{ia\cos\theta}\, e^{-i(2k-1)\theta}$$
$$= (e^{i\pi k}/\sin\tfrac{1}{2}\mu)\exp[-i(r^2+r'^2)\cot\tfrac{1}{2}\mu]J_{2k-1}(a) \quad (1.17)$$

where $a \equiv (rr'/\sin\tfrac{1}{2}\mu)$

[J_{2k-1} is the Bessel function].[8] In checking that the proper boundary condition is obeyed at $\mu = 0$, one needs the identity

$$\lim_{t\to 0}(\pi i t)^{-1/2} e^{iz^2/t} = \delta(z). \quad (1.18)$$

In a similar manner, the kernel $M^{(k,+)}$ may be found by starting from the Green's function for the "antiharmonic" oscillator which in one dimension has the Hamiltonian $-\frac{1}{2}(d/dx^2) - \frac{1}{2}x^2$. The Green's function for two such uncoupled oscillators obeys

$$-\frac{1}{4}\left(r^2 + \frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial\theta^2}\right)\mathcal{G}'(r,r';\theta,\theta';\nu)$$
$$= -i\frac{\partial}{\partial\nu}\mathcal{G}'(r,r';\theta,\theta';\nu),$$
$$\mathcal{G}'(r,r';\theta,\theta';0) = (1/r)\delta(r-r')\delta(\theta-\theta') \quad (1.19)$$

and is explicitly given by

$$\mathcal{G}'(r,r';\theta,\theta';\nu) = [(i/2\pi)\sinh\tfrac{1}{2}\nu)$$
$$\times \exp\{-i[(r^2+r'^2)\cosh\tfrac{1}{2}\nu - 2rr'\cos(\theta-\theta')]/2\sinh\tfrac{1}{2}\nu\} \quad (1.20)$$

Once again, the term going like $e^{i(2k-1)(\theta-\theta')}$ in the Fourier development of $\mathcal{G}'$ yields the required kernel $M^{(k,+)}$ obeying the second pair of Eqs. (1.14):

$$M^{(k,+)}(r,r';\nu) = (i/2\pi\sinh\tfrac{1}{2}\nu)\exp[-\tfrac{1}{2}i(r^2+r'^2)\coth\tfrac{1}{2}\nu]$$
$$\times \int_0^{2\pi} d\theta\, e^{ib\cos\theta}\, e^{-i(2k-1)\theta}$$
$$= (e^{i\pi k}\sinh\tfrac{1}{2}\nu)\exp[-\tfrac{1}{2}i(r^2+r'^2)\coth\tfrac{1}{2}\nu]J_{2k-1}(b) \quad (1.21)$$

where $b = (rr'/\sinh\tfrac{1}{2}\nu)$.

The basic equations needed for this description of the UIR's of type D_k^+ are now obtained. For all values of k, the Hilbert space consists of complex functions on the positive real line with the scalar product given in Eq. (1.11). Thus it has a form independent of k. The one-parameter subgroup generated by J_2 has a simple action given in Eq. (1.12), and is again k-independent. But the transformations generated by J_0 and J_1 are nonlocal, and the k-dependent kernels, involving Bessel functions, are given by Eqs. (1.17) and (1.21) respectively.[9]

The results for the UIR's D_k^- can be obtained quite easily from those given above. If the three operators J_0, J_1, J_2 give rise to the UIR D_k^+ it follows that the operators $-J_0, -J_1, J_2$ give rise to the UIR D_k^-. So the description of the UIR D_k^- assumes this form: The Hilbert space, as well as the action of the transformation $e^{i\zeta J_2}$, are the same as before, being given by the first line of Eq. (1.11) and Eq. (1.12), respectively. The unitary operators $e^{i\mu J_0}$ and $e^{i\nu J_1}$ are described again by equations having the same appearance as Eq. (1.13), the only difference being that we now have new kernels $L^{(k,-)}(r,r';\mu)$ and $M^{(k,-)}(r,r';\nu)$ on the right. These new kernels are given in terms of the earlier ones by

$$L^{(k,-)}(r,r';\mu) \equiv L^{(k,+)}(r,r';-\mu),$$
$$M^{(k,-)}(r,r';\nu) \equiv M^{(k,+)}(r,r';-\nu) \quad (1.22)$$

We conclude this section with the following comment. It is interesting to see that our construction of the discrete UIR's of $SU(1,1)$ has yielded the two UIR's $D_{1/2}^{\pm}$ along with all the others for $k \geq 1$, even though the latter alone appear in the Plancherel formula for expanding a square integrable function on $SU(1,1)$ (along with the continuous nonexceptional UIR's).[10] In contrast, in the next section we will see that similar constructions using oscillator operators lead to the continuous nonexceptional UIR's alone, and not the exceptional series UIR's.

2. REPRESENTATIONS OF THE CONTINUOUS SERIES

To obtain the remaining representations belonging to the principal series, we consider in place of Eqs. (1.4) the following set of generators:

$$J_0 = \tfrac{1}{2}(a_1^+ a_1 - a_2^+ a_2),$$
$$J_1 = \tfrac{1}{4}[(a_1^+)^2 + (a_1)^2 - (a_2^+)^2 - (a_2)^2], \quad (2.1)$$
$$J_2 = -\tfrac{1}{4}i[(a_1^+)^2 - (a_1)^2 + (a_2^+)^2 - (a_2)^2].$$

These are also Hermitian and obey the CR's of $SU(1.1)$. J_2 is the same as before, while J_0 and J_1 are different; in particular, J_0 is no longer positive-definite. The Hermitian operator S defined as

$$S = -\tfrac{1}{2}i(a_1^+ a_2^+ - a_1 a_2) \quad (2.2)$$

can be seen to commute with all the J's and, in fact, the Casimir operator Q is a function of S:

$$Q \equiv (J_1)^2 + (J_2)^2 - (J_0)^2 = \tfrac{1}{4} + S^2. \quad (2.3)$$

This shows that we are dealing here with the continuous (nonexceptional) series of UIR's of $SU(1,1)$. To reduce this representation of $SU(1,1)$, we must note that in addition to commuting with S, the generators defined in Eq. (2.1) are invariant under the unitary transformation that takes a_j, a_j^+ into $-a_j, -a_j^+$, respectively. Using the representation equation (1.8) for the oscillator operators, this transformation is just reflection in the plane, $x_j \to -x_j$. [Invariance under this reflection, which certainly obtains for the generators of the discrete series Eq. (1.4), was not explicitly mentioned there because this operation is the same as rotation in the plane by an amount π, and those generators already commute with the operator that generates rotations, namely, $(2K-1)$]. Using Eq. (1.8), the Hilbert space appears as all functions $f(x_1, x_2)$ for which

$$\|f\|^2 \equiv \int_{-\infty}^{\infty} dx_1 \int_{-\infty}^{\infty} dx_2\, |f(x_1,x_2)|^2 < \infty \quad (2.4)$$

and the operator J_0, for instance, takes the form

$$J_0 = \frac{1}{4}\left(x_1^2 - \frac{\partial^2}{\partial x_1^2} - x_2^2 + \frac{\partial^2}{\partial x_2^2}\right). \quad (2.5)$$

The total Hilbert space $\mathcal{H}$ splits into two orthogonal subspaces $\mathcal{H}_+$ and $\mathcal{H}_-$ consisting, respectively, of functions $f(x_1, x_2)$ that are even and odd under the operation $x_1 \to -x_1$, $x_2 \to -x_2$; both $\mathcal{H}_+$ and $\mathcal{H}_-$ are invariant under the transformations of $SU(1,1)$. On the other hand, since J_0 is one half the difference of two harmonic oscillator Hamiltonians, a general eigenvalue of J_0 is of the form $\frac{1}{2}(n_1 - n_2)$, with each n_1, n_2 taking one of the values $0, 1, 2, \ldots$. The corresponding eigenfunctions, which form a basis for $\mathcal{H}$, are just products of eigenfunctions for the individual oscillators. But since these individual eigenfunctions possess definite parity properties, it is clear that the product eigenfunction corresponding to the pair (n_1, n_2) goes into $(-1)^{n_1+n_2}$ times itself under $x_j \to -x_j$ while, acting on it, $e^{2\pi i J_0}$ has the eigenvalue $(-1)^{n_1-n_2} = (-1)^{n_1+n_2}$. Putting these facts together, it follows that in $\mathcal{H}_+$, J_0 has only integer eigenvalues; and in $\mathcal{H}_-$, J_0 has half odd integer eigenvalues. As a consequence, restriction of the original representation of $SU(1,1)$ to $\mathcal{H}_+(\mathcal{H}_-)$ and its subsequent reduction will lead to the continuous integral (half-integral) series of UIR's of this group.

To perform this reduction, it is convenient to divide the $x_1 - x_2$ plane into four equal regions, and introduce hyperbolic variables in each:

$$\begin{aligned}
&1: x_2 \geq 0, \quad -x_2 \leq x_1 \leq x_2, \quad && x_2 = r\cosh\eta, \quad x_1 = r\sinh\eta;\\
&2: x_1 \geq 0, \quad -x_1 \leq x_2 \leq x_1, \quad && x_1 = r\cosh\eta, \quad x_2 = r\sinh\eta;\\
&1': x_2 \leq 0, \quad x_2 \leq x_1 \leq -x_2, \quad && x_2 = -r\cosh\eta, \quad x_1 = -r\sinh\eta;\\
&2': x_1 \leq 0, \quad x_1 \leq x_2 \leq -x_1, \quad && x_1 = -r\cosh\eta, \quad x_2 = -r\sinh\eta.
\end{aligned} \tag{2.6}$$

In each region, $0 \leq r < \infty$, $-\infty < \eta < \infty$. A function $f(x_1, x_2)$ is now replaced by a set of four functions $f_1(r,\eta), f_2(r,\eta), f_{1'}(r,\eta), f_{2'}(r,\eta)$. If it belongs to $\mathcal{H}_\epsilon$, $\epsilon = \pm 1$, it obeys $f_{1'}(r,\eta) = \epsilon f_1(r,\eta), f_{2'}(r,\eta) = \epsilon f_2(r,\eta)$. From now on, we need only deal with one subspace $\mathcal{H}_\epsilon$ at a time. The integration measure $dx_1 dx_2$ becomes, of course, $r dr d\eta$. Uniformly in all regions, the operator S takes the form

$$S = -\frac{i}{2}\frac{\partial}{\partial\eta}. \tag{2.7}$$

Therefore, by considering functions belonging to $\mathcal{H}_\epsilon$ and whose dependence on η is $e^{2is\eta}$, $-\infty < s < \infty$, which therefore amount to *pairs* of functions $(f_1(r), f_2(r))$, and evaluating the effect on them of the finite transformations generated by the operators of Eq. (2.1), we will obtain the continuous series UIR for which $Q = \frac{1}{4} + s^2$; it will belong to the integral class if $\epsilon = +1$, to the half-integral class otherwise. The action of J_2 is, as before, quite simple. For every continuous series UIR, of whatever kind, the Hilbert space, $\mathcal{H}^{(c)}$ say, consists of pairs $f \to (f_1(r), f_2(r))$ with

$$\|f\|^2 = \int_0^\infty r dr[\,|f_1(r)|^2 + |f_2(r)|^2] < \infty \tag{2.8}$$

and

$$[e^{i\zeta J_2} f]_i(r) = e^{\zeta/2} f_i(r e^{\zeta/2}), \quad i = 1, 2. \tag{2.9}$$

For the other two relevant one-parameter subgroups, we write, for given ϵ and $Q = \frac{1}{4} + s^2$,

$$\begin{aligned}
[e^{i\mu J_0} f]_i(r) &= \sum_{j=1}^{2} \int_0^\infty r' dr' L_{ij}^{(s,\epsilon)}(r, r'; \mu) f_j(r'),\\
[e^{i\nu J_1} f]_i(r) &= \sum_{j=1}^{2} \int_0^\infty r' dr' M_{ij}^{(s,\epsilon)}(r, r'; \nu) f_j(r').
\end{aligned} \tag{2.10}$$

We prefer to obtain these kernels first, and then mention the corresponding differential equations and boundary values, analogous to Eq. (1.14). In the case of J_0, for example, we start with the form given in Eq. (2.5). We know that the solution to the boundary value problem

$$\begin{aligned}
&\frac{1}{4}\left(x_1^2 + \frac{\partial^2}{\partial x_1^2} - x_2^2 + \frac{\partial^2}{\partial x_2^2}\right) \mathcal{G}(x_1, x_1'; x_2, x_2'; \mu)\\
&\quad = -i\frac{\partial}{\partial\mu}\mathcal{G}(x_1, x_1'; x_2, x_2'; \mu)\\
&\quad \mathcal{G}(x_1, x_1'; x_2, x_2'; 0) = \delta(x_1 - x_1')\delta(x_2 - x_2')
\end{aligned} \tag{2.11}$$

is given by

$$\begin{aligned}
\mathcal{G}(x_1, x_1'; x_2, x_2'; \mu) &= (-1/2\pi\sin\tfrac{1}{2}\mu)\\
&\times \exp\{-i[(x_1^2 + x_1'^2 - x_2^2 - x_2'^2)\cos\tfrac{1}{2}\mu\\
&- 2(x_1x_1' - x_2x_2')]/2\sin\tfrac{1}{2}\mu\},
\end{aligned} \tag{2.12}$$

so the effect of $e^{i\mu J_0}$ [J_0 as in Eq. (2.5)], on a function $f(x_1, x_2)$ in $\mathcal{H}_\epsilon$ can be given explicitly. Using Eq. (2.6) this reads

$$\begin{aligned}
&[e^{i\mu J_0} f]_1(r,\eta) = -(2\pi\sin\tfrac{1}{2}\mu)^{-1}\int_0^\infty r' dr' \Big(\exp[\tfrac{1}{2}i(r^2 + r'^2)\\
&\times \cot\tfrac{1}{2}\mu]\int_{-\infty}^{\infty} d\eta'(e^{-ia\cosh(\eta-\eta')} + \epsilon e^{ia\cosh(\eta-\eta')}) f_1(r', \eta')\\
&+ \exp[\tfrac{1}{2}i(r^2 - r'^2)\cot\tfrac{1}{2}\mu]\\
&\times \int_{-\infty}^{\infty} d\eta'[e^{ia\sinh(\eta-\eta')} + \epsilon e^{-ia\sinh(\eta-\eta')}] f_2(r', \eta')\Big)\\
&[e^{i\mu J_0} f]_2(r,\eta) = -(2\pi\sin\tfrac{1}{2}\mu)^{-1}\int_0^\infty r' dr' \Big(\exp[-\tfrac{1}{2}i(r^2 - r'^2)\\
&\times \cot\tfrac{1}{2}\mu]\int_{-\infty}^{\infty} d\eta'(e^{ia\sinh(\eta-\eta')} + \epsilon e^{-ia\sinh(\eta-\eta')}) f_1(r', \eta')\\
&+ \exp\{-\tfrac{1}{2}i(r^2 + r'^2)\cot\tfrac{1}{2}\mu)\\
&\times \int_{-\infty}^{\infty} d\eta'[e^{ia\cosh(\eta-\eta')} + \epsilon e^{-ia\cosh(\eta-\eta')}] f_2(r', \eta')\Big),
\end{aligned}$$

where again $a \equiv (rr'/\sin\tfrac{1}{2}\mu)$. (2.13)

If we specialize to functions $f_i(r,\eta)$ whose η dependence is $e^{2is\eta}$ and carry out the η integrations, we get directly the kernels $L_{ij}^{(s,\epsilon)}(r, r'; \mu)$. For example, from the first of Eqs. (2.13), we have

$$\begin{aligned}
L_{11}^{(s,\epsilon)}(r, r'; \mu) &= -(2\pi\sin\tfrac{1}{2}\mu)^{-1}\exp[\tfrac{1}{2}i(r^2 + r'^2)\cot\tfrac{1}{2}\mu]\\
&\times \int_{-\infty}^{\infty} d\eta\, e^{2is\eta}(e^{-ia\cosh\eta} + \epsilon e^{ia\cosh\eta}).
\end{aligned} \tag{2.14}$$

This and similar integrals can be expressed in terms of Hankel and Macdonald functions.[11] Thus, we find

$$\begin{aligned}
L_{11}^{(s,\epsilon)}(r, r'; \mu) &= (i/2\sin\tfrac{1}{2}\mu)\exp[\tfrac{1}{2}i(r^2 + r'^2)\cot\tfrac{1}{2}\mu]\\
&\times [e^{\pi s} H_{2is}^{(2)}(a) - \epsilon e^{-\pi s} H_{2is}^{(1)}(a)],\\
L_{12}^{(s,\epsilon)}(r, r'; \mu) &= -(\pi\sin\tfrac{1}{2}\mu)^{-1}\exp[\tfrac{1}{2}i(r^2 - r'^2)\cot\tfrac{1}{2}\mu]\\
&\times (e^{\pi s} + \epsilon e^{-\pi s}) K_{2is}(a),
\end{aligned}$$

$$L_{21}^{(s,\epsilon)}(r,r';\mu) = -(\pi \sin\tfrac{1}{2}\mu)^{-1}\exp[-\tfrac{1}{2}i(r^2-r'^2)\cot\tfrac{1}{2}\mu] \times (e^{-\pi s}+\epsilon e^{\pi s})K_{2is}(a),$$

$$L_{22}^{(s,\epsilon)}(r,r';\mu) = (-i/2\sin\tfrac{1}{2}\mu)\exp[-\tfrac{1}{2}i(r^2-r'^2) \times \cot\tfrac{1}{2}\mu][e^{-\pi s}H_{2is}^{(1)}(a) - \epsilon e^{\pi s}H_{2is}^{(2)}(a)]. \quad (2.15)$$

Similar calculations can be made to obtain the functions $M_{ij}^{(s,\epsilon)}$. We need now to work with the Green's functions for antiharmonic oscillators. Omitting details, the results are

$$M_{11}^{(s,\epsilon)}(r,r';\nu) = (i/2\sinh\tfrac{1}{2}\nu)\exp[\tfrac{1}{2}i(r^2+r'^2)\coth\tfrac{1}{2}\nu] \times [e^{\pi s}H_{2is}^{(2)}(b) - \epsilon e^{-\pi s}H_{2is}^{(1)}(b)],$$

$$M_{12}^{(s,\epsilon)}(r,r';\nu) = -(\pi\sinh\tfrac{1}{2}\nu)^{-1}\exp[\tfrac{1}{2}i(r^2-r'^2)\coth\tfrac{1}{2}\nu] \times (e^{\pi s}+\epsilon e^{-\pi s})K_{2is}(b), \quad (2.16)$$

$$M_{21}^{(s,\epsilon)}(r,r';\nu) = -(\pi\sinh\tfrac{1}{2}\nu)^{-1}\exp[-\tfrac{1}{2}i(r^2-r'^2)\coth\tfrac{1}{2}\nu] \times (e^{-\pi s}+\epsilon e^{\pi s})K_{2is}(b),$$

$$M_{22}^{(s,\epsilon)}(r,r';\nu) = (-i/2\sinh\tfrac{1}{2}\nu)\exp[-\tfrac{1}{2}i(r^2+r'^2)\coth\tfrac{1}{2}\nu] \times [e^{-\pi s}H_{2is}^{(1)}(b) - \epsilon e^{\pi s}H_{2is}^{(2)}(b)],$$

where $b \equiv (rr'/\sinh\tfrac{1}{2}\nu)$

With these kernels, our construction of the continuous series UIR's of $SU(1,1)$ is complete. For every one of these UIR's the form of the Hilbert space is given by Eq. (2.8) and the action of $e^{i\zeta J_2}$, which is a point transformation, by Eq. (2.9). The other two one-parameter subgroups act as nonlocal transformations according to Eq. (2.10), ϵ being + for the integral and − for the half-integral case.[12]

As stated in the Introduction, the Hilbert spaces in which we have constructed both the discrete series and the continuous series UIR's are very similar; all that has happened in going from the former to the latter is a kind of doubling of the space. This much difference has to be allowed for, considering the well-known properties of the hyperbolic generator J_2: In any UIR of the discrete series, every real number is a possible (generalized) eigenvalue for J_2, and there is just one (generalized) eigenvector per eigenvalue; on the other hand, in any UIR of the continuous series, the eigenvalues are the same but we have two eigenvectors per eigenvalue.[13]

Now we may explain why in this section we preferred to obtain first the kernels L and M, and delayed writing down their defining differential equations. The point is that these kernels are different in the integral and the half-integral cases for one and the same value of Q, as they must be; and the ϵ dependent terms in Eqs. (2.15) and (2.16) show the differences. But if we write out the boundary conditions and differential equations for these kernels in a purely formal way, the parameter ϵ makes no appearance anywhere. Using the transformations of Eq. (2.6) appropriately, and working purely formally, we easily get the following:

$$-i\frac{\partial}{\partial\mu}L_{ij}^{(s,\epsilon)}(r,r';\mu) = \frac{1}{4}\left(-r^2+\frac{\partial^2}{\partial r^2}+\frac{1}{r}\frac{\partial}{\partial r}+\frac{4s^2}{r^2}\right)(\sigma_3)_{ik}L_{kj}^{(s,\epsilon)}(r,r';\mu),$$
$$L_{ij}^{(s,\epsilon)}(r,r';0) = (1/r)\delta(r-r')\delta_{ij},$$

$$-i\frac{\partial}{\partial\nu}M_{ij}^{(s,\epsilon)}(r,r';\nu) = -\frac{1}{4}\left(r^2+\frac{\partial^2}{\partial r^2}+\frac{1}{r}\frac{\partial}{\partial r}+\frac{4s^2}{r^2}\right)(\sigma_3)_{ik}M_{ij}^{(s,\epsilon)}(r,r';\nu),$$
$$M_{ij}^{(s,\epsilon)}(r,r';0) = (1/r)\delta(r-r')\delta_{ij}. \quad (2.17)$$

As mentioned, the parameter ϵ makes no specific appearance here, which may seem somewhat strange. The difference between the integral and the half-integral cases lies in the delicate distinction to be drawn between the domains of these generators in the two cases, namely the kinds of function pairs $(f_1(r), f_2(r))$ to which these formal differential operators may be applied. In order to avoid having to solve this problem of domains, we preferred to compute directly the effect of the finite group transformations, for each class of continuous series UIR's.

---

[1]The standard reference for the UIR's of SU(1,1) is V. Bargmann, Ann. Math. **48**, 568 (1947). A detailed account of the construction of these UIR's may also be found in I. M. Gel'fand, M. I. Graev, and N. Ya. Vilenkin, *Generalized Functions* (Academic, New York, 1966), Vol. 5, Chap. VII.

[2]The following papers describe the SU(1,1) representations in a noncompact basis: (a) N. Mukunda, J. Math. Phys. **8**, 2210 (1968); (b) J. G. Kuriyan, N. Mukunda, and E. C. G. Sudarshan, J. Math. Phys. **9**, 2100 (1968); (c) A. O. Barut and E. C. Phillips, Commun. Math. Phys. **8**, 52 (1968); (d) N. Mukunda, J. Math. Phys. **10**, 2086, 2092 (1969); (e) G. Linblad and B. Nagel, Royal Institute of Technology Preprint, Sweden (1969).

[3]I. M. Gel'fand *et al.*, cf. Ref. 1.

[4]See, for example (a) A. O. Barut and C. Fronsdal, Proc. R. Soc. A **287**, 532 (1965); (b) Y. Nambu, Prog. Theor. Phys. **37**, 368 (1966); Phys. Rev. **160**, 1171 (1967). (c) W. J. Holman and L. C. Biedenharm, Ann. Phys. (N.Y.) **39**, 1 (1966); (d) A. O. Barut and E. C. Phillips, Ref. 2(c).

[5]This construction of operators is simply related to the one used in Ref. 4(c).

[6]This representation of the generators is the same, up to a similarity transformation, as the one obtained in Ref. 2(b).

[7]R. P. Feyman and A. R. Hibbs, *Quantum mechanics and path integrals* (McGraw-Hill, New York, 1965), p. 63.

[8]*Higher transcendental functions*, edited by A. Erdelyi (McGraw-Hill, New York, 1953), Vol. II.

[9]Using the asymptotic expressions for the Bessel functions (see Ref. 8), one may verify that the boundary conditions of Eq. (1.14) for the Kernels are fulfilled.

[10]Cf. V. Bargmann, Ref. 1.

[11]*Integral transforms*, edited by A. Erdelyi (McGraw-Hill, New York, 1954), Vol. II, p. 82.

[12]Using the asymptotic expressions for the Hankel and MacDonald functions (Ref. 8), one may verify that the boundary conditions for $L_{ij}^{(s,\epsilon)}$ and $M_{ij}^{(s,\epsilon)}$ stated in Eq. (2.17) are satisfied.

[13]Cf. V. Bargmann, Ref. 1 (Appendix) and N. Mukunda, Ref. 2(a).

Description of a system of interacting linear chains by means of the electron network model

M. Weger and S. Alexander
Department of Physics, The Hebrew University, Jerusalem, Israel

G. Della Riccia
Department of Mathematics and School of Applied Sciences and Technology, The Hebrew University, Jerusalem, Israel
(Received 14 January 1971)

The electron network model of Coulson and Della Riccia is applied to describe a simple lattice consisting of two families of linear chains, coupled together. It is shown that the model displays resonances, at which the wavefunction is preferentially situated at one or the other family of chains, and that at these resonances, there is a $1/\sqrt{E}$ divergence in the electronic density of states.

INTRODUCTION

There has been considerable interest in the appearance of anomalously high peaks in the electronic density of states of some metals which are much narrower than the respective matrix elements would seem to justify. In particular, the very sharp peak observed near the Fermi surface in the β-W structure and its relationship to the chainlike one-dimensional features of this structure have been discussed by several authors[1] in an attempt to relate the observed peaks to the well-known $1/\sqrt{E}$ singularities at the edges of one-dimensional band, making use of the tight binding approximation for the d band of these alloys. Due to the presence of a conduction band, mixing thoroughly with the d band, the description of the d band by means of the tight binding approximation is questionable, and it might perhaps be here described by resonances in the conduction band as suggested by Friedel[2] and Heine.[3]

Following these considerations, it seemed of interest to get some more insight into the way such singularities related to lower dimensionality can appear in a system.

We have used the directed bond electron network model[4] (ENM) in the form suggested by Coulson[5] and Della Riccia[6] to calculate band structures and the density of states for a simple two-dimensional lattice related to the β-W structure. It was found that the bands exhibit one-dimensional $(1/\sqrt{E})$ singularities in the density of states. We then show that such singularities are a fairly general feature of the ENM and result from the appearance of polygons and polyhedra as constant energy surfaces in two and three dimensions, respectively. We show that such energy surfaces always give divergences in the density of states, and discuss their nature.

I. BAND STRUCTURE OF ENM MODEL

The two-dimensional model lattice is described in Fig. 1. It consists of two families of linear chains, one in the x direction and one in the y direction. The junctions on each chain are equally spaced (with spacing a) and at each junction there is a cross link to one junction on the perpendicular family (of length d and relative strength ν). In the ENM model,[4] the wavefunction between junctions is a solution of the one-dimensional Schrödinger equation with a constant potential so that

$$\psi_{\alpha\beta}=\epsilon_\alpha\frac{\sin q(l-x_{\alpha\beta})}{\sin ql}+\epsilon_\beta\frac{\sin qx_{\alpha\beta}}{\sin ql} \tag{1}$$

is the wavefunction on the bond connecting junctions α and β.[7] In Eq. (1), ϵ_α and ϵ_β are the amplitudes of the wavefunction at junctions α and β, respectively, l is the length of the bond, $x_{\alpha\beta}$ is the distance along the bond, from junction α, and $\epsilon=q^2$ is the (one-dimensional) kinetic energy along the bond. Clearly q can be either real or imaginary.

Equation (1) is supplemented by the continuity requirements at the junction[5]

$$\epsilon_\alpha=\text{const}, \tag{2a}$$

$$\sum_\beta \nu_{\alpha\beta}\frac{\partial\psi_{\alpha\beta}}{\partial x_{\alpha\beta}}=\lambda_\alpha\epsilon_\alpha, \tag{2b}$$

where λ_α is a parameter representing a δ function potential and the summation is over all bonds connected to junction α.

This is in fact a slight generalization of the usual model. The parameters $\nu_{\alpha\beta}$ allow us to give different strengths to different inequivalent bonds. It is fairly straightforward to see that such parameters arise naturally in the limiting process of Ref. 4 if the cross sections of the respective channels have a constant ratio in the limiting process.

In addition because of the periodicity, we have

$$\epsilon_\alpha=e^{i\mathbf{k}\mathbf{R}_{\alpha\alpha'}}\epsilon_{\alpha'}, \tag{3}$$

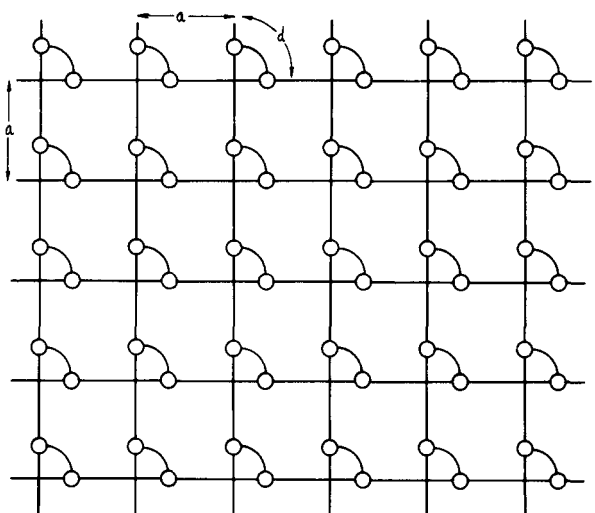

FIG. 1. A two-dimensional network of two families of linear chains, coupled together.

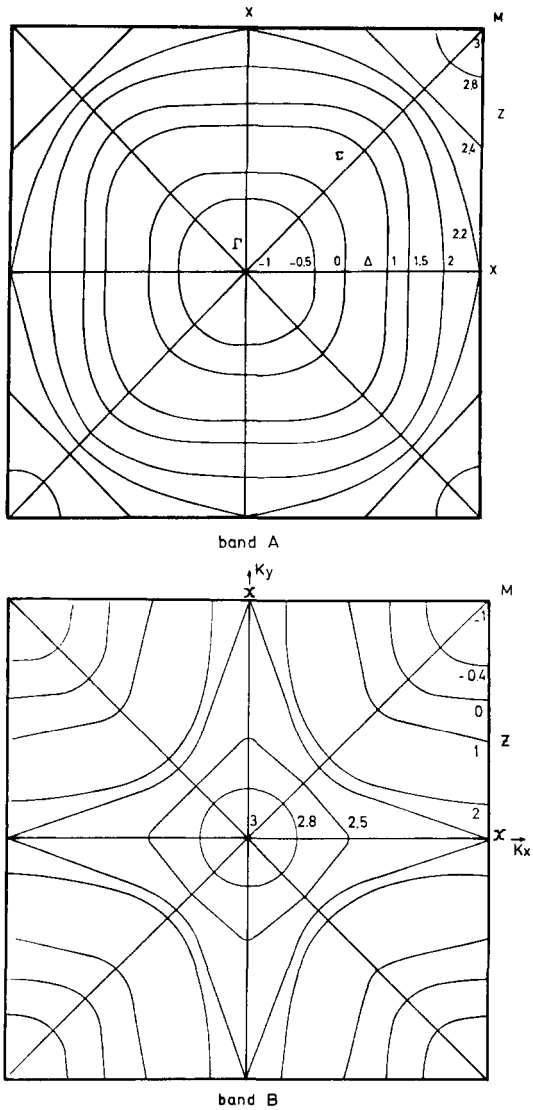

FIG. 2. Typical constant energy surfaces for the network of Fig. 1. The surfaces are frequently in the form of "squares" or "crosses" around Γ.

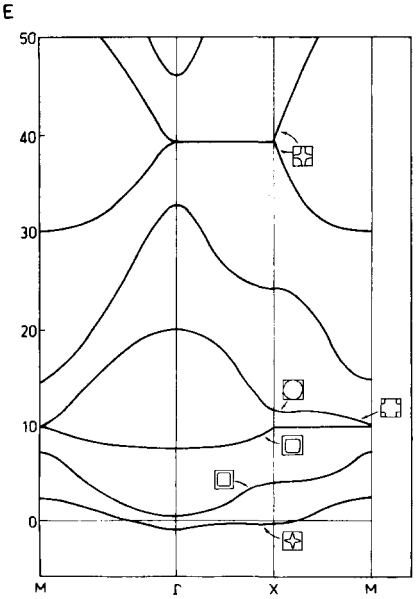

FIG. 3. The energy as function of the momentum along the symmetry lines $\Gamma-X$, $X-M$, $\Gamma-M$.

where $\mathbf{R}_{\alpha\alpha}$ is a lattice translation, α' a junction related to α by this translation, and $\mathbf{k}$ the crystal momentum.

For the model lattice described in Fig. 1 this leads to the equations

$$[-2q\cot qa+2q(\cos k_x a/\sin qa)-\nu q\cot qd]\epsilon_x + (\nu q/\sin qd)\epsilon_y=\lambda\epsilon_x, \quad (4a)$$

$$[-2q\cot qa+2q(\cos k_y a/\sin qa)-\nu q\cot qd]\epsilon_y + (\nu q/\sin qd)\epsilon_x=\lambda\epsilon_y \quad (4b)$$

for the amplitudes in the cell $R=0$. This leads to a transcendental secular equation which is most conveniently written in the form

$$[A(q)-\cos k_x a][A(q)-\cos k_y a]-B(q)^2=0, \quad (5)$$

where

$$A(q)=\cos qa+\frac{\nu}{2}\frac{\cos qd\sin qa}{\sin qd}+\frac{\lambda\sin qa}{2q}, \qquad B(q)=-\frac{\nu\sin qa}{\sin qd}. \quad (6)$$

In general the transcendental equation (5) will have negative energy solutions as well as an infinite number of solutions with positive energy (real q). The most convenient way of treating Eq. 5 was to set the energy (i.e., q) and regard the secular equation as an equation for the constant energy surface in k space. Typical constant-energy surfaces are shown in Fig. 2. It is seen that these are generally in the form of "crosses" or "squares" around point Γ.

In Fig. 3 we have plotted the energy as a function of $\mathbf{k}$ along the main symmetry directions in the Brillouin zone. The energy is given along the line from the origin (Γ) to the center of the zone boundary (X). From there to the corner (M) and along $\Gamma-M$. The bands were calculated for the special case

$$-\lambda=\nu=1, \quad d=1.22a. \quad (7)$$

It is of interest to note the formal similarity of the ENM secular equations, such as Eq. (5), with analogous tight binding equations with energy dependent overlap integrals. This comes out most clearly if we compare Eq. (5) with the tight binding secular equation

$$(\epsilon_0-2J_1\cos k_x a)(\epsilon_0-2J_1\cos k_y a)-J_2^2=0, \quad (8)$$

where

$$\epsilon_0=A\Big/\frac{dA}{d\epsilon}, \quad J_1=\left(2\frac{dA}{d\epsilon}\right)^{-1}, \quad J_2=B\Big/\left(2\frac{dA}{d\epsilon}\right). \quad (9)$$

This definition assures that the effective masses are equal. The energy dependent tight binding parameters ϵ_0, J_1, and J_2 are plotted as a function of energy in Fig. 4.

II. ONE DIMENSIONALITY

For values of q such that

$$B(q)\ll 1, \quad (10)$$

the coupling between the two families of chains becomes

weak. As a result the bands display strong one dimensionality. A convenient measure of this is the parameter

$$\eta = [(\epsilon_x/\epsilon_y) - (\epsilon_y/\epsilon_x)]^2. \tag{11}$$

It can be seen from Eq. 4 that

$$\eta = [(\cos k_x a - \cos k_y a)/B(q)]^2. \tag{12}$$

The numerator of (12) is of order unity for most of the Brillouin zone so that (10) leads to large η.

In Fig. 5 the one dimensionality (η) is plotted as a function of energy for k along the lines $\Gamma - X - M$. It can be seen that η has strong resonances and minima.

III. DIVERGENCES IN THE DENSITY OF STATES

Of particular interest are the divergences in η. They occur when $B(q)$ vanishes, i.e., for values of q such that

$$\sin qa = 0, \quad qa = n\pi. \tag{13}$$

It is then possible to find nontrivial solutions of the basic equations of the ENM which vanish identically at all vertices. For example, there are solutions for which the wavefunction vanishes on the y chains and d bonds but is nonzero on the x chains. It can be seen that such solutions are a fairly general feature of the ENM model and will occur whenever there are paths traversing the lattice which use only one type of bond. These solutions are obviously "one dimensional." In many cases they also lead to divergences in the density of states. The density of states $N(E)$ is plotted as function of the energy in Fig. 6. At values of q for which $\sin qa = 0$, $q \neq 0$, the density of states diverges like $1/\sqrt{E}$.

It is of interest to investigate this a little more closely.

When solutions of Eq. (13) have a definite k along the chain, the energy becomes independent of the components of k perpendicular to the direction of the path. In two dimensions this leads to a constant energy surface which is a straight line segment and in three dimensions a plane. The complete surface is then a polygon or polyhedron. In our model lattice these are the boundaries of the Brillouin zone when $qa = \pi$ and the lines $(k_x, 0)(0, k_y)$ when $qa = 2\pi$.

On such a plane surface we have

$$\left(\frac{\partial^n E(\mathbf{k})}{(\partial k_{\parallel})^n}\right) \equiv 0, \tag{14}$$

where $k_{\parallel}$ is a component of k in the plane.

In many special cases the constant energy surfaces obtained in this way have special symmetry properties so that

$$\left(\frac{\partial E}{\partial k_{\perp}}\right) = 0 \tag{15}$$

from symmetry. The leading term in the expansion of the energy is then proportional to $k_{\perp}^2$ (where $k_{\perp}$ is the perpendicular component of momentum measured from the surface). This leads to a one-dimensional relation and therefore to a $(\epsilon - \epsilon_s)^{-1/2}$ divergence in the density of states at ϵ_s. Note that this holds for both two and three-dimensional lattices.

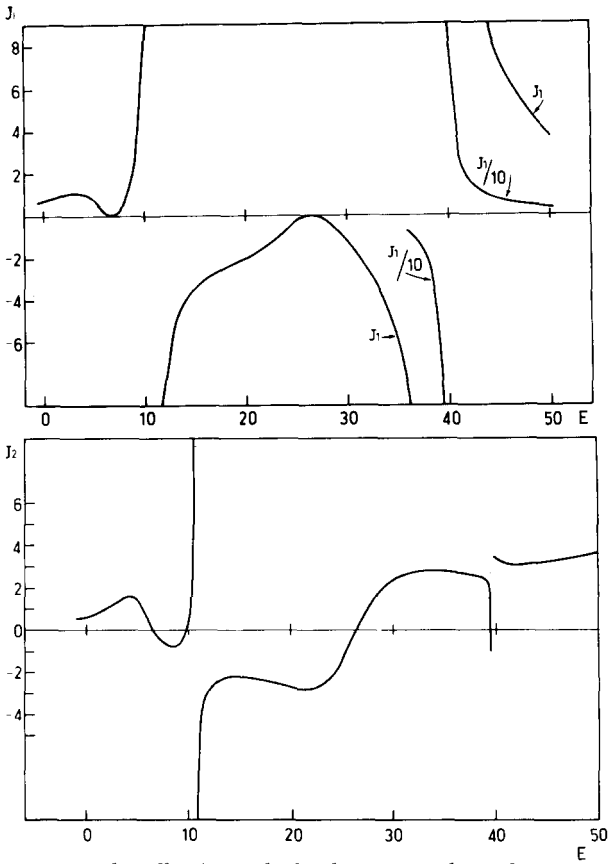

FIG. 4. The effective tight-binding integrals as function of energy. There are violent oscillations near $\sin qa = 0$, $qa \neq 0$. (Here, $E = \pi^2, 4\pi^2$.)

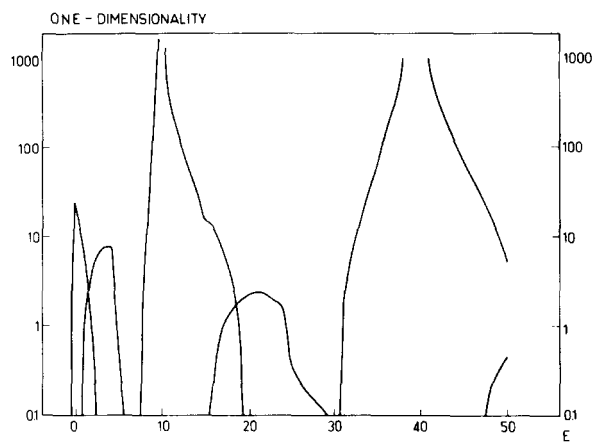

FIG. 5. The "one dimensionality", as defined in Eq. (11) as function of energy, for k along the symmetry lines $\Gamma - X$, $X - M$. When two bands overlap, the "one dimensionality" of the individual bands is shown. There are divergences when $\sin qa = 0$, $qa \neq 0$.

In the general case (15) does not hold and there is no singularity of this type. There is, however, still a divergence in the density of states which results from the corners of the polygon in two dimensions and from the edges of the polyheder in three dimensions.

In two dimensions it is obvious that each corner of the constant energy polygon is an extremal point of the

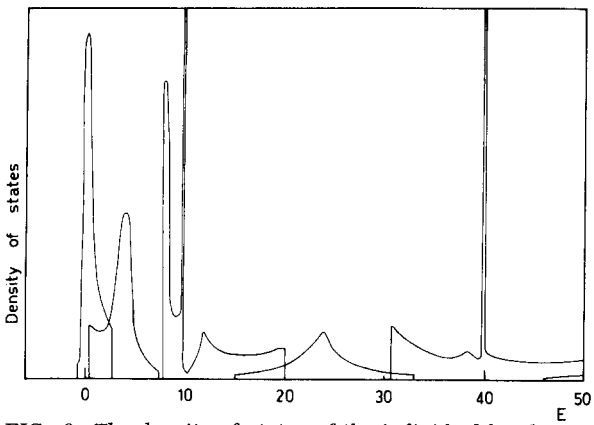

FIG. 6. The density of states of the individual bands, as function of energy. There are $1/\sqrt{E}$ divergences when $\sin qa = 0$, $qa \neq 0$. (Here, at $E = \pi^2$ and $E = 4\pi^2$.)The other peaks in the density of states correspond to "regular" Van Hove singularities, which should be logarithmic infinities for a two-dimensional system. The numerical accuracy of the present calculation is not sufficient to display those "soft" divergences.

energy because the derivatives vanish along two different directions. Moreover the extremum cannot be a simple minimum or maximum because the constant energy surfaces cannot cross the polygon edges.

The simplest extremum is therefore a saddle point which gives a logarithmic singularity in the density of states. When the extremum is not a saddle point it has to be of higher order and gives a $1/\sqrt{E}$ or even stronger singularity in the density of states. This will happen, for example, when a saddle point is ruled out by symmetry or when more than two lines cross at one point (e.g., the origin).

In three dimensions the situation is similar. Each edge of a constant energy polyhedron can be considered as a string of degenerate two-dimensional extremal points in the plane perpendicular to the edge. It is easy to see that this leads to a divergence in the three-dimensional density of states which is of the same type as the one resulting in two dimensions from the respective extremal point. A well-known case is the logarithmic singularity in the density of states of a bcc lattice in the tight binding approximation considering only nearest neighbors.[8]

IV. CONCLUSION

It is seen that the model lattice in the ENM displays resonances (peaks) and antiresonances (zeros) in the "one dimensionality"; and the peaks in the one dimensionality are associated with $1/\sqrt{E}$ divergences in the density of states. The question arises, as to whether these properties of the mathematical model, have any physical reality. The essential feature of the ENM that gives rise to these properties, seems to be the absence of direct interaction between the "chemical bonds," the interaction being only indirect, via the vertices ("atoms"). In reality electrons cannot be assumed to be localized in infinitely narrow channels, and wavefunctions belonging to different bonds should overlap. This should smear out the resonances. Thus, the ENM has a chance to work only if bonds are reasonably far apart or otherwise isolated. In the $\beta - W$ lattice, the distance between equivalent parallel bonds is rather large (twice the nearest neighbor distance): Therefore, the ENM may have some chance to be applicable there.

ACKNOWLEDGMENTS

One of us (M.W.) had many stimulating discussions with Professor J. Friedel, and benefited from valuable suggestions by Professor V. Heine regarding the description of the d band as a resonance in the conduction bands. Mr. I. Goldberg assisted with the computations.

[1]M. Weger, Rev. Mod. Phys. **36,** 175 (1964); J. Phys. Chem. Solids **31,** 1621 (1970); J. Labbe and J. Friedel, Phys. Rev. **158,** 647 (1967); J. Phys. (Paris) **27,** 153 (1966); J. Phys. (Paris) **27,** 303 (1966).
[2]J. Friedel, *Metallic solid solutions* (Benjamin, New York, 1963).
[3]V. Heine, Phys. Rev. **153,** 673 (1967).
[4]K. Rudenberg and C. W. Scherr, J. Chem. Phys. **21,** 1965 (1953); J. Chem. Phys. **22,** 151 (1959). J. R. Platt *et al., Free electron theory of conjugated molecules* (Wiley, New York, 1964).
[5]C. A. Coulson, Proc. Phys. Soc. Lond. **67,** 608 (1954).
[6]G. Della Riccia, Proceedings Intl. Conf. on Semiconductors, Exeter (1962), p. 570.
[7]A somewhat more general form of the model has recently been discussed by Montroll in J. Math. Phys. **11,** 635 (1970).
[8]N. F. Mott and H. Jones, *Theory and properties of metals and alloys* (Dover, New York, 1958).

Structure of the combinatorial generalization of hypergeometric functions for $SU(n)$ states. II

Cheng-Sheng Huang and Alfred C. T. Wu

Department of Physics, The University of Michigan, Ann Arbor, Michigan 48104

(Received 8 August 1972)

In the construction of the general $SU(5)$ states, the action of *each* individual lowering operators (raised to a power) operating on the semimaximal state leads to an operator-valued polynomial which is shown to belong to the class of generalized hypergeometric functions in the sense of Gel'fand (namely, they are Radon transform of linear forms). Three new functions are found at the $SU(5)$ level and their content in terms of known lower-hierarchy functions are explicitly exhibited. The structure of the general $SU(n)$ states due to the combined action of all lowering operators is quite complicated, but the action of each individual lowering operator taken one at a time may still be manageable for higher n, and, in the spirit of boson operator formalism, this may be one systematical way of producing high-hierarchy generalized hypergeometric functions.

I. INTRODUCTION

Previous work[1-4] shows that the combinatorics of the boson operator formalism in the construction of the $SU(n)$ states provides a natural scheme for the appearance of certain generalized hypergeometric functions. We recall that a general state is obtained by operating an appropriate string of lowering operators L_j^i (raised to a power) on the so-called semimaximal state, the latter being expressed as products of certain (antisymmetrized) creation operators acting on the vacuum state. As a result of pushing the lowering operators through the creation operators, the nonvanishing commutators thus yield an operator-valued polynomial (operating on the vacuum). For the $SU(3)$ state, this operator-valued polynomial is simply expressed as the Gauss hypergeometric function ${}_2F_1(a,b;c;x)$, as pointed out by Baird and Biedenharn,[1] namely,

$$|\text{general } SU(3) \text{ state}\rangle = \text{const (product of antisymmetrized creation operators)} \times {}_2F_1(a,b;c;x)|0\rangle. \tag{1}$$

Or, symbolically, the relevant ingredient reads

$$SU(3): (L_2^1)^n[aa] \to \text{Gauss } {}_2F_1, \tag{2}$$

where each factor of a in the bracket stands for an antisymmetrized $(a_{i_1 i_2 \cdots i_s})^\beta$ that the lowering operator has to negotiate with.

What is the generalization of the statement (1)? It was found[3,4] that a general $SU(4)$ state which is obtained via a product of three lowering operators $(L_3^1)^n$, $(L_3^2)^n$, $(L_2^1)^n$ does not have a simple form, but may be regarded as folded products of known functions. In other words, at the $SU(4)$ level the action of each individual lowering operator still yields a recognizable function, namely

$SU(4)$: operator	resulting N-fold sum	content	Gel'fand criterion: Radon transform of linear forms
$(L_3^1)^n[aaaa]$	2	Appell function F_2	Yes
$(L_3^2)^n[aa]$	1	Gauss function ${}_2F_1$	Yes
$(L_2^1)^n[aaaa]$	3	Lauricella function $F_D^{(3)}$	Yes

(3)

For higher-rank $SU(n)$ states ($n \geq 5$), it turns out that our present repertory of generalized hypergeometric functions clearly is not adequate to accommodate even the action of each individual lowering operator. One has to either invent new names for these generalized hypergeometric functions if one adopts the viewpoint that the boson operator formalism is a good way of generating (hopefully systematically) such functions, or alternatively one may try to exhibit the inner structure thereof in terms of known functions.

In this paper, we examine the structure of the general $SU(5)$ states, obtained by pushing through a set of six lowering operators, $L_4^1, L_4^2, L_4^3, L_3^1, L_3^2$, and L_2^1 (each raised to a power). Their individual action can be summarized as follows (the details are given in Sec. III):

$SU(5)$: operator	resulting N-fold sum	content	Gel'fand criterion: Radon transform of linear forms
$(L_4^1)^n[aaaaaa]$	3	Appell $F_2 \times {}_3F_2$	Yes
$(L_4^2)^n[aaaaa]$	3	Appell $F_2 \times {}_3F_2$	Yes
$(L_4^3)^n[aa]$	1	Gauss ${}_2F_1$	Yes
$(L_3^1)^n[aaaaaaa]$	8	Appell F_2 × Lauricella $F_D^{(3)}$ × Lauricella $F_B^{(3)}$	Yes
$(L_3^2)^n[aaa]$	2	Appell F_1	Yes
$(L_2^1)^n[aaaaaaa]$	6	Lauricella $F_D^{(6)}$	Yes

(4)

The following remarks are obvious at the $SU(5)$ level:

(a) The operator $(L_4^3)^n[aa]$ yields the Gauss ${}_2F_1$ function. This result is analogous to the action of $(L_2^1)^n[aa]$ at the $SU(3)$ level, or that of $(L_3^2)^n[aa]$ at the $SU(4)$ level.

(b) The operator $(L_3^2)^n[aaa]$ yields F_1, the Appell function of the first kind (in 2-variables).

(c) The operator $(L_2^1)^n[aaaaaaa]$ yields $F_D^{(6)}$, the Lauricella function of the fourth kind in 6-variables. Basically this is rather similar to the case (b) above, except that $(L_2^1)^n$ here has to push through seven factors of a's. Evidently, the action of $(L_p^{p-1})^n$ [$(s+1)$ factors of a] would yield $F_D^{(s)}$, and Lauricella function of the fourth kind in s-variables. Note that $F_D^{(1)} \equiv$ Gauss ${}_2F_1$, $F_D^{(2)} \equiv$ Appell F_1.

(d) The operators $(L_4^1)^n, (L_4^2)^n, (L_3^1)^n$ yield three essentially new functions of several variables. Two of them involve tripple sums and the other an eightfold sum. Instead of giving new names to these functions, we have exhibited their content as folded products of known functions. They are shown, however, to belong to the class of generalized hypergeometric functions in the sense of Gel'fand[5] as being the Radon transform of linear forms.

II. GENERAL $SU(5)$ STATES

As is well known, a general $SU(5)$ state may be constructed by applying a set of appropriate lowering operators to the semimaximal state.

$$|\text{general } SU(5) \text{ state}\rangle \equiv \left|\begin{pmatrix} m_{15} & m_{25} & m_{35} & m_{45} & 0 \\ & m_{14} & m_{24} & m_{34} & m_{44} \\ & & m_{13} & m_{23} & m_{33} \\ & & & m_{12} & m_{22} \\ & & & & m_{11} \end{pmatrix}\right\rangle$$

$$= \text{const}\,(L_2^1)^{n_{12}}(L_3^2)^{n_{23}}(L_3^1)^{n_{13}}(L_4^3)^{n_{34}}(L_4^2)^{n_{24}}(L_4^1)^{n_{14}} \times \left|\begin{pmatrix} m_{15} & m_{25} & m_{35} & m_{45} & 0 \\ & m_{14} & m_{24} & m_{34} & m_{44} \\ & & m_{14} & m_{24} & m_{34} \\ & & & m_{14} & m_{24} \\ & & & & m_{14} \end{pmatrix}\right\rangle$$

$$= \text{const}\,(L_2^1)^{n_{12}}(L_3^2)^{n_{23}}(L_3^1)^{n_{13}}(L_4^3)^{n_{34}}(L_4^2)^{n_{24}}(L_4^1)^{n_{14}} \times (a_{1234})^{\nu_{44}}(a_{1235})^{n_{45}}(a_{123})^{\nu_{34}}(a_{125})^{n_{35}}(a_{12})^{\nu_{24}} \times (a_{15})^{n_{25}}(a_1)^{\nu_{14}}(a_5)^{n_{15}}|0\rangle. \quad (5)$$

The set of lowering operators L_j^i are defined in Ref. 6. Those with $i < j \leq 3$ appeared in the discussion of $SU(4)$ case.[3,4] L_4^1 reads explicitly

$$L_4^1 \equiv \mathcal{E}_{12}\mathcal{E}_{13}E_{41} + \mathcal{E}_{13}E_{42}E_{21} + \mathcal{E}_{12}E_{43}E_{31} + E_{43}E_{32}E_2 \quad (6)$$

The exponents n_{ij}, ν_{ij} in Eq. (5) are shorthand notations as before,[4] namely

$$n_{ij} \equiv m_{ij} - m_{ij-1}, \qquad \nu_{ij} \equiv m_{ij} - m_{i+1,j+1}. \quad (7)$$

III. ACTION OF EACH INDIVIDUAL LOWERING OPERATOR

By a straightforward calculation, the action of each $(L_j^i)^n$ operator on the relevant set of creation operators turns out to be as follows:

Step 1, L_4^1:

$$A \equiv (a_{1235})^{n_{45}}(a_{123})^{\nu_{34}}(a_{125})^{n_{35}}(a_{12})^{\nu_{24}}(a_{15})^{n_{25}}(a_1)^{\nu_{14}}, \quad (8)$$

$$(L_4^1)^{n_{14}}A|0\rangle = \text{const}(w_0)^{n_{14}}A \times \sum_{k_1,k_2,k_3} \frac{(-n_{14})_{k_1+k_2+k_3}(-n_{25})_{k_1}(-n_{35})_{k_2}(-n_{45})_{k_3}(-s_1-1)_{k_3}}{(1+\nu_{14}-n_{14})_{k_1}(-s_2-1)_{k_2+k_3}(-s_3-2)_{k_3}} \times \frac{w_1^{k_1}}{k_1!}\frac{w_2^{k_2}}{k_2!}\frac{w_3^{k_3}}{k_3!}|0\rangle \quad (9a)$$

$$= \text{const}(w_0)^{n_{14}}A \times \sum_{k_1,k_2} \frac{(-n_{14})_{k_1+k_2}(-n_{25})_{k_1}(-n_{35})_{k_2}}{(1+\nu_{14}-n_{14})_{k_1}(-s_2-1)_{k_2}}\frac{w_1^{k_1}}{k_1!}\frac{w_2^{k_2}}{k_2!} \times \sum_{k_3} \frac{(-n_{14}+k_1+k_2)_{k_3}(-n_{45})_{k_3}(-s_1-1)_{k_3}}{(-s_3-2)_{k_3}(-s_2-1+k_2)_{k_3}} \times \frac{w_3^{k_3}}{k_3!}|0\rangle, \quad (9b)$$

where

$$\text{const} \equiv [\nu_{14}!/(\nu_{14}-n_{14})!][(s_2+1)!/(s_2+1-n_{14})!] \times [(s_3+2)!/(s_3+2-n_{14})!], \quad (10)$$

$$s_1 \equiv \nu_{14} + \nu_{24} + n_{25}, \qquad s_2 \equiv \sum_{i=1}^{2}(\nu_{i4} + n_{i+1,5}), \qquad s_3 \equiv \sum_{i=1}^{3}(\nu_{i4} + n_{i+1,5}), \quad (11)$$

$$w_0 \equiv a_4/a_1, \quad w_1 \equiv a_1a_{45}/a_4a_{15}, \quad w_2 \equiv a_{124}a_5/a_{125}a_4, \quad w_3 \equiv a_{1234}a_5/a_{1235}a_4. \quad (12)$$

As a generalized hypergeometric series in three variables, the expression (9a) does not seem to be a known function. Alternatively, Eq. (9b) shows that it may be written as a folded product of an Appell F_2 function (in two variables) with a ${}_3F_2$ function (in one variable).

Step 2, L_4^2:

$$B \equiv (a_{1235})^{n_{45}-k_3}(a_{123})^{\nu_{34}}(a_{125})^{n_{35}-k_2}(a_{124})^{k_2}(a_{12})^{\nu_{24}}, \quad (13)$$

$$(L_4^2)^{n_{24}}B|0\rangle = \text{const}(u_0)^{n_{24}}B \times \sum_{l_1l_2l_3} \frac{(-n_{24})_{l_1+l_2+l_3}(-n_{45}+k_3)_{l_2+l_3}}{(1+\nu_{24}-n_{24})_{l_1+l_3}} \times \frac{(-n_{35}+k_2)_{l_1}(-k_2)_{l_3}}{(-s_4-1)_{l_2+l_3}}\frac{u_1^{l_1}}{l_1!}\frac{u_2^{l_2}}{l_2!}\frac{u_3^{l_3}}{l_3!}|0\rangle \quad (14a)$$

$$= \text{const}(u_0)^{n_{24}}B \times \sum_{l_3} \frac{(-n_{24})_{l_3}(-n_{45}+k_3)_{l_3}(-k_2)_{l_3}}{(1+\nu_{24}-n_{24})_{l_3}(-s_4-1)_{l_3}}\frac{u_3^{l_3}}{l_3!} \times \sum_{l_1l_2} \frac{(-n_{24}+l_3)_{l_1+l_2}(-n_{35}+k_2)_{l_1}(-n_{45}+k_3+l_3)_{l_2}}{(1+\nu_{24}-n_{24}+l_3)_{l_1}(-s_4-1+l_3)_{l_2}} \times \frac{u_1^{l_1}}{l_1!}\frac{u_2^{l_2}}{l_2!}|0\rangle, \quad (14b)$$

where

$$\text{const} \equiv [\nu_{24}!/(\nu_{24}-n_{24})!][(s_4+1)!/(s_4+1-n_{24})!], \quad (15)$$

$$s_4 \equiv \nu_{24} + \nu_{34} + n_{35} + n_{45} - k_3, \quad (16)$$

$$u_0 \equiv a_{14}/a_{12}, \quad u_1 \equiv a_{12}a_{145}/a_{14}a_{125}, \quad u_2 \equiv a_{15}a_{1234}/a_{14}a_{1235}, \quad u_3 \equiv a_{12}a_{145}a_{1234}/a_{14}a_{124}a_{1235}. \quad (17)$$

The expression (14a) in three variables does not seem

to be a known function, but Eq. (14b) shows that it has the structure of a folded product of Appell F_2 function with a ${}_3F_2$ function.

Step 3, L_4^3:

$$C \equiv (a_{1235})^{n_{45}-k_3-l_2-l_3}(a_{123})^{\nu_{34}}, \tag{18}$$

$$(L_4^3)^{n_{34}} C\,|\,0\rangle = \text{const}\left(\frac{a_{124}}{a_{123}}\right)^{n_{34}} C \sum_{k_4} \times \frac{(-n_{34})_{k_4}(-n_{45}+k_3+l_2+l_3)_{k_4}}{(1+\nu_{34}-n_{34})_{k_4}} \frac{w_4^{k_4}}{k_4!}\,|\,0\rangle \tag{19a}$$

$$= \text{const}\left(\frac{a_{124}}{a_{123}}\right)^{n_{34}} C \times {}_2F_1(-n_{34}, -n_{45}+k_3+l_2+l_3; 1+\nu_{34}-n_{34}; w_4)\,|\,0\rangle, \tag{19b}$$

where

$$\text{const} \equiv \nu_{34}/(\nu_{34}-n_{34})!, \tag{20}$$

$$w_4 \equiv a_{123}a_{1245}/a_{124}a_{1235}. \tag{21}$$

For the purpose of the subsequent steps, it will be convenient to rewrite (21) with the aid of the identity $a_{123}a_{1245} = a_{124}a_{1235} - a_{125}a_{1234}$ as

$$w_4^{k_4} = \sum_{k_5} \frac{(-k_4)_{k_5}}{k_5!}\left(\frac{a_{125}a_{1234}}{a_{124}a_{1235}}\right)^{k_5}. \tag{22}$$

This has the effect of simplifying the expressions (23), (29), and (33) in not having to include the factor $(a_{1245})^{k_4}$ (which does not commute with L_3^1, L_3^2, nor a_{1345} with L_2^1).

Step 4, L_3^1:

$$D \equiv (a_{124})^{n_{34}+k_2-l_3-k_5}(a_{125})^{n_{35}-k_2-l_1+k_5}(a_{145})^{l_1+l_3}(a_{12})^{\nu_{24}-n_{24}+l_1+l_3}(a_{14})^{n_{24}-l_1-l_2-l_3}(a_{15})^{n_{25}-k_1+l_2}(a_1)^{\nu_{14}-n_{14}+k_1}, \tag{23}$$

$$(L_3^1)^{n_{13}} D\,|\,0\rangle = \text{const}(v_0)^{n_{13}} D \sum_{\sigma} \frac{(-1)^{\sigma_6+\sigma_8}(-n_{13})_{\sigma_1+\sigma_2+\cdots+\sigma_8}(-n_{25}+k_1-l_2)_{\sigma_1+\sigma_6}}{(1+\nu_{14}-n_{14}-n_{13}+k_1)_{\sigma_1+\sigma_2+\sigma_3+\sigma_6+\sigma_7+\sigma_8}}$$

$$\times \frac{(-n_{24}+l_1+l_2+l_3)_{\sigma_2+\sigma_7}(-l_1-l_3)_{\sigma_3+\sigma_8}(-n_{34}-k_2+l_3+k_5)_{\sigma_4+\sigma_6}}{(-s_5-1)_{\sigma_4+\sigma_5+\sigma_6+\sigma_7+\sigma_8}}$$

$$\times (-n_{35}+k_2+l_1+k_5)_{\sigma_5+\sigma_7}(-\nu_{24}+n_{24}-l_1-l_3)_{\sigma_8} \prod_{j=1}^{8} \frac{v_j^{\sigma_j}}{\sigma_j!}\,|\,0\rangle \tag{24a}$$

$$= \text{const}(v_0)^{n_{13}} D \sum_{\sigma_4,\sigma_5} \frac{(-n_{13})_{\sigma_4+\sigma_5}(-n_{34}+k_2+l_3-k_5)_{\sigma_4}(-n_{35}+k_2+l_1+k_5)_{\sigma_5}}{(-s_5-1)_{\sigma_4+\sigma_5}} \frac{v_4^{\sigma_4}}{\sigma_4!}\frac{v_5^{\sigma_5}}{\sigma_5!}$$

$$\times \sum_{\hat\sigma_1\hat\sigma_2\hat\sigma_3} \frac{(-n_{13}+\sigma_4+\sigma_5)_{\hat\sigma_1+\hat\sigma_2+\hat\sigma_3}(-n_{25}+k_1-l_2)_{\hat\sigma_1}(-n_{24}+l_1+l_2+l_3)_{\hat\sigma_2}(-l_1-l_3)_{\hat\sigma_3}}{(1+\nu_{14}-n_{14}-n_{13}+k_1)_{\hat\sigma_1+\hat\sigma_2+\hat\sigma_3}} \frac{v_1^{\hat\sigma_1}v_2^{\hat\sigma_2}v_3^{\hat\sigma_3}}{\hat\sigma_1!\hat\sigma_2!\hat\sigma_3!}$$

$$\times \sum_{\sigma_6,\sigma_7,\sigma_8} \frac{(-1)^{\sigma_7}(-n_{34}-k_2+l_1+k_5+\sigma_4)_{\sigma_6}(-n_{35}+k_2+l_1+k_5+\sigma_5)_{\sigma_7}(-\nu_{24}+n_{24}-l_1-l_3)_{\sigma_8}}{(-s_5-1+\sigma_4+\sigma_5)_{\sigma_6+\sigma_7+\sigma_8}}$$

$$\times (-\hat\sigma_1)_{\sigma_6}(-\hat\sigma_2)_{\sigma_7}(-\hat\sigma_3)_{\sigma_8} \frac{\hat v_6^{\sigma_6}\hat v_7^{\sigma_7}\hat v_8^{\sigma_8}}{\sigma_6!\sigma_7!\sigma_8!}\,|\,0\rangle, \tag{24b}$$

where

$$\text{const} \equiv [(\nu_{14}-n_{14}+k_1)!/(\nu_{14}-n_{14}-n_{13}+k_1)!] \times [(s_5+1)!/(s_5+1-n_{13})!], \tag{25}$$

$$s_5 \equiv n_{25}+n_{34}+n_{35}+\nu_{14}+\nu_{24}-n_{14}, \tag{26}$$

$$v_0 \equiv a_3/a_1, \quad v_1 \equiv a_1a_{35}/a_3a_{15}, \quad v_2 \equiv a_1a_{34}/a_3a_{14}, \quad v_3 \equiv a_1a_{345}/a_3a_{145}, \quad v_4 \equiv a_{123}a_4/a_{124}a_3, \quad v_5 \equiv a_{123}a_5/a_{125}a_3, \quad v_6 \equiv a_1a_{45}a_{123}/a_3a_{15}a_{124}, \quad v_7 \equiv a_1a_{45}a_{123}/a_3a_{14}a_{125}, \quad v_8 \equiv a_1a_{45}a_{123}/a_3a_{12}a_{145}. \tag{27}$$

In Eq. (24b), we have for $j = 1, 2, 3$

$$\hat\sigma_j \equiv \sigma_j + \sigma_{j+5}, \quad \hat v_{j+5} \equiv v_{j+5}/v_j. \tag{28}$$

The expression (24a) does not seem to correspond to a known function. On the other hand, Eq. (24b) shows that it has the following structure: Appell function (in v_4, v_5), Lauricella $F_D^{(3)}$ (in v_1, v_2, v_3), and Lauricella $F_B^{(3)}$ (in v_6, v_7, v_8). The last which is a generalization of the Appell F_3 function makes its first appearance at the $SU(5)$ level.

Step 5, L_3^2:

$$E \equiv (a_{124})^{n_{34}+k_2-l_3-k_5-\sigma_4-\sigma_6}(a_{125})^{n_{35}-k_2-l_1+k_5-\sigma_5-\sigma_7} \times (a_{12})^{\nu_{24}-n_{24}+l_1+l_3-\sigma_8} \tag{29}$$

$$(L_3^2)^{n_{23}} E\,|\,0\rangle = \text{const}(\mu_0)^{n_{23}} E \times \sum_{r_1,r_2} \frac{(-n_{23})_{r_1+r_2}(-n_{34}-k_2+l_3+k_5+\sigma_4+\sigma_6)_{r_1}}{(1+\nu_{24}-n_{24}+l_1+l_3-\sigma_8)_{r_1+r_2}}$$

$$\times (-n_{35}+k_2+l_1+k_5+\sigma_5+\sigma_7)_{r_2}\frac{\mu_1^{r_1}\mu_2^{r_2}}{r_1!r_2!}|0\rangle \tag{30a}$$

$$= \text{const}(\mu_0)^{n_{23}}E \times F_1(-n_{23};-n_{34}-k_2+l_3+k_5+\sigma_4+\sigma_6, -n_{35}+k_2+l_1+k_5+\sigma_5+\sigma_7;\mu_1,\mu_2)|0\rangle. \tag{30b}$$

where

$$\text{const} \equiv (\nu_{24}-n_{24}+l_1+l_3-\sigma_8)!/(\nu_{24}-n_{24}-n_{23}+l_1+l_3-\sigma_8)!, \tag{31}$$

$$\mu_0 \equiv a_{13}/a_{12}, \quad \mu_1 \equiv a_{12}a_{134}/a_{13}a_{124}, \quad \mu_2 \equiv a_{12}a_{135}/a_{13}a_{125}. \tag{32}$$

Step 6, $L_{\frac{1}{2}}$:

$$F \equiv (a_{134})^{r_1}(a_{135})^{r_2}(a_{145})^{l_1+l_3-\hat{\sigma}_3}(a_{13})^{n_{23}-r_1-r_2} \times (a_{14})^{n_{24}-l_1-l_2-l_3-\hat{\sigma}_2}(a_{15})^{n_{25}-k_1+l_2-\hat{\sigma}_1} \times (a_1)^{\nu_{14}-n_{14}-n_{13}+k_1+\hat{\sigma}_1+\hat{\sigma}_2+\hat{\sigma}_3} \tag{33}$$

$$(L_{\frac{1}{2}})^{n_{12}}F|0\rangle = \text{const}(a_2/a_1)^{n_{12}}F \times F_D^{(6)}(-n_{12};-n_{25}+k_1-l_2+\hat{\sigma}_1, -n_{24}+l_1+l_2+l_3+\hat{\sigma}_2,-l_1-l_3+\hat{\sigma}_3, -n_{23}+r_1+r_2,-r_1,-r_2; 1+\nu_{14}-n_{14}-n_{13}-n_{12}+k_1+\hat{\sigma}_1+\hat{\sigma}_2+\hat{\sigma}_3; x_1,x_2,x_3,x_4,x_5,x_6)|0\rangle \tag{34}$$

where

$$\text{const} \equiv (\nu_{14}-n_{14}-n_{13}+k_1+\hat{\sigma}_1+\hat{\sigma}_2+\hat{\sigma}_3)!/(\nu_{14}-n_{14}-n_{13}-n_{12}+k_1+\hat{\sigma}_1+\hat{\sigma}_2+\hat{\sigma}_3)!, \tag{35}$$

$$x_1 \equiv a_1a_{25}/a_2a_{15}, \quad x_2 \equiv a_1a_{24}/a_2a_{14}, \quad x_3 \equiv a_1a_{23}/a_2a_{13}, \quad x_4 \equiv a_1a_{245}/a_2a_{145}, \quad x_5 \equiv a_1a_{235}/a_2a_{135}, \quad x_6 \equiv a_1a_{234}/a_2a_{134}. \tag{36}$$

IV. GEL'FAND CRITERION: RADON TRANSFORM OF LINEAR FORMS

One class of generalized hypergeometric functions has the property that they are Radon transforms of *linear* forms. It so happens that all the known low-hierarchy functions such as Gauss, Appell, and Lauricella functions satisfy this Gel'fand criterion.[5]

From the expression (4), the simple functions associated with the action of each individual operator $(L_{\frac{3}{4}})^n$, $(L_{\frac{2}{3}})^n$, $(L_{\frac{1}{2}})^n$ obviously have this property. For the others, it is not apparent from their contents as folded products of simple functions. In general, the Gel'fand criterion which holds for each constituent may not be preserved under folded multiplication. However, it is rather remarkable that the functions associated with the action of each operator $(L_{\frac{1}{4}})^n$, $(L_{\frac{2}{4}})^n$, $(L_{\frac{1}{3}})^n$, at the $SU(5)$ level *still* satisfy the Gel'fand criterion. The proof of this statement, which consists of using well-known integral representation for each constituent and a simple change of variables, is left for the reader.

ACKNOWLEDGMENT

The paper was written while the second-named author (A. Wu) was at the University of Michigan Cyclotron Laboratory. He wishes to thank Professor W. C. Parkinson for the hospitality extended to him.

[1]G. E. Baird and L. C. Biedenharn, J. Math. Phys. **4**, 1449 (1963).
[2]M. Ciftan and L. C. Biedenharn, J. Math. Phys. **10**, 221 (1969). W. J. Holman III, J. Math. Phys. **10**, 1710 (1969).
[3]M. Ciftan, J. Math. Phys. **10**, 1635 (1969).
[4]A. C. T. Wu, J. Math. Phys. **12**, 437 (1971), referred to here as I.
[5]I. M. Gelfand, M. I. Graev, and N. Ya. Vilenkin, *Generalized functions* (Academic, New York, 1966), Vol. V.
[6]J. G. Nagel and M. Moshinsky, J. Math. Phys. **6**, 682 (1965).

Nonlinear stability problems for the sine–Gordon equation*

Andrew J. Callegari and Edward L. Reiss
Courant Institute of Mathematical Sciences, New York University, New York, New York 10012
(Received 11 July 1972; revised manuscript received 3 October 1972)

Two stability problems for the nonlinear sine–Gordon equation are studied. The stability of a class of time independent (static) solutions is studied using linear dynamic stability theory. An asymptotic approximation of the nonlinear transient response to small disturbances of an unstable static state is obtained by the two time method. Interpretations of the results are given for the continuous pendulum problem and for the Josephson tunnel junction. A proof of the validity of the asymptotic approximation is given.

1. INTRODUCTION AND FORMULATION

The one-dimensional sine–Gordon equation,

$$\Phi_{tt} - \Phi_{xx} + \lambda^2 \sin\Phi = 0, \tag{1.1}$$

has been used as a model to describe the dynamics of Josephson tunnel junctions and in a variety of other problems; see, e.g., Refs. 1–3 and references given there. The Josephson junction consists of two superconductor strips separated by a thin dielectric film. If the film is sufficiently thin, then there is a coherence between Cooper pairs of electrons on opposite sides of the film, and a superconducting tunnel current passes through the film. Josephson proposed that the tunneling current is given by $J \sin\Phi$. Here J is a constant that depends on the thickness of the film and the temperature. Φ is the difference in phases on opposite sides of the film of the Cooper pairs in the two superconductors. Dimensionless variables are used in (1.1). For example, distance is scaled by the length D of the junction and time is scaled by $D(cl)^{1/2}$, where l is the transmission line equivalent series inductance per unit area and c is the capacitance of the film per unit area.[1]

The parameter λ^2 in (1.1) is given by

$$\lambda^2 \equiv (2e/\hbar)D^2 l J, \tag{1.2}$$

where $(\hbar/2e)$ is the flux quantum.

Scott[2] has constructed a mechanical system whose equations of motion are approximated by (1.1). The system consists of a series of pendulums with a common horizontal axis of rotation. The axis is fixed in space. The pendulums oscillate in parallel planes. Furthermore, the pendulums are interconnected by an axial torsion spring. The spring resists the relative rotation of neighboring pendulums. We refer to this system as a continuous pendulum. We shall also refer to it as the pendulum when there can be no confusion with a single pendulum. Equation (1.1) is the limit of the equations of motion of the pendulums as the distances between them vanish. $\Phi(x,t)$ is the angle of rotation of the continuous pendulum and x is the dimensionless axial distance. The time scale is $lD\sqrt{m/c}$ where m is mass per unit axial length D, l is the hanging length, c is the torsional spring constant, and λ^2 is given by

$$\lambda^2 \equiv lmgD^2/c, \tag{1.3}$$

where g is the gravitational acceleration.

We generalize (1.1) by introducing linear dissipation. Thus we consider the damped sine–Gordon equation,

$$\Phi_{tt} - \Phi_{xx} + \Gamma\Phi_t + \lambda^2 \sin\Phi = 0. \tag{1.4}$$

Here $\Gamma \geq 0$ is proportional to the damping constant. Dissipation can occur in the pendulum from friction in the rotation axis or from air resistance. In the Josephson junction the dissipation term corresponds to losses due to tunneling of nonsuperconducting electrons.[1]

We shall consider the following stability problems. The continuous pendulum is uniformly rotated by an angle of π so that it is balanced in the upward equilibrium position, $\Phi(x,t) = \pi$. In the first problem, the ends of the continuous pendulum are not restrained against rotation. A small disturbance is then applied to the pendulum. We wish to study the resulting motion. Since a single pendulum with a fixed pivot is unstable in the upward position, the continuous pendulum with unconstrained ends is unstable with respect to any initial disturbance with a nonzero mean value. For such a disturbance, the continuous pendulum would then "fall" from its upward position and oscillate about $\Phi = 0$. However, we shall study the response of the continuous pendulum to small initial disturbances with zero mean value.

For the Josephson junction problem, $\Phi = \pi$ means that the phases of the Cooper pairs in the two superconductor strips differ by π radians. We shall study the transient response of the junction in this state to small initial disturbances with zero mean values.

We define a new dependent variable $u(x,t)$ by,

$$\Phi = \pi + u. \tag{1.5}$$

Then, by substituting (1.5) into (1.4), we obtain

$$u_{tt} - u_{xx} + \Gamma u_t - \lambda^2 \sin u = 0. \tag{1.6a}$$

The mathematical statement of the stability problem is to determine solutions of (1.6a) in the region $0 \leq x \leq 1$, $t \geq 0$ that satisfy the following boundary and initial conditions:

$$u_x(0,t) = u_x(1,t) = 0, \quad \text{for} \quad t \geq 0, \tag{1.6b}$$

$$u(x,0) = F(x), \quad u_t(x,0) = G(x), \quad \text{for} \quad 0 \leq x \leq 1. \tag{1.6c}$$

The boundary conditions (1.6b) imply, for the pendulum problem, that the ends $x = 0, 1$ are free to rotate. For the superconductor problem they imply that the currents in the superconductor strips vanish at the ends since these currents are proportional to Φ_x. The functions $F(x)$ and $G(x)$ are the specified initial disturbances. The conditions that the initial data have zero mean values are

$$\int_0^1 F(x)dx = \int_0^1 G(x)dx = 0. \tag{1.7}$$

In the second problem that we study, we wish to solve (1.6) with the boundary conditions (1.6b) replaced by

$$u(0,t) = u(1,t) = 0. \tag{1.8}$$

The conditions (1.8) imply that the ends of the pendulum are restrained from rotating from their upward position. For the superconductor problem (1.8) implies that the voltage across the film vanishes at the ends of the conductor. Since the ends are restrained from rotating, the conditions (1.7) are no longer required. We analyze (1.6) in detail. The corresponding results for the problem (1.6a), (1.6c), (1.8) are discussed briefly in Secs. 6C and 7.

We shall summarize the results of this paper in terms of the unrestrained pendulum problem. A brief interpretation of the results in terms of the Josephson tunnel junction is given in Sec. 7. We study the solution of (1.6) as λ increases from zero. Since λ is defined by (1.3), increases in λ can be achieved either by increasing the mass and/or the axial length of the pendulum and by decreasing the spring's stiffness. The upward equilibrium position $u(x,t) = U_0(x) \equiv 0$ is a static solution for all values of λ. It is stable with respect to disturbances satisfying (1.7) according to the linear dynamic stability theory if $\lambda < \pi$. It is unstable for $\lambda \geq \pi$. This means that for $\lambda < \pi$ the stiffness of the spring is sufficiently large (or the mass is sufficiently small) to prevent the pendulum from moving away from the upward position. A small disturbance will cause the pendulum to execute a small amplitude oscillation about the upward position. If $\Gamma > 0$, the oscillations will eventually de damped and the pendulum will remain in the upward position.

At $\lambda = \pi$, two other static solutions, $\pm U_1(x) \not\equiv 0$, differing only in their sign, branch from U_0. They exist for all $\lambda \geq \pi$. For λ near π, U_1 is approximately proportional to $\cos\pi x$. Thus when λ exceeds the critical value of π, the mass of the pendulum is sufficiently large and/or the spring is sufficiently weak so that the upward state is unstable and new, deflected equilibrium positions can occur. These deflected states are unstable for arbitrary initial disturbances according to the linear dynamic stability theory. However, as we show in Sec. 3, if the disturbances are suitably restricted they are stable for all $\lambda > \pi$.

For each $\lambda = n\pi$, $n = 2, 3, \cdots$, two new static states $\pm U_n(x)$ branch from U_0 and exist for all $\lambda > n\pi$. U_n has n internal nodes. It is linearly unstable for $n \geq 2$ even with the restricted data for which U_1 is stable.

In Secs. 4, 5, and 6 we study the transient motion of the pendulum when λ is increased slightly above π and a small initial disturbance, with zero mean value, is applied to the pendulum in its upward equilibrium state. We employ a two-time method to formally obtain an asymptotic approximation of the solution of (1.6) and (1.7) as $\epsilon \to 0$. Here ϵ is a small parameter that is proportional to the amplitude of the initial data and to the magnitude of $(\lambda - \pi)^{1/2}$. This method and variations of it, have been used previously in other nonlinear stability problems; see Refs. 4–6 and references given there. A proof of the validity of the asymptotic approximation in any finite time interval is given in the Appendix.

2. THE STATIC STATES

The static states (equilibrium states) are the time independent solutions $u(x,t) = U(x)$ of (1.6a), (1.6b). Thus $U(x)$ satisfies

$$U'' + \lambda^2 \sin U = 0, \quad \text{for} \quad 0 < x < 1, \qquad U'(0) = U'(1) = 0. \tag{2.1}$$

A prime denotes differentiation with respect to the argument of U. The boundary value problem (2.1) is identical with Euler's elastica theory for the buckling of slender, simply supported rods subjected to a compressive end thrust proportional to λ^2. Since the solutions of (2.1) are well known[7] we shall merely summarize some of their properties.

Property 1: $U = U_0 \equiv 0$ and $U = \pm\pi$ are solutions of (2.1) for all λ^2. If $U(x)$ is a solution of (2.1) for some value of λ^2, then for the same value of λ^2, $\pm U(x) + 2m\pi$ are solutions for any integer m.

Thus, without loss of generality, we can assume that

$$0 \leq U(0) \equiv \alpha < \pi. \tag{2.2}$$

It follows from (2.1) that if (2.2) is satisfied then for each λ^2, solutions occur in pairs $\pm U(x)$.

We linearize (2.1) for small U and obtain the eigenvalue problem

$$U'' + \lambda^2 U = 0, \quad \text{for} \quad 0 < x < 1, \qquad U'(0) = U'(1) = 0. \tag{2.3}$$

It has nonzero solutions if and only if $\lambda^2 = \lambda_n^2$ and $U = Z_n(x)$. The eigenvalues λ_n^2 and the eigenfunctions $Z_n(x)$ are given by

$$\lambda_n^2 \equiv (n\pi)^2, \quad Z_n \equiv A_n \psi_n(x), \quad n = 1, 2, \cdots, \tag{2.4a}$$

where the A_n are arbitrary constants and the ψ_n are defined by

$$\psi_n(x) \equiv \cos n\pi x, \quad n = 1, 2, \cdots. \tag{2.4b}$$

Property 2: $U \equiv 0$ is the only solution of (2.1) with the restriction (2.2) for $0 < \lambda^2 < \lambda_1^2$.

Property 3: A nontrivial solution $U_n(x)$ of (2.1) branches from $U \equiv 0$ at $\lambda^2 = \lambda_n^2$, $n = 1, 2, \cdots$, and exists for all $\lambda^2 \geq \lambda_n^2$. On each branch λ and U_n are given by

$$\lambda = \lambda_{(n)}(k) \equiv 2nK(k), \quad n = 1, 2, \ldots, \qquad U_n(x;k) \equiv \cos^{-1}\{1 - 2k^2 \mathrm{sn}^2[(1 + 2nx)K(k)]\}, \tag{2.5}$$

and K and k are defined by[8]

$$k \equiv \sin\frac{\alpha}{2}, \quad K(k) \equiv \int_0^{\pi/2} (1 - k^2 \sin^2 s)^{-1/2} ds. \tag{2.6}$$

Here $K(k)$ is the complete elliptic integral of the first kind. It is a monotonically increasing function of the modulus, k. The linear eigenvalue problem (2.3) yields the bifurcation points of the solutions of the nonlinear problem (2.1) since $K(0) = \pi/2$, $\lambda_{(n)}(0) = 2nK(0) = \lambda_n$, and $U_n(x;0) = 0$. We refer to the solutions (2.5) as the deflected static states. We need not consider the eigenvalue $\lambda_0 = 0$ of (2.3) since there is no solution of (2.1) with $\lambda^2 > 0$ that branches from $\lambda_0 = 0$.

Property 4: $U_n(x;k)$ has exactly n internal nodes.

Thus for any λ^2 in $(\lambda_n^2, \lambda_{n+1}^2)$, there are exactly $(2n + 2)$ static states. They are the upward position $U_0 \equiv 0$, the downward position $U_0 \equiv \pi$ and the $2n$ deflected static states that correspond to the branches emanating from $\lambda_1^2, \lambda_2^2 \ldots, \lambda_n^2$; see Fig. 1.

We interpret these properties for the pendulum problem. Property 1 asserts that the vertical positions, upward or downward, are possible static states for all λ^2. Additional static states can be obtained by rotating the pendulum from a given static state by any multiple of 2π. Property 2 implies that there are no deflected static states for sufficiently small lmD^2/c. For a fixed length D this implies that the spring is too stiff compared to the mass and hence the pendulum is restrained from deflecting from the upward position. Or equivalently, for fixed m/c this implies that the pendulum is sufficiently short so that it is too stiff to deflect. As λ^2 increases, the ability of the springs to restrain deflections is diminished. Thus, at the critical value of $\lambda^2 = \lambda_1^2$, two new equilibrium states, $\pm U_1(x)$, occur.

According to Property 4, these states have exactly one internal node. As λ^2 increases further, these are the only deflected equilibrium states until $\lambda^2 = \lambda_2^2$. Then two more equilibrium states, $\pm U_2(x)$, can occur. They have exactly two internal nodes. They could conceivably be attained by restraining the pendulums in the upward position until $\lambda^2 = \lambda_2^2$. Additional static states exist as λ^2 increases as asserted by Property 3.

3. THE LINEAR DYNAMIC STABILITY THEORY

We shall use the linear dynamic theory to test the stability of the static state $U_n(x)$ for a fixed value of λ. Thus we consider one parameter families of initial data, $F(x;\eta)$ and $G(x;\eta)$, and solutions $u(x,t;\eta)$ of (1.6) such that

$$\begin{aligned} &u(x,t;0) = U_n(x), \\ &F(x;0) = U_n(x), \qquad G(x;0) = 0. \end{aligned} \tag{3.1}$$

The linear stability theory for $V(x,t) \equiv \partial u(x,t;0)/\partial\eta$ is the variational problem of (1.6) with respect to η. It is given by

$$\begin{aligned} &V_{tt} - V_{xx} + \Gamma V_t - \lambda_{(n)}^2 (\cos U_n)V = 0, \\ &V_x(0,t) = V_x(1,t) = 0, \\ &V(x,0) = F_\eta(x;0) \equiv F^0(x), \\ &V_t(x,0) = G_\eta(x;0) \equiv G^0(x). \end{aligned} \tag{3.2}$$

We first test the stability of the upward equilibrium position, $U = U_0 \equiv 0$, with respect to initial disturbances F^0 and G^0 that satisfy (1.7). We assume that F^0 and G^0 have the convergent Fourier expansions,

$$F^0(x) = \sum_{k=1}^{\infty} F_k^0 \psi_k(x), \qquad G^0(x) = \sum_{k=1}^{\infty} G_k^0 \psi_k(x). \tag{3.3}$$

The solution of (3.2) with $U_n \equiv 0$ and $\lambda_{(0)}^2 = \lambda^2$ is

$$V(x,t) = \sum_{k=1}^{\infty} S_k(t)\psi_k(x), \tag{3.4}$$

where $S_k(t)$ are the solutions of

$$\begin{aligned} &S_k'' + \Gamma S_k' + (\lambda_k^2 - \lambda^2)S_k = 0, \\ &S_k(0) = F_k^0, \qquad S_k'(0) = G_k^0, \end{aligned} \tag{3.5}$$

for $k = 1, 2, \cdots$. Primes denote differentiation with respect to t. An analysis of the solutions of (3.5) shows that the upward state is stable if

$$\begin{aligned} &\lambda < \lambda_1, \quad \text{for} \quad \Gamma = 0, \\ &\lambda \le \lambda_1, \quad \text{for} \quad \Gamma > 0. \end{aligned} \tag{3.6}$$

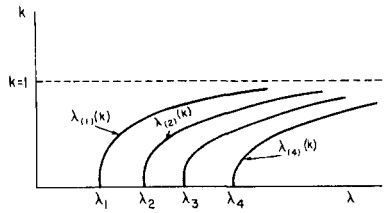

FIG. 1. A sketch of the response curves $\lambda = \lambda_{(n)}(k)$ for the first four nonlinear static states. The curves are shown only for $k \ge 0$ since they are symmetric with respect to $k = 0$. The curves branch from $k = 0$ at the eigenvalues λ_n of the linearized static theory (2.4a) Thus $\lambda_n(0) = \lambda_n$ and $\lim\lambda_{(n)}(k) = \infty$ as $k \to 1$. The downward equilibrium state $U = \pi$ corresponds to the line $k = 1$; see (2.6).

It is unstable for all other values of $\lambda \ge 0$. If the initial data (3.3) do not satisfy (1.7), then they contain the Fourier components F_0^0 and/or G_0^0. The upward state is then unstable for all values of $\lambda \ge 0$. A similar analysis shows that the downward state $U = \pi$ is stable for all values of λ^2.

We now test the stability of the deflected static states, $U = U_n$, $n = 1, 2, \cdots$. The solution $V(x,t;n,k)$ of (3.2) is then

$$V = \sum_{j=0}^{\infty} [a_j^+ \exp(ip_j^+ t) + a_j^- \exp(ip_j^- t)]\phi_j(x), \tag{3.7}$$

where $\phi_j, j = 0, 1, \cdots$, are the eigenfunctions of

$$\phi_j'' + [\mu_j + \lambda_{(n)}^2 \cos U_n]\phi_j = 0, \tag{3.8a}$$

$$\phi_j'(0) = \phi_j'(1) = 0. \tag{3.8b}$$

Primes on ϕ_j denote differentiation with respect to x. The eigenvalues $\mu_j(n,k)$ and the eigenfunctions $\phi_j(x;n,k)$ depend on the branch number n and the position on the branch characterized by the parameter k; see (2.5) and (2.6). In general, they are not known explicitly. The exponents in (3.7) are the two roots of

$$p_j^2 - i\Gamma p_j = \mu_j, \qquad j = 0, 1, \cdots. \tag{3.9}$$

Thus the $p_j^\pm$ are also functions of n and k. The coefficients in (3.7) are defined by

$$\begin{aligned} &a_j^-(n,k) \equiv (p_j^+ \hat{F}_j^0 + i\hat{G}_j^0)/(p_j^+ - p_j^-), \\ &a_j^+(n,k) = -(p_j^- \hat{F}_j^0 + i\hat{G}_j^0)/(p_j^+ - p_j^-), \quad j = 0, 1, \ldots, \end{aligned} \tag{3.10}$$

where $\hat{F}_j^0(n,k)$ and $\hat{G}_j^0(n,k)$ are Fourier coefficients that are defined by

$$F^0(x) = \sum_{j=0}^{\infty} \hat{F}_j^0 \phi_j(x), \qquad G^0(x) = \sum_{j=0}^{\infty} \hat{G}_j^0 \phi_j(x). \tag{3.11}$$

The expansions (3.7) and (3.11) are valid for sufficiently smooth F^0 and G^0 since for each n, (3.8) is a Sturm–Liouville problem. Consequently, there exists an infinite sequence $\mu_0, \mu_1, \cdots$ of eigenvalues and corresponding eigenfunctions $\phi_0, \phi_1, \cdots$ such that $\lim_{j\to\infty} \mu_j = \infty$. The eigenfunctions form a complete set. Furthermore,[9] ϕ_j has precisely j zeros on $(0,1)$, $j = 0, 1, \cdots$.

The stability results that we shall establish are summarized in

Theorem 3.1: A static state $U_n(x;k)$, $n = 1, 2, \cdots$, is linearly dynamically stable for $\lambda^2 = \lambda_{(n)}^2(k)$ if and only if

$$a_j^-(n,k) = 0, \qquad j = 0, 1, \ldots, n-1. \tag{3.12}$$

The proof of Theorem 3.1 follows from Lemmas 3.1, 3.2, and 3.3 which we now establish.

Lemma 3.1: The nth eigenvalue $\mu_n(n,k)$ for the nth static branch and the corresponding eigenfunction $\phi_n(x;n,k)$ of (3.8) are[10]

$$\mu_n(n,k) = \lambda_{(n)}^2(k)k^2, \quad \phi_n(x;n,k) = sn(z), \quad z \equiv (1+2nx)K(k), \quad n = 1,2,\cdots. \tag{3.13}$$

Proof: It is easy to show by direct substitution into (3.8) that μ_n and ϕ_n, given by (3.13), are respectively an eigenvalue and the corresponding eigenfunction. Solutions in terms of Jacobian elliptic functions are to be expected since (3.8a) is essentially a Lamé equation. Since $\mathrm{sn}(z)$ has n zeros on $(0,1)$, μ_n given in (3.13) must be the nth eigenvalue.

Lemma 3.2: The eigenvalues of (3.8) do not vanish for any k in $(0,1)$.

Proof: To prove the lemma, we show by contradiction that $\mu = 0$ is not an eigenvalue of (3.8) for any k in $(0,1)$ and for any $n = 1,2,\cdots$. Thus we assume that there is an integer n and a value of k in $(0,1)$ such that $\mu = 0$ is an eigenvalue of (3.8). Then (3.8) is reduced to

$$\phi'' + (\lambda_{(n)}^2 \cos U_n)\phi = 0, \tag{3.14a}$$

$$\phi'(0) = \phi'(1) = 0. \tag{3.14b}$$

We observe that

$$\phi = y_1(x) \equiv \mathrm{cn}(z). \tag{3.15}$$

is a solution of (3.14a), where z is defined in (3.13). We obtain a second linearly independent solution $y_2(x)$ corresponding to y_1 by standard methods. We substitute the general solution formed from y_1 and y_2 into (3.14b). Then we conclude that if (3.14) has an eigenfunction corresponding to $\mu = 0, k$ must satisfy

$$-k^2 \int_0^{\pi/2} \frac{\cos^2\theta}{(1-k^2\sin^2\theta)^{1/2}} d\theta = 0. \tag{3.16}$$

Since the left side of (3.16) is negative for any k in $(0,1)$, the contradiction is established and the lemma is proved.

Lemma 3.3: For $n = 1,2,\cdots$ and all k in $[0,1)$ we have

$$\mu_j(n,k) \begin{cases} < 0, & \text{for } j = 0,1,\ldots,n-1, \\ \geq 0, & \text{for } j = n \quad (\text{the equality holds only for } k = 0), \\ > 0, & \text{for } j = n+1,\cdots. \end{cases} \tag{3.17}$$

Proof: We recall from (2.5) and (2.6) that $\lambda_{(n)}^2(0) = (n\pi)^2$ and $\cos U_n(x;0) = 1$. Then we can solve (3.8) for $k = 0$. We obtain

$$\mu_j(n,0) = (j^2 - n^2)\pi^2, \tag{3.18}$$

$$\phi_j(x;n,0) = A_j\psi_j(x), \tag{3.19}$$

where A_j are arbitrary constants. Since (3.18) satisfies (3.17) with $k = 0$ and the eigenvalues depend continuously on k, the proof is completed by reference to Lemmas 3.1 and 3.2.

Proof of Theorem 3.1: The roots of (3.9) are

$$ip_j^{\pm} = -\tfrac{1}{2}\Gamma \mp \tfrac{1}{2}(\Gamma^2 - 4\mu_j)^{1/2}, \quad j = 1,2,\cdots. \tag{3.20}$$

If $\Gamma^2 - 4\mu_j \leq 0$, then $\exp(ip_j^{\mp}t)$ decays to zero as $t \to \infty$ when $\Gamma > 0$. It is oscillatory and hence bounded as $t \to \infty$ if $\Gamma = 0$. If $\Gamma^2 - 4\mu_j > 0$, then $\exp(ip_j^{\mp}t)$ approaches zero as $t \to \infty$ if $\mu_j > 0$. If $\mu_j > 0$, then $\exp(ip_j^{-}t)$ grows since $ip_j^{-} > 0$ and $\exp(ip_j^{+}t)$ decays. The theorem follows from Lemma 3.3 since by linear dynamic stability of the static states we mean that V is bounded as $t \to \infty$ for arbitrary initial data possessing the expansions (3.11).

All the static states U_n, $n = 1,2,\cdots$, are therefore dynamically unstable for arbitrary disturbances. However, as we see from Theorem 3.1, they are stable for suitably restricted disturbances. In particular, *if we assume that*

$$\hat{F}_0^0(1,k) = \hat{G}_0^0(1,k) = 0, \quad \text{for} \quad 0 \leq k \leq 1, \tag{3.21}$$

and the data (3.11) is otherwise arbitrary, then (3.10) and Theorem 3.1 imply that U_1 is dynamically stable and U_n, $n \geq 2$, are dynamically unstable for all k in $(0,1)$. It is easy to verify that

$$\phi_0(x;n,k) = dn(z), \quad \mu_0(n,k) = 4n^2(k^2-1)K^2(k), \tag{3.22}$$

are the principal eigenfunction and eigenvalue of (3.8). Thus (3.21) are equivalent to

$$\int_0^1 F^0 dn(z)dx = \int_0^1 G^0 dn(z)dx = 0. \tag{3.23}$$

4. THE TRANSIENT MOTION

We shall now study the transient motion that results when a small disturbance satisfying (1.7) is applied to the unstable static state $U_0 \equiv 0$ for λ slightly greater than $\lambda_1 = \pi$. Thus we consider (1.6) with the data

$$\lambda^2 = \lambda_1^2 + \epsilon^2, \tag{4.1a}$$

$$F(x) \equiv \epsilon f(x), \quad G(x) \equiv \epsilon g(x). \tag{4.1b}$$

The small parameter $\epsilon > 0$ is defined either by (4.1a) or (4.1b). By comparing (4.1a) and (2.5) with $n = 1$, we conclude the $\epsilon = O(k)$. We assume that the damping is small, i.e., that

$$\Gamma = \epsilon\gamma, \tag{4.2}$$

where γ is independent of ϵ. Furthermore, we assume that the initial data, which satisfies (1.7), has the Fourier expansion[11]

$$f(x) = \sum_{m=1}^{\infty} f_m\psi_m(x), \quad g(x) = \sum_{m=1}^{\infty} g_m\psi_m(x). \tag{4.3}$$

We refer to the initial, boundary value problem (1.6) with the data (4.1), (4.2), and (4.3) as Problem N.

We shall obtain a formal asymptotic expansion of the solution of Problem N as $\epsilon \to 0$ by a two-time method. The method was applied to a related stability problem in Ref. 5. Thus we assume that the solution of Problem N depends on two time scales: The fast time t and the slow time θ which is defined by

$$\theta = \epsilon t. \tag{4.4}$$

We seek an asymptotic expansion of the solution in the form

$$u(x,t;\epsilon) = \epsilon \sum_{j=0}^{\infty} u^j(x,t,\theta)\epsilon^j. \tag{4.5}$$

We assume that the expansion coefficients $u^j(x,t,\theta)$, $j = 0, 1, \cdots$, are bounded functions of x, t and θ for x in $[0,1]$ and t and θ in $[0,\infty)$. To determine them, we substitute (4.1)–(4.5) into Problem N. We assume that (4.5) can be differentiated term by term. Then by equating coefficients of the same powers of ϵ we find that for $j = 0, 1, \cdots$ the u^j satisfy

$$\begin{aligned} &u^j_{tt} - u^j_{xx} - \lambda_1^2 u^j = r^j, \\ &u^j_x(0,t,\theta) = u^j_x(1,t,\theta) = 0, \\ &u^j(x,0,0) = \delta_{j0} f(x), \qquad u^j_t(x,0,0) = \delta_{j0} g(x) - u^{j-1}_\theta(x,0,0). \end{aligned} \tag{4.6}$$

Here δ_{j0} is the Kronecker delta and $u^{-1} = u^{-2} \equiv 0$. The inhomogeneous terms r^j are functions of $u^0, u^1, \ldots, u^{j-1}$ and their derivatives.

The u^j are successively determined by solving the linear, initial, boundary value problems (4.6) for $j = 0, 1, \cdots$. Since the u^j are required to be bounded functions of t and θ, resonance producing terms must be eliminated at each step from the r^j. The analysis is similar to Ref. 5, although slightly more complicated. We shall omit the details of the calculations and present the results. We find that u^0 is given by

$$u^0(x,t,\theta) = b(\theta)\psi_1(x) + e^{-\gamma\theta/2} \times \sum_{m=2}^{\infty} \left(f_m \cos\omega_m t + \frac{g_m}{\omega_m} \sin\omega_m t \right) \psi_m(x). \tag{4.7}$$

Here ψ_m is defined in (2.4b) and $b(\theta)$ is the solution of the initial value problem:

$$b_{\theta\theta} + \gamma b_\theta + c(\sigma b + b^3) = 0, \qquad b(0) = f_1, \qquad b_\theta(0) = 0. \tag{4.8}$$

$\sigma(\theta)$ and c are defined by

$$\sigma(\theta) \equiv \frac{2}{\lambda_1^2}\left\{ -1 + 3e^{-\gamma\theta} \sum_{n=2}^{\infty} \left[f_n^2 + \left(\frac{g_n}{\omega_n} \right)^2 \right] \right\}, \qquad c \equiv \frac{\lambda_1^2}{8}. \tag{4.9}$$

We refer to (4.8) as the slowly varying amplitude problem. In obtaining (4.7), we required that the initial data satisfy the condition

$$g_1 = 0. \tag{4.10}$$

A similar restriction occurred in Ref. 5. If the initial data violates (4.16), then a different asymptotic expansion of the solution is required. We assume that the Fourier coefficients f_m and g_m decay sufficiently rapidly so that the series in (4.7) converges and represents the solution of (4.6) with $j = 0$.

In the Appendix we establish the validity of the asymptotic approximation u^0. Specifically we prove the following theorems.

Theorem 4.1: If $u(x,t;\epsilon)$ is a solution of Problem N, then

$$\max_{(x,t)\in D_T} |u(x,t;\epsilon)| = O(\epsilon) \quad \text{as } \epsilon \to 0.$$

Theorem 4.2: If u^0 is the approximation given by (4.7), then $u = \epsilon u^0 + \epsilon R$, where $R = O(\epsilon)$ as $\epsilon \to 0$ uniformly in D_T for any $T < \infty$.

In Theorems 4.1 and 4.2, the region D_T is defined for any positive $T < \infty$ by

$$D_T \equiv \{x,t \mid 0 \le x \le 1, \quad 0 \le t \le T\} \tag{4.11}$$

5. THE SLOWLY VARYING AMPLITUDE PROBLEM

We shall now summarize the qualitative behavior of the solution of the slowly varying amplitude problem (4.8) (cf. Ref. 5).

First, we consider (4.8) with $\gamma = 0$, i.e., we consider Problem N with no damping. Then (4.8) is an autonomous differential equation since σ, which is defined by (4.9), is a constant. The qualitative features of the solution are easily obtained from a phase plane analysis. The number of singular points of (4.8) depends upon the sign of σ. If $\sigma \ge 0$, then the origin, $b = b_\theta = 0$, of the phase plane is the unique singular point. It is a center. The trajectories in the phase plane are then closed curves (see Fig. 2), and hence $b(\theta)$ is periodic. The period depends on the magnitude of the initial data f_1.

If $\sigma < 0$, then there are three singular points,

$$b = b_\theta = 0, \tag{5.1}$$

$$b = \pm B_0, \qquad b_\theta = 0, \qquad B_0 \equiv \sqrt{-\sigma}. \tag{5.2}$$

The singular point (5.1) is a saddle and the singular points (5.2) are centers. The origin corresponds to the upward equilibrium state and the points (5.2) correspond to the two deflected equilibrium states that branch from $\lambda = \lambda_1$ (see Fig. 1). The trajectories of (4.8) with $\sigma < 0$ are closed curves (see Fig. 3). Thus, $b(\theta)$ is a periodic function. There are two types of periodic motion depending on the magnitude of f_1. If $|f_1| < f_1^*$, where f_1^* is defined by

$$f_1^* \equiv \sqrt{-2\sigma}, \tag{5.3}$$

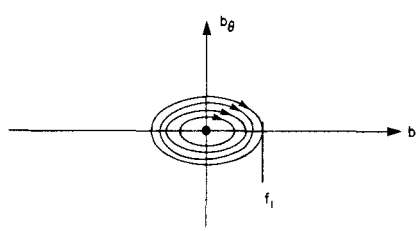

FIG. 2. Phase plane diagram for the slowly varying amplitude problem (4.8) with $\sigma \ge 0$. The origin is the unique singular point. It is a center. The trajectories in the phase plane are closed curves. They represent periodic solutions on the slow time scale $\theta = \epsilon t$.

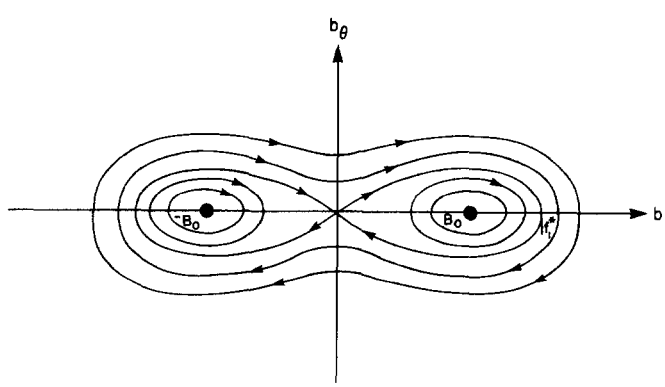

FIG. 3. Phase plane diagram for (4.8) with $\sigma < 0$.

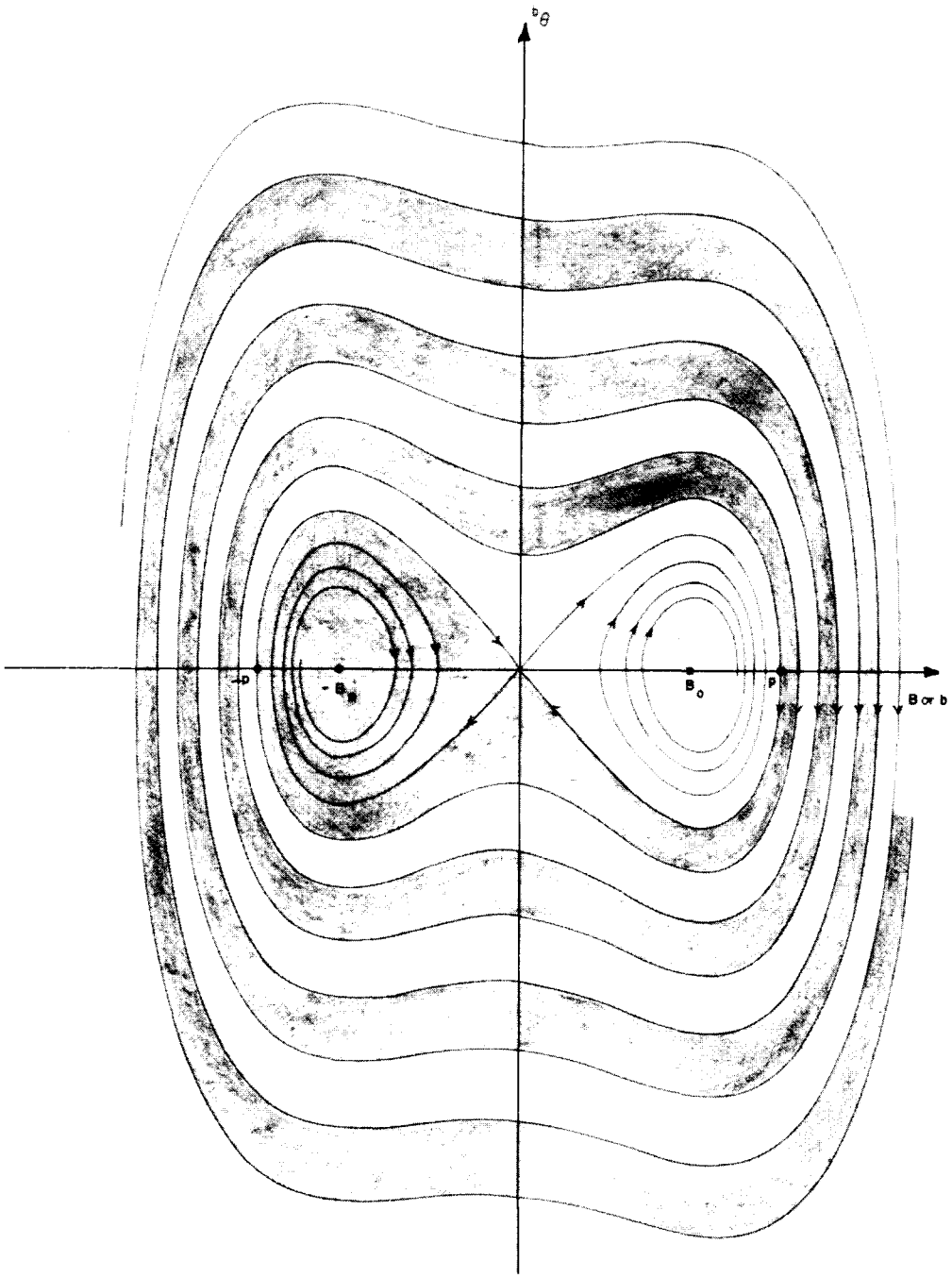

FIG. 4. Phase plane diagram for (4.8) with the monochromatic data (5.4) and for a representative value of γ. The graph was obtained by Dr. Louis Bauer by solving (4.8) numerically.

then $b(\theta)$ is a periodic motion about one of the singular points (5.2). We refer to this periodic motion as a polarized mode of vibration. The motion is about the positive (negative) singular point if $f_1 > 0$ (< 0). If $|f_1| = f_1^*$, the initial point lies on the separatrix curve. The solution then approaches the origin as $\theta \to \infty$. If $|f_1| > f_1^*$, the resulting periodic motion is referred to as the swaying mode of vibration. Then $b(\theta)$ oscillates between neighborhoods of the two singular points (5.2); see Fig. 3.

We now consider (4.8) with $\gamma > 0$ and the monochromatic initial data:

$$f_1 \neq 0, \quad f_n = g_n = 0, \quad n = 2, 3, \cdots. \tag{5.4}$$

Then from (4.9) σ and $c\sigma$ are given by

$$\sigma = -2/\lambda_1^2 < 0, \quad c\sigma = -1/4. \tag{5.5}$$

Since $\sigma < 0$, (4.8), which is an autonomous equation, has the three singular points (5.1) and (5.2). The origin is a saddle and (5.2) are stable spirals (stable nodes) if $\gamma^2 + 8c\sigma = \gamma^2 - 2\pi^4 < 0$ (> 0). A sketch of the solutions of (4.8) that pass through the origin of the phase plane, i.e., the separatrices, is given in Fig. 4 for $\gamma^2 < 2\pi^4$. The points $b = \pm p, b_\theta = 0$ (see Fig. 4), are critical values of f_1 in the following sense. If $|f_1| < p$, then $b(\theta)$ will execute a damped, polarized oscillation about one of the singular points (5.2) and approach the singular point as $\theta \to 0$. The choice of the singular point depends on

the sign of f_1; see Fig. 4. If $|f_1| \geq p$ and f_1 lies on a separatrix, then the solution will approach the origin as $\theta \to \infty$. It will oscillate several times in the swaying mode before monotonically approaching the origin. The number of oscillations depends on the size of f_1. If $|f_1| \geq p$ and f_1 is not on a separatrix, then $b(\theta)$ will execute a finite number of swaying oscillations before it is "captured" by one of the singularities (5.2). It will be captured as a damped oscillation approaching the singular point as $\theta \to \infty$. The singular point that $b(\theta)$ approaches depends on the value of f_1. If f_1 is in a shaded (unshaded) region in Fig. 4, then the final state will be at $-B_0(+B_0)$. Thus $b(\infty)$ need not be "near" $b(0)$.

For $\gamma > 0$ and initial data violating (5.4), $\sigma \neq$ const, and, consequently, (4.8) is a nonautonomous equation. Thus it is not possible to use a phase plane analysis. However, we observe from (4.9) that $\lim_{\theta\to\infty}\sigma(\theta) = -2/\lambda_1^2$ and the limit is approached exponentially. Thus it is likely that for sufficiently large θ the qualitative behavior of the solution with initial data violating (5.4) is similar to the behavior of the solution with data satisfying (5.4). A limited numerical study of (4.8) suggests that this true for all θ.

6. QUALITATIVE FEATURES OF THE APPROXIMATION

We shall discuss some of the qualitative features of the approximation (4.7) and interpret them in terms of the pendulum problem. A brief interpretation for the superconductor problem is given in Sec. 7. We introduce the following terminology to simplify the discussion; cf. Ref. 5. The leading term $f_1 \cos\pi x$ ($g_1 = 0$) in the initial data (4.3) is called the primary data. The remaining sums in (4.3) are called the secondary data. The leading term in the approximation (4.7), $b(\theta)\cos\pi x$, is called the primary motion and the sum in (4.7) is called the secondary motion.

The primary motion is independent of t. The slow time enters in the secondary motion only as a slow time decay. The secondary motion is independent of θ if $\gamma = 0$. It is essentially a high frequency motion. By using obvious trigonometric identities, we can express the sum in the secondary motion as the sum of periodic progressing waves traveling in the positive and negative x directions with phase functions

$$Q_m^{\pm}(x,t) \equiv m\pi x \pm \omega_m t. \tag{6.1}$$

Thus the secondary motion, is within the factor $e^{-\gamma\theta/2}$, an almost periodic function of two variables. We observe that the secondary motion depends only on the secondary initial data.[12] The primary motion depends on the primary data through the initial conditions (4.8) and on the secondary data through the factor σ; see (4.9).

If the amplitude of the secondary motion is "small" compared to the primary motion, then the pendulum motion as approximated by ϵu^0 is a small amplitude, high frequency oscillation (the secondary motion), superposed on the predominant slow time periodic primary motion. If, in addition, $\sigma < 0$, the slow time motion will be either polarized about one of the deflected static states or it will be in the swaying mode, depending on the magnitude of f_1; see Figs. 3 and 4. However, if $\sigma > 0$, then the primary motion will be periodic about the upward static state. The secondary motion will appear as a quivering of the pendulum as it executes its primary periodic motion. Of course, if the amplitude of the secondary motion is sufficiently large, then the high-frequency secondary motion is the main motion. The effect of the primary motion may appear as a slowly varying drift of the pendulum as it oscillates on the fast time scale.

Theorems 4.1 and 4.2 insure that for any finite time, the solution of Problem N is approximated by ϵu^0. That is, for sufficiently small ϵ, the pendulum oscillates as described either about $\pm U_1$ or about the upward vertical position. However the downward equilibrium position $U = \pm\pi$ is stable for all λ^2. Thus it is conceivable that as $t \to \infty$ the equilibrium state U_1, which is unstable for arbitrary disturbances (see Sec. 3), will repulse the transient motion. Then the pendulum will eventually "fall" and oscillate about $U = \pi$. However, if Theorems 4.1 and 4.2 were true on D_∞, then it would be impossible for the pendulum to fall to $U = \pi$.

There is further heuristic evidence to indicate that the pendulum does not fall as $t \to \infty$. First, if $\gamma > 0$, the secondary motion will eventually be damped. In fact, we have from (4.9), (4.7), and (5.2) that

$$\lim_{t\to\infty} u^0(x,t,\epsilon t) = \pm(\sqrt{2}/\lambda_1)\cos\pi x. \tag{6.2}$$

The sign in (6.2) is determined by the value of f_1; see Fig. 4 and the discussion in Sec. 5. The right side of (6.2) is precisely the coefficient of ϵ in the power series expansion of $\pm U_1(x; k(\epsilon))$, as is easily demonstrated from (2.5). This suggests that within terms of $O(\epsilon^2)$, the solution of Problem N approaches the deflected static state branching from λ_1 as $t \to \infty$.

Finally, it is not difficult to show that if the initial data (1.6c) are odd functions about $x = 1/2$, then the solution of (1.6) is also an odd function. For sufficiently large t, the solution u will provide a "disturbance" to the static states $\pm U_1$. Specifically at $t = t_0$ this disturbance is

$$F^0(x) = u(x,t_0) - U_1(x), \qquad G^0(x) = u_t(x,t_0). \tag{6.3}$$

Since U_1 and u are odd about $x = 1/2$, F^0 and G^0 satisfy (3.23). Consequently $\pm U_1$ are stable with respect to the disturbances (6.3). Thus if the transient motion for $\lambda > \lambda_1$ is started with odd initial data, the resulting motion will not be repelled by $\pm U_1$.

A. Periodic solutions

To further interpret the secondary motion, we shall now obtain, for $\gamma = 0$, periodic solutions of (1.6a), (1.6b) near $u = U_1(x)$ by the Poincaré–Linstedt method. This method was developed for ordinary differential equations. An extension of this method to partial differential equations and numerous applications are given in Ref. 13.

Thus we seek formal periodic solutions of circular frequency Ω near the static solution $U_1(x)$ in the form,

$$u(x,t,\delta) = U_1(x) + \delta\sum_{j=0}^{\infty} w^j(x,\Omega t)\delta^j, \qquad \Omega = \sum_{j=0}^{\infty}\Omega_j\delta^j, \tag{6.4}$$

where $\delta \geq 0$ is a small parameter.

A sequence of linear problems to determine the expansion coefficients is obtained by sutstituting (6.4) into (1.6a), (1.6b) and equating to zero the coefficients of each power of δ. An analysis of the first of these problems using Lemma 3.3 shows that for $s = 1, 2, \cdots$

$$w^0 = w_s^0 \equiv [B_s \sin(\sqrt{\mu}_s + \cdots)t + C_s\cos(\sqrt{\mu}_s + \cdots)t]\phi_s(x) \tag{6.5}$$

are the leading terms in the expansion (6.4) of periodic solutions of (1.6a), (1.6b). Here B_s and C_s are arbitrary constants, μ_s are the positive eigenvalues of (3.8), and ϕ_s are the corresponding eigenfunctions. Successive coefficients can be evaluated by studying the remaining linear problems. We shall not determine them since the approximations (6.5) suffice for our purposes.

We now approximate w_s^0 given by (6.5) for $\lambda_{(1)}(k)$ near λ_1. Let ϵ be defined as in (4.1). Then this is equivalent to approximating the eigenvalues and eigenfunctions of (3.8) for small ϵ by power series in ϵ. The leading terms in these power series are given by (3.18). We insert (3.18) into (6.5). Then by appropriately choosing the arbitrary constants B_s and C_s in (6.5), we observe that w_s^0, $s = 2, 3, \cdots$, are the components of the secondary motion in (4.7) with $\gamma = 0$.

B. Periodic spatial waves on the infinite interval

The asymptotic expansion (4.5) of the solution of Problem N also yields asymptotic expansions of solutions of the sine-Gordon equation which are periodic on $|x| < \infty$ for $\lambda^2 = \lambda_1^2 + \epsilon^2$. To show this, we extend the initial data, $F(x)$ and $G(x)$, and the solution of Problem N as even functions for $x < 0$ and $x > 1$. Then the extended function is a solution of (1.6a) that is periodic in x of period 2 and satisfies the initial conditions (1.6c). The initial data is also of period 2. Thus (4.7) is an asymptotic approximation of the solution of (1.6a) as $\epsilon \to 0$ for periodic initial data of period 2 on the infinite line and subject to the conditions (4.1), (4.2), and (4.3).

If we consider initial data which are even on $[0, 1]$ with respect to $x = \frac{1}{2}$, then it is not difficult to show that the resulting solution of (1.6a) is also spatially periodic of period 1. The proof requires some of the estimates derived in the Appendix. It is elementary and will not be given. Consequently, asymptotic expansions of spatially periodic solutions of period 1 can be obtained from (4.5) by specializing Problem N to initial data even about $x = \frac{1}{2}$ and by extending the data and the solution as even functions for $x < 0$ and $x > 1$.

C. Generalizations

(1) Problem (1.6) with the boundary conditions (1.6b) replaced by (1.8) can be analyzed in the same way as (1.6). There are deflected static states branching from each eigenvalue of the linearized static theory. The downward position $U = \pm \pi$ is not a possible equilibrium state with boundary conditions (1.8). If we define ν by

$$\nu \equiv U'(0), \tag{6.6}$$

then the branches $\lambda = \lambda_{(n)}(\nu)$, $n = 1, 2, \cdots$, are obtained by solving the implicit relationships,

$$\lambda = 2nK(\nu/2\lambda). \tag{6.7}$$

The solutions on the branches are given by

$$U(x; \nu) = 2 \sin^{-1}[(\nu/2\lambda)\ \mathrm{sn}(\lambda x; \nu/2\lambda)]. \tag{6.8}$$

We can show that $\lambda_{(n)}(0) = n\pi$, $\lambda_{(n)}(\nu)$ is a monotonic increasing function and for $n = 1, 2, \cdots$

$$\begin{aligned} &\lambda_{(n)}(\nu) > \nu/2, \quad \text{for } 0 \le \nu < \infty, \\ &\lambda_{(n)}(\nu) \to \nu/2 \quad \text{as } \nu \to \infty. \end{aligned} \tag{6.9}$$

The states $\pm U_1(x)$ are stable according to the linear dynamic stability theory without restrictions on the initial data. This is intuitively clear since the boundary conditions (1.8) imply that the ends of the pendulum are fixed in the upward vertical position. All other deflected static states are unstable. If the transient motion approaches a steady state as $t \to \infty$, then this steady state must be $\pm U_1(x)$. The two-time method can be used to study the transient motion for λ^2 slightly greater than the lowest eigenvalue. Results similar to those presented in Secs. 4–6 can be obtained.

(2) The formal transient analysis of Sec. 4 can be extended to initial boundary value problems for more general wave equations where $\sin u$ in (1.6a) is replaced by an arbitrary function $H(u)$. In particular, if $H(u)$ satisfies the conditions

$$H(0) = H''(0) = 0, \quad H'(0) = 1, \quad H'''(0) = -H_0 < 0, \tag{6.10}$$

then it follows directly from the calculations in Sec. 4 that the leading term in the asymptotic expansion of the solution is given by (4.7) and (4.8), where σ is defined now by

$$\sigma \equiv \frac{2}{\lambda_1^2}\left\{-\frac{1}{H_0} + 3e^{-\gamma\theta}\sum_{n=2}^{\infty}\left[f_n^2 + \left(\frac{g_n}{\omega_n}\right)^2\right]\right\}. \tag{6.11}$$

If, in addition to (6.10), there exist constants K_1, K_2 such that $H(u)$ satisfies the following conditions:

$$|H'(u)| \le K_1, \quad |H'(u) - 1| \le K_2|u|, \tag{6.12}$$

then the proof in the Appendix of the validity of the asymptotic approximation is directly applicable. If $H(u)$ violates (6.10) and (6.12), then a similar analysis may still be possible but it cannot be obtained trivially from the present results. A new computation may be required.

7. THE JOSEPHSON JUNCTION

We shall now briefly describe the results in terms of the Josephson tunnel junction. The tunneling current per unit length $j_T(x, t)$, is given by

$$j_T = -J \sin u. \tag{7.1}$$

The constant J depends on the dielectric film thickness and the temperature. It increases rapidly as the thickness of the film decreases and it increases as the temperature decreases.[3] Since λ^2 is given by (1.2), increasing values of λ can be achieved by, imagining a sequence of junctions with decreasing film thickness,[14] or by lowering the temperature of the junction, or by increasing the length of the junction.

The voltage across the film at point x and time t is proportional to $u_t(x, t)$ and the current in the superconductor strips is proportional to $u_x(x, t)$; see Ref. 1. The static solutions $U(x)$ are therefore equivalent to states with zero voltage. We refer to them as dc states. Time dependent states are called ac states. We shall refer to the conditions (1.6b) and (1.8) as the current and the voltage boundary conditions, respectively.

We denote the total tunneling current corresponding to the solution $u(x, t)$ by $I_u(t)$. It is given by

$$I_u = -JD\int_0^1 \sin u(x, t)dx. \tag{7.2}$$

By integrating (1.6a) and using (7.2), we obtain

$$I_u = \frac{JD}{\lambda^2}\left(-\int_0^1 (u_{tt} + \Gamma u_t)dx + u_x(1,t) - u_x(0,t)\right). \quad (7.3)$$

Thus corresponding to a static state $U(x)$ we have

$$I_U = (JD/\lambda^2)[U'(1) - U'(0)]. \quad (7.4)$$

If $U(x)$ satisfies the current boundary conditions, then $I_U = 0$. This implies that all dc states satisfying (1.6b) have zero total tunneling current. The pointwise tunneling current $j_T(x)$ does not vanish, in general. If the dc state U satisfies the voltage boundary conditions, then I_U vanishes if and only if $U'(1) = U'(0)$. We conclude from (6.8) that U_1 corresponding to the voltage boundary conditions is symmetric about $x = \frac{1}{2}$. Consequently, $U'(1) = -U'(0) = -\nu$, and we have from (7.4)

$$I_{U_1} = -(2JD/\lambda^2)\nu. \quad (7.5)$$

We employ (6.9) in (7.5) and obtain

$$|I_{U_1}| < I_{U_1}^* \equiv 4JD/\lambda = 2[(2h/e)(J/l)]^{1/2}, \quad (7.6)$$

where we have used (1.2). Furthermore, we conclude from (6.9) that $\lim_{\nu\to\infty} I_{U_1} = I_{U_1}^*$ is independent of D. Thus, if λ is increased by increasing D and J is held fixed, then $I_{U_1}^*$ is the maximum total current that the dc state U_1 can support.

The dc state $U = U_0 \equiv 0$ is a zero-tunneling-current state since $\sin U_0 \equiv 0$. It is unstable for $\lambda > \lambda_1$. New dc states $U_n(x)$ appear in pairs at $\lambda = \lambda_n$, $n = 1, 2, \cdots$. In Secs. 4–6 we studied the electronic switching (ac states) that occurs when the junction is disturbed from U_0 for λ slightly greater than λ_1. If $\gamma > 0$, then the ac effect is eventually dissipated and the junction reaches[15] one of the dc states $\pm U_1$. If $\gamma = 0$, then the ac effect persists, as described in Secs. 5 and 6. According to the approximation (4.7), the pointwise voltages and tunneling currents in the ac states do not vanish. If u satisfies the current boundary conditions, then we obtain from (4.5), (4.7), and (7.3) that $I_u(t) = O(\epsilon^2)$. If u satisfies (1.8), then, in general, $I_u(t) = O(\epsilon)$.

The discussion in this section implies that the voltage boundary conditions are more significant if the Josephson junction is to be used in an electronic switching device.

ACKNOWLEDGMENTS

The authors benefited from discussions with Professor D. McLaughlin. Figures 2, 3 and 4 are slightly modified versions of figures that first appeared in Ref. 5. They are published here with the permission of the editors of the Quarterly of Applied Mathematics.

APPENDIX: PROOFS OF THEOREMS 4.1 AND 4.2

Problem N is to determine a solution $u(x, t; \epsilon)$ in D_T of (1.6) with the data (4.1)–(4.3). We assume that there is class of sufficiently smooth initial data such that: Problem N has a unique solution $u(x, t, \epsilon)$ in D_T for every $T < \infty$ and all sufficiently small $\epsilon > 0$; u_{xxt} and u_{ttt} exist and are continuous in D_T; and finally $u^0_{\theta\theta\theta}$, $u^0_{\theta\theta t}$, and $u^0_{\theta tt}$ exist and are continuous in D^T, where u^0 is the solution of (4.6) with $j = 0$. The latter condition will be satisfied if the Fourier coefficients f_m and g_m in the series in (4.7) decay sufficiently rapidly. The proof of Theorem 4.1 will follow directly from Sobolev's lemma[16] and Lemmas A2 and A3. First we state and prove the lemmas.

We use the following notation for any L_2 function $v(x, t)$ in D_T:

$$\|v\|^2 \equiv \int_0^1 v^2(x,t)dx, \qquad \|v\|_t^2 \equiv \int_0^t \|v\|^2 dt. \quad (A1)$$

We shall frequently employ the elementary inequality

$$2vw \le v^2 + w^2 \quad (A2)$$

for any two L_2 functions $v(x, t)$ and $w(x, t)$.

Lemma A1: For all t in $[0, T]$, the solution u of Problem N satisfies

$$\|u\|_t^2 \le t(2\epsilon^2\|f\|^2 + t\|u_t\|_t^2). \quad (A3)$$

Proof: The proof follows by squaring the equality

$$u(x,t) = u(x,0) + \int_0^t u_s(x,s)ds \quad (A4)$$

and then using Schwarz's inequality and the initial conditions (1.6c), (4.1b).

Lemma A2: $\|u\|_t$, $\|u_x\|_t$, and $\|u_t\|_t$ are $O(\epsilon)$ uniformly in D_T.

Proof: We multiply (1.6a) by u_t, integrate with respect to x by parts, and use the boundary conditions (1.6b). This gives

$$\frac{dW}{dt} = 2\lambda^2\int_0^1 (\sin u)u_t dx - 2\epsilon\gamma\|u_t\|^2 \le \lambda^2(\|\sin u\|^2 + \|u_t\|^2) \le \lambda^2(\|u\|^2 + \|u_t\|^2), \quad (A5)$$

where $W(t)$ is defined by

$$W(t) \equiv \|u_x\|^2 + \|u_t\|^2. \quad (A6)$$

We integrate (A5) and obtain

$$W(t) \le \lambda^2(\|u\|_t^2 + \|u_t\|_t^2) + W(0). \quad (A7)$$

Lemma A1 applied to (A7) gives

$$W(t) \le \lambda^2 t(2\epsilon^2\|f\|^2 + t\|u_t\|_t^2) + \lambda^2\|u_t\|_t^2 + W(0) \le (\epsilon^2 2\lambda^2 T\|f\|^2 + W(0)) + \lambda^2(1 + T^2)\int_0^t W dt. \quad (A8)$$

Since $W(0) = O(\epsilon^2)$ as determined by the initial data (4.1b), we conclude from Gronwall's inequality applied to (A8) that

$$W(t) = O(\epsilon^2), \quad (A9)$$

uniformly in D_T. Since, from (A6), $W(t) \ge \|u_x\|^2$ and $W(t) \ge \|u_t^2\|$, the proof follows from (A9) and the definitions in (A1).

Lemma A3: $\|u_{xx}\|_t$, $\|u_{tt}\|_t$, and $\|u_{xt}\|_t$ are $O(\epsilon)$ uniformly in D_T.

Proof: Since u_{xxt} and u_{ttt} exist, we can differentiate (1.6) with respect to t and obtain an initial-boundary value problem similar to (1.6) for $w = u_t$. The coefficients in the resulting differential equation involve u.

We establish the required estimates of $u_{tt} = w_t$ and $u_{tx} = w_x$ by applying the same techniques used in the proof of Lemma 2 to the new initial-boundary value problem. Therefore, we shall omit the details of the proof. The estimate on u_{xx} is then obtained from (1.6a).

Proof of Theorem 4.1: Lemmas A2 and A3 give L_2 estimates in D_T of u and all its first and second derivatives. Sobolev's lemma[16] then establishes the maximum norm estimate of the theorem.

Proof of Theorem 4.2: The proof of this theorem is similar to the proof of Theorem 4.1. We obtain L_2 estimates of

$$\epsilon R \equiv u - \epsilon u^0 \tag{A10}$$

and its first and second derivatives in D_T and then we use Sobolev's lemma to show that $R = O(\epsilon)$ uniformly in D_T. For simplicity, we give the proof for $\gamma = 0$. It is also true for $\gamma > 0$. We insert (A10) into (1.6). Then R satisfies

$$R_{tt} - R_{xx} - \lambda^2 R = E, \tag{A11a}$$

$$R_x(0, t) = R_x(1, t) = 0, \tag{A11b}$$

$$R(x, 0) = 0, \quad R_t(x, 0) = -\epsilon u^0_\theta(x, 0, 0). \tag{A11c}$$

The inhomogeneous term in (A11a) is defined by

$$E(x, t) \equiv (\lambda^2/\epsilon)(\sin u - u) + \epsilon^2(u^0 - u^0_{\theta\theta}). \tag{A11d}$$

In deriving (A11) we have used (4.1a), (4.6) with $j = 0$ and the fact that, for $\gamma = 0$, $u^0_{t\theta} = 0$. The L_2 estimates of R are obtained by multiplying (A11a) by R_t, integrating with respect to x and using (A11b). Then, integrating the result with respect to t and using (A2), we get

$$N(t) \equiv \|R_t\|^2 + \|R_x\|^2 \le \|E\|_t^2 + \|R_t\|_t^2 + \lambda^2\|R\|^2 + N(0). \tag{A12}$$

Since $R(x, 0) = 0$ it is easy to show (see Lemma A1) that

$$\|R\|^2 \le t\|R_t\|_t^2 \le T\|R_t\|_t^2. \tag{A13}$$

Then from (A12) and (A13) we obtain

$$N(t) \le \|E\|_t^2 + N(0) + (1 + \lambda^2 T)\int_0^t N(s)ds. \tag{A14}$$

In order to apply Gronwall's inequality to (A14) we first estimate $\|E\|_t$.

Since $u = O(\epsilon)$ in D_T, $\sin u - u = O(u^3) = O(\epsilon^3)$ in D_T and

$$(\lambda^4/\epsilon^2)\|\sin u - u\| = O(\epsilon^4). \tag{A15}$$

Furthermore, we observe from (4.7) and (4.8) that u^0 and $u^0_{\theta\theta}$ are uniformly bounded in D_T. Thus (A11d) and (A15) imply that

$$\|E\|_t^2 = O(\epsilon^4), \quad \text{uniformly in } D_T^0. \tag{A16}$$

Then (A16) and Gronwall's inequality applied to (A14) show that $N(t) = O(\epsilon^2)$, uniformly in $[0, T]$ since $N(0) = O(\epsilon^2)$ by (A11c). Consequently, from the definition of $N(t)$ and (A13) we conclude that $\|R\|_t$, $\|R_t\|_t$, and $\|R_x\|_t$ are $O(\epsilon)$ uniformly in D_T. We prove that $\|R_{xx}\|_t$, $\|R_{tt}\|_t$, and $\|R_{xt}\|_t$ are also $O(\epsilon)$ by the same argument that we used to establish Lemma 3. Sobolev's lemma now yields

$$\max_{(x,t)\in D_T} |R(x, t)| = O(\epsilon),$$

which completes the proof of Theorem 4.2.

---

*This work was supported in part by the Air Force Office of Scientific Research under Grant No. AFOSR-71-2107 and by the National Science Foundation under Grant No. GP-27223.

[1]A. C. Scott, *Active and Nonlinear Wave Propagation in Electronics* (Interscience, New York, 1970).
[2]A. C. Scott, Am. J. Phys. **37**, 52 (1969).
[3]D. J. Langenberg, D. J. Scalapino, and B. N. Taylor, Proc. IEEE **54**, 560 (1966).
[4]S. Kogelman and J. B. Keller, SIAM J. Appl. Math. (Soc. Ind. Appl. Math.) **20**, 619 (1971).
[5]E. L. Reiss and B. J. Matkowsky, Q. Appl. Math. **29**, 245 (1971).
[6]B. Matkowsky, Bull. Am. Math. Soc. **17**, 620 (1970).
[7]E. L. Reiss, in *Bifurcation Theory and Nonlinear Eigenvalue Problems*, edited by J. B. Keller and S. Antman (Benjamin, New York 1969), pp. 1–16.
[8]We shall employ the standard notation sn(z), cn(z), etc. for the Jacobian elliptic functions [P. F. Byrd and M. D. Friedman, *Handbook of Elliptic Integrals for Engineers and Scientists* (Springer-Verlag, Heidelberg, 1971), 2nd ed.].
[9]E. A. Coddington and N. Levinson, *Theory of Ordinary Differential Equations* (McGraw-Hill, New York 1955).
[10]$K(k)$ is defined in (2.6).
[11]There are actually three independent small parameters in (4.1) and (4.2). They are $\lambda^2 - \lambda_1^2$, the amplitude of the initial data, and the magnitude of Γ. We assume that they are related as indicated in (4.1) and (4.2). Other relations between the parameters may lead to different representations for the solution of (1.6).
[12]This need not be true for higher order terms in the asymptotic expansion (4.5).
[13]J. B. Keller and L. Ting, Commun. Pure Appl. Math. **19**, 371 (1966).
[14]λ^2 is proportional to l. However, l decreases relatively slowly as the thickness decreases (Ref. 1).
[15]Cf. the discussions in Sec. 6.3 and at the end of Sec. 6.
[16]R. Courant and D. Hilbert, *Method of Mathematical Physics* (Interscience, New York, 1962), Vol. II.

Coset symmetrization operators

Robert Gilmore
Physics Department, University of South Florida, Tampa, Florida 33620
(Received 3 October 1972; final revision received 21 July 1972)

Coset symmetrization operators are presented and their properties discussed. Applications are indicated.

I. INTRODUCTION

Symmetrization operators based on group sums[1-4] are useful for constructing states with a specified kind of symmetry. The usual group symmetrization operator is given by

$$P^{\lambda}_{\mu\nu} = (n^{\lambda}/|G|) \sum_{g\in G} \Gamma^{\lambda}_{\mu\nu}(g)^{*} g, \tag{1}$$

where G is a finite group of order $|G|$, Γ^{λ} is a unitary irreducible representation of G of dimension n^{λ}, and the sum is over all group elements $g \in G$. These operators have a number of drawbacks:

1. A spurious index occurs in the projected state;

2. the projected states are generally neither orthogonal nor normalized to unity;

3. they involve sums over an entire group, rather than just those group elements which produce distinct linearly independent states;

4. matrix representatives for each group operation must be constructed explicitly.

II. MATHEMATICAL PRELIMINARIES

For many problems of physical interest, the application of (1) can be considerably simplified. This occurs whenever the vector space V, on which $P^{\lambda}_{\mu\nu}$ is applied, carries a permutation[5] representation of G induced[6,7] from the identity representation of a subgroup $H \subset G$.

More specifically, let V be a linear vector space with basis vectors $v_1, v_2, \ldots, v_d$ and inner product $(v_i, v_j) = \delta_{ij}$. Assume that G acts transitively[8] on the basis vectors and permutes them among themselves. Let $H \subset G$ be the stability subgroup of some particular basis vector, say v_1. Then $d = \dim V = |G|/|H|$, where $|H|$ is the order of H. The coset representatives $C_1, C_2, \ldots, C_d$ in G/H act effectively[8] on v_1. Assume that the basis vectors for $\Gamma^{\lambda}(G)$ are chosen so that the subduced[6,7] representation $\Gamma^{\lambda}(G) \downarrow H$ is in fully reduced block diagonal form. We indicate the rows and columns of $\Gamma^{\lambda}(G)$ in which the identity representation $\gamma^{id}(H)$ occurs in the restriction $G \downarrow H$, by $0_r, r = 1, 2, \ldots, n(\lambda, id)$. Here $n(\lambda, id)$ is the number of times $\gamma^{id}(H)$ occurs in $\Gamma^{\lambda}(G) \downarrow H$.

III. RESULTS

It is a simple matter to show that $P^{\lambda}_{\mu\nu}$ projects the vector 0 from v_1 unless $\nu = 0_r$. It is also straightforward to show that the coset symmetrization operators

$$S^{\lambda}_{\mu,0_r} = \left(\frac{n^{\lambda}}{|G|/|H|}\right)^{1/2} \sum_{i=1}^{d} \Gamma^{\lambda *}_{\mu,0_r}(C_i) C_i \tag{2}$$

project fully symmetrized vectors from v_1. Moreover, using the orthogonality and completeness relations[1-4,9-12] for the representations of G and H, it is straightforward to show that the vectors projected from v_1 (for all possible values of λ, μ, r) span V, are orthogonal, and are normalized to unity. The operator (2) involves none of the difficulties possessed by (1):

1. There are no spurious indices. The state projected from v_1 using $S^{\lambda}_{\mu,0_r}$ transforms like μ^{th} partner in a basis for an irreducible representation $\Gamma^{\lambda}(G)$. There are $n(\lambda, id)$ distinct invariant subspaces in V which carry irreducible representations $\Gamma^{\lambda}(G)$.

2. The projected states are orthogonal and normalized to unity.

3. The sum in (2) includes the minimal number of group elements necessary to construct symmetrized states.

4. Only the matrix elements in the columns 0_r need be constructed explicitly. Moreover, they need be constructed only for the d coset representatives C_i, rather than for all the $|G|$ group operations g.

The operator (2) may easily be extended to the case where G is compact and H is a closed subgroup.[9-12]

IV. APPLICATIONS

The operator (2) can be used to facilitate the construction of symmetrized many-electron states and symmetrized states for multilevel atomic systems, and for projecting symmetrized states suitable for energy band calculations from plane waves. As an example, the symmetrized spin states given by Schiff[13] can be written down immediately using the matrix elements for representations of S_3 given by Hamermesh[14], together with the appropriate normalizing factor $(n^{\lambda}|H|/|G|)^{1/2}$.

V. CONCLUSION

For vector spaces with the properties listed in Sec. II, the usual group symmetrization operator (1) simplifies (up to a constant multiple) to the coset symmetrization operator (2). The properties of this operator have been derived[15] and described.

ACKNOWLEDGMENTS

I wish to thank Professor G. F. Koster for useful discussions. I also wish to thank Professor H. J. Foster for the hospitality extended at Alabama A. and M. University, where this work was performed.

[1] E. P. Wigner, *Group Theory and Its Application to the Quantum Mechanics of Atomic Spectra* (Academic, New York, 1959).
[2] H. Weyl, *The Theory of Groups and Quantum Mechanics* (Dover, New York, 1931).
[3] M. Hamermesh, *Group Theory and Its Application to Physical Problems* (Addison-Wesley, Reading, Mass., 1962).
[4] L. Jansen and M. Boon, *Theory of Finite Groups* (Wiley, New York, 1967).
[5] G. de B. Robinson, *Representation Theory of the Symmetric Group* (University of Toronto Press, Toronto, 1961).
[6] G. W. Mackey, *Induced Representations of Groups and Quantum Mechanics* (Benjamin, New York, 1968).
[7] A. J. Coleman, "Induced and Subduced Representations," in *Group Theory and Its Applications*, edited by E. M. Loebl, Academic, New York, 1968).
[8] R. Hermann, *Lie Groups for Physicists* (Benjamin, New York, 1966).
[9] N. Ja. Vilenkin, *Special Functions and the Theory of Group Representations*, Transl. of Mathematical Monographs, Vol. 22 (American Mathematical Society, Providence, R.I., 1968).

[10]W. Miller, *Lie Theory and Special Functions* (Academic, New York, 1968).
[11]J. D. Talman, *Special Functions: A Group Theoretic Approach* (Benjamin, New York, 1968).
[12]R. L. Anderson and K. B. Wolf, J. Math. Phys. **11,** 3176 (1970).
[13]L. I. Schiff, *Quantum Mechanics* (McGraw-Hill, New York, 1968), 3rd edition, p. 377.
[14]Ref. 3, p. 224.
[15]R. Gilmore (unpublished).

The Nagel–Moshinsky operators for $U(p, 1) \supset U(p)$

J. Patera
Centre de recherches mathématiques, Université de Montréal, Montréal, Canada
(Received 16 August 1972)

It is shown that the operators of Nagel and Moshinsky which lower and raise the irreducible spaces of $U(n-1)$ contained in an irreducible space of $U(n)$ are also the operators which lower and raise the irreducible spaces of $U(n-1)$ contained in an irreducible space of $U(n-1,1)$. The validity of this conclusion is demonstrated for all discrete, and continuous principal and supplementary series of representations of $U(n-1,1)$.

I. INTRODUCTION

Since all nontrivial unitary representations of real noncompact semisimple Lie groups are infinite dimensional, standard problems (construction of bases with particular properties, calculation of matrix elements of operators, Clebsch–Gordan series and coefficients, etc.) are almost invariably more complicated than the same problems for the corresponding compact groups. It is worth noting, therefore, whenever the solution of a problem is valid for both the compact and the noncompact cases.

The purpose of this paper is to demonstrate that the lowering and raising operators $\mathcal{L}_n^m$ and $\mathcal{R}_m^n$, constructed by Nagel and Moshinsky[1] for the group-subgroup pair $U(n) \supset U(n-1)$, coincide with similar operators for the pair $U(n-1,1) \supset U(n-1)$. These operators are of interest because the pairs $U(p,1) \supset U(p)$ (throughout the paper $n = p+1$) are frequently considered in connection with problems in quantum physics.[2]

Our demonstration depends on a suitable choice of formalism. Indeed, we have to give a new meaning to the symbols (generators and patterns) which appear in the definition of $\mathcal{L}_n^m$ and $\mathcal{R}_m^n$, as well as in their construction and final form.

Many authors have described the irreducible unitary representations of the Lie algebras of $U(n)$ and $U(n-1,1)$ in various degrees of completeness (cf. Refs. 3–5 and further references therein). A suitable formalism for our purpose would be either that of Gel'fand and Graev,[3] or that of Ottoson.[5] Both of them use the chain of subgroups $U(n-1,1) \supset U(n-1) \supset U(n-2) \supset \cdots \supset U(1)$, and each one is an extension of the pattern formalism of Gel'fand and Tseitlin[1,3] for the compact $U(n)$. They differ by a choice of some of the generators which results in a different definition of the top line of the $U(n-1,1)$ patterns. The formalism of Gel'fand and Graev is remarkably simple, while that of Ottoson is formulated for all discrete and continuous series including the supplementary one missing in Ref. 3. We choose the formalism of Gel'fand and Graev, and as one of the results of this paper, we describe the supplementary series in the spirit of Ref. 3. This turns out to be an easy task once we know the results of Gel'fand and Graev for all other series. Using standard methods,[4–6] one easily verifies that the representations of supplementary series, as described in Sec. II are, indeed, unitary, irreducible, and nonequivalent in the usual sense.[6]

In the formalism of Ref. 3, the definition and construction of the lowering and raising operators of Nagel and Moshinsky can be repeated step-by-step without modification for the $U(n-1,1)$ case. The only differences between the Nagel–Moshinsky case and the present one, are the range of the variables h_λ and q_μ which appear in the normalization functions of $\mathcal{L}_n^m$ and $\mathcal{R}_m^n$ and, of course, the Hilbert space in which the operators act.

The present extension of the results for $U(n) \supset U(n-1)$ to $U(n-1,1) \supset U(n-1)$ is a straightforward one because in both cases the relevant subgroup is the same compact $U(n-1)$. For $U(p,q) \supset U(p,q-1), q > 1$, a similar extension would require considerable modifications of the definition of the lowering and raising operators as well as their derivation because in this case the subgroup is the noncompact $U(p,q-1)$. Nevertheless, such modifications appear feasible at least for the lower and upper bounded discrete series of $U(p,q)$ for which the Gel'fand–Graev formalism is known.[3] Similar difficulties would arise for any pair $U(n-1,1) \supset G$ or $U(n) \supset G$, where $G \neq U(n-1)$.

In Sec. II we describe the $U(p,1)$ representations. More precisely, we introduce orthonormal complete bases of patterns which span Hilbert spaces irreducible with respect to representations of the Lie algebra $L^{p,1}$ of $U(p,1)$. Matrix elements of generators of the representations are given explicitly. Our description of discrete series is a particular case of Ref. 3, and the description of the principal continuous series is that of Ref. 3, whereas the description of the supplementary series appears for the first time in this form. A brief comparison is made with the pattern formalism for the compact $U(n)$. For the sake of clarity our description of these series is self contained; our notations are that of Ref. 1. Since no new derivation is necessary, Sec. III contains only comments on the derivation of Nagel and Moshinsky and their final form of $\mathcal{L}_n^m$ and $\mathcal{R}_m^n$. Two examples are considered in Sec. IV.

II. IRREDUCIBLE UNITARY REPRESENTATIONS OF $U(p,1)$

Following Gel'fand and Graev[3] we describe here discrete and continuous series of irreducible unitary representations of the Lie algebra $L^{p,1}$ of the group $U(p,1)$ of complex matrices which leave invariant the bilinear form

$$|x_1|^2 + |x_2|^2 + \cdots + |x_p|^2 - |x_{p+1}|^2. \tag{1}$$

A representation of $L^{p,1}$ is unitary, if all operators of the representations are anti-Hermitian. A representation operator is a linear combination (with real coefficients) of the generators

$$\begin{aligned} &iC_\mu^\mu, \quad \mu = 1,2,\ldots,p+1, \\ &H_{\lambda\mu}^+ = C_\mu^\lambda - C_\lambda^\mu, \quad H_{\lambda\mu}^- = i(C_\mu^\lambda + C_\lambda^\mu), \quad \lambda,\mu \leq p, \\ &H_{p+1,p}^+ = C_{p+1}^p + C_p^{p+1}, \quad H_{p+1,p}^- = i(C_{p+1}^p - C_p^{p+1}) \end{aligned} \tag{2}$$

of the representation of $L^{p,1}$. Here C_μ^λ satisfy the commutation relations

$$[C_\mu^\lambda, C_\nu^\kappa] = \delta_\nu^\lambda C_\mu^\kappa - \delta_\mu^\kappa C_\nu^\lambda, \quad \lambda,\mu,\kappa,\nu = 1,2,\ldots,p+1. \tag{3}$$

The unitarity requirement implies that the generators C_μ^λ transform under Hermitian conjugation as follows:

$$\begin{aligned} &(C_\mu^\mu)^\dagger = C_\mu^\mu, \\ &(C_\mu^\lambda)^\dagger = C_\lambda^\mu, \quad \lambda \neq \mu, \lambda, \mu = 1, 2, \ldots, p, \\ &(C_{p+1}^p)^\dagger = -C_p^{p+1}. \end{aligned} \tag{4}$$

A representation of the generators C_μ^λ is defined by their action on an orthonormal complete basis of patterns.[3]

$$C_\mu^\mu \left| \begin{matrix} h_{1n} \cdots h_{nn} \\ \cdots \\ h_{11} \end{matrix} \right\rangle = \left(\sum_{\lambda=1}^{\mu} h_{\lambda\mu} - \sum_{\lambda=1}^{\mu-1} h_{\lambda,\mu-1} \right) \left| \begin{matrix} h_{1n} \cdots h_{nn} \\ \cdots \\ h_{11} \end{matrix} \right\rangle,$$

$$C_\lambda^{\lambda-1} \left| \begin{matrix} h_{1n} \cdots h_{nn} \\ \cdots \\ h_{11} \end{matrix} \right\rangle = a_{\lambda-1}^1 \left| h_{1,\lambda-1} - 1 \right\rangle + a_{\lambda-1}^2 \left| h_{2,\lambda-1} - 1 \right\rangle + \cdots + a_{\lambda-1}^{\lambda-1} \left| h_{\lambda-1,\lambda-1} - 1 \right\rangle,$$

$$C_{\lambda-1}^\lambda \left| \begin{matrix} h_{1n} \cdots h_{nn} \\ \cdots \\ h_{11} \end{matrix} \right\rangle = b_{\lambda-1}^1 \left| h_{1,\lambda-1} + 1 \right\rangle + b_{\lambda-1}^2 \left| h_{2,\lambda-1} + 1 \right\rangle + \cdots + b_{\lambda-1}^{\lambda-1} \left| h_{\lambda-1,\lambda-1} + 1 \right\rangle, \tag{5}$$

where $|h_{\lambda\mu} \pm 1\rangle$ denotes the original pattern in which the value of the parameter $h_{\lambda\mu}$ was raised/lowered by one; and

$$a_{\lambda-1}^\mu = \left(\frac{\prod_{j=1}^{\lambda} (h_{j\lambda} - h_{\mu,\lambda-1} - j + \mu + 1) \prod_{j=1}^{\lambda-2} (h_{j,\lambda-2} - h_{\mu,\lambda-1} - j + \mu)}{\prod_{j\neq\mu} (h_{j,\lambda-1} - h_{\mu,\lambda-1} - j + \mu + 1)(h_{j,\lambda-1} - h_{\mu,\lambda-1} - j + \mu)} \right)^{1/2}, \tag{6}$$

$$b_{\lambda-1}^\mu = \left(\frac{\prod_{j=1}^{\lambda} (h_{j\lambda} - h_{\mu,\lambda-1} - j + \mu) \prod_{j=1}^{\lambda-2} (h_{j,\lambda-2} - h_{\mu,\lambda-1} - j + \mu - 1)}{\prod_{j\neq\mu} (h_{j,\lambda-1} - h_{\mu,\lambda-1} - j + \mu)(h_{j,\lambda-1} - h_{\mu,\lambda-1} - j + \mu - 1)} \right)^{1/2}. \tag{6'}$$

The properties of other generators besides $C_\mu^\mu, C_\lambda^{\lambda-1}$, and $C_{\lambda-1}^\lambda$ are obtained from (5) by means of (3).

In order that the representation (5) is completely defined, one has to specify the series, to fix the irreducible representation within the series, and to give the range of values the parameters $h_{\lambda\mu}$ of a pattern can take in accordance with the unitarity condition (4), irreducibility, and pairwise nonequivalence of the representations. On the right side of (5) it is understood that a pattern $|h_{\lambda\mu} \pm 1\rangle$ differs from zero only if its elements are still within the range of values admissible for the pattern on the left side of (5)

A discrete series is defined by the value of a parameter $S \in (0, 1, 2, \ldots, p)$. An irreducible representation of a discrete series S is specified by the integers

$$h_{1,p+1} \geq h_{2,p+1} \geq \cdots \geq h_{p+1,p+1}. \tag{7}$$

The parameters $h_{\lambda\mu}$ are integers and satisfy the following inequalities

$$h_{\lambda,\mu+1} \geq h_{\lambda\mu} \geq h_{\lambda+1,\mu+1}, \quad \lambda \leq \mu < p, \tag{8}$$

and

$$\begin{aligned} &h_{\lambda-1,p+1} + 2 > h_{\lambda p} > h_{\lambda,p+1} \\ &\quad \text{if} \quad \lambda \leq S, \quad \text{where} \quad h_{0,p+1} = \infty, \\ &h_{\lambda+1,p+1} > h_{\lambda p} > h_{\lambda+2,p+1} - 2 \\ &\quad \text{if} \quad \lambda > S, \quad \text{where} \quad h_{p+2,p+1} = -\infty. \end{aligned} \tag{9}$$

The integers (7) specify an irreducible representation of the series S only if they satisfy also the sharp inequalities (9).

The commutation relations (3) are satisfied by (5), where the range of parameters $h_{\lambda\mu}$ is defined according to (7)–(9), provided that the following phase convention is adopted: Let M_μ^λ and $\hat{M}_\mu^\lambda$ be the number of factors (brackets) in (6) and (6'), respectively, which change their signs whenever the $U(p+1)$ range of $h_{\lambda\mu}$,

$$h_{\lambda,\mu+1} \geq h_{\lambda\mu} \geq h_{\lambda+1,\mu+1}, \quad \lambda \leq \mu \leq p, \tag{10}$$

is replaced by (8) and (9); then $M_\mu^\lambda = \hat{M}_\mu^\lambda$ and

$$\arg a_{\lambda-1}^\mu = \arg b_{\lambda-1}^\mu = M_\mu^\lambda \frac{\pi}{2}. \tag{11}$$

An irreducible unitary representation of the principal continuous series is specified by fixing the numbers

$$h_{1,p+1} = -\tfrac{1}{2}p + z, \quad h_{2,p+1} \geq h_{3,p+1} \geq \cdots \geq h_{p,p+1}, \quad h_{p+1,p+1} = \tfrac{1}{2}p + z^*, \tag{7'}$$

where z and z^* are complex conjugate numbers with $\mathrm{Im}z > 0$. Instead of (9), one has

$$\infty > h_{1p} \geq h_{2p} \geq \cdots \geq h_{pp} > -\infty. \tag{9'}$$

Equations (5), (6), and (8) hold as before.

An irreducible unitary representation of the supplementary continuous series is given when the numbers

$$h_{1,p+1} = -\tfrac{1}{2}p + \rho, \quad h_{2,p+1} \geq h_{3,p+1} \geq \cdots \geq h_{p,p+1}, \quad h_{p+1,p+1} = p - \tfrac{1}{2}\rho, \tag{7''}$$

where $0 < \rho < \frac{1}{2}p$ are fixed. The remaining $h_{\lambda\mu}$ are restricted by the inequalities (8) and (9'). Equations (5) and (6) are valid as before.

For our purposes it is essential to notice the difference between the representations of $L^{p,1}$ and L^{p+1}. Equations (3), (5), (6), and (6') are common to both. For the compact L^{p+1} one has

$$(C_\mu^\lambda)^\dagger = C_\lambda^\mu, \quad \lambda, \mu = 1, 2, \ldots, p+1, \tag{12}$$

instead of (4), and (10) instead of (7), (8), and (9).

III. REMARKS ABOUT THE DERIVATION OF $\mathfrak{L}_n^m$ and $\mathfrak{R}_m^n$

The derivation of the raising and lowering operators of

Nagel and Moshinsky does not need to be modified in order to make their operators valid for the representations of $U(p,1)$ as described in Sec. II. Therefore we limit ourselves to the definitions of $\mathcal{L}_n^m$ and $\mathcal{R}_m^n$ and to pointing out the properties of the representations which are used in their derivation and are common to $U(n)$ and $U(p,1)$. For completeness, the final form of the operators is also shown. An equality from Ref. 1 is recalled by its number. There is no ambiguity because our numbering of equations never contains the decimal point.

Let us first show that the definitions of the operators which lower and raise the irreducible spaces of $U(n-1)$ contained in an irreducible space of $U(n)$, given by Nagel and Moshinsky, can be used as the definitions of the operators which lower and raise the irreducible spaces of $U(n-1)$ contained in a $U(n-1,1)$ irreducible space.

The normalized operators $\mathcal{L}_n^m$ and $\mathcal{R}_m^n$ are defined by the following two requirements [cf. (2.9′) and (2.9″)]:

$$\text{A.}\quad \begin{aligned} w_\mu(\mathcal{L}_n^m \mid h_{\lambda\mu}\rangle) &= w_\mu(\mid h_{\lambda\mu}\rangle) - \delta_\mu^m, \\ w_\mu(\mathcal{R}_m^n \mid h_{\lambda\mu}\rangle) &= w_\mu(\mid h_{\lambda\mu}\rangle) + \delta_\mu^m. \end{aligned} \qquad (1 \leq \mu < n)$$

Here δ_μ^m is the Kronecker symbol, $\mathcal{L}_n^m \mid h_{\lambda\mu}\rangle$ and $\mathcal{R}_m^n \mid h_{\lambda\mu}\rangle$ are some linear transformations of the pattern $\mid h_{\lambda\mu}\rangle$, and w_μ is the μth component of the weight of the pattern.[7] It is given by

$$w_\mu(\mid h_{\lambda\mu}\rangle) = \sum_{\lambda=1}^{\mu} h_{\lambda\mu} - \sum_{\lambda=1}^{\mu-1} h_{\lambda,\mu-1}. \qquad (13)$$

$$\text{B.}\quad \mathcal{L}_n^m \left| \begin{matrix} h \\ q \end{matrix} \right\rangle = \left| \begin{matrix} h_1 & h_2 & \cdots & h_n \\ & q_1 & q_2 \cdots q_m - 1 \cdots & q_{n-1} \end{matrix} \right\rangle,$$

$$\mathcal{R}_m^n \left| \begin{matrix} h \\ q \end{matrix} \right\rangle = \left| \begin{matrix} h_1 & h_2 & \cdots & h_n \\ & q_1 & q_2 \cdots q_m + 1 \cdots & q_{n-1} \end{matrix} \right\rangle,$$

where $\left| \begin{matrix} h \\ q \end{matrix} \right\rangle$ is a pattern

$$\left| \begin{matrix} h \\ q \end{matrix} \right\rangle \equiv \left| \begin{matrix} h_1 & h_2 & \cdots & h_n \\ & q_1 & q_2 \cdots q_{n-1} & \end{matrix} \right\rangle$$

$$\equiv \left| \begin{matrix} h_1 & h_2 & \cdots & h_n \\ q_1 & q_2 & \cdots & q_{n-1} \\ q_1 & q_2 & \cdots & q_{n-2} \\ & \cdots & & \\ & q_1 & q_2 & \\ & & q_1 & \end{matrix} \right\rangle, \qquad (14)$$

with $h_{\lambda n} \equiv h_\lambda$ and $h_{\lambda\mu} = h_{\lambda,n-1} \equiv q_\lambda$ for $\lambda \leq \mu = 1, 2, \ldots, n-1$.

The requirement A can be applied to the patterns of $U(n-1,1)$ as well as $U(n)$ because w_μ is the same function (13) of $h_{\lambda\mu}$. Similarly the requirement B can be used because patterns (14) are found among patterns of every series of $U(n-1,1)$, as can be verified by inspection of the inequalities for $h_{\lambda\mu}$ in Sec. II. Because the irreducible space of $U(n-1)$ contained in that of $U(n-1,1)$, is given by q_λ $(\lambda = 1, 2, \ldots, n-1)$, the operators $\mathcal{L}_n^m$ and $\mathcal{R}_m^n$, indeed, raise and lower these spaces.

The generators C_μ^λ of $U(n-1,1)$ and $U(n)$ satisfy the same commutation relations (3) and (2.2); they act on a $U(n-1,1)$ or $U(n)$ pattern in the same way, (5) and (5′); their matrix elements, (6) and (6′), are the same functions of the elements $h_{\lambda\mu}$ of the pattern. The elements $h_{\lambda\mu}$ of the pattern. The elements $h_{\lambda\mu}$ of the lower $n-1$ levels of any $U(n)$ or $U(n-1,1)$ pattern satisfy the same inequalities (8). Therefore one has also the important properties (2.7 and 2.8)

$$\begin{aligned} C_\lambda^\lambda \left| \begin{matrix} h \\ q \end{matrix} \right\rangle &= q_\lambda \left| \begin{matrix} h \\ q \end{matrix} \right\rangle, \quad \lambda < n, \\ C_\mu^\lambda \left| \begin{matrix} h \\ q \end{matrix} \right\rangle &= 0, \quad \mu < \lambda < n, \end{aligned} \qquad (15)$$

of patterns (14) regardless of whether $\left| \begin{matrix} h \\ q \end{matrix} \right\rangle$ belongs to a representation of $U(n)$ or to any series of $U(n-1,1)$. Consequently, the unnormalized operators L_n^m and R_m^n for $U(n)$ and $U(n-1,1)$ are the same polynomials of C_μ^λ [cf. (2.27a′) and (2.27b′) and (2.27a″) and (2.27b″)]:

$$\begin{aligned} L_n^m = &\left(\sum_{p=0}^{n-m-1} \sum_{\mu_p > \mu_{p-1} > \cdots > \mu_1 = m+1}^{n-1} \right. \\ &\left. \times C_{\mu_1}^m C_{\mu_2}^{\mu_1} \cdots C_{\mu_p}^{\mu_{p-1}} C_n^{\mu_p} \prod_{i=1}^{p} \mathcal{E}_{m\mu_i}^{-1} \right) \prod_{\mu=m+1}^{n-1} \mathcal{E}_{m\mu} \\ = &\prod_{n=m+1}^{n-1} \mathcal{E}_{m\mu} \sum_{p=0}^{n-m-1} \sum_{\mu_p > \mu_{p-1} > \cdots > \mu_1 = m+1}^{n-1} \\ &\times \left(\prod_{i=1}^{p} \mathcal{E}_{m\mu_i}^{-1} \right) C_n^{\mu_p} C_{\mu_p}^{\mu_{p-1}} \cdots C_{\mu_2}^{\mu_1} C_{\mu_1}^m, \end{aligned} \qquad (16)$$

$$\begin{aligned} R_m^n = &\left(\sum_{p=0}^{m-1} \sum_{\mu_p > \mu_{p-1} > \cdots > \mu_1 = 1}^{m-1} \right. \\ &\left. \times C_m^{\mu_p} C_{\mu_p}^{\mu_{p-1}} \cdots C_{\mu_2}^{\mu_1} C_{\mu_1}^n \prod_{i=1}^{p} \mathcal{E}_{m\mu_i}^{-1} \right) \prod_{\mu=1}^{m-1} \mathcal{E}_{m\mu} \\ = &\prod_{\mu=1}^{m-1} \mathcal{E}_{m\mu} \sum_{p=0}^{m-1} \sum_{\mu_p > \mu_{p-1} > \cdots > \mu_1 = 1}^{m-1} \\ &\times \left(\prod_{i=1}^{p} \mathcal{E}_{m\mu_i}^{-1} \right) C_{\mu_1}^n C_{\mu_2}^{\mu_1} \cdots C_{\mu_p}^{\mu_{p-1}} C_m^{\mu_p}, \end{aligned} \qquad (17)$$

where $1 \leq m < n$, and

$$\mathcal{E}_{\mu\lambda} = C_\mu^\mu - C_\lambda^\lambda + \lambda - \mu,$$

and also

$$\prod_{\mu=1}^{0} \mathcal{E}_{1\mu} \equiv 1.$$

The normalization functions $N_{q_m-1}^{q_m}$ and $N_{q_m+1}^{q_m}$ are defined by (5.2′ and 5.2″):

$$\begin{aligned} &\left| \begin{matrix} h_1 & & & h_n \\ & q_1 \cdots q_m - 1 \cdots & & q_{n-1} \end{matrix} \right\rangle = \mathcal{L}_n^m \left| \begin{matrix} h \\ q \end{matrix} \right\rangle \\ &= (N_{q_m-1}^{q_m})^{-1} L_n^m \left| \begin{matrix} h \\ q \end{matrix} \right\rangle, \\ &\left| \begin{matrix} h_1 & & & h_n \\ & q_1 \cdots q_m + 1 \cdots & & q_{n-1} \end{matrix} \right\rangle \\ &= \mathcal{R}_m^n \left| \begin{matrix} h \\ q \end{matrix} \right\rangle = (N_{q_m+1}^{q_m})^{-1} R_m^n \left| \begin{matrix} h \\ q \end{matrix} \right\rangle. \end{aligned} \qquad (18)$$

Since L_n^m and R_m^n are polynomials in C_μ^λ, the normalization functions are linear combinations of products of matrix elements of generators C_μ^λ. But these matrix elements are the same functions (6) and (6′) of the elements $h_{\lambda\mu}$ of patterns regardless of whether we deal with a representation of $U(n)$ or $U(n-1,1)$. The particular patterns (14) which appear in (18) contain only elements h_λ $(\lambda = 1,2,\ldots,n)$ and q_μ $(\mu = 1,2,\ldots,n-1)$. Therefore $N_{q_m^{-1}}^{q_m}$ and $N_{q_m^{+1}}^{q_m}$ are the same functions of h_λ and q_μ as the normalization functions of Nagel and Moshinsky. They are equal to [cf. (5.11′) and (5.11″)]:

$$N_{q_m^{-1}}^{q_m} = \left[-\left(\prod_{\mu=m+1}^{n-1} q_{m\mu} \Big/ \prod_{\mu=1}^{m-1} (q_{m\mu}-1)\right) \times \prod_{\mu=1}^{n} (q_m - h_n + \mu - m - 1)\right]^{1/2},$$

$$N_{q_m^{+1}}^{q_m} = (-)^{m-1}\left[-\left(\prod_{\mu=1}^{m-1} q_{m\mu} \Big/ \prod_{\mu=m+1}^{n-1} (q_{m\mu}+1)\right) \times \prod_{\mu=1}^{n} (q_m - h_\mu + \mu - m)\right]^{1/2}, \tag{19}$$

where

$$q_{\lambda\mu} = q_\lambda - q_\mu + \mu - \lambda.$$

The phase in (19) is determined according to the convention of Sec. II.

IV. TWO EXAMPLES

Although the operators $\mathfrak{L}_2^1$ and $\mathfrak{R}_1^2$ for $U(1,1)$ are trivially derived, they are a transparent illustration of our result. The $U(1,1)$ and $U(2)$ patterns are all of the type (14). From (5) and (6), we have

$$C_2^1 \left| \begin{matrix} h_1 & & h_2 \\ & q & \end{matrix} \right\rangle = [(q-h_2)(h_1-q+1)]^{1/2} \left| \begin{matrix} h_1 & & h_2 \\ & q-1 & \end{matrix} \right\rangle,$$

$$C_1^2 \left| \begin{matrix} h_1 & & h_2 \\ & q & \end{matrix} \right\rangle = [(h_1-q)(q-h_2+1)]^{1/2} \left| \begin{matrix} h_1 & & h_2 \\ & q+1 & \end{matrix} \right\rangle. \tag{20}$$

Comparing (20) with the definition of $\mathfrak{L}_2^1$ and $\mathfrak{R}_1^2$, we have immediately

$$\mathfrak{L}_2^1 = [(q-h_2)(h_1-q+1)]^{-1/2} C_2^1,$$
$$\mathfrak{R}_1^2 = [(h_1-q)(q-h_2+1)]^{-1/2} C_1^2, \tag{21}$$

which is exactly the form of $\mathfrak{L}_2^1$ and $\mathfrak{R}_1^2$ found in Ref. 1. Operators (21) are those of $U(2)$, if h_1, h_2 and q are integers such that $h_1 \geq q \geq h_2$. They belong to one of the discrete series if either $q > h_1 \geq h_2$ or $h_1 \geq h_2 > q$ for integer h_1, h_2, and q. They belong to the principal series if $h_1 = -\frac{1}{2} + z$, $h_2 = \frac{1}{2} + z^*$, and q is any integer from $(-\infty, +\infty)$. Finally they are the Nagel-Moshinsky operators for the supplementary series of $U(1,1)$ if $h_1 = -\frac{1}{2} + \rho, h_2 = \frac{1}{2} - \rho$ and q is any integer. One easily notices that the normalization functions N_{q-1}^q and N_{q+1}^q are real for $U(2)$ and purely imaginary for $U(1,1)$. Together with (4) and (12), it means that $\mathfrak{L}_2^1$ and $\mathfrak{R}_1^2$ transform under Hermitian conjugation as $(\mathfrak{L}_2^1)^+ = \mathfrak{R}_1^2$.

Our second example is form $U(2,1)$. We take the operators $\mathfrak{L}_3^m$ and $\mathfrak{R}_m^3$ of Sec. III, and verify by a straightforward computation that they satisfy the requirements A and B of Sec. III for every series.

There are five unitary series of $L^{2,1}$ described in Sec. II by the following inequalities:

I. $S = 0$; $h_{13} \geq h_{23} \geq h_{33}$
$h_{12} \geq h_{11} \geq h_{22}$, $h_{23} > h_{12} > h_{33} - 2, h_{33} > h_{22} > -\infty$.

II. $S = 1$; $h_{13} \geq h_{23} \geq h_{33}$,
$h_{12} \geq h_{11} \geq h_{22}$, $\infty > h_{12} > h_{13}, h_{33} > h_{22} > -\infty$.

III. $S = 2$; $h_{13} \geq h_{23} \geq h_{33}$,
$h_{12} \geq h_{11} \geq h_{22}$, $\infty > h_{12} > h_{13}, h_{13} + 2 > h_{22} > h_{23}$.

IV. $h_{13} = -1 + z$, h_{23} integer, $h_{33} = 1 + z^*$,
$h_{12} \geq h_{11} \geq h_{22}$, $h_{12} \geq h_{23} \geq h_{22}$.

V. $h_{13} = -1 + \rho$, h_{23} integer,
$h_{33} = 1 - \rho$, $0 < \rho < 1$
$h_{12} \geq h_{11} \geq h_{22}$, $h_{12} \geq h_{23} \geq h_{22}$.

Our subsequent reasoning holds not only for the series I–V of $U(2,1)$, but also for the $U(3)$ one:

VI. finite dimensional;
integers $h_{13} \geq h_{12} \geq h_{23} \geq h_{22} \geq h_{33}$,
$h_{12} \geq h_{11} \geq h_{22}$, $h_{13} \geq h_{12} \geq h_{23} \geq h_{22} \geq h_{33}$.

Let us now recall the definition of C_μ^λ for $U(2,1)$ and $U(3)$. Omitting the top line $h_{13} h_{23} h_{33}$ of each pattern, one has from the general rules (5) and (6):

$$C_1^1 \left| \begin{matrix} h_{12} & & h_{22} \\ & h_{11} & \end{matrix} \right\rangle = h_{11} \left| \begin{matrix} h_{12} & & h_{22} \\ & h_{11} & \end{matrix} \right\rangle,$$

$$C_2^2 \left| \begin{matrix} h_{12} & & h_{22} \\ & h_{11} & \end{matrix} \right\rangle = (h_{12} + h_{22} - h_{11}) \left| \begin{matrix} h_{12} & & h_{22} \\ & h_{11} & \end{matrix} \right\rangle,$$

$$C_3^3 \left| \begin{matrix} h_{12} & & h_{22} \\ & h_{11} & \end{matrix} \right\rangle = (h_{13} + h_{23} + h_{33} - h_{12} - h_{22}) \left| \begin{matrix} h_{12} & & h_{22} \\ & h_{11} & \end{matrix} \right\rangle,$$

$$C_2^1 \left| \begin{matrix} h_{12} & & h_{22} \\ & h_{11} & \end{matrix} \right\rangle = [(h_{11} - h_{22})(h_{12} - h_{11} + 1)]^{1/2} \left| \begin{matrix} h_{12} & & h_{22} \\ & h_{11} - 1 & \end{matrix} \right\rangle,$$

$$C_1^2 \left| \begin{matrix} h_{12} & & h_{22} \\ & h_{11} & \end{matrix} \right\rangle = [(h_{12} - h_{11})(h_{11} - h_{22} + 1)]^{1/2} \left| \begin{matrix} h_{12} & & h_{22} \\ & h_{11}+1 & \end{matrix} \right\rangle,$$

$$C_3^2 \left| \begin{matrix} h_{12} & & h_{22} \\ & h_{11} & \end{matrix} \right\rangle = a_2^1 \left| \begin{matrix} h_{12}-1 & & h_{22} \\ & h_{11} & \end{matrix} \right\rangle + a_2^2 \left| \begin{matrix} h_{12} & & h_{22}+1 \\ & h_{11} & \end{matrix} \right\rangle, \tag{22}$$

$$C_2^3 \left| \begin{matrix} h_{12} & & h_{22} \\ & h_{11} & \end{matrix} \right\rangle = b_2^1 \left| \begin{matrix} h_{12}+1 & & h_{22} \\ & h_{11} & \end{matrix} \right\rangle + b_2^2 \left| \begin{matrix} h_{12} & & h_{22}+1 \\ & h_{11} & \end{matrix} \right\rangle,$$

$$C_3^1 \left| \begin{matrix} h_{12} & & h_{22} \\ & h_{11} & \end{matrix} \right\rangle = a_{21}^{11} \left| \begin{matrix} h_{12}-1 & & h_{22} \\ & h_{11}-1 & \end{matrix} \right\rangle + a_{21}^{21} \left| \begin{matrix} h_{12} & & h_{22}-1 \\ & h_{11}-1 & \end{matrix} \right\rangle,$$

$$C_1^3 \left| \begin{matrix} h_{12} & & h_{22} \\ & h_{11} & \end{matrix} \right\rangle = b_{21}^{11} \left| \begin{matrix} h_{12}+1 & & h_{22} \\ & h_{11}+1 & \end{matrix} \right\rangle + f_{21}^{21} \left| \begin{matrix} h_{12} & & h_{22}+1 \\ & h_{11}+1 & \end{matrix} \right\rangle,$$

where for the last two equalities we have used also (3).
Here

$$a_2^1 = \left(\frac{(h_{13}-h_{12}+1)(h_{12}-h_{23})(h_{12}-h_{33}+1)(h_{12}-h_{11})}{(h_{12}-h_{22})(h_{12}-h_{22}+1)}\right)^{1/2},$$

$$a_2^2 = \left(\frac{(h_{13}-h_{22}+2)(h_{23}-h_{22}+1)(h_{22}-h_{33})(h_{11}-h_{22}+1)}{(h_{12}-h_{22}+2)(h_{12}-h_{22}+1)}\right)^{1/2},$$

$$b_2^1 = \left(\frac{(h_{13}-h_{12})(h_{12}-h_{23}+1)(h_{12}-h_{33}+2)(h_{12}-h_{11}+1)}{(h_{12}-h_{22}+1)(h_{12}-h_{22}+2)}\right)^{1/2},$$

$$b_2^2 = \left(\frac{(h_{13}-h_{22}+1)(h_{23}-h_{22})(h_{22}-h_{33}+1)(h_{11}-h_{22})}{(h_{12}-h_{22}+1)(h_{12}-h_{22})}\right)^{1/2},$$

$$a_{21}^{11} = \left(\frac{(h_{11}-h_{22})(h_{13}-h_{12}+1)(h_{12}-h_{23})(h_{12}-h_{33}+1)}{(h_{12}-h_{22})(h_{12}-h_{22}+1)}\right)^{1/2}, \tag{23}$$

$$a_{21}^{21} = -\left(\frac{(h_{12}-h_{11}+1)(h_{13}-h_{22}+2)(h_{23}-h_{22}+1)(h_{22}-h_{33})}{(h_{12}-h_{22}+2)(h_{12}-h_{22}+1)}\right)^{1/2},$$

$$b_{21}^{11} = \left(\frac{(h_{11}-h_{22}+1)(h_{13}-h_{12})(h_{12}-h_{23}+1)(h_{12}-h_{33}+2)}{(h_{12}-h_{22}+1)(h_{12}-h_{22}+2)}\right)^{1/2},$$

$$b_{21}^{21} = -\left(\frac{(h_{12}-h_{11})(h_{13}-h_{22}+1)(h_{23}-h_{22})(h_{22}-h_{33}+1)}{(h_{12}-h_{22}+1)(h_{12}-h_{22})}\right)^{1/2},$$

The equalities (22) and (23) are valid for the series I–VI. Hence without specifying the particular series, i.e., by showing the range of parameters h_{12}, h_{22}, and h_{11}, we *can* verify that the two unnormalized operators L_3^1, L_3^2, R_1^3, and R_2^3 have the two required properties.

Let us start with the lowering operator L_3^1. From (16) it follows that

$$L_3^1 = C_3^1(C_1^1 - C_2^2 + 1) + C_2^1 C_3^2 = (C_1^1 - C_2^2 + 1)C_3^1 + C_3^2 C_2^1. \tag{24}$$

From (22) and (23), we find

$$L_3^1 \left| \begin{matrix} h_{12} & & h_{22} \\ & h_{11} & \end{matrix} \right\rangle = (h_{11}-h_{22}+1)\left(\frac{(h_{13}-h_{12}+1)(h_{12}-h_{23})(h_{12}-h_{33}+1)(h_{11}-h_{22})}{(h_{12}-h_{22})(h_{12}-h_{22}+1)}\right)^{1/2} \left| \begin{matrix} h_{12}-1 & & h_{22} \\ & h_{11}-1 & \end{matrix} \right\rangle$$
$$+ (h_{11}-h_{12})\left(\frac{(h_{12}-h_{11}+1)(h_{13}-h_{22}+2)(h_{23}-h_{22}+1)(h_{22}-h_{33})}{(h_{12}-h_{22}+2)(h_{12}-h_{22}+1)}\right)^{1/2} \left| \begin{matrix} h_{12} & & h_{22}-1 \\ & h_{11}-1 & \end{matrix} \right\rangle. \tag{25}$$

The property A is readily verified from (13) and (25). Indeed,

$$w_1\left(\left| \begin{matrix} h_{12}-1 & & h_{22} \\ & h_{11}-1 & \end{matrix} \right\rangle\right) = w_1\left(\left| \begin{matrix} h_{12} & & h_{22}-1 \\ & h_{11}-1 & \end{matrix} \right\rangle\right)$$
$$= w_1\left(\left| \begin{matrix} h_{12} & & h_{22} \\ & h_{11} & \end{matrix} \right\rangle\right) - 1,$$

$$w_2\left(\left| \begin{matrix} h_{12}-1 & & h_{22} \\ & h_{11}-1 & \end{matrix} \right\rangle\right) = w_2\left(\left| \begin{matrix} h_{12} & & h_{22}-1 \\ & h_{11}-1 & \end{matrix} \right\rangle\right)$$
$$= w_1\left(\left| \begin{matrix} h_{12} & & h_{22} \\ & h_{11} & \end{matrix} \right\rangle\right). \tag{26}$$

In order to verify the requirement B, we use L_3^1 to a pattern (14) for which

$$h_{13} \equiv h_1, \quad h_{23} \equiv h_2, \quad h_{33} \equiv h_3,$$
$$h_{12} \equiv q_1, \quad h_{22} \equiv q_2, \quad h_{11} = q_1. \tag{27}$$

Clearly the values (27) of h_{λ_μ} do not contradict the inequalities for any of the $U(2,1)$ and $U(3)$ series of representations. Inserting (27) into (25), we have

$$L_3^1 \left| \begin{matrix} q_1 & & q_2 \\ & q_1 & \end{matrix} \right\rangle = [(q_1-q_2+1)(h_1-q_1+1)(q_1-h_2)$$
$$\times (q_1-h_3+1)]^{1/2} \left| \begin{matrix} q_1-1 & & q_2 \\ & q_1-1 & \end{matrix} \right\rangle. \tag{28}$$

Consequently,

$$\mathfrak{L}_3^1 = [(q_1-q_2+1)(h_1-q_1+1)(q_1-h_3+1)]^{-1/2} L_3^1. \tag{29}$$

That is precisely the form of the $U(3)$ operator $\mathfrak{L}_3^1$ of Nagel and Moshinsky which follows from (16) and (19). The same conclusion concerning $\mathfrak{R}_2^3$, is proved similarly.

For the remaining operators $\mathfrak{L}_3^2$ and $\mathfrak{R}_1^3$ the proof is elementary. As in the previous example they are just the generators C_3^2 and C_1^3 multiplied by the normalization functions equal respectively to $(a_2^2)^{-1}$ and $(b_2^1)^{-1}$ in which (27) is inserted.

ACKNOWLEDGMENTS

I would like to thank Professor M. Moshinsky for turning my attention to the results of Ref. 1, for helpful discussions, and for reading of the manuscript. I am also grateful for the hospitality of the Aspen Center for Physics where this work was completed.

---

[1]J. G. Nagel and M. Moshinsky, J. Math. Phys. **6,** 682 (1965).
[2]N. Mukunda, L. O'Raifeartaigh, and E. C. G. Sudarshan, Phys. Rev. Lett. **15,** 1041 (1965); G. Cocho, C. Fronsdal, I. T. Grodsky, and R. White, Phys. Rev. **162,** 1662 (1967); C. S. Kalman, Can. J. Phys. **50,** 481 (1972).
[3]I. M. Gel'fand and M. I. Graev, Izv. Akad. Nauk SSSR Ser. Math. **29,** 1329 (1965) [Am. Math. Soc. Transl. **64,** 116 (1967) Ser. 2].
[4]L. C. Biedenharn, *Non-Compact Groups in Particle Physics,* edited by Ytze Chow (Benjamin, New York, 1966); R. M. Santilli, Nuovo Cimento **51,** 75 (1967); A. Chakrabarti, J. Math. Phys. **9,** 2087 (1968).
[5]U. Ottoson, Commun. Math. Phys. **10,** 114 (1968).
[6]N. Ya. Vilen'kin, *Special Functions and the Theory of Group Representations* (Amer. Math. Soc., Providence, R.I., 1968).
[7]It follows from the definition of the weight that a linear combination of patterns has a definite weight only if every pattern of the combination has the weight.

Algebraically special perturbations of the Schwarzschild metric*

W. E. Couch and E. T. Newman
Department of Physics, University of Pittsburgh, Pittsburgh, Pennsylvania 15213
(Received 25 September 1972)

Algebraically special perturbations of the Schwarzschild metric are found and expressed in a simple form. They become singular on the event horizon.

I. INTRODUCTION

The nature of gravitation collapse without spherical symmetry is a problem currently under widespread investigation. Price[1] and others have shown that with appropriate boundary conditions first-order perturbations of the Schwarzschild field with $l \geq 2$ [the perturbations are separable in spin-weighted spherical harmonics ${}_sY_{lm'}(\theta, \phi)$] decay to zero as the Schwarzschild time increases to infinity. Here we investigate perturbations of the Schwarzschild metric not by the *a priori* imposition of specific boundary conditions, but by restricting our considerations to algebraically special vacuum fields. We use the equation given by Lind[2] and Talbot[3] that are obeyed by fields of the algebraically special class to derive the equations governing the algebraically special perturbations of Schwarzschild. These special perturbations can easily be solved for and expressed in simple form. In the general case when the degenerate principal null congruence has nonvanishing twist, the special perturbations, for $l \geq 2$, have terms involving both exponential growth and decay and thus the solution does not decay back to Schwarzschild (at least in this linearized perturbation). This does not contradict the result of Price due, apparently, to the fact that the restriction to algebraically special fields violates Price's initial conditions.

II. THE PERTURBATIONS

The metric for algebraically special fields has the form[2]

$$ds^2 = 2(l_\mu dx^\mu)\left(dr - \mathrm{Re}\,\frac{\omega d\zeta}{\rho P} - U l_\nu dx^\nu\right) - \frac{d\zeta d\bar{\zeta}}{2\rho\bar{\rho}P^2}, \quad (2.1)$$

where $l_\mu dx^\mu = du + Re(Ld\zeta/P)$ and $\rho = -1/(r + i\Sigma)$. The coordinate r is an affine parameter for the degenerate principal null congruence with tangent vector l^μ; and ζ and $\bar{\zeta}$ are complex angular coordinates related to θ and ϕ by $\zeta = e^{i\phi}\cot\theta/2$. The divergence and twist of l^μ are given by the real and imaginary parts of ρ respectively. The real quantities U and $\Sigma(u, \zeta, \bar{\zeta})$ and the complex quantity ω are given in terms of the real function $P(u, \zeta, \bar{\zeta})$ and the complex auxiliary quantity $L(u, \zeta, \bar{\zeta})$ by

$$\begin{aligned}
\omega &= \omega^0\bar{\rho} + \dot{L} - L\,\dot{P}/P,\\
U &= U^0 + (\dot{P}/P)r - Re(\rho\psi_2^0),\\
2i\Sigma &= (\eth\bar{L} + L\dot{\bar{L}}) - (\text{c.c.}),\\
i\omega^0 &= (\eth\Sigma + L\dot{\Sigma} + 2\Sigma(\dot{L} - L\,\dot{P}/P),\\
-2U^0 &= (\eth N + L\dot{N} - \dot{P}/P\,LN) + (\text{c.c.}),\\
N &= \dot{\bar{L}} + \bar{\eth}\log P,
\end{aligned} \quad (2.2)$$

where the degree superscript denotes the absence of r dependence, the dot denotes $\partial/\partial u$, and $\eth$ and $\bar{\eth}$ are defined as operators on quantities of spin weight[4] s by

$$\eth\eta = 2P^{1-s}\frac{\partial}{\partial\zeta}(P^s\eta), \quad \bar{\eth}\eta = 2P^{1+s}\frac{\partial}{\partial\bar{\zeta}}(P^{-s}\eta); \quad (2.3)$$

$\eth\eta$ and $\bar{\eth}\eta$ have spin weights $s + 1$ and $s - 1$, respectively.

The spin weights of L, N, ψ_2^0, Σ, and P are $1, -1, 0, 0, 0$ respectively. The complex quantity ψ_2^0 and the quantities L and P satisfy the differential equations

$$\begin{aligned}
&\eth\psi_2^0 + L\dot{\psi}_2^0 + 3\dot{L}\,\psi_2^0 - 3L(\dot{P}/P)\psi_2^0 = 0,\\
&\dot{\psi}_2^0 - 3(\dot{P}/P)\psi_2^0 = \eth\psi_3^0 + L\dot{\psi}_3^0 + 2\dot{L}\psi_3^0 - 3L(\dot{P}/P)\psi_3^0,\\
&\psi_2^0 - \bar{\psi}_2^0 = -4i\,\Sigma U^0\\
&\qquad + 2i\,\mathrm{Re}[\eth W + L\dot{W} + \dot{L}W - 2L(\dot{P}/P)W],
\end{aligned} \quad (2.4)$$

where

$$\begin{aligned}
W &\equiv \bar{\eth}\Sigma + \bar{L}\dot{\Sigma} + \dot{\bar{L}}\Sigma - (\dot{P}/P)\bar{L}\Sigma,\\
\psi_3^0 &= \eth R + L\dot{R} - 2(\dot{P}/P)LR,\\
R &\equiv \bar{\eth}N + \bar{L}\dot{N} - (\dot{P}/P)N\bar{L} + N^2 - 2N\bar{\eth}\log P.
\end{aligned}$$

The Schwarzschild solution is given by

$$\psi_2^0 = 2\sqrt{2}m \equiv \mu, \quad P = P_0 \equiv \tfrac{1}{2}(1 + \zeta\bar{\zeta}), \quad L = \Sigma = 0, \quad (2.5a)$$

where m is the mass; and the metric takes the form

$$ds^2 = 2(1 - \mu/r)\,du^2 + 2dudr - r^2\,d\zeta d\bar{\zeta}/2P^2. \quad (2.5b)$$

We denote the Schwarzschild value of ψ_2^0 by μ.

We now assume that all quantities are given by their Schwarzschild value plus a small perturbation. In what follows instead of introducing additional notation for the perturbations, the symbol for each quantity will always mean that quantity's perturbation from its Schwarzschild value. When the perturbation expansion is put into Eqs. (2.4) and terms of second order (and higher) are dropped, we find the equations governing the perturbations,

$$\begin{aligned}
&\eth_0\psi_2^0 + 3\mu\dot{L} = 0,\\
&\psi_2^0 - 3\mu I = \eth_0\psi_3^0,\\
&\psi_2^0 - \bar{\psi}_2^0 = 2i(\eth_0\bar{\eth}_0 + 2)\Sigma,
\end{aligned} \quad (2.6)$$

where $I \equiv P/P_0$, $\psi_3^0 = \eth_0\bar{\eth}_0(\dot{\bar{L}} + \bar{\eth}_0 I)$, $2i\Sigma = \eth_0\bar{L} - \bar{\eth}_0 L$, $\eth_0$ is defined by $\eth_0\eta \equiv 2P_0^{1-s}(\partial/\partial\zeta)(P_0^s\eta)$, and similarly for $\bar{\eth}_0$.

By a coordinate transformation of the form $u' = f(u, \zeta, \bar{\zeta})$, $r' = rf^{-1}$, $\zeta' = \zeta$ [preserving the form of the metric given by Eq. (2.1)] we impose the condition that the "electric" part of L vanish (i.e., $\mathrm{Re}\bar{\eth}_0 L = 0$). This coordinate condition is the linearized version of the canonical coordinates of Aronson, Lind, Messmer, and Newman.[5]

With this condition, the most general solution to Eqs. (2.6) that is expandable in ${}_sY_{lm'}$, $l \geq 2$, is a linear combination with respect to l, of solutions of the form,

$$\begin{aligned}
\psi_2^0 &= e^{\alpha u}\sum_{m'=-l}^{l} c_{lm'}\,{}_0Y_{lm'},\\
I &= \frac{-1}{3\mu}\,e^{-\alpha u}\sum_{m'=-l}^{l} d_{lm'}\,{}_0Y_{lm'},
\end{aligned} \quad (2.7)$$

$$L = \frac{-1}{3\mu\alpha}\,\eth_0\psi_2^0,$$

$$\Sigma = \frac{i\psi_2^0}{(l-1)(l+2)},$$

where α is defined by $3\mu\alpha = (l-1)\,l(l+1)(l+2)$ and the constants $c_{lm'}$ and $d_{lm'}$ must satisfy $c_{lm'} = \bar{c}_{l,-m'}$, $d_{lm'} = \bar{d}_{l,-m'}$. The condition on $c_{lm'}$ means that $\mathrm{Re}\psi_2^0 = 0$, that is, the "electric" part of ψ_2^0 vanishes. The condition on $d_{lm'}$ merely reflects the fact that I is real. Note that the "electric" part of the field has $e^{-\alpha u}$ time dependence and the "magnetic" part has $e^{\alpha u}$ dependence. One might picture this behavior as arising from a body whose radial motion decays while its differential rotation grows.

The $1/r$ part of the Weyl tensor (i.e., the radiation field at future null infinity) is given by

$$\psi_4^0 = -\frac{\sqrt{(l-1)\,l(l+1)(l+2)}\,\alpha}{3\mu} \times \sum_{m'=-l}^{l} (c_{lm'}e^{\alpha u} + d_{lm'}e^{-\alpha u})_{-2}Y_{lm'}.$$

The radiation field (incoming) at past null infinity, to the approximation used here, vanishes.

Perturbations with $l = 0$ or $l = 1$ are, of course, also possible. The former is merely an addition to the mass and the latter is linearized Kerr solution. These are of no interest to us here.

From a comparison of the theory of algebraically special solutions[2,3] with the theory of characteristic data[6] for the gravitational field, it is obvious that the imposition of algebraic specialness on the field must involve giving part of the data in a specific fashion. One can easily show that our results imply that the data for perturbations of Schwarzschild that are required to be algebraically special (and twisting with $l \geq 2$) are necessarily singular on the event horizon $u = \infty$. Hence such perturbations cannot arise from data that differs from Schwarzschild data by a small amount.

Because of the exponential growth with increasing u present in the perturbation, the linearized perturbation analysis done here must become invalid as an approximation after some time. This exponential behavior therefore does not preclude the logical possibility that with the nonlinearities taken into account there may exist well-behaved algebraically special fields that are initially close to Schwarzschild, deviate far from it, and finally return to Schwarzschild (or Kerr).

A different possibility that can be imagined is that a nonsingular algebraically special solution initially close to Schwarzschild might correspond to a source whose small asymmetries cause it to bifurcate after some time.

We have also considered the algebraically special perturbations of the Einstein–Maxwell equations about the Reissner–Nordstrom solution. The results are similar to those just described. The perturbations expandable in ${}_sY_{lm'}$ have both $e^{\alpha' u}$ and $e^{-\alpha' u}$ dependence where $\alpha' = \alpha 2(\sqrt{1+\delta} - 1)/\delta$, $\delta \equiv 32\epsilon^2(l-1)(l+2)/9\mu^2$, ϵ is the charge.

*Research supported in part by the National Science Foundation.

[1]R. H. Price, "Nonspherical Perturbations of Relativistic Gravitational Collapse," preprint.
[2]R. W. Lind, Ph.D. thesis (University of Pittsburgh, 1970).
[3]C. J. Talbot, Commun. Math. Phys. **13**, 45 (1969).
[4]J. N. Goldberg, *et al.*, J. Math. Phys. **8**, 2155 (1967).
[5]B. Aronson, R. Lind, J. Messmer, and E. Newman, J. Math. Phys. **12**, 2462 (1971).
[6]R. K. Sachs, J. Math. Phys. **3**, 908 (1962).

Asymptotic estimation of Fourier transforms and light cone dominance

T. W. Chen
Department of Physics, New Mexico State University, Las Cruces, New Mexico 88001
(Received 6 April 1972; first revised received 21 July 1972; second revised received 3 October 1972)

The usual method of asymptotic estimation of Fourier transforms, the phase oscillation method, is examined. It is found that the usual argument of phase oscillation can be misleading due to either locations of singularities or absolute integrability of functions. As a by-product we have found that the scaling law for the inelastic e–p scattering should also be valid as the energy change $\nu\to\infty$ regardless of the momentum transferred squared κ, according to the usual approach of light cone dominance.

I. INTRODUCTION

One of the most frequently used tools in discussing extreme high energy interactions in recent years has been the phase oscillation argument. This is simply because we have to deal with Fourier integrals most of the time in physics. Fourier integrals contain oscillating exponential functions: And the phase oscillation argument which says that the most contribution comes from the region where the phase is bounded comes into play handily when high momentum cases are considered. In fact it enables us to conclude that the high momentum scattering processes can be realized through the behavior of currents near the light cone. Wilson's expansion[1] of current products near the light cone thus becomes a powerful suggestion in these processes. In this manner, many interesting physical consequences, including the celebrated scaling behavior, have been reported in the past two years by many authors.[1-3]

Unfortunately, the usual statement on the phase oscillation argument is not quite precise. Recently objections against this argument were raised; counter-examples were illustrated.[4] As this argument has been proven to be of great use, it is rather surprising to see that its justification has left much to be desired.

In this paper this method of asymptotic estimation of Fourier transforms is reexamined. The precise statement is searched for through a fundamental theorem, the Riemann–Lebesgue lemma. After brief introduction of the phase oscillation argument in Sec. II, it is pointed out in Sec. III that a one-dimensional Fourier integral $G(k)$ at $k\to\infty$ of a function $F(z)$ is mainly controlled by $F(z)$ near singular points, as it is clearly implied by the lemma. In fact, the usual phase oscillation argument makes sense only if the singularities of $F(z)$ are at $z=0$ and $F(z)$ with singularities subtracted away is absolutely integrable.

The two-dimensional case in which we are most interested is not so straightforward. In Sec. IV the Fourier integral $G(\alpha,\beta)$ of a function $F(u,v)$ is considered at the limit $\alpha\to\infty$ and $\beta\to\infty$ as well as the limit $\alpha\to\infty$ but β fixed. Again based on the Riemann-Lebesgue lemma, it is concluded that $G(\alpha,\beta)$ at $\alpha\to\infty$ *and* $\beta\to\infty$ is controlled not by the origin of the coordinate system but $F(u,v)$ near singular regions. More interestingly it is shown that the same conclusion still follows at $\alpha\to\infty$ *but* β fixed. It is then demonstrated in Sec. V that the usual conclusion of light cone dominance through the phase oscillation argument is valid only if the singularities in the configuration space are only on the light cone and a condition of absolute integrability is satisfied. The result that the same asymptotic behavior should be observed at $\alpha\to\infty$ and β fixed is rather significant; it implies, among others, that *the scaling behavior of structure functions* $W(\kappa,\nu)$ *should exist as long as* $\nu\to\infty$ *regardless of* κ. Some of the immediate impacts of this result are discussed.

Finally in Sec. VI comments regarding the phase oscillation method are made. The controversial counter example raised by Sucher and Woo[4] is also examined.

II. PHASE OSCILLATION APPROACH

In analyzing high energy behaviors of the inelastic e–p scattering process, we must consider the matrix element of current (for simplicity scalar) commutators between single-nucleon states, spin averaged

$$W(\kappa,\nu)=\int d^4x\, e^{-iqx}\,\langle p|[j(x),j(o)]|p\rangle, \tag{1}$$

where $\kappa=q^2$ and $\nu=-pq/M$.

In recent literature[2,3] it is customary to conclude through phase oscillation argument that the major contribution to the function W at $\nu\to\infty$ and $\omega=2M\nu/\kappa$ fixed, comes from the current commutator near the light cone. The argument goes as follows. In the rest frame where $p=(\mathbf{0},M)$ and $q=(0,0,\sqrt{\kappa+\nu^2},\nu)$, the oscillating function at $\nu\to\infty,\omega$ fixed is given by

$$e^{-iqx}\simeq e^{-i[\nu(z-t)+(M/\omega)z]}. \tag{2}$$

Since most of the contribution to (1) comes from the region where the phase is bounded, we see that the function W is dominated by the current commutator in the region where

$$|z-t|<1/\nu \tag{3}$$

and

$$|z|<\omega/M, \tag{4}$$

thus, for time-like x,

$$0\le -x^2=(t-z)(t+z)-x^2-y^2\lesssim\omega/M\nu\to 0. \tag{5}$$

This region, therefore, is near the light cone.

The above argument is indeed not a proof. Recently several objections have been raised, using counter examples.[4] In fact, the above argument can be true only if $F(x)$ is subject to certain properties as we shall see later.

Since interesting physical consequences follow from the light cone dominance argument,[2,3] it is important to clear up the above confusion. There are two crucial points we must investigate. Firstly, more rigorous statement and proof ought to be searched. Secondly, since we are dealing with two-dimensional F.t. (Fourier transform) $W(\kappa,\nu)$ at $\nu\to\infty$ with $\kappa+\infty$ or fixed, the correlation between these two limiting processes must be studied carefully. We shall first study the first point in the next section.

III. ASYMPTOTIC ESTIMATION OF F.t.'S (FOURIER TRANSFORMS) IN ONE DIMENSION

Consider an F.t. $G(k)$ of a generalized function $F(z)$ in one dimension. The behavior of $G(k)$ at $k \to \infty$ can be best analyzed by the following theorem, a generalized version of the Riemann-Lebesgue lemma.[5]

Theorem: If the generalized function $F(z)$ has a finite number of singularities $z = z_1, z_2, \ldots, z_M$, and if (for each m from 1 to M) $F(z) - f_m(z)$ has absolutely integrable Nth derivative in an interval including z_m, where $f_m(z)$ is a linear combination of functions of the type[6]

$$|z - z_m|^\beta, \quad |z - z_m|^\beta \epsilon(z - z_m), \quad |z - z_m|^\beta \log|z - z_m|,$$
$$|z - z_m|^\beta \log|z - z_m| \ \epsilon(z - z_m)$$

and $\delta^{(p)}(z - z_m)$, for different values of β and p, and if $F^{(N)}(z)$ is well behaved at infinity, the F.t. $G(k)$ satisfies

$$G(k) = \sum_{m=1}^{M} g_m(k) + o(|k|^{-N}) \quad \text{as } |k| \to \infty,$$

where $g_m(k)$ is the F.t. of $f_m(z)$.

For our present purpose it is enough to consider $F(z)$ of the above theorem with singularities only at $z = 0$. The asymptotic behavior of its F.t. is then

$$G(k) = g_0(k) + o(|k|^{-N}), \quad \text{as } |k| \to \infty, \tag{6}$$

where $g_0(k)$ is the F.t. of functions which are singular at $z = 0$.

We can now further see that the result in (6) is mainly controlled by the behavior of $F(z)$ in $|z| \lesssim 1/|k|$. This can be most easily seen if one notes that

$$G(k) = \tfrac{1}{2}\left(\int_{-\infty}^{-z_0} + \int_{-z_0}^{z_0} + \int_{z_0}^{\infty}\right) dz\ e^{ikz}\left[F(z) - F\left(z + \frac{\pi}{k}\right)\right], \tag{7}$$

where $z_0 \lesssim \pi/|k|$. Since $F(z)$ is absolutely integrable in $(-\infty, -z_0)$, it follows that

$$\int_{-\infty}^{-z_0} dz\ e^{ikz}\left[F(z) - F\left(z + \frac{\pi}{k}\right)\right] \leq \int_{-\infty}^{-z_0} dz \left| F(z) - F\left(z + \frac{\pi}{k}\right)\right| \to 0, \quad |k| \to \infty$$

and the first integral in (7) vanishes as $|k| \to \infty$. The same is true for the third integral. Thus, $G(k)$ is mainly determined by the second integral in (7), that is, by the behavior of $F(z)$ in the interval where

$$|z| \lesssim 1/|k| \to 0, \tag{8}$$

where "$\lesssim$" means "smaller or of the order of."

In closing this section, the following comments are in order: (1) Customarily we say, by looking at the oscillating exponential function e^{ikz} in an F.t., that the most contribution to the F.t. comes from $|z| \lesssim 1/|k|$ at $|k| \to \infty$. This can only be true if singularities are at $z = 0$ only. It should be clear from above that the major contribution comes from the regions where $|z - z_m| \lesssim 1/|k|, m = 1, \ldots, M$, if singularities are at $z = z_1, \ldots, z_M$. (2) Contributions of singularities of $F(z)$ to asymptotic behavior of $G(k)$ are indeed most important. However, the contribution of the nonsingular region of $F(z)$ should not be totally ignored if we are interested in contributions from higher orders in $1/|k|$. Its effect to $G(k)$ at $|k| \to \infty$ depends on the existence of absolutely integrable derivatives in an interval, as is clearly stated in the theorem. In all we see that the precise statement regarding the usual phase oscillation argument is essentially in the Riemann–Lebesgue lemma.

IV. ASYMPTOTIC ESTIMATION OF F.t.'S IN TWO DIMENSIONS

Consider now the F.t. of a function $F(u, v)$

$$G(\alpha, \beta) = \int e^{i(\alpha u - \beta v)} F(u, v)\, du\, dv. \tag{9}$$

We are interested in the asymptotic behavior of $G(\alpha, \beta)$ at the limit: $\alpha \to \infty$ and $\beta \to \infty$ as well as at $\alpha \to \infty$ but β fixed. For our purpose, it is enough to consider $F(u, v)$ which has singularities only on lines $u = 0$ and $v = 0$. We shall assume that $F(u, v)$ is absolute integrable in the region away from singularities and also that $\alpha \geq \beta \geq 0$.

The two-dimensional case is a straightforward generalization of the one-dimensional case. We can again show that the major contribution to asymptotic $G(\alpha, \beta)$ comes from $F(u, v)$ in the neighborhood of singular lines. We shall show this at the two limits (1) $\alpha \to \infty$ and $\beta \to \infty$ and (2) $\alpha \to \infty$ but β fixed separately.

(1) At $\alpha \to \infty$ and $\beta \to \infty$. To do this we write (9) as

$$G(\alpha, \beta) = \int dv\ e^{-i\beta v}\, g(v, \alpha), \tag{10}$$

where

$$g(v, \alpha) = \int du\ e^{i\alpha u}\, F(u, v). \tag{11}$$

Thus for some fixed v, $g(v, \alpha)$ is an F.t. in one dimension. Taking into account that, for fixed v, $F(u, v)$ has singularities only at $u = 0$, we can take over the result of the preceding section and conclude that the major contribution to $g(v, \alpha)$ at $\alpha \to \infty$ comes from $F(u, v)$ in the region where, for some v,

$$|u| \lesssim 1/\alpha. \tag{12}$$

Now $G(\alpha, \beta)$ at $\alpha \to \infty$ is again an F.t. of one dimension of variable β. Since the only singularities in $g(v, \alpha)$ are at $v = 0$, we conclude in exact analogy to the above argument that the major contribution to $G(\alpha \to \infty, \beta)$ at $\beta \to \infty$ comes from $g(v, \alpha)$ in the region where

$$|v| \lesssim 1/\beta. \tag{13}$$

Combining (12) and (13), we have finally that most of $G(\alpha, \beta)$ at $\alpha \to \infty, \beta \to \infty$ comes from $F(u, v)$ in the region where

$$|uv| \lesssim 1/\alpha\beta, \tag{14}$$

which indeed implies that the important region is the neighborhood of singular lines where $|uv| \sim 0$.

The above result is quite straightforward. However, let us note that the above proof cannot go through if not both α and β go to ∞. Next we shall consider just the case when $\alpha \to \infty$ but β fixed.

(2) At $\alpha \to \infty$ but β fixed. For this case, we shall consider

(9) in a rotated frame where

$$u' = \cos\theta\, u + \sin\theta\, v, \quad v' = -\sin\theta\, u + \cos\theta\, v, \tag{15}$$

where θ is some fixed angle, $\frac{1}{2}\pi > \theta > 0$. (9) becomes

$$G(\alpha,\beta) = \int e^{i(\alpha' u' - \beta' v')} F(u', v')\, du'\, dv', \tag{16}$$

where

$$\alpha' = \alpha\cos\theta - \beta\sin\theta, \quad \beta' = \alpha\sin\theta + \beta\cos\theta.$$

It should be noted that the singularities of the function F in the new coordinate system are on those two lines:

$$u' = \tan\theta\, v' \tag{17}$$

and

$$u' = -\cot\theta\, v'. \tag{18}$$

Furthermore, the limit $\alpha \to \infty, \beta$ fixed, says that $\alpha' \to \infty$ *and* $\beta' \to \infty$ in this rotated frame. We can now proceed in analogy to the case (1) by writing (16) as

$$G(\alpha', \beta') = \int dv'\ e^{-i\beta' v'}\ g(v', \alpha'), \tag{19}$$

where

$$g(v', \alpha') = \int du'\ e^{i\alpha' u'} F(u', v'). \tag{20}$$

Since $g(v', \alpha')$ for fixed v' is an F.t. of one dimension and $F(u', v')$ for fixed v' has singularities only at $u' = \tan\theta\ v'$ and $u' = \cot\theta\ v'$, we must conclude that the major contribution to $g(v', \alpha')$ at $\alpha' \to \infty$ comes from $F(u', v')$ in the region where

$$|u' - \tan\theta\, v'| \lesssim 1/|\alpha'| \tag{21}$$

and the region where

$$|u' + \cot\theta\, v'| \lesssim 1/|\alpha'|. \tag{22}$$

Again $G(\alpha', \beta')$ at $\alpha' \to \infty$ is an F.t. in one dimension of the variable β. The singularities of $g(v'; \alpha')$ are now only at $v' = 0$. Thus, we conclude that the major contribution to $G(\alpha' \to \infty, \beta')$ at $\beta' \to \infty$ comes from $g(v'; \alpha' \to \beta)$ in the region where

$$|v'| \lesssim 1/|\beta'|. \tag{23}$$

Therefore, at $\alpha \to \infty, \beta$ fixed, there are two important regions: One is specified by (21) and (23) and the other by (22) and (23). In terms of original coordinates system this means the following two regions:

$$\text{(A)}\quad |u| \lesssim \frac{1}{\alpha}, \ |-\sin\theta\ u + \cos\theta\ v| \lesssim \frac{1}{\alpha\sin\theta} \tag{24}$$

and

$$\text{(B)}\quad |v| \lesssim \frac{\sin\theta}{\alpha\cos\theta}, \ |-\sin\theta\ u + \cos\theta\ v| \lesssim \frac{1}{\alpha\sin\theta}. \tag{25}$$

It is not difficult to see that in both regions,

$$|uv| \to 0 \text{ as } \alpha \to \infty, \quad \text{for } \tfrac{1}{2}\pi > \theta > 0.$$

We thus have shown that at $\alpha \to \infty, \beta$ fixed, the major contribution to $G(\alpha, \beta)$ comes from $F(u, v)$ in the neighborhood of the assumed singular lines. In fact, we have shown that this is the case as long as $\alpha \to \infty$ independent of the other variable β.

V. LIGHT CONE DOMINANCE

We are now ready to see some physical consequences out of the mathematical results of preceding sections. Let us go back to the structure function given in Sec. II, viz.

$$W(\kappa, \nu) = \int d^4x\ e^{-iqx} F(x),$$

where

$$F(x) = \langle p|[j(x), j(o)]|p\rangle,$$

which in the rest frame of p is given by

$$W(\kappa, \nu) = \int dz\ dt\ e^{-i(q_3 z - q_4 t)} F(z, t), \tag{26}$$

where

$$q_3 = \sqrt{\kappa + \nu^2}, \quad q_4 = \nu,$$

and

$$F(z, t) = \int d\mathbf{x}\, dy\, F(x).$$

The structure function is therefore expressed as an F.t. in two dimensions. In the infinite momentum frame it becomes

$$W(\kappa, \nu) = \int du\ dv\ e^{i(\alpha u - \beta v)} F(u, v), \tag{27}$$

where

$$u = -\tfrac{1}{2}\sqrt{2}(z - t), \quad v = \tfrac{1}{2}\sqrt{2}(z + t),$$

and

$$\alpha = \tfrac{1}{2}\sqrt{2}[(\kappa + \nu^2)^{1/2} + \nu] \quad \text{and} \quad \beta = \tfrac{1}{2}\sqrt{2}[(\kappa + \nu^2)^{1/2} - \nu].$$

Our purpose here is to examine the asymptotic behavior of $W(\kappa, \nu)$ at $\nu \to \infty$ with κ or ω fixed. From the result of Sec. IV, this asymptotic behavior must be mainly determined by singularities of $F(u, v)$. In physical cases it is natural to expect that the singularities of current commutators are only on the light cones. Thus, the singularities of $F(u, v)$ should be only on the lines $u = 0$ and $v = 0$. The results of the preceding section immediately imply that $W(\kappa, \nu)$ at $\nu \to \infty$ is mainly determined by $F(u, v)$ in the region where

$$|uv| \sim 0, \tag{28}$$

if $F(u, v)$ is absolutely integrable in the region away from singularities. Note that the above statement is true regardless of κ. It should be clear that $\beta \to M/\omega$ if $\nu \to \infty$ and ω fixed, whereas $\beta \to 0$ if $\nu \to \infty$ but κ fixed. Nevertheless, since $\alpha \to \infty$ at $\nu \to \infty$ with ω fixed (i.e., the Bjorken limit) or at $\nu \to \infty$ and κ fixed (i.e., the Regge limit), the above conclusion follows. Taking (28) together with the causal property, we can easily come to the conclusion that the major contribution to $W(\kappa, \nu)$ at $\nu \to \infty$ comes from $F(x)$ near the light cone where $x^2 \sim 0$.[2,3]

It should be noted that the above conclusion follows only for $F(x)$ satisfying the conditions given in the stated Riemann-Lebesgue lemma. In particular, it should be emphasized that $F(x)$ must be absolutely integrable.

It is well known that the light cone expansion of $F(x)$ naturally gives rise to the experimentally conformed phenomenon of scaling behavior.[2,3] This is indeed a direct result of the above conclusion that the asymptotic behavior of $W(\kappa, \nu)$ is mainly determined by $F(x)$ near

the light cone. The result of our careful study, in fact, is much more than the usual scaling law which says that $W(\kappa, \nu)$ becomes a function of ω only at the Bjorken limit (i.e., $\nu \to \infty$ and ω fixed); our study shows that this should be valid *as long as* $\nu \to \infty$ *regardless of* κ.[7] This has rather strong physical implications. One of the immediate ones is that Bjorken's scaling behavior can now be compared with the Regge asymptotic behavior at large ν but any κ. Another interesting one is that at $\nu \to \infty$ $W(\kappa, \nu)$ for any κ is actually given by $W(\kappa, \nu)$ evaluated at $\kappa = 0$, $\nu \to \infty$.[8] All of them, thus, put the usual approach of light cone dominance[2,3] to a severe test. We shall state these and other physical applications of our results elsewhere in order that we can confine ourselves only to the mathematical aspect of the asymptotic properties of F.t.'s in this paper.

VI. DISCUSSION

In summary we have shown that only the neighborhoods of singularities of a function $F(u, v)$ are significant to its F.t. $G(\alpha, \beta)$ at $\alpha \to \infty$. Consequently it makes sense to consider Taylor's expansions of $F(u, v)$ around singularities in order to estimate $G(\alpha, \beta)$ at $\alpha \to \infty$. A good application of this result is to make light cone expansions of current commutators in order to estimate the asymptotic behavior of the structure functions for the inelastic e–p scattering.[2] We have concluded, as a result, that $W(\kappa, \nu)$ of (1) yields the scaling law at $\nu \to \infty$ *regardless of* κ provided that the associated function $F(x)$ in configuration space has finite number of singularities and satisfies the appropriate condition of absolute integrability as precisely stated in the Riemann-Lebesque lemma.

The following observations should be worth noting:

(1) As the asymptotic behavior of an F.t. is essentially decided at singularities, the usual phase oscillation argument can be misleading. The conclusion of the phase oscillating argument has nothing to do with the origins of coordinate systems. In fact, it follows only when there are singularities at origins (or axes in multi-dimensional cases) and nowhere else. The asymptotic behavior of the F.t. of a function cannot be estimated at all without some knowledge of the function in the configuration space.

(2) According to the theorem in Sec. III, the contributions to F.t. $G(k)$ at $k \to \infty$ come from two parts: (a) functions around singularities and (b) absolutely integrable functions obtained after subtracting away appropriate singular functions. The second contribution is of $o(|k|^{-N})$ if the functions possess absolutely integrable Nth derivatives. Thus, the usual argument that leading singularities dominate Fourier integrals in asymptotic regions can also be misleading; in order that there are no (or small) contributions from the nonsingular functions, it is necessary that they have absolutely integrable derivatives of order N which might be high enough so that $o(|k|^{-N})$ is small compared with the contributions due to the singular functions.

As an example, consider the F.t. (26) with $F(x)$ given by the function of Sucher and Woo,[4] viz.

$$F(x) = (x^2 - i\epsilon)^{-1} u\,(x\cdot p) + e^{-ip\cdot x} x^2 \Delta_F(x^2, m^2), \quad (29)$$

where $u(\lambda) \to$ constant as $\lambda \to \infty$. The first term is singular. Nevertheless, as is pointed out by Sucher *and* Woo, the contribution to $G(\kappa, \nu)$ due to the first term does not necessarily dominate at $\nu \to \infty$ *and* ω fixed. In fact, it does not if $\nu \to \infty$ and $\omega = 1$. That this is not surprising can be seen if we note that the second term, even though continuous across the light cone, is not absolutely integrable [the exponential phase factor $\exp(-ipx)$ is really irrelevant]. The second term can, therefore, even make larger contribution than the singular term.

(3) As the phase oscillation argument is based on the Riemann-Lebesgue lemma, the conditions on singularities as well as the condition of absolute integrability should not be overlooked. Unfortunately these conditions are merely sufficient, not necessary, conditions. Consequently, if the condition of absolute integrability is not satisfied, the phase oscillation argument is simply not applicable even though the result in some situations may still be consistent with it. The above example illustrates this point clearly: $F(x)$ of (29) does give the consistent result as the phase oscillation argument in the Regge limit as well as in the Bjorken limit for $\omega > 1$, but not so if $\omega = 1$.

To illustrate this point further, let us consider Eq. (29) without the exponential factor $\exp(-ipx)$. Again the second term is not absolutely integrable. The Riemann-Lebesgue lemma again fails to work. Nevertheless, the second term vanishes in the Bjorken limit, but not in the Regge limit.

(4) It is well known that the infinite momentum frame has many advantages in dealing with extreme high energy interactions. However, it is interesting to note, that if we stay in the infinite momentum frame (i.e., to take $\theta = 0$ in Sec. IV), we can only come to the conclusion of the light cone dominance at the Bjorken limit where $\alpha\beta = \kappa \to \infty$, but not at the Regge limit where $\alpha\beta = \kappa$ fixed. In coming to the same conclusion at the Regge limit we must go off the infinite momentum frame as is shown in Sec. IV. This shortcoming of the infinite momentum frame can be avoided if we use the almost infinite momentum frame[9] which is essentially the particular case of rotated frame (u', v') in Sec. IV when θ is infinitesimally small.

(5) As experimental data strongly suggest the scaling behavior even at low κ,[10] our result that the region near the light cone dominates at $\nu \to \infty$ regardless of κ appears very interesting indeed. Impact of this new consequence to various physical processes is under investigation.

---

[1] K. G. Wilson, Phys. Rev. **179**, 1499 (1969). For a quick introductory review, see R. Jackiw, Phys. Today **25**, No. 1, 23 (1972); see also H. Fritzsch and M. Gell-Mann, talk presented at the Coral Gables Conference, Univ. of Miami, Coral Cables, Florida (1971).

[2] R. A. Brandt, Phys. Rev. Lett. **23**, 1260 (1969); Phys. Rev. D **1**, 2808 (1970); Y. Frishman, Phys. Rev. Lett. **25**, 966 (1970); J. M. Cornwall and R. Jackiw, Phys. Rev. D **4**, 367 (1971).

[3] T. W. Chen, Phys. Rev. D **3**, 1989 (1971); Phys. Rev. D **3**, 2257 (1971).

[4] J. Sucher and C. H. Woo, Phys. Rev. Lett. **27**, 696 (1971).

[5] The theorem stated below is exactly the one in M. J. Lighthill, *Fourier Analysis and Generalized Functions* (Cambridge U. P., Cambridge, England, 1964), p. 52. Proof of the theorem can be found there.

[6] Note that this type of singularity is enough to take care of most of the physical cases.

[7] Our conclusion is based on rather general principles; however, this is not exactly new; by direct calculation using some spectral function representations, the same conclusion has been reached by R. Jackiw, R. Van Royen, and G. B. West, Phys. Rev. D **2**, 2473 (1970). See also R. A. Brandt, second paper in Ref. 2, and J. M. Cornwall, D. Corrigan, and R. E. Norton, Phys. Rev. D **3**, 536 (1971).

[8]As is pointed out by papers in Ref. 7, this is not inconsistent with Regge pole theory.

[9]Applications of the almost infinite momentum frame are discussed in Ref. 3 based on the usual phase oscillation argument.

[10]E. D. Bloom *et al.*, Phys. Rev. Lett. **23**, 930 (1969); M. Breidenbach *et al.*, Phys. Rev. Lett. **23**, 935 (1969); E. D. Bloom *et al.*, Kiev report, SLAC PUB 796 (1970); T. J. Braunstein, *et al.*, SLAC PUB 858 (1971); J. Moritz *et al.*. DESY PUB 71/61 (1971).

Introduction to nonstandard analysis

A. Voros*

Service de Physique Théorique, Centre d'Etudes Nucléaires de Saclay, BP n° 2-9 1 190, Gif-sur-Yvette, France

(Received 7 April 1972; revised manuscript received 15 September 1972)

Nonstandard analysis is a recent branch of mathematics in which usual notions about analysis and topology can be formulated in an attractive and condensed manner. The main feature of this theory is that it introduces the concept of infinitely large or small numbers and that it allows one to compute with them in exactly the same way as in ordinary analysis. We believe that in the long run this new language might be usefully applied to physics. However, we only present a few examples of a rather trivial nature, our purpose being to give a short introduction of the subject to physicists. Thus we omit most of the technical subtleties especially those of a metamathematical nature, and we concentrate our attention on the practical side of the theory.

Nonstandard analysis is a recent branch of mathematics in which usual notions about analysis and topology can be formulated in an attractive and condensed manner. We believe that in the long range this new language might be usefully applied to physics. This article will show a few examples; but its main purpose is to introduce the subject briefly to physicists. Thus, we omit some technical subtleties and many proofs, but we must assume that the reader knows the basic facts and notations of set theory, algebra,[1] and topology.[2] For more precise physical applications, see Ref. 3.

1. THE NONSTANDARD REAL NUMBERS

The 17th century analysts thought of infinitesimals as of numbers smaller (in absolute value) than any positive real number, yet reproducing all the properties of the real numbers on an infinitely reduced scale. Although this intuitive view was useful to create analysis, it was later discarded as inconsistent by Cauchy. The theory presented here, the work of Professor Robinson,[4] provides a coherent background to the intuitive approach.[5]

We begin by constructing an *enlargement* of the set of real numbers $\mathbb{R}$, i.e., a set $^*\mathbb{R}$ containing $\mathbb{R}$ and also quantities interpreted as infinitely small or large numbers; moreover $^*\mathbb{R}$ satisfies all algebraic and order related properties of $\mathbb{R}$ (except the Archimedian property). In this framework, ordinary real numbers are called *standard* (in short: S), while the "generalized" numbers in $^*\mathbb{R}$ are called *nonstandard* (NS). Thus the qualifier "nonstandard" should *not be* understood as the negation of "standard" but as its generalization. We also call $^*\mathbb{R}$ the nonstandard extension of $\mathbf{R}$. We want to stress that there is *no simple relation* at all between this $^*\mathbf{R}$ (of algebraic origin) and the various compactifications of $\mathbb{R}$ that also introduce "points of infinity" by topological methods.

A. Construction of the nonstandard real field $^*\mathbb{R}$

We shall start from Cauchy's conception of infinitely small or large quantities as a means of representing the limiting properties of real-valued functions. The simplest example of such a problem is the study of sequences $\{U_n\}_{n\in\mathbb{N}}$ as $n\to\infty$, and it turns out that, for the practical purpose of constructing $^*\mathbb{R}$, nothing else is necessary.

Let $\mathcal{E}$ be the set of sequences of real numbers, turned into a commutative ring by the operations of termwise addition and multiplication. If $^*\mathbb{R}$ exists, then with any $\{U_n\}\in\mathcal{E}$, we want to associate a nonstandard (NS) number denoted $[U_n]$ ($[U_n]\in{}^*\mathbb{R}$) that will, in some sense, describe $\{U_n\}$ asymptotically. We impose the following rules:

(1) $^*\mathbb{R}$ is a *commutative ring*, and the mapping $\{U_n\}\xrightarrow{\sigma}[U_n]$ is a ring homomorphism from $\mathcal{E}$ onto $^*\mathbf{R}$.

(2) $\mathbb{R}$ is contained in $^*\mathbf{R}$, as a *subring*, and the number $a\in\mathbb{R}$ is the image by σ of the constant sequence $\{U_n = a\}$.

(3) If for some n_0: $(n > n_0 \Rightarrow U_n = V_n)$ then $[U_n] = [V_n]$.

The following method, called the *ultrapower construction*, gives a concrete realization of $^*\mathbb{R}$ that has the extra property of being a *totally* (= *linearly*) *ordered field*, like $\mathbb{R}$.

Choose once for all a free ultrafilter[6] $\mathfrak{U}$ on $\mathbb{N}$ (see Appendix) and define the relation for $\{U_n\},\{V_n\}\in\mathcal{E}$:

$$U_n \overset{\text{a.e.}}{=\!=} (\text{"almost everywhere"}) \Leftrightarrow \{n\,|\,U_n = V_n\}\in\mathfrak{U} \Leftrightarrow \mu\{n\,|\,U_n\neq V_n\} = 0.$$

From measure theory it follows that $(\overset{\text{a.e.}}{=\!=})$ is an equivalence relation and the quotient set $\mathcal{E}/\overset{\text{a.e.}}{=\!=}$ is a ring. We define

$$^*\mathbf{R} = \mathcal{E}/\overset{\text{a.e.}}{=\!=} \text{ and } [U_n] = \sigma\{U_n\} = \text{equivalence class of } \{U_n\}.$$

This just means that $[U_n]$ and $[V_n]$ coincide iff $U_n \overset{\text{a.e.}}{=\!=} V_n$. This definition satisfies rules (1) to (3).

Moreover, $^*\mathbb{R}$ is a *field:* every $a\in{}^*\mathbb{R}, a\neq 0$ has an inverse.

Proof: $a = [U_n]$ for some $\{U_n\}\in\mathcal{E}$, and $a\neq 0 \Leftrightarrow \{n\,|\,U_n = 0\}\notin\mathfrak{U} \Leftrightarrow$ [property (D) of the Appendix is essential here] $\{n\,|\,U_n\neq 0\}\in\mathfrak{U}$—so that $\{1/U_n\}$ can be defined a.e. and $1/a = [1/U_n]$ because $U_n\times 1/U_n \overset{\text{a.e.}}{=\!=} 1$. Also $^*\mathbb{R}$ can be *ordered* by the relation

$$[U_n]\leq[V_n] \Leftrightarrow U_n \text{a.e.} \leq V_n.$$

Again, property (D) is crucial to prove that the order is *total*, i.e., that any two $a,b\in{}^*\mathbb{R}$ can be compared ($a\leq b$ or $b\leq a$).

B. Structure and properties of $^*\mathbb{R}$

We just saw that $^*\mathbb{R}$ is a totally ordered field, like $\mathbf{R}$. Besides, $\mathbb{R}$ is a subset of $^*\mathbb{R}$, and its ordered field structure is the natural restriction of that of $^*\mathbb{R}$. *All algebraic rules on $\mathbb{R}$ remain true on $^*\mathbb{R}$*, and will be written the same way. For instance, "absolute value" will also denote the function on $^*\mathbb{R}$:

$$|a| = a \text{ if } 0\leq a, \quad |a| = -a \text{ otherwise.}$$

Definitions: $a\in{}^*\mathbb{R}$ is *standard* $\Leftrightarrow a\in\mathbb{R}$; $a\in{}^*\mathbf{R}$ is *finite* $\Leftrightarrow \exists\, b\in\mathbb{R}: |a| < b$.

The set of finite numbers forms a ring noted M_0, or $\mathcal{O}(1)$. Otherwise, $a \in {}^*\mathbb{R}$ is *infinite* $\Longleftrightarrow \forall\, b \in \mathbb{R}$: $|a| > b$. For instance $[\lambda \log n], [\lambda n^\alpha], [\lambda e^{\alpha n^\beta}]$ (where λ, α, β are standard; $\lambda \neq 0; \alpha, \beta > 0$) are infinite numbers:

$a \in {}^*\mathbb{R}$ is *infinitesimal* $\Longleftrightarrow \forall\, b \in \mathbb{R}_+ - \{0\}$: $0 \leq |a| < b$.

For instance, $[n.\log n.\log(\log n)\ldots(\log_p n)]^{-1}$ (for λ, standard and p integer) is a well-known scale of infinitesimal numbers.

The set of infinitesimals forms a ring noted M_1 (or $o(1)$)]. We have the inclusions: $M_1 \subset M_0 \subset {}^*\mathbb{R}$; $\mathbb{R} \subset M_0 \subset {}^*\mathbf{R}$.

$a, b \in {}^*\mathbb{R}$ are *infinitely close* $\Longleftrightarrow a - b \in M_1$,

this relation will be noted $a \sim b$.

For any S-number $x \in \mathbb{R}$, we define a subset of ${}^*\mathbb{R}$, called the *monad of x*: $\mu(x) = \{y \in {}^*\mathbb{R} \mid y \sim x\}$.

Theorem: Every finite number $x \in M_0$ lies in the monad of one (and only one) S-number ${}^0x = st(x) \in R$, called the *standard part* of x. The mapping (st) is a ring homomorphism with kernel M_1. For Proof see Ref. 4, p. 56.

Intuitively the structure of ${}^*\mathbb{R}$ is shown in *Fig. 1:* Starting from $\mathbb{R}$, around every S-number we add the cluster of its infinitely close NS numbers (its monad). This way we get *all* the finite numbers. Then, on the negative and positive side, we add the infinite numbers. Here are now a few rules for the calculus on ${}^*\mathbb{R}$:

Infinitesimals: If $x, y \in M_1$ and $z \in M_0$: $x + y \in M_1$ and $xz \in M_1$. (M_1 is an ideal of the ring M_0.)

If x and y are S numbers and $x < y$, then $\mu(x)$ and $\mu(y)$ are disjoint, with any number in $\mu(x)$ smaller than any number in $\mu(y)$.

Infinite numbers: (1) If ω is infinite and $a \in \mathbb{R}$: $(a + \omega)$ and $(a \times \omega)$ are infinite (except for $0 \times \omega = 0$). This is a precise formulation of the well-known "rules": "$a + \infty = \infty$" and "$a \times \infty = \infty$", but the reader should realize that, in general, $a + \omega \neq \omega$ and $a \times \omega \neq \omega$, because the rules of algebra remain valid. (2) If ω and ω' are infinite positive, $(\omega + \omega')$ and $(\omega\omega')$ are infinite positive. But $(\omega - \omega')$ can *a priori* lie anywhere in ${}^*\mathbb{R}$: This is a precise interpretation of the indetermination of "$\infty - \infty$"; however, in the ${}^*\mathbb{R}$ calculus, $(\omega - \omega')$ has a value that is perfectly determined (by the values of ω and ω'), and its computation corresponds to lifting the indetermination in standard calculus. Other rules can be worked out by the reader; they always extend the rules valid on $\mathbb{R}$.

C. Enlargements related to ${}^*\mathbb{R}$

We can perform the same ultrapower construction as above with sequences $\{U_n\}_{n\in\mathbb{N}}$ where the condition $U_n \in \mathbb{R}$ is changed.

(1) If $U_n \in \mathbb{C}$, we obtain the NS complex field ${}^*\mathbb{C}$ that has the same relation to ${}^*\mathbb{R}$ as $\mathbb{C}$ has to $\mathbb{R}$.

(2) If $U_n \in \mathbb{R}^p$ (p *finite*, fixed), we obtain the NS vector space ${}^*(\mathbb{R}^p)$ over the field ${}^*\mathbb{R}$. Obviously, it is the direct sum of p times the space ${}^*\mathbb{R}$, i.e. ${}^*(\mathbb{R}^p) = ({}^*\mathbb{R})^p$.

(3) If $U_n \in \mathbb{Z}$ (set of integers), we obtain the set of NS integers, ${}^*\mathbb{Z}$, obviously contained in ${}^*\mathbb{R}$. The structure

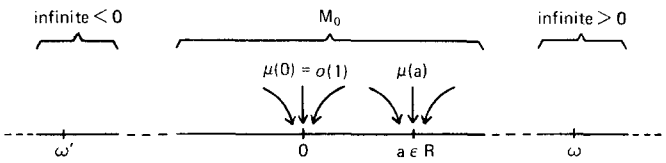

FIG. 1. The nonstandard real line.

of ${}^*\mathbb{Z}$ is shown in *Fig. 2:* The set of finite integers $({}^*\mathbb{Z} \cap M_0)$ coincides with the standard set $\mathbb{Z}$ because no infinitesimal integers exist. The remaining numbers of ${}^*\mathbb{Z}$ are infinite (positive or negative); they are also called $*$-*finite* to recall that, in the NS frame, they have all the properties of finite integers. The set ${}^*\mathbb{Z}$ is uncountable.

2. ENLARGEMENTS IN SET THEORY

For any set E that can be defined according to Axiomatic Set theory, an enlarged set ${}^*E$ can be found by the ultrapower construction, and the rules of formal logic show that:

(i) Many theorems true about the set E remain true about the set ${}^*E$ and vice versa;
(ii) propositions dealing about limits (i.e., topological propositions) in E can be translated into simpler propositions about NS objects in ${}^*E$.

Mathematical propositions about E can refer to elements of E ("*individuals*"), but also to *relations* on E of any type. Until now, we have only considered individuals; since many theorems cannot be expressed in terms of individuals alone, we shall get deeper results by working with the *higher-order structure on E*, that consists of all individuals and relations of E.

A. Description of the higher-order structure on E

The relations on E are the elements of any of the following sets: E; $E \times E$ (the set of ordered pairs of individuals), $\mathcal{P}(E)$ (the set of subsets of E), and all sets deductible from E by finite successive applications of $\times$ and $\mathcal{P}(\ )$, such as $\mathcal{P}(E \times \mathcal{P}(E \times E) \times \mathcal{P}(\mathcal{P}(E)))$, for instance. Any relation belongs to one of those sets, and the order of succession of $\mathcal{P}$ and $\times$ for that set defines the *type* of the relation.

Examples on $E = \mathbb{R}$: The order relation $\leq$ is a binary relation on $\mathbb{R}$, and it is defined by its graph: $G(\leq) = \{(x, y) \in \mathbb{R}^2 \mid x \leq y\}$: a subset of $(\mathbb{R} \times \mathbb{R})$; therefore $\leq$ is considered an element of $\mathcal{P}(\mathbb{R} \times \mathbb{R})$. The same is true for any function from $\mathbb{R}$ to $\mathbb{R}$ (a function is always a relation). The law of multiplication on $\mathbb{R}$, seen as a function on two variables, is an element of $\mathcal{P}(\mathbb{R} \times \mathbb{R} \times \mathbb{R})$. The set $L^2(\mathbb{R})$ is an element of $\mathcal{P}(\mathcal{P}(\mathbb{R} \times \mathbb{R}))$. These examples show that all current mathematical objects are relations, including individuals, subsets, mappings, and relations in everybody's sense.

Enlargement of E: The ultrapower construction is carried out simultaneously on the individuals *and* relations of E.

(i) If $\{e_n\}$ is a sequence of individuals of E, its equi-

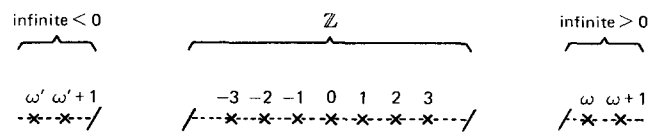

FIG. 2. The nonstandard integers.

$$\left.\begin{array}{l} f(x) = o(1) \\ f(x) = \mathfrak{G}(1) \\ f \text{ continuous} \\ f \text{ has derivative } C \end{array}\right\} \text{ at } x_0 \in \mathbb{R} \Longleftrightarrow \forall x \in {}^*\mathbb{R}: x \sim x_0 \Rightarrow \begin{cases} f(x) \in M_1 \\ f(x) \in M_0 \\ f(x) \sim f(x_0) \\ \dfrac{f(x) - f(x_0)}{x - x_0} \sim C \end{cases}$$

f is uniformly continuous on $[a,b] \Longleftrightarrow \forall x, y \in {}^*[a,b]$: $f(x) \sim f(y)$

FIG. 3.

valence class (for the relation $\{e_n\} \overset{\text{a.e.}}{=\!=} \{e'_n\}$) defines an individual of $^*E$ (as in the case of $\mathbb{R}$, above) noted $[e_n]$.

(ii) If $\{(e_n, f_n)\}$ is a sequence of pairs, $(e_n, f_n) \in E \times E$, its equivalence class $[(e_n, f_n)]$ for the relation $\{(e_n, f_n)\} \overset{\text{a.e.}}{=\!=} \{(e'_n, f'_n)\}$ is identified with the pair $([e_n], [f_n]) \in {}^*E \times {}^*E$. All elements of $({}^*E \times {}^*E)$ can be obtained this way (as in the case of $\mathbb{R}^p$ above).

(iii) If $\{E_n\}$ is a sequence of subsets, $E_n \in \mathcal{P}(E)$, its equivalence class $[E_n]$ (for the relation $\{E_n\} \overset{\text{a.e.}}{=\!=} \{E'_n\}$ can be identified with the subset of all individuals $[e_n]$ such that $e_n \overset{\text{a.e.}}{\in} E_n$, therefore we write $[E_n] \in \mathcal{P}(^*E)$. However, *all subsets of $^*E$ cannot be obtained* this way. Those that cannot are called *external*; those that can are called *internal*; among the latter, those that are equivalence classes of constant sequences are called *standard*. Generalizing:

If $\{r_n\}$ is a sequence of relations on E *of a fixed type*, its class $[r_n]$ can be identified with a relation on $^*E$ of the same type. The relations on $^*E$ are external or internal (and among these: standard).
By definition, all individuals are internal.

Examples on $E = \mathbb{R}$:

Subsets: if $\{E_n\}$ is the constant sequence $E_n = \mathbb{R}$ (or $\mathbb{Z}$), then $[E_n] = {}^*\mathbb{R}$ (or $^*\mathbb{Z}$) (as defined above); if $E_n = [a,b]$ (or $(a,b]$, etc$\cdots$) where $a, b \in \mathbb{R}$, then $[E_n] = {}^*[a,b]$, the closed interval of $^*\mathbb{R}$ with endpoints a and b (or: $^*(a,b]$, etc$\cdots$). Therefore, $^*\mathbb{R}$, $^*\mathbb{Z}$, $^*[a,b]$, $^*(a,b]$, etc$\cdots$ (if $a, b \in \mathbb{R}$) are *standard* subsets, the nonstandard extensions to $^*\mathbb{R}$ of the corresponding sets in $\mathbb{R}$.

If a, b are NS we can only say that the intervals of $^*\mathbb{R}$: $[a,b]$, $(a,b]$, etc$\cdots$ are *internal* subsets. The subsets $\mathbb{R}$, $M_0, M_1, \mu(a)$ (for $a \in \mathbb{R}$) are *external*.

Relations, functions: If $\{r_n\}$ is a constant sequence of relations with value $+, \times, \leq$, or any function $\mathbb{R}^m \underset{\rightarrow}{f} \mathbb{R}^{m'}$ (m, m' finite), then $[r_n]$ is a *standard* relation, we choose to denote it by the same symbol $+, \times, \leq$ or f,[7] remembering that it is the NS extension to $^*\mathbb{R}$ of the corresponding relation on $\mathbb{R}$. The functions $x \to x + \omega$, $x \to x^\omega$, etc$\cdots$ (where $\omega \in {}^*\mathbb{R}$) are *internal*. The function $x \to st(x)$ is *external*.

The main theorem: If $^*E$ is an enlargement of E then the following are true:

(1) Every mathematical notion that is meaningful for E, is meaningful for $^*E$.

$$\begin{array}{ll} x \in T \text{ is an isolated point} & \Longleftrightarrow \ \mu(x) = \{x\} \\ E \subset T \text{ is an open set} & \Longleftrightarrow \ \forall x \in E: \mu(x) \subset {}^*E \\ E \subset T \text{ is a closed set} & \Longleftrightarrow \ \forall x \in T\text{-}E: \mu(x) \cap {}^*E = \phi \\ T \text{ is a Hausdorff space} & \Longleftrightarrow \ \forall x, y \in T, x \neq y: \mu(x) \cap \mu(y) = \phi \\ T \text{ is a compact space} & \Longleftrightarrow \ \forall x \in {}^*T: x \text{ is "near-standard"}^{11} \\ T \underset{\rightarrow}{f} T' \text{ is} \left\{\begin{array}{l}\text{continuous} \\ \text{bicontinuous}\end{array}\right\} \text{at } x \in T & \Longleftrightarrow \left\{\begin{array}{l} f(\mu_T(x)) \subset \mu_{T'}(f(x)) \\ f(\mu_T(x)) = \mu_{T'}(f(x)) \end{array}\right\} \end{array}$$

$$\lim_{n\to\infty} f_n = b \quad \left\{\begin{array}{l}\text{pointwise} \Longleftrightarrow \forall x \in T \\ \text{uniformly} \Longleftrightarrow \forall x \in {}^*T\end{array}\right\} \forall n \text{ infinite: } f_n(x) \sim f(x)$$

FIG. 4.

(2) Every proposition that is true for E, is true for $^*E$ if we restrict our attention to the *internal* entities of the higher-order structure of $^*E$.
(3) Every proposition that is true for the internal entities of $^*E$ and is meaningful for E, is true for E.
(4) Additional remarks: All individuals of $^*E$ are internal; $^*E$ is strictly bigger than E iff E is not a set with a finite number of elements.

3. APPLICATIONS TO ANALYSIS AND TOPOLOGY

A. Convergence of sequences

A sequence $\{a_n\}_{n\in\mathbb{N}}$ is a function $\mathbb{N} \to \mathbb{R}$. Its NS extension is a *standard* function $^*\mathbb{N} \to {}^*\mathbb{R}$ denoted by the same symbol $\{a_n\}_{n\in{}^*\mathbb{N}}$. Now a_n is defined for infinite n, too. Properties about $\{a_n\}_{n\in\mathbb{N}}$ can be translated into properties of $\{a_n\}_{n\in{}^*\mathbb{N}}$ as follows[8]:

$$\lim_{\substack{n\to\infty \\ n\in N}} a_n = a \Longleftrightarrow \forall n \text{ infinite: } a_n \sim a,$$

$\{a_n\}$ is a Cauchy sequence $\Longleftrightarrow \forall m, n$ infinite: $a_m \sim a_n$.

B. Local properties of functions

Let f be a function $\mathbb{R} \underset{\rightarrow}{f} \mathbb{R}$; the same symbol will denote its NS extension to a standard function $^*\mathbb{R} \underset{\rightarrow}{f} {}^*\mathbb{R}$. Again, properties can be translated (see *Fig. 3*).[9]

differentials: If f is a standard function, df is defined as the internal function with two arguments:

$$df(x,h) = f(x+h) - f(x), \text{ where } x \in {}^*\mathbb{R} \text{ and } h \in M_1.$$

Riemann integration: If f is a function $[a,b] \underset{\rightarrow}{f} \mathbb{R}$ where $[a,b]$ is a finite interval, then f is Riemann-integrable and $\int_a^b f = I \Longleftrightarrow$ for any subdivision $(x_0 = a, x_1, \ldots, x_n = b)$, where $n \in {}^*\mathbb{N}$ is infinite and $x_j - x_{j-1} = (b-a)/n$, we have

$$I = st\left(\frac{1}{n}\sum_1^n f(x_j)\right).$$

Lebesgue integration: Equidistant subdivisions as above are too coarse to give the correct integral if we apply the formula to any measurable function. However, a more refined set of points $\{x_1, \ldots, x_n\}$ (for some infinite $n \in {}^*\mathbb{N}$)[10] can be found (but not written explicitly) so that for every $f \in L^1[a,b]$,

$$\int_a^b f = st\left(\frac{1}{n}\sum_1^n f(x_j)\right).$$

C. Topological spaces

Let T be a space with a topology defined by the family $\mathcal{T}$ of open sets, and let $^*T$ be an enlargement of T. For any (standard) $x \in T$, the *monad* of x is the subset of $^*T$:

$$\mu(x) = \underset{\substack{V\in\mathcal{T} \\ x\in V}}{{}^*V}$$

The reader can check that this definition, in the case of $\mathbb{R}$ with the usual topology, coincides with the one already given. The monads are essentially dependent on the topology (the finer the latter, the smaller the former) and they describe the topological properties of T (*Fig. 4*).

metric spaces: If T has a metric $d(x,y)$, this has a nonstandard extension to a function $^*T \times {}^*T \underset{\rightarrow}{d} {}^*\mathbb{R}_+$.

Theorem: For $x \in T$: $\mu(x) = \{y \in {}^*T \mid d(x,y) \sim 0\}$. All properties in Fig. 4 remain valid; in addition:

(i) T is a bounded space $\Longleftrightarrow$ all points of ${}^*T$ are at finite distance from T. (Such points are called *finite*.)

(ii) T is a complete space $\Longleftrightarrow$ for any standard Cauchy sequence $\{x_n\}$ there exists an infinite n and an $x \in T$ such that $x_n \sim x$.

(iii) Properties of sequences of standard functions $\{f_n\}$ are translated in *Fig. 5*.

Hilbert spaces: If H is a separable Hilbert space on $\mathbb{C}$ and $\{e_n\}_{n\in\mathbb{N}}$ is a basis of H, then ${}^*H$ is a Hilbert space on ${}^*\mathbb{C}$ with basis $\{e_n\}_{n\in{}^*\mathbb{N}}$. Let P^ω be any internal projector of ${}^*H$ of rank ω (ω infinite, $\omega \in {}^*\mathbb{N}$) satisfying

$$\forall x \in {}^*H\text{: } x \text{ is near-standard} \Longrightarrow \|x - P_\omega x\| \sim 0.$$

Then the space $H_\omega = P_\omega({}^*H)$ is finite-dimensional in the sense of ${}^*H$ (because its dimension is a well-defined number of ${}^*\mathbb{N}$) and approximates all points of H. This is interesting because all the machinery of finite-dimensional linear algebra applies to H_ω. Many choices of H_ω are possible. For example,

(i) take H_ω = subspace generated by $\{e_1, \ldots, e_\omega\}$;

(ii) in the case of $H = L^2(\mathbb{R})$, choose $\epsilon \sim 0$ positive such that $\epsilon\omega$ is infinite, and define, for $k \in \mathbb{R}$, and $f \in {}^*H$:

$$(P'_\omega)(k) = \frac{1}{\epsilon}\int_{n\epsilon}^{(n+1)\epsilon} f(k')dk'$$

if $|k| < \epsilon\omega$ and $k \in [n\epsilon, (n+1)\epsilon)$

$= 0$ otherwise.

(H_ω is then a space of step functions.)

If T is a bounded operator on H, we define: $T_\omega = P_\omega T P_\omega$ acting on H_ω, and conversely: $T = (st \circ T_\omega)_{|H}$, so that all the information about T is contained in the "finite-dimensional" T_ω.

4. APPLICATIONS TO PHYSICS

It could be suitable for mathematical physics to assume that the position or the momentum space, for instance, are NS spaces and to perform all computations in the NS formalism, provided we take the standard part of the result at the end (this corresponds to the finite precision of the physical measuring process). We give two quantum mechanical examples of a trivial nature; but examples in statistical mechanics or other fields might be interesting too.

A. The Dirac formalism

First we recall well-known results: Given a self-adjoint operator A on a Hilbert space H, we consider its spectral decomposition, i.e., the unique family $E(\lambda)_{\lambda\in\mathbb{R}}$ of projectors of H satisfying[12]

(1) for all $\lambda, \mu \in \mathbb{R}$: $E(\lambda)E(\mu) = E(\min(\lambda,\mu))$;

(2) $E(-\infty) = 0, E(+\infty) = 1, E(\lambda + 0) = E(\lambda)$ for all λ;

(3) $A = \int_{\mathbb{R}} \lambda \cdot dE(\lambda)$;

where the operator limits involved in (2) and (3) are strong limits.

The support of the measure $dE(\lambda)$ is the spectrum σ of A, which consists of a pure point (eigenvalue) part σ_{pp} and an absolutely continuous spectrum σ_{ac}. We can then

the family $\{f_n\}$ is

continuous	$\Longleftrightarrow \forall n \in N \ \forall x \in T$	$\forall y \in {}^*S$: $y \sim x \Longrightarrow f_n(y) \sim f_n(x)$
uniformly continuous	$\Longleftrightarrow \forall n \in N \ \forall x \in {}^*T$	
equicontinuous	$\Longleftrightarrow \forall n \in {}^*N \ \forall x \in T$	
uniformly equicontinuous	$\Longleftrightarrow \forall n \in {}^*N \ \forall x \in {}^*T$	

FIG. 5.

decompose H as:

$$H = \left(\int_\sigma dE(\lambda)\right)(H)$$

$$= \left(\int_{\sigma_{pp}} dE(\lambda)\right)(H) \oplus \left(\int_{\sigma_{ac}} dE(\lambda)\right)(H) = H_{pp} \oplus H_{ac}.$$

An essential property of H_{pp} is that it is spanned by the eigenvectors of A; this result is widely used in quantum mechanics. A similar description of H_{ac} would be useful; but the corresponding "eigenstates" of A cannot qualify as vectors of H because they should have infinite norm. Such "generalized eigenstates" are defined by the use of nested Hilbert spaces, as a rule,[13,14] but NS analysis provides an alternative description where all eigenstates, proper or generalized, are vectors of ${}^*H$.

Consider the NS spectral decomposition ${}^*A = \int_{{}^*\mathbb{R}} \lambda \cdot dE(\lambda)$ and define for some infinite integer ω and all $k \in {}^*\mathbb{Z}$ the projectors $E_{k/\omega} = E[(k + \frac{1}{2})/\omega] - E[(k - \frac{1}{2})/\omega]$ the subspaces $H_{k/\omega} = E_{k/\omega}({}^*H_{ac})$, and the operator A_ω by the formulas

$$A_\omega = {}^*A_{pp} \text{ on } {}^*H_{pp}, \quad A_\omega = k/\omega . 1 \text{ on } H_{k/\omega}.$$

Then ${}^*H$ decomposes as a direct sum of *eigenspaces* of A_ω:

$${}^*H = {}^*H_{pp} \oplus \left[\bigoplus_k H_{k/\omega}\right]; A_\omega = {}^*A_{pp} + \sum_k \frac{k}{\omega} E_{k/\omega}$$

and the operator norm of $(A - A_\omega)$ is $1/2\omega \sim 0$, which implies that for all near-standard $x \in {}^*H$, $\|({}^*A - A_\omega)x\| \sim 0$.

On any standard interval where the continuous spectrum of A has finite multiplicity[12] m, we can moreover choose projectors $E'_{k/\omega}$ *of rank* m onto subspaces $H'_{k/\omega}$ of $H_{k/\omega}$ in such a way that the operator $A'_\omega = {}^*A_{pp} + \sum_k k/\omega E'_{k/\omega}$ still satisfies $\|({}^*A - A'_\omega)x\| \sim 0$ for all near-standard x; and A'_ω also reproduces the original multiplicity of A. The proof is left to the reader.

The nonstandard treatment of this problem is in no way unique; but all models should give the same numerical results provided these are finite.

As an example, take the momentum representation in one-dimensional wave mechanics: $H = L^2(\mathbb{R}, dp)$; A = multiplication by p. Then ${}^*H_{pp} = \{0\}$, and $H'_{k/\omega}$ is the one-dimensional space generated on ${}^*\mathbb{C}$ by the function equal to 1 on the interval $((k - \frac{1}{2})/\omega, (k + \frac{1}{2})/\omega]$, to 0 elsewhere; A'_ω is multiplication by k/ω on $H'_{k/\omega}$. This approximation is similar to the finite box cutoff used in constructive quantum field theory,[15] but our cutoff is infinite.

B. The canonical commutation relations (CCR)

Gårding and Wightman[16] have written down all the representations of the following CCR: $[a_k, a_l^+] = \delta_{kl}$ ($k, l \in \mathbb{N}$). For any representation, the Hilbert space H is explicited

as a direct integral. Its NS extension $^*H$ is then a representation space for: $[a_k, a_l^\dagger] = \delta_{kl}(k, l \in {}^*\mathbb{N})$.

For any *-finite $\omega \in {}^*\mathbb{N}$, we can cancel all the oscillators except $a_0, \ldots, a_\omega$ by applying some projector P_ω on $^*H$; P_ω has all the properties mentioned in the section on Hilbert spaces (except that its rank is not *-finite). Moreover, the space $P_\omega(^*H)$ carries a representation of the CCR with a *-finite number of degrees of freedom, and the theorems about the finite case can be carried over: This reduced representation, an approximation of the original one, is a direct sum of equivalent irreducible representations.

We hope that a more refined use of NS analysis might help classifying the representations of the CCR.

ACKNOWLEDGMENTS

The author wishes to thank Professor A. Jaffe (Harvard University) and Dr. P. Kelemen (NYU) for their encouragements and advice.

APPENDIX

An ultrafilter on $\mathbb{N}$ is any family $\mathfrak{U}$ of subsets of $\mathbb{N}$ satisfying the following properties[17]:

(A): $\phi \notin \mathfrak{U}, \mathbb{N} \in \mathfrak{U}$

(B): $X_1, X_2 \in \mathfrak{U} \Longrightarrow X_1 \cap X_2 \in \mathfrak{U}$

(C): $X \in \mathfrak{U}, X \subset X' \Longrightarrow X' \in \mathfrak{U}$

(D): for any $X \subset \mathbb{N}$: either X or its complement is an element of $\mathfrak{U}$ [both cannot be, because of (A) and (B)]. $\mathfrak{U}$ is called a *free ultrafilter* if it satisfies:

(E): if $X \subset \mathbb{N}$ is a finite set, its complement is an element of $\mathfrak{U}$.

As a consequence of the axiom of choice (or Zorn's lemma) there exists a free ultrafilter on $\mathbb{N}$. There even exists an infinity of them, but they are all equivalent for our practical purposes.

Connection with measure theory: If $\mathfrak{U}$ is a free ultrafilter, then define, for any subset $X \subset \mathbb{N}$:

$$\mu(X) = 0 \text{ iff } X \notin \mathfrak{U}, \quad \mu(X) = 1 \text{ iff } X \in \mathfrak{U}.$$

Properties (A)-(D) imply that this is a *measure*[18] defined on all subsets of $\mathbb{N}$; (E) implies that any finite subset has measure 0 and the whole set $\mathbb{N}$ has measure 1.

---

*Détaché du CNRS.

[1]See P. Halmos, *Naive Set Theory* (Van Nostrand, Princeton, N. J., 1960).

[2]See N. Bourbaki, *Topologie générale* (Hermann, Paris, 1971), Chap 1, 2, 9.

[3]P. Kelemen and A. Robinson, J. Math. Phys. **13,** 1870 (1972).

[4]A. Robinson, *Non-Standard Analysis* (North-Holland, Amsterdam, 1966).

[5]A shorter account is given by W. Luxemburg, in *Lecture Notes on NS Analysis* (Caltech, Pasadena, Calif., 1966).

[6]If we took an ultrafilter on an uncountable set, we would obtain a larger structure: Here we shall get the smallest nontrivial enlargement, which is sufficient for most purposes.

[7]However, if a relation F is viewed as a subset, its extension is traditionally written $^*F$.

[8]See Ref. 4, pp. 60–63.

[9]See Ref. 4, pp. 66–81.

[10]Here a bigger $^*\mathbb{N}$ is needed; see Ref. 5.

[11]This means: $x \sim y$ for some $y \in T$.

[12]See K. Yosida, *Functional Analysis* (Springer-Verlag, Berlin, 1968), Chap. XI.

[13]See J. E. Roberts, J. Math. Phys. **7,** 1097 (1966).

[14]See I. M. Gel'fand and N. Vilenkin, *Generalized Functions* (Academic, New York, 1964), Vol. 4, Chap. 1.

[15]See J. Glimm and A. Jaffe, Less Houches Summer School Lectures (1970), (Gordon and Breach, New York) (to appear).

[16]In Proc. Natl. Acad. Sci. USA **40,** 622 (1954).

[17]See G. Choquet, *Lectures on Analysis* (Benjamin, New York, 1969), Chap. 1, Sec. 4.

[18]This measure is additive, but not countably additive.